

The Paradox of Financial Fire Sales: the Role of Arbitrage Capital

JAMES DOW and JUNGSUK HAN*

Journal of Finance forthcoming

ABSTRACT

How can fire sales for financial assets happen when the economy contains well capitalized, but non-specialist investors? Our explanation combines rational expectations equilibrium and “lemons” models. When specialist (informed) market participants are liquidity-constrained, prices become less informative. This creates an adverse selection problem, decreasing the supply of high-quality assets, and lowering valuations by non-specialist (uninformed) investors, who become unwilling to supply capital to support the price. In normal times, arbitrage capital can “multiply” itself by making uninformed capital function as informed capital, but in a crisis this stabilizing mechanism fails.

In a fire sale, sellers are forced to sell assets at deep discounts because no one is willing to buy them at fair prices. Sellers can be forced to sell because of financial distress, credit

*Dow is from London Business School and Han is from Stockholm School of Economics. We thank three anonymous referees, an Associate Editor, and Bruno Biais (the Editor) for very helpful comments. We also thank Brandon Daley, Alex Edmans, Marco Di Maggio, Julian Franks, Simon Gervais, Vincent Glode, John Kuong, Pete Kyle, Paolo Sodini, Per Strömberg, and Jörgen Weibull, as well as seminar participants at AFA 2016, Brevan Howard/FTG conference 2016, Cambridge Corporate Finance Theory Symposium 2015, FIRS 2016, Jackson Hall finance conference 2016, BI Norwegian Business School, INSEAD, Korea University, London Business School, SKKU, SKK GSB, Stockholm School of Economics, University of Gothenburg, University of Mannheim, and Yonsei University. Dow declares that he has not received any financial support for the research described in this paper, and he also has no relevant or material financial interests that relate to the research described in this paper. Han declares that he has not received any financial support for the research described in this paper, and he also has no relevant or material financial interests that relate to the research described in this paper.

market frictions, regulation, margin calls, etc.¹ Why do fire sales happen? What makes investors avoid buying assets that are apparently cheap? One explanation is given by Shleifer and Vishny (1992): if industry experts with higher private valuations do not have enough liquidity, assets are bought at a discount by non-experts who cannot use them efficiently. This argument applies naturally to real assets rather than financial securities. But since financial securities typically require the holder only to collect cash flows, not to operate the assets, there should not be significant differences in private valuations.²

Allen and Gale (1994), Shleifer and Vishny (1997), Gromb and Vayanos (2002), Brunnermeier and Pedersen (2009), and Duffie (2010) give an alternative explanation based on limited arbitrage capital. They argue that complex securities are normally held by specialized investors who can understand them properly, so that if those investors become severely capital constrained the available pool of specialized capital is too small to pay full value for the outstanding stock of assets. It's true that if all specialized investors are capital constrained, then they are not able to buy undervalued assets. But what about non-specialized investors? Many episodes described in the literature as fire sales (such as the LTCM crisis and the collapse of the mortgage-backed securities (MBS) markets during the financial crisis of 2007-2008) occurred even though there were plenty of well-capitalized investors somewhere in the world economy (e.g., Warren Buffett, or sovereign wealth funds). Why wouldn't well-capitalized outside investors want to step in and buy undervalued assets, thereby preventing large price drops? This seems paradoxical.

A possible resolution of the paradox is based on "lemons" problems as described by Akerlof (1970). Non-specialized investors could be exposed to adverse selection perpetrated by industry insiders. This could happen if insiders choose to hold back high-quality assets, and sell only low-quality assets, causing a classic lemons problem in which anybody who is forced to sell high-quality assets would suffer a loss. But this explanation has a major flaw because it does not explain why the lemons problem would suddenly increase during a crisis. In a crisis, asset sellers should have *less* discretion on selling, thus, liquidity shortages should even mitigate the lemons problem (e.g., Malherbe (2014)). It

¹See, for example, Shleifer and Vishny (2011) for a survey.

²Some differences in valuation for financial securities do exist. For example, some holders are able to repo the security and others are not. Also, expert investors may understand the risk of a security better, allowing them to hedge their overall portfolio more effectively. However, these differences in private valuations should be minor.

seems that the paradox remains.

We give an intuitive answer to this question by examining the role of informed trading in preventing adverse selection. We argue that informed traders who can *buy* assets help to make market prices informative. This prevents high-quality assets from trading at similar prices to low-quality assets. Hence, informed buyers remove the incentive for sellers to sell low-quality assets. But in case of a large shock to the market, this mechanism breaks down. A severe need for liquidity affecting specialized traders prevents them from using their private information to bid up undervalued assets. This can make prices uninformative, leading to adverse selection in which sellers predominantly supply overvalued assets to the market. This in turn leads uninformed agents, who are potential buyers for those assets, to withdraw from the market even though they are not wealth-constrained. The result is a market freeze for high-quality assets: they are not traded, except when their owners are subject to a severe liquidity shortage and therefore forced to sell at fire sale prices.

To formalize this argument, we develop an information-based theory of fire sales using a noisy rational expectations equilibrium (REE) framework with endogenous adverse selection. We aim to achieve the following goals in our paper. First, we explain the role of informed trading in fire sales and market freezes. Unlike the traditional literature on adverse selection that only features informed sellers, we highlight the role of informed traders who compete to exploit mispricing, and thereby make the price more informative. Second, we answer the question of why fire sales occur even when, somewhere in the economy, there are traders with enough capital to correct prices. In particular, we show that a liquidity shortage for informed traders can deter the uninformed, well capitalized traders from buying, thereby allowing asset prices to fall. This also sheds light on the paradoxical nature of fire sales in which capital moves out of the market when it is needed most and would seemingly earn higher returns. Third, we explain why a market freeze happens at the same time as fire sales. That is, our paper explains the “double whammy” situation where fire sales and low trading volume occur together (see, for example, Tirole (2011) for a discussion of the financial crisis of 2007-2008).

Consider a two-period model with some participants who are informed, but financially constrained, and some who are unconstrained, but uninformed. There exists a marketable asset with risky payoffs whose value is only known to the informed partici-

pants.³ There are two types of informed participants: a distressed seller who needs to raise liquidity, and arbitrageurs who aim to make trading profits. The seller is forced to meet liquidity needs, which he can do either by selling his holdings of the marketable asset, or alternatively by liquidating another asset which is non-marketable (such as profit-generating operations).⁴ This choice leads to the lemons problem; the seller only sells the marketable asset when it is overvalued (a lemon), or when liquidating the non-marketable asset is very costly. There are two types of uninformed participants: hedgers and risk-neutral investors. Hedgers trade to hedge against their risk exposure which is correlated with the payoff of the marketable asset. The randomness in hedgers' risk exposure creates trading noise which prevents full revelation of informed participants' private information.

In normal market conditions where the arbitrageurs have enough liquidity, the price reveals the fundamental value of the asset because informed trading volume overwhelms the impact of noise. This allows the seller in liquidity shortage to fund itself by selling the marketable asset at intrinsic value regardless of the quality of the asset. This in turn makes uninformed investors willing to absorb the supply of asset without worrying about adverse selection. On the other hand, in a crisis situation where the arbitrageurs are liquidity-constrained, prices are less informative. This makes the seller willing to supply only low-quality assets to the market, except in cases where it is effectively forced to sell because the alternative of liquidating the non-marketable asset is very costly. This adverse selection problem in turn makes uninformed investors unwilling to absorb the supply of assets unless there is a drop in price to reflect a lemons discount. It also creates a market freeze in the sense that the supply of high-quality assets decreases. In summary, arbitrage capital can “multiply” itself by making uninformed capital effectively function as informed capital in normal market, but this stabilizing mechanism can be undone during a crisis.

An initial liquidity reduction to the arbitrageurs (modeled as a comparative statics analysis where the amount of arbitrage capital is varied) decreases their capacity to trade, so prices become less informative. This discourages the seller from selling high-quality assets. This lemons problem causes prices to fall on average, further reinforcing the seller's reluctance to sell high-quality assets.

³In practice they are likely to be financial intermediaries such as banks and hedge funds.

⁴The non-marketable asset would produce low levels of output if it is liquidated early. Similar assumptions are standard in the literature, for example, the assumption of illiquid assets held by banks in Diamond and Dybvig (1983).

We also show that economic welfare is maximized when prices are close to fundamental value. Intuitively, welfare depends on both expected total wealth and risk sharing. On the expected total wealth side, mispricing causes inefficient liquidation of otherwise valuable assets: in case traded assets are underpriced, the seller raises funds to meet liquidity shortages by inefficiently liquidating non-marketable assets, rather than selling traded assets. On the risk sharing side, market freezes accompanied by fire sales reduce the hedgers' expected utility by preventing them from hedging away their risk exposure. Consequently, shortage in arbitrage capital can reduce economic welfare through fire sales and market freezes.

It is worth noting that, in our model, the majority of the supply of the asset can be absorbed by the uninformed investors rather than the informed arbitrageurs. Even a small amount of informed capital is able to promote liquidity in the market as long as it can make prices sufficiently informative. Informed capital facilitates the movement of uninformed capital from markets with excess liquidity to those with low liquidity. This potency of arbitrage capital, while apparently attractive, is however a double-edged sword. A small reduction in informed capital can trigger fire sales and market freezes by driving away the demand of uninformed investors. This implies that this “multiplier” effect of arbitrage capital can actually serve as the source of financial *instability* rather than financial stability.

Our model provides a useful tool for understanding fire-sale episodes, such as the financial crisis of 2007-2008 during which many financial institutions were forced to unwind their positions or to reduce leverage (See, for example, Brunnermeier (2009) and Shleifer and Vishny (2011)). Our theory suggests how reductions in informed trading can act as a transmission mechanism of liquidity shocks to fire sales in unrelated asset classes such as the repo market (e.g., Gorton and Metrick (2012)) and residential mortgage-backed securities (RMBS) (e.g., Merrill, Nadauld, Stulz, and Sherlund (2014)); it also sheds light on market freezes such as the collapse of the non-agency RMBS market (e.g., Vickery and Wright (2013)).

The organization of the paper is as follows. Section I relates our paper to the literature. Section II describes the basic model. Section III solves for the equilibrium. Section IV describes our main results about fire sales and market freezes. Section V extend our baseline model by endogenizing arbitrageurs' margin constraints. Section VI studies implications of our paper about financial stability (including policy implication),

and the financial crisis of 2007-2008. Section VII concludes.

I. Literature Review

There is a large literature on both the theory and the empirics of fire sales.⁵ On the empirical side, there is evidence of fire sales across various classes of assets and securities: (i) real assets (e.g., Pulvino (1998), Schlingemann, Stulz, and Walkling (2002)), (ii) equities (e.g., Coval and Stafford (2007), Jotikasthira, Lundblad, and Ramadorai (2012)), (iii) bonds (e.g., Ellul, Jotikasthira, and Lundblad (2011), Jotikasthira, Lundblad, and Ramadorai (2012)), (iv) structured products (e.g., Merrill, Nadauld, Stulz, and Sherlund (2014)), and (v) repos (e.g., Duarte and Eisenbach (2014)).

As discussed in the introduction, the previous theoretical literature suggests that fire sales occur because of liquidity shocks to industry experts (e.g., Shleifer and Vishny (1992)), and limits to arbitrage (e.g., Shleifer and Vishny (1997), Gromb and Vayanos (2002), Brunnermeier and Pedersen (2009)). Several papers also focus on the amplification mechanism of fire sales as in Gromb and Vayanos (2002), Geanakoplos (2003), Brunnermeier and Pedersen (2009), Krishnamurthy (2010), Greenwood, Landier, and Thesmar (2015), and Kuong (2015). For example, Krishnamurthy (2010) shows that fire sales can occur due to feedback effects between asset prices and balance sheets. Counterparty risk can also contribute to the amplification mechanism of fire sales (e.g., Krishnamurthy (2010), Caballero and Simsek (2013)). For example, Caballero and Simsek (2013) study a model where fire sales occur because of counterparty risk in a complex network. We agree with the main thrust of this literature that financial assets are typically held by specialized investors who help to set prices. But the explanations given in the literature do not detail the economic mechanisms that prevent non-specialized investors from trading the assets. For example, He and Krishnamurthy (2013) assume they cannot participate, while in the learning model given in Duffie (2010) they first need to become specialized, which takes time, before participating. Thus, we complement this literature by explaining why and when nonspecialized investors will stay away.

In his seminal paper, Akerlof (1970) shows that a market collapse (or freeze) can happen due to adverse selection. If the quality of assets is only known to sellers, assets of different quality trade at the same price. This makes owners of high quality assets

⁵Shleifer and Vishny (2011) and Tirole (2011) provide surveys on the literature.

withdraw from the market, and subsequently lowers buyers' expectation about the quality, and this in turn makes more sellers withdraw from the market if they own assets with the highest quality among the remaining sellers. This process can continue until there is no seller left, thus the market collapses. This intuition has been extended and applied in the finance literature. In particular, several papers have emphasized the role of fire sales where sellers are forced to sell due to distress. However, buyers cannot tell whether the supply is coming from liquidity-driven sales or information-driven sales, if other sellers try to sell when they get bad signals. This type of adverse selection can be the source of a market freeze.⁶ For example, Eisfeldt (2004) shows that market illiquidity endogenously arises due to lemons problem in a dynamic consumption economy where agents trade for both informational reasons and liquidity needs.⁷ Bolton, Santos, and Scheinkman (2009) show that distressed sellers may choose to sell earlier at fire sale prices to avoid potential adverse selection problems in the future. Dang, Gorton, and Holmstrom (2012) suggest that a market freeze can occur in the debt market by extending the argument of Myers and Majluf (1984). Although debt securities are information-insensitive relative to equity, they become information sensitive when approaching a default state. This causes increased information acquisition, but such information asymmetries may lead to adverse selection and a market freeze. Malherbe (2014) focuses on the self-fulfilling nature of liquidity dry-ups. If sellers are forced to sell their assets due to liquidity needs, and do not choose to sell for informational reasons, adverse selection problems do not arise. Because prices will not reflect adverse selection, sellers do not need to hoard liquidity. On the other hand, if asset sales are driven by information, prices will reflect adverse selection. Anticipating this, sellers will need to hoard liquidity. This creates multiple equilibria.

Our theory is in stark contrast with existing theories of fire sales and market freezes. Akerlof (1970) shows how a market freeze can happen in the presence of information

⁶Market illiquidity can also arise due to adverse selection in different contexts than distress: Daley and Green (2012) show that information release can create a market freeze by creating delayed sales of high-quality assets. Glode, Green, and Lowery (2012) show that financial intermediaries tend to overinvest in expertise, thereby creating more information asymmetries which contribute to illiquidity through adverse selection. Guerrieri and Shimer (2014) show that high-quality assets may be traded at low frequencies to signal their types. Fishman and Parker (2015) show that buying only good assets by early stage investors through more information acquisition leaves more bad assets to buy by later stage, uninformed investors, thereby lowering asset prices on average. This price externality leads to strategic complementarity creating multiple equilibria.

⁷Kurlat (2013) also shows that information asymmetries can amplify negative shocks in a dynamic consumption economy through the channel of rising interest rates that in turn lower asset prices and create further adverse selection problems.

asymmetries. But, there is no informed trader who can correct market prices in that argument. On the other hand, Grossman and Stiglitz (1980) show that the existence of informed traders can improve the informativeness of prices, but the supply of assets is inelastic to market prices. In our paper, Akerlof (1970) meets Grossman and Stiglitz (1980), and this gives a mechanism that creates fire sales. This idea is illustrated in Figure 1.

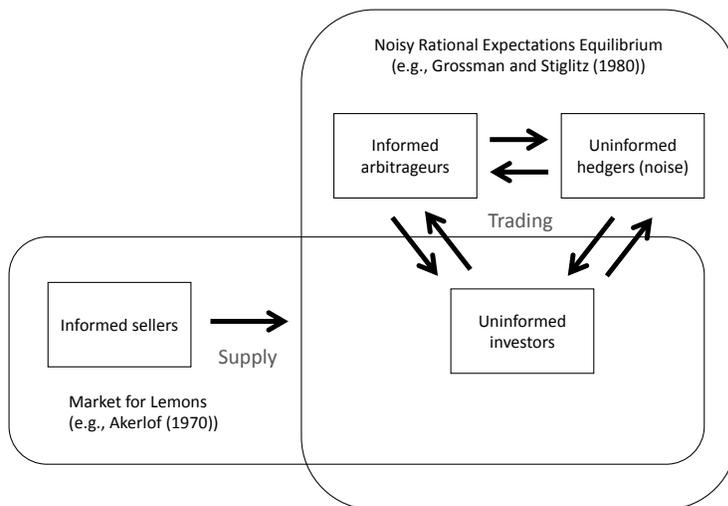


Figure 1. An illustration of our model. Akerlof (1970) meets Grossman and Stiglitz (1980).

Our theory is different from Shleifer and Vishny (1992) because it does not require private valuations. Our theory differs from Shleifer and Vishny (1997), Gromb and Vayanos (2002), and Brunnermeier and Pedersen (2009) because it does not rely on exogenous exclusion of deep-pocketed investors (also, our model generates adverse selection). Our theory differs from Malherbe (2014) because a market freeze (or fire sale) occurs when informed traders are liquidity constrained. Specifically, Malherbe (2014) argues that fire sales occur because of increased adverse selection. However, for this argument to hold, traders should be only moderately distressed (i.e., more funding liquidity for sellers) during fire sales so that they can choose to sell lemons instead of good assets. Yet fire sales typically occur when traders are highly distressed.

Recent papers have extended the limits-to-arbitrage argument, and suggest that “Slow-moving capital” could be a reason why fire sales still occur even when there is enough capital in the economy (e.g., Mitchell, Pedersen, and Pulvino (2007), Duffie

(2010)). However, it is hard to see why uninformed investors would be slow to buy undervalued assets.⁸ What stops capital flowing to correct mispricing? Our paper contributes to the discussion of slow-moving capital by suggesting that the presence of sufficient informed capital can facilitate movement of uninformed capital. However, the market can quickly become illiquid if informed capital providers are liquidity-constrained, as the resulting lemons problem prevents entry of uninformed capital.

Our paper is also related to the literature on financial markets with intermediaries. The usual setup is that only intermediaries participate in asset markets, thus, they trade assets on behalf of consumers or outside investors (e.g., Shleifer and Vishny (1997), Gromb and Vayanos (2002), Allen and Gale (2004), Brunnermeier and Pedersen (2009), He and Krishnamurthy (2011), He and Krishnamurthy (2013), Vayanos and Woolley (2013), Dow and Han (2015)). The difference in the objectives of intermediaries and delegating investors, which is often endogenously determined in the model, is the key to generate mispricing or shock amplification in asset prices. For example, He and Krishnamurthy (2011) and He and Krishnamurthy (2013) study dynamic general equilibrium economies in which households, who do not have access to a long-lived risky asset, invest in the equity of capital-constrained financial intermediaries (or specialists), who have access to the risky asset. They show that shocks are amplified through the channel of financial intermediation; as intermediary capital shrinks up to the point where their capital constraint binds, risk premia of the risky asset rise. More generally in this line of literature (e.g., Shleifer and Vishny (1997), Gromb and Vayanos (2002), Brunnermeier and Pedersen (2009)), a negative shock to intermediary capital restricts investment by outside investors and hence lowers asset prices. Our paper differs from this line of literature in three main respects. First, other models in this line of literature assume that the marginal buyers are informed participants, while we allow the marginal buyer to be determined endogenously in equilibrium—in contrast, it turns out to be the uninformed participants. Second, we focus on how the interaction between informed capital and uninformed capital determines liquidity in financial markets. We study why and how uninformed traders endogenously inject or withdraw capital in response to reductions in informed capital. This allows us to analyze the multiplier effect of informed capital in financial markets. Third, the main mechanism of our model involves an endogenous lemons problem resulting from the degree of informed trading activity. This creates a

⁸We accept that there are a small number of financial markets which are prohibited to unregulated outside investors. For examples, only registered insurers can provide insurance.

market freeze, which is typically absent in this line of literature, even though investors have enough capital to absorb the asset supply in the market.

II. Model

A. Basic Setup

Consider a two-period economy ($t = 1, 2$) with a risk-free asset in infinitely elastic supply with an exogenously-given return r_f .⁹¹⁰ There is also a risky asset that is tradable by all the participants in the financial market (henceforth, “the marketable asset”). The marketable asset pays a random liquidation value of $v \equiv \theta + \epsilon$ at $t = 2$. The first component θ represents the “quality” of the asset which is given by

$$\theta = \begin{cases} \theta_H & \text{with probability } \rho; \\ \theta_L & \text{with probability } 1 - \rho; \end{cases}$$

where $\theta_H > \theta_L > 0$ and $0 < \rho < 1$. We call the marketable asset a “high-quality” asset in the event $\theta = \theta_H$, and “low-quality” otherwise. The second component ϵ is a residual uncertainty which is either positive ($\bar{\epsilon}$) or negative ($-\bar{\epsilon}$) with equal probabilities.¹¹ The two components θ and ϵ are independent of each other, and the realization of θ is observable to traders with expertise whereas that of ϵ is unobservable to anyone in the economy. Examples of the marketable asset include most tradable financial securities such as equities, corporate bonds, MBS, or CDO. For example, in the case of a fixed income security such as corporate bond, one can interpret $\theta_H + \bar{\epsilon}$ as the promised payoff (face value plus coupon) of the security, and other realizations as possible recovery values.

There are four classes of participants: (i) a seller, (ii) arbitrageurs, (iii) hedgers, and (iv) investors. There is a unit mass of arbitrageurs, hedgers, and investors in the economy. We denote \mathcal{A} , \mathcal{H} , and \mathcal{I} to be the set of arbitrageurs, hedgers, and investors, respectively. The seller and the arbitrageurs are capital-constrained, and informed about

⁹See Appendix A for a list of symbols used in this paper.

¹⁰The exogenous risk-free rate is a simplifying assumption, which is a standard assumption in noisy rational expectations model. Endogenizing the risk-free rate by incorporating intertemporal consumption would amplify shocks in the presence of adverse selection (e.g., Kurlat (2013)). Therefore, it could possibly create an even greater magnitude of fire sales due to amplified shocks.

¹¹The assumption that ϵ follows a Bernoulli distribution is for simplicity of exposition. It can generally be replaced with any random variable with mean zero with a finite support.

the quality of the marketable asset. On the other hand, the hedgers and the investors are unconstrained and uninformed.

The seller participates to sell his endowment of the marketable asset to meet liquidity needs. One can consider the seller to be a representative bank in the economy which makes loans and sells securities made out of them.¹² The seller is endowed with \bar{x} unit of the marketable asset. Besides the marketable asset, he is also endowed with a “non-marketable asset” which cannot be traded or transferred to other participants. The non-marketable asset can be considered as the profit-generating operations of the firm.¹³ The non-marketable asset generates a return y per unit of investment with probability density function $f_y(\cdot)$ and cumulative distribution function $F_y(\cdot)$ with support $[r_f, \infty)$. The seller’s liquidity shortage is given by l , and it can be met by either selling the marketable asset or liquidating the non-marketable asset. For simplicity, we further assume that the seller decides to sell either none or the entire holdings of the marketable asset. Any remaining proceeds from selling the marketable asset can be reinvested in the non-marketable asset. The seller can observe the realization of θ and y .

The arbitrageurs participate to generate profits by exploiting any mispricing by trading the asset. They are informed about the value of the marketable asset. Each arbitrageur $a \in \mathcal{A}$ has liquidity position w_A which is strictly positive. One can interpret w_A as the arbitrageurs’ cash position and available credit (e.g., cash from pledging their inventory of other assets). As in the case of the sellers, the arbitrageurs observe the realization of θ . We assume that the arbitrageurs are subject to margin requirements. In a similar fashion to Brunnermeier and Pedersen (2009), we assume that arbitrageur a ’s total margin on his position x^a cannot exceed his available capital w_A :

$$|m^+ x^{a,+}| + |m^- x^{a,-}| \leq w_A,$$

where m^+ and m^- is the dollar margin on the long and short position, and $x^{a,+}$ and

¹²Notice that there is only a single seller. The reason is as follows: In Akerlof’s lemons market for used cars, each car could be of different quality, and hence in principle is a different good with its own market and a single seller (i.e., the existing owner). Because there is no private information on the buy-side of the market, all cars will trade at the same price, so it is possible to treat these different markets as a single representative market. With private information among buyers, this is not possible. Hence, in our model, we have a single market with a single seller.

¹³The non-marketability in this context is equivalent to assuming that cash flow from operations cannot be also pledged to the full extent due to certain frictions such as moral hazard and information asymmetries. Higher realized returns of the non-marketable asset represent the states of the world where the seller is forced to raise liquidity by selling even a high-quality asset even at cheaper prices.

$x^{a,+}$ and $x^{a,-}$ are the positive and negative part of arbitrageur a 's position in units of the asset (i.e., $x^a = x^{a,+} + x^{a,-}$). The dollar margins m^+, m^- are exogenously-given parameters, but are endogenized later in Section V.

The hedgers participate to hedge their non-tradable risk exposure that is correlated with the payoff of the marketable asset. Such hedging needs may arise from a common risk factor that affects both the future wealth of a hedger (such as their labor income) and the payoff of the marketable asset.¹⁴ We assume that η portion of hedgers are affected by a wealth shock (thus, $1 - \eta$ portion of hedgers are unaffected by it) where η is a random variable that follows a probability distribution with probability density function $f_\eta(\cdot)$ with support on $[0, 1]$. Each hedger h observes his own individual wealth shock of vz^h where $z^h = -\bar{z} < 0$ if he is affected, and $z^h = 0$ otherwise. The hedgers are unconstrained and have an identical constant absolute risk aversion (CARA) utility with a risk aversion parameter γ .

The investors participate to transfer their liquidity to the future, and also potentially to make trading profits. They are risk-neutral, and unconstrained in their capacity to invest in available investment opportunities. The arbitrageurs and the hedgers submit demand schedule conditioning on prices. The investors observe the aggregate order flows and set the price that clears the market. This is equivalent to assuming that there exists an auctioneer who finds market clearing prices given demand schedules of all market participants including the investors.¹⁵ The hedgers and the investors are uninformed about the realization of θ , but in equilibrium they will learn about it from the price of the marketable asset.

Intuitively, the seller can be considered to be a firm that invests in projects and creates financial securities. A bank that originates and distributes loans is an example. If the bank writes down some assets in its balance sheet, capital requirements force it to either shrink its balance sheet (by selling assets) or raise capital. The arbitrageurs are firms that specialize in securities trading and have particular expertise in trading the marketable asset but have limited capital. In practice such investors are typically firms rather than individuals. Examples include investment banks that specialize in market-making, hedge funds that specialize in investing in the marketable asset and similar

¹⁴We assume that the wealth shock is perfectly negatively correlated with the payoff of the marketable asset for simplicity. The results are qualitatively unaffected by the assumption of a partial and positive correlation.

¹⁵See, for example, Vives (1995) for a further discussion on this equivalence.

assets, etc. Such arbitrageurs can be subject to liquidity constraints in the form of regulatory requirements, margin calls, client capital withdrawals, etc. Elsewhere in the economy, capital is not limited but the investors deploying this capital have no expertise in the tradable asset. Such investors could be pension funds, insurance companies, financial institutions in other countries, sovereign wealth funds, wealthy investors like Warren Buffett, etc—in other words, any investors who do not have particular expertise in valuing and trading the marketable asset. A more conceptual interpretation is that the investors and the hedgers are the “representative consumer” of the economy who, although he may choose to hold assets via intermediaries, has the option of investing directly.

The market opens if the seller decides to sell his holdings of the marketable asset. If there is no supply of the marketable asset from the seller, the market does not open. Once the market is open, arbitrageurs and investors then condition their demands on the price. The market clears by equating the supply of the seller to the demand from the other traders (arbitrageurs, hedgers, and investors). We assume that the hedgers cannot open the market by themselves unlike the seller because they are infinitesimal.

B. Optimization Problems

The seller maximizes future expected profit by deciding whether to sell his endowment of marketable asset or liquidate the non-marketable asset. We denote x^s to be the units of marketable asset sold by the seller. Then, the seller solves an optimal trading problem given θ and y :

$$\max_{x^s \in \{0, \bar{x}\}} E[v(\bar{x} - x^s) + (1 + y)(px^s - l) | \theta, y], \quad (1)$$

where l is the size of liquidity shortage.¹⁶

Each arbitrageur $a \in \mathcal{A}$ maximizes expected trading profit by choosing a portfolio of the marketable asset and the risk-free asset under the liquidity constraint. Arbitrageur a solves the following constrained optimization problem given θ and p :

$$\max_{x^a} E[vx^a + (1 + r_f)(w_A - px^a) | \theta, p],$$

¹⁶Notice that the seller’s objective function is independent of l because the non-marketable asset has constant returns to scale.

subject to

$$|m^+ x^{a,+}| + |m^- x^{a,-}| \leq w_A, \quad (2)$$

where equation (2) is the margin constraint.

Each hedger $h \in \mathcal{H}$ maximizes expected utility of future wealth by choosing a portfolio that consists of the marketable asset and the risk-free asset. Investor i solves the following optimization problem given the price p and the wealth shock z^h :

$$\max_{x^h} E \left[-\frac{1}{\gamma} \exp \left(-\gamma [w_H(1+r_f) + (v-p(1+r_f))x^h + vz^h] \right) \middle| p, z^h \right],$$

where $z^h = -\bar{z}$ if hedger h receives the wealth shock, and $z^h = 0$ otherwise.

III. Equilibrium

A. Definition of Equilibrium

Equilibrium is defined in a standard manner for noisy rational expectations equilibrium models with a mass of risk-neutral uninformed investors (e.g., Vives (1995), Dow and Rahi (2000), Dow and Rahi (2003)): competitive market participants (arbitrageurs and hedgers) submit orders conditioning on prices; risk-neutral, competitive investors set prices to clear the market conditioning on the net supply of the asset (they are analogous to market makers in Kyle (1985) model).

DEFINITION 1: *An equilibrium is a price p , the supply of the marketable asset x^s , and trades $((x^a)_{a \in \mathcal{A}}, (x^h)_{h \in \mathcal{H}})$ such that (i) x^s solves the seller's problem, (ii) x^a solves each arbitrageur a 's problem, (iii) x^h solves each hedger h 's problem, (iv) the net supply to the investors is given by*

$$X \equiv x^s - \int_{h \in \mathcal{H}} x^h dh - \int_{a \in \mathcal{A}} x^a da,$$

and (v) the price satisfies

$$p = \frac{1}{1+r_f} E[v|p, X]. \quad (3)$$

We define a market freeze to be a situation in which the market fails to open with a positive probability because the seller may sometimes prefer not to sell the marketable asset. Define

$$\begin{aligned}\mu_H &\equiv Pr\left(x^s > 0 \mid \theta_H\right); \\ \mu_L &\equiv Pr\left(x^s > 0 \mid \theta_L\right).\end{aligned}$$

Because the low-quality asset is circulated in the market with probability one, μ_H can be also interpreted as the relative circulation rate of the high-quality asset to that of the low-quality asset.

DEFINITION 2: There is a market freeze if the high-quality asset fails to fully circulate in the market, that is, $\mu_H < 1$.

We define a fire sale to be an event in which the seller sells his holdings at a price so that he would never want to sell unless he is forced to sell for non-informational reasons (such as liquidity shortage).

DEFINITION 3: The seller engages in a fire sale if he sells his holdings of the high-quality asset at a discount rate greater than the risk-free rate, that is, $x^s > 0$ when $E[p \mid \theta_H] < \frac{E[v \mid \theta_H]}{1+r_f}$.

The intuition behind these definitions is as follows. Since sellers are informed, one could conjecture that they will normally sell overvalued assets and not sell undervalued assets. Hence, a seller who sells an undervalued asset must be trying to raise cash, not selling for informational reasons. We call this a fire sale. When prices are lower than $\frac{E[v \mid \theta_H]}{1+r_f}$, the seller would not want to sell unless the realization of y is high enough that he wants to sell the apparently-undervalued high-quality asset, so as not to give up more profitable non-marketable asset.

Turning to the volume of sales, it may seem natural in this model for the seller to always sell his holdings because the returns on the non-marketable asset dominate the returns required by the investors (i.e., y is greater than or equal to r_f). So, if the seller is keeping back his holdings from the market, this suggests that there is some kind of malfunction in the market. As we will show, it is impossible for the marketable asset to trade in equilibrium at less than the present value of the low-quality asset. This implies that the seller will always sell the low-quality asset. But if the price is not revealing, the

seller may hold back the high-quality asset for classic lemons motives. This is what we describe as a market freeze.

B. Solving for Equilibrium

B.1. Demand and Supply

We first describe the seller's supply of the marketable asset given the value of the marketable asset and the return on the non-marketable asset. The following is immediate from the seller's problem (1):

LEMMA 1: *Given θ and y , the seller's supply is as follows:*¹⁷

$$x^s \in \begin{cases} 0 & \text{if } E[p|\theta] < \frac{E[v|\theta]}{1+y}; \\ \bar{x} & \text{if } E[p|\theta] \geq \frac{E[v|\theta]}{1+y}. \end{cases}$$

The seller sells his holdings of the marketable asset if the expected price exceeds the value of the asset, discounted at the opportunity cost of capital. (This is the seller's return on the non-marketable asset, since we have assumed that it is higher than or equal to the risk-free rate.) In other words, when the non-marketable asset has a high return, the seller will sell his endowment of the marketable asset to meet liquidity needs. When the non-marketable asset has a low return, the seller will instead meet his liquidity needs by liquidating the non-marketable asset.

Once the market opens the arbitrageurs, hedgers, and investors trade with the seller. Equation (3) implies that a high-quality asset are never undervalued whereas a low-quality asset are never undervalued. Therefore, arbitrageurs are always willing to buy a high-quality asset, and sell a low-quality asset:

¹⁷We assume a tie-breaking rule that the seller chooses to sell in case the seller is indifferent between choices. This assumption is purely to simplify exposition, and does not affect the result. The other possible tie-breaking rule (i.e., not selling his holdings when indifferent) results in an identical outcome because it is a measure-zero event.

LEMMA 2: *The arbitrageurs' aggregate order flow, $X_A \equiv \int_{a \in \mathcal{A}} x^a da$, is as follows:*¹⁸

$$X_A \in \begin{cases} \frac{w_A}{m^+} & \text{if } \theta = \theta_H; \\ -\frac{w_A}{m^-} & \text{if } \theta = \theta_L. \end{cases}$$

For notational convenience, we denote $X_A(\theta)$ to be arbitrageurs' aggregate order flows given θ because it effectively depends only on θ .

Each hedger observes his own wealth shock, then participate in trading at the market. We get the following aggregate order flow of the hedgers:

LEMMA 3: *The hedgers' aggregate order flow, $X_H \equiv \int_{h \in \mathcal{H}} x^h dh$, is given by*

$$X_H = X^{CARA}(p) + \eta \bar{z},$$

where $X^{CARA}(p)$ is a function of p which is independent of η .

Proof: See Appendix B.

Q.E.D.

The first component is a standard CARA demand function given p , and the second component is the aggregate hedging needs of those hedgers who are affected by the wealth shock negatively correlated with the payoff of the marketable asset. Each hedger attempts to infer θ from the market clearing price p , thus, the demand of the asset is based on their posterior belief. As we will show in the next section, there is no speculative trading benefit under the equilibrium price in equation (3), thus, the first component becomes zero.

B.2. Learning

Lemma 1 implies that the probability of the market opening for the high- and the low-quality asset, respectively, is given by:

$$\mu_H = 1 - F_y \left(\frac{\theta_H}{E[p|\theta_H]} - 1 \right), \quad (4)$$

$$\mu_L = 1. \quad (5)$$

¹⁸As in the supply of the seller, we assume a tie-breaking rule that the arbitrageurs choose to buy a good asset and sell in case they are indifferent between choices.

Because the market is less likely to be open for the high-quality asset relative to the low-quality asset, the fact that the investors are participating in the market delivers some information about the quality of the traded asset. Conditional on the market being open, the investors' probability assessment that the asset is of high quality is given by¹⁹

$$\hat{\rho} \equiv Pr(\theta = \theta_H | x^s > 0) = \frac{\rho\mu_H}{\rho\mu_H + (1 - \rho)}. \quad (7)$$

Uninformed agents (i.e., hedgers and investors) assess that the quality of the traded asset is likely to be poorer when they expect a greater undervaluation for the high-quality asset (i.e., $E[p|\theta_H]$ is lower). This reflects the classic lemons intuition that good assets are in smaller supply than lemons. This adverse selection problem becomes more severe as prices become more dislocated. Therefore, uninformed agents' prior belief of the quality of the asset being high is adjusted to be $\hat{\rho}$ rather than ρ .

Unlike the seller who influences the average quality of the traded asset, arbitrageurs influence the informativeness of prices and trading volumes. Arbitrageurs' order flow, $X_A(\theta)$, partially reveal the quality of the traded asset, θ , due to the existence of noise, $\eta\bar{z}$, in the net supply. That is, the investors infer a noisy signal, $\xi \equiv X_A(\theta) + \eta\bar{z}$, which is a sufficient statistic for the net supply ξ , then update their belief about θ using Bayes' rule:²⁰

LEMMA 4: *Given the prior belief $\hat{\rho}$, the investors' posterior belief that the traded asset is of a high quality conditional on ξ is given by*

$$q(\xi) = \begin{cases} \frac{\hat{\rho}f_{\eta}(\eta)}{\hat{\rho}f_{\eta}(\eta) + (1 - \hat{\rho})f_{\eta}(\eta + \Delta X/\bar{z})}; & \text{if } \theta = \theta_H; \\ \frac{\hat{\rho}f_{\eta}(\eta - \Delta X/\bar{z})}{\hat{\rho}f_{\eta}(\eta - \Delta X/\bar{z}) + (1 - \hat{\rho})f_{\eta}(\eta)}; & \text{if } \theta = \theta_L, \end{cases}$$

where ΔX is the maximum difference in the information component of net supply for

¹⁹Using Bayes' rule, we have

$$Pr(\theta = \theta_H | x^s > 0) = \frac{\rho Pr(x^s > 0 | \theta = \theta_H)}{\rho Pr(x^s > 0 | \theta = \theta_H) + (1 - \rho) Pr(x^s > 0 | \theta = \theta_L)}. \quad (6)$$

Then, equation (7) is immediate from equation (6) because $Pr(x^s > 0 | \theta = \theta_H) = 1 - F_y\left(\frac{\theta_H}{E[p|\theta_H]}\right) - 1$ and $Pr(x^s > 0 | \theta = \theta_L) = 1$.

²⁰Notice that the hedgers also share the same posterior belief as the investors because they can infer the investors' posterior belief from the price given in equation (3)

the high- and low- quality assets:

$$\Delta X \equiv X_A(\theta_H) - X_A(\theta_L) = \frac{w_A}{m^+} + \frac{w_A}{m^-}.$$

Proof: See Appendix B.

Q.E.D.

When the arbitrageurs have more capital (or liquidity), net supply reveals fundamental value better. Because ΔX measures the variability in the net supply due to arbitrage activities, we call it the “arbitrage trading variation.” Lemma 4 implies that the investors’ posterior belief becomes more accurate with a greater ΔX ; higher arbitrage trading variation generally increases the chance of revealing the fundamental value of the marketable asset. Because the noise in the supply, $\eta\bar{z}$, has bounded support, prices will fully reveal the quality of the traded asset when arbitrage trading variation, ΔX , is sufficiently large. As ΔX increases, $q(\xi)$ tends to approach one if $\theta = \theta_H$, and tends to approach zero if $\theta = \theta_L$.

Figure 2 illustrates this. When arbitrage trading variation, ΔX , is large as in Figure 2.(a), any realization of ξ will not overlap between the two cases with θ_H and θ_L . On the other hand, when arbitrage trading variation is small as in Figure 2.(b), a large portion of realizations of ξ are likely to overlap between the two cases with θ_H and θ_L . This results in a noisier price.

B.3. Equilibrium

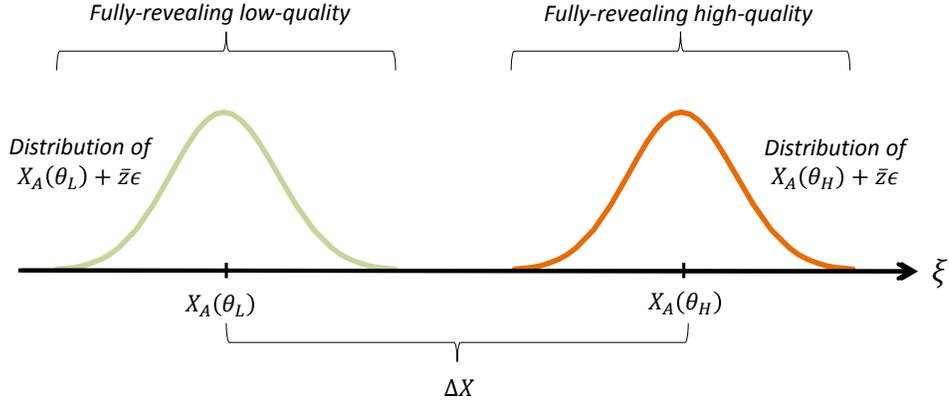
The following proposition shows that the model always has an equilibrium, and also provides its characteristics as well as its uniqueness condition:

PROPOSITION 1: (i) (*Existence of equilibrium*) *There always exists an equilibrium, that is, there exists an average asset quality of the traded asset $\hat{\rho} \in (0, \rho]$ that satisfies*

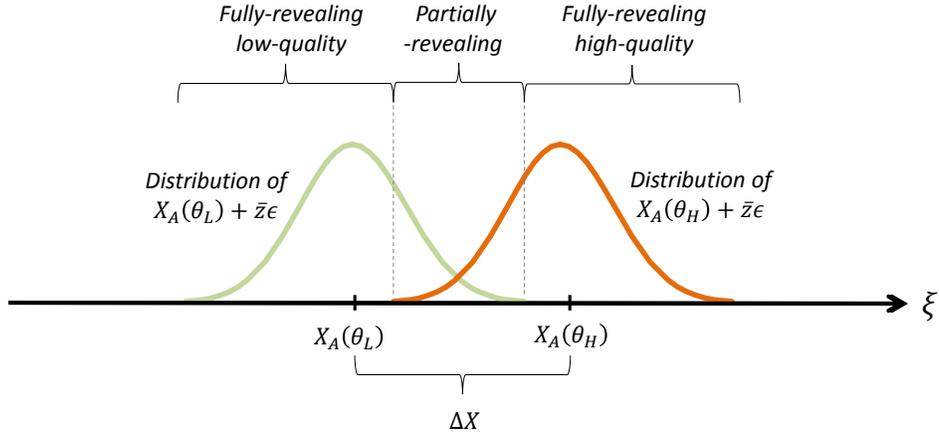
$$\hat{\rho} = H(\hat{\rho}), \tag{8}$$

where

$$H(\hat{\rho}) \equiv \frac{\rho(1 - F_y(\frac{\theta_H}{E[p(\theta, \eta; \hat{\rho})|\theta_H]} - 1))}{\rho(1 - F_y(\frac{\theta_H}{E[p(\theta, \eta; \hat{\rho})|\theta_H]} - 1)) + (1 - \rho)}.$$



(a) large arbitrage trading variation ΔX



(b) small arbitrage trading variation ΔX

Figure 2. An illustration of price informativeness. This figure depicts the degree of price informativeness under large and small arbitrage trading variation given $\xi \equiv X_A(\theta) + \eta z$ which is a sufficient statistic for net supply X .

The equilibrium price p given θ, η and $\hat{\rho}$ is given by

$$p(\theta, \eta; \hat{\rho}) = \frac{1}{1 + r_f} (q(\theta, \eta; \hat{\rho})\theta_H + (1 - q(\theta, \eta; \hat{\rho}))\theta_L), \quad (9)$$

with the weight $q(\theta, \eta; \hat{\rho})$ such that

$$q(\theta_H, \eta; \hat{\rho}) \equiv \frac{\hat{\rho}f_\eta(\eta)}{\hat{\rho}f_\eta(\eta) + (1 - \hat{\rho})f_\eta(\eta + \Delta X/\bar{z})}; \quad (10)$$

$$q(\theta_L, \eta; \hat{\rho}) \equiv \frac{\hat{\rho}f_\eta(\eta - \Delta X/\bar{z})}{\hat{\rho}f_\eta(\eta - \Delta X/\bar{z}) + (1 - \hat{\rho})f_\eta(\eta)}. \quad (11)$$

(ii) (*Uniqueness of equilibrium*) The equilibrium is unique if (a) $H'(\rho) < 1$ if $H(\rho) = \rho$, and (b) $H''(\hat{\rho}) \geq 0$ for all $\hat{\rho} \in [\rho', \rho]$ if $H''(\rho') > 0$.

Proof: See Appendix B.

Q.E.D.

Proposition 1.(i) shows existence of equilibrium by proving that there exists a fixed point for $\hat{\rho}$ that solves equation (8). Proposition 1.(ii) provides a sufficient condition for uniqueness of equilibrium that depends on the distribution of y and η . It is well known that the market for lemons can generally have multiple equilibria (e.g., Wilson (1979, 1980)). When low prices are expected, the seller is less inclined to sell a high quality asset, thereby lowering prices even further. This feedback effect may result in self-fulfilling prophecy that creates multiple equilibria.

A complete characterization of the set of necessary and sufficient conditions under which equilibrium is unique is beyond the scope of the paper. However, we need to address the question of uniqueness, since we will provide figures illustrating comparative statics of our model, based on numerical calculations of equilibrium. For these comparative statics to be more convincing, we should check that the equilibrium is unique in these calculations. The conditions in part (ii) of the propositions can easily be checked, so they allow us to verify this. Conditions (a) and (b) together rule out coexistence of two fixed points either on the interior or at the corner $\hat{\rho} = \rho$. Condition (a) rules out the possibility of a corner fixed point at $\hat{\rho} = \rho$ with a slope of $H(\cdot)$ greater than or equal to one. Condition (b) makes sure that $H(\cdot)$ does not become concave for higher values of $\hat{\rho}$ once it was convex for lower values of $\hat{\rho}$.²¹ All the comparative statics presented in this paper satisfy these sufficient conditions for uniqueness.

²¹For example, any concave function or any convex function satisfy condition (b). Also, any function which is concave for lower values, but convex for higher values satisfies it.

IV. Main Results

A. Fire Sales and Market Freezes

In this subsection, we show that fire sales and market freezes can occur due to reductions in the arbitrageurs' capital. When the arbitrageurs have enough capital, we find that there is no price dislocation. Intuitively, uninformed investors are willing to provide liquidity and supply capital to buy the asset at normal prices. In that case, the seller always sells his holdings of the marketable asset rather than giving up on their profitable non-marketable assets. Therefore, the supply of the marketable asset is insensitive to the quality of the asset, and is driven only by the liquidity needs of the seller.

On the other hand, in case arbitrageurs' capital is scarce, prices are dislocated. Intuitively, uninformed investors are not willing to provide liquidity at normal prices, but, they are only willing to supply capital to buy the asset at a fire sale price. In that case, seller's decisions depend on the asset quality. If the marketable asset is of low quality, he sells it. However, if the marketable asset is of high quality, he only sells it when divesting from the non-marketable asset is very costly. Therefore, average asset quality per trading volume goes down. This deterioration of average asset quality further depresses prices, thereby creating fire sales in case seller has to sell. To summarize, the reduction in arbitrageurs' capital leads to uninformative prices, which in turn cause further price drops through adverse selection. This idea is illustrated in Figure 3.

We now state our main results about fire sales and market freezes.

PROPOSITION 2: (i) (Fire sales) fire sales occur if and only if arbitrage capital is below $\bar{w}_A \equiv \bar{z} \left(\frac{m^+ m^-}{m^+ + m^-} \right)$, that is,

$$E[p|\theta_H] \in \begin{cases} \left\{ \frac{\theta_H}{1+r_f} \right\} & \text{if } w_A \geq \bar{w}_A; \\ \left[\frac{\theta_L}{1+r_f}, \frac{\theta_H}{1+r_f} \right) & \text{if } w_A < \bar{w}_A, \end{cases}$$

(ii) (Lemons problem) the average quality of the traded asset becomes poorer if and only if arbitrage capital is below \bar{w}_A , that is,

$$\hat{\rho} \in \begin{cases} \{\rho\} & \text{if } w_A \geq \bar{w}_A; \\ [0, \rho) & \text{if } w_A < \bar{w}_A. \end{cases}$$

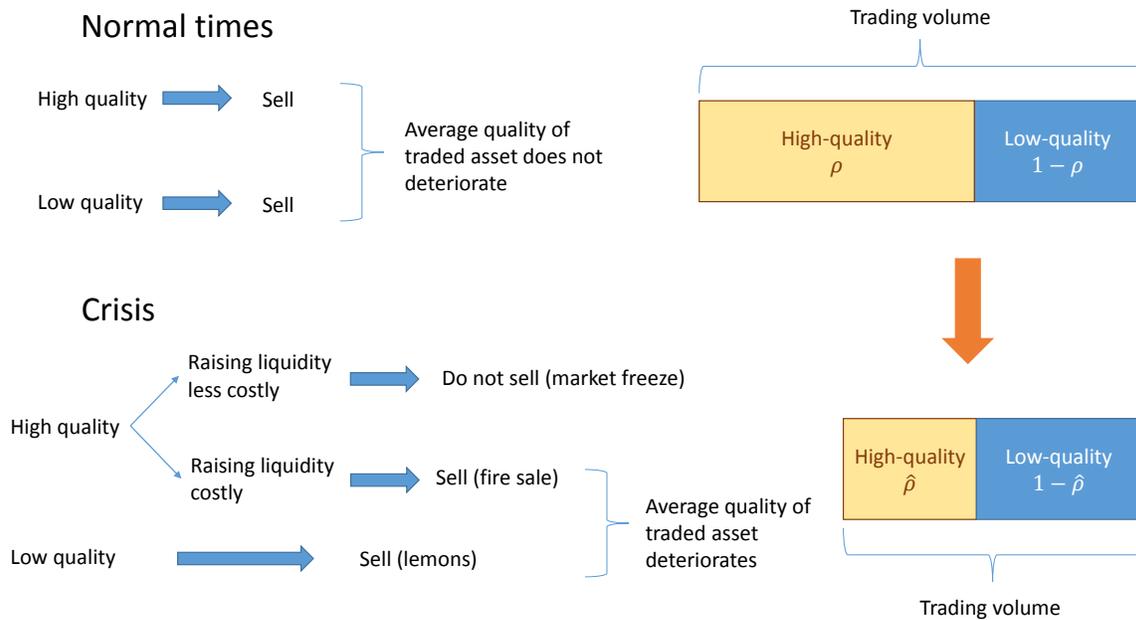


Figure 3. An illustration of endogenous lemons problem. This figure depicts how a crisis causes a lemons problem.

Proof: See Appendix B.

Q.E.D.

Proposition 2 shows that there is neither a fire sale nor a market freeze when there is enough capital for the arbitrageurs. \bar{w}_A is the level of arbitrage capital that eliminates the possibility of any fire sale or market freeze. We call this level of arbitrage capital the “fire-sale-free” level. On the other hand, insufficient arbitrage capital can create a double whammy of fire sales and market freezes. In case of the high-quality asset, a reduction in arbitrage capital creates a large price reduction because it affects both supply and demand. That is, the investors are less willing to absorb the net supply at any given level of prices.

Why do trades of high-quality assets shrink so much when there is a reduction in arbitrage capital? Answering this question is the main purpose of our analysis in this subsection. Consider the expected price at which a good quality asset can be sold. Then, a reduction in arbitrage capital has both a direct effect and an indirect effect that reduce demand, thereby precipitating price falls.

The direct effect, which we call the “price informativeness effect”, is related to price

dislocation. This effect is captured by our noisy REE framework that connects price informativeness to informed trading. The reduction of arbitrage capital creates an initial price fall for high-quality assets because the investors cannot infer the quality of the traded asset as much as before.

The indirect effect, which we call the “adverse selection effect”, is related to feedback effects between price dislocation and the lemons problem. When prices are noisy, the seller will not sell the high-quality asset unless they are effectively forced by very pressing liquidity needs (i.e., high liquidation costs for the non-marketable asset). That is, the initial price dislocation due to uninformative prices (i.e., the direct effect) worsens adverse selection of the seller, thereby lowering the average quality of the traded asset poorer. Because the average quality gets poorer, prices fall further, but this in turn lowers the quality further by worsening the adverse selection problem, and so on. Therefore, price becomes more dislocated as the overall quality of traded assets becomes poorer. This feedback mechanism causes the “double whammy” of a large fire sale discount and a market freeze for the high-quality asset.

When arbitrage capital is reduced from \bar{w}_A to w'_A , the fire sale discount (or the change in prices) can be decomposed into two components as follows:

$$\underbrace{\frac{E[v|\theta_H]}{1+r_f} - E[p'|\theta_H]}_{\text{fire sale discount}} = \underbrace{E\left[\frac{(\theta_H - \theta_L)}{1+r_f} \left(q(\theta_H, \eta; \rho, \bar{w}_A) - q(\theta_H, \eta; \rho, w'_A)\right)\right]}_{\text{price informativeness effect}} + \underbrace{E\left[\frac{(\theta_H - \theta_L)}{1+r_f} \left(q(\theta_H, \eta; \rho, w'_A) - q(\theta_H, \eta; \hat{\rho}', w'_A)\right)\right]}_{\text{adverse selection effect}},$$

where $q(\theta_H, \eta; \hat{\rho}, w_A)$ is the value of the weight $q(\theta_H, \eta; \hat{\rho})$ at the given level of w_A , and $E[p'|\theta_H]$ and $\hat{\rho}'$ are the expected price of the high-quality asset and the average quality of the traded assets given w'_A , respectively.

This idea is illustrated in Figure 4. The “demand” shown in Figure 4 can be viewed as an inverse demand curve as follows. For a given quantity offered for sale to the uninformed investors, the curve shows the price at which they are willing to buy that amount. Because the uninformed investors are risk-neutral and are making inferences from the quantity (net supply), this “inverse demand” interpretation is more appropriate than the usual “demand” interpretation as a quantity that will be bought by investors at a given price. First consider an equilibrium prior to the reduction in arbitrage capital.

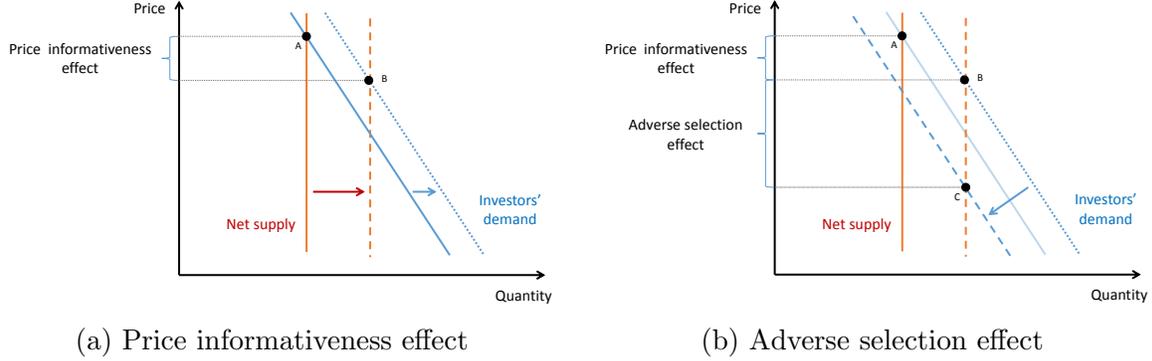


Figure 4. Price informativeness effect vs. adverse selection effect. This figure depicts fire sales of the high-quality asset.

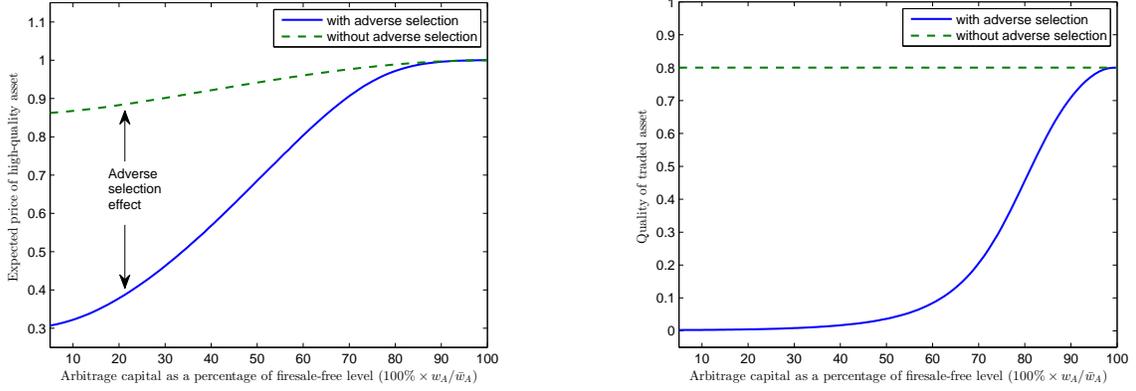
The equilibrium is determined at point A where demand and supply are matched given the initial size of arbitrage capital. Now, suppose that there is a reduction in arbitrage capital, and this shifts the supply curve outward. Because the shock to arbitrage capital is common knowledge among all participants, the investors incorporate that information, thus, their demand shifts outward to accommodate the change. However, the demand will not shift enough to support the same price as at point A because prices are now less revealing. Therefore, the price will be lowered to point B . This in turn results in a further decrease in prices through the lemons problem. That is, the reduction in arbitrage capital lowers price informativeness, thus, this increases adverse selection. Therefore, demand shrinks as a result of the increased lemons problem. As a result of the decrease in demand, the equilibrium is determined at point C rather than B .

For numerical examples, we choose the following parameter values unless stated otherwise: $\theta_H = 1, \theta_L = 0.3, \rho = 0.8, \bar{\epsilon} = 1, \bar{x} = 1, m^+ = m^- = 1, \gamma = 1, \bar{z} = 2, r_f = 0$. We also assume that η follows a beta distribution with shape parameters $\alpha_\eta = 2$ and $\beta_\eta = 2$, and $\hat{y} \equiv y - r_f$, which is the excess return of the non-marketable asset over the risk-free rate, follows a log-normal distribution with location parameter $\mu_y = -5$ and scale parameter $\sigma_y = 1.8$.²²

Figure 5 illustrates the impact of a reduction in arbitrage capital on the expected price and the average quality of the traded asset in case of the high-quality asset. When

²²The probability density function of η given parameters α and β is given by

$$f_\eta(\eta; \alpha, \beta) = \frac{\eta^{\alpha-1}(1-\eta)^{\beta-1}}{\int_0^1 \eta^{\alpha-1}(1-\eta)^{\beta-1} du}, \text{ for all } [0, 1];$$



(a) Expected price of high-quality asset ($E[p|\theta_H]$) (b) Average quality of traded asset ($\hat{\rho}$)

Figure 5. Comparative statics. This figure depicts the impact of a reduction in arbitrage capital on the expected price of the high-quality asset and the average quality of the traded asset (the x-axis is scaled as the percentage of \bar{w}_A).

there is enough arbitrage capital, the marketable asset is in full supply and price approaches fundamental value. As arbitrage capital falls, however, the fire sale discount increases (see the left panel of Figure 5). Also, the average quality of the traded asset falls because the supply of the high-quality asset decreases (see the right panel of Figure 5). That is, market freezes and fire sales occur together due to adverse selection. Furthermore, notice that the increase in the fire sale discount starts accelerating due to the adverse selection effect even for a relatively small reduction in arbitrage capital (i.e., where $\frac{w_A}{\bar{w}_A} \approx 90\%$).

The left panel of Figure 6 illustrates the impact of a reduction in arbitrage capital on the circulation rates of the high- and low- quality asset, respectively. When there is enough arbitrage capital, the seller sells the marketable asset regardless of its quality (i.e., $\mu_H = \mu_L = 1$). When there is a reduction in arbitrage capital, however, the seller is more likely to sell lemons rather than the high-quality asset (i.e., $\mu_H < 1$ and $\mu_L = 1$). The right panel of Figure 6 illustrates the impact of a reduction in arbitrage capital

and that of \hat{y} given parameter μ_y, σ_y is given by

$$f_y(\hat{y}; \mu_y, \sigma_y) = \frac{1}{\hat{y}\sigma_y\sqrt{2\pi}} \exp\left(-\frac{(\log(\hat{y}) - \mu_y)^2}{2\sigma_y^2}\right), \text{ for all } [0, \infty).$$

Our results are robust with other values of parameters or other types of distributions for both η and y such as truncated normal distributions, truncated exponential distributions, etc.

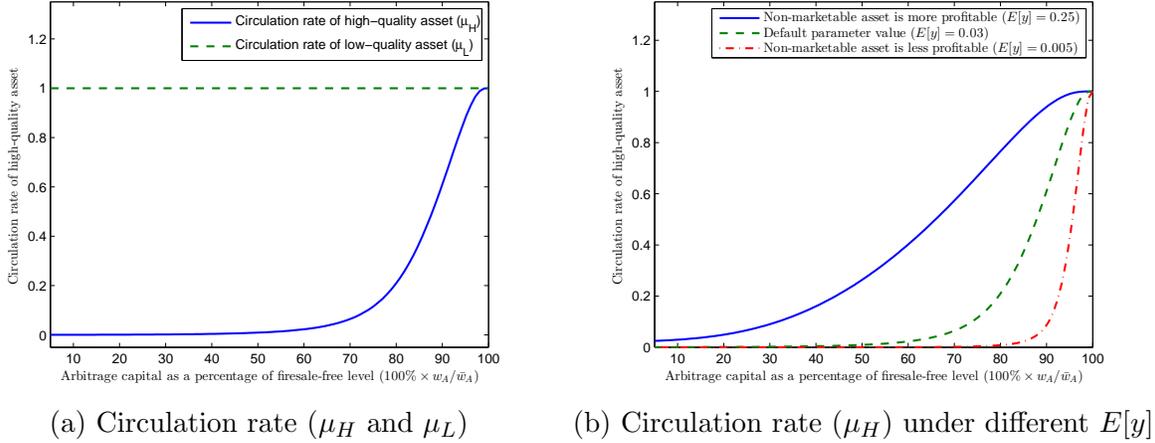


Figure 6. Comparative statics. This figure depicts the impact of a reduction in arbitrage capital on the circulation rate of the marketable asset (the x-axis is scaled as the percentage of \bar{w}_A).

on the circulation rate of the high-quality asset given different distributions of y (the return on the non-marketable asset).²³ When the non-marketable asset is less profitable (y is on average lower), the seller is more likely to liquidate the non-marketable asset instead of selling the high-quality asset. Therefore, there is more adverse selection with lower y (notice that the circulation rate falls down much faster when y is on average lower). The adverse selection problem is costly for the seller with the high-quality asset because it causes inefficient liquidation of the otherwise valuable non-marketable asset. Therefore, the circulation rate of the high-quality asset, μ_H , reflects how easily the seller can raise liquidity by selling the asset in the market. That is, μ_H proxies for the degree of liquidity in the market. By comparison to the classical informed trading literature (such as Grossman and Stiglitz (1980) and Kyle (1985)), our model contributes to the literature by suggesting an alternative mechanism that can create market illiquidity. It suggests that reductions in funding liquidity can decrease market liquidity through the lemons problem.

Figure 7 illustrates the impact of a reduction in arbitrage capital on the expected price of the high-quality asset under different parameter values the probability of being high quality, ρ , and the payoff of the marketable asset in the low state, θ_L . In the left

²³In the right panel of Figure 6, the value of location parameter μ_y for random variable y is given by $-3, -5, -7$, which yields the mean of y equal to 0.25, 0.03, 0.005, respectively.

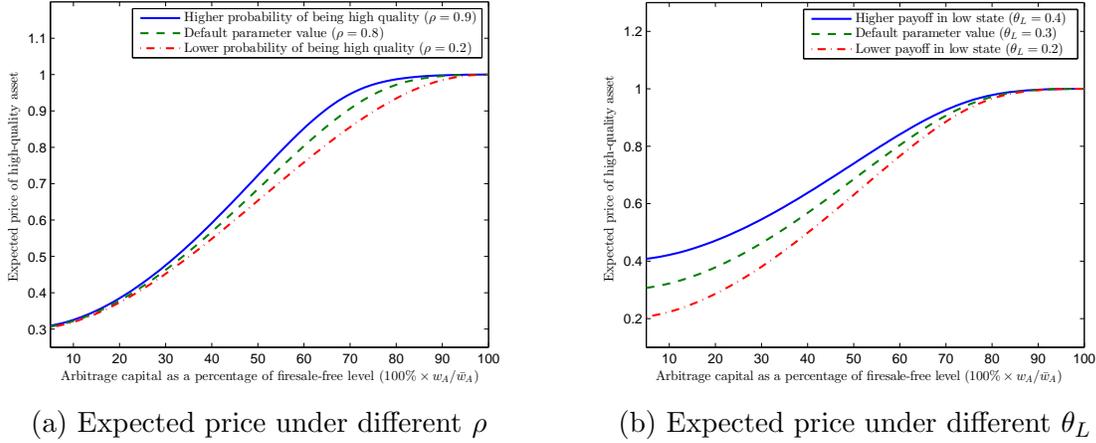


Figure 7. Comparative statics. This figure depicts the impact of a reduction in arbitrage capital on the expected price of the high-quality asset under various levels of parameter values (the x-axis is scaled as the percentage of \bar{w}_A).

panel of Figure 7, the fire sale discount increases faster when the marketable asset is more likely to be of low quality (i.e., ρ is lower). In case of fixed income securities, we can interpret that speculative-grade securities are more vulnerable to fire sales under small reductions in arbitrage capital. Notice, however, that fire sales eventually occur even for those assets with higher ρ (or investment-grade securities) as arbitrage capital decreases further. In the right panel of Figure 7, the fire sale discount increases faster when the payoff of the low-quality asset is on average lower (i.e., θ_L is lower). In case of fixed income securities, we can interpret that securities with smaller recovery value are more vulnerable to fire sales under small reductions in arbitrage capital.

B. Welfare Analysis

Because the marketable asset merely changes hands through trades, changes in welfare only arise from (i) allocations between the non-marketable assets and the risk-free asset, and (ii) sharing the risk of hedgers' wealth shocks. Recall that the non-marketable assets are ex-ante operating at an optimal level in the absence of liquidity constraints, so liquidation destroys value. Therefore, expected total wealth is maximized when liquidation of the non-marketable assets is minimized.²⁴ On the risk sharing side, the hedgers are

²⁴Mispricing caused by fire sales creates incentives to sacrifice the efficiency of investment. Then, those in liquidity shortage are forced to liquidate otherwise profitable non-marketable assets (even

risk-averse, thus, they want to eliminate any unwanted risk exposure as long as they have trading opportunities. Because market freezes prevent risk sharing, the hedgers are worse off when less arbitrage capital is around. However, there is some subtlety in the hedgers' expected utility: they may sometimes be better off with noisy prices due to the Hirshleifer (1971) effect. They are unable to insure against risk whose realizations are already reflected in prices (see, for example, Dow and Rahi (2000) for a further discussion). Such offsetting effects are dominated when the residual uncertainty of the marketable asset is large enough. Therefore, the hedgers' expected utility is also maximized with more informed arbitrage capital under such a condition.

We can show that the component of expected total wealth that is relevant to w_A is given by

$$W \equiv \left[\rho \left(E \left[y \mid y \geq \frac{\theta_H}{E[p|\theta_H]} - 1 \right] - r_f \right) E[p|\theta_H] + (1 - \rho)(E[y] - r_f)E[p|\theta_L] \right] \bar{x},$$

whose derivation is provided in the appendix. Notice that W represents the economic surplus created by the non-marketable asset, which may be destroyed in case of its liquidation. For risk sharing, we use the expected utility of the hedgers, which is given by

$$U^H \equiv \int_{h \in \mathcal{H}} E \left[\exp \left(-\gamma [(v - p(1 + r_f))x^h + vz^h] \right) \right] dh.$$

In line with Proposition 2, we find that the presence of enough arbitrage capital maximizes economic welfare by preventing fire sales and market freezes.²⁵

PROPOSITION 3: (i) (Expected total wealth) W is maximized when $w_A \geq \bar{w}_A$. (ii) (Risk sharing) U^H is maximized when $w_A \geq \bar{w}_A$ for sufficiently large $\bar{\epsilon}$ under the condition that $f_\eta(\cdot)$ is bounded away from zero on its support.

Proof: See Appendix B.

Q.E.D.

when they can improve firm value by further increasing investment in non-marketable assets). For example, Ivashina and Scharfstein (2010) find that bank lending decreased significantly after the onset of the financial crisis.

²⁵The condition that $f_\eta(\cdot)$ is strictly positive on the support $[0, 1]$ is a technical condition needed due to the Hirshleifer effect. The condition is stronger than necessary. It ensures that the possibility of hedgers having profits does not become unboundedly large relative to that of having losses at any level of arbitrage capital.

Proposition 3.(i) implies expected total wealth decreases with more mispricing due to misallocated resources. Figure 8 illustrates this with numerical examples.²⁶ The results show that negative spillover effects to expected total wealth are larger (a) when liquidating the non-marketable asset is less likely costly ($E[y]$ is lower), and (b) when the payoff of a low-quality asset is smaller (θ_L is lower). Similarly, Proposition 3.(ii) implies the hedgers' expected utility decreases with more mispricing due to reduced risk sharing opportunities. The figures demonstrate that negative spillover effects to risk sharing are larger (c) when the hedgers are more risk-averse (γ is higher), and (d) when the residual uncertainty of the marketable asset is higher ($\bar{\epsilon}$ is greater).

V. Extension: Endogenous Margin Constraints

In this section, we extend our baseline model by endogenizing margin constraints. Although margin constraints can be endogenized in various ways, we choose one of the simplest ways of doing it without introducing new frictions such as transaction costs, moral hazard, or other additional information asymmetries.²⁷ The purpose of our exercise here is (i) to show that our results are robust with endogenous margin constraints, and (ii) to further explore the implications of our baseline model related to endogenous margin constraints.

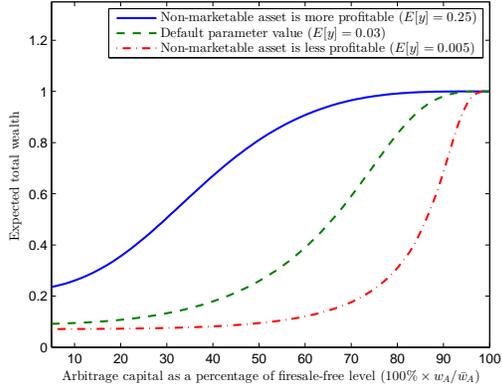
We consider the case where constrained margin traders (i.e., arbitrageurs) get margin loans from unconstrained participants in the economy (i.e., hedgers and investors). For analytical simplicity, we assume that margin loans are risk-free; margin constraints are endogenously determined so that margin loans offer the same rate of return as the risk-free rate to both lenders and borrowers.²⁸ The initial margin, which is the capital pledged by a margin trader as collateral, should be sufficient to cover any loss caused by his margin trading in every possible state of the world.²⁹ That is, the trader's pledged

²⁶The measure of total expected wealth is positive, whereas the measure of risk sharing (the expected utility of hedgers) is negative. Therefore, the normalized maximum of total expected wealth is equal to one, and the normalized maximum of risk sharing measure is equal to minus one.

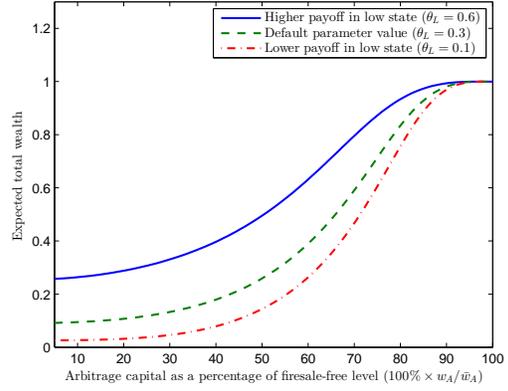
²⁷For example, Biais, Heider, and Hoerova (2016) endogenize margin constraints by introducing moral hazard of neglecting risk management by trading counterparties. They show that moral hazard can limit risk sharing, but margin calls can mitigate it, therefore, enhance risk sharing.

²⁸In case margin loans are risky, the interest rate charged on margin loans will reflect the risk of its default. We consider the risk-free loan for technical simplicity, but also to avoid confounding effects arising from the default.

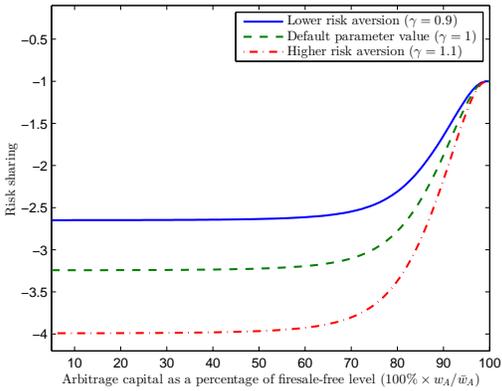
²⁹Therefore, lenders provide margin loans to any market participant as long as one can pledge some



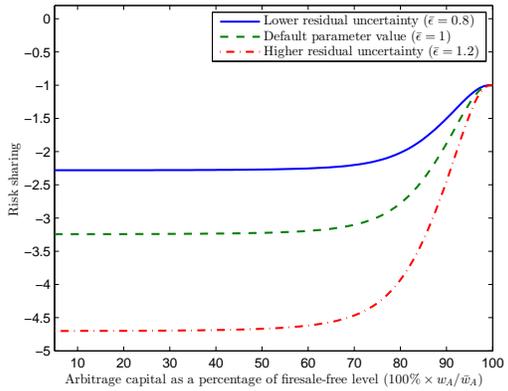
(a) Expected total wealth under different $E[y]$



(b) Expected total wealth under different θ_L



(c) Risk sharing under different γ



(d) Risk sharing under different $\bar{\epsilon}$

Figure 8. Comparative statics. This figure depicts the impact of a reduction in arbitrage capital on the economic welfare under various levels of parameter values (the x-axis is scaled as the percentage of \bar{w}_A , and the y-axis is scaled as the fraction of the maximum level of the economic welfare measures).

margin should be larger than his maximum possible loss due to low (high) payoffs of the marketable asset in case of long (short) positions. We further assume that capital in the margin account grows at the risk-free rate; If an arbitrageur keeps w_A of his capital in his margin account, it grows to $(1 + r_f)w_A$ at $t = 2$. Other than the setup involving

capital as the initial margin whether the participant possess expertise on the marketable asset or not. Alternatively, we can assume that lenders know whether the borrower possesses expertise or not. With that setup, we would also need to assume that $\bar{\epsilon} > \theta_H - \theta_L$ to generate risk of losing money even when they are informed about the quality of the marketable asset. This will simply relax margin constraints further by reducing the potential loss that informed margin traders can suffer, but will not affect our results in any other qualitative way.

margin constraints, we keep all other assumptions the same as in our baseline model.

We can show that margins are endogenously determined given the price p as follows:

LEMMA 5: *Given the price p , the long and short margins m^+, m^- are determined to be*

$$\begin{aligned} m^+ &= p - \frac{\theta_L - \bar{\epsilon}}{1 + r_f}; \\ m^- &= \frac{\theta_H + \bar{\epsilon}}{1 + r_f} - p. \end{aligned} \tag{12}$$

Proof: See Appendix.

Q.E.D.

Notice that margin traders are allowed greater margin trading capacity to buy assets as prices fall, and to short assets as prices rise. The intuition is that potential loss from a long (short) position becomes small when prices are already low (high). Furthermore, equation (12) reveals that such an increase in margin trading capacity will accelerate as prices approach either the lowest or the highest possible values (i.e., $\frac{\theta_L}{1+r_f}$ and $\frac{\theta_H}{1+r_f}$). Then, arbitrage trading variation, ΔX , given p is represented as

$$\Delta X \equiv \frac{w_A}{m^+} + \frac{w_A}{m^-} = \frac{w_A}{p - \frac{\theta_L - \bar{\epsilon}}{1+r_f}} + \frac{w_A}{\frac{\theta_H + \bar{\epsilon}}{1+r_f} - p}.$$

It is easy to verify that ΔX has an inverse hump-shaped relationship with p (i.e., it is maximized at either corner of the interval of $p \in [\frac{\theta_L}{1+r_f}, \frac{\theta_H}{1+r_f}]$, but is minimized in the middle of the interval). This implies that an increase in mispricing will reinforce itself by reducing trading capacity as prices start moving away from the true fundamental. The endogenous margin constraint is the source of such a feedback effect, but it does not qualitatively change our results in other ways. Moreover, all our previous results under the exogenous margin constraint in our baseline model are even further pronounced under the endogenous margin constraint (in other words, the magnitude of fire sales under the endogenous margin constraint is even greater than that under the exogenous margin constraint given the same percentage of reduction in arbitrage capital from the fire-sale-free level assuming all other things equal).

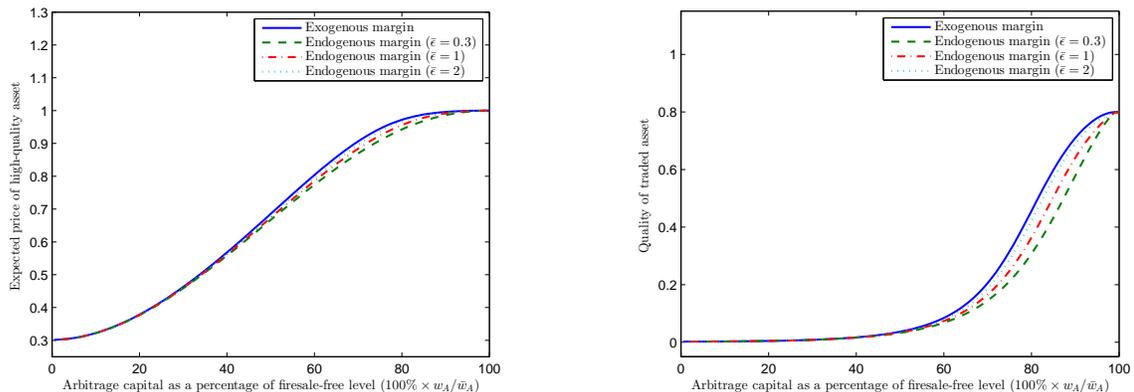
The rest of analysis under the endogenous margin constraint is identical to that of our baseline model, and we can further show that a model with the endogenous margin constraint becomes equivalent to those with the exogenous margin as the residual uncertainty becomes larger.

PROPOSITION 4: *As $\bar{\epsilon}$ becomes sufficiently large, prices of the marketable asset under the endogenous margin constraint and the exogenous margin constraint become asymptotically identical given any percentage decrease from the corresponding fire-sale-free level of arbitrage capital.*

Proof: See Appendix.

Q.E.D.

Proposition 4 implies any equilibrium property or comparative statics (e.g., quality of traded assets and welfare of participants) with respect to all the parameters of our model will also become identical regardless of assumptions on margin constraints as the residual uncertainty becomes large enough. Therefore, a setup with the exogenous margin constraint can be understood as the limit case of a setup with the endogenous margin constraint. Figure 9 confirms these observations using numerical examples with various parameter values of the residual uncertainty.



(a) Expected price of high-quality asset ($E[p|\theta_H]$)

(b) Average quality of traded asset ($\hat{\rho}$)

Figure 9. Comparative statics. This figure depicts the impact of a reduction in arbitrage capital on the expected price of the high-quality asset and the average quality of the traded asset under various assumptions on the margin constraint (the x-axis is scaled as the percentage of \bar{w}_A).

VI. Discussion

A. *Financial Fragility and Policy Implications*

We have shown that arbitrage capital can minimize mispricing and promote liquidity. It is worth stressing that arbitrageurs do not have to absorb the majority of the supply from the seller in our model. On the contrary, the uninformed investors absorb the major share of the supply of the risky asset whereas the arbitrageurs simply play a supporting role of setting the price close to the fundamental value. The role of arbitrage capital is, however, crucial because the investors absorb the supply due to the presence of arbitrage capital. A moderate amount of arbitrage capital is enough to make the market efficient and stable. We draw the conclusion that informed capital facilitates the movement of uninformed capital from the markets with excessive liquidity to those with lack of liquidity.

In our model, marketable assets can be traded without fire sales even with an arbitrarily large supply \bar{x} as long as arbitrage capital is enough to offset the effect of hedgers' wealth shock. The reason is that the uninformed investors with deep pockets (or high liquidity) are willing to absorb any amount of supply at the fair value as long as prices are set correctly by informed arbitrageurs.³⁰ In other words, even a small amount of informed capital can make the price efficient. Therefore, arbitrage capital creates information spillover effects in the market that multiplies its price-setting ability; the revelation of private information through prices makes otherwise-uninformed capital work like informed capital.

This efficiency of arbitrage capital may seem advantageous for financial stability, but it is in fact a double-edge sword. The potency of a small amount of arbitrage capital may actually be the reason for financial fragility because it means prices will be sensitive to a reduction of arbitrage capital. That is, small capital shocks to arbitrageurs in the market can trigger fire sales and market freezes by driving away the demand of uninformed investors. Indeed, as we have seen in Figure 5, small shocks to arbitrage capital create fire sales by exacerbating lemons problems in the market. Notice that in the figure, large drops in prices are triggered by small changes in arbitrage holdings

³⁰If the investors are risk-averse, they would require risk premium for holding a larger amount of risky assets. But, they would be willing to absorb more volume with uncertainty resolved by informed arbitrageurs, so, the effect of risk aversion would be relatively small in this argument.

(relative to total holdings). The “multiplier” effects of arbitrage capital can actually serve as the source of financial *instability* rather than financial stability.

Fire sales and market freezes create negative externalities and distort resource allocations. Our paper highlights the value of arbitrage capital in facilitating the efficient functioning of markets. This has policy implications for capital adequacy regulation.

Recent papers such as Mitchell, Pedersen, and Pulvino (2007) and Duffie (2010) suggest that institutional impediments such as search frictions, taxes, regulations, and market segmentation can slow down the speed of arbitrage capital, thereby creating fire sales and market freezes during market stress. But isn't there anybody who can try to arbitrage away obvious mispricing? A very few markets are truly closed to outsiders by regulatory fiat (writing insurance is an example) but most markets are open to investors who want to participate. Our results suggest an explanation for the underlying assumption of “slow-moving capital”; the presence of sufficient informed capital can facilitate movement of uninformed capital during normal times, but the market can quickly become illiquid if informed capital providers are liquidity constrained, as the resulting lemons problem prevents entry of uninformed capital. Indeed, recent empirical evidence seems to support our explanations. For example, Huang, Ringgenberg, and Zhang (2016) find that mutual fund fire sales, which are caused by a flow shock, create price pressure because fund managers choose to sell low-quality stocks. The resulting lemons problem depresses asset prices for a prolonged period of time due to unresolved information asymmetries.

Since the financial crisis, there has been an extensive debate about how to prevent fire sales because they can lead to negative welfare consequences in the economy (e.g., Shleifer and Vishny (2011), Tirole (2011)). The suggested remedies for fire sales can be usually categorized into two types. One is an ex-ante approach that reduces the possibility of fire sales (e.g., Krishnamurthy (2010), Diamond and Rajan (2011), Perotti and Suarez (2011)), and the other is an ex-post approach that mitigates the magnitude of fire sales and the following adverse effects (e.g., Shleifer and Vishny (2010), Diamond and Rajan (2011), Tirole (2012), Guerrieri and Shimer (2014)). Our theory has implications for both ex-ante and ex-post measures against fire sales.

On the ex-post side, our results imply that asset purchase programs may not be effective if asset purchase programs merely aim to reduce the net supply of assets sold in fire sales. Asset purchase programs might be effective if fire sales were caused simply

by cash-in-the-market pricing that does not involve information asymmetries. However, they will not be effective if fire sales are caused by lemons problems in the market because reducing asset supply in itself does not improve price informativeness.³¹ Therefore, uninformed capital would not move into the market in fire sales simply because of an asset purchase program. On the other hand, extending liquidity to arbitrageurs can improve price informativeness, thereby restoring the price mechanism that allows uninformed capital to participate in the market. Of course, incentive problems should be properly addressed because financial institutions that receive liquidity support may have very different objectives from the government agency that provides liquidity to them.

On the ex-ante side, our results imply that regulations on arbitrageurs such as capital requirements may or may not achieve desired effects of stabilizing the market. We have argued that lemons problems are endogenously determined by the availability of arbitrage capital, and that they have an intrinsic vulnerability to shocks. In Section IV, we showed that enough arbitrage capital will prevent fire sales, but that arbitrageurs do not have incentives to provide enough capital because they do not internalize negative spillover effects. In that sense, requiring financial institutions to keep enough capital for potential crisis could have a direct effect of lowering the frequency of fire sales. However, such tightening of capital requirements would lower the return on arbitrage capital. One might consider that in practice, this could have negative indirect effects, for example by inducing arbitrageurs to exit entirely.³²

B. Interpretation of the Financial Crisis of 2007-2008

In this section, we discuss the recent financial crisis in the light of our results in this paper, and find some implications on the policy or regulations. The recent financial crisis is often viewed as a consequence of the collapse of the housing bubble that grew through the 2000s. Along with real estate, other assets and securities were also overvalued. The aftermath of the collapse during 2007-2008 induced many financial institutions to unwind their positions as well as to reduce their leverage. This process created liquidity

³¹In Tirole (2012), asset purchase programs can still be effective in the presence of a lemons problem if the government is able to “clean up” the market by removing low-quality assets.

³²For example, capital may flow to unregulated sectors such as non-bank financial intermediaries rather than to regulated sectors. Therefore, fire sales would not be eliminated even with tightened capital requirements because capital is costly. The argument that a fire-sale-free equilibrium is impossible is parallel with the argument in Grossman and Stiglitz (1980) who find that informationally-efficient markets are not possible in equilibrium if information acquisition is costly.

shortage among leveraged institutions, thus created further illiquidity spillover effects in other markets. There was market freeze across a large class of assets in particular for structured financial products.³³

Gorton and Metrick (2012) document that banks suffered liquidity shortage during the financial crisis. They further find that relatively small amount of subprime risk lead to spread rises of unrelated asset classes. Our theory suggests a reduction in informed trading can act as a transmission mechanism of liquidity shocks to unrelated asset classes. During the onset of the recent financial crisis, there was a liquidity crunch among the financial institutions as a result of downturn in housing markets (and subsequent price falls in structured products on housing mortgages). The liquidity constraints affected the arbitrage trading variation of financial institutions which are likely to be more informed about many securities they were trading. Merrill, Nadauld, Stulz, and Sherlund (2014) find evidence about fire sales of RMBS in the aftermath of the subprime crisis.

In our model, the seller decides whether to bring their endowment of the marketable asset to the market or not. Bringing the asset to the market can be interpreted as trading the asset in the market, but it can be also alternatively interpreted as creating the asset. Vickery and Wright (2013) document that issuance of non-agency RMBS (or private-label RMBS) decreased significantly relative to agency RMBS since mid-2007 and, during this period, secondary markets for trading non-agency RMBS were extremely illiquid. Non-agency RMBS are less regulated and tend to include riskier underlying loans (i.e., subprime mortgages) relative to agency RMBS, thus, they are more subject to information asymmetries than agency RMBS. Our theory can explain why there was a collapse of non-agency RMBS market during the recent crisis.

VII. Conclusions

In our paper, we have developed an information-based theory of fire sales using a noisy REE framework with endogenous lemons problem. Our model combines limits to arbitrage and adverse selection to provide a plausible mechanism of financial fire sales and market freezes. In a situation when informed market participants are not highly liquidity-constrained, arbitrage activity is high and prices are informative. This allows uninformed investors to absorb the supply of assets without worrying about adverse

³³See, for example, Brunnermeier (2009).

selection. Therefore, arbitrage capital can “multiply” itself by making uninformed capital effectively function as informed capital in normal market. When arbitrageurs are liquidity-constrained, however, arbitrage activity is reduced and prices become less informative. This creates an adverse selection problem, increasing the supply of low-quality assets. This lemons problem makes well-capitalized uninformed market participants unwilling to supply capital to support the price, thereby freezing the market. This can explain the “double whammy” in which fire sales and market freezes occur together. Our results shed light on the paradoxical nature of fire sales in which capital moves out of the market when it is needed most and apparently would earn higher returns. Furthermore, our results show how a financial fire sale can reduce economic welfare by reducing expected total wealth as well as risk sharing.

The model has implications for financial fragility. Only a small amount of arbitrage capital is required to make the market work efficiently because, so long as it ensures assets are priced efficiently, this can facilitate the movement of uninformed capital. However, this is a double-edge sword because a small reduction in arbitrage capital can create a lemons problem and slow down the movement of uninformed capital, causing a fire sale. Therefore, the multiplier effect of arbitrage capital can also serve as the source of financial *instability* rather than financial stability.

Appendix A. List of symbols

v	payoff of marketable asset
θ	observable component of v where $\theta \in \{\theta_H, \theta_L\}$
ϵ	unobservable component of v where $\epsilon \in \{\bar{\epsilon}, -\bar{\epsilon}\}$
ρ	probability that marketable asset pays θ_H
\bar{x}	seller's endowment of marketable asset
l	seller's liquidity shortage
r_f	risk-free rate
y	return on non-marketable asset
w_A	liquidity position of arbitrageur
\bar{w}_A	fire-sale-free level of arbitrage capital
m^+	dollar margin on arbitrageur's long position
m^-	dollar margin on arbitrageur's short position
γ	risk aversion parameter of hedger
η	portion of hedgers who are affected by random wealth shock
x^a	arbitrageur $a \in \mathcal{A}$'s order flow where \mathcal{A} is the set of arbitrageurs
x^h	hedger $h \in \mathcal{H}$'s order flow where \mathcal{H} is the set of hedgers
z^h	hedger $h \in \mathcal{H}$'s wealth shock where $z^h \in \{0, -\bar{z}\}$
x^s	seller's supply of marketable asset
X_A	arbitrageurs' aggregate order flow where $X_A \equiv \int_{x^a \in \mathcal{A}} x^a da$
X_H	hedgers' aggregate order flow where $X_H \equiv \int_{x^h \in \mathcal{H}} x^h dh$
X	net supply of marketable asset where $X \equiv x^s - X_A - X_H$
ξ	sufficient statistic for X where $\xi \equiv X_A + \eta\bar{z}$
ΔX	arbitrage trading variation
μ_H	circulation rate of high-quality asset
μ_L	circulation rate of low-quality asset
$\hat{\rho}$	average quality of traded asset
q	investors' posterior belief that $\theta = \theta_H$
W	expected total wealth
U^H	hedgers' expected utility (or, measure of risk sharing)

Appendix B.

Proof of Lemma 3: Each hedger $h \in \mathcal{H}$ maximizes his expected utility of future consumption given price p and wealth shock vz^h :

$$\max_{x^h} \sum_{\theta \in \{\theta_H, \theta_L\}} \sum_{\epsilon \in \{\bar{\epsilon}, -\bar{\epsilon}\}} \left(\frac{Pr(\theta|p)}{2} \right) u(x^h, z^h; \theta, \epsilon),$$

where $u(x^h, z^h; \theta, \epsilon)$ denotes hedger h 's utility with portfolio choice x^h given the realizations of z^h, θ, ϵ

$$u(x^h, z^h; \theta, \epsilon) = -\frac{1}{\gamma} e^{-\gamma[w_H(1+r_f) + (\theta + \epsilon - p(1+r_f))x^h + (\theta + \epsilon)z^h]}, \quad (\text{B.1})$$

and $Pr(\theta|p)$ denotes the hedgers' posterior belief that the quality of the traded asset is θ conditional on p .³⁴

The first order condition is

$$-\gamma \sum_{\theta, \epsilon} \left(\frac{Pr(\theta|p)}{2} \right) (\theta + \epsilon - p(1+r_f)) u(x^h, z^h; \theta, \epsilon) = 0, \quad (\text{B.2})$$

and the second order condition is always satisfied:

$$\gamma^2 \sum_{\theta, \epsilon} \left(\frac{Pr(\theta|p)}{2} \right) (\theta + \epsilon - p(1+r_f))^2 u(x^h, z^h; \theta, \epsilon) < 0.$$

Therefore, there exists a unique, interior, solution for this optimization problem.

By defining $\hat{x}^h \equiv x^h + z^h$, we can alternatively express equation (B.1) as

$$u(x^h, z^h; \theta, \epsilon) = e^{-\gamma p(1+r_f)z^h} u(\hat{x}^h, 0; \theta, \epsilon),$$

so the optimization problem becomes

$$\max_{\hat{x}^h} e^{-\gamma p(1+r_f)z^h} \sum_{\theta, \epsilon} \left(\frac{Pr(\theta|p)}{2} \right) u(\hat{x}^h, 0; \theta, \epsilon). \quad (\text{B.3})$$

³⁴Observing z^h does not help inferring the quality of the traded asset because each individual hedger is infinitesimal and the total mass of affected hedgers, η , is random.

This implies that solving for an optimization problem with wealth shock z^h is equivalent to solving it without wealth shock ($z^h = 0$), then, subtracting the size of the shock z^h from the optimal solution $\hat{x}^{h,*}$ that satisfies the following first-order condition for the alternative optimization problem (B.3):

$$-\gamma e^{-\gamma p(1+r_f)z^h} \sum_{\theta, \epsilon} \left(\frac{Pr(\theta|p)}{2} \right) (\theta + \epsilon - p(1+r_f)) u(\hat{x}^h, 0; \theta, \epsilon) = 0. \quad (\text{B.4})$$

Notice that the left-hand side of equation (B.4) exceeds zero when \hat{x}^h is sufficiently small and is strictly less than zero when \hat{x}^h is sufficiently large.³⁵ Because the left-hand side of equation (B.4) is monotone decreasing, there exists a unique solution $X^{CARA}(p)$ for equation (B.4), which is a function of p but is independent of z^h . Therefore, solving equation (B.2) for x^h gives the optimal portfolio given p and z^h as follows:

$$x^{h,*} = X^{CARA}(p) - z^h.$$

By aggregating each individual demand x^h across all hedgers in \mathcal{H} , we obtain the aggregate demand of hedgers. Q.E.D.

Proof of Lemma 4: The investors observe the net supply $X = \bar{x} - (X^{CARA}(p) + \eta\bar{z}) - X_A(\theta)$, and use this to update their beliefs about θ . Because \bar{x} and $X^{CARA}(p)$ are known quantities for the investors, they can infer $\xi \equiv X_A(\theta) + \eta\bar{z}$, which is a noisy signal about θ . Using Bayes' rule, the investors' posterior beliefs that the traded asset is of high quality can be derived as follows:

$$q(\xi) = \frac{\hat{\rho} f_\eta((\xi - X_A(\theta_H))/\bar{z})}{\hat{\rho} f_\eta((\xi - X_A(\theta_H))/\bar{z}) + (1 - \hat{\rho}) f_\eta((\xi - X_A(\theta_L))/\bar{z})}. \quad (\text{B.5})$$

Now, suppose $\theta = \theta_H$. Then, we have $\xi = X_A(\theta_H) + \eta\bar{z}$. Equation (B.5) implies

$$q(\xi) = \frac{\hat{\rho} f_\eta(\eta)}{\hat{\rho} f_\eta(\eta) + (1 - \hat{\rho}) f_\eta(\eta + (X_A(\theta_H) - X_A(\theta_L))/\bar{z})}. \quad (\text{B.6})$$

³⁵Recall that p in equation (3) is between $\frac{\theta_L}{1+r_f}$ and $\frac{\theta_H}{1+r_f}$ due the risk-neutrality of the investors.

Likewise, suppose $\theta = \theta_L$. Then, we have $\xi = X_A(\theta_L) + \eta\bar{z}$. Equation (B.5) implies

$$q(\xi) = \frac{\hat{\rho}f_\eta(\eta - (X_A(\theta_H) - X_A(\theta_L))/\bar{z})}{\hat{\rho}f_\eta(\eta - (X_A(\theta_H) - X_A(\theta_L))/\bar{z}) + (1 - \hat{\rho})f_\eta(\eta)}. \quad (\text{B.7})$$

Substituting $\Delta X \equiv X_A(\theta_H) - X_A(\theta_L)$ into equations (B.6) and (B.7) completes the proof. Q.E.D.

Proof of Proposition 1: (i) (Existence of equilibrium) We prove existence of equilibrium by construction in two steps. First, we fix the quality of the traded asset $\hat{\rho}$ as given, allowing us to solve for a unique price p that clears the market at any given level of θ and η . Second, we show that there exists a solution for $\hat{\rho}$ given the expected price $E[p|\theta]$ using the price function derived in the first step.

We start with the first step by fixing $\hat{\rho}$. Because the investors are competitive and risk-neutral, Lemma 4 implies that the equilibrium price given θ and η should be uniquely determined as follows:

$$p(\theta, \eta; \hat{\rho}) = \frac{1}{1 + r_f} (q(\theta, \eta; \hat{\rho})\theta_H + (1 - q(\theta, \eta; \hat{\rho}))\theta_L),$$

where the weight $q(\theta, \eta; \hat{\rho})$ is defined by

$$q(\theta_H, \eta; \hat{\rho}) \equiv \frac{\hat{\rho}f_\eta(\eta)}{\hat{\rho}f_\eta(\eta) + (1 - \hat{\rho})f_\eta(\eta + \Delta X/\bar{z})}; \quad (\text{B.8})$$

$$q(\theta_L, \eta; \hat{\rho}) \equiv \frac{\hat{\rho}f_\eta(\eta - \Delta X/\bar{z})}{\hat{\rho}f_\eta(\eta - \Delta X/\bar{z}) + (1 - \hat{\rho})f_\eta(\eta)}. \quad (\text{B.9})$$

Then, the expected price of the high-quality asset is given by

$$E[p(\theta, \eta; \hat{\rho})|\theta_H] = \frac{1}{1 + r_f} \left[\theta_L + (\theta_H - \theta_L) \left(1 - F_\eta(\eta^*) + \int_0^{\eta^*} q(\theta_H, \eta; \hat{\rho})f_\eta(\eta)d\eta \right) \right], \quad (\text{B.10})$$

where

$$\eta^* \equiv \max\left(1 - \frac{\Delta X}{\bar{z}}, 0\right). \quad (\text{B.11})$$

Now, we turn to the second step that proves existence of the equilibrium supply from the seller. From equations (4) and (7), in equilibrium the quality of the traded asset $\hat{\rho}$

should satisfy:

$$\hat{\rho} = H(\hat{\rho}), \quad (\text{B.12})$$

where

$$H(\hat{\rho}) \equiv \frac{\rho(1 - F_y(\frac{\theta_H}{E[p(\theta, \eta; \hat{\rho})|\theta_H]} - 1))}{\rho(1 - F_y(\frac{\theta_H}{E[p(\theta, \eta; \hat{\rho})|\theta_H]} - 1)) + (1 - \rho)}.$$

Notice that equation (B.8) implies that $E[p(\theta, \eta; \hat{\rho})|\theta_H] = \frac{\theta_L}{1+r_f}$ when $\hat{\rho} = 0$. Therefore, we have

$$H(0) = \frac{\rho(1 - F_y(r_f + \frac{(1+r_f)(\theta_H - \theta_L)}{\theta_L}))}{\rho(1 - F_y(r_f + \frac{(1+r_f)(\theta_H - \theta_L)}{\theta_L})) + (1 - \rho)} > 0.$$

We can also easily verify that $H(\rho) \leq \rho$ using the definition of $H(\cdot)$. Because $H(\cdot)$ is continuous in $\hat{\rho}$, there must exist a fixed point solving equation (B.12) on the set $(0, \rho]$.

(ii) (Uniqueness of equilibrium) To prove that the equilibrium is unique, it suffices to show that $H(\cdot)$ has a single fixed point either in the interior $(0, \rho)$ or at the corner $\hat{\rho} = \rho$. We first prove the following lemma that will be used later:

LEMMA B.6: $H(\hat{\rho})$ monotone increases in $\hat{\rho}$.

Proof: From equation (B.10), differentiating $E[p(\theta, \eta; \hat{\rho})|\theta_H]$ with respect to $\hat{\rho}$ yields

$$\frac{\partial E[p(\theta, \eta; \hat{\rho})|\theta_H]}{\partial \hat{\rho}} = \frac{\theta_H - \theta_L}{1 + r_f} \int_0^{\eta^*} \frac{\partial q(\theta_H, \eta; \hat{\rho})}{\partial \hat{\rho}} f_\eta(\eta) d\eta, \quad (\text{B.13})$$

where η^* is the threshold of η defined in equation (B.11). Notice that, for all $\eta \in [0, \eta^*]$, we have

$$\frac{\partial q}{\partial \hat{\rho}} = \frac{f_\eta(\eta) f_\eta(\eta + \Delta X/\bar{z})}{(\hat{\rho} f_\eta(\eta) + (1 - \hat{\rho}) f_\eta(\eta + \Delta X/\bar{z}))^2} > 0,$$

which implies that $\frac{\partial E[p(\theta, \eta; \hat{\rho})|\theta_H]}{\partial \hat{\rho}} > 0$. Thus, the following is immediate:

$$\frac{\partial H(\hat{\rho})}{\partial \hat{\rho}} = \frac{\rho(1 - \rho) f_y(\frac{\theta_H}{E[p(\theta, \eta; \hat{\rho})|\theta_H]} - 1)}{[\rho(1 - F_y(\frac{\theta_H}{E[p(\theta, \eta; \hat{\rho})|\theta_H]} - 1)) + (1 - \rho)]^2} \frac{\theta_H}{E[p|\theta_H]^2} \frac{\partial E[p|\theta_H]}{\partial \hat{\rho}} > 0.$$

Q.E.D.

Next, we show that there can exist only one fixed point that solves $H(\hat{\rho}) = \hat{\rho}$ under

the conditions provided in Proposition 1.(ii). Suppose that those conditions are true, that is, (a) $H'(\rho) < 1$ if $H(\rho) = \rho$, and (b) $H''(\hat{\rho}) \geq 0$ for all $\hat{\rho} \in [\rho', \rho]$ if $H''(\rho') > 0$.

We first prove that there is no interior fixed point, if there exists a corner fixed point:

LEMMA B.7: *Under conditions (a) and (b), if $H(\rho) = \rho$, then, there is no $\hat{\rho} \in (0, \rho)$ such that $H(\hat{\rho}) = \hat{\rho}$.*

Proof: Assume that $H(\rho) = \rho$ but there also exists at least one interior fixed point. Let $\hat{\rho}^* \in (0, \rho)$ be the smallest fixed point among those interior fixed points. Because $H(0) > 0$, Lemma B.6 implies $H'(\hat{\rho}^*) \in [0, 1)$ under conditions (a) and (b).³⁶ Given condition (b), there are two possibilities: either $H(\cdot)$ is concave throughout the interval of $[\hat{\rho}^*, \rho]$, or becomes convex at some point on the interval. First, suppose that $H''(\hat{\rho}) \leq 0$ for all $\hat{\rho} \in [\hat{\rho}^*, \rho]$. Then, it implies $H(\hat{\rho}) < \hat{\rho}$ for all $\hat{\rho} \in (\hat{\rho}^*, \rho]$ because $H(\hat{\rho}^*) = \hat{\rho}^*$ and $H'(\hat{\rho}^*) < 1$. It contradicts with the assumption that $H(\rho) = \rho$. Now, suppose that there exists $\tilde{\rho} \in [\hat{\rho}^*, \rho]$ such that $H''(\hat{\rho}) > 0$ for all $\hat{\rho} \in [\tilde{\rho}, \rho]$. But, condition (b) together with $H(\rho) = \rho$ would then imply $H'(\rho) \geq 1$. This contradicts with condition (a). Therefore, there cannot exist an interior fixed point if $H(\rho) = \rho$. Q.E.D.

Now, we show that there can only be at most one interior fixed point:

LEMMA B.8: *Under conditions (a) and (b), if there exist $\hat{\rho}^*, \hat{\rho}^{**} \in (0, \rho)$ such that $H(\hat{\rho}^*) = \hat{\rho}^*$ and $H(\hat{\rho}^{**}) = \hat{\rho}^{**}$, then, $\hat{\rho}^* = \hat{\rho}^{**}$.*

Proof: Assume that $\hat{\rho}^* \neq \hat{\rho}^{**}$. Let $\hat{\rho}^*$ be the smaller one between the two, that is, $\hat{\rho}^* < \hat{\rho}^{**}$. As in the proof of Lemma B.7, the given conditions imply $H'(\hat{\rho}^*) \in [0, 1)$ with two possibilities. First, suppose $H''(\hat{\rho}) \leq 0$ for all $\hat{\rho} \in [\hat{\rho}^*, \rho]$. Then, it implies $H(\hat{\rho}) < \hat{\rho}$ for all $\hat{\rho} \in (\hat{\rho}^*, \rho]$ because $H(\hat{\rho}^*) = \hat{\rho}^*$ and $H'(\hat{\rho}^*) < 1$. It contradicts with the assumption that $H(\hat{\rho}^{**}) = \hat{\rho}^{**}$ and $\hat{\rho}^* < \hat{\rho}^{**}$. Now, suppose that there exists $\tilde{\rho} \in [\hat{\rho}^*, \rho]$ such that $H''(\hat{\rho}) > 0$ for all $\hat{\rho} \in [\tilde{\rho}, \rho]$. Then, it is clear that $\tilde{\rho} \in [\hat{\rho}^*, \hat{\rho}^{**})$ and $H'(\hat{\rho}^{**}) > 1$. Condition (b) implies that $H'(\hat{\rho}) > 1$ for all $\hat{\rho} \in [\hat{\rho}^{**}, \rho]$, which in turn implies $H(\rho) > \rho$. But, this contradicts with $H(\rho) \leq \rho$. Therefore, it should be the case that $\hat{\rho}^* = \hat{\rho}^{**}$. Q.E.D.

Finally, Lemma B.7 and Lemma B.8 together imply that there cannot exist more than one fixed points on the set $(0, \rho]$. Therefore, conditions (a) and (b) are sufficient

³⁶Suppose $H'(\hat{\rho}^*) \geq 1$. Condition (b) implies that $H''(\hat{\rho}^*) > 0$ because $H(0) > 0$ and $H(\cdot)$ monotone increases in $\hat{\rho}$. Under condition (b), however, this in turn implies $H'(\rho) \geq 1$, which contradicts with condition (a).

to ensure uniqueness of equilibrium.

Q.E.D.

Proof of Proposition 2: From equation (9), we can represent the expected price of the high-quality asset given w_A as follows:

$$E[p(\theta, \eta; \hat{\rho})|\theta_H] = \frac{1}{1+r_f} \left[\theta_L + (\theta_H - \theta_L) \left(1 - F_\eta(\eta^*) + \int_0^{\eta^*} q(\theta_H, \eta; \hat{\rho}) f_\eta(\eta) d\eta \right) \right], \quad (\text{B.14})$$

where

$$\eta^* \equiv \max \left(1 - \frac{\Delta X}{\bar{z}}, 0 \right) = \max \left(1 - \frac{w_A}{\bar{z}} \left(\frac{1}{m^+} + \frac{1}{m^-} \right), 0 \right).$$

Notice that $q(\theta_H, \eta; \hat{\rho}) < 1$ whenever $\eta < \eta^*$.³⁷ Therefore, equation (B.14) implies that $E[p(\theta, \eta; \hat{\rho})|\theta_H] = \frac{\theta_H}{1+r_f}$ if and only if $\eta^* = 0$, or equivalently,

$$w_A \geq \bar{w}_A \equiv \bar{z} \left(\frac{m^+ m^-}{m^+ + m^-} \right).$$

Q.E.D.

Proof of Proposition 3: (i) (Expected total wealth): We can represent the ex-ante expectation of the aggregate final wealth of each type of participants (in the order of seller, arbitrageurs, hedgers, and investors) as follows:

$$\begin{aligned} W^S &\equiv E[v(\bar{x} - x^s) + (1+y)(px^s - l)]; \\ W^A &\equiv \int_{a \in \mathcal{A}} E[(v - p(1+r_f))x^a] da + (1+r_f)w_A; \\ W^H &\equiv \int_{h \in \mathcal{H}} E[(v - p(1+r_f))x^h + vz^h] dh + (1+r_f)w_H; \\ W^I &\equiv \int_{i \in \mathcal{I}} E[(v - p(1+r_f))x^i] di + (1+r_f)w_I. \end{aligned}$$

where w_A, w_H, w_I denote the initial wealth of arbitrageurs, hedgers, and investors, respectively. Let y^* denote the threshold of y above which the seller sells the high-quality

³⁷From equation (10), it is easy to see that $q(\theta_H, \eta; \hat{\rho}) < 1$ if and only if $\eta + \Delta X/\bar{z} < 1$ (or equivalently, $\eta < \eta^*$).

marketable asset, and below which he does not, that is,

$$y^* \equiv \frac{\theta_H}{E[p(\theta, \eta; \hat{\rho})|\theta_H]} - 1.$$

Summing up the ex-ante expected final wealth of all the participants gives us the ex-ante expectation of the aggregate wealth in the economy, $AW \equiv W^S + W^A + W^H + W^I$:³⁸

$$AW = W + W_0, \tag{B.15}$$

where

$$\begin{aligned} W &\equiv \bar{x} \left[\rho \mu_H E[p|\theta_H] (E[y|y \geq y^*] - r_f) + (1 - \rho) E[p|\theta_L] (E[y] - r_f) \right]; \\ W_0 &\equiv E[v] E[\eta] \bar{z} - (1 + E[y]) l + (1 + r_f) (w_A + w_H + w_I). \end{aligned}$$

The first term, W , in equation (B.15) is used as a measure of expected total wealth (or economic surplus) at the given level of arbitrage capital; it measures the ex-ante expectation of extra wealth created by the non-marketable asset at the given level of arbitrage capital w_A . It is because the way that expected wealth can increase is via less frequent liquidation of the non-marketable asset, as represented in W . The other term, W_0 , captures other components irrelevant to economic surplus such as reallocations of the initial wealth.³⁹

Let $W(w_A)$ be a function of w_A which is equal to W given w_A in equation (B.15). For any $w_A \geq \bar{w}_A$, there is no fire sale in any state of the world, thus, we have $W(w_A) = W(\bar{w}_A)$ where

$$W(\bar{w}_A) = \frac{\bar{x}}{1 + r_f} \left[\rho \theta_H + (1 - \rho) \theta_L \right] (E[y] - r_f). \tag{B.16}$$

³⁸Recall that we have the following equation in equilibrium:

$$\bar{x} = \int_{a \in \mathcal{A}} x^a da + \int_{h \in \mathcal{H}} x^h dh + \int_{i \in \mathcal{I}} x^i di.$$

³⁹There are redistributions, captured in the first two terms of W_0 , but their total does not change: $E[v] E[\eta] \bar{z} - (1 + E[y]) l$ is unaffected by w_A . We also exclude the last term of W_0 , $(1 + r_f) (w_A + w_H + w_I)$, because it mechanically reflects the increase in wealth from the initial wealth - in other words, when performing comparative statics on wealth creation caused by increasing w_A , we net off the wealth change that results directly.

Using equation (B.16), we can represent $W(\bar{w}_A)$ as

$$\begin{aligned}
W(\bar{w}_A) &= \frac{\rho\theta_H\bar{x}}{1+r_f} \left[(1-\mu_H)(E[y|y < y^*] - r_f) + \mu_H(E[y|y \geq y^*] - r_f) \right] \\
&\quad + \frac{(1-\rho)\theta_L\bar{x}}{1+r_f} (E[y] - r_f) \\
&= \frac{\rho\theta_H\bar{x}}{1+r_f} \left[(1-\mu_H)(E[y|y < y^*] - r_f) + \mu_H(E[y|y \geq y^*] - E[y]) \right] \\
&\quad + \frac{\bar{x}}{1+r_f} \left[\rho\mu_H\theta_H + (1-\rho)\theta_L \right] (E[y] - r_f).
\end{aligned}$$

For any $w_A < \bar{w}_A$, we can calculate the difference between $W(\bar{w}_A)$ and $W(w_A)$ as follows:

$$W(\bar{w}_A) - W(w_A) = A_0(w_A) + A_1(w_A), \quad (\text{B.17})$$

where

$$\begin{aligned}
A_0(w_A) &\equiv \frac{\rho(1-\mu_H)\theta_H\bar{x}}{1+r_f} (E[y|y < y^*] - r_f) + \rho\mu_H\bar{x} \left(\frac{\theta_H}{1+r_f} - E[p|\theta_H] \right) (E[y|y \geq y^*] - E[y]); \\
A_1(w_A) &\equiv \bar{x} \left[\rho\mu_H \left(\frac{\theta_H}{1+r_f} - E[p|\theta_H] \right) + (1-\rho) \left(\frac{\theta_L}{1+r_f} - E[p|\theta_L] \right) \right] (E[y] - r_f).
\end{aligned}$$

In the following lemma, we show that the second term, $A_1(w_A)$, is zero for any w_A .

LEMMA B.9: *For any w_A , we have*

$$\rho\mu_H \left(\frac{\theta_H}{1+r_f} - E[p|\theta_H] \right) + (1-\rho) \left(\frac{\theta_L}{1+r_f} - E[p|\theta_L] \right) = 0. \quad (\text{B.18})$$

Proof: Equation (B.18) can be represented as

$$\hat{\rho} \int_0^{1-\Delta X/\bar{z}} (1 - q(\theta_H, \eta; \hat{\rho})) f_\eta(\eta) d\eta - (1 - \hat{\rho}) \int_{\Delta X/\bar{z}}^1 q(\theta_L, \eta; \hat{\rho}) f_\eta(\eta) d\eta = 0, \quad (\text{B.19})$$

where

$$q(\theta_H, \eta; \hat{\rho}) = \frac{\hat{\rho}f_\eta(\eta)}{\hat{\rho}f_\eta(\eta) + (1 - \hat{\rho})f_\eta(\eta + \Delta X/\bar{z})};$$

$$q(\theta_L, \eta; \hat{\rho}) = \frac{\hat{\rho}f_\eta(\eta - \Delta X/\bar{z})}{\hat{\rho}f_\eta(\eta - \Delta X/\bar{z}) + (1 - \hat{\rho})f_\eta(\eta)}.$$

Using integration by substitution, we can again rewrite equation (B.19) as

$$\hat{\rho} \int_0^{1-\Delta X/\bar{z}} \left[(1 - q(\theta_H, \eta; \hat{\rho}))f_\eta(\eta) - (1 - \hat{\rho})q(\theta_H, \eta; \hat{\rho})f_\eta(\eta + \Delta X/\bar{z}) \right] d\eta = 0. \quad (\text{B.20})$$

Because $\hat{\rho}(1 - q(\theta_H, \eta; \hat{\rho}))f_\eta(\eta) = (1 - \hat{\rho})q(\theta_H, \eta; \hat{\rho})f_\eta(\eta + \Delta X/\bar{z})$, it is immediate that equation (B.20) is true, which in turn implies that equation (B.19) is true. Q.E.D.

It is easy to check that the first term of equation (B.17), $A_0(w_A)$, is positive if $w_A < \bar{w}_A$, and zero otherwise. Therefore, Lemma B.9 implies that $W(\bar{w}_A)$ is strictly greater than $W(w_A)$ when $w_A < \bar{w}_A$, and is equal to $W(w_A)$ when $w_A \geq \bar{w}_A$. Q.E.D.

(ii) (Risk sharing): For notational convenience, we set w_H equal to zero without loss of generality (because the hedgers' maximization problem is unaffected by it). Given w_A , the ex-ante expected utility of a hedger can be represented as

$$U^H(w_A) = -\frac{1}{\gamma} \left[1 - E[\eta] + \rho(1 - \mu_H)E[\eta] \exp(\gamma\theta_H\bar{z})\Upsilon \right. \\ \left. + \rho\mu_H \int_0^1 \eta \exp(\gamma(1 + r_f)p(\theta_H, \eta; \hat{\rho})\bar{z}) f_\eta(\eta) d\eta \right. \\ \left. + (1 - \rho) \int_0^1 \eta \exp(\gamma(1 + r_f)p(\theta_L, \eta; \hat{\rho})\bar{z}) f_\eta(\eta) d\eta \right],$$

where

$$\Upsilon \equiv \frac{1}{2} \left(\exp(\gamma\bar{e}\bar{z}) + \exp(-\gamma\bar{e}\bar{z}) \right).$$

For any $w_A \geq \bar{w}_A$, there is no fire sale in any state of the world, thus, we have $U^H(w_A) = U^H(\bar{w}_A)$ where

$$U^H(\bar{w}_A) = -\frac{1}{\gamma} \left[1 - E[\eta] + E[\eta] \left(\rho \exp(\gamma\theta_H\bar{z}) + (1 - \rho) \exp(\gamma\theta_L\bar{z}) \right) \right].$$

We claim that $U^H(\bar{w}_A)$ is the highest value which $U^H(w_A)$ can take for any $w_A \in [0, \infty)$. We let ΔU^H be the difference between $U^H(\bar{w}_A)$ and $U^H(w_A)$, that is,

$$\Delta U^H(w_A) \equiv U^H(\bar{w}_A) - U^H(w_A) = U_1(w_A) + U_2(w_A), \quad (\text{B.21})$$

where

$$\begin{aligned} U_1(w_A) &\equiv -\frac{1}{\gamma} \rho (1 - \mu_H) E[\eta] \exp(\gamma \theta_H \bar{z}) (1 - \Upsilon); \\ U_2(w_A) &\equiv -\frac{1}{\gamma} \left[\rho \mu_H \int_0^{1 - \Delta X / \bar{z}} \eta \left(\exp(\gamma \theta_H \bar{z}) - \exp(\gamma (1 + r_f) p(\theta_H, \eta; \hat{\rho}) \bar{z}) \right) f_\eta(\eta) d\eta \right. \\ &\quad \left. + (1 - \rho) \int_{\Delta X / \bar{z}}^1 \eta \left(\exp(\gamma \theta_L \bar{z}) - \exp(\gamma (1 + r_f) p(\theta_L, \eta; \hat{\rho}) \bar{z}) \right) f_\eta(\eta) d\eta \right]. \end{aligned}$$

Then, the claim is proven if $\Delta U^H(w_A) > 0$ for all $w_A < \bar{w}_A$.

LEMMA B.10:

$$\lim_{w_A \uparrow \bar{w}_A} \frac{dE[p|\theta_H]}{dw_A} > 0.$$

Proof: Differentiating $E[p|\theta_H]$ with respect to w_A yields

$$\frac{dE[p|\theta_H]}{dw_A} = \frac{\partial E[p|\theta_H]}{\partial w_A} + \frac{\partial E[p|\theta_H]}{\partial \hat{\rho}} \frac{d\hat{\rho}}{dw_A}. \quad (\text{B.22})$$

From equation (B.13), we can easily verify that

$$\lim_{w_A \uparrow \bar{w}_A} \frac{\partial E[p|\theta_H]}{\partial w_A} = \frac{1}{1 + r_f} (\theta_H - \theta_L) (1 - q(\theta_H, 0; \rho)) f_\eta(0) \left(\frac{m^+ + m^-}{m^+ m^- \bar{z}} \right) > 0; \quad (\text{B.23})$$

$$\lim_{w_A \uparrow \bar{w}_A} \frac{\partial E[p|\theta_H]}{\partial \hat{\rho}} = 0. \quad (\text{B.24})$$

Furthermore, we can also verify that $\lim_{w_A \uparrow \bar{w}_A} \frac{d\hat{\rho}}{dw_A}$ is finite. To see this, we let $K(\hat{\rho}) \equiv H(\hat{\rho}) - \hat{\rho}$.⁴⁰ Then, it is clear that $K(\hat{\rho}) = 0$ if the economy is in equilibrium. By the implicit function theorem, we have

$$\frac{d\hat{\rho}}{dw_A} = -\frac{\frac{\partial K}{\partial w_A}}{\frac{\partial K}{\partial \hat{\rho}}}, \quad (\text{B.25})$$

⁴⁰See equation (B.12) for the definition of $H(\cdot)$

where

$$\begin{aligned}\frac{\partial K}{\partial w_A} &= \frac{\rho(1-\rho)f_y\left(\frac{\theta_H}{E[p(\theta,\eta;\hat{\rho})|\theta_H]} - 1\right)}{\left[\rho\left(1 - F_y\left(\frac{\theta_H}{E[p(\theta,\eta;\hat{\rho})|\theta_H]} - 1\right)\right) + (1-\rho)\right]^2} \frac{\theta_H}{E[p|\theta_H]^2} \frac{\partial E[p|\theta_H]}{\partial w_A}, \\ \frac{\partial K}{\partial \hat{\rho}} &= \frac{\rho(1-\rho)f_y\left(\frac{\theta_H}{E[p(\theta,\eta;\hat{\rho})|\theta_H]} - 1\right)}{\left[\rho\left(1 - F_y\left(\frac{\theta_H}{E[p(\theta,\eta;\hat{\rho})|\theta_H]} - 1\right)\right) + (1-\rho)\right]^2} \frac{\theta_H}{E[p|\theta_H]^2} \frac{\partial E[p|\theta_H]}{\partial \hat{\rho}} - 1.\end{aligned}$$

Using equations (B.23), (B.24) and (B.25), we can show that $\lim_{w_A \uparrow \bar{w}_A} \frac{d\hat{\rho}}{dw_A}$ is finite and positive. Then, the result is immediate from equation (B.22). Q.E.D.

In the following lemma, we prove that $\Delta U^H(w_A)$ is positive for any w_A arbitrarily close to \bar{w}_A .

LEMMA B.11: *There exists \underline{w}_A such that $\Delta U^H(w_A) > 0$ for all $[w_A, \bar{w}_A)$.*

Proof: Differentiating $U_1(w_A)$ with respect to w_A yields

$$\frac{dU_1}{dw_A} = \frac{\rho}{\gamma} \frac{d\mu_H}{dw_A} E[\eta] \exp(\gamma\theta_H\bar{z})(1 - \Upsilon),$$

where

$$\frac{d\mu_H}{dw_A} = f_y\left(\frac{\theta_H}{E[p(\theta,\eta;\hat{\rho})|\theta_H]} - 1\right) \frac{\theta_H}{E[p|\theta_H]^2} \frac{dE[p|\theta_H]}{dw_A}.$$

Lemma B.10 implies $\lim_{w_A \uparrow \bar{w}_A} \frac{d\mu_H}{dw_A} > 0$, which in turn implies $\lim_{w_A \uparrow \bar{w}_A} \frac{dU_1}{dw_A}$ is strictly negative because $\Upsilon > 1$. Similarly, it can also be shown that

$$\lim_{w_A \uparrow \bar{w}_A} \frac{dU_2}{dw_A} = \frac{1-\rho}{\gamma} \left(\frac{m^+ + m^-}{m^+ m^- \bar{z}} \right) \left(\exp(\gamma\theta_L\bar{z}) - \exp(\gamma(1+r_f)p(\theta_L, 1; \rho)\bar{z}) \right) f_\eta(1) < 0.$$

Therefore, $\Delta U^H(w_A)$ strictly decreases in w_A arbitrarily close to \bar{w}_A . Because $\Delta U^H(\bar{w}_A)$ is equal to zero by construction, the continuity of $\Delta U^H(\cdot)$ implies it is positive for any w_A arbitrarily close to \bar{w}_A . Q.E.D.

From the definition of $U_2(w_A)$ in equation (B.21), it is easy to see that $U_2(w_A)$ is

bounded below; for any w_A , we have

$$U_2(w_A) \geq -\frac{1}{\gamma} \rho E[\eta] \left(\exp(\gamma \theta_H \bar{z}) - \exp(\gamma \theta_L \bar{z}) \right).$$

At any given level of $w_A < \bar{w}_A$, however, $U_1(w_A)$ strictly increases in $\bar{\epsilon}$ in an unbounded manner. Therefore, there exists a constant $\bar{\epsilon}^*$ such that, for any $\bar{\epsilon} \geq \bar{\epsilon}^*$, $\Delta U^H(w_A)$ is positive for all $w_A \leq \underline{w}_A$. From Lemma B.11, $\Delta U^H(w_A)$ is also positive for any $w_A \in [\underline{w}_A, \bar{w}_A]$. Consequently, given $\bar{\epsilon} \geq \bar{\epsilon}^*$, $\Delta U^H(w_A)$ is zero only when $w_A \geq \bar{w}_A$, and positive otherwise. Q.E.D.

Proof of Lemma 5: We start with the long margin m^+ . Suppose arbitrageur a (or any margin trader) takes $x^{a,+}$ unit of long positions by pledging w_A as the initial margin. The maximum loss occurs when the marketable asset pays the lowest possible payoff, $\theta_L - \bar{\epsilon}$. For the margin loan to be risk-free, the final balance for the arbitrageur's margin trading should be non-negative. It is because any loss should be covered by the initial margin in that case. Given $x^{a,+}$, the final balance for arbitrageur a 's margin trading is non-negative if and only if

$$(\theta_L - \bar{\epsilon} - p(1 + r_f))x^{a,+} + (1 + r_f)w_A \geq 0, \tag{B.26}$$

where $(\theta_L - \bar{\epsilon} - p(1 + r_f))x^{a,+}$ is the maximum possible trading loss, and $(1 + r_f)w_A$ is the capital in the margin account at $t = 2$. When the arbitrageur chooses the maximum possible long position (i.e., $x^{a,+} = \frac{w_A}{m^+}$) given his initial capital of w_A and the long margin of m^+ , equality holds in equation (B.26) as follows:

$$(\theta_L - \bar{\epsilon} - p(1 + r_f)) \frac{w_A}{m^+} + (1 + r_f)w_A = 0,$$

whose solution for m^+ is given by

$$m^+ = p - \frac{\theta_L - \bar{\epsilon}}{1 + r_f}.$$

Now, we turn to the short margin m^- . Suppose arbitrageur a takes $x^{a,-}$ unit of short positions by pledging w_A as the initial margin. The maximum loss occurs when the marketable asset pays the highest possible payoff, $\theta_H + \bar{\epsilon}$. Given $x^{a,-}$, the final

balance for arbitrageur a 's margin trading is non-negative if and only if

$$(p(1 + r_f) - \theta_H - \bar{\epsilon})x^{a,-} + (1 + r_f)w_A \geq 0. \quad (\text{B.27})$$

Similarly as in the long margin, we can obtain the following using equation (B.27):

$$m^- = \frac{\theta_H + \bar{\epsilon}}{1 + r_f} - p.$$

Q.E.D.

Proof of Proposition 4: We define \bar{w}_A^{en} to be the fire-sale-free level of arbitrage capital under the endogenous margin constraint given residual uncertainty parameter $\bar{\epsilon}$, and also define \bar{w}_A^{ex} to be the fire-sale-free level of arbitrage capital under the exogenous margin constraint given margin parameters m^+, m^- . Then, it is easy to show that

$$\begin{aligned} \bar{w}_A^{en} &= \frac{(1 + r_f)(\theta_H - \theta_L + \bar{\epsilon})\bar{\epsilon}\bar{z}}{\theta_H - \theta_L + 2\bar{\epsilon}}; \\ \bar{w}_A^{ex} &= \bar{z} \left(\frac{m^+ m^-}{m^+ + m^-} \right). \end{aligned}$$

Let δ be the fraction of decrease in arbitrage capital from the fire-sale-free level, that is, $\delta \equiv (\bar{w}_A - w_A)/\bar{w}_A$ for any $w_A \leq \bar{w}_A$. When there is a δ fraction of decrease in arbitrage capital, the arbitrage trading variation under the endogenous and the exogenous margin constraint is, respectively, given by

$$\begin{aligned} \Delta X^{en} &= (1 - \delta)\bar{z} \left(\frac{(\theta_H - \theta_L + \bar{\epsilon})\bar{\epsilon}}{(q(\theta_H - \theta_L) + \bar{\epsilon})((1 - q)(\theta_H - \theta_L) + \bar{\epsilon})} \right); \\ \Delta X^{ex} &= (1 - \delta)\bar{z}. \end{aligned}$$

Notice that $\Delta X^{en} \leq \Delta X^{ex}$, and also that $\Delta X^{en} \rightarrow \Delta X^{ex}$ as $\bar{\epsilon} \rightarrow \infty$. Due to Proposition 1, the equilibrium prices under the endogenous and the exogenous margin constraints should become asymptotically identical as their arbitrage trading variations become arbitrarily close. This in turn implies that all the equilibrium quantities such as the quality of traded asset, the circulation rate of the high-quality asset, etc should be also arbitrarily close. Q.E.D.

REFERENCES

- Akerlof, George A., 1970, The market for “lemons”: Quality uncertainty and the market mechanism, *Quarterly Journal of Economics* 84, 488–500.
- Allen, Franklin, and Douglas Gale, 1994, Limited market participation and volatility of asset prices, *American Economic Review* 84, 933–955.
- Allen, Franklin, and Douglas Gale, 2004, Financial intermediaries and markets, *Econometrica* 72, 1023–1061.
- Biais, Bruno, Florian Heider, and Marie Hoerova, 2016, Risk-sharing or risk-taking? counterparty risk, incentives, and margins, *Journal of Finance* 71, 1669–1698.
- Bolton, Patrick, Tano Santos, and Jose A. Scheinkman, 2009, Market and public liquidity, *American Economic Review: Papers & Proceedings* 99, 594–599.
- Brunnermeier, Markus K, 2009, Deciphering the liquidity and credit crunch 2007-2008, *Journal of Economic Perspective* 23, 77–100.
- Brunnermeier, Markus K., and Lasse Heje Pedersen, 2009, Market liquidity and funding liquidity, *Review of Financial Studies* 22, 2201–2199.
- Caballero, Ricardo J, and Alp Simsek, 2013, Fire sales in a model of complexity, *Journal of Finance* 68, 2549–2587.
- Coval, Joshua, and Erik Stafford, 2007, Asset fire sales (and purchases) in equity markets, *Journal of Financial Economics* 86, 479–512.
- Daley, Brendan, and Brett Green, 2012, Waiting for news in the market for lemons, *Econometrica* 80, 1433–1504.
- Dang, Tri Vi, Gary B Gorton, and Bengt Holmstrom, 2012, Ignorance, debt and financial crises, Working paper.
- Diamond, Douglas W., and Philip H. Dybvig, 1983, Bank runs, deposit insurance, and liquidity, *Journal of Political Economy* 91, 401–419.
- Diamond, Douglas W., and Raghuram G. Rajan, 2011, Fear of fire sales, illiquidity seeking, and credit freezes, *Quarterly Journal of Economics* 126, 557–591.

- Dow, James, and Jungsuk Han, 2015, Contractual incompleteness, limited liability and asset price bubbles, *Journal of Financial Economics* 116, 383–409.
- Dow, James, and Rohit Rahi, 2000, Should speculators be taxed?, *Journal of Business* 73, 89–107.
- Dow, James, and Rohit Rahi, 2003, Informed trading, investment and welfare, *Journal of Business* 76, 439–454.
- Duarte, Fernando, and Thomas M. Eisenbach, 2014, Fire-sale spillovers and systemic risk, Working paper.
- Duffie, Darrell, 2010, Asset price dynamics with slow-moving capital, *Journal of Finance* 65, 1237–1267.
- Eisfeldt, Andrea L, 2004, Endogenous liquidity in asset markets, *Journal of Finance* 59, 1–30.
- Ellul, Andrew, Chotibhak Jotikasthira, and Christian Lundblad, 2011, Fear of fire sales, illiquidity seeking, and credit freezes, *Journal of Financial Economics* 101, 596–620.
- Fishman, Michael J., and Jonathan A. Parker, 2015, Valuation, adverse selection, and market collapses, *Review of Financial Studies* 28, 2575–2607.
- Geanakoplos, John, 2003, Liquidity, default, and crashes: Endogenous contracts in general equilibrium, in Mathias Dewatripont, Lars Peter Hansen, and Stephen J. Turnovsky, eds., *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Volume 2* (Cambridge University Press, Cambridge).
- Glode, Vicent, Richard C Green, and Richard Lowery, 2012, Financial expertise as an arms race, *Journal of Finance* 67, 1723–1759.
- Gorton, Gary, and Andrew Metrick, 2012, Securitized banking and the run on repo, *Journal of Financial Economics* 104, 425–451.
- Greenwood, Robin, Augustin Landier, and David Thesmar, 2015, Vulnerable banks, *Journal of Financial Economics* 115, 471–485.
- Gromb, Denis, and Dimitri Vayanos, 2002, Equilibrium and welfare in markets with financially constrained arbitrageurs, *Journal of Financial Economics* 66, 361–407.

- Grossman, Sanford J., and Joseph E. Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393–408.
- Guerrieri, Veronica, and Robert Shimer, 2014, Dynamic adverse selection: A theory of illiquidity, fire sales, and flight to quality, *American Economic Review* 104, 1875–1908.
- He, Zhiguo, and Arvind Krishnamurthy, 2011, A model of capital and crises, *Review of Economic Studies* 79, 735–777.
- He, Zhiguo, and Arvind Krishnamurthy, 2013, Intermediary asset pricing, *American Economic Review* 103, 732–770.
- Hirshleifer, Jack, 1971, The private and social value of information and the reward to inventive activity, *American Economic Review* 61, 561–574.
- Huang, Sheng, Matthew C. Ringgenberg, and Zhe Zhang, 2016, The information in fire sales, WFA - Center for Finance and Accounting Research Working Paper No. 15/07.
- Ivashina, Victoria, and David Scharfstein, 2010, Bank lending during the financial crisis of 2008, *American Economic Review* 97, 319–338.
- Jotikasthira, Chotibhak, Christian Lundblad, and Tarun Ramadorai, 2012, Asset fire sales and purchases and the international transmission of funding shocks, *Journal of Finance* 67, 2015–2050.
- Krishnamurthy, Arvind, 2010, Amplification mechanisms in liquidity crises, *American Economic Journal: Macroeconomics* 3, 1–30.
- Kuong, John Chi-Fong, 2015, Self-fulfilling fire sales: Fragility of collateralised short-term debt markets, Working paper.
- Kurlat, Pablo, 2013, Lemons markets and the transmission of aggregate shocks, *American Economic Review* 103, 1463–1489.
- Kyle, Albert S., 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–1335.
- Malherbe, Frederic, 2014, Self-fulfilling liquidity dry-ups, *Journal of Finance* 69, 947–970.

- Merrill, Craig B., Taylor Nadauld, Rene M. Stulz, and Shane M. Sherlund, 2014, Were there fire sales in the RMBS market?, Working paper.
- Mitchell, Mark, Lasse Heje Pedersen, and Todd Pulvino, 2007, Slow moving capital, *American Economic Review* 97, 215–220.
- Myers, Stewart C., and Nicholas S. Majluf, 1984, Corporate financing and investment decisions when firms have information that investors do not have, *Journal of Financial Economics* 13, 187–221.
- Perotti, Enrico, and Javier Suarez, 2011, A pigouvian approach to liquidity regulation, *International Journal of Central Banking* 7, 3–41.
- Pulvino, Todd C., 1998, Do asset fire sales exist? an empirical investigation of commercial aircraft transactions, *Journal of Finance* 53, 939–978.
- Schlingemann, Frederik P., René M. Stulz, and Ralph A. Walkling, 2002, Divestitures and the liquidity of the market for corporate assets, *Journal of Financial Economics* 64, 117–144.
- Shleifer, Andrei, and Robert W. Vishny, 1992, Liquidation values and debt capacity: A market equilibrium approach, *Journal of Finance* 47, 1343–1366.
- Shleifer, Andrei, and Robert W. Vishny, 1997, The limits of arbitrage, *Journal of Finance* 52, 35–55.
- Shleifer, Andrei, and Robert W. Vishny, 2010, Asset fire sales and credit easing, *American Economic Review: Papers & Proceedings* 100, 46–50.
- Shleifer, Andrei, and Robert W. Vishny, 2011, Fire sales in finance and macroeconomics, *Journal of Economic Perspectives* 25, 29–48.
- Tirole, Jean, 2011, Illiquidity and all its friends, *Journal of Economic Literature* 49, 287–325.
- Tirole, Jean, 2012, Overcoming adverse selection: How public intervention can restore market functioning, *American Economic Review* 102, 29–59.
- Vayanos, Dimitri, and Paul Woolley, 2013, An institutional theory of momentum and reversal, *Review of Financial Studies* 26, 1087–1145.

Vickery, James, and Joshua Wright, 2013, Tba trading and liquidity in the agency mbs market, *Economic Policy Review* 19, 1–18.

Vives, Xavier, 1995, The speed of information revelation in a financial market mechanism, *Journal of Economic Theory* 67, 178–204.

Wilson, Charles A., 1979, Equilibrium and adverse selection, *American Economic Review: Papers & Proceedings* 69, 313–317.

Wilson, Charles A., 1980, The nature of equilibrium in markets with adverse selection, *The Bell Journal of Economics* 11, 108–130.