

Liquidity and the Threat of Fraudulent Assets*

Yiting Li

Guillaume Rocheteau

Pierre-Olivier Weill

October 3, 2011

Abstract

We study an over-the-counter (OTC) market with bilateral meetings and bargaining where the usefulness of assets, as means of payment or collateral, is limited by the threat of fraudulent practices. We assume that agents can produce fraudulent assets at a positive cost, which generates endogenous upper bounds on the quantity of each asset that can be sold, or posted as collateral in the OTC market. Each endogenous, asset-specific, resalability constraint depends on the vulnerability of the asset to fraud, on the frequency of trade, and on the current and future prices of the asset. In equilibrium, the set of assets can be partitioned into three liquidity tiers, which differ in their resalability, their prices, their sensitivity to shocks, and their responses to policy interventions. The dependence of an asset's resalability on its price creates a pecuniary externality, which leads to the result that some policies commonly thought to improve liquidity can be welfare reducing.

J.E.L. Classification: D82, D83, E40, E50, G10

Keywords: search, payments, collateral, fraud, liquidity premia.

*A previous version of this paper was circulated under the title "Liquidity Constraints." We thank two anonymous referees and Rob Shimer for comments that greatly improved the paper. We also thank Veronica Guerrieri for an insightful discussion, Andy Atkeson, Bill Branch, Dean Corbae, Manolis Galenianos, Hugo Hopenhayn, Thomas Philippon, Shouyong Shi, and seminar participants at the Bank of Canada, the Federal Reserve Bank of Chicago, the Federal Reserve Bank of Cleveland, the National University of Seoul, the National Taiwan University, the Singapore Management University, the University of California at Riverside and Santa Cruz, the University of Wisconsin, and the Wharton school at the University of Pennsylvania for useful comments. We thank Monica Crabtree-Reusser for editorial assistance. Yiting Li: National Taiwan University, yitingli@ntu.edu.tw; Guillaume Rocheteau: University of California, Irvine and Federal Reserve Bank of Cleveland, grochete@uci.edu; Pierre-Olivier Weill: University of California, Los Angeles, NBER, and CEPR, poweill@econ.ucla.edu.

1 Introduction

Liquidity premia, or convenience yields, are key determinants of asset prices. This point is uncontroversial for fiat money, which derives its value solely from its liquidity services. According to Krishnamurthy and Vissing-Jorgensen (2010), the same is true for government securities, high-grade corporate bonds, and agency bonds. In this paper we present a theory of asset liquidity and convenience yields, based on the following premise: an asset’s liquidity—the extent to which it can facilitate exchange, as means of payment or as collateral—depends on its vulnerability to fraud. We address a class of questions related to the cross-sectional dispersion and time-variation of liquidity premia, such as what fundamental characteristics make some assets have higher turnover and lower yields than others? What shocks prompt investors to suddenly shift their portfolios towards the most liquid assets, which leads to widening yield spreads? Are liquid assets more susceptible of exhibiting excess volatility? And, what types of open-market operations and financial regulations are effective to mitigate aggregate liquidity shortages?

The threat of fraud has been a pervasive friction throughout history. Classical examples include the clipping of coins in ancient Rome and medieval Europe, and the counterfeiting of banknotes during the first half of the 19th century in the United States (Sargent and Velde, 2002; Mihm, 2007). Modern financial assets are no less susceptible to fraud. Intangible means of payment suffer from identity thefts (Schreft, 2007), and mortgage-backed securities are subject to moral hazard problems and lax incentives that plague the process of securitization (see, among others, Keys, Mukherjee, Seru, and Vig, 2010).¹ Similarly, the fact that some investors can spend resources to cherry-pick the collateral used to secure risk-sharing arrangements is a concern for participants in OTC derivative markets.²

We introduce the threat of fraud into a search-theoretic model of asset markets, building on

¹The 2010 “Performance and Activity Report” of the SEC details many cases of financial fraud related to mortgage-based securities. Frauds and moral hazard problems in the mortgage market are not new. Snowden (2010) describes the US mortgage crisis of the late 20s and 30s and the earlier forms of securitization in the 20s. Real estate bond houses were overappraising properties, they violated underwriting standards, and they substituted bad loans for performing mortgages in their mortgage pools.

²The International Swap and Derivatives Association (2010) reported that over 78% of OTC derivatives trades are collateralized. Importantly, market participants consider some asset classes (e.g., cash or government securities) to be of higher collateral quality than others. Collateral quality depends on various factors such as volatility, credit risk, and pricing ease.

recent work in monetary and financial economics (e.g., Lagos and Wright, 2005; Duffie, Gârleanu, and Pedersen, 2005). In the first period, agents trade an arbitrary number of assets in a competitive market. In the second period, they trade goods and services in an over-the-counter (OTC) market, with bilateral meetings and bargaining. Because of the frictions caused by a lack of commitment and limited enforcement, agents use assets as means of payment, or as collateral, in the OTC market. However, the extent to which an asset can play such a role is limited by the threat of fraud: after incurring an asset-specific fixed cost, an agent can produce fraudulent assets, which are worthless and indistinguishable from their genuine counterparts. In order to solve the resulting OTC bargaining problem under asymmetric information, we assume that the asset holder makes a take-it-or-leave-it offer, and we use the recent methodology of In and Wright (2011) for signaling games with hidden choices to select an equilibrium.

A key insight of our analysis is that the threat of fraud generates asset-specific, endogenous resalability constraints. While there are no exogenous restrictions on the transfer of assets in bilateral matches in the OTC market, if the quantity of an asset offered is above some threshold, then the trade is rejected with positive probability because of the rational fear that the asset might be fraudulent. In equilibrium, agents never find it optimal to offer more of the asset than what can be accepted with certainty, which prevents fraud from taking place. The resulting endogenous resalability constraint has three determinants: the asset's vulnerability to fraud, the difference between the asset's price and the discounted value of its cash flows, and the frequency of trades in OTC markets. We emphasize three main implications of these endogenous resalability constraints below.

First, because an asset's resalability depends on its own vulnerability to fraud, prices and measures of liquidity vary across assets with identical cash flows. We obtain an endogenous three-tier categorization of assets: illiquid, partially liquid, and liquid assets, which differ in their resalability, their price, as well as their sensitivity to shocks and policy interventions. While the price of an illiquid asset is equal to the present value of its cash flows, the price of a partially liquid or liquid asset is strictly larger than the present value of its cash flows; i.e., this asset enjoys a liquidity premium. This premium increases with the asset's recognizability but decreases with its supply, which is consistent with the downward-sloping aggregate demand for Treasury debt documented

in Krishnamurthy and Vissing-Jorgensen (2010). Finally, while the prices of illiquid and partially liquid assets are constant in the absence of shocks to fundamentals, the prices of liquid assets can exhibit self-fulfilling fluctuations.

Second, in a similar spirit as Guerrieri and Shimer (2011), our model identifies shocks that generate phenomena akin to flights to liquidity, whereby investors shift their asset demands from less liquid to more liquid assets, widening the liquidity spread between the two types of assets (see Longstaff, 2004, and Dick-Nielsen, Feldhutter, and Lando, 2010). For instance, we consider an increase in the frequency of liquidity needs in the OTC market that results in higher demand for collateral. Such a shock increases the value of holding assets, as they are more likely to be used as means of payment or as collateral, but it also has the countervailing effect of increasing fraud incentives. We show that the first effect dominates for liquid assets and raises their prices, while the second effect dominates for partially liquid assets and lowers their prices. Moreover, the set of liquid assets endogenously shrinks, meaning that agents shift their demand to the most recognizable assets, in accordance with a flight to liquidity. The same phenomenon can be generated in our model by a shock that raises the threat of fraud for some partially liquid or liquid assets, thereby reducing their resalability.

The third main implication of our results concerns policies aimed at managing the aggregate supply of liquidity through open-market operations or financial regulations. In our model, an open-market operation has a positive welfare effect if and only if it increases a simple measure of aggregate liquidity—a weighted sum of asset supplies. Therefore, a substitution of liquid assets with other liquid assets is irrelevant. An open-market purchase of illiquid assets with liquid ones, on the other hand, raises aggregate liquidity and output. However, under a balanced budget requirement, a purchase of partially liquid assets with liquid ones *reduces* aggregate liquidity, the yield of liquid assets, and output. This paradoxical result arises because of a "pecuniary externality," according to which an increase in the price of an asset reduces its resalability, which in turn can lower its liquidity premium *below* the true marginal social value of its liquidity services. Due to this externality, a balanced budget open-market purchase syphons out more liquidity than it is injecting in. This result can shed some light on quantitative easing, which consists of injecting reserves in exchange for less liquid assets (Krishnamurthy and Vissing-Jorgensen, 2011). According to our model, for

such policies to successfully increase aggregate liquidity, they must target the most illiquid assets. In a similar vein, we study retention requirements that were introduced by the Dodd-Frank Act to mitigate moral hazard problems in the securitization process. In the context of our model, such requirements are welfare improving only if applied to illiquid assets.

1.1 Literature review

Kiyotaki and Moore (2001, 2005) study limited resalability by assuming that each period, agents cannot sell more than an exogenous proportion of their asset holdings. While such exogenous resalability constraints can be chosen to replicate our distribution of asset prices, they generate markedly different comparative statics and policy recommendations (see Supplementary Appendix E). For instance, with proportional resalability constraints, an increase in the frequency of trading needs weakly increases the prices of all assets, while in our model it has asymmetric effects: it increases the prices of liquid assets, and decreases the prices of partially liquid assets, consistent with evidence on flight to liquidity. As another example, with proportional liquidity constraints, an open-market purchase of partially liquid assets with liquid ones increases liquidity, asset yields, and welfare. In our model, because of a new pecuniary externality, we obtain the opposite effects, consistent with evidence on quantitative easing.

In Holmstrom and Tirole's (2011, and references therein) corporate finance model, a moral hazard problem generates endogenous borrowing constraints, i.e., resalability constraints in the primary asset market. In the secondary market, corporate claims with identical cash flows enjoy the same liquidity premium. In our model, by contrast, we focus on moral hazard in secondary markets. We highlight the fact that agents' incentives to take hidden actions depend on contemporaneous secondary market prices and on OTC market frictions, and we generate cross-sectional differences in liquidity premia between assets with identical cash flows.

The search-theoretic literature on the liquidity structure of asset returns includes, e.g., Wallace (1998, 2000), Weill (2008), and Lagos (2010), and related work on the rate-of-return-dominance puzzle. Our approach goes beyond this earlier search literature by showing how cross-sectional differences in liquidity arise from fraud-based endogenous resalability constraints.³ Lester, Postle-

³Wallace (1998, 2000) emphasizes assets' indivisibilities, Weill (2008) assumes increasing returns in the matching

waite, and Wright (2011) consider a private information problem where agents can recognize the quality of an asset at some cost, but to determine the terms of trade under asymmetric information they make the simplifying assumption that unrecognized assets are not accepted in a bilateral match.⁴ They address this issue in an extension that follows our methodology closely.

There is a literature that emphasizes adverse selection problems in asset markets with search frictions (e.g., Hopenhayn and Werner, 1996). The most closely related papers are Rocheteau (2009) who introduces an adverse selection problem in a monetary model to explain the illiquidity of risky assets, and Guerrieri, Shimer, and Wright (2010), who consider a competitive search environment to illustrate how trading delays emerge endogenously to screen high- and low-quality assets. Guerrieri and Shimer (2011) extend the previous paper to a general equilibrium framework and, among other results, provide an explanation for flights to liquidity based on a dynamic adverse selection problem. While the distinction between adverse selection and moral hazard in asset markets is often subtle, the methodologies for capturing the two frictions differ profoundly. We take the view that informational asymmetries in asset markets often result from strategic behavior, which allows us to focus the model more squarely on the effects of the threat of fraud on asset liquidity. At a more theoretical level, an important distinction between adverse selection and moral hazard is that the type distribution is exogenous with the former, but is endogenous with the latter. With an exogenous type distribution, under some conditions, agents can mitigate the asymmetric information friction by holding broadly diversified asset portfolios. As our model demonstrates, when the type distribution is endogenous, the asymmetric information friction remains relevant.

The next section presents the model. Section 3 solves the bargaining game under the threat of fraud. Section 4 solves for asset prices, and Section 5 presents three main implications. The appendix contains omitted proofs, and the supplementary appendix presents additional results and extensions.

technology, and Lagos (2010) introduces exogenous restrictions on the use of some assets as means of payment. Similarly, Shi (2008) studies the pricing of bonds in a search economy where exogenous legal restrictions prevent bonds from being used in payments in a fraction of trades.

⁴There is also a related literature on counterfeiting, e.g., Green and Weber (1996), Williamson and Wright (1994), and Nosal and Wallace (2007). In those studies, there is a single asset, asset holdings are restricted to $\{0, 1\}$, and assets are indivisible, while those restrictions are all relaxed in our model.

2 The model

The economy lasts for two periods, $t \in \{0, 1\}$, and is populated by a continuum of agents who trade sequentially in two markets: in a centralized market (CM) at $t = 0$, and in a decentralized over-the-counter market (DM) at $t = 1$. There are two perfectly divisible and perishable goods. The first good, which we take to be the *numéraire*, is produced and consumed at $t = 0$ and at the end of $t = 1$. The second good, labeled the DM good, is produced and consumed in bilateral meetings in the DM. There is a finite set of assets indexed by $s \in S$. Each asset pays off at the end of $t = 1$ a dividend normalized to one unit of the numéraire.

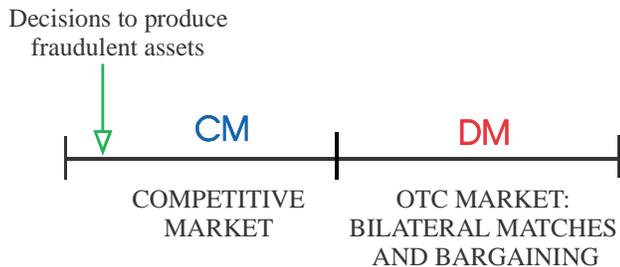


Figure 1: Timing of the game

Agents are divided evenly into two types, called *buyers* and *sellers*. Buyers wish to consume in the DM but cannot produce, while sellers have the technology to produce goods in the DM but do not want to consume. Together with frictions described below, this preference structure creates a need for liquidity: buyers will acquire assets in the CM in order to finance the purchase of goods produced by sellers in the DM. The utility of a buyer is:

$$x_0 + \beta [u(q_1) + x_1], \quad (1)$$

where $x_t \in \mathbb{R}$ is the consumption of the numéraire good at time t , with $x_t < 0$ being interpreted as production, $q_1 \in \mathbb{R}_+$ is the consumption of the DM good, and $\beta \equiv (1 + r)^{-1} \in (0, 1)$ is a discount factor. The utility function, $u(q)$, over the DM good is twice continuously differentiable, with $u(0) = 0$, $u'(q) > 0$, $u'(0) = \infty$, $u'(\infty) = 0$, and $u''(q) < 0$. The utility of a seller is:

$$x_0 + \beta (-q_1 + x_1), \quad (2)$$

where q_1 is the seller’s production in a pairwise meeting in the DM. Let $q^* = \arg \max_q [u(q) - q] > 0$ denote the output level that maximizes the match surplus, so $u'(q^*) = 1$.

The CM is a perfectly competitive market, where agents trade the numéraire good and assets. The DM, on the other hand, is an over-the-counter market, where a fraction $\sigma \in (0, 1]$ of buyers are matched bilaterally and at random with an equal fraction of sellers. Because of a lack of commitment and limited enforcement, buyers purchase DM goods with assets or, equivalently, with loans collateralized by assets (see footnote 11).

Terms of trade in pairwise meetings in the DM are determined according to a simple bargaining game, in which the buyer makes a take-or-leave-it offer.⁵ The buyer, whose asset holdings are private information, asks for a given amount of the DM good in exchange for some specified portfolio of assets.⁶ The seller accepts or rejects the offer. If the seller accepts the offer, then the trade is implemented, provided that the asset transfer is feasible given the buyer’s asset holdings. Matched agents split apart before assets pay off.

We introduce the possibility of asset fraud as follows. In the CM at $t = 0$, a buyer can pay a fixed cost $k(s) > 0$ to produce any quantity of fraudulent asset of type s . Fraudulent assets have zero terminal value and, in the DM, cannot be distinguished by sellers from genuine assets.

2.1 Interpretations

Counterfeiting of a means of payment. A literal interpretation of the model concerns assets used as means of payment, such as coins or banknotes, for which the fraud consists of producing counterfeits.⁷ During the first half of the 19th century, the fixed cost to produce fake banknotes included the cost to acquire plates and dies. See, e.g., Mihm (2007). Nowadays, this cost corresponds

⁵In her discussion of our paper, Veronica Guerrieri investigated a version of the model with competitive search and showed that this alternative pricing mechanism generates the same liquidity constraint as the one obtained under our simple bargaining game.

⁶By assuming that asset holdings are unobservable, we reduce the set of signals sent by a buyer, which simplifies the analysis. As shown in the earlier version of our working paper, results are robust to alternative assumptions regarding the observability of asset holdings. Also, we do not allow buyers to offer lotteries over allocations. In our context we conjecture that it is with no loss in generality, but such lotteries could be useful in the presence of alternative cost structures of producing fraudulent assets—see Supplementary Appendix C on variable costs.

⁷Our model can accommodate fraud on unsecured credit in bilateral matches. In this case, an agent has the option to produce a fake identity in the CM at a fixed cost (e.g., the cost incurred by a computer hacker to steal the identity of someone else) and he can issue an IOU in the DM if matched. The repayment of genuine IOUs can be enforced in the following CM. In contrast, IOUs based on fake identities are not repaid.

to the price of photo-editing software and copy machines.

Collateral fraud. An alternative interpretation is that buyers use assets as collateral to secure loans to be repaid at the end of $t = 1$. If the asset is a house, the transaction in the DM is an equity extraction loan to finance consumption. An example of mortgage fraud that closely resembles our model is the property flipping scheme, whereby a buyer obtains a high-loan-to-value mortgage based on a fake property appraisal, and the bank is left with worthless collateral.⁸ In this example, the cost of producing fraudulent assets represents the cost of creating false documentation about the borrower and the property. The DM can also be interpreted as an OTC market for credit derivatives, such as the market for credit default swaps or interest rate swaps. In that context, the goods traded in the DM are risk-sharing services, and collateral is used to mitigate counterparty risk.⁹ The cost of producing fraudulent assets is the informational cost incurred by the buyer to identify bad collateral. This cost is related to the complexity of the asset, its issuer, and the quality and quantity of information released about the asset's cash flows.

Securitization fraud. In this context buyers represent mortgage securitizers who originate and package loans in the CM. Sellers represent final asset holders who acquire securitized assets in the DM. There are gains from trading assets in the DM because it allows mortgage securitizers to spread the risk of the underlying loans to final asset holders.¹⁰ In this example, the cost of producing fraudulent assets is the cost of generating false documentation about the underlying security, bribing an agency for a good rating, or engaging in accounting frauds.

3 Bargaining under the threat of fraud

In this section we solve for the equilibrium of the game between a buyer and a seller matched at random. The game starts in the CM at $t = 0$ and ends in the DM at $t = 1$. For now we take as

⁸See http://www.fbi.gov/about-us/investigate/white_collar/mortgage-fraud/mortgage_fraud.

⁹In Supplementary Appendix G, we provide an explicit model of risk-sharing arrangements, where the DM good can be interpreted as risk-sharing services.

¹⁰In Supplementary Appendix H, we provide such a model of securitization, where agents have Constant Absolute Risk Aversion (CARA) utilities. This model confirms, albeit with different functional forms, that $u(q)$ can be interpreted as the utility of reducing the securitizer's risk position, and $q = c(q)$ can be interpreted as the cost of increasing the final asset holder's risk position.

given asset prices in the CM, $\phi(s)$, $s \in S$, and we anticipate that, in equilibrium, they will satisfy $\phi(s) \geq \beta$; i.e., the rate of return of asset s is no greater than the discount rate, which would be the “fundamental price” of the asset in a frictionless economy.

The sequence of moves is as follows: (i) In the CM at $t = 0$, the buyer chooses a portfolio of $\{a(s)\}$ genuine and $\{\tilde{a}(s)\}$ fraudulent assets, subject to $a(s) \geq 0$ and $\tilde{a}(s) \geq 0$; (ii) In the DM at $t = 1$, the buyer is matched with a seller with probability σ , in which case he makes an offer $(q, \{d(s)\})$, where q represents the output produced by the seller and $d(s)$ is the transfer of assets of type s (genuine or fraudulent) from the buyer to the seller; (iii) The seller decides whether to accept the offer; (iv) If the offer is accepted, the seller delivers q units of goods to the buyer, and the buyer delivers $\tau(s) \in [0, a(s)]$ genuine and $\tilde{\tau}(s) \in [0, \tilde{a}(s)]$ fraudulent units of asset s to the seller, with $\tau(s) + \tilde{\tau}(s) = d(s)$.¹¹ The extensive form of the game, for the $\sigma = 1$ case, is illustrated in the left panel of Figure 2. Arcs indicate that the action set at a given node is infinite, while a dotted line represents an information set.

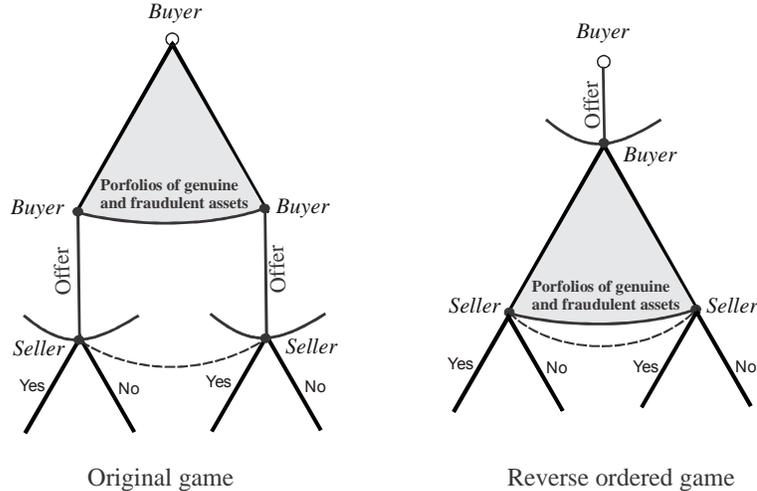


Figure 2: Game trees

¹¹We can reinterpret the payment, $(q, \{d(s)\})$, as a fully collateralized loan, where the buyer promises to repay $\sum_{s \in S} d(s)$ units of the CM output at the end of period 1. In order to secure the repayment of the loan, the buyer posts $d(s)$ units of asset s as collateral with a third party. If one asset is fraudulent, then the buyer will choose to default on his obligation, in which case the seller seizes the assets that serve as collateral. If all assets are genuine, then the buyer is indifferent between repaying his debt or defaulting.

Payoffs. The Bernoulli payoff of the buyer is:

$$-\sum_{s \in S} \left\{ k(s) \mathbb{I}_{\{\tilde{a}(s) > 0\}} + \phi(s) a(s) \right\} + \beta \mu \left\{ u(q) + \sum_{s \in S} [a(s) - \tau(s)] \right\} + \beta(1 - \mu) \sum_{s \in S} a(s),$$

where $\mathbb{I}_{\{\tilde{a}(s) > 0\}} = 1$ if the buyer produces fraudulent assets of type s , $\tilde{a}(s) > 0$, and zero otherwise. In the above, $\mu = 1$ if the buyer meets a seller who accepts his offer, and $\mu = 0$ otherwise. The first term is the payoff of the buyer at $t = 0$. In order to accumulate $\tilde{a}(s) > 0$ fraudulent units of asset s , the buyer must incur the fixed cost $k(s)$. In order to accumulate $a(s)$ units of genuine asset s , he must produce $\phi(s)a(s)$ units of the numéraire good in the CM. The second term is the discounted payoff at $t = 1$ if $\mu = 1$; i.e., if the buyer meets a seller in the DM and his offer is accepted. He then enjoys the utility of DM good consumption, $u(q)$, as well as the payoff from his net holding of genuine assets, $a(s) - \tau(s)$, the initial amount purchased net of the asset transfer to the seller, keeping in mind that each unit of genuine asset pays off one unit of the numéraire good at the end of $t = 1$. The last term is, similarly, the discounted payoff of the buyer at $t = 1$ if $\mu = 0$. Collecting terms, we can rewrite the payoff as

$$-\sum_{s \in S} \left\{ k(s) \mathbb{I}_{\{\tilde{a}(s) > 0\}} + \left[\phi(s) - \beta \right] a(s) \right\} + \beta \mu \left\{ u(q) - \sum_{s \in S} \tau(s) \right\}. \quad (3)$$

Similarly, the Bernoulli payoff of the seller is

$$\beta \mu \left\{ -q + \sum_{s \in S} \tau(s) \right\}, \quad (4)$$

where we anticipate that, in equilibrium, sellers will not find it optimal to accumulate assets in the CM.¹² If the seller accepts the offer ($\mu = 1$), he suffers the disutility of producing, q , and receives $\tau(s)$ genuine units of asset s .

Equilibrium concept. Our equilibrium concept is Perfect Bayesian Equilibrium: actions are sequentially rational following every history, and beliefs accord with Bayes's rule whenever it is possible. The notion of Perfect Bayesian Equilibrium imposes little discipline on the seller's belief in the DM regarding the decision of the buyer in the initial stage of the game to produce fraudulent assets, conditional on an off-equilibrium offer being made. Our approach to circumvent this

¹²Sellers have no strict incentives to accumulate assets if $\phi(s) \geq \beta$, because their asset holdings are not observable and hence do not affect the terms of trade offered by the buyer.

difficulty consists of adopting a notion of strategic stability, according to which any equilibrium of the original game should also be an equilibrium of the reverse-ordered game, whose timing is shown in Figure 2: (i) The buyer determines his DM offer, $(q, \{d(s)\})$, before making any decision in the CM (e.g., one interpretation is that he posts an offer at the beginning of the CM for the next DM); (ii) He chooses his portfolio composed of genuine and fraudulent assets; (iii) He is matched with a seller who chooses whether to accept or reject the offer.¹³ This reordered game captures the idea that upon seeing the buyer's offer, the seller will infer that the buyer's unobservable actions (portfolio and production of fraudulent assets) were chosen optimally with the offer in mind. The refinement is intuitive in that it selects an equilibrium of the original game that yields the highest payoff to the player making the offer, in our case the buyer. Moreover, it improves tractability as subgame perfection becomes sufficient to solve the game.

Solving for equilibrium. The analysis of the game can be simplified by making two observations. First, because of the fixed cost, the buyer will either produce the quantity of fraudulent assets that is necessary to execute the offer in a match or he will produce no fraudulent asset at all. Consequently, $\tilde{\tau}(s) = [1 - \chi(s)]d(s)$ and $\tau(s) = \chi(s)d(s)$, where $\chi(s) = 0$ if the buyer produces fraudulent assets, and $\chi(s) = 1$ otherwise. Moreover, the buyer must be able to cover his intended transfer of genuine assets; i.e., $a(s) \geq \chi(s)d(s)$.

Second, we can solve for the buyer's optimal asset demand before solving for equilibrium offers. Indeed, if $\phi(s) = \beta$, it follows from the buyer's payoff, (3), that any demand satisfying the constraint $a(s) \geq \chi(s)d(s)$ is optimal. If $\phi(s) > \beta$, it is costly to hold assets, and so it is optimal to demand $a(s) = \chi(s)d(s)$. In both cases, substituting the optimal asset demands into the objective amounts

¹³The re-ordering methodology, called the *reordering invariance refinement*, was developed by In and Wright (2011) for signaling games with unobservable choices. This refinement is based on the invariance condition of strategic stability from Kohlberg and Mertens (1986), which requires that the solution of a game should also be the solution of any game with the same reduced normal form. (The intuitive criterion does not apply to our game because in contrast to standard signaling games types are endogenous.) Beside being powerful in selecting equilibria and tractable (because subgame perfection becomes sufficient to solve the game), this equilibrium notion has a strong decision-theoretic justification and nice normative properties. Specifically, in our model the reordered game captures the idea that upon seeing the buyer's offer, the seller will infer that the buyer's unobservable actions (portfolio and production of fraudulent assets) were chosen optimally with the offer in mind. (This forward induction logic is reminiscent to the one of most refinements in the signaling literature.) From a normative viewpoint, this refinement has the appealing property of selecting an equilibrium of the original game that yields the highest payoff to the buyer, the agent making the offer. A more detailed description of the merits of this approach is provided in In and Wright (2011).

to replacing $a(s)$ with $\chi(s)d(s)$.

With these observations in mind, a buyer's strategy specifies the following two objects: the offer, $(q, \{d(s)\})$, and conditional on any offer, a probability distribution over $\{\chi(s)\} \in \{0, 1\}^S$, denoted by η . The seller's strategy specifies, conditional on any offer $(q, \{d(s)\})$, the probability of accepting, denoted by π .

The game is solved by backward induction. Following an offer, $(q, \{d(s)\})$, the seller's decision to accept a trade must be optimal given the buyer's decision to produce fraudulent assets; i.e.,

$$\pi \in \arg \max_{\hat{\pi} \in [0, 1]} \hat{\pi} \left\{ -q + \sum_{s \in S} \eta(s)d(s) \right\}, \quad (5)$$

where $\eta(s)$ denotes the marginal probability of bringing genuine assets of type s .¹⁴ The seller's value of accepting the offer depends on the disutility of producing q units of goods and on the expected quality of the asset transfer, determined by η .

Similarly, following an offer $(q, \{d(s)\})$, the buyer's decision to bring genuine or fraudulent assets is optimal given the seller's probability of accepting; i.e.,

$$\{\eta(s)\} \in \arg \max_{\{\hat{\eta}(s)\}} - \sum_{s \in S} \left\{ k(s) [1 - \hat{\eta}(s)] + [\phi(s) - \beta] \hat{\eta}(s)d(s) + \beta\sigma\pi\hat{\eta}(s)d(s) \right\}, \quad (6)$$

where the expression that is maximized consists of the terms in the buyer's payoff that depend on η . It shows that there are two gains from producing fraudulent assets: the savings in the holding cost, $\phi(s) - \beta$, and the savings in the expected cost of transferring genuine assets to a seller.

Finally, given equilibrium decision rules $\{\eta(s)\}$ and π , the optimal offer, $(q, \{d(s)\})$, maximizes the following objective

$$- \sum_{s \in S} \left\{ k(s) [1 - \eta(s)] + [\phi(s) - \beta] \eta(s)d(s) \right\} + \beta\sigma\pi \left\{ u(q) - \sum_{s \in S} \eta(s)d(s) \right\}. \quad (7)$$

A perfect Bayesian equilibrium that satisfies the reordering invariance refinement is a pair of buyer's and seller's strategies satisfying (5), (6), and (7). The next proposition provides a simple joint characterization of the asset demands and the offers made in any equilibrium.

¹⁴Note that, after replacing $a(s)$ and $\tau(s)$ with $\chi(s)d(s)$ in (3) and (4), the payoffs of buyers and sellers become linear functions of the binary actions $\{\chi(s)\}$. Therefore, taking expectations with respect to η amounts to replacing $\chi(s)$ with the marginal probability $\eta(s)$.

Proposition 1 *The asset demands, $\{a(s)\}$, and the equilibrium offers, $(q, \{d(s)\})$, solve:*

$$\max_{q, \{a(s), d(s)\}} \left\{ - \sum_{s \in S} [\phi(s) - \beta] a(s) + \beta \sigma [u(q) - q] \right\} \quad (8)$$

$$s.t. \quad \sum_{s \in S} d(s) - q = 0 \quad (9)$$

$$d(s) \leq \frac{k(s)}{\phi(s) - \beta + \beta \sigma}, \quad \text{for all } s \in S \quad (10)$$

$$d(s) \in [0, a(s)], \quad \text{for all } s \in S. \quad (11)$$

Moreover, following any equilibrium offer, the buyer transfers genuine assets with probability one, $\eta(s) = 1$ for all s , and the seller accepts the offer with probability one, $\pi = 1$.

Proposition 1 shows that equilibrium asset demands and offers maximize the buyer's expected utility subject to three constraints. First is the individual rationality constraint, (9), which states that the seller must be indifferent between accepting and rejecting the offer, given that the buyer's assets are genuine. The seller's expected payoff is zero since the bargaining protocol specifies that the buyer makes a take-it-or-leave-it offer. Second is the incentive compatibility constraint, (10), which states that the buyer must find it optimal to accumulate genuine assets with probability one, given that the seller accepts with probability one. Third is the feasibility constraint, (11), which states that the buyer must hold enough genuine assets to cover his transfer to the seller.

To understand why the buyer finds it optimal to bring genuine assets with probability one, consider a candidate equilibrium in which he brings genuine assets of type s_0 with a probability $\eta(s_0) \in (0, 1)$.¹⁵ In this candidate equilibrium, the buyer's payment capacity is slack. To see this, notice that the buyer could deviate and demand higher consumption in the DM, $q' > q$, keep the same $\{d(s)\}$, and compensate the seller by bringing genuine assets of type s_0 with higher probability, $\eta'(s_0) > \eta(s_0)$. This deviation would not change the buyer's expected cost of transferring assets, since he is indifferent between genuine or fraudulent assets of type s_0 . Moreover, by (6), indifference implies:

$$k(s_0) = [\phi(s_0) - \beta + \beta \sigma \pi] d(s_0) \quad \implies \quad \pi = \frac{k(s_0) - [\phi(s_0) - \beta] d(s_0)}{\beta \sigma d(s_0)};$$

i.e., the seller's probability of acceptance, π , is pinned down by the transfer $d(s_0)$, and is unaffected by the increase in q . Taken together, these observations mean that the buyer could increase his

¹⁵Looking at $\eta(s_0) > 0$ is without loss. See the proof of Proposition 1 for details.

payoff by raising his offer q without changing his expected cost of transferring the asset, and without changing the seller's acceptance probability.

Lastly, the proposition shows that, in equilibrium, the buyer always finds it optimal to make an offer that is accepted with probability one. This result is not obvious because offering more assets than the threshold of equation (10) has two effects going in opposite directions. The positive effect is that the buyer can demand a higher q in exchange for a higher d . The negative effect is that a larger offer increases fraud incentives, and hence it has a positive probability of being rejected. Our proof shows that, with the fixed cost of producing fraudulent assets, the negative effect always dominates.¹⁶

Endogenous resalability constraints. Perhaps the most important result of Proposition 1 is that the incentive-compatibility constraints, (10), take the form of resalability constraints, specifying upper bounds on the transfer of assets from buyers to sellers.¹⁷ The resalability constraints depend on the cost of producing fraudulent assets, $k(s)$, the holding cost of an asset, $\phi(s) - \beta$, and the frequency of trades in the DM, σ .

From (10), an asset which is more susceptible to fraud is subject to a more stringent resalability constraint. To illustrate this point, suppose that there are no search frictions, $\sigma = 1$. Then, the resalability constraint of asset s is $\phi(s)d(s) \leq k(s)$. The real value of the asset that can be transferred in a bilateral match is simply the cost of producing fraudulent assets. In accordance with the Wallace (1998) dictum, the liquidity of an asset depends on its intrinsic properties, which here are captured by the ease of producing fraudulent assets.

The resalability constraints also depend on the frequency of trade in the DM. Increasing the frequency of trade exacerbates the threat of fraud because the trade surplus of a con artist, $u(q)$, is greater than the match surplus of an honest buyer, $u(q) - q$. Therefore, the upper bound must

¹⁶In Supplementary Appendix C, we show that the negative effect also dominates if we add proportional costs of producing fraudulent assets provided that those costs are not too large. If the proportional costs are large relative to the fixed costs, then there can be situations where fraud generates rationing both at the intensive margin (the quantity of assets that can be transferred in a match) and at the extensive margin (the number of matches in which trade occurs).

¹⁷If the asset is interpreted as an IOU (see Footnote 7), $s = \ell$, then one can set $\phi(\ell) = \beta$ since an IOU is issued in the DM and there is no cost of holding it. In this case the incentive-compatibility constraint, (10), takes the form of a borrowing constraint, $d(\ell) \leq \frac{k(\ell)}{\beta\sigma}$.

be lowered to keep incentives in line. To give a concrete example, if the process of securitization implies that an asset can be retraded more frequently, then an increase in securitization raises the threat of fraud and makes resalability constraints more likely to bind.¹⁸

Finally, the holding cost of the asset, $\phi(s) - \beta$, enters the resalability constraint, because lack of commitment forces agents to accumulate assets before liquidity needs occur. An increase in the asset price raises the holding cost, which raises the buyer's incentives to produce fraudulent versions of the asset for a given size of the trade.

4 The liquidity structure of asset returns

In this section we study the implications of our model for cross-sectional liquidity premia. We endogenize asset prices in the CM and show that the endogenous resalability constraints derived in Proposition 1 create liquidity and price differences across assets, even if they have the same cash flows. Our results help explain differences in asset prices that cannot be fully accounted for by risk, and shed light on a variety of evidence on the positive relationship between liquidity and asset prices.¹⁹

4.1 The liquidity-return trade-off

Assume that each asset $s \in S$ comes in fixed supply, denoted by $A(s)$. We define a symmetric equilibrium to be a collection of prices, $\{\phi(s)\}$, asset demands, $\{a(s)\}$, and a DM offer, $(q, \{d(s)\})$, such that the asset demands and the offer solve the buyer's problem (8)-(11) given prices, and the asset market clears; i.e., $a(s) = A(s)$ for all $s \in S$.²⁰

Guessing that $a(s) \geq 0$ and $d(s) \geq 0$ do not bind, the first-order conditions of the buyer's

¹⁸Keys, Mukherjee, Seru, and Vig (2010) establish evidence that the securitization of subprime loans led to lax screening. Purnanandam (2009) finds that banks involved highly in the originate-to-distribute market, where the originator of loans sells them to third parties, originated excessively poor-quality mortgages.

¹⁹Since Amihud and Mendelson (1986), liquidity (level and risk) has been shown to explain risk-adjusted asset return differentials. For recent studies, see, e.g., Chordia, Huh, and Subrahmanyam (2009).

²⁰The symmetry restriction that all buyers have the same asset demands serves to pin down portfolios when some assets are priced at their fundamental values, $\phi(s) = \beta$.

problem are:

$$\xi = \beta\sigma [u'(q) - 1] = \lambda(s) + \nu(s) \quad (12)$$

$$\phi(s) = \beta + \nu(s), \quad (13)$$

for all $s \in S$, where $\xi \geq 0$ is the Lagrange multiplier of the seller's participation constraint, (9), $\lambda(s) \geq 0$ is the multiplier of the resalability constraint, (10), and $\nu(s) \geq 0$ is the multiplier of the feasibility constraint, (11). The multiplier, ξ , measures the net utility of spending an additional unit of asset in the DM, if matched with a seller with probability σ . The increased consumption yields marginal utility $u'(q)$ to the buyer, and the asset transfer has an opportunity cost equal to one.

Taken together, (12) and (13) imply the following bounds on asset prices:

$$\beta \leq \phi(s) \leq \beta + \xi. \quad (14)$$

The upper bound is the present value of the asset's cash flow, β , which we refer to as the "fundamental value" of the asset, augmented by the net utility of spending an additional unit of the asset in the DM, ξ . The lower bound is the "fundamental value" of the asset, β , since a buyer can always hold onto any unit of the asset and consume its cash flow at the end of $t = 1$. Assuming for now that $q < q^*$, so that $\xi > 0$, these first-order conditions imply that there are three categories of assets.

Liquid assets. For this type of asset, the feasibility constraint is binding, $\nu(s) > 0$, but the resalability constraint is slack, $\lambda(s) = 0$. Therefore, the asset price is equal to the upper bound, $\beta + \xi$. The asset is said to be perfectly liquid in the following sense: if the buyer holds an additional unit of the asset, he would spend it in the DM. Substituting the market clearing condition, $a(s) = A(s)$, and the price, $\phi(s) = \beta + \xi$, into the binding feasibility constraint and the slack resalability constraint, we obtain $d(s) = A(s) \leq \frac{k(s)}{\xi + \beta\sigma}$. This last inequality can be equivalently written as $\kappa(s) \geq \beta\sigma + \xi$, where $\kappa(s) \equiv k(s)/A(s)$ is the cost of fraud per unit of the asset.

partially liquid assets. For this type of asset, both the resalability and feasibility constraints bind, $\lambda(s) > 0$ and $\nu(s) > 0$. In equilibrium, a buyer spends all his holdings of the asset. However, if

he were to acquire an additional unit, he would choose not to spend it in the DM, for otherwise there would be a positive probability of the trade being rejected. The asset is thus said to be partially liquid and its price must be lower than the upper bound. From (10), $d(s) = A(s) = \frac{k(s)}{\phi(s) - \beta + \beta\sigma}$, which leads to $\phi(s) = \beta + \kappa(s) - \beta\sigma$, keeping in mind that $\kappa(s) = k(s)/A(s)$. The conditions $\lambda(s) = \xi + \beta - \phi(s) > 0$ and $\nu(s) = \phi(s) - \beta > 0$ can be written as $\beta\sigma < \kappa(s) < \beta\sigma + \xi$.

Illiquid assets. Lastly, there are assets for which the resalability constraint binds, $\lambda(s) > 0$, but the feasibility constraint is slack, $\nu(s) = 0$. In equilibrium the buyer does not spend a fraction of his asset holdings even though he is liquidity constrained. Therefore, the asset is said to be illiquid, and its price is equal to the lower bound, $\phi(s) = \beta$. The binding resalability constraint implies that $d(s) = \frac{k(s)}{\beta\sigma}$. Substituting this expression into the slack feasibility constraint, we obtain that $\kappa(s) \leq \beta\sigma$.

The next step is to determine ξ and verify that $q < q^*$. From the above, we have:

$$d(s) = \min \left[A(s), \frac{k(s)}{\beta\sigma} \right] = \theta(s)A(s), \quad \text{where} \quad \theta(s) = \min \left[1, \frac{\kappa(s)}{\beta\sigma} \right].$$

That is, the buyer either transfers all his holdings of asset s , or the maximum holding consistent with the resalability constraint and the no-arbitrage restriction that $\phi(s) \geq \beta$. Substituting the expression for $d(s)$ into the seller's binding participation constraint, (9), we obtain

$$q = L \equiv \sum_{s \in S} \theta(s)A(s). \tag{15}$$

The aggregate liquidity, L , is a weighted average of asset supplies, with endogenous weights depending on trading frictions and assets' recognizability characteristics.²¹ Given q , the convenience yield of liquid assets, ξ , is determined by (12). One can easily verify that, if $L < q^*$, the above asset prices, offer, and asset demands constitute a symmetric equilibrium. The condition $L < q^*$ means that the aggregate liquidity is not large enough to satiate buyers' liquidity needs, represented by q^* . If $L \geq q^*$, then the equilibrium has $q = q^*$ and $\phi(s) = \beta$ for all $s \in S$. Summarizing:

²¹This approach is consistent with a definition of the quantity of money suggested by Friedman and Schwartz (1970) as "the weighted sum of the aggregate value of all assets, the weights varying with the degree of moneyness." Our definition of aggregate liquidity is also related to the Divisia monetary aggregates (e.g., Barnett, Fisher, and Serletis, 1992). A key difference is that in our approach the weight assigned to an asset in order to calculate liquidity changes is not equal to its holding cost, which has normative implications that we discuss in Section 5.

Proposition 2 (The liquidity-return relationship) *There exists a unique symmetric equilibrium. If $L \geq q^*$, then $q = q^*$ and $\phi(s) = \beta$ for all $s \in S$. If $L < q^*$, then $q < q^*$, $\xi \equiv \beta\sigma [u'(q) - 1] > 0$. Letting $\underline{\kappa} \equiv \beta\sigma$, and $\bar{\kappa} \equiv \beta\sigma + \xi$, there are three categories of assets:*

1. *Liquid assets: for any $s \in S$, such that $\kappa(s) \geq \bar{\kappa}$,*

$$\phi(s) = \beta + \xi \quad (16)$$

$$\theta(s) = 1. \quad (17)$$

2. *Partially liquid assets: for any $s \in S$, such that $\kappa(s) \in (\underline{\kappa}, \bar{\kappa})$,*

$$\phi(s) = \beta + [\kappa(s) - \beta\sigma] \quad (18)$$

$$\theta(s) = 1. \quad (19)$$

3. *Illiquid assets: for any $s \in S$, such that $\kappa(s) \leq \underline{\kappa}$,*

$$\phi(s) = \beta \quad (20)$$

$$\theta(s) = \frac{\kappa(s)}{\beta\sigma} < 1. \quad (21)$$

The central implication of Proposition 2 is that, whenever there is a liquidity shortage, $L < q^*$, assets with identical cash flows can have different prices. See Figure 3 for a graphical representation of these price differences. This departure from the no-arbitrage principle is another formulation of the rate-of-return dominance puzzle, according to which monetary assets coexist with other assets with similar risk characteristics that generate a higher yield. In our model price differentials across assets are attributed to differences in the cost of fraud. An asset which is more recognizable—in the sense of not being sensitive to fraudulent activities—as captured by a high cost of fraud, is used more intensively to finance random spending opportunities. Relative to assets that are less recognizable, this asset generates some non-pecuniary liquidity services, $\nu(s) = \phi(s) - \beta$, also referred to as a convenience yield, and is sold at a higher price.²²

²²To see why the price differentials do not represent arbitrage opportunities, relax the short-selling constraint and assume that, in order to sell an asset he does not own, an agent has to borrow it from someone else in exchange for a fee, to be determined in equilibrium. The agent who borrows the asset can use it in the DM, but the agent who lends it cannot. The equilibrium remains unchanged, and the fee clearing the market for borrowing asset $s \in S$ is equal to its convenience yield, $\phi(s) - \beta$. Indeed, an agent who borrows a liquid or partially liquid asset must compensate the lender for his forgone liquidity services in the DM.

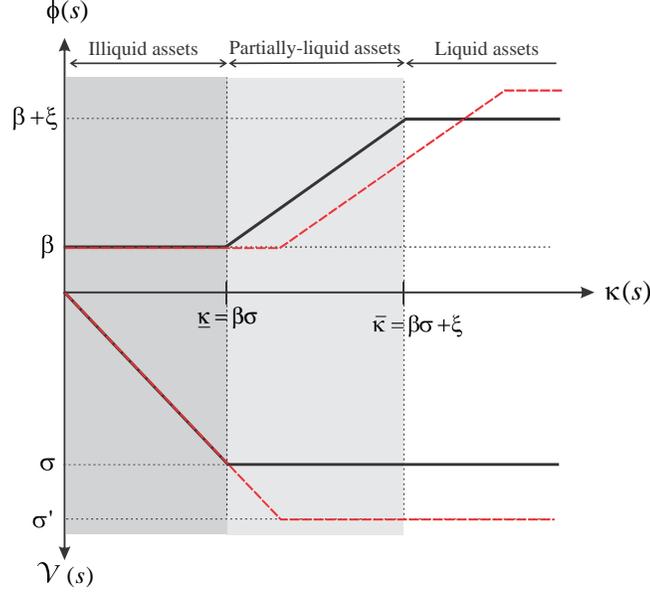


Figure 3: Liquidity structure of asset returns ($L < q^*$)

Krishnamurthy and Vissing-Jorgensen (2010, 2011) document the existence of convenience yields for Treasury securities and, to a lesser extent, highly-rated bonds. They argue that a safety-premium (which they view as distinct from a standard risk premium) is an important component of asset prices. Through the lens of our model, we can interpret this safety premium as the premium offered by assets that are highly recognizable and that are less sensitive to informational asymmetries and moral hazard considerations. Similarly, Vickery and Wright (2010) argue about the existence of a liquidity premium for agency mortgage-backed securities, which are better protected against the informational asymmetries that plague the process of securitization.

Proposition 2 also has insights for cross-sectional differences in transaction velocity, a standard measure of liquidity in monetary economies. In our model, transaction velocity in the DM is $\mathcal{V}(s) \equiv \frac{\sigma d(s)}{A(s)} = \sigma \theta(s)$. Proposition 2 predicts a positive relationship between the price of an asset and its velocity. The most liquid assets (i.e., any asset s such that $\kappa(s) \geq \bar{\kappa}$) trade at the highest price, and their velocity is maximum and equal to the frequency of spending opportunities in the DM, σ . Illiquid assets (i.e., any asset s such that $\kappa(s) < \underline{\kappa}$), however, have the highest rate of return, equal to the rate of time preference, and the lowest velocities, less than σ . This result is consistent

with the view that bonds that are used more intensely as collateral in OTC markets tend to have higher prices (Duffie, 1996).

In reality, a myriad of assets are not used as means of payment or collateral. This observation is consistent with our results if there is a mass of assets that do not circulate in the DM, $\theta(s) = 0$. From (21) such assets must be characterized by $\kappa(s) = 0$: these are assets for which agents have so little knowledge about their mere existence or attributes, that even simple, costless frauds can be deceptive.²³

5 Applications and Extensions

In this last section we apply our model of the liquidity structure of asset returns to analyze flight-to-liquidity phenomena and to assess the effectiveness of aggregate liquidity management policies. Moreover, we extend the model to an infinite time horizon in order to study time variations in liquidity premia.

5.1 Flights to liquidity

A flight to liquidity occurs when market participants seek to reallocate their portfolios toward highly liquid assets, which leads to a widening yield spread between liquid and less liquid assets.²⁴ In what follows, we apply our analysis on the liquidity structure of asset returns to identify the shocks that can generate a simultaneous increase in the prices of the most-liquid assets and a reduction in the prices of less-liquid ones—a phenomenon resembling a flight to liquidity.

According to our model, a flight to liquidity can be explained by an exogenous reduction in $k(s)$ for some initially liquid or partially liquid assets that make them become illiquid. For instance, agents might realize that some assets (e.g., MBS) can be subject to a broader set of fraudulent

²³That assets, or claims on those assets, can be counterfeited at no cost has been the standard explanation in monetary theory for why capital goods are illiquid, since Freeman (1985), and more recently, Lester, Postlewaite, and Wright (2011).

²⁴During the 1998 Russian-default crisis, many investors shifted their funds into the more liquid U.S. Treasury market, widening the yield spread between Treasury bonds and less-liquid debt instruments (Longstaff, 2004). Evidence also shows that, during the subprime crisis, the flight-to-quality was confined to AAA-rated bonds, and the illiquidity component of the rate of return of bonds with lower grades rose sharply (Longstaff, 2010; Dick-Nielsen, Feldhutter, and Lando, 2010).

practices than previously thought.²⁵ The resalability and velocity of these assets decrease, which causes aggregate liquidity and output to fall, and the liquidity premium on liquid assets, ξ , to increase.²⁶ The prices of partially liquid assets do not change, except for the ones that are characterized by a lower cost of fraud. In addition, an increase in the threat of fraud can shrink the set of liquid assets, while it expands the set of illiquid and partially liquid ones. Indeed, the threshold $\bar{\kappa} = \beta\sigma + \xi$ and the interval $\bar{\kappa} - \underline{\kappa} = \xi$ are increasing functions of the size of the liquidity premium on liquid assets. Therefore, during a flight to liquidity, market demand for assets is concentrated on a smaller set of highly recognizable assets.

An alternative explanation for a flight to liquidity is an increase in σ that formalizes an aggregate liquidity demand shock, e.g., an increase in counterparty risk, leading to an increase in the demand for collateral for OTC transactions.²⁷ From (16) and (18) when σ increases the prices of liquid assets rise, whereas the prices of partially liquid assets fall. The increase in the prices of liquid assets occurs due to two effects going in the same direction. There is a direct effect according to which liquid assets are used more often as collateral or means of payment, which raises their liquidity value. The indirect effect is to reduce aggregate liquidity: from (15), an increase in σ lowers the weights of illiquid assets in L , which reduces the output in bilateral matches and makes liquid assets even more useful; i.e., the term $\beta[u'(q) - 1]$ in (12) goes up. For partially liquid assets the increase in σ has the additional markedly different effect of exacerbating fraud incentives. As a result, their prices have to fall so that their resalability constraints hold, re-establishing buyers' incentives to bring genuine assets. As shown in Figure 3, the set of illiquid and partially liquid assets expands (because $\underline{\kappa}$ increases with σ and $\bar{\kappa} - \underline{\kappa}$ increases with ξ) while the set of liquid assets shrinks (because $\bar{\kappa}$ increases in ξ).

²⁵For some prominent economists this type of shock is a central explanation for the financial crisis of 2008. In an interview to the Wall Street journal (09/24/2011), Robert Lucas argued that "the shock came because complex mortgage-related securities minted by Wall Street and certified as safe by rating agencies had become part of the effective liquidity supply of the system. All of a sudden, a whole bunch of this stuff turns out to be crap".

²⁶Some recent studies (e.g., Ajello, 2010; Shi, 2011) formulate the hypothesis that recessions are driven by liquidity shocks formalized by a reduction in the exogenous resalability of some assets. In contrast to our approach, these models have the counterfactual implication that the prices of the assets that become more difficult to resell increase.

²⁷Suppose, for instance, that a fraction σ_u of the trades in the DM can be financed with unsecured debt (e.g., because commitment/enforcement is available in those meetings) while a fraction σ_s of the trades require collateral to be posted because of counterparty risk (e.g., sellers in those meetings cannot commit or cannot be forced to repay their debt.) An increase in counterparty risk can be formalized as an increase in σ_s such that $\sigma_s + \sigma_u$ is unchanged.

5.2 Liquidity management

In this section we use our model to study the effectiveness of policies aimed at managing the supply of liquidity in the economy. These policies can take the form of open-market operations by the central bank, which are intended to substitute liquid assets for less-liquid ones, or regulatory measures that reduce the threat of frauds and relax resalability constraints for some assets.

Measuring the social value of assets' liquidity services. Much of the analysis that follows is based on the following theoretical observation. In competitive models with reduced-form demand for liquidity (e.g., cash-in-advance or money-in-the-utility function), the convenience yield of an asset not only measures the *marginal private value* of its liquidity services, but also its *marginal social value*.²⁸ In our model this property holds true for illiquid and liquid assets, but fails to hold for partially liquid assets.

The marginal social value of the liquidity services provided by a unit of asset s is $\frac{\partial L}{\partial A(s)}\xi$, which is equal to ξ for liquid *and* partially liquid assets, and 0 for illiquid assets. Therefore, the convenience yield of partially liquid assets, $\phi(s) - \beta < \xi$, underestimates the true marginal social value of their liquidity services. The reason for this discrepancy is that an increase in the price of an asset reduces its demand in two ways: by raising the holding cost, $\phi(s) - \beta$, and by tightening the resalability constraint. The latter effect creates a negative “pecuniary externality,” which can depress asset prices below the marginal social value of the asset’s liquidity services.²⁹ As we show below, this observation implies that liquidity management policies targeting partially liquid assets can be welfare reducing, because they underestimate these assets’ true contribution to aggregate liquidity. By contrast, when targeting illiquid assets, the same policies are welfare improving.

Open-market purchases. Central banks routinely engage in aggregate liquidity management, by issuing (or withdrawing) reserves, the most liquid assets, in exchange for Treasuries and, in recent

²⁸This logic is underlying the calculation for the welfare cost of inflation in Lucas (2000), the measure of the liquidity services provided by Treasuries in Krishnamurthy and Vissing-Jorgensen (2010), and Barnett, Fisher, and Serletis’s (1992) definition of Divisia monetary aggregates.

²⁹By contrast, with the exogenous proportional resalability constraint, there is no such pecuniary externality, and asset convenience yields coincide with the marginal social value of the asset’s liquidity services. See Supplementary Appendix E.

years, a wider range of less liquid assets, including agency bonds and mortgage-backed securities. Consider a policy-maker in the CM, who sells a quantity, $\Delta A(s)$, of some liquid asset s from his portfolio, and simultaneously purchases a quantity, $\Delta A(s')$, of some other asset s' . A small open-market operation has a small effect on prices, so that the budget constraint of the policy-maker is, to a first-order approximation, $\phi(s)\Delta A(s) + \phi(s')\Delta A(s') = 0$. The welfare effect of such a policy is

$$\Delta L \times \xi = \left[\frac{\partial L}{\partial A(s)} \Delta A(s) + \frac{\partial L}{\partial A(s')} \Delta A(s') \right] \xi = \left[1 - \frac{\partial L}{\partial A(s')} \frac{\phi(s)}{\phi(s')} \right] \Delta A(s) \times \xi.$$

Suppose first that $\kappa(s') > \bar{\kappa}$, so both s and s' are liquid assets. Then, $\phi(s) = \phi(s')$, $\frac{\partial L}{\partial A(s')} = 1$, and $\Delta L = 0$. Such an open-market operation is irrelevant: it does not change aggregate liquidity and welfare, and hence it has no effect on output and asset prices. So liquidity management has real effects only if it involves assets with different degrees of liquidity.

Suppose next that $\kappa(s') < \underline{\kappa}$, asset s' is illiquid. In this case aggregate liquidity does increase because the purchase of illiquid assets has no consequence on aggregate liquidity; i.e., $\Delta L \times \xi = \Delta A(s) \times \xi > 0$. Thus, welfare increases, the price of liquid assets decreases, and the price of illiquid assets is unaffected.

Finally, suppose that $\kappa(s') \in (\underline{\kappa}, \bar{\kappa})$; i.e., asset s' is partially liquid. Then, $\phi(s') < \phi(s)$ and $\Delta L \times \xi = \left[1 - \frac{\phi(s)}{\phi(s')} \right] \Delta A(s) \times \xi < 0$, implying that such a policy *reduces* aggregate liquidity and welfare. The intuition is in line with our earlier observation: while partially liquid and liquid assets have different prices, they contribute *equally* to aggregate liquidity. At the same time, because it has a higher price, one share of a liquid asset buys more than one share of a partially liquid one. Thus a balanced-budget open-market operation ends up syphoning out more liquidity than it is injecting in; i.e., aggregate liquidity is reduced. The welfare effect of this open-market operation is of the opposite sign of the yield difference between the asset that is withdrawn and the asset that is injected, and the prices of both assets s and s' increase.

The results above can help interpret some of the findings in Krishnamurthy and Vissing-Jorgensen (2011) regarding the effect of quantitative easing. They find that the purchases of Treasuries, agency bonds, and highly-rated corporate bonds in exchange for reserves led to a drop in interest rates but it did not affect the yields on relatively illiquid assets (Baa corporate bonds). This finding is consistent with our results if we interpret Baa corporate bonds as illiquid assets,

Treasuries and highly rated bonds as partially liquid, and reserves as fully liquid. Furthermore, according to our findings, the drop in interest rates indicates that quantitative easing reduced liquidity and welfare.

Regulatory measures. Some of the leading regulatory measures of the Dodd-Frank Act aim to curb fraud incentives in the securitization industry.³⁰ One of these measures is a requirement for securitizers to retain at least 5 percent of the credit risk they originate. Importantly, some asset-backed securities, deemed of higher quality, are exempted from this requirement. In this section, we study the optimality and welfare impact of retention requirements. We show that the regulator faces a trade-off between the role these requirements play as a discipline mechanism and the distortion they introduce by increasing the costs of holding assets. We demonstrate that the first effect dominates for illiquid assets, while the second effect dominates for partially liquid and liquid assets. Hence, our model suggests that retention requirements should be confined to the least liquid assets, i.e., the ones more susceptible to fraud.

Under a retention requirement policy, a buyer who wishes to transfer $d(s)$ units of asset s in the DM must hold $1 + \rho(s)$ units of the asset; i.e., $d(s) \leq \frac{a(s)}{1+\rho(s)}$, where $\rho(s)$ is the retention rate associated with asset s . The policy imposes that the asset kept in retention is the exact same asset as the one transferred in a match, i.e., if the asset transferred is fraudulent, so is the asset in retention.³¹ The cost of producing $d(s)$ units of fraudulent asset is of the form $k_f(s) + k_v(s)d(s)$, where the variable cost component, $k_v(s)d(s)$, was introduced to provide a channel through which the regulatory measure can reduce agents' incentive to commit fraud. We let $k_f(s) > 0$ and take $k_v(s)$ to be small enough so that, as before, equilibrium offers are accepted with probability one (see Supplementary Appendix B). The resalability constraint of asset s becomes:

$$k_f(s) + k_v(s) [1 + \rho(s)] d(s) \geq [\phi(s) - \beta] [1 + \rho(s)] d(s) + \beta \sigma d(s). \quad (22)$$

The left side of (22) is the cost of fraud on $d(s)$ units of asset s . If $k_v(s) > 0$, then policy increases

³⁰The Dodd-Frank Act, enacted in July 2010 in response to the 2007-08 financial crisis, institutes a wide array of new regulations for the financial services industry.

³¹In the context of securitization (see Supplementary Appendix H), a retention requirement means that the securitizer (represented in the model by the buyer) needs to retain assets from the same issue of asset-based securities he is offering to the general public (represented in the model by the seller).

the cost of fraud and, therefore, reduces fraud incentives. The right side of (22) is the cost of holding $[1 + \rho(s)]d(s)$ genuine units of asset s . Thus, if the asset is liquid or partially liquid, $\phi(s) - \beta > 0$, the retention requirement generates a distortion by increasing the effective holding cost of the asset.

In Supplementary Appendix B, we solve for equilibrium following the same steps as before. We show that a retention requirement has asymmetric effects on the resalability of an asset, depending on its liquidity status. For illiquid assets, equilibrium resalability becomes:

$$\theta(s) = \frac{k_f(s)/A(s)}{\beta\sigma - k_v(s)[1 + \rho(s)]}. \quad (23)$$

It is an *increasing* function of $\rho(s)$ because when $\phi(s) = \beta$ retention rates raise the cost of committing fraud but do not affect assets' effective holding costs. For liquid and partially liquid assets resalability becomes:

$$\theta(s) = \frac{1}{1 + \rho(s)}, \quad (24)$$

which is a *decreasing* function of $\rho(s)$. Thus, for liquid or partially liquid assets, the distortionary effect of retention rates dominates the incentive effect, reducing velocity and welfare. In the case of liquid assets, this result is straightforward since the threat of fraud is not a binding constraint. In the case of partially liquid assets, retention requirements have the partial equilibrium effect of relaxing resalability constraints. But this effect simultaneously increases the demand for partially liquid assets. Therefore, for asset markets to clear, the prices of partially liquid assets must increase, tightening back the resalability constraints and eliminating the positive incentive effect of the retention policy. Taken together, the above results suggest that retention requirements should target illiquid assets. Other, more recognizable assets, should be exempted, in line with the prescriptions of the Dodd Frank Act.

5.3 Dynamics of liquidity premia

This section provides a dynamic extension of our static model. We show that the prices of all liquid assets covary with a common liquidity premium.³² This common liquidity premium can be

³²References on the empirical literature on the co-movements of liquidity across assets are included in Acharya and Pedersen (2005) who develop an asset pricing model in which a per-share cost of selling securities can vary over time.

the subject of self-fulfilling fluctuations, creating excess volatility in the price of liquid assets. The prices of illiquid and partially liquid assets are, however, immune to such fluctuations.

Given agents' quasilinear preferences, it is straightforward to introduce an infinite time horizon using the setup of Lagos and Wright (2005). Time is indexed by $t \in \mathbb{N}$. Each period is divided into two subperiods, a DM followed by a CM. Each unit of the asset s pays off a dividend normalized to one unit of the numéraire at the beginning of each CM. The technology to produce fraudulent assets in period $t - 1$ becomes obsolete in period t , and all fraudulent assets produced in period t are confiscated by the government before agents enter the CM of period t .³³

As shown in supplementary Appendix D, these assumptions allow us to apply the analysis of the static model, where the terminal value of the asset is equal to the cum-dividend value, $1 + \phi_t(s)$, of reselling the asset in the CM in period t . Focusing on equilibrium with $q_t < q^*$, Proposition 2 generalizes as follows.³⁴ There are three classes of assets, of which the prices solve:

$$\phi_{t-1}(s) = [1 + \phi_t(s)] \times \begin{cases} \beta + \xi_t & \text{if } \kappa_t(s) \geq \bar{\kappa}_t \\ \beta + [\kappa_t(s) - \beta\sigma] & \text{if } \kappa_t(s) \in (\underline{\kappa}, \bar{\kappa}_t) \\ \beta & \text{if } \kappa_t(s) \leq \underline{\kappa} \end{cases}, \quad (25)$$

where $\kappa_t(s) \equiv \frac{k(s)}{[1 + \phi_t(s)]A(s)}$, $\underline{\kappa} \equiv \beta\sigma$, $\bar{\kappa}_t \equiv \xi_t + \beta\sigma$, and

$$\xi_t = \beta\sigma [u'(q_t) - 1] \quad (26)$$

$$q_t = L = \sum_{s \in S} \theta_t(s) [1 + \phi_t(s)] A(s), \quad \text{where } \theta_t(s) = \min \left[1, \frac{\kappa_t(s)}{\beta\sigma} \right]. \quad (27)$$

The equilibrium equations are the same as in the static model, but with an endogenous terminal value of $1 + \phi_t(s)$. This difference is substantial because expectations of future liquidity premia, capitalized in $\phi_t(s)$, feed back into the current liquidity premium, $\phi_{t-1}(s) - \beta [1 + \phi_t(s)]$.³⁵

We now characterize equilibria in a neighborhood of the unique steady state, $\langle \{\bar{\phi}(s)\}, \bar{q}, \bar{\xi} \rangle$. In such a neighborhood, the sets of liquid, partially liquid, and illiquid assets do not change. Moreover,

³³This assumption, borrowed from Nosal and Wallace (2007), is made for tractability to prevent fraudulent assets from circulating across periods.

³⁴If aggregate liquidity is abundant, there exists a unique equilibrium in which the resalability constraint and the feasibility constraint do not bind at any date, $q_t = q^*$, and each asset is priced at its fundamental value; i.e., $\phi_t(s) = 1/r$.

³⁵This effect, as is well known, can lead to an equilibrium in which an asset has positive value even if it pays no dividend; i.e., a positive liquidity premium can be a self-fulfilling phenomenon. In Supplementary Appendix D we consider such an economy with fiat money.

from (25), one can verify that $\frac{d\phi_{t-1}(s)}{d\phi_t(s)} \in [0, 1)$ for all $\kappa_t(s) < \bar{\kappa}_t$, so that the prices of illiquid and partially liquid assets are equal to their steady-state values in any dynamic equilibrium. This need not be the case for liquid assets. To see this point, let us linearize the equilibrium equations near the steady state. We obtain, from (25), that the price of liquid assets solves:

$$\hat{\phi}_{t-1}(s) = (\beta + \bar{\xi}) \hat{\phi}_t(s) + [\bar{\phi}(s) + 1] \hat{\xi}_t, \quad (28)$$

where $\hat{\phi}_t(s) \equiv \phi_t(s) - \bar{\phi}(s)$ and $\hat{\xi}_t \equiv \xi_t - \bar{\xi}$. The first term on the right side of (28) is the discounted value of the future price of the asset, with the discount rate augmented by a liquidity premium; the second term captures the change in the liquidity premium. Linearizing (26) and (27) in the neighborhood of the steady state:

$$\hat{\xi}_t = \beta \sigma u''(\bar{q}) \hat{q}_t \quad \text{where} \quad \hat{q}_t = \sum_{s: \kappa(s) \geq \bar{\kappa}} \hat{\phi}_t(s) A(s), \quad (29)$$

and $\hat{q}_t \equiv q_t - \bar{q}$. From (29), the size of the liquidity premium, relative to its steady-state value, depends negatively on changes in the market capitalization of liquid assets.

Multiplying both sides of (28) by $A(s)$ and taking the sum over all liquid assets, we obtain

$$\hat{\xi}_{t-1} = \gamma \hat{\xi}_t, \quad \text{where} \quad \gamma = \beta + \bar{\xi} + \beta \sigma u''(\bar{q}) \sum_{s: \kappa(s) \geq \bar{\kappa}} [\bar{\phi}(s) + 1] A(s). \quad (30)$$

The nature of the dynamics depends on γ . If $\gamma > -1$, then $\hat{\xi}_t = \bar{\xi}$ for all t , and the liquidity premium is constant over time. If $\gamma < -1$, in contrast, there exists a continuum of equilibria indexed by the initial value of $\hat{\xi}$ in the neighborhood of zero that converges to the steady state. Along these equilibria, $\hat{\xi}_t$ alternates between positive and negative values. The price of liquid assets covaries and exhibits excessive volatility relative to fundamentals, whereas the prices of partially liquid assets and illiquid ones remain constant. As can be seen from (29), the fluctuating liquidity premium is a self-fulfilling phenomenon. If agents anticipate that the liquidity premium will be high in the future so that $\hat{\phi}_t > 0$, then aggregate liquidity and output are high, $\hat{q}_t > 0$. But, at the margin, agents do not value much the asset's liquidity services, and so the current liquidity premium is low, $\hat{\xi}_t < 0$. If the constant multiplying $\hat{\xi}_t$ in (28) is large enough, then $\hat{\phi}_{t-1} < 0$. The same reasoning implies that the liquidity premium one period before will be high, $\hat{\xi}_{t-1} > 0$, and the fluctuations will continue.

6 Conclusion

In this paper we have proposed a theory of the cross-sectional distribution and time-variation of liquidity premia by taking seriously the possibility of asset fraud in an economy with limited commitment and enforcement. We have shown the emergence of asset-specific resalability constraints that take the form of upper bounds on the transfer of assets in OTC market trades. These bounds are not invariant to policy shifts (e.g., the composition of asset supplies and regulation on assets' retention requirements), and they depend on some characteristics of the assets such as their vulnerability to fraud, as well as the frequency of trading opportunities. Our model generates a liquidity structure of asset returns based on a three-tier classification of assets. This classification is relevant for the comparative statics of asset prices, the dynamics of liquidity premia, explanations of flight-to-liquidity phenomena, and the analysis of open-market operations and regulations of the OTC market.

References

- [1] Ajello, Andrea (2010). “Financial Intermediation, Investment Dynamics and Business Cycle,” manuscript.
- [2] Acharya, Viral V. and Lasse Heje Pedersen (2005). “Asset Pricing with Liquidity Risk,” *Journal of Financial Economics*, Elsevier 77, 375-410.
- [3] Amihud, Yakov and Haim Mendelson (1986). “Asset Pricing and the Bid Ask Spread,” *Journal of Financial Economics* 17, 223-249.
- [4] Barnett, William A., Douglas Fisher, and Apostolos Serletis (1992). “Consumer Theory and the Demand for Money,” *Journal of Economic Literature* 30, 2086-2119.
- [5] Chordia, Tarun, Huh Sahn-Wook and Avanidhar Subrahmanyam (2009). “Theory-Based Illiquidity and Asset Pricing,” *Review of Financial Studies* 22, 3629-3668.
- [6] Dick-Nielsen, Jens, Peter Feldhutter, and David Lando (2010). “Corporate Bond Liquidity Before and After the Onset of the Subprime Crisis,” manuscript.
- [7] Duffie, Darel (1996). “Special Repo Rates” *Journal of Finance* 51, 493-526.
- [8] Duffie, Darrell, Nicolae Garleanu and Lasse Heje Pedersen (2005). “Over-the-counter Markets,” *Econometrica* 73, 1815-1847.
- [9] Freeman, Scott (1985). “Transactions Costs and the Optimal Quantity of Money,” *Journal of Political Economy* 93, 146-157.
- [10] Green, Edward and Warren Weber (1996). “Will the New \$100 Bill Decrease Counterfeiting?” *Federal Reserve Bank of Minneapolis Quarterly Review* 20, 3-10.
- [11] Guerrieri, Veronica and Robert Shimer (2011). “Dynamic Adverse Selection: A Theory of Illiquidity, Fire Sales, and Flight to Quality,” mimeo.
- [12] Guerrieri, Veronica, Robert Shimer, and Randall Wright (2010). “Adverse Selection in Competitive Search Equilibrium,” *Econometrica* 78, 1823-1862.

- [13] Holmstrom, Bengt and Jean Tirole (2011). *Inside and Outside Liquidity*, Princeton, Princeton University Press.
- [14] Hopenhayn, Hugo A. and Ingrid M. Werner (1996). “Information, Liquidity, and Asset Trading in a Random Matching Game,” *Journal of Economic Theory* 68, 349-379.
- [15] In, Younghwan and Julian Wright (2011). “Signaling Private Choices,” manuscript.
- [16] Keys, Benjamin J., Tanmoy K. Mukherjee, Amit Seru, and Vikrant Vig (2010). “Did securitization lead to lax screening? Evidence from subprime loans,” *The Quarterly Journal of Economics* 125, 307-62.
- [17] Kiyotaki, Nobuhiro and John Moore (2001). “Liquidity, Business Cycles and Monetary Policy,” manuscript.
- [18] Kiyotaki, Nobuhiro and John Moore (2005). “Liquidity and asset prices,” *International Economic Review* 46, 317-349.
- [19] Kohlberg, Elon and Jean-Francois Mertens (1986). “On the strategic stability of equilibria,” *Econometrica* 54, 1003-1037.
- [20] Krishnamurthy, Arvind and Annette Vissing-Jorgensen (2010). “The aggregate demand for Treasury debt,” manuscript.
- [21] Krishnamurthy, Arvind and Annette Vissing-Jorgensen (2011). “The Effects of Quantitative Easing on Interest Rates,” manuscript.
- [22] Lagos, Ricardo (2010). “Asset prices and liquidity in an exchange economy,” *Journal of Monetary Economics* 57, 913-930.
- [23] Lagos, Ricardo and Randall Wright (2005). “A unified framework for monetary theory and policy analysis,” *Journal of Political Economy* 113, 463–484.
- [24] Lester, Benjamin, Andrew Postlewaite and Randall Wright (2011). “Information, Liquidity, Asset Prices, and Monetary Policy,” *Review of Economic Studies*, forthcoming.

- [25] Longstaff, Francis A. (2004). “The Flight-to-Liquidity Premium in U.S. Treasury Bond Prices,” *Journal of Business* 77, 511-526.
- [26] Longstaff, Francis A. (2010). “The subprime credit crisis and contagion in financial markets,” *Journal of Financial Economics* 97, 436-450.
- [27] Lucas, Robert E. (2000). “Inflation and Welfare,” *Econometrica* 68, 247-274.
- [28] Friedman, Milton and Anna J. Schwartz (1970). *Monetary statistics of the United States: estimates, sources, methods*, NBER Studies in business cycles, no. 20, New York: National Bureau of Economic Research.
- [29] Mihm, Stephen (2007). *A nation of counterfeiters: Capitalists, con men, and the making of the United States*, Cambridge, Harvard University Press.
- [30] Nosal, Ed and Neil Wallace (2007). “A model of (the threat of) counterfeiting,” *Journal of Monetary Economics* 54, 229-246.
- [31] Purnanandam, Amiyatosh K. (2009). “Originate-to-Distribute Model and the Subprime Mortgage Crisis,” manuscript.
- [32] Rocheteau, Guillaume (2009). “A Monetary Approach to Asset Liquidity,” Working Papers 09-01, Federal Reserve Bank of Cleveland.
- [33] Sargent, Thomas J. and Francois R. Velde (2002). *The Big Problem of Small Change*, Princeton: Princeton University Press.
- [34] Schreft, Stacey L. (2007). “Risks of identity theft: Can the market protect the payment system?” *Federal Reserve Bank of Richmond, Economic Review*, Fourth Quarter, 5-40.
- [35] Shi, Shouyong (2008). “Efficiency improvement from restricting the liquidity of nominal bonds,” *Journal of Monetary Economics* 55, 1025– 1037.
- [36] Shi, Shouyong (2011). “Liquidity, Assets and Business Cycles,” manuscript.

- [37] Snowden, Kenneth A. (2010) “Covered Farm Mortgage Bonds in the Late Nineteenth Century U.S,” NBER Working Papers 16242
- [38] Vickery, James and Joshua Wright (2010). “TBA Trading and Liquidity in the Agency MBS Market,” manuscript.
- [39] Wallace, Neil (1998). “A dictum for monetary theory,” FRB Minneapolis Quarterly Review 22: 20-26.
- [40] Wallace, Neil (2000). “A model of the liquidity structure based on asset indivisibility,” Journal of Monetary Economics 45, 55-68.
- [41] Weill, Pierre-Olivier (2008). “Liquidity premia in dynamic bargaining markets,” Journal of Economic Theory 140, 66-96.
- [42] Williamson, Stephen and Randall Wright (1994). “Barter and Monetary Exchange under Private Information,” American Economic Review 84, 104-123.

A Proof of Proposition 1

We define an outcome of the game as an offer $(q, \{d(s)\})$ made by the buyer, probabilities $\{\eta(s)\}$ of bringing genuine assets, and a probability $\pi \in [0, 1]$ that the seller accepts the offer. Let us consider the auxiliary problem of choosing an outcome in order to maximize the expected utility of a buyer,

$$-\sum_{s \in S} \left\{ k(s) [1 - \eta(s)] + [\phi(s) - \beta] \eta(s) d(s) \right\} + \sigma \beta \pi \left[u(q) - \sum_{s \in S} \eta(s) d(s) \right], \quad (31)$$

subject to the constraint that the probabilities π and $\{\eta(s)\}$ are the basis of an equilibrium in the sub-game following offer $(q, \{d(s)\})$; that is:

$$\pi \in \arg \max_{\hat{\pi} \in [0,1]} \hat{\pi} \left\{ -q + \sum_{s \in S} \eta(s) d(s) \right\} \quad (32)$$

$$\eta(s) \in \arg \max_{\hat{\eta} \in [0,1]} \hat{\eta} \left\{ k(s) - [\phi(s) - \beta + \beta \sigma \pi] d(s) \right\}, \text{ for all } s \in S. \quad (33)$$

We start by showing that:

Claim 1 *Any solution of the auxiliary problem has the property that $\eta(s) = 1$ and $q = \sum_{s \in S} d(s)$.*

Proof. Consider first any feasible outcome $\langle q, d, \eta, \pi \rangle$ such that $\eta(s_0) < 1$ for some s_0 . If $\eta(s_0) = 0$, then consider the alternative outcome, $\langle q', d', \eta', \pi' \rangle$, such that: (i) $q' = q$, $d'(s) = d(s)$ for all $s \neq s_0$, $d'(s_0) = 0$; (ii) $\eta'(s) = \eta(s)$ for all $s \neq s_0$ and $\eta'(s_0) = 1$; (iii) $\pi' = \pi$. The incentive constraint of the seller, (32), is satisfied since it only depends on the product $\eta(s)d(s)$. The incentive constraint of the buyer, (33), is obviously satisfied for $s \neq s_0$. For $s = s_0$ we have $k(s_0) > [\phi(s) - \beta + \beta \sigma] d'(s_0) = 0$ and so $\eta'(s_0) = 1$ is optimal for the buyer. One can then verify that, with this alternative outcome, the expected utility of the buyer increases by $k(s_0) > 0$.

Next, consider any feasible outcome such that $\eta(s_0) \in (0, 1)$: the buyer is indifferent between counterfeiting asset s_0 or not. We then increase $\eta(s_0)$ by $\varepsilon \in (0, 1]$ and q by $\varepsilon d(s_0)$, which is positive since the incentive constraint of the buyer, (33), binds. The incentive constraint of the seller, (32) is satisfied because his payoff conditional on accepting the offer does not change. Because the buyer is indifferent between counterfeiting asset s_0 or not, an increase in $\eta(s_0)$ affects neither his payoff, (31), nor his incentive constraint, (33). The corresponding increase in q , however, increases his payoff strictly.

Next, consider any feasible outcome $\langle q, d, \eta, \pi \rangle$ such that $\eta(s) = 1$ for all s , but $q < \sum_{s \in S} d(s)$. Then the alternative outcome with $q' = \sum_{s \in S} d(s) > q$, $\eta'(s) = 1$, and $\pi' = \pi$, increases the expected payoff to the buyer by $\sigma\beta\pi [u(q') - u(q)] > 0$ and satisfies all the constraints. ■

This claim implies that we can rewrite the auxiliary problem as

$$\max_{q, d, \pi} - \sum_{s \in S} [\phi(s) - \beta] d(s) + \sigma\beta\pi [u(q) - q] \quad (34)$$

$$\text{s.t.} \quad \sum_{s \in S} d(s) - q = 0 \quad (35)$$

$$k(s) \geq [\phi(s) - \beta + \beta\sigma\pi] d(s), \text{ for all } s \in S. \quad (36)$$

The second condition is the first-order necessary and sufficient condition for (33) evaluated at $\eta(s) = 1$. Next, we show:

Claim 2 *Any solution of the auxiliary problem, (31)-(33), has the property that $u'(q) \geq 1$ and $\pi = 1$.*

Proof. The first claim holds because otherwise one could reduce the quantity produced, increase the expected utility of the buyer, and satisfy all the constraints. To prove the second claim suppose, towards a contradiction, that $\pi < 1$. Note first that the value of the auxiliary problem must be positive: a small offer $q' = d'(s_0) > 0$, $d'(s) = 0$ for $s \neq s_0$, and $\pi' = 1$ yields a positive payoff. This implies that both $q > 0$ and $\pi > 0$. Moreover, at least one of the incentive constraints, (36), must be binding since otherwise one could increase π without violating any of the incentive constraints, and improve the objective. Let $S_B \subseteq S$ be the set of binding IC constraints. Since $[\phi(s) - \beta + \beta\sigma\pi] d(s) = k(s)$ for all $s \in S_B$, it follows that $d(s) > 0$. Now consider the following variational experiment: increase π by some small ε and decrease the payments $d(s)$, for all $s \in S_B$, so that all the incentive constraints continue to bind. Up to second-order terms, the decrease in $d(s)$ is equal to $m(s)\varepsilon$, where

$$m(s) \equiv \frac{\beta\sigma d(s)}{\phi(s) - \beta + \beta\sigma\pi},$$

is the marginal rate of substitution between π and $d(s)$ in the IC constraint for asset $s \in S_B$. Lastly, to satisfy the participation constraint, the decrease in q must be, up to second order terms,

$\sum_{s \in S_B} m(s)\varepsilon$. The change in the buyer's expected utility is, up to second-order terms, equal to $\Delta U \times \varepsilon$, where

$$\begin{aligned} \Delta U &= \sum_{s \in S_B} [\phi(s) - \beta] m(s) - \beta\sigma\pi [u'(q) - 1] \sum_{s \in S_B} m(s) + \beta\sigma [u(q) - q] \\ &> \sum_{s \in S_B} [\phi(s) - \beta] m(s) + \beta\sigma [u'(q) - 1] \left[q - \pi \sum_{s \in S_B} m(s) \right] \\ &\geq \sum_{s \in S_B} [\phi(s) - \beta] m(s) + \beta\sigma [u'(q) - 1] \sum_{s \in S_B} [d(s) - \pi m(s)] = \sum_{s \in S_B} [\phi(s) - \beta] m(s) u'(q) \geq 0, \end{aligned}$$

where we move from the first line to the second line using $u(q) - q > q [u'(q) - 1] \geq 0$ (the equality is strict because of two facts: first, $u(q)$ is strictly concave and second, $q > 0$, since the value of the auxiliary problem is positive); from the second line to the third line using $q \geq \sum_{s \in S_B} d(s)$; and from the third to the fourth line by noting that $d(s) - \pi m(s) = [\phi(s) - \beta] m(s) / (\beta\sigma)$. ■

From Claims 1-2 and the result according to which $a(s) \geq \chi(s)d(s)$ if $\phi(s) = \beta$, and $a(s) = \chi(s)d(s)$ if $\phi(s) > \beta$, it follows that the auxiliary problem, (31)-(33), reduces to the maximization problem of Proposition 1, (8)-(11). Now we note that the solution to the auxiliary problem is an upper bound on the value of the buyer in any equilibrium of the game. Let $(\tilde{q}, \{\tilde{d}(s)\})$ be one solution of the auxiliary problem. Because, as argued above, the value of the auxiliary problem is positive, it must satisfy $\tilde{q} > 0$ and $\tilde{d}(s) > 0$ for some $s \in S$. Consider, for any $\varepsilon > 0$ small enough, the offer $d^\varepsilon(s) = \max\{\tilde{d}(s) - \varepsilon, 0\}$ and $q^\varepsilon = \tilde{q} - (S+1)\varepsilon$. By construction, this offer is such that $[\phi(s) - \beta + \beta\sigma] d^\varepsilon(s) < k(s)$, and $q^\varepsilon < \sum_{s \in S} d^\varepsilon(s)$. Thus, $\pi = 1$ and $\eta(s) = 1$ is the unique equilibrium in the subgame following $(q^\varepsilon, \{d^\varepsilon(s)\})$. By letting ε go to zero and making the above offer, the buyer can achieve a value arbitrarily close to that of the auxiliary problem. Therefore, in any equilibrium, the buyer's value must be equal to that of the auxiliary problem. Moreover, any equilibrium outcome satisfies (32) and (33). Therefore, any equilibrium outcome must solve the auxiliary problem.

Supplementary Appendix to “Liquidity and the Threat of Fraudulent Assets”

This supplementary appendix establishes results to complement and extend the main analysis of the paper. Each section is self-contained and can be read separately.

- Appendix B, page 37, derives the equilibrium with retention requirements and variable costs of fraud, that is used in the discussion of regulatory measures, in Section 5.2 of the paper.
- Appendix C, page 44, shows that with variable costs of fraud one can generate situations where sellers accept buyers’ offer with probability less than one, creating rationing at the extensive margin.
- Appendix D, page 46, provides the detailed derivation of the dynamic equilibrium presented in Section 5.3.
- Appendix E, page 52, shows that our model has different implications than a model with exogenous resalability constraints, à-la Kiyotaki and Moore.
- Appendix F, page 63, shows that our model has different implications than corporate finance-based models of liquidity constraints.
- Appendix G, page 65, provides an over-the-counter market version of our model, where agents buy assets in the CM in order to collateralize fixed-for-floating risk-sharing contracts in the DM.
- Appendix H, page 75, shows how to embed our model into the standard Capital Asset Pricing Model (CAPM). This model provides a simple representation of the securitization process: buyers are interpreted as professional investors who buy risky assets in the CM, package them into portfolios and re-sell them to consumers in the DM. But DM transactions are subject to the threat of fraud.

B Proofs omitted in Section 5.2

This section derives the equilibrium with retention requirements and variable costs of fraud, that is used in the discussion of regulatory measures, in Section 5.2.

In all what follows we assume that the cost of producing $d(s)$ units of fraudulent assets is $k_f(s) + k_v(s)d(s)$; i.e., it includes a variable cost. We consider collections of retention requirements $\rho \equiv \{\rho(s)\}$ belonging to the set $R = [0, \bar{\rho}]^S$, where $\bar{\rho}$ such that

$$(1 + \bar{\rho}) \max_{s \in S} \left\{ \frac{k_f(s)}{A(s)} \right\} > \beta\sigma.$$

Following the same steps as in the proof of Proposition 1 (Appendix A, page 33) we find that the participation constraint of the seller is binding, and that the buyer brings genuine assets with probability one. With these preliminary results in mind, we prove existence of equilibrium in two steps. First, given small enough variable costs, and assuming that the seller accepts the offer with probability $\pi = 1$, we solve for the equilibrium allocation and prices in Section B.1. Second, given that these prices prevail, and given small enough variable costs, we prove in Section B.2 that a buyer always finds it optimal to make an offer that is accepted with probability one. In Section B.3, we fix some small enough k_v and solve for the optimal retention requirement in $[0, \bar{\rho}]^S$.

B.1 Equilibrium assuming that $\pi = 1$

Assuming $\pi = 1$, the asset demands, $\{a(s)\}$, and offers, $(q, \{d(s)\})$, maximize:

$$-\sum_{s \in S} [\phi(s) - \beta] a(s) + \beta\sigma [u(q) - q]$$

s.t.:

$$q = \sum_{s \in S} d(s) \tag{37}$$

$$d(s) \leq \frac{k_f(s)}{[\phi(s) - \beta - k_v(s)] [1 + \rho(s)] + \beta\sigma} \tag{38}$$

$$[1 + \rho(s)] d(s) \leq a(s). \tag{39}$$

The first-order conditions with respect to q , $d(s)$, and $a(s)$ are:

$$\begin{aligned}\xi &= \beta\sigma [u'(q) - 1] \\ \phi(s) &= \beta + \nu(s) \\ \xi &= \lambda(s) + [1 + \rho(s)]\nu(s),\end{aligned}$$

where, as in the paper, ξ is the Lagrange multiplier on the participation constraint, (37), $\lambda(s)$ is the multiplier on the resalability constraint (38), and $\nu(s)$ is the multiplier on the feasibility constraint, (39). Assuming for now that $q < q^*$, where q^* is such that $u'(q^*) = 1$, and following the same step as in the text, we obtain our usual three-tier categorization of assets:

Liquid assets. For these assets $\lambda(s) = 0$ and $\nu(s) > 0$, implying that:

$$\begin{aligned}d(s) &= \frac{A(s)}{1 + \rho(s)} \\ \phi(s) &= \beta + \frac{\xi}{1 + \rho(s)} \\ \kappa(s) &\geq \frac{\beta\sigma + \xi}{1 + \rho(s)},\end{aligned}$$

where $\kappa(s) \equiv k_f(s)/A(s) + k_v(s)$ is, as before, the average cost of fraud on asset s .

Partially liquid assets. For these assets, $\lambda(s) > 0$ and $\nu(s) > 0$, implying that:

$$\begin{aligned}d(s) &= \frac{A(s)}{1 + \rho(s)} \\ \phi(s) &= \beta + \kappa(s) - \frac{\beta\sigma}{1 + \rho(s)} \\ \frac{\beta\sigma}{1 + \rho(s)} &< \kappa(s) < \frac{\beta\sigma + \xi}{1 + \rho(s)}.\end{aligned}$$

Illiquid assets. For these assets, $\lambda(s) > 0$ and $\nu(s) = 0$, implying that:

$$\begin{aligned}d(s) &= \frac{k_f(s)}{-k_v(s)[1 + \rho(s)] + \beta\sigma} \\ \phi(s) &= \beta \\ \kappa(s) &\leq \frac{\beta\sigma}{1 + \rho(s)}.\end{aligned}$$

From (37), we then obtain that $q < q^*$ if

$$q = L = \sum_{s \in S} \min \left\{ \frac{k_f(s)}{-[1 + \rho(s)] k_v(s) + \beta\sigma}, \frac{A(s)}{1 + \rho(s)} \right\} < q^*.$$

Conversely, if the above condition is satisfied, and if the variable costs are chosen small enough so that $k_v(s) [1 + \rho(s)] < \beta\sigma$ for all $s \in S$, then the above price and allocation constitute an equilibrium.

B.2 Buyers choose $\pi = 1$

We now prove that, for small enough variable costs, given the prices derived in the previous section, a buyer always finds it optimal to make an offer that is accepted with probability one; i.e., $\pi = 1$. To that end, let us consider some collection of variable costs $k_v \equiv \{k_v(s)\}_{s \in S}$ in the compact set K_v defined by $0 \leq k_v(s) [1 + \bar{\rho}] \leq \frac{\beta\sigma}{2}$ for all $s \in S$. These restrictions on k_v ensures that the price $\phi(s)$ derived in the previous section are well defined continuous functions of k_v and ρ , given that $\rho(s) \leq \bar{\rho}$. Consider, then, the auxiliary problem given the candidate equilibrium prices derived in the previous section. Instead of solving the full auxiliary problem, let us start by solving it for any given $\pi \in [0, 1]$. That is, let us consider:

$$v(\pi, k_v, \rho) = \sup_{q, \{d(s)\} \geq 0} - \sum_{s \in S} [\phi(s) - \beta] [1 + \rho(s)] d(s) + \beta\sigma\pi [u(q) - q] \quad (40)$$

s.t.

$$q = \sum_{s \in S} d(s) \quad (41)$$

$$k_f(s) + [1 + \rho(s)] k_v(s) d(s) \geq \{ [\phi(s) - \beta] [1 + \rho(s)] + \beta\sigma\pi \} d(s), \quad (42)$$

where the prices, $\phi(s)$, implicitly depend on (k_v, ρ) . We proceed with the following results. First, we apply the Theorem of the Maximum and show:

Claim 3 *The function $v(\pi, k_v, \rho)$ is continuous, satisfies $v(0, k_v, \rho) = 0$ and $v(\pi, k_v, \rho) > 0$ for all $\pi \in (0, 1]$.*

The proof follows standard arguments and is relegated to Section B.4. Next, show that:

Claim 4 *There exists some $\underline{\pi} > 0$ such that, for all $(k_v, \rho) \in K_v \times R$, $\pi < \underline{\pi}$ implies that $v(\pi, k_v, \rho) < \sup_{\pi \in [0, 1]} v(\pi, k_v, \rho)$.*

The claim, shown in Section B.4, follows from the continuity of $v(\pi, k_v, \rho)$ and from the fact that $v(0, k_v, \rho) = 0$. Thus, for the purpose of maximizing $v(\pi, k_v, \rho)$ with respect to $\pi \in [0, 1]$, we can without loss of generality restrict attention on $\pi \in [\underline{\pi}, 1]$. Having bounded π below, we can also bound the quantity offered below:

Claim 5 *There exists $\underline{q} > 0$ such that, for all $\pi \in [\underline{\pi}, 1]$ and all $(k_v, \rho) \in K_v \times R$, the output q maximizing (40)-(42) is greater than \underline{q} .*

The claim is shown in Section B.4. With this in mind, for $(\pi, k_v, \rho) \in [\underline{\pi}, 1] \times K_v \times R$, consider a maximizer of (40)-(42) and perform the variational experiment in the proof of Proposition 1 (Claim 2 in the appendix of the paper). Following the same steps, we obtain that, when increasing π by some small $\varepsilon > 0$ the change in utility is, up to second order terms, equal to $\Delta U(\pi, k_v, \rho) \times \varepsilon$ where

$$\Delta U(\pi, k_v, \rho) = \beta\sigma [u(q) - q] + \sum_{s \in S_B} [\phi(s) - \beta] [1 + \rho(s)] m(s) - \beta\sigma\pi [u'(q) - 1] \sum_{s \in S_B} m(s),$$

S_B is the set of binding IC constraints, and

$$m(s) = \frac{\beta\sigma d(s)}{[\phi(s) - \beta - k_v(s)] [1 + \rho(s)] + \beta\sigma\pi} > 0.$$

To conclude the proof, we show that:

Claim 6 *There is some neighborhood N_v of zero variable costs such that, for all $(\pi, k_v, \rho) \in [\underline{\pi}, 1] \times N_v \times R$, $\Delta U(\pi, k_v, \rho) > 0$.*

First we note that

$$\beta\sigma [u(q) - q] = \beta\sigma [u(q) - u'(q)q] + \beta\sigma [u'(q)q - q] \geq \beta\sigma [u(\underline{q}) - u'(\underline{q})\underline{q}] + \beta\sigma [u'(q) - 1] q,$$

where the inequality follows from the function $q \mapsto u(q) - u'(q)q$ being increasing, and $q \geq \underline{q}$. Plugging this inequality back into the expression for $\Delta U(\pi, k_v, \rho)$, and using that $q \geq \sum_{s \in S_B} d(s)$,

we obtain after some manipulations:

$$\begin{aligned} \Delta U(\pi, k_v, \rho) &\geq \beta\sigma [u(\underline{q}) - u'(\underline{q})\underline{q}] + \sum_{s \in S_B} m(s) [1 + \rho(s)] \left\{ -k_v(s) [u'(\underline{q}) - 1] + u'(\underline{q}) [\phi(s) - \beta] \right\} \\ &\geq \beta\sigma [u(\underline{q}) - u'(\underline{q})\underline{q}] - \sum_{s \in S_B} m(s) k_v(s) [1 + \bar{\rho}] [u'(\underline{q}) - 1], \end{aligned}$$

where the second inequality follows because $\phi(s) \geq \beta$ for all $s \in S$, because $\rho(s) \leq \bar{\rho}$, and because $u'(q) \leq u'(\underline{q})$ as $q \geq \underline{q}$. Now restricting attention to variable costs such that $0 \leq k_v(s) [1 + \bar{\rho}] \leq \beta\sigma\bar{\pi}/2$, we can bound $m(s)$ above by setting $d(s) = q^*$, $\pi = \bar{\pi}$, $k_v(s) = [1 + \bar{\rho}]^{-1} \beta\sigma\bar{\pi}/2$, and ignoring the $\phi(s) - \beta$ term at the denominator. This gives the upper bound $m(s) \leq 2q^*/\bar{\pi}$. Plugging this back in the above we find that:

$$\Delta U(\pi, k_v, \rho) \geq \beta\sigma [u(\underline{q}) - u'(\underline{q})\underline{q}] - \sum_{s \in S} k_v(s) \frac{2q^*}{\bar{\pi}}.$$

Since $u(q)$ is strictly concave, $u(0) = 0$, and $\underline{q} > 0$, the first term is strictly positive. The second term can be made arbitrarily close to zero by choosing k_v close enough to zero. This proves the claim and shows that the auxiliary problem is solved at $\pi = 1$.

B.3 Optimal retention requirements

Fix some $k_v \in N_v$ and consider retention requirements $\rho \in R$. By construction of N_v , for all $\rho \in R$, there exists an equilibrium of the form shown in Section B.1. Focusing on such equilibria, it is clear that an optimal retention requirement maximizes L . If the asset is partially liquid or liquid, in that $\kappa(s) \geq \beta\sigma$, then $\rho(s) = 0$. If the asset is illiquid, in that $\kappa(s) < \beta\sigma$, then the optimal retention requirement solves

$$\frac{k_f(s)}{-[1 + \rho(s)] k_v(s) + \beta\sigma} = \frac{A(s)}{1 + \rho(s)} \iff [1 + \rho(s)] \left[\frac{k_f(s)}{A(s)} + k_v(s) \right] = \beta\sigma,$$

which, by (??) is strictly less than $\bar{\rho}$.

B.4 Omitted Proofs

Proof of Claim 3. We use the Theorem of the Maximum as stated in Chapter 3.3 of Stokey, and Lucas (1989). First, note that we can without loss of generality restrict attention to $q \in [0, Sq^*]$,

$d(s) \in [0, q^*]$, for all $s \in S$. Indeed, if $d(s) > q^*$, then by (41) it follows that $q > q^*$. Thus $u'(q) < 1$ and reducing $d(s)$ and q improves the objective (40) strictly. Second, we need to show that the correspondence $\Gamma(\pi, k_v, \rho)$ defined by (41), (42), $q \in [0, Sq^*]$ and $d(s) \in [0, q^*]$ for all $s \in S$, is continuous.

To show that it is upper hemi continuous (uhc), let us first note that it is non-empty as it contains the zero offer, $d(s) = q = 0$. Next, consider a sequence (π^n, k_v^n, ρ^n) converging to some (π, k_v, ρ) , and some associated sequence $(d^n, q^n) \in \Gamma(\pi^n, k_v^n, \rho^n)$. Because for any (π, k_v, ρ) , $\Gamma(\pi, k_v, \rho)$ is a subset of the compact $[0, q^*]^S \times [0, Sq^*]$, the sequence (d^n, q^n) has a converging subsequence. Given that the correspondence $\Gamma(\cdot)$ is defined by weak inequalities satisfied by continuous functions, the limit of the converging sequence must belong to $\Gamma(\pi, k_v, \rho)$.

To show that it is lower hemi continuous (lhc), consider some $(d, q) \in \Gamma(\pi, k_v, \rho)$, and some sequence (π^n, k_v^n, ρ^n) converging to (π, k_v, ρ) . We need to find some sequence (d^n, π^n) converging to (π, k_v, ρ) and such that $(d^n, q^n) \in \Gamma(\pi^n, k_v^n, \rho^n)$ for all n large enough. We proceed as follows. If for some $s \in S$, the constraint (42) is not binding for $d(s)$ with (π, k_v, ρ) then, by continuity, for n large enough it is not binding for $d(s)$ with (π^n, k_v^n, ρ) . Thus we can pick $d^n(s) = d(s)$. If the constraint (42) is binding for some s with (π, k_v, ρ) , then we distinguish three sub cases.

1. If $[1 + \rho(s)] k_v(s) < [\phi(s) - \beta] [1 + \rho(s)] + \beta\sigma\pi$, then by continuity the same strict inequality holds for (π^n, k_v^n) and n large enough, and we can choose $d^n(s)$ so that the constraint (42) holds with equality, that is:

$$d^n(s) = \frac{k_f(s)}{[\phi(s) - \beta] [1 + \rho^n(s)] - [1 + \rho^n(s)] k_v^n(s) + \beta\sigma\pi^n} \longrightarrow d(s).$$

2. If $[(1 + \rho(s)] k_v(s) > [\phi(s) - \beta] [1 + \rho(s)] + \beta\sigma\pi$, then by continuity the same strict inequality holds for (π^n, k_v^n, ρ^n) and n large enough, implying that the associated IC constraint, (42), holds for all $d \in [0, q^*]$. In particular, we can choose $d^n(s) = d$.
3. If $[(1 + \rho(s)] k_v(s) = [\phi(s) - \beta] [1 + \rho(s)] + \beta\sigma\pi$ then the left-hand side and the right-hand side coefficients multiplying d in the IC constraint (42) can be made arbitrarily close to each other for n large enough. Thus, for n large enough, the IC constraint holds for all $d \in [0, q^*]$, and we can choose $d^n(s) = d(s)$.

Since the correspondence $\Gamma(\pi, k_v, \rho)$ is compact valued, uhc, and lhc, and since the objective (40) is continuous, it follows from the Theorem of the Maximum that $v(\pi, k_v, \rho)$ is continuous. That $v(0, k_v, \rho) = 0$ follows since, when $\pi = 0$, the objective is non-positive, and since the objective is zero with offer $d(s) = q = 0$. That $v(1, k_v, \rho) > 0$ follows because, when $\pi > 0$, a small offer is feasible and, since $u'(0) = \infty$, achieves a positive value.

Proof of Claim 4. Let $\underline{\pi} \equiv \inf \{ \pi : \exists (k_v, \rho) \in K_v \times R, v(\pi, k_v, \rho) \geq v(1, k_v, \rho)/2 \}$. To see that $\underline{\pi} > 0$, suppose by contradiction that $\underline{\pi} = 0$. By definition of $\underline{\pi}$, there exists a sequence (π^n, k_v^n, ρ^n) such that $\pi^n \rightarrow \underline{\pi} = 0$ and $v(\pi^n, k_v^n, \rho^n) > v(1, k_v^n, \rho^n)/2$. Without loss of generality since $K_v \times R$ is compact, we can assume that (k_v^n, ρ^n) has some limit (k_v, ρ) . By continuity of $v(\cdot)$, we thus obtain that $v(0, k_v, \rho) \geq v(1, k_v, \rho)$, a contradiction. Clearly, for all $\pi \in [0, \underline{\pi}]$ and all $(k_v, \rho) \in K_v \times R$, we have:

$$v(\pi, k_v, \rho) < \frac{v(1, k_v, \rho)}{2} < \sup_{\pi \in [0, 1]} v(\pi, k_v, \rho).$$

Proof of Claim 5. First note that the program (40)-(42) has a maximizer because the objective is continuous and the constraint set compact. Moreover, since the objective is strictly concave in q , all maximizers offer the same level of output q . Now to prove the claim, suppose to the contrary that we can find a sequence $(\pi^n, k_v^n, \rho^n, d^n, q^n) \in [\underline{\pi}, 1] \times K_v \times R \times [0, q^*]^S \times [0, Sq^*]$, such that, for each n , (d^n, q^n) maximizes (40)-(42), and such that $q^n \rightarrow 0$. Clearly, since the sequence is bounded, we can assume without loss of generality that it converges to some (π, k_v, ρ, d, q) , where $q = 0$. Since the correspondence $\Gamma(\cdot)$ defining feasible offers is continuous, we must have $(d, q) \in \Gamma(\pi, k_v, \rho)$. Moreover, since the objective (40) evaluated at (d^n, q^n) achieves the maximum value $v(\pi^n, k_v^n, \rho)$ for each n , and since $v(\cdot)$ is continuous, it follows by taking the limit that (d, q) achieves the maximum value $v(\pi, k_v, \rho)$. But since $q = 0$ this means that $v(\pi, k_v, \rho) = 0$, which is a contradiction given that $\pi \geq \underline{\pi}$.

C Variable costs of fraud and rationing at the extensive margin

Throughout this paper we assume that the production of fraudulent assets involves a fixed cost only. This assumption is realistic to describe counterfeiting of assets and frauds in payments. It also makes the model simple and tractable. We want to explore how our results would be affected if the cost of producing fraudulent assets had a linear component. We will show that under alternative cost structures our model can generate resalability constraints that affect both the intensive margin (the quantity of assets transferred in a match) and the extensive margin (the frequency of trades).

In order to make our point we describe a single-asset economy where the cost of producing d units of fraudulent asset is $k_f + k_v d$. Following the same reasoning as in Proposition 1, it can be shown that an outcome of the game between a buyer and a seller solves

$$\max_{q,d,\pi} \{ -(\phi - \beta)d + \beta\sigma\pi[u(q) - d] \} \quad (43)$$

$$\text{s.t.} \quad -q + d = 0 \quad (44)$$

$$k_f + [k_v - \phi + \beta(1 - \sigma\pi)]d \geq 0, \quad (45)$$

where (q, d) is the equilibrium offer, π is the probability that the offer is accepted. If $k_v \geq \phi - \beta(1 - \sigma)$, then the liquidity constraint (45) does not bind and $\pi = 1$. In the following we focus on the case where $k_v < \phi - \beta(1 - \sigma)$.

Consider first the case where $k_f = 0$ and $k_v > 0$. If $k_v \leq \phi - \beta$, then the per unit cost of producing fraudulent assets is less than that of acquiring genuine assets. As a consequence, the DM shuts down, $d = q = 0$. Suppose next that $\phi - \beta < k_v < \phi - \beta(1 - \sigma)$. The solution to (43)-(45) is

$$\begin{aligned} \pi &= \frac{k_v - \phi + \beta}{\beta\sigma} \in (0, 1). \\ q &= u'^{-1} \left(\frac{k_v}{k_v - \phi + \beta} \right). \end{aligned}$$

The resalability constraint takes the form of a positive probability that an offer is rejected. The fraction of the meetings where the asset is accepted increases with the cost of producing fraudulent assets, but decreases with the price of the asset and the frequency of the trading opportunities.

In order to endogenize asset prices, suppose that there is a fixed supply of the asset, A . If $k_v \geq \beta\sigma u'(q)$ where $q = \min(A, q^*)$, then the asset is liquid, in which case $\phi = \beta + \beta\sigma[u'(q) - 1]$

and $\pi = 1$. If $k_v < \beta\sigma u'(q)$, then

$$\begin{aligned}\phi &= \beta + k_v \frac{[u'(A) - 1]}{u'(A)} \\ \pi &= \frac{k_v}{u'(A)\beta\sigma}.\end{aligned}$$

The price of a partially liquid asset depends on the proportional cost of fraud but it does not depend on the frequency of trading opportunities. In contrast, the probability that a trade is accepted decreases with the frequency of trading opportunities so that the asset velocity is constant.

In the more general case where $k_f > 0$ and $k_v > 0$, the resalability constraint can affect both the quantity of assets transferred in a match and the frequency of matches. To see this, suppose that $\phi = \beta$ and $k_f < (\beta\sigma - k_v)q^*$. The solution to (43)-(45) is such that

$$\begin{aligned}\pi &= \frac{k_f + k_v q}{\beta\sigma q} \\ q &= \arg \max_{q \in [\underline{q}, q^*]} \left[\frac{k_f}{q} + k_v \right] [u(q) - q].\end{aligned}$$

where $\underline{q} = \frac{k_f}{\beta\sigma - k_v}$ (assuming $\beta\sigma - k_v > 0$). First, $q < q^*$ so that agents trade less than the surplus-maximizing quantity. Second, if

$$k_v > \frac{k_f(1 - \underline{\eta})}{\underline{q}\underline{\eta}},$$

where $\underline{\eta} = \frac{q[u'(q) - 1]}{u(q) - q} \in (0, 1)$, then $q > \underline{q}$ and $\pi < 1$. If prices are endogenous, such an outcome is an equilibrium provided that $A \geq q$.

Kiyotaki and Moore (2005) assume that agents can only use a fraction of their capital holdings to advance investment opportunities, while in Lagos (2010) agents can spend their assets in an exogenously given fraction of trading opportunities. Our model can capture both effects with various setups on the cost structure of producing fraudulent assets.

D Proof Omitted in Section 5.3

We derive the equilibrium for the dynamic extension of our model. Section D.1 provides the main equilibrium equations and the pricing of assets in period $t - 1$, taking as given the price prevailing in period t . Section D.2 derives the steady-state equilibrium. Section D.3 studies the special case of fiat money. Section D.4 provides the one-stage deviation principle allowing us to reduce agents inter-temporal problem to a sequence of static problems.

D.1 Equilibrium equations

In Section D.4 below we show that a one-stage deviation principle applies to agents' problem. Together with quasilinearity, this implies that agents' intertemporal optimization problem reduces to a sequence of static problems. These static problems are identical to the one presented in the analysis of the static model, except for one difference: the terminal value of the asset is no longer normalized to one, but is instead equal to the cum-dividend value of reselling the asset in next period CM. With this in mind, a buyer in the CM of $t - 1$ choosing his portfolio of assets faces the following problem:

$$\max_{a_t(s), d_t(s), q_t} - \sum_{s \in S} \left\{ \phi_{t-1}(s) - \beta [\phi_t(s) + 1] \right\} a_t(s) + \sigma \beta [u(q_t) - q_t] \quad (46)$$

$$\text{s.t.} \quad -q_t + \sum_{s \in S} [1 + \phi_t(s)] d_t(s) = 0 \quad (47)$$

$$d_t(s) \leq \frac{k(s)}{\phi_{t-1}(s) - \beta(1 - \sigma) [1 + \phi_t(s)]} \quad (48)$$

$$d_t(s) \in [0, a_t(s)]. \quad (49)$$

The first-order conditions associated with the problem (46)-(49) are

$$\begin{aligned} u'(q) &= 1 + \frac{\xi_t}{\sigma \beta} \\ \phi_{t-1}(s) &= \beta [1 + \phi_t(s)] + \nu_t(s), \quad \text{for all } s \in S \\ \lambda_t(s) + \nu_t(s) &= \xi_t [1 + \phi_t(s)], \quad \text{for all } s \in S, \end{aligned}$$

where ξ_t , $\lambda_t(s)$ and $\nu_t(s)$ are the Lagrange multipliers associated with the seller's participation constraint, (47), the resalability constraint, (48), and the feasibility constraint, (49), respectively.

The complementary slackness conditions are

$$\begin{aligned}\lambda_t(s) \left(\frac{k(s)}{\phi_{t-1}(s) - \beta(1-\sigma)[1 + \phi_t(s)]} - d_t(s) \right) &= 0 \\ \nu_t(s) [a_t(s) - d_t(s)] &= 0.\end{aligned}$$

Consider first an equilibrium with abundant liquidity, where $\xi_t = 0$ for all t . Then, $q_t = q^*$ and $\lambda_t(s) = \nu_t(s) = 0$ for all t and all $s \in S$. From the asset pricing equation, $\phi_{t-1}(s) = \beta [1 + \phi_t(s)]$ for all s . Under the transversality condition $\lim_{t \rightarrow \infty} \beta^t \phi_t(s) = 0$, the solution to this first-order difference equation is $\phi_t(s) = 1/r$.

Consider next "liquidity-constrained" equilibria, $\xi_t > 0$ for some t . As in the static model we distinguish three classes of assets: illiquid assets ($\nu_t(s) = 0$, $\lambda_t(s) > 0$), partially liquid assets ($\nu_t(s) > 0$, $\lambda_t(s) > 0$), and liquid assets ($\nu_t(s) > 0$, $\lambda_t(s) = 0$). The price of illiquid assets solves $\phi_{t-1}(s) = \beta [1 + \phi_t(s)]$. The condition $\nu_t(s) = 0$ implies $d_t(s) \leq a_t(s)$, which yields $\frac{k(s)}{\beta\sigma[1+\phi_t(s)]} \leq a_t(s)$. Therefore, the fraction of the stock of the asset that can be resold is $\theta_t(s) = \frac{\kappa_t(s)}{\beta\sigma} \leq 1$, where $\kappa_t(s) \equiv \frac{k(s)}{[1+\phi_t(s)]A(s)}$, which implies $\kappa_t(s) \leq \bar{\kappa} \equiv \beta\sigma$. The price of liquid assets solves $\phi_{t-1}(s) = (\beta + \xi_t) [\phi_t(s) + 1]$. The conditions $\nu_t(s) = 0$ and $\lambda_t(s) > 0$ imply that the buyer spends all his asset, while the resalability constraint, (48), is slack; that is, $d(s) = A(s) \leq \frac{k(s)}{(\xi_t + \beta\sigma)[1 + \phi_t(s)]}$. Therefore, $\kappa_t(s) \geq \bar{\kappa}_t \equiv \xi_t + \beta\sigma$. Finally, the price of partially liquid assets solves $\phi_{t-1}(s) - \beta(1-\sigma)[1 + \phi_t(s)] = \frac{k(s)}{A_t(s)}$ since $\lambda_t(s) > 0$. Similarly, from $\lambda > 0$ and $\nu > 0$, one can derive $\kappa_t(s) \in (\underline{\kappa}_t, \bar{\kappa}_t)$. Rearranging these expressions leads to the pricing equations shown in the text.

D.2 Steady-state equilibrium

From (25) asset prices at the steady state solve

$$\bar{\phi}(s) = \begin{cases} \frac{\beta + \bar{\xi}}{1 - \beta - \bar{\xi}} \\ \frac{\beta}{1 - \beta} \\ \frac{\beta(1 - \sigma) + \frac{k(s)}{A(s)}}{1 - \beta(1 - \sigma)} \end{cases} \quad \text{if} \quad \begin{cases} \frac{(1 - \beta - \bar{\xi})k(s)}{A(s)} \geq \bar{\kappa} \\ \frac{(1 - \beta)k(s)}{A(s)} \leq \underline{\kappa} \\ \frac{[1 - \beta(1 - \sigma)]k(s)}{[A(s) + k(s)]} \in (\underline{\kappa}, \bar{\kappa}) \end{cases} \quad (50)$$

where $\underline{\kappa} = \beta\sigma$ and $\bar{\kappa} = \beta\sigma + \bar{\xi}$. The output traded in the DM is

$$\begin{aligned} \bar{q} = & \sum_{s: \frac{(1-\beta)k(s)}{A(s)} \leq \beta\sigma} \frac{k(s)}{\beta\sigma} + \frac{A(s) + k(s)}{1 - \beta(1 - \sigma)} \mathbb{I} \left\{ \frac{[1-\beta(1-\sigma)]k(s)}{[A(s)+k(s)]} \in (\underline{\kappa}, \bar{\kappa}) \right\} \\ & + \sum_{s: \frac{(1-\beta)k(s)}{A(s)} > \beta\sigma} \frac{A(s)}{1 - \beta - \bar{\xi}} \mathbb{I} \left\{ \frac{[1-\beta(1-\sigma)]k(s)}{[A(s)+k(s)]} \geq \bar{\kappa} \right\}. \end{aligned} \quad (51)$$

From (52) \bar{q} is a non-decreasing function of $\bar{\xi}$. Finally, $\bar{\xi}$ is the unique solution to

$$\bar{\xi} = \beta\sigma [u'(\bar{q}) - 1].$$

Such a solution exists if $L < q^*$. If $L \geq q^*$, then the steady state is the “liquidity-abundant” equilibrium described above.

D.3 A special case fiat money

As a special case, we now study an economy with a single asset, fiat money, that pays no dividend. This case is worthwhile studying as counterfeiting of currency has been prevalent throughout history. We will ask whether the threat of counterfeiting affects the existence of a monetary equilibrium and the design of the optimal monetary policy. In order to study inflation policy we assume that the supply of money, A_t , is growing at a constant rate, $\gamma \equiv \frac{A_{t+1}}{A_t} > \beta$. Money is injected through lump-sum transfers to buyers, $T = \phi_t(A_{t+1} - A_t)$. Because of quasi-linear preferences, these transfers do not affect the buyer’s problem in the CM.

We focus on stationary equilibria where the real value of money is constant over time, $\phi_{t+1}A_{t+1} = \phi_t A_t$. The rate of return of money is $\frac{\phi_{t+1}}{\phi_t} = \gamma^{-1} < \beta$. From the paper, the resalability constraint can be rewritten as $\phi_t d \leq \frac{k}{\gamma - \beta(1 - \sigma)}$. One novelty is that policy, through the growth rate of the money supply, has a direct effect on the resalability of the asset. An increase in the inflation rate, γ , makes it more costly to hold genuine money, which makes it more attractive to produce counterfeits. If the resalability constraint does not bind, then the value of money solves

$$\phi_t = \frac{q}{A_t}. \quad (52)$$

From (25) and the fact that $\frac{\phi_{t-1}}{\phi_t} = \gamma$, we obtain $\xi = \gamma - \beta$. Hence q solves

$$u'(q) = 1 + \frac{\gamma - \beta}{\beta\sigma}. \quad (53)$$

When the resalability constraint binds, we obtain

$$\phi_t = \frac{k}{[\gamma - \beta(1 - \sigma)] A_t} \quad (54)$$

$$q = \frac{k}{\gamma - \beta(1 - \sigma)}. \quad (55)$$

Let $\bar{k} \equiv [\gamma - \beta(1 - \sigma)] \hat{q}$ be the threshold for the counterfeiting cost below which the resalability constraint is binding, where \hat{q} solves $u'(\hat{q}) = 1 + \frac{\gamma - \beta}{\beta\sigma}$. If $k \geq \bar{k}$, then q solves (53); otherwise q is determined by (55). Thus, there exists a monetary equilibrium with $\phi_t A_t > 0$ if and only if

$$u'(0) > 1 + \frac{\gamma - \beta}{\beta\sigma}. \quad (56)$$

Fiat money can be valued even though sellers do not have the technology to distinguish genuine units of money from counterfeits. The possibility of counterfeiting does not threaten the existence of a monetary equilibrium since the cost of producing counterfeits, $k > 0$, is absent from (56).

In terms of policy implications, the Friedman rule ($\gamma \searrow \beta$) is optimal because, from (53) and (55), q is a decreasing function of γ . It achieves the first best if and only if $k \geq \bar{k}$ at $\gamma = \beta$ and $q = q^*$; i.e.,

$$k \geq \beta\sigma q^*. \quad (57)$$

That is, the Friedman rule achieves the first-best allocation if and only if the cost of producing counterfeits is large enough.

D.4 The one-stage deviation principle

For the seller. A seller chooses a plan $\{a_t(s) : s \in S\}_{t=1}^\infty$ and $\{\pi_t\}_{t=1}^\infty$, where the choice of asset holdings, $a_t(s)$, is a function of the history up to the CM of $t - 1$, and the acceptance rule, $\pi_t(o_t)$, that specifies the probability that an offer, $o_t = (q_t, \{d_t(s)\})$, is accepted is a function of the history up to the DM of t . The expected discounted utility of the seller up to period T with initial asset holdings $\{a_0(s)\}$ is

$$\begin{aligned} U_s^T &= \sum_{s \in S} \phi_0(s) [a_0(s) - a_1(s)] + \mathbb{E}_0 \left[\sum_{t=1}^T \beta^t \pi_t(o_t) \left(-q_t + \sum_{s \in S} [\phi_t(s) + 1] d_t(s) \right) \right] \\ &\quad + \mathbb{E}_0 \left[\sum_{t=1}^T \beta^t \sum_{s \in S} \{[\phi_t(s) + 1] a_t(s) - \phi_t(s) a_{t+1}(s)\} \right], \end{aligned}$$

where the expectation is with respect to the different histories up to T , and the meeting probability is taken into account by the expectation operator. Rearrange the terms to obtain:

$$\begin{aligned}
U_s^T &= \sum_{s \in S} \phi_0(s) a_0(s) + \mathbb{E}_0 \left[\sum_{t=1}^T \beta^t \pi_t(o_t) \left(-q_t + \sum_{s \in S} [\phi_t(s) + 1] d_t(s) \right) \right] \\
&\quad + \mathbb{E}_0 \left[\sum_{t=0}^{T-1} \beta^t \sum_{s \in S} \{ -\phi_t(s) + \beta [\phi_{t+1}(s) + 1] \} a_{t+1}(s) \right] \\
&\quad - \sum_{s \in S} \beta^T \phi_T(s) a_{T+1}(s).
\end{aligned}$$

Taking the limit as T goes to infinity, and using the assumptions that $a_{T+1}(s)$ is bounded above and $\lim_{T \rightarrow \infty} \beta^T \phi_T(s) = 0$, we have

$$\begin{aligned}
U_s^\infty &= \sum_{s \in S} \phi_0(s) a_0(s) + \mathbb{E}_0 \left[\sum_{t=1}^{\infty} \beta^t \pi_t(o_t) \left(-q_t + \sum_{s \in S} [\phi_t(s) + 1] d_t(s) \right) \right] \\
&\quad + \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \sum_{s \in S} \{ -\phi_t(s) + \beta [\phi_{t+1}(s) + 1] \} a_{t+1}(s) \right].
\end{aligned}$$

It is then clear from the above expression that the one-stage deviation principle applies. In the DM of period t the seller accepts or rejects an offer depending on the sign of $-q_t + \sum_{s \in S} [\phi_t(s) + 1] \mathbb{E}_t d_t(s)$. (Remember that the offer is independent of the seller's asset holdings.) In the CM of period t the seller chooses his asset holdings to maximize $\{ -\phi_t(s) + \beta [\phi_{t+1}(s) + 1] \} a_{t+1}(s)$. Since $-\phi_t(s) + \beta [\phi_{t+1}(s) + 1] \leq 0$, it is optimal for sellers not to hold assets.

For the buyer. A buyer chooses a plan $\{a_t(s) : s \in S\}_{t=0}^\infty$ and a sequence of offers, $\{q_t, \{d_t(s) : s \in S\}\}_{t=1}^\infty$, where the choice of asset holdings, $a_t(s)$, is a function of the history up to the CM of $t - 1$, and the offer, o_t , is a function of the history up to the DM of t . The expected discounted utility of the buyer up to T is

$$\begin{aligned}
U_b^T &= \sum_{s \in S} \phi_0(s) [a_0(s) - a_1(s)] + \mathbb{E}_0 \left[\sum_{t=1}^T \beta^t \pi_t(o_t) \left\{ u(q_t) - \sum_{s \in S} [\phi_t(s) + 1] d_t(s) \right\} \right] \\
&\quad + \mathbb{E}_0 \left[\sum_{t=1}^T \beta^t \sum_{s \in S} \{ [\phi_t(s) + 1] a_t(s) - \phi_t(s) a_{t+1}(s) \} \right].
\end{aligned}$$

Rearrange the terms to obtain:

$$U_b^T = \sum_{s \in S} \phi_0(s) a_0(s) + \mathbb{E}_0 \left[\sum_{t=0}^{T-1} \beta^t \mathcal{U}_t(\{a_t(s)\}, q_t, \{d_t(s)\}) \right] - \mathbb{E}_0 \left[\sum_{s \in S} \beta^T \phi_T(s) a_{T+1}(s) \right],$$

where

$$\mathcal{U}_t = \sum_{s \in S} \left\{ -\phi_t(s) + \beta [\phi_{t+1}(s) + 1] \right\} a_{t+1}(s) + \beta \pi_{t+1}(o_{t+1}) \left(u(q_{t+1}) - \sum_{s \in S} [\phi_{t+1}(s) + 1] d_{t+1}(s) \right).$$

As before, we can take the limit as T goes to obtain

$$U_b^\infty = \sum_{s \in S} \phi_0(s) a_0(s) + \mathbb{E}_0 \left[\sum_{t=0}^{T-1} \beta^t \mathcal{U}_t(\{a_t(s)\}, q_t, \{d_t(s)\}) \right].$$

The buyer faces a sequence of independent problems in each period, where he has to maximize $\mathbb{E}_{t-1} [\mathcal{U}_t(\{a_t(s)\}, q_t, \{d_t(s)\})]$, taking as given the acceptance rule of sellers.

E Comparison: models with exogenous resalability constraints

In this section we show that the implications of our model with endogenous resalability constraints differ from the implications of leading models with exogenous resalability constraints.

E.1 Proportional resalability constraints

We start with the proportional resalability constraints of Kiyotaki and Moore (2005, henceforth KM): agents can re-sell a fraction $\delta(s) \in [0, 1]$ of their holdings of asset $s \in S$. We first show an observational equivalence result: it is possible to choose asset supplies and resalability constraints in order to obtain the same cross-sectional asset prices as in our fraud-based model. However, the comparative statics and policy implications often go in opposite directions than in KM. Notable differences between the two models include:

1. Flight to liquidity. With KM resalability constraints, an increase in σ increases the price of partially liquid and liquid assets, and increases welfare. In our model, increasing the frequency of trades has the additional effect of increasing incentives to commit fraud. As a result, the prices of liquid assets increase, but the prices of partially liquid assets decrease. In some case, welfare can also decrease.
2. Open-market operations. With KM resalability constraints, a budget-balanced open market purchase of partially liquid assets in exchange for liquid assets leads to an increase in the interest rate, as measured by the yield of liquid assets. Moreover, it increases welfare. In our model we obtain the opposite effects: the interest rate goes down, and welfare decreases.
3. Aggregate liquidity measurement. With KM resalability constraints, an increase in the Divisia liquidity aggregate is always associated with an increase in welfare. In our model, we can have the opposite effect. This is because, in our model, the Divisia method underestimates the true contribution of partially liquid assets to aggregate liquidity.
4. Excess volatility. With KM resalability constraints, the prices of all partially liquid and liquid assets can be indeterminate and exhibit endogenous fluctuations and excess volatility. In contrast, in our fraud-based model only the prices of liquid assets can be indeterminate.

Therefore, even though the model with KM resalability constraints can generate the same cross-sectional asset prices as in our model, it has different implications for dynamics.

5. Shocks to assets' resalability. Several authors (e.g., Ajello, 2010; Shi, 2011) formulate the hypothesis that recessions are driven by liquidity shocks formalized by a reduction of the exogenous resalability of some assets. Under the KM specification, however, these models can have the counterfactual implication that the prices of the assets that become more difficult to resale increase. In contrast, in our fraud-based model we can formalize such a shock as a reduction in $k(s)$, which reduces the resalability of the asset. If the asset that becomes less recognizable is a partially liquid asset, then its price decreases.

The model. Assume that buyers in the DM can only re-sell a fraction $\delta(s) \in [0, 1]$ of their holding of asset $s \in S$. These resemble the exogenous resalability constraints of Kiyotaki Moore (2001, 2005). Following the same steps as in the paper, the asset demands and offers of buyers solve:

$$\max_{a(s), q, d(s)} - \sum_{s \in S} [\phi(s) - \beta] a(s) + \beta \sigma [u(q) - q] \quad (58)$$

subject to

$$\sum_{s \in S} d(s) = q \quad (59)$$

$$0 \leq d(s) \leq \delta(s)a(s), \quad \text{for all } s \in S. \quad (60)$$

Note that, since $\delta(s) \in [0, 1]$, we do not need to add the feasibility constraint. An equilibrium is a collection of prices, $\{\phi(s)\}$, asset demands, $\{a(s)\}$, and DM offer, $(q, \{d(s)\})$, such that the asset demands and the offer solve the buyer's problem (58)-(60) given prices, and the asset market clears; i.e., $a(s) = A(s)$ for all $s \in S$.

Let ξ be the multiplier on the seller's participation constraint, (59), and $\lambda(s)$ the multiplier on the resalability constraint, (60). The first-order conditions are:

$$\xi = \beta \sigma [u'(q) - 1] = \lambda(s) \quad (61)$$

$$\phi(s) = \beta + \lambda(s)\delta(s). \quad (62)$$

Following similar calculations as in the paper, we obtain the following Proposition.

Proposition 3 *Let $L \equiv \sum_{s \in S} \delta(s)A(s)$. If $L \geq q^*$, then $q = q^*$ and $\phi(s) = \beta$ for all $s \in S$. If $L < q^*$, then*

$$q = L = \sum_{s \in S} \delta(s)A(s) \quad (63)$$

$$\phi(s) = \beta + \xi \delta(s), \quad (64)$$

for all $s \in S$, where $\xi = \beta\sigma [u'(q) - 1]$.

Partial observational equivalence. We show that the model with exogenous resalability constraint is observationally equivalent to our fraud-based model: one can set asset supplies and resalability constraints, $\delta(s)$, so as to obtain the same cross-sectional asset prices as in our model.

Consider the fraud-based model described in the paper, with asset supplies $\tilde{A}(s)$, equilibrium output \tilde{q} , and equilibrium asset prices $\tilde{\phi}(s)$. Let us start with the liquid assets, i.e., assets such that $\tilde{\kappa}(s) \geq \beta\sigma + \tilde{\xi}$ in our fraud-based model. The price of such assets is $\tilde{\phi}(s) = \beta + \tilde{\xi}$. Assuming $q = \tilde{q}$ and $\xi = \tilde{\xi}$, one can set $\delta(s) = 1$ so that $\phi(s) = \tilde{\phi}(s)$. Consider next illiquid assets, i.e., assets such that $\tilde{\kappa}(s) \leq \beta\sigma$ in our fraud-based model. The price of such assets is $\tilde{\phi}(s) = \beta$. Therefore, we have to set $\delta(s) = 0$ to obtain $\phi(s) = \tilde{\phi}(s)$. Finally, consider partially liquid assets, i.e., assets such that $\beta\sigma < \tilde{\kappa}(s) < \beta\sigma + \tilde{\xi}$. The prices of such assets are $\tilde{\phi}(s) = \beta + [\tilde{\kappa}(s) - \beta\sigma]$. The condition $\phi(s) = \tilde{\phi}(s)$ implies $\delta(s) = \frac{\tilde{\kappa}(s) - \beta\sigma}{\tilde{\xi}}$. See Figure 4. In order to guarantee that $q = \tilde{q}$ and $\xi = \tilde{\xi}$, one can choose the supplies of assets such that $\sum_{s \in S} \delta(s)A(s) = \tilde{q}$.

Even though the equilibrium distribution of asset prices in the model with KM resalability constraints is the same as in our fraud-based model, asset velocities are different: in our model, an asset with $\phi(s) = \beta$ is partially traded in the DM, whereas $\delta(s) = 0$ for such an asset in the KM model; i.e., it is not traded at all. Moreover, we show below that the comparative statics and welfare analysis often go in the opposite direction in the two models.

Flight to liquidity. From Proposition 3, when σ increases in the KM model, q is unaffected but ξ rises. Because, from (58), the liquidity premium of all assets are proportional to ξ , it follows that the prices of the partially liquid and liquid assets go up. In the fraud-based model, an increase in σ has an additional effect by raising incentives to produce fraudulent assets, which tightens

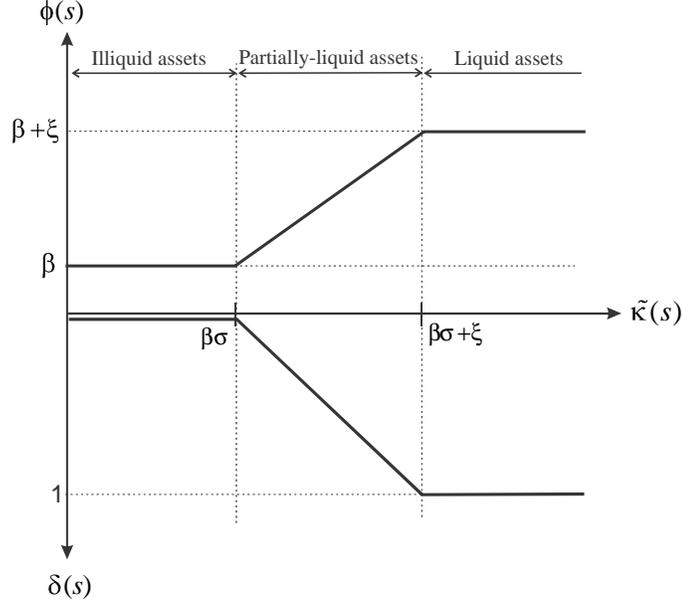


Figure 4: Observational equivalence: Asset prices and proportional liquidity constraints.

resalability constraints. As a consequence, the prices of partially liquid assets fall. Moreover, in the fraud-based model, which assets are liquid or illiquid is endogenously determined, and depends on the frequency of trades: as σ increases, the set of illiquid assets expands and the set of liquid assets shrinks. Such an effect cannot be captured in a model with exogenous liquidity constraints.

Open market operations. Suppose the Central Bank purchases some partially liquid asset, s , by selling some more liquid asset, s' , with $\delta(s') > \delta(s)$. We consider a small open market operation that is budget balanced at $t = 0$, which implies:

$$\phi(s)\Delta A(s) + \phi(s')\Delta A(s') = 0 \quad \Rightarrow \quad \Delta A(s) = -\frac{\phi(s')}{\phi(s)}\Delta A(s'), \quad (65)$$

where $\Delta A(s') > 0$ and $\Delta A(s) < 0$. The increase in liquidity in the DM is:

$$\delta(s)\Delta A(s) + \delta(s')\Delta A(s') = \frac{1}{\xi} \left\{ [\phi(s) - \beta] \Delta A(s) + [\phi(s') - \beta] \Delta A(s') \right\}.$$

Substituting the above equation into (65), we obtain the change in liquidity as

$$\frac{\beta \phi(s') - \phi(s)}{\xi \phi(s)} \Delta A(s').$$

The change in welfare is the change in liquidity multiplied by ξ ,

$$\Delta W = \beta \underbrace{\left(\frac{1}{\phi(s)} - \frac{1}{\phi(s')} \right)}_{\text{yield spread}} \underbrace{\phi(s') \Delta A(s')}_{\text{quantity sold}} > 0.$$

Thus, exchanging liquid for illiquid assets is always welfare improving. This is intuitive: the Central Bank wants to have as many liquid assets in circulation as possible.

Perhaps surprisingly, our model has the opposite implication. The reason for this difference is that in our fraud-based model partially liquid assets are underpriced, in the sense that their prices is low relative to the “true” amount of liquidity services they provide in the DM. When the Central Bank buys partially liquid assets, it syphons out assets that provide good liquidity at a low price. But, when it sells liquid assets, it injects assets that provide good liquidity at a high price. If one insists on budget balanced intervention, the net effect on aggregate liquidity and welfare is strictly negative. The prices of partially liquid and liquid assets go up instead of going down.

Aggregate liquidity measurement. From (63) and (64), changes in aggregate liquidity in the KM model can be calculated using the standard Divisia method:

$$\Delta L^{\text{divisia}} = \sum_{s \in S} \frac{\phi(s) - \beta}{\xi} \Delta A(s). \quad (66)$$

i.e., the marginal social value of asset s is proportional to its holding cost, or convenience yield. In the KM model, an increase in the Divisia aggregate, $\Delta L^{\text{divisia}} > 0$ in equation (66), is always associated with an increase in welfare.

This is not so in our fraud-based model. For instance, the open-market which sells liquid assets in exchange for partially liquid ones increases the Divisia liquidity aggregate; i.e., $\Delta L^{\text{divisia}} > 0$ in equation (66). At the same time, it decreases welfare. The reason is that, in our model, partially liquid assets contribute their full supply to aggregate liquidity, even though their convenience yields are smaller than that of the most liquid assets. In other word, the convenience yields of partially liquid asset underestimate their true contribution to aggregate liquidity.

Resalability constraints. Suppose that $\delta(s)$ decreases; i.e., the resalability of asset s decreases. From (63), output falls and ξ rises, but the effect on the price of asset s is ambiguous. To see this,

suppose that there is a single asset and the utility function is $u(q) = \frac{(q+b)^{1-\alpha} - b^{1-\alpha}}{1-\alpha}$, with b close to 0. From (64), the asset price is

$$\phi = \beta + \beta\sigma \{ \delta^{1-\alpha} A^{-\alpha} - \delta \}.$$

If $\alpha > 1$, then ϕ decreases with δ : as the resalability constraint becomes tighter, the price of the asset increases. In our fraud-based model, however, a decrease in $k(s)$ tightens the resalability constraint, and so the price of partially liquid assets declines.

Asset supply effects. In the KM model, the liquidity premia of all assets are proportional to ξ . From (63), an increase in the supply of any asset with $\delta(s) > 0$ and $\phi(s) > \beta$ results in a higher q , which reduces ξ , and consequently, prices of all assets traded in the DM fall. In our fraud-based model, by contrast, the prices of partially liquid assets, $s' \neq s$, are affected by their own supply, and do not respond to changes in the supply of asset s .

Furthermore, our model is better suited to address the observation that "asset demand curves slope down" (see Schleifer, 1986, and the limit-to-arbitrage literature more generally). To see this, suppose that there is a continuum of assets in infinitesimal supply. Then, an increase in the supply of any asset has no effect on its price. In our model, by contrast, an increase in the supply of a partially liquid asset affects its price, even in the limit where the supply of all assets is small.

Optimal liquidity provision. In the model with KM resalability constraints, aggregate liquidity, $L = \sum_{s \in S} \delta(s)A(s)$, is strictly increasing with the supply of liquid or partially liquid assets. The monetary authority thus can always satiate liquidity needs in the DM by issuing sufficient liquid assets. In a version of the model with fiat money, this finding suggests that the Friedman rule implements the first best. In our fraud-based model, however, aggregate liquidity has an upper bound independent of asset supplies:

$$L = \sum_{s \in S} \min \left[A(s), \frac{k(s)}{\beta\sigma} \right] \leq \sum_{s \in S} \frac{k(s)}{\beta\sigma}.$$

Therefore, even if assets are abundant, the liquidity they provide may not be sufficient for the economy needs, due to the recognizability problem of assets. (Some researchers, e.g., Caballero

(“On the macroeconomics of asset shortages”, 2006) , have emphasized the shortage of safe assets to explain the recent crisis and dynamics of asset prices.) This result implies that, in the context of the model with fiat money, the Friedman rule may not implement the first best if the cost to produce counterfeits is sufficiently low.

Dynamics. We now show that the model with KM resalability constraints generates different dynamics than our fraud-based model. In order to make our point, it suffices to consider an economy with a single asset. This asset is infinitely lived and generates a dividend equal to ζ units of numéraire good in every CM. In the DM an agent can only spend a constant fraction, δ , of his asset holdings. A buyer in the CM of $t - 1$ faces the following problem:

$$\max_{a_t, d_t, q_t} \left\{ - \sum_{s \in S} [\phi_{t-1} - \beta(\phi_t + \zeta)] a_t + \sigma \beta [u(q_t) - q_t] \right\} \quad (67)$$

$$\text{s.t.} \quad -q_t + (\phi_t + \zeta) d_t = 0 \quad (68)$$

$$d_t \in [0, \delta a_t]. \quad (69)$$

The first-order conditions are

$$\begin{aligned} u'(q_t) &= 1 + \frac{\xi_t}{\sigma \beta} \\ \phi_{t-1} &= \beta(\phi_t + \zeta) + \delta \lambda_t \\ \lambda_t &= \xi_t(\phi_t + \zeta), \end{aligned}$$

where ξ_t and λ_t are the Lagrange multipliers associated with the seller’s participation constraint, (68), and the resalability constraint, (69), respectively. Following the same method as in the paper, the dynamics of asset prices are

$$\phi_{t-1} = (\phi_t + \zeta) \times \begin{cases} (\beta + \delta \xi_t) & \text{if } \delta > 0 \\ \beta & \text{if } \delta = 0 \end{cases} \quad (70)$$

In a liquidity-constrained equilibrium, output is given by

$$q_t = \delta(\phi_t + \zeta) A.$$

The unique solution for the price of illiquid assets is $\phi_t = 1/r$.

Suppose next that the asset is at least partially liquid, $\delta > 0$. From (70) and using $\xi_t = \beta\sigma [u'(q_t) - 1]$, the asset price solves

$$\phi_{t-1} = \beta \{1 + \delta\sigma \{u'[\delta(\phi_t + \zeta)A] - 1\}\} (\phi_t + 1). \quad (71)$$

Consider the limit when $\zeta \rightarrow 0$ and take the functional form $u(q) = \frac{(q+b)^{1-\alpha} - b^{1-\alpha}}{1-\alpha}$ with b close to 0. The difference equation (71) can be reexpressed as

$$\phi_{t-1} = \beta \{(1 - \delta\sigma)\phi_t + \delta^{1-\alpha}\sigma A^{-\alpha}\phi_t^{1-\alpha}\}.$$

The steady-state value of the asset price is

$$\bar{\phi} = \frac{1}{A} \left(\frac{\delta^{1-\alpha}\beta\sigma}{1 - \beta + \delta\sigma\beta} \right)^{\frac{1}{\alpha}}.$$

With this in mind we obtain:

$$\left. \frac{\partial\phi_{t-1}}{\partial\phi_t} \right|_{\phi_t=\bar{\phi}} = 1 - \alpha [1 - \beta(1 - \sigma\delta)].$$

If $\left. \frac{\partial\phi_{t-1}}{\partial\phi_t} \right|_{\phi_t=\bar{\phi}} < -1$ then the equilibrium is indeterminate, which occurs if $\alpha > \frac{2}{1-\beta(1-\sigma\delta)}$. The price of a partially liquid asset can be indeterminate and exhibit endogenous fluctuations provided that agents are sufficiently risk-averse. In our fraud-based model only liquid assets exhibit fluctuations in prices.

E.2 Match-specific resalability constraints

We consider a version of the model where some assets are accepted in a fraction of the matches, as in Lagos (2010) or Lester, Postlewaite, and Wright (2011). For tractability, and following the above mentioned papers, we assume that there are only two assets, $S = \{1, 2\}$. Asset 1, which is accepted in all matches, is said to be liquid, while asset 2 is accepted in a fraction $\delta \in [0, 1]$ of the matches, and is called a partially liquid asset. We show that, just as KM resalability constraints, match-specific resalability constraints have different implications than the endogenous resalability constraints studied in the paper:

1. Flight to liquidity. In the model with match specific resalability constraint constraints, increasing the frequency of trades always increase asset prices and welfare. We can obtain the opposite in our model.

2. Open market operations. With match specific resalability constraints, a budget balanced open market purchase of partially liquid assets in exchange for liquid assets leads to an increase in welfare. In our model we obtain the opposite effect.
3. Increasing the resalability of the partially liquid asset. With match specific resalability constraint, increasing δ can reduce the price of liquid assets. In contrast, in our model if the recognizability of a partially liquid asset improves; i.e., $k(s)$ increases, its price increases but the price of liquid assets is unaffected.

The model. Let $\{\underline{q}, \underline{d}(s)\}$ denote the offer in a match where only asset 1 is accepted, and $\{\bar{q}, \bar{d}(s)\}$ the offer in a match where both assets are accepted. The buyer's problem in the CM is

$$\max_{a(s), \{\underline{q}, \underline{d}(s)\}, \{\bar{q}, \bar{d}(s)\}} - \sum_{s \in S} [\phi(s) - \beta] a(s) + \beta \sigma \{ \delta [u(\bar{q}) - \bar{q}] + (1 - \delta) [u(\underline{q}) - \underline{q}] \}$$

subject to

$$\begin{aligned} d(1) + d(2) &= \bar{q} \\ d(1) &= \underline{q} \\ d(s) &\in [0, a(s)], \quad \text{for all } s \in S. \end{aligned}$$

The first-order conditions are

$$\begin{aligned} \bar{\xi} &= \beta \sigma \delta [u'(\bar{q}) - 1] \\ \underline{\xi} &= \beta \sigma (1 - \delta) [u'(\underline{q}) - 1] \\ \phi(s) &= \beta + \nu(s), \quad s \in \{1, 2\} \\ \nu(1) &= \bar{\xi} + \underline{\xi} \\ \nu(2) &= \bar{\xi}, \end{aligned}$$

where $\nu(s)$ is the Lagrange multiplier associated with the feasibility constraint for asset s , and $\bar{\xi}$ and $\underline{\xi}$ are the Lagrange multipliers associated with the seller's participation constraints in the two types of matches.

The marginal welfare effect of an open-market operation that consists in the Central bank selling liquid assets and purchasing less liquid ones is measured by

$$\begin{aligned}\Delta W &= \beta\sigma \{ \delta [u'(\bar{q}) - 1] [\Delta A(1) + \Delta A(2)] + (1 - \delta) [u'(\underline{q}) - 1] \Delta A(1) \} \\ &= [\phi(2) - \beta] \Delta A(2) + [\phi(1) - \beta] \Delta A(1).\end{aligned}$$

The change in welfare is equal to a weighted sum of the changes in asset supplies where the weight assigned to each asset is its holding cost. We already saw that this measure was inaccurate in our fraud-based model because the holding cost of an asset has a direct effect on its resalability that is ignored by reduced form models.

We distinguish three regimes.

Regime 1: $A(1) \geq q^*$ Then liquidity is abundant in all matches. In this case, $\phi(1) = \phi(2) = \beta$ and $\underline{q} = \bar{q} = q^*$. The first best is achieved and all assets are priced at their fundamental value.

Regime 2: $A(1) < q^*$ and $A(1) + A(2) \geq q^*$ Then liquidity is scarce in a fraction of the matches. In this case, $\bar{q} = q^*$, $\underline{q} = A(1)$, and

$$\begin{aligned}\phi(1) &= \beta + \beta\sigma(1 - \delta) [u'(\underline{q}) - 1] \\ \phi(2) &= \beta.\end{aligned}$$

Asset 2 is illiquid (at the margin) while asset 1 is liquid. The price of the liquid asset increases with σ and decreases with its own supply. A policymaker who purchases asset 2 in exchange for asset 1 raises output in matches where liquidity is scarce. These findings are consistent with our fraud-based model. An increase in σ raises the velocity of illiquid assets whereas in our fraud-based model the liquidity of illiquid assets is unaffected by σ . Indeed, in our model σ reduces the resalability of illiquid assets.

Regime 3: $A(1) < q^*$ and $A(1) + A(2) < q^*$ Then liquidity is scarce in all matches. In this case, $\bar{q} = A(1) + A(2)$, $\underline{q} = A(1)$, and

$$\begin{aligned}\phi(1) &= \beta + \beta\sigma [u'(\underline{q}) - 1] + \beta\sigma\delta [u'(\bar{q}) - u'(\underline{q})] \\ \phi(2) &= \beta + \beta\sigma\delta [u'(\bar{q}) - 1].\end{aligned}$$

According to our terminology, asset 1 is liquid while asset 2 is partially liquid. Asset prices increase with σ and decrease with the supply of liquid assets. These findings are consistent with those of a model with proportional liquidity constraints. In contrast, in our fraud-based model the prices of partially liquid assets decreases with σ because σ affects incentives to commit frauds and resalability. The supply of illiquid assets only affects $\phi(2)$. If asset 2 can be used more often, δ increases, its price increases, but the price of asset 1 decreases. In contrast, in our fraud-based model the resalability of partially liquid assets does not affect the price of liquid assets.

Suppose the Central Bank purchases some asset 2 by selling some asset 1. Budget balance implies:

$$\phi(1)\Delta A(1) + \phi(2)\Delta A(2) = 0 \quad \Rightarrow \quad \Delta A(2) = -\frac{\phi(1)}{\phi(2)}\Delta A(1),$$

where $\Delta A(1) > 0$ and $\Delta A(2) < 0$. The change in expected welfare in the DM is:

$$\begin{aligned} \Delta W &= \beta\sigma \left\{ \delta [u'(\bar{q}) - 1] [\Delta A(1) + \Delta A(2)] + (1 - \delta) [u'(\underline{q}) - 1] \Delta A(1) \right\} \\ &= \beta\sigma \left\{ \delta [u'(\bar{q}) - 1] \left[\frac{\phi(2) - \phi(1)}{\phi(2)} \right] \Delta A(1) + (1 - \delta) [u'(\underline{q}) - 1] \Delta A(1) \right\} \\ &= \beta \left(\frac{1}{\phi(2)} - \frac{1}{\phi(1)} \right) \phi(1)\Delta A(1) > 0. \end{aligned}$$

As before, the change in welfare is the product of the yield spread and the quantity sold. In contrast, in our fraud based model such an open market operation can lead to a reduction in social welfare by reducing aggregate liquidity.

F Comparison: corporate-finance-based models

Holmstrom and Tirole (2011, Section 1.2), HT henceforth, proposed a model of investment with moral hazard. The baseline economy is composed of a firm and competitive outside investors that are risk neutral. There are two periods. In the initial period the firm receives an opportunity to invest. The cost of the investment is I . In the second period, the investment pays off R if it is a success and 0 otherwise. The probability of success depends on an unobserved action taken by the entrepreneur of where to invest I . There is an efficient technology with probability of success p_h and an inefficient technology with probability of success $p_\ell < p_h$. The use of the inefficient technology allows the entrepreneur to capture some private benefit, B . HT shows that the maximum amount the firm can borrow against the returns of its investment, the pledgeable income, is

$$p_h \left(R - \frac{B}{p_h - p_\ell} \right).$$

Our model and HT have similarities: (i) They both propose a theory of endogenous liquidity constraints; (ii) Both models incorporate some hidden actions: in our model, the decision to produce fraudulent assets, in HT the decision of an investment technology. (iii) Opportunistic behavior (default or fraud) does not materialize in equilibrium; (iv) Both theories have implications for how liquidity affect asset prices. In our view, the most salient differences between the two approaches are the following:

1. Timing of hidden actions. The timing of the moral hazard problem differs in a crucial way. In HT, the hidden action follows the investment opportunity. In our model, by contrast, the hidden action is taken before the agent receives a spending opportunity. This is because lack of commitment and enforcement forces agents to hold assets in advance of spending opportunities, and because the hidden action determines the quality of these assets. As a consequence, in our model the holding cost of the asset as well as the frequency of trades, σ , determine the incentives to take different actions. This would not be the case if the action was taken after the consumption opportunity is realized. As shown in the paper, our novel comparative statics arise precisely because the frequency of trades, σ , and the asset holding cost enter the resalability constraints.

2. Liquidity differences in secondary asset markets. While HT endogenizes inside liquidity (the aggregate supply of corporate claims) our model endogenizes differences in outside assets liquidity. In particular:
 - (a) Our model can explain a shortage of liquidity even when assets are abundant.
 - (b) Our model can generate violations of the no-arbitrage principle: assets with identical cash flows can be traded at different prices. By contrast, in the Liquidity Asset Pricing Model (see Chapter 4 in HT's book), a no-arbitrage principle holds: the price of all outside assets can be found by applying some common stochastic discount factor to their cash flow.
3. Market structure. We consider different market structures. While HT stays as close as possible to Arrow-Debreu, we consider an over-the-counter market with bilateral meetings and bargaining. The frequency of trades in the OTC market matter for incentives, resalability constraints, and asset prices. Moreover, the presence of bilateral meetings provides a natural framework to discuss issues related to private information and signaling.

G Risk sharing in OTC markets

This appendix provides a simple model of an OTC market for bilateral risk sharing, such as the market of interest-rate swaps. We show that, up to small modifications, this OTC-market model can be mapped into our benchmark model.

G.1 The model

There are two dates, $t \in \{0, 1\}$, one good, and a finite set S of assets. Asset $s \in S$ supply is $A(s)$ and pays off at the end of $t = 1$ a dividend normalized to one. The economy is populated by two kinds of traders who seek to share risk with each others: "buyers", who seek to purchase risk-sharing services, and sellers, who seek to provide such services. At time zero, buyers can purchase assets in a perfectly competitive centralized market (CM), where asset $s \in S$ is sold at price $\phi(s)$. At $t = 1$, a fraction $\sigma \in (0, 1]$ of buyers and sellers are matched at random in pairs in an OTC market.

We specify the demand for risk sharing services in the OTC market so that the contracts signed by buyers and sellers resemble "swaps," whereby buyers and sellers make a fixed-for-floating exchange of cash flow streams. Precisely, we assume that a buyer enters the OTC market with an obligation to make a floating payment to some outside counterparty at time $t = 1$. Formally, the buyer is endowed with one share of some risky tree with payoff x at the end of $t = 1$, for some negative random variable x with twice continuously differentiable moment generating function. In a pairwise meeting, the buyer makes an offer to the seller specifying that, at the end of $t = 1$, the seller will make the floating payment $-qx$ to the buyer, and the buyer will make some fixed payment to the seller. We assume one sided limited commitment: the seller can commit to his payment at the end of $t = 1$ but the buyer cannot. Therefore, the buyer's promise to pay at the end of $t = 1$ has to be fully collateralized by assets acquired at $t = 0$. As in the paper, we assume that the buyer can transfer fraudulent assets. Namely, at $t = 0$, the buyer can produce any quantity of fraudulent assets of type s at a fixed cost $k(s)$. The terminal value of a fraudulent asset at the end of $t = 1$ is zero.

G.2 Payoffs

Using the same notations as in the paper, the sum of all payoffs of a buyer is:

$$c_B = - \sum_{s \in S} \left\{ k(s) [1 - \chi(s)] + [\phi(s) - 1] a(s) \right\} + \mu \left(x \cdot (1 - q) - \sum_{s \in S} \tau(s) \right) + (1 - \mu)x, \quad (72)$$

and the sum of all payoffs of a seller is:

$$c_S = \mu \left(x \cdot q + \sum_{s \in S} \tau(s) \right). \quad (73)$$

To simplify the analysis, we assume that buyers and sellers evaluate the utility associated with their random payoffs in the following way. As in Shi (1997) and Shimer (2010), we organize buyers and sellers in large families that share all the idiosyncratic risk created by random matching and mixed strategies. We assume that all the floating payments, x , are perfectly correlated across buyers, since otherwise all payment risks would be diversified away within each family, eliminating families' demand for risk sharing services.

In our OTC context, it is natural to interpret a family as a financial institution employing a large number of traders. The systematic payment risk, represented by x , could represent the risk associated with fluctuations in interest rates. Financial institutions share such risk by trading in OTC markets for interest rate swaps.

We assume that families derive CARA utility over the aggregate payoff of all their traders. Restricting attention to symmetric equilibria where all buyers make the same offer in the OTC market, and aggregating the payoffs (72) across all the traders in a buyer's family, it follows that a buyer's family has a payoff of the form:

$$\text{deterministic constant} + \sigma \bar{\pi} (1 - \bar{q}) \cdot x + (1 - \sigma \bar{\pi}) \cdot x,$$

where \bar{q} is the offer of a representative buyer, and $\bar{\pi}$ is the probability that a representative seller accepts the offer. The first term contains the payoffs associated with asset transfers and fraudulent asset productions. These payoffs are deterministic when aggregated at the family level. The second term is the quantity of risk collectively held by buyers whose offer was accepted by some seller. The third term is the quantity of risk collectively held by the buyers who did not meet a seller or

whose offer was rejected. Now, let α be the coefficient of absolute risk aversion of a buyer's family, and define

$$\gamma_B \equiv \alpha [\sigma \bar{\pi}(1 - \bar{q}) + (1 - \sigma \bar{\pi})]. \quad (74)$$

It then follows that a buyer's family has a marginal utility

$$\Lambda_B(x) = \text{deterministic constant} \times e^{-\gamma_B x}. \quad (75)$$

over payoffs. As is standard in large family models, an individual trader in a buyer's family uses the marginal utility $\Lambda_B(x)$ as a stochastic discount factor for his payoff. That is, an individual buyer in the family evaluates his payoff according to $E[\Lambda_B(x)c_B]$, taking $\Lambda_B(x)$ as given, where c_B is defined in equation (72). Similarly, an individual seller in a seller's family evaluates his payoff according to $E[\Lambda_S(x)c_S]$, where $\Lambda_S(x) = \text{deterministic constant} \times e^{-\gamma_S x}$, and

$$\gamma_S \equiv \alpha \sigma \bar{\pi} \bar{q}.$$

Now, let us calculate the expected payoffs of buyers and sellers, conditional on all random variables except x . Note that, conditional on all other random variables, both c_B and c_S can be written as affine functions $K_0 + K_1 x$, for some constants K_0 and K_1 . Therefore, to obtain the expected payoff of a trader conditional on all random variables except x , all we need to do is to calculate

$$\mathbb{E}[e^{-\gamma x}(K_0 + K_1 x)] = \mathbb{E}[e^{-\gamma x}] \times [K_0 - z(\gamma)K_1], \quad \text{where } z(\gamma) \equiv -\frac{\mathbb{E}[xe^{-\gamma x}]}{\mathbb{E}[e^{-\gamma x}]}$$

Moreover, we show in subsection G.5.1 that $z(\gamma)$ is positive and increasing. This immediately leads to the following Lemma:

Lemma 1 *Let $z(\gamma) \equiv -E[xe^{-\gamma x}]/E[e^{-\gamma x}]$. Then $z(\gamma)$ is a positive and increasing function. Moreover, conditional on all random variables except x , the expected payoff of a buyer is, up to some positive constant of proportionality:*

$$-\sum_{s \in S} \left\{ k(s) \mathbb{I}_{\{\bar{a}(s) > 0\}} + [\phi(s) - 1] a(s) \right\} + \mu \left(z(\gamma_B) \cdot q - \sum_{s \in S} \tau(s) \right) - z(\gamma_B),$$

and the expected payoff of a seller is, up to some positive constant of proportionality:

$$\mu \left(-z(\gamma_S) \cdot q + \sum_{s \in S} \tau(s) \right).$$

This Lemma shows that there is a very simple mapping between the model of the paper and the present OTC market setup: the utility of a buyer is $z(\gamma_B) \cdot q$, and the cost of a seller is $z(\gamma_S) \cdot q$. The two constants that individual buyers and sellers take as given, γ_B and γ_S , are endogenous and depend on the amount of risk that a typical family holds after trading in the OTC market.

G.3 Equilibrium

Proposition 4 *In any symmetric equilibrium: $z(\gamma_B) \geq z(\gamma_S)$; $\phi(s) \geq 1$ for all $s \in S$; genuine assets are acquired with probability one; the asset demands and offer solve:*

$$\max_{q, \{a(s), d(s)\}} \left\{ - \sum_{s \in S} [\phi(s) - 1] a(s) + \sigma [z(\gamma_B) - z(\gamma_S)] q \right\} \quad (76)$$

$$\text{s.t.} \quad -z(\gamma_S)q + \sum_{s \in S} d(s) = 0 \quad (77)$$

$$d(s) \leq \frac{k(s)}{\phi(s) - (1 - \sigma)}, \quad \text{for all } s \in S \quad (78)$$

$$0 \leq d(s) \leq a(s), \quad \text{for all } s \in S. \quad (79)$$

$$0 \leq q \leq 1. \quad (80)$$

If, in addition, $\gamma_B > \gamma_S$, then the seller accepts the offer with probability one.

To understand the first restriction, note that if $\gamma_B < \gamma_S$, then there would be negative gains from transferring $q > 0$ trees, implying that $q = 0$ or $\pi = 0$ in any pairwise meeting. But then, from (74) and (75), it follows that $\gamma_B = \alpha > \gamma_S = 0$, a contradiction. The restriction that $\phi(s) \geq 1$ follows from elementary no-arbitrage reasoning: indeed, if the asset price were less than its terminal payoff, $\phi(s) < 1$, then the asset demand would be infinite.

The asset demands derived above imply the following equilibrium asset prices. We use the same notations as in the paper and we let: $\kappa(s) \equiv k(s)/A(s)$, $\theta(s) \equiv \min\{1, \kappa(s)/\sigma\}$, and $L \equiv \sum_{s \in S} \theta(s)A(s)$. First, consider the case when there is enough liquidity to achieve the efficient amount of risk sharing in the OTC market.

Proposition 5 *Suppose*

$$L > z(\alpha\sigma q^*) \cdot q^*, \quad \text{where } \sigma q^* = \min \left\{ \sigma, \frac{1}{2} \right\}. \quad (81)$$

Then, in equilibrium: $\phi(s) = 1$ for all $s \in S$. Moreover, if $\sigma > 1/2$, then $\sigma\pi q = 1/2$ and $\gamma_S = \gamma_B$. If, on the other hand, $\sigma < 1/2$, then $q = \sigma$ and $\gamma_S < \gamma_B$.

The second case is:

Proposition 6 *Suppose $L < z(\alpha\sigma q^*)q^*$. Then, in equilibrium, q solves $z(\alpha\sigma q)q = L$, $\gamma_B = 1 - \sigma q$, and $\gamma_S = \sigma q$. Letting $\kappa(s) \equiv k(s)/A(s)$, asset prices are given by:*

1. *Liquid assets: if $\kappa(s) \geq \sigma \frac{z(\gamma_B)}{z(\gamma_S)}$, $\phi(s) = 1 + \sigma \left[\frac{z(\gamma_B)}{z(\gamma_S)} - 1 \right]$.*
2. *Partially liquid assets: if $\sigma < \kappa(s) < \sigma \frac{z(\gamma_B)}{z(\gamma_S)}$, then $\phi(s) = 1 - \sigma + \kappa(s)$.*
3. *Illiquid assets: if $\kappa(s) \leq \sigma$, then $\phi(s) = 1$.*

G.4 Comments

The results of this Appendix confirm the robustness of our benchmark model: Proposition 4 is almost identical to Proposition 1 in the main body of the paper, and the asset-pricing formulas of Propositions 5 and 6 are essentially the same as in Proposition 2 in the main body of the paper. In addition, the OTC model offers a few novel insights that we describe below.

The efficient trade. Because of the large family assumption, the efficient trade in a bilateral meeting depends on the matching technology. To see why, note that the total amount of risk transferred by the buyer's family is equal to σq . Given that all buyers hold one tree, the maximum feasible risk transfer is σ . In addition, on aggregate, a family of buyers will not find it optimal to transfer more than $1/2$ shares of risky tree – i.e., the amount that results in perfect risk sharing. Taken together, this implies that the efficient trade in a bilateral meeting solves:

$$\sigma q^* = \min \left\{ \sigma, \frac{1}{2} \right\}. \quad (82)$$

The condition (81) for assets to be priced at their fundamental value. The condition has the same interpretation as in the paper: if there is enough liquidity to realize the efficient trade, in equation (82), then assets are priced at their fundamental value. To derive it explicitly, note that the cost to an individual seller of providing the efficient amount of insurance is $z(\gamma_S)q^*$. But if

all other sellers in the family provide that efficient amount of insurance, then the cost is $z(\alpha\sigma q^*)q^*$, where q^* is defined in equation (82). For this to be the basis of an equilibrium, we need that a buyer's liquidity is sufficient to cover that cost; i.e. $L \geq z(\alpha\sigma q^*)q^*$, as stated in condition (81).

Some new insights coming out of condition (81). A difference between condition (81) and its counterpart in the benchmark model is that σ affects fraud incentives in two ways: as in the paper, an increase in σ increases the incentives to commit fraud, because traders are more likely to encounter counterparties in the DM, who cannot distinguish between fraudulent and genuine assets. But there is now another effect that can go in the opposite direction: if $\sigma \geq 1/2$ an increase in σ allows a buyer's family to transfer less risk per match and achieve the same aggregate amount of risk transfer. Therefore, each buyer is in charge of a smaller trade, and so he has less incentives to commit fraud.

The frequency of trades, σ , and the liquidity premium. When $\phi(s) > 1$ for some s , then the maximum liquidity premium is $\sigma \left(\frac{z(\gamma_B)}{z(\gamma_S)} - 1 \right)$. But $\gamma_B = \alpha(1 - \sigma q)$, $\gamma_S = \sigma q$, and $z(\alpha\sigma q)q = L$. So σq and γ_S are increasing functions of σL , while γ_B is a decreasing function of σL . Taken together, this means that $z(\gamma_B)/z(\gamma_S)$ is a decreasing function of σL , so we can write the maximum liquidity premium as:

$$\sigma \times (\text{some decreasing function of } \sigma L).$$

One sees that, contrary to the benchmark model, the effect of σ on the maximum liquidity premium is ambiguous. One of the effect of the benchmark model is still present: an increase in σ means that a trader finds a match with higher probability, which makes liquidity more valuable and this increases the liquidity premium. But there is another effect going in the opposite direction. When σ is higher, the family can share more risk in other matches: this lowers the value of sharing risk in any particular match, makes liquidity less valuable, and thus decreases the liquidity premium. Either effect can dominate: if $\gamma_B \simeq \gamma_S$, and in this sense liquidity is plentiful, then the second effect dominates. If every $k(s)$ is less than $\sigma A(s)$, and in this sense liquidity is scarce, then σL is independent on σ , and the first effect dominates.

G.5 Proofs

G.5.1 Proof of Lemma 1

All we need to show is that $z(\gamma)$ is positive and increasing. Positivity follows because, since $\gamma \geq 0$ and $x < 0$, $xe^{-\gamma x} \leq x < 0$. Taking expectations on both sides leads to $E[xe^{-\gamma x}] < 0$. Now, differentiating $z(\gamma)$ with respect to γ leads to:

$$z'(\gamma) = \frac{\mathbb{E}[x^2 e^{-\gamma x}] \mathbb{E}[e^{-\gamma x}] - \mathbb{E}[xe^{-\gamma x}]^2}{\mathbb{E}[e^{-\gamma x}]^2}$$

But note that:

$$\mathbb{E}[xe^{-\gamma x}]^2 = \mathbb{E}[-xe^{-\gamma x}]^2 = \mathbb{E}\left[(x^2 e^{-\gamma x})^{1/2} (e^{-\gamma x})^{1/2}\right]^2 < \mathbb{E}[x^2 e^{-\gamma x}] \mathbb{E}[e^{-\gamma x}]$$

where the inequality follows from the Cauchy-Schwarz inequality, given that $x^2 e^{-\gamma x}$ and $e^{-\gamma x}$ are linearly independent. Plugging this back shows that $z'(\gamma) > 0$.

G.5.2 Proof of Proposition 4

The proof that $\gamma_B \geq \gamma_S$ and $\phi(s) \geq 1$ follows from the argument stated in the text. Next, we follow the same steps as in the paper. Claims 1 through 4 follow identically, and we can rewrite the auxiliary problem as:

$$\begin{aligned} & \max_{q, d, \pi} - \sum_{s \in S} [\phi(s) - 1] d(s) + \sigma \pi [z(\gamma_B) - z(\gamma_S)] q \\ & \text{s.t.} \quad -z(\gamma_S)q + \sum_{s \in S} d(s) = 0 \\ & \quad k(s) \geq [\phi(s) - 1 + \sigma \pi] d(s). \end{aligned}$$

To conclude the proof, there are three cases to consider. The first case is when $\gamma_B > \gamma_S$ and $\phi(s_0) < 1 + \sigma \left[\frac{z(\gamma_B)}{z(\gamma_S)} - 1 \right]$ for some $s_0 \in S$. Then the value of the auxiliary problem is positive – for instance, a small offer $q = d(s_0)/z(\gamma_S) > 0$, $d(s) = 0$ for $s \neq s_0$, and $\pi = 1$ would yield a positive payoff. Applying the arguments following Claim 4 in the paper, we then obtain that, for any equilibrium of the game, $\pi = 1$, $\eta(s) = 1$ for all s , and the asset demands and offer solve the problem of Proposition 4.

The second case is when $\gamma_B > \gamma_S$ and $\phi(s) \geq 1 + \sigma \left[\frac{z(\gamma_B)}{z(\gamma_S)} - 1 \right]$ for all $s \in S$. Then the objective of the buyer becomes, after substituting in the participation constraint:

$$\sum_{s \in S} \left(1 + \sigma \pi \left[\frac{z(\gamma_B)}{z(\gamma_S)} - 1 \right] - \phi(s) \right) d(s) \leq 0.$$

Regardless of π , the buyer can achieve this zero upper bound by offering $q' = d'(s) = 0$. Therefore, the buyer's value in any equilibrium must be zero. Given that any equilibrium outcome is feasible for the auxiliary problem, it thus follows that any equilibrium outcome solves the auxiliary problem. In particular, by Claim 4, $\eta(s) = 1$ for all s . Moreover, in a symmetric asset market equilibrium, all buyers hold $a(s) = A(s) > 0$. Since $\phi(s) > 1$ for all $s \in S$, the equilibrium offer must be $d(s) = A(s) > 0$ and $z(\gamma_S)q = \sum_{s \in S} A(s)$. Substituting the previous expression into the buyer's objective derived above yields: $\sum_{s \in S} \left(1 + \sigma \pi \left[\frac{z(\gamma_B)}{z(\gamma_S)} - 1 \right] - \phi(s) \right) A(s) = 0$. It thus follows that $\pi = 1$ and $\phi(s) = 1 + \sigma \left[\frac{z(\gamma_B)}{z(\gamma_S)} - 1 \right]$. Clearly, the equilibrium asset demands and offer solve the problem of Proposition 4.

The third case is when $\gamma_B = \gamma_S$. Then, the value of the auxiliary problem must be zero. As above, any equilibrium must solve the auxiliary problem, and so by Claim 4 we have that $\eta(s) = 1$ for all s . Also note that in a symmetric asset market equilibrium, $a(s) = A(s)$ for all buyers. If $\phi(s) > 1$ for some s , then $d(s) = a(s) = A(s)$, implying that the value of the auxiliary problem is negative. Thus, $\phi(s) = 1$ for all s . Clearly, the equilibrium asset demand and offer solve the problem of Proposition 4.

G.5.3 Proofs of Proposition 5 and 6

Case 1: if $\gamma_B = \gamma_S$. In that case, the value of the auxiliary problem is zero and, as shown at the end of the proof of Proposition 4, $\phi(s) = 1$ for all $s \in S$. Also, using the formulas (74) and (75) for γ_B and γ_S , one obtains that $2\sigma\pi q = 1$. Since $\pi \in [0, 1]$, this implies that $\sigma \geq 1/2$. Plugging this back into (75), we find $\gamma_S = \alpha/2$. Using the participation constraint of a seller, we find:

$$z(\gamma_S)q = z \left(\frac{\alpha}{2} \right) \frac{1}{2\sigma\pi} = \sum_{s \in S} d(s) \leq \sum_{s \in S} \min \left\{ \frac{k(s)}{\sigma}, A(s) \right\} = L \implies \sigma L \geq z \left(\frac{\alpha}{2} \right) \frac{1}{2}.$$

where: the inequality follows from the resalability constraint, (78), evaluated at $\phi(s) = 1$, and the feasibility constraint (79); the implication follows from the fact that $\pi \in [0, 1]$. Taken together, this

means that a necessary condition for $\gamma_B = \gamma_S$ is $\sigma \geq 1/2$ and $\sigma L \geq z\left(\frac{\alpha}{2}\right)\frac{1}{2}$. Conversely, one easily verifies that, if these two conditions hold, then one can construct an equilibrium with $\gamma_B = \gamma_S$.

Moving on to the case when $\gamma_B > \gamma_S$, we let ξ be the multiplier on the participation constraint (77), $\lambda(s)$ the multiplier on the resalability constraint (78), $\nu(s)$ the multiplier on the feasibility constraint (79), and ψ the multiplier on the constraint (80) that the tree transfer cannot exceed one unit. The first-order conditions can be written:

$$\xi = \sigma \left(\frac{\sigma\gamma_B - \psi}{\sigma\gamma_S} - 1 \right), \quad \xi = \lambda(s) + \nu(s), \quad \text{and} \quad \phi(s) = 1 + \nu(s).$$

We consider to cases:

Case 2: if $\gamma_B > \gamma_S$ and $\xi = 0$. Then $\phi(s) = 1$ for all $s \in S$. Moreover, $\psi > 0$ and so $q = 1$. But then $\gamma_S < \gamma_B$ implies that $\sigma < 1/2$. Also, from the participation constraint, $z(\gamma_S)q = z(\alpha\sigma) = \sum_{s \geq S} d(s) \leq L$, which can be rearranged as $\sigma L \geq z(\alpha\sigma)\sigma$. Conversely, if $\sigma < 1/2$ and $\sigma L \geq z(\alpha\sigma)\sigma$, then the above is the basis of an equilibrium with $\gamma_B > \gamma_S$ and $\xi = 0$.

Case 3: if $\gamma_B > \gamma_S$ and $\xi > 0$. Following the same arguments as in the paper, we obtain the following characterization of asset prices:

1. If $\kappa(s)/\sigma \leq 1$, then $\phi(s) = 1$.
2. If $\kappa(s)/\sigma \geq \frac{\sigma\gamma_B - \psi}{\sigma\gamma_S}$, then $\phi(s) = 1 + \sigma \left(\frac{\sigma\gamma_B - \psi}{\sigma\gamma_S} - 1 \right)$.
3. Otherwise, $\phi(s) = 1 + \sigma + \kappa(s)$.

In addition, $\sum_{s \in S} d(s) = L$; i.e., all liquidity is used up in the OTC transaction. Using the participation constraint $L = z(\gamma_S)q$ and keeping in mind that $\gamma_S = \alpha\sigma q$, we obtain that $z(\alpha\sigma q)q = L$. If $q < 1$, we have that $\psi = 0$, and prices are the same as in Proposition 6. If $q = 1$, then we can pick any $\psi \in \left[0, \sigma [z(\gamma_B) - z(\gamma_S)] \right]$. In all case, $z(\alpha\sigma q)q = L$ and $q \leq 1$ implies that $z(\alpha\sigma) \geq L$, or $\sigma L \leq z(\alpha\sigma)\sigma$. Moreover, $\gamma_B > \gamma_S$ also implies that $q < 1/(2\sigma)$. Plugging this into $z(\alpha\sigma q)q = L$, we obtain that $\sigma L < z\left(\frac{\alpha}{2}\right)\frac{\alpha}{2}$. Conversely, if $\sigma L \leq z(\alpha\sigma)\sigma$ and $\sigma L < z\left(\frac{\alpha}{2}\right)\frac{\alpha}{2}$, then one can construct an equilibrium where $\gamma_B > \gamma_S$ and $\xi > 0$.

Wrapping up we find the following characterization, which implies Proposition 5 and 6:

- There exists an equilibrium with $\xi = 0$ and $\gamma_B = \gamma_S$ if and only if $\sigma L \geq z \left(\frac{\alpha}{2}\right) \frac{\alpha}{2}$ and $\sigma \geq \frac{1}{2}$.
- There exists an equilibrium with $\xi = 0$ and $\gamma_B > \gamma_S$ if and only if $\sigma L \geq z(\alpha\sigma)\sigma$ and $\sigma < \frac{1}{2}$.
- There exists an equilibrium with $\xi > 0$ and $\gamma_B > \gamma_S$ if and only if $\sigma L \leq z(\alpha\sigma)\sigma$ and $\sigma L < z \left(\frac{\alpha}{2}\right) \frac{\alpha}{2}$. If, moreover, the first inequality is strict, then the prices are the ones given in Proposition 6.

H A fraud-based CAPM

In this Appendix we embed our fraud friction into the standard Capital Asset Pricing Model (CAPM) of Sharpe (1964). We view the CM as a competitive market where professional investors buy and sell risky assets among each others. The DM is a decentralized market where these professional investors can re-sell these assets to consumers. The gains from trade in the DM arise because professional investors spread the risk of assets to the broader investment base represented by consumers; i.e., in the language of the paper, the DM is a market where professional investors buy risk sharing services from consumers. As in the main body of the paper, DM trades are threatened by fraud.

Interpretations. This fraud-based CAPM provides a simple representation of the securitization process. Intermediaries can be viewed, say, as securitizers who compete to purchase the mortgages of homeowners. The DM stands in for the securitization process, whereby mortgages are pooled together into mortgage based securities and sold to the general public. Exposure to fraud vary across types of mortgages, capturing the commonly held view that subprime mortgages are more susceptible to fraud than prime mortgages.

Stepping further away from the main paper, we can also think of professional investors buying and selling a broader range of assets, setting up an investment fund, and offering consumers to buy shares in their investment funds. In that case, fraudulent assets represent claims to assets that were not actually acquired, as in many cases of fraud filed by the Securities and Exchanges Commission (SEC). The cost of fraud can differ across assets because it may be easier to defraud investors on some part of a portfolio than on others: for instance, one may argue that it is easier to commit fraud on private equity, on opaque or complex assets, or on novel assets.

Main results. This appendix confirm the robustness of our results and show how they can be integrated in a textbook asset pricing model. We find that, if the cost of fraud is large enough relative to the gains from transferring risk in the DM, then it is easy for professional investors to re-sell assets to consumers, and there is full risk sharing between them. Moreover, assets are priced according to their fundamental value, which in this context means that the standard CAPM

formula holds. If the cost of fraud is not large enough, then there is partial risk sharing. Professional investors hold more risk than consumers: they cannot re-sell as much asset as they would like in the absence of fraud frictions. Just as in the main body of the paper, we obtain a three-tier characterization of assets, holding fundamentals (risk, dividend) fixed:

1. Liquid assets: if an asset cost of fraud is large enough, then it is fully re-sold to consumers in the DM, and the resalability constraint is slack. Such an asset is priced according to its risk contribution to consumers' portfolio. Thus, it ends up with a high price because consumers hold less risk than professional investors. Note that, for this class of assets, a modified version of CAPM holds after replacing the market portfolio by the portfolio of consumers.
2. Partially liquid assets: if an asset cost of fraud takes on intermediate values, then the asset is also fully re-sold to consumers in the DM. But the resalability constraint binds: as in the paper this means that the equilibrium price has to fall below the price of an otherwise identical liquid asset. The price of such an asset decreases with its own supply, even if this asset is negligible relative to the aggregate supply of risk – a feature that is reminiscent of empirical evidence from the literature on limits to arbitrage (see Schleifer, 1986). For this class of assets, no version of CAPM hold.
3. Illiquid assets: if an asset cost of fraud is low enough, then the asset is not fully re-sold to consumers. Professional investors have to retain some of it in their portfolio. Such an asset is priced according to its risk contribution to professional investors' portfolio. Thus, it ends up with a low price because professional investors hold more risk than consumers. Note that, for this class of assets, a modified version of CAPM holds after replacing the market portfolio by the portfolio of professional investors. The pricing formula is reminiscent of evidence about market segmentations, whereby agents who specialize in trading some assets (dealers, professional investors, etc...) somehow have difficulties re-selling these assets to non specialists. As a result, they end up holding too much of the risk of these assets, and they require a higher return for holding them (Collin-Dufresne, Goldstein, and Martin, 2001, and Gabaix, Krishnamurthy, and Vigneron, 2007).

H.1 The formal model

There two dates, $t \in \{0, 1\}$. The economy is populated by a measure one of professional investors and a measure one of consumers. An professional investor (consumer) is a coalition of n_P (n_C) risk averse agents with CARA utility

$$-\frac{n_P}{\theta} \mathbb{E} \left[e^{-\frac{\theta}{n_P}(c_0+c_1)} \right]$$

over consumption streams $(c_0, c_1) \in \mathbb{R}^2$, with a negative consumption being interpreted as production. The utility of a consumer is defined similarly, with n_P being replaced by n_C .

There is a finite set S of assets. The supply of asset $s \in S$ is $A(s)$, and at the end of $t = 1$ it pays off the dividend:

$$y(s) = \bar{y}(s) + \gamma(s)f,$$

where $\bar{y}(s)$ is a constant component and $\gamma(s)$ is the exposure to f , a random component affecting all assets, which we take to be normally distributed with mean zero and variance 1. For simplicity we assume that $\gamma(s) > 0$ and we impose the technical condition:

$$\bar{y}(s) - \gamma(s) \frac{\theta}{\min\{n_P, n_C\}} \sum_{s' \in S} \gamma(s') A(s') \geq 0.$$

We model the process through which assets are allocated from professional investors to consumers in the following way. We assume that, at $t = 0$, there is a Walrasian market (the CM) where only professional investors participate and buy the entire supply of financial assets. But professional investors only hold these assets temporarily: at $t = 1$ there is a decentralized market (the DM) where professional investors meet consumers and sell them these assets. Precisely, we assume that, at time $t = 1$ in the DM, there is full matching of professional investors and consumers.³⁶ A professional investor offers a consumer to buy up to $a_C = \{a_C(s)\}$ shares of each asset. She also offers a price schedule: if the consumer decides to purchase fractions $\pi = \{\pi(s)\} \in [0, 1]^S$ of the maximum offers, $\{a_C(s)\}$, then she has to pay $\tau(\pi)$, for some function such that $\tau(0) = 0$.

As in the main body of the paper, there is a moral hazard problem: the professional investor can defraud the consumer. Namely, instead of purchasing the entire offer $a_C(s)$ of asset s in the

³⁶Handling partial matching in this setup is made difficult because it creates an additional source of risk. One could circumvent that difficulty using the framework of Appendix G .

CM, she can buy only a fraction $\eta(s) \in [0, 1]$ and commit fraud on the remaining fraction, $1 - \eta(s)$, at cost $k(s)[1 - \eta(s)]$. That is, $k(s)$ is for asset s the cost of fraud per $\bar{y}(s)$ dollars of asset s face value. Summarizing, the timing of move is as follows:

- The professional investor first commits to an offer $\{a_C(s)\}$ and $\tau(\pi)$, and also chooses a portfolio $\{a_P(s)\}$ of asset to acquire for its own proprietary investment.
- Then, the professional investor chooses, for each asset s , the fractions $\eta(s)$ of $a_C(s)$ to purchase in the DM, and the fraction $1 - \eta(s)$ on which to commit fraud.
- The professional investor is matched with a consumer, and the consumer decides the fractions $\{\pi(s)\}$ to acquire. If the consumer accepts a fraction $\pi(s)$ of the offer, she receives $\eta(s)\pi(s)a_C(s)$ genuine assets.

The payoff of a family of professional investor, $c_0 + c_1$, is:

$$\tau(\pi) - \sum_{s \in S} \left\{ p(s) \left[a_P(s) + \eta(s)a_C(s) \right] + [1 - \eta(s)]k(s) \right\} + \sum_{s \in S} \left[a_P(s) + \eta(s)(1 - \pi(s))a_C(s) \right] y(s)$$

After plugging this expression into the CARA utility function, we find that the certainty-equivalent payoff can be written $\tau(\pi) + U_P(a_P, a_C, \eta, \pi)$ where

$$U_P(a_P, a_C, \eta, \pi) = - \sum_{s \in S} \left\{ -p(s) \left[a_P(s) + \eta(s)a_C(s) \right] + [1 - \eta(s)]k(s) \right\} + \sum_{s \in S} \left[a_P(s) + \eta(s)(1 - \pi(s))a_C(s) \right] \bar{y}(s) - \frac{\theta}{2n_P} \left\{ \sum_{s \in S} \left[a_P(s) + \eta(s)(1 - \pi(s))a_C(s) \right] \gamma(s) \right\}^2.$$

Similarly, the certainty-equivalent payoff of a consumers' family is $-\tau(\pi) + U_C(a_P, a_C, \eta, \pi)$, where:

$$U_C(a_P, a_C, \eta, \pi) = \sum_{s \in S} \left[\eta(s)\pi(s)a_C(s) \right] \bar{y}(s) - \frac{\theta}{2n_C} \left\{ \sum_{s \in S} \eta(s)\pi(s)a_C(s)\gamma(s) \right\}^2.$$

H.2 Equilibrium

Just as in the original model, we consider the auxiliary problem of maximizing the payoff of the professional investor:

$$U_P(a_P, a_C, \eta, \pi) + \tau(\pi)$$

with respect to (a_P, a_C, η, π) and τ , subject to the constraint that the fractions (η, π) are the basis of an equilibrium in the subgame following offer (a_C, τ) :

$$\begin{aligned} U_P(a_P, a_C, \eta, \pi) + \tau(\pi) &\geq U_P(a_P, a_C, \tilde{\eta}, \pi) + \tau(\pi) \\ U_C(a_P, a_C, \eta, \pi) - \tau(\pi) &\geq U_C(a_P, a_C, \eta, \tilde{\pi}) - \tau(\tilde{\pi}), \end{aligned}$$

for all alternative $\tilde{\eta}$ and $\tilde{\pi}$. We proceed in two steps. First, we simplify the auxiliary problem and show the analogue of Proposition 1 in the main body of the paper. Second, we solve for asset prices, and provide the analogue of Proposition 2 in the main body of the paper.

Proposition 7 *The equilibrium value of a professional investor coincide with the value of the auxiliary problem, and is given by the maximum of*

$$-\sum_{s \in S} p(s) \left[a_P(s) + a_C(s) \right] + \sum_{s \in S} \left[a_P(s) + a_C(s) \right] \bar{y}(s) - \frac{\theta}{2n_P} \left\{ \sum_{s \in S} a_P(s) \gamma(s) \right\}^2 - \frac{\theta}{2n_C} \left\{ \sum_{s \in S} a_C(s) \gamma(s) \right\}^2$$

with respect to $a_P(s) \geq 0$ and $a_C(s) \geq 0$, and subject to the incentive constraint $k(s)/p(s) \geq a_C(s)$.

With this in mind, we let an equilibrium be a collection of prices $\{p(s)\}$ and asset demands, $\{a_P(s), a_C(s)\}$, such that: the asset demands solve the problem of a professional investor, as given in Proposition 7; and asset markets clear; i.e., $a_P(s) + a_C(s) = A(s)$. As before, there are two cases to consider.

Proposition 8 (Full risk sharing) *Let $p^*(s) = \bar{y}(s) - \theta \gamma(s) \sum_{s' \in S} \gamma(s') A(s')$ denote the CAPM price of asset $s \in S$ and suppose that:*

$$\sum_{s' \in S} \gamma(s') \min \left\{ A(s'), \frac{k(s')}{p^*(s')} \right\} \geq n_C \sum_{s' \in S} \gamma(s') A(s'). \quad (83)$$

Then CAPM holds: $p(s) = p^*(s)$ for all $s \in S$ and there is full risk sharing:

$$\frac{1}{n_I} \sum_{s \in S} \gamma(s) a_I(s) = \frac{1}{n_C} \sum_{s \in S} \gamma(s) a_C(s).$$

The other case is:

Proposition 9 (Partial risk sharing) *Suppose that condition (83) does not hold. Then, in equilibrium, there is partial risk sharing, in that $\frac{1}{n_I} \sum_{s \in S} \gamma(s) a_I(s) > \frac{1}{n_C} \sum_{s \in S} \gamma(s) a_C(s)$. Letting $\underline{p}(s) \equiv \bar{y}(s) - \frac{\theta}{n_I} \gamma(s') \sum_{s \in S} \gamma(s) a_I(s)$ and $\bar{p}(s) \equiv \bar{y}(s) - \frac{\theta}{n_C} \gamma(s') \sum_{s \in S} \gamma(s) a_C(s)$, asset prices are given by:*

1. *Liquid assets: if $k(s)/A(s) \geq \bar{p}(s)$, then $p(s) = \bar{p}(s)$.*
2. *Partially liquid assets: if $k(s)/A(s) \in (\underline{p}(s), \bar{p}(s))$, then $p(s) = k(s)/A(s)$.*
3. *Illiquid assets: if $k(s)/A(s) \leq \underline{p}(s)$, then $p(s) = \underline{p}(s)$.*

H.3 Arbitrage

In the above, we have derived conditions under which identical assets can trade at different prices: an asset that is relatively easier to re-sell (large $k(s)$, small $A(s)$) to a consumer trades at a premium, even if it has the exact same fundamental characteristics (expected dividend and risk). One would think, then, that obvious arbitrage opportunities would arise if we relaxed the short-selling constraint.³⁷ Our results are, however, robust to relaxing the short-selling constraint

To see why, assume that, in order to short an asset in the CM, a professional investor has to borrow this asset from some other professional investor. Crucially, a professional investor who lends an asset cannot re-sell it in the DM; i.e., in order to transfer an asset to an outside investor, one needs the physical stock certificate. The intuition for why this is sufficient to preclude arbitrage is intuitive: essentially, this means that an asset on loan cannot be used to transfer risk to consumer. If an asset is liquid or partially liquid, in that $p(s) > \underline{p}(s)$, then this constraint creates a strictly positive opportunity cost, equal to $p(s) - \underline{p}(s)$. Clearly, in equilibrium, the lending fee of that asset has to be greater than the opportunity cost. Now we argue that it has to be exactly equal to the opportunity cost. Indeed, each asset is in strictly positive supply, which implies that there are more securities held long than short.³⁸ Hence, only a fraction of the supply of asset is lent, which

³⁷Consider two assets (s, s') with identical fundamental characteristics but different prices, say $p(s) > p(s')$. Then, one would think that a professional investor should be able to make arbitrage profits by shorting one share of s , buying one share of s' : he would pocket the price difference, $p(s) - p(s') > 0$ in the CM, and use the dividend of asset s to repay the dividend of asset s' .

³⁸This follows from the market clearing condition: shares held long = exogenous supply + shares held short.

implies in turns that holders of the asset have to be indifferent between lending or not (Graveline and McBrady, 2010).

H.4 Proofs

H.4.1 Proof of Proposition 7

As before we proceed in several steps.

Claim 7 *To solve the auxiliary problem, we can restrict attention to price schedules such that $\tau(\pi) = U_C(a_P, a_C, \eta, \pi)$.*

Indeed, take any feasible (a_P, a_C, η, π) and τ , and consider the alternative price schedule $\hat{\tau}(\tilde{\pi}) \equiv U_C(a_P, a_C, \eta, \tilde{\pi})$ for all $\tilde{\pi}$. Together with (a_P, a_C, η, π) , this price schedule satisfies the two IC constraints. Moreover, evaluating the IC constraint of the consumer at $\tilde{\pi} = 0$, we obtain that $\tau(\pi) \leq U_C(a_P, a_C, \eta, \pi) = \hat{\tau}(\pi)$. Hence, with $\hat{\tau}(\tilde{\pi})$ instead of $\tau(\tilde{\pi})$, the objective of the professional investor's auxiliary problem increases weakly.

Claim 8 *To solve the auxiliary problem, we can restrict attention to η such that $\eta(s) > 0$ for all $s \in S$.*

Consider any feasible (a_P, a_C, η, π) with $\tau(\tilde{\pi}) = U_C(a_P, a_C, \eta, \tilde{\pi})$. Suppose that $\eta(s_0) = 0$ for some $s_0 \in S$. Then consider the alternative choice obtained by setting $\eta(s_0) = 1$ and $a_C(s_0) = 0$, and keeping everything else the same. Clearly, the IC constraint of the consumer is satisfied, since it only depends on $\eta(s)a_C(s)$. The IC constraint of the intermediary continues to hold for $\eta(s) \neq s_0$, and it also holds for s_0 because $a_C(s_0) = 0$.³⁹

Claim 9 *The IC constraint of the professional investor can be written:*

$$\begin{aligned} & -p(s)a_C(s) + k(s) + (1 - \pi(s))a_C(s)\bar{y}(s) \\ & - \frac{\theta}{n_P}(1 - \pi(s))a_C(s)\gamma(s) \sum_{s' \in S} \gamma(s') \left[a_I(s') + (1 - \pi(s'))\eta(s')a_C(s') \right] \geq 0 \end{aligned}$$

for all $s \in S$, with equality if $\eta(s) \in (0, 1)$.

³⁹Since $\eta \mapsto U_P(a_P, a_C, \eta, \pi)$ is concave, it is sufficient to check the IC constraint coordinate-per-coordinate.

Since $\eta \mapsto U_P(a_P, a_C, \eta, \pi)$ is concave, first-order conditions are necessary and sufficient in the IC constraint. The result follows by taking derivative with respect to $\eta(s)$, and keeping in mind that $\eta(s) > 0$.

Claim 10 *To solve the auxiliary problem, we can restrict attention to η such that $\eta(s) = 1$ for all $s \in S$.*

Consider any feasible (a_P, a_C, η, π) with $\tau(\tilde{\pi}) = U_C(a_P, a_C, \eta, \tilde{\pi})$. Suppose that $\eta(s_0) \in (0, 1)$ for some s_0 . Then the corresponding IC constraint holds with equality. In particular, since $k(s_0) > 0$, it follows that $a_C(s_0) > 0$ and

$$p(s_0) - (1 - \pi(s_0))\bar{y}(s_0) + \frac{\theta}{n_P}\gamma(s_0)(1 - \pi(s_0)) \sum_{s' \in S} \left[a_P(s') + (1 - \pi(s'))\eta(s')a_C(s') \right] = 0.$$

Then consider the alternative choice obtained by setting $\eta'(s_0) = 1$, $a'_C(s) = \eta(s_0)a_C(s_0)$, and keeping everything else the same. The right-hand side of the s_0 IC constraint of the professional investor decreases, and so this IC constraint holds. The other IC constraints, only depends on the product $\eta(s)a_C(s)$, which stays the same, and so they are all satisfied. Lastly, the objective of the professional investor increases strictly.

Taken together the above claims imply that, to solve for the value of the auxiliary problem, it is sufficient to maximize

$$\begin{aligned} & - \sum_{s \in S} p(s) \left[a_P(s) + a_C(s) \right] + \sum_{s \in S} \left[a_P(s) + a_C(s) \right] \bar{y}(s) \\ & - \frac{\theta}{2n_P} \left\{ \sum_{s \in S} \left[a_P(s) + (1 - \pi(s))a_C(s) \right] \gamma(s) \right\}^2 - \frac{\theta}{2n_C} \left\{ \sum_{s \in S} \pi(s)a_C(s)\gamma(s) \right\}^2 \end{aligned}$$

with respect to $\pi(s) \in [0, 1]$, $a_I(s) \geq 0$, $a_C(s) \geq 0$ and subject to the IC constraint:

$$k(s) \geq a_C(s) \left\{ p(s) - (1 - \pi(s))\bar{y}(s) + \frac{\theta}{n_P}(1 - \pi(s))\gamma(s) \sum_{s' \in S} \gamma(s') \left[a_P(s') + (1 - \pi(s'))a_C(s') \right] \right\}.$$

Note that a lower bound for the value of the problem is zero since it is feasible to choose $a_P(s) = a_C(s) = \pi(s) = 0$. Also, if both $a_P(s)$ and $a_C(s)$ are large enough, then the objective is strictly negative, for any other choice of $a_P(s')$, $a_C(s')$ and $\pi(s')$. So, to find the value of the problem, we can restrict attention to a compact set of $a_P(s)$ and $a_C(s)$. We are left maximizing

a continuous function on a compact: we thus know the problem has a solution. Now, if at this solution, $\pi(s) = 0$ for some $s \in S$, then the same value is achieved by $\pi'(s) = 1$, $a'_C(s) = 0$, and $a'_P(s) = a_P(s) + a_C(s)$, and keeping everything else the same. So we can always consider solutions such that $\pi(s) > 0$ for all s .

Now, in a neighborhood of such a solution, we make the change of variable

$$\begin{aligned}\tilde{a}_P(s) &\equiv a_P(s) + (1 - \pi(s))a_C(s) \\ \tilde{a}_C(s) &\equiv \pi(s)a_C(s).\end{aligned}$$

The objective becomes:

$$-\sum_{s \in S} p(s) \left[\tilde{a}_I(s) + \tilde{a}_C(s) \right] + \sum_{s \in S} \left[\tilde{a}_I(s) + \tilde{a}_C(s) \right] \bar{y}(s) - \frac{\theta}{2n_P} \left\{ \sum_{s \in S} \tilde{a}_P(s) \gamma(s) \right\}^2 - \frac{\theta}{2n_C} \left\{ \sum_{s \in S} \tilde{a}_C(s) \gamma(s) \right\}^2$$

and the IC constraint becomes:

$$\pi(s)k(s) \geq \tilde{a}_C(s) \left\{ p(s) - (1 - \pi(s))\bar{y}(s) + \frac{\theta}{n_P} (1 - \pi(s)) \gamma(s) \sum_{s' \in S} \gamma(s') \tilde{a}_P(s') \right\}.$$

We now show that:

Claim 11 *To solve the auxiliary problem, we can restrict attention to $\pi(s)$ such that $\pi(s) = 1$ for all $s \in S$.*

Start from a solution of the auxiliary problem and, for all s , replace $\pi(s)$ by the the largest $\pi'(s)$ such that the IC constraint holds. Clearly, this does not change the value of the objective, so this is also a solution of the problem. If this results in $\pi'(s) = 1$ for all s , we are done. Otherwise, if $\pi'(s) < 1$ for some s , then the IC constraint must bind, and furthermore:

$$k(s) < \tilde{a}_C(s) \left[\bar{y}(s) - \frac{\theta}{n_P} \gamma(s) \sum_{s' \in S} \gamma(s') \tilde{a}_P(s') \right].$$

Otherwise it would be possible to increase $\pi(s)$ further without violating the IC constraint. Multiplying both sides by $\pi'(s)$ and replacing the resulting strict inequality into the binding IC constraint, we obtain that:

$$p(s) < \bar{y}(s) - \frac{\theta}{n_P} \gamma(s) \sum_{s' \in S} \gamma(s') \tilde{a}_P(s'). \quad (84)$$

But then for all s such that $\pi'(s) < 1$, we increase $\tilde{a}_P(s)$ by ε , and simultaneously lower $\pi'(s)$ so that all the corresponding IC constraints hold. Note that these increases do not impact the IC constraint for those s such that $\pi'(s) = 1$. Thus, this deviation is feasible. The marginal increase in the objective is:

$$\sum_{s:\pi'(s)<1} \left[-p(s) + \bar{y}(s) - \frac{\theta}{n_P} \gamma(s) \sum_{s' \in S} \gamma(s') \tilde{a}_P(s') \right],$$

which is strictly positive because of (84), a contradiction. Lastly, we have

Claim 12 *In any equilibrium, the value of an intermediary is equal to the value of the auxiliary problem.*

We first note that the value of the auxiliary problem is an upper bound of the value of the professional investor in any equilibrium. Keeping this in mind, let a_P and a_C be solution of the auxiliary problem and consider any candidate equilibrium such that the professional achieves a strictly lower value. The intermediary could offer $\max\{a_C - \varepsilon, 0\}$, $\tau(1) = U_C(a_P, a_C - \varepsilon, 1, 1) - \varepsilon$, and $\tau(\tilde{\pi}) = U_C(a_P, a_C - \varepsilon, 1, \tilde{\pi})$. The unique equilibrium in the corresponding subgame is then $\eta = 1$ and $\pi = 1$. By letting ε go to zero, one sees that the professional investor can guarantee herself a value arbitrarily close to that of the auxiliary problem. In particular, she can achieve a value strictly higher than the one in the candidate equilibrium, a contradiction.

H.4.2 Proof of Propositions 8 and 9

We solve for the unique equilibrium in steps.

Claim 13 *If an asset demand $\{a_P(s), a_C(s)\}_{s \in S}$ solves the professional investor's problem, then*

$$\frac{1}{n_P} \sum_{s \in S} \gamma(s) a_P(s) \geq \frac{1}{n_C} \sum_{s \in S} \gamma(s) a_C(s). \quad (85)$$

Indeed, if the opposite equality were true, it would be possible to increase the intermediary's objective by lowering some $a_C(s)$ and increase some $a_P(s)$.

Claim 14 *In a "full risk sharing" equilibrium, where (85) holds with equality:*

$$\frac{1}{n_P} \sum_{s \in S} \gamma(s) a_P(s) = \sum_{s \in S} \gamma(s) A(s) = \frac{1}{n_C} \sum_{s \in S} \gamma(s) a_C(s).$$

In a “partial risk sharing” equilibrium, where (85) holds with inequality:

$$\frac{1}{n_P} \sum_{s \in S} \gamma(s) a_P(s) > \sum_{s \in S} \gamma(s) A(s) > \frac{1}{n_C} \sum_{s \in S} \gamma(s) a_C(s).$$

This follows directly from combining (85) with the market clearing condition $a_I(s) + a_C(s) = A(s)$. Now, letting $\lambda(s)$ be the multiplier on the IC constraint for asset $s \in S$, the first-order necessary and sufficient conditions are:

$$p(s) \geq \bar{y}(s) - \frac{\theta}{n_P} \gamma(s) \sum_{s' \in S} \gamma(s') a_P(s'), \quad \text{with “=” if } a_P(s) > 0 \quad (86)$$

$$p(s) \geq \bar{y}(s) - \frac{\theta}{n_C} \gamma(s) \sum_{s' \in S} \gamma(s') a_C(s') - \lambda(s), \quad \text{with “=” if } a_C(s) > 0, \quad (87)$$

and $\lambda(s) (k(s)/p(s) - a_C(s)) = 0$, with complementary slackness. Based on this, we have:

Claim 15 *In a full risk-sharing equilibrium, CAPM holds in that*

$$p(s) = p^*(s) = \bar{y}(s) - \theta \gamma(s) \sum_{s' \in S} \gamma(s') A(s').$$

In equilibrium, either $a_P(s) > 0$ or $a_C(s) > 0$. If $a_P(s) > 0$, then the claim follows from (86). If $a_C(s) > 0$, then (87) holds with equality, implying that $p(s) \leq p^*(s)$. But, in a full risk-sharing equilibrium, (86) writes $p(s) \geq p^*(s)$. Taken together, these inequalities imply that $p(s) = p^*(s)$ as claimed.

Claim 16 *There exists a full risk-sharing equilibrium if and only if (83) holds.*

If there exists a full risk-sharing equilibrium, then $p(s) = p^*(s)$. By market clearing, $a_C(s) \leq A(s)$, and by the IC constraint $a_C(s) \leq k(s)/p^*(s)$. Thus, $a_C(s) \leq \min\{A(s), k(s)/p^*(s)\}$. Plugging this into the condition that $\sum_{s \in S} \gamma(s) a_C(s) = n_C \sum_{s \in S} \gamma(s) A(s)$, we obtain inequality (83). Conversely, suppose that (83) holds. Then it is possible to find $\{a_C(s)\}_{s \in S}$ so that $a_C(s) \leq A(s)$, $a_C(s) \leq k(s)/p(s)$, and $\sum_{s \in S} \gamma(s) a_C(s) = n_C \sum_{s \in S} \gamma(s) A(s)$. One then easily verifies that these asset demands are the basis of a full risk sharing equilibrium, together with $p(s) = p^*(s)$ and $a_P(s) = A(s) - a_C(s)$.

Claim 17 *In a partial risk-sharing equilibrium, $a_C(s) > 0$ for all $s \in S$.*

If $a_C(s) = 0$ for some $s \in S$, then it would be possible to increase the intermediary's utility by increasing $a_C(s)$ (this does not violate the IC constraint) and decreasing $a_P(s)$.

For the next claim, note that the FOC implies that the price of asset s , $p(s)$, must lie in between the following two bounds:

$$\underline{p}(s) \equiv \bar{y}(s) - \frac{\theta}{n_I} \gamma(s') \sum_{s \in S} \gamma(s) a_I(s) \quad \text{and} \quad \bar{p}(s) \equiv \bar{y}(s) - \frac{\theta}{n_C} \gamma(s') \sum_{s \in S} \gamma(s) a_C(s)$$

With these definitions in mind, we have:

Claim 18 *In a partial risk-sharing equilibrium:*

1. If $k(s)/A(s) \leq \underline{p}(s)$, then $p(s) = \underline{p}(s)$.
2. If $k(s)/A(s) \geq \bar{p}(s)$, then $p(s) = \bar{p}(s)$.
3. Otherwise, $p(s) = k(s)/A(s)$.

Suppose $k(s)/A(s) \leq \underline{p}(s)$ but $p(s) > \underline{p}(s)$. Then, from (86), $a_P(s) = 0$. Together with market clearing, this implies that $a_C(s) = A(s)$. Plugging this into the IC constraint, we find that $p(s)A(s) \leq k(s) \Leftrightarrow p(s) \leq \underline{p}(s)$, a contradiction.

Suppose $k(s)/A(s) \geq \bar{p}(s)$ but $p(s) < \bar{p}(s)$. Then, from (87), we have that the IC constraint must bind, and $p(s)a_C(s) = k(s)$. Together with the market-clearing condition, this implies that $p(s)A(s) \geq k(s) \Leftrightarrow p(s) \geq \bar{p}(s)$, a contradiction.

Now suppose that $\underline{p}(s) < k(s)/A(s) < \bar{p}(s)$. If $p(s) = \underline{p}(s)$, then $p(s) < k(s)/A(s)$, so the IC constraint does not bind, and $p(s) = \bar{p}(s)$, a contradiction. If, on the other hand, $p(s) = \bar{p}(s)$, then $p(s) \geq k(s)/A(s)$ and so the IC constraint binds if $a_C(s) \geq A(s)$. Therefore, $a_C(s) < A(s)$ and, by market clearing, $a_I(s) > 0$. But this means that $p(s) = \underline{p}(s)$, a contradiction. Since $p(s) > \underline{p}(s)$, we have that $a_I(s) = 0$ and thus, by market clearing, $a_C(s) = A(s)$. Since $p(s) < \bar{p}(s)$, the IC constraint binds, and therefore $a_C(s) = A(s) = k(s)/p(s) \Rightarrow p(s) = k(s)/p(s)$.

Claim 19 *There exists a partial risk sharing equilibrium if and only if (83) does not hold.*

Suppose a partial risk sharing equilibrium exists, and let $x \equiv \frac{1}{n_P} \sum_{s' \in S} \gamma(s') a_P(s')$, so that $\underline{p}(s) = \bar{y}(s) - \gamma(s)\theta x$. Note that, if $p > \underline{p}(s)$, then $a_P(s) = 0$, and $k(s)/A(s) > \underline{p}(s)$. If $p(s) = \underline{p}(s)$,

then the IC constraint binds and so $a_C(s) = k(s)/\underline{p}(s)$, and $a_P(s) = A(s) - k(s)/\underline{p}(s)$. Taken together, this implies that:

$$a_P(s) = \max \left\{ 0, A(s) - \frac{k(s)}{\bar{y}(s) - \gamma(s)\theta x} \right\}.$$

Multiplying by $\gamma(s)$ and adding up, we obtain:

$$n_P x = \sum_{s \in S} \gamma(s) \max \left\{ 0, A(s) - \frac{k(s)}{\bar{y}(s) - \gamma(s)\theta x} \right\}. \quad (88)$$

In a partial risk-sharing equilibrium, this equation must be satisfied for some $x > 1/n_P \sum_{s \in S} \gamma(s)A(s)$. But note that the left-hand side of the above equation is increasing, while the right-hand side is non-increasing. Thus, if the equation is satisfied for some $x > \sum_{s \in S} \gamma(s)A(s)$, then the left-hand side must be less than the right-hand side when evaluated at $1/n_P \sum_{s \in S} \gamma(s)A(s)$. That is:

$$\begin{aligned} n_P \sum_{s \in S} \gamma(s)A(s) &< \sum_{s \in S} \gamma(s) \max \left\{ 0, A(s) - \frac{k(s)}{p^*(s)} \right\} \\ \Leftrightarrow n_P \sum_{s \in S} \gamma(s) \left(A(s) + \frac{k(s)}{p^*(s)} \right) &< \sum_{s \in S} \gamma(s) \max \left\{ A(s), \frac{k(s)}{p^*(s)} \right\} \\ \Leftrightarrow n_P \sum_{s \in S} \gamma(s) \left(A(s) + \frac{k(s)}{p^*(s)} \right) &< \sum_{s \in S} \gamma(s) \left(A(s) + \frac{k(s)}{p^*(s)} - \min \left\{ A(s), \frac{k(s)}{p^*(s)} \right\} \right) \\ \Leftrightarrow \sum_{s \in S} \gamma(s) \min \left\{ A(s), \frac{k(s)}{p^*(s)} \right\} &< n_C \sum_{s \in S} \gamma(s)A(s). \end{aligned}$$

where: the second line follows by adding $\sum_{s \in S} \gamma(s)k(s)/p(s)$ on both sides of the equation; the third line follows from the identity $\max\{a, b\} = a + b - \min\{a, b\}$; the fourth line follows from re-arranging keeping in mind that $n_P + n_C = 1$. This establishes the ‘‘if’’ part of the proposition.

For the ‘‘only if’’, note that if condition (83) does not hold, then the fixed point equation for x , (88), has a solution greater than $\sum_{s \in S} \gamma(s)A(s)$. Then, one easily checks that the following is an equilibrium: $\underline{p}(s) \equiv \bar{y}(s) - \theta\gamma(s)x$, $a_P(s) \equiv \max \left\{ 0, A(s) - \frac{k(s)}{\underline{p}(s)} \right\}$, $a_C(s) \equiv A(s) - a_P(s) = \min \left\{ A(s), \frac{k(s)}{\underline{p}(s)} \right\}$, $\bar{p}(s) \equiv \bar{y}(s) - \frac{\theta}{n_C} \gamma(s) \sum_{s' \in S} \gamma(s')a_C(s')$, and $p(s)$ as given in the claim above.

References

- [1] Ajello, Andrea (2010). “Financial Intermediation, Investment Dynamics and Business Cycle,” manuscript.
- [2] Caballero, Ricardo J. (2006). “On the Macroeconomics of Asset Shortages,” MIT Department of Economics Working Paper No. 06-30.
- [3] Collin-Dufresne, Pierre, Robert S. Goldstein, and J. Spencer Martin (2001). “The Determinants of Credit Spread Changes,” *Journal of Finance* 56, 2177-2207.
- [4] Gabaix, Xavier, Arvind Krishnamurthy, and Olivier Vigneron (2007). “Limits of Arbitrage: Theory and Evidence from the Mortgage-Backed Securities Market,” *Journal of Finance* 62, 557-595.
- [5] Graveline, Jeremy and Matthew R. McBrady (2010). “Who Makes On-the-Run Treasuries Special?” manuscript.
- [6] Holmstrom, Bengt and Jean Tirole (2011). *Inside and outside liquidity*, Princeton, Princeton University Press.
- [7] Kiyotaki, Nobuhiro and John Moore (2001). “Liquidity, Business Cycles and Monetary Policy,” manuscript.
- [8] Kiyotaki, Nobuhiro and John Moore (2005). “Liquidity and asset prices,” *International Economic Review* 46, 317-349.
- [9] Lagos, Ricardo (2010). “Asset prices and liquidity in an exchange economy,” *Journal of Monetary Economics* 57, 913-930.
- [10] Sharpe, William F. (1964). “Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk,” *Journal of Finance* 19, 425-442.
- [11] Shi, Shouyong (1997). “A Divisible Search Model of Fiat Money,” *Econometrica* 65, 75-102.
- [12] Shi, Shouyong (2011). “Liquidity, Assets and Business Cycles,” manuscript.

[13] Shimer, Robert (2010). “Wage Rigidities and Jobless Recoveries,” manuscript.

[14] Shleifer, Andrei (1986). “Do Demand Curves for Stocks Slope Down?,” *Journal of Finance* 41, 579-90.