

A Model of Capital and Crises*

Zhiguo He

Arvind Krishnamurthy

August 2008

Abstract

We develop a model in which the capital of the intermediary sector plays a critical role in determining asset prices. The model is cast within a dynamic general equilibrium economy, and the role for intermediation is derived endogenously based on optimal contracting considerations. Low intermediary capital reduces the risk-bearing capacity of the marginal investor. We show how this force helps to explain patterns during financial crises. The model replicates the observed rise during crises in Sharpe ratios, conditional volatility, correlation in price movements of assets held by the intermediary sector, and fall in riskless interest rates. In a dynamic context, we show that aversion to drops in intermediary capital can generate a two-factor asset pricing model with a role for both a market factor and a liquidity factor.

JEL Codes: G12, G2, E44

Keywords: Liquidity, Hedge Funds, Delegation, Financial Institutions.

*University of Chicago, Graduate School of Business, zhiguo.he@chicagogsb.edu; Northwestern University, Kellogg School of Management and NBER, a-krishnamurthy@northwestern.edu. We thank Andrea Eisfeldt, Oleg Bondarenko, Darrell Duffie, Mike Fishman, Ravi Jagannathan, Haitao Li, John Moore, Andrea Prat, Dimitri Vayanos, Adrien Verdelhan, Wei Xiong and participants at seminars at Northwestern, LSE, UIC, NBER Summer Institute, CICF 2006, and SITE 2006 for helpful comments. A previous version of this paper was titled “Intermediation, Capital Immobility, and Asset Prices.”

1 Introduction

Financial crises, such as the hedge fund crisis of 1998 or the 2007/2008 subprime crisis, have several common characteristics: risk premia rise, interest rates fall, conditional volatilities of asset prices rise, correlations between assets rise, and investors “fly to the quality” of a riskless liquid bond. This paper offers an account of a financial crisis in which intermediaries play the central role. Intermediaries are the marginal investors in our model. The crisis occurs because shocks to the capital of intermediaries reduce their risk-bearing capacity, leading to a dynamic that replicates each of the afore-mentioned regularities. These results are developed within a dynamic general equilibrium model with a contractual micro-foundation for intermediation.

The intermediation model also offers insights into financial behavior outside of crises. A number of recent papers have documented the existence of a priced liquidity risk factor (see Amihud, 2002; Acharya and Pedersen, 2005; Pastor and Stambaugh, 2003; and Sadka, 2006). That is, these papers show that assets whose payoffs are low during times of marketwide illiquidity carry high ex-ante risk premia. The financial crisis of our model can be readily viewed as an illiquidity episode. We show that intermediaries, who are central to the dynamics of a financial crisis, will demand assets that help them hedge against a financial crisis. This hedging behavior, since the intermediaries are also marginal in pricing assets, leads to a priced liquidity risk factor.

Our paper makes two principal contributions: (1) we show that modeling intermediaries can help to explain a collection of asset market facts both inside and outside of financial crises; and, (2) we offer a model of intermediation and crises that is fully dynamic and less stylistic than some of the existing models in the literature.¹

There is a large literature on intermediation and asset pricing, ranging from banking models to models of portfolio delegation.² Our paper is closest to the banking models in that we emphasize capital effects. Allen and Gale (1994) present a model in which the amount of “cash” of the marginal investors affects asset prices. This “cash-in-the-market” can be linked to the balance sheet position of intermediaries, and Allen and Gale (2005) draw such a connection more explicitly. Holmstrom and Tirole (1997) present a model in which there is a role for intermediary capital, and

¹In a companion paper (He and Krishnamurthy, 2008), we develop this second point by incorporating additional realistic features into the model so that it can be calibrated. We show that the calibrated model can quantitatively match crisis and non-crisis asset market behavior.

²Some of the papers in the literature include Allen and Gorton (1993), Brennan (1993), Dow and Gorton (1994), Grossman and Zhou (1996), Shleifer and Vishny (1997), Dasgupta, Prat, and Verardo (2008), Brunnermeier and Pedersen (2008), and Guerrieri and Kondor (2008).

changes in this capital affect asset prices (the interest rate in Holmstrom and Tirole). The models in these papers are stylized one or two period models, which we go beyond. The micro-foundation for intermediation in our model draws from the Holmstrom and Tirole model.

Xiong (2001), Kyle and Xiong (2001), and Vayanos (2005) develop dynamic models to study crises and illiquidity. Both Xiong (2001) and Kyle and Xiong (2001) papers model a capital effect for asset prices and show that this effect can help to explain some of the crises regularities we have noted. These papers model an “arbitrageur” sector using a shorthand log utility assumption. In contrast, we develop a role for intermediation within the model, derive the constraints endogenously from an explicit principal-agent problem, and are thereby better able to articulate the part of intermediaries in crises.³ These models also do not speak to the issue of liquidity risk. Vayanos (2005) more explicitly models intermediation and offers an explanation for the pricing of liquidity risk. His model generates a risk premium on a volatility factor, which he argues may be what the empirical studies on liquidity risk are picking up. Vayanos’ model introduces an open-ending friction, rather than a capital friction, into a model of intermediation.⁴

Empirically, the evidence for an intermediation capital effect comes in two forms. First, by now it is widely accepted that the fall of 1998 crisis was due to negative shocks to the capital of intermediaries (hedge funds, market makers, trading desks, etc.). These shocks led intermediaries to liquidate positions, which lowered asset prices, further weakening intermediary balance sheets.⁵ Similar capital-related phenomena have been noted in the 1987 stock-market crash (Mitchell, Pedersen, and Pulvino, 2007), the mortgage-backed securities market following an unexpected prepayment wave in 1994 (Gabaix, Krishnamurthy, and Vigneron, 2006), as well the corporate bond market following the Enron default (Berndt, et al., 2004). Froot and O’Connell (1999), and Froot (2001) present evidence that the insurance cycle in the catastrophe insurance market is due to fluctuations in the capital of reinsurers. Duffie (2007) discusses some of these cases in the context of search costs

³The same distinction exists between our paper and Pavlova and Rigobon (2008), who study a model with log-utility agents facing exogenous portfolio constraints and use the model to explore some regularities in exchange rates and international financial crises. Like us, their model shows how contagion and amplification can arise endogenously. While their application to international financial crises differs from our model, at a deeper level the models are related.

⁴Gromb and Vayanos (2002) and Liu and Longstaff (2004) study settings in which an arbitrageur with limited wealth and facing a capital constraint trades to exploit a high Sharpe-ratio investment. Liu and Longstaff show that the capital constraint can substantially affect the arbitrageur’s optimal trading strategy. Gromb and Vayanos show that the capital constraints can have important asset pricing effects. Both of these papers point to the importance of a capital effect for asset pricing.

⁵Other important asset markets, such as the equity or housing market, were relatively unaffected by the turmoil. The dichotomous behavior of asset markets suggests that the problem was hedge fund capital specifically, and not capital more generally. Investors did not bypass the distressed hedge funds in a way as to undo any asset price impact of the hedge fund actions. They also did not restore the hedge funds’ capital.

and slow movement of capital into the affected intermediated markets. One of the motivations for our paper is to reproduce asset market behavior during crisis episodes.

Although the crisis evidence is dramatic, crisis episodes are rare and do not lend themselves to systematic study. The second form of evidence for the existence of intermediation capital effects come from studies examining the cross-sectional/time-series behavior of asset prices within a particular asset market. Gabaix, Krishnamurthy, and Vigneron (2006) study a cross-section of prices in the mortgage-backed securities market and present evidence that the marginal investor who prices these assets is a specialized intermediary rather than a CAPM-type representative investor. Similar evidence has been provided for index options (Bates, 2003; Garleanu, Pedersen, and Poteshman, 2005), and corporate bonds and default swaps (Collin-Dufresne, Goldstein, and Martin, 2001; Berndt, et al., 2004). These studies are particularly good motivation for our model because the markets they consider tend to be ones dominated by intermediaries. Thus they reiterate the relevance of intermediation capital for asset prices.

This paper is laid out as follows. Section 2 describes the model and derives the capital constraint based on agency considerations. Section 3 solves for asset prices in closed form, and studies the implications of intermediation capital on asset pricing. Section 4 explains the parameter choices in our numerical examples. Section 5 concludes.

2 The Model

2.1 Agents and Assets

We consider an infinite-horizon, continuous-time, economy with a single perishable consumption good, along the lines of Lucas (1978). We use the consumption good as the numeraire. There are two assets, a riskless bond in zero net supply, and a risky asset that pays a risky dividend. We normalize the total supply of the risky asset to be one unit.

The risky asset pays a dividend of D_t per unit time, where $\{D_t : 0 \leq t < \infty\}$ follows a geometric Brownian motion,

$$\frac{dD_t}{D_t} = gdt + \sigma dZ_t \quad \text{given } D_0; \tag{1}$$

and $g > 0$ and $\sigma > 0$ are constants. Throughout this paper $\{Z\} = \{Z_t : 0 \leq t < \infty\}$ is a standard Brownian motion on a complete probability space $(\Omega, \mathcal{F}, \mathcal{P})$ with an augmented filtration $\{\mathcal{F}_t : 0 \leq t < \infty\}$ generated by the Brownian motion $\{Z\}$.

We denote the progressively measurable processes $\{P_t : 0 \leq t < \infty\}$ and $\{r_t : 0 \leq t < \infty\}$ as the

risky asset price and interest rate processes, respectively. They will be determined in equilibrium.

There are two classes of agents in the economy, households and specialists. Without loss of generality, we set the measure of each agent class to be one. We are interested in studying an intermediation relationship between households and specialists. To this end, we assume that the risky asset payoff comprises a set of complex investment strategies that the specialist has a comparative advantage in managing, and therefore intermediates the households' investments into the risky asset. Throughout this paper we will think of the dividend process from the risky asset as corresponding to a representative "intermediated" asset. This asset is an amalgam of payoffs from mortgage-backed securities investments, emerging-market investments, investments in long-short liquidity provision strategies, etc. In particular, the risky asset should not be thought of as the S&P 500 stock index.

As in the literature on limited market participation (e.g., Mankiw and Zeldes, 1991; Allen and Gale, 1994; Basak and Cuoco, 1998; and Vissing-Jorgensen, 2002), we make the extreme assumption that the household cannot directly invest in the risky asset and can directly invest only in the bond market. Following the limited participation literature, we motivate this assumption by appealing to "informational" transaction costs that households face in order to invest directly in the risky asset market.

We depart from the limited participation literature by allowing specialists to invest in the risky asset on behalf of the households. Households allocate some funds to intermediaries that are run by specialists. We can think of an intermediary as a hedge fund or bank investing in mortgage-backed securities or emerging markets' sovereign bonds. The specialist plays the role of insider/manager of the intermediary.

Both specialists and households are infinitely lived and have log preferences over date t consumption. Denote c_t (c_t^h) as the specialist's (household's) consumption rate. The specialist maximizes

$$E_0 \left[\int_0^\infty e^{-\rho t} \ln c_t dt \right],$$

while the household maximizes

$$E_0 \left[\int_0^\infty e^{-\rho^h t} \ln c_t^h dt \right],$$

where the positive constants ρ and ρ^h are the specialist's and household's time-discount rates, respectively. Throughout we use the superscript "h" to indicate households. Note that ρ may differ from ρ^h ; this flexibility is useful when specifying the boundary condition for the economy.

2.2 Intermediaries and Intermediation Contract

At every t , households invest in a continuum of intermediaries that are run by specialists. As detailed in Section 2.5, the market for intermediation is competitive with specialists providing intermediation services, while households purchasing these services. We will think of an intermediary as being invested in by a continuum of identical households, although for ease of exposition we sometimes describe the contracting as between a representative specialist and household.

After the time- t intermediation decisions are taken by specialists and households, the specialists trade in a Walrasian stock and bond market on behalf of the intermediaries, and the households trade in only the bond market. The intermediation relation is short-term, and at $t + dt$ the intermediation market repeats itself.

Consider one of these intermediaries. It is run by the specialist who makes all of the investment decisions. Absent proper incentives, the specialist will shirk some of his investment tasks in order to enjoy a private benefit. Thus, there is a moral hazard problem that must be alleviated by writing a financial contract between specialist and household.

The household is the principal in this relationship and the specialist is the agent. A financial contract dictates how much funds each party contributes to the intermediary, and how much each party is paid as a function of realized returns.

Consider a specialist with wealth W and a household with wealth W^h . In equilibrium, these wealth levels evolve endogenously. To save notation, we are omitting time subscripts on these wealth levels.

The specialist contributes $T \in [0, W]$ into the intermediary. We focus on contracts in which any remaining specialist wealth $W - T$ earns the riskless interest rate of r_t . This restriction is similar to, but weaker than, the usual one of no private savings by the agent.⁶ The household contributes $T^h \in [0, W^h]$ into the intermediary. We refer to $T^I = T + T^h$ as the total capital of the intermediary.

The intermediary is run by the specialist. We formalize the moral hazard problem by assuming that the specialist makes an *unobserved portfolio choice decision* and an *unobserved due-diligence decision of “shirking” or “working.”* For any given portfolio choice, if the specialist shirks, the return on the portfolio falls by $x dt$, but the specialist gets a private benefit (in units of the consumption good) of $bT^I dt$, where $x > b > 0$ can be state-dependent, e.g., increasing with risk

⁶This assumption can be relaxed further. Our analysis goes through as long as the specialist cannot short stocks through his personal account. Given the moral hazard issue, this assumption seems reasonable.

premia.⁷ Throughout we assume that it is always optimal for households to write a contract that implements working from the specialist.

We denote \mathcal{E}^I as the intermediary's portfolio choice (a decision made by the specialist) measured in units of money invested in the risky asset. In other words, \mathcal{E}^I is the intermediary's dollar exposure in the risky asset. If the specialist works, the intermediary's total dollar return as a function of the asset position \mathcal{E}^I is,

$$T^I \widetilde{dR}_t(\mathcal{E}^I) = \mathcal{E}^I (dR_t - r_t dt) + T^I r_t dt, \quad (2)$$

where dR_t is the return on the risky asset (specified in equation (9) of the next section). Note that when $\mathcal{E}^I > T^I$, the intermediary is shorting the bond (or borrowing) in the Walrasian bond market.

At the end of the intermediation relationship, the fund is liquidated and each party gets paid based on the contract terms and the return on the fund. We denote $\beta \in [0, 1]$ as the share of returns that goes to the specialist, and $1 - \beta$ as the share to the household. The specialist may also be paid a fee of $\hat{K} dt$ to manage the intermediary. Note that since our model is set in continuous time and there is only one source of risk, it follows from spanning arguments that focusing on linear-share/fixed-fee contract is not a substantive restriction. Any nonlinear contract looks like an affine contract in this setting. The substantive restriction imposed by our analysis is that we do not consider contracts where one specialist's performance is benchmarked to another's.⁸

The household offers a contract $\Pi = (T, T^h, \beta, \hat{K}) \in [0, W] \times [0, W^h] \times [0, 1] \times \mathbb{R}$ to the specialist. Given the contract Π , the specialist makes three decisions: (1) whether to participate in the contract or not; (2) portfolio choice \mathcal{E}^I ; (3) shirk/work. The household must design the contract Π in consideration of these specialist decisions.

2.3 Reducing the Problem

We first reduce the contracting space. Let us write the dynamic budget constraints for both specialist and household:

$$dW = \beta T^I \widetilde{dR}_t(\mathcal{E}^I) + (W - T) r_t dt + \hat{K} dt,$$

and,

$$dW^h = (1 - \beta) T^I \widetilde{dR}_t(\mathcal{E}^I) + (W^h - T^h) r_t dt - \hat{K} dt.$$

⁷We think of shirking as failing to execute trades in an efficient manner. If one specialist shirks and his portfolio return falls by $x dt$, the other investors in the risky asset collectively gain $x dt$. Since each specialist is infinitesimal, the other specialists' gain is infinitesimal. Shirking only leads to transfers and not a change in the aggregate endowment.

⁸This form of contractual incompleteness is often present in macroeconomic models of credit market frictions. See Krishnamurthy (2003) for further discussion of this point.

Substituting from equation (2) we rewrite these equations as,

$$\begin{cases} dW = \beta \mathcal{E}^I (dR_t - r_t dt) + (\beta T^I + W - T) r_t dt + \hat{K} dt, \\ dW^h = (1 - \beta) \mathcal{E}^I (dR_t - r_t dt) + ((1 - \beta) T^I + W^h - T^h) r_t dt - \hat{K} dt. \end{cases}$$

For any given (β, T, T^h) we can define an appropriate $K = (\beta T^I - T) r_t + \hat{K}$ so that these budget constraints become:

$$\begin{cases} dW = \beta \mathcal{E}^I (dR_t - r_t dt) + W r_t dt + K dt, \\ dW^h = (1 - \beta) \mathcal{E}^I (dR_t - r_t dt) + W^h r_t dt - K dt. \end{cases}$$

That is, it is without loss of generality to restrict attention to contracts that only specifies a pair (β, K) .

Reducing the problem in this way highlights the nature of the gains from intermediation in our economy. The specialist offers the household exposure to the excess return on the risky asset, which the household cannot directly achieve due to his limited market participation. This is the first term in the household's budget constraint (i.e., $(1 - \beta) \mathcal{E}^I$). The second term in the household's budget constraint is standard; it is the risk-free interest that the household earns on his wealth. Similar interpretations hold for the specialist's budget equation. The third term is the transfer between the household and the specialist. In Section 2.5, we will come to interpret this transfer as a price that the household pays to the specialist for the intermediation service .

2.4 Incentive Compatibility and Household's Maximum Exposure

We next discuss how varying the contract term β affects the household and the specialist in the intermediary. The unobservability of the portfolio choice decision \mathcal{E} has important implications for our problem. With a slight abuse of notation, denote \mathcal{E}^I as the intermediary's optimal position (chosen by the specialist) in the risky asset given a contract (β, K) , and similarly denote $\mathcal{E}^{I'}$ as the optimal position given a different contract (β', K') . We must have the following relation:

$$\beta \mathcal{E}^I = \beta' \mathcal{E}^{I'} = \mathcal{E}^*,$$

where \mathcal{E}^* is the specialist's optimal exposure in the risky asset from the perspective of his own investment/portfolio problem. This relation, which we refer to as "undoing," implies that the contract terms (β, K) do not have any effect on the specialist's ultimate exposure to the risky asset. The reason is that if β is changed, the specialist adjusts his portfolio choice so that his net exposure of $\beta \mathcal{E}^I$ remains the same.⁹

⁹One may consider whether it is possible to induce the specialist to choose a different portfolio by varying the transfer K . On the cost side, giving the specialist a larger transfer K costs the household in the order of dt . On the

While undoing implies a portfolio exposure for the specialist that does not depend on the contract, it does not imply the same for the household. For any β , the household's post-undoing exposure to the risky asset is,

$$(1 - \beta)\mathcal{E}^I = \frac{1 - \beta}{\beta}\mathcal{E}^*. \quad (3)$$

The household can vary the contract terms, β , to achieve his desired exposure to the risky asset. Setting β to one provides zero exposure to the risky asset, and decreasing β increases the household's exposure to the risky asset.

Incentive compatibility places a limit on how low β can fall. For any total capital T^I and return dR_t , if the specialist shirks, the intermediary's return falls by $x dt$ and the specialist earns a private benefit of $bT^I dt$. For working to be incentive compatible, β must be such that:

$$\beta x T^I dt \geq b T^I dt \quad \Rightarrow \beta \geq \frac{b}{x}. \quad (4)$$

We call the condition (4) the incentive-compatible constraint, and assume that $x > b$ are sufficiently high so that it is always preferable for households to implement working. As the surplus to the household of implementing working rather than shirking depends on the state (e.g., the risk premium), our assumption implicitly requires that x and b may be state-dependent (e.g., increasing with risk premia). For simplicity, we assume that the ratio $\frac{b}{x}$, which plays the central role in the analysis, is a constant.

From (3), the household's risk exposure is simply $\mathcal{E}^h = (1 - \beta)\mathcal{E}^I = \frac{1 - \beta}{\beta}\mathcal{E}^*$. The maximum portfolio exposure by the household is achieved when β is set to the minimum value of $\frac{b}{x}$. Therefore, the maximum exposure is,

$$\frac{1 - \frac{b}{x}}{\frac{b}{x}}\mathcal{E}^* = m\mathcal{E}^*, \quad (5)$$

where we have defined a constant $m \equiv \frac{x}{b} - 1$. The above constraint says that household's exposure to the risky asset (i.e., $(1 - \beta)\mathcal{E}$) is constrained to be less than m times that of the specialist (i.e., \mathcal{E}^*). The inverse of m measures the severity of agency problems. That is, a lower m implies a more severe agency problem and a smaller maximum exposure.

This maximum exposure constraint

$$\mathcal{E}^h \leq m\mathcal{E}^*, \quad (6)$$

benefit side, the difference in the household's portfolio exposure induced by varying K , via changing the specialist's wealth, is of order dt . This implies that any potential gain due to the change in the risky asset exposure will only be in the order of $(dt)^2$. Therefore it is not profitable to affect the exposure through the transfer K .

which is rooted in the specialist’s incentive compatible constraint, is critical for our model. Because of the underlying friction of limited market participation, the households are gaining exposure to the risky asset through intermediaries. However, due to agency considerations, the risk exposure of households, who are considered “outsiders” in the intermediary, must be capped by the maximum exposure $m\mathcal{E}^*$, which is m times that of the specialists’, or “insiders’,” risk exposure.

In our model, households know the specialist’s wealth W , his preferences, and the stochastic processes for asset returns. Therefore, even though they cannot directly observe the specialist’s portfolio choice decisions, they can compute the optimal exposure \mathcal{E}^* of a given specialist. A specialist with a greater \mathcal{E}^* (which we will see to be linear in his wealth W due to log preferences) can offer a greater maximum exposure. Of course, because the specialists are identical in the model, it is true that along the equilibrium path all specialists have the same \mathcal{E}^* at any time.

2.5 Equilibrium Intermediation Contracts

2.5.1 Competitive Intermediation Market

We model the competitive intermediation market as follows. At time t , households offer intermediation contracts (β, K) ’s to the specialists; and then the specialists can accept the offer, or opt out of the intermediation market and manage their own wealth. In addition, any number of households are free to form coalitions with some specialists. At $t + dt$ the relationship is broken and the intermediation market repeats itself.

Definition 1 *In the intermediation market, households make offers (β, K) to specialists, and specialists can accept/reject the offers. A contract equilibrium in the intermediation market at date t satisfies the following two conditions:*

1. β is incentive compatible for each specialist.
2. There is no coalition of households and specialists, such that some other contracts can make households strictly better off while specialists weakly better off.

2.5.2 Equilibrium Contracts

Denote \mathcal{E}^h as the exposure of a household to the risky asset. We argue that given condition (2) in Definition 1, the equilibrium has to be symmetric with every specialist receiving the fee K , and every household obtaining exposure \mathcal{E}^h and paying a total fee of K . The argument we present here borrows from the core’s “equal-treatment” property in the study of the equivalence between the core

and Walrasian equilibrium (see Mas-Colell, Whinston, and Green (1995) Chapter 18, Section 18.B). Suppose that the equilibrium is asymmetric. We choose the household who is doing the worst—i.e. receiving lowest utility at some exposure \mathcal{E}^h and paying a fee K —and match him with the specialist who is doing the worst—i.e. receiving the lowest fee. This household-specialist pair can do strictly better by matching and forming an intermediation relationship. The only equilibrium in which such a deviating coalition does not exist is the symmetric equilibrium.

Next, we argue that in equilibrium, when purchasing risk exposure from the specialists, households are price takers who face a per-unit-exposure price

$$k = \frac{K}{\mathcal{E}^h}.$$

Thus a household that chooses exposure \mathcal{E}^h pays $k\mathcal{E}^h$ to obtain this exposure. The argument is as follows. Suppose that a measure of n (symmetric) households consider reducing their per-household exposure by ϵ relative to the equilibrium level \mathcal{E}^h . To do so, they reduce the measure of specialists in the coalition by $\frac{n\epsilon}{\mathcal{E}^h}$, thereby saving total fees of $\frac{n\epsilon}{\mathcal{E}^h}K$. Since the allocation is symmetric, each household reduces his fees, per unit ϵ , by $\frac{K}{\mathcal{E}^h}$. A similar argument implies that the households can raise their exposure at a price of k .

Consider a household's portfolio choice problem in investing in an intermediary given this price k . Suppose that each dollar of the risk exposure to the risky asset generates a risk premium of π_R . Then, paying the fee of k reduces the household's after-fee return to be $\pi_R - k$. It is obvious that the households' demand for risk exposure $\mathcal{E}^{h*}(k)$ is decreasing in k .

We have so far discussed how k enters into the household's investment decisions. For the specialist, since he has an outside option to trade on his own, it must be that $k \geq 0$ (i.e., $K \geq 0$) in equilibrium.¹⁰ We next argue that there are two distinct equilibria that arise: one with $k > 0$, and the other with $k = 0$. Which equilibrium arises depends on whether the incentive-compatibility constraint (4) is binding or not.

Suppose that the incentive-compatibility constraint (4) is slack, i.e., $\beta > \frac{b}{x}$. Note that each specialist just earns a profit of $K = k\mathcal{E}^h$, and households prefer a contract with a lower per-household delegation transfer. Then it implies that the equilibrium exposure price k has to be zero.

¹⁰There is a further result here that affects how specialists view the transfer K . Even though in equilibrium, each specialist receives the fee K , the specialist actually receives a fee that is linear in his maximum risk exposure $m\mathcal{E}^*$. Thus for example if one specialist had a maximum exposure of $2m\mathcal{E}^*$, he would receive a fee of $2K$. Since our model is symmetric, this cannot occur in equilibrium. However, when making dynamic decisions the specialist accounts for this dependence in considering how decisions alter $m\mathcal{E}^*$. We will explain this later in the paper when deriving the specialist's Euler equation.

Otherwise, by forming a coalition with n measure of households and $n - \epsilon$ measure of specialists, and reducing the specialists' share β to $\frac{(n-\epsilon)\beta}{n-\epsilon\beta} < \beta$ (so the households' total exposure remains at $\frac{1-\beta}{\beta}n\mathcal{E}^*$ in (3)) without changing the transfer K per-specialist, the new coalition can maintain the same per-household risk exposure at $\frac{1-\beta}{\beta}\mathcal{E}^*$, lower the per-household transfer, while keep the specialists indifferent. This deviation is strictly profitable unless the transfer K becomes zero, i.e., the exposure price $k = 0$.

We classify this case as the *unconstrained equilibrium*, or *unconstrained region*, where the incentive-compatibility constraint (4) is slack and the per-unit-exposure price k is zero. We can also think about this case in terms of the demand and supply of intermediation. Denote the households' aggregate demand for the risk exposure, given the free intermediation service, as $\mathcal{E}^{h*}(k = 0)$. The zero-delegation-price equilibrium arises when $\mathcal{E}^{h*}(k = 0)$ is below the maximum risk exposure $m\mathcal{E}^*$ available in the economy. When this occurs, the economy is in the unconstrained equilibrium,

When $\mathcal{E}^{h*}(k = 0)$ exceeds the aggregated maximum exposure $m\mathcal{E}_s^*$ provided by the specialists, we are at the *constrained equilibrium*, or *constrained region*. In this case, specialists earn a positive rent $K = km\mathcal{E}^* > 0$ for their scarce service. Following the deviating coalition/contract argument as above implies that the incentive-compatibility constraint (4) for every specialist must be binding. i.e., $\beta = \frac{b}{x}$. Otherwise, invoking our previous argument, households could indeed form a coalition with a specialist whose incentive constraint is slack, thereby lowering their price k .

We summarize these results regarding the equilibrium classification in the following proposition:

Proposition 1 *At any date t , the economy is in one of two equilibria:*

1. *The intermediation unconstrained equilibrium occurs when,*

$$\mathcal{E}^{h*}(k = 0) \leq m\mathcal{E}^*.$$

In this case, the incentive-compatibility constraint of every specialist is slack.

2. *The intermediation constrained equilibrium occurs when there exists a positive exposure price k , such that*

$$\mathcal{E}^{h*}(k > 0) = m\mathcal{E}^*.$$

In this case, the incentive-compatibility constraint is binding for all specialists. β is equal to its minimum value of $\frac{b}{x}$.

2.6 Implementation

In Section 2.4, we see that the heart of the agency friction imposes a restriction on the maximum risk exposure that the households can obtain through intermediaries, in that $\mathcal{E}^h \leq m\mathcal{E}^*$ in (6). From a slightly different angle, because \mathcal{E}^* is the specialist's exposure to the risky asset, this restriction dictates a risk-sharing rule between the household and the specialist in the intermediary. In the language of equity contracts, the restriction can be interpreted as one in which the households, as outsiders of the intermediary, cannot hold more than $\frac{m}{1+m}$ (equity) shares of the intermediary.

Therefore, the somewhat abstract (β, K) contract can be implemented and interpreted readily in terms of equity contributions by households and specialists, with the maximum exposure constraint interpreted as an equity capital constraint—i.e., given the specialist's equity contribution W , households can make at most mW equity contributions to the intermediary. Moreover, households pay the specialist an intermediation fee f per-unit of wealth that is invested in the intermediary; then the delegation transfer K can be interpreted as the households' total intermediation cost when they seek equity investment in intermediaries.

The following definition gives the equity implementation of our optimal contract. This is the language we use in the rest of the paper when discussing the contracting problem.

Definition 2 (Equity Implementation)

The equity implementation of the optimal contract is as follows:

1. A specialist contributes all his wealth W_t into an intermediary, and household(s) contribute $T_t^h \leq W_t^h$.¹¹
2. Both parties purchase equity shares in the intermediary. The specialist owns $\frac{W_t}{W_t + T_t^h}$ fraction of the equity of intermediary, while the households own $\frac{T_t^h}{W_t + T_t^h}$.
3. Equity contributions must satisfy the **capital constraint**

$$T_t^h \leq mW_t.$$

4. Households pay the specialist an intermediation fee of f_t per dollar invested in the intermediary. The total transfer K paid by the households is $T_t^h f_t$.

¹¹Note that on point (1), the specialist is indifferent between contributing and not contributing all of his wealth to the intermediary. We can also consider implementations in which the specialist contributes a fraction $\gamma \in (0, 1]$ of his wealth to the intermediary. Our results will be identical, with a suitable redefinition of the capital constraint parameter m to be m/γ . The primitive incentive constraint is invariant to the value of γ .

The counterpart of Proposition 1, which describes the equilibrium conditions in the equity implementation, is

Proposition 2 *At any date t , the economy is in one of two equilibria:*

1. *In the unconstrained region, the capital constraint is slack, $T_t^h < mW_t$, and we have zero intermediation fee $f_t = 0$.*
2. *In the constrained region, the capital constraint is binding, $T_t^h = mW_t$, and we have a positive intermediation fee $f_t > 0$.*

The equity implementation of our model makes it clear that, along the equilibrium path, the specialists have to absorb no less than $\frac{1}{1+m}$ of the aggregate risk in this economy, independent of the specialists' wealth. Therefore, under unfavorable economic conditions when their wealth is low, specialists have to bear disproportionately large risk, and as a result asset prices have to adjust to make the greater risk exposure optimal. This tension drives our asset pricing results throughout the paper.

Our modeling of intermediation and the derivation of the capital requirement closely follows Holmstrom and Tirole (1997). We have adapted the Holmstrom and Tirole assumptions to a setting with risk averse agents and no limited liability, but still recover the capital requirement as the key aspect of intermediation contracts.

We think of the incentive constraint that emerges from the model as similar to the explicit and implicit incentives across many modes of intermediation. For example, a hedge fund manager is typically paid 20% of the return on his fund. We may think of this 20% as corresponding to the minimum fraction β that has to be paid to the hedge fund manager in order to provide investment incentives.

The equity implementation and constraint, as argued in Holmstrom and Tirole, is also similar to the capital constraints faced by commercial banks. Stretching the interpretation a bit more, we may also think of the incentive constraint as capturing implicit incentives in the mutual fund industry. There is a well established relation between past performance and mutual fund flows (see, e.g., Warther (1995)). We can think of this performance-flow relation as reflecting an implicit incentive constraint. As W falls, the households contribution T^h falls. Shleifer and Vishny (1997) present a model with a similar feature: the supply of funds to an arbitrageur in their model is assumed to be a function of the previous period's return by the arbitrageur.

The key feature of the model, which we think is robustly reflected across many modes of intermediation in the world, is the feedback between losses suffered by an intermediary (drop in W) and exit by the investors of that intermediary. Our model captures this feature through the capital constraint, when it is binding.

2.7 Decisions and Equilibrium

The decision problem of a specialist is to choose his consumption rate c_t and the portfolio share in the risky asset α_t for the intermediary. The share choice α_t is isomorphic to the exposure choice \mathcal{E}^I described in Section 2.2, but it is more convenient to work with the former under the equity implementation.

Denote the cumulative return delivered by the intermediary as \widetilde{dR}_t . The specialist contributes all of his wealth to the intermediary and earns the return \widetilde{dR}_t plus the fee of $f_t T_t^h dt$. Thus, the specialist's problem is:

$$\max_{\{c_t, \alpha_t\}} E \left[\int_0^\infty e^{-\rho t} \ln c_t dt \right] \quad s.t. \quad dW_t = -c_t dt + W_t \widetilde{dR}_t(\alpha_t) + f_t T_t^h dt, \quad (7)$$

where the return delivered by intermediaries \widetilde{dR}_t , as a function of α_t , is

$$\widetilde{dR}_t(\alpha_t) = \alpha_t (dR_t - r_t dt) + r_t dt.$$

Note that the intermediary's portfolio share α_t is also the portfolio share on the specialist's own wealth.

The household chooses his consumption rate c_t^h and funds for delegation T_t^h , given his wealth W_t^h . Following the equity implementation of the intermediation contract with delegation fee, the fraction of wealth that is invested with an intermediary is $\frac{T_t^h}{W_t^h}$, which earns a net return of $\widetilde{dR}_t - f_t dt$. Then the return on the household's wealth is,

$$\widetilde{dR}_t^h = \left(1 - \frac{T_t^h}{W_t^h}\right) r_t dt + \frac{T_t^h}{W_t^h} (\widetilde{dR}_t - f_t dt),$$

The optimization problem for a household is:

$$\max_{\{c_t^h, T_t^h\}} E \left[\int_0^\infty e^{-\rho^h t} \ln c_t^h dt \right] \quad s.t. \quad dW_t^h = -c_t^h dt + W_t^h \widetilde{dR}_t^h. \quad (8)$$

Definition 3 An equilibrium is a set of progressively measurable price processes $\{P_t\}$, $\{r_t\}$, and $\{f_t\}$, and decisions $\{T_t^h, c_t, c_t^h, \alpha_t\}$ such that,

1. Given the price processes, decisions solve (7) and (8).

2. The intermediation decisions satisfy the equilibrium conditions of Proposition 2.

3. The stock market clears:

$$\alpha_t(W_t + T_t^h) = P_t.$$

4. The goods market clears:

$$c_t + c_t^h = D_t.$$

Given market clearing in risky asset and goods markets, the bond market clears by Walras' law. The market clearing condition for the risky asset market reflects that the intermediary is the only direct holder of the risky asset, and the total holding of the risky asset by the intermediary must equal the supply of the risky asset.

3 Asset Market Equilibrium

We look for a stationary Markov equilibrium where the state variables are (W_t, D_t) . It is clear that D_t must be one of the state variables, because the dividend process is the fundamental driving force in the economy. Intermediation frictions imply that the distribution of wealth between households and specialists affects equilibrium as well. For example, whether capital constraints bind or not depends on the relative wealth of households and specialists. We have some freedom in choosing how to define the wealth distribution state variable. We choose to use the specialist's wealth W_t to emphasize the effects of intermediary capital.

The intrinsic scale invariance (the log preferences and the log-normal dividend process) in our model implies that the scaled specialist's wealth $w = W/D$ is the only state variable to characterize our economy. Indeed, we will see that the equilibrium price/dividend ratio P/D , the risk premium π_R , the interest rate r , and the intermediation fee f are functions of w only.

We write the total return on the risky asset as,

$$dR_t = \frac{D_t dt + dP_t}{P_t} = \mu_R dt + \sigma_R dZ_t, \tag{9}$$

where μ_R is the risky asset's expected return and σ_R is the volatility. The risky asset's risk premium π_R is simply $\mu_R - r$.

3.1 Risky Asset Price

A simple argument due to log preferences for both agents allows us to derive the equilibrium risky asset price P_t in closed form. For the household with wealth W_t^h , his optimal consumption is

$$c_t^h = \rho^h W_t^h.$$

Likewise the optimal consumption for the specialist is $c_t = \rho W_t$. But since the debt is in zero net supply, the aggregated wealth has to equal the market value of the risky asset, i.e.,

$$W_t^h + W_t = P_t.$$

Invoking the goods market clearing condition $c_t + c_t^h = D_t$, we solve for the equilibrium price of the risky asset.

Proposition 3 *The equilibrium risky asset price as a function of the state variables is:*

$$P_t = \frac{D_t}{\rho^h} + \left(1 - \frac{\rho}{\rho^h}\right) W_t. \quad (10)$$

It follows that the price/dividend ratio is $\frac{P_t}{D_t} = \frac{1}{\rho^h} + \left(1 - \frac{\rho}{\rho^h}\right) w_t$.

Taking the limit where the specialist wealth goes to zero, we observe that the asset price P_t approaches D_t/ρ^h . Loosely speaking, this is the asset price for an economy only consisting of households. At the other limit, as the households wealth goes to zero (i.e., W_t approaches P_t), the asset price/dividend ratio approaches D_t/ρ .

We assume throughout that $\rho^h > \rho$. Then, the asset price is lowest when households make up all of the economy, and increases linearly from there with specialist wealth, W_t . This is a simple way of capturing a low “liquidation value” of the asset, which becomes relevant when specialist wealth falls and there is disintermediation. Note that liquidation is an off-equilibrium thought experiment, since in our model, asset prices adjust so that the asset is never liquidated by the specialist.

3.2 Capital Constraint and Specialist’s Portfolio Share

The specialist chooses the portfolio share α_t of the risky asset for the intermediary, which is also the portfolio share for the specialist’s own wealth invested in the risky asset. We can use the market clearing condition for the risky asset to pin down α_t . As the capital constraint affects the specialist’s exposure to the risky asset, we have to consider two regions depending on whether the capital constraint is binding or not.

First, we argue that if $mW_t > W_t^h$, then the capital constraint is slack, and we are at the unconstrained region as defined in Proposition 2. To see this, we only need to check that the zero intermediation fee $f_t = 0$ leads to an intermediation demand T_t^h lower than mW_t . In fact, we argue that the household's intermediation demand at zero fee is his entire wealth, i.e., $T_t^h = W_t^h < mW_t$.

The argument is as follows. When $f_t = 0$, both household and specialist face identical investment opportunities. As a result, by purchasing $T_t^h = W_t^h < mW_t$ amount of equity, the household obtains the same portfolio share as the specialist. Because the specialist makes the portfolio share decision for the specialist, which is therefore the optimal portfolio choice for the specialist, this portfolio choices must also be optimal for the household. In short, when $mW_t > W_t^h$, households can invest 100% of their wealth into intermediaries, obtaining their optimal exposure to the risky asset.

Therefore, the economy is in the unconstrained region when $mW_t > W_t^h$. In this case, both household and specialist must have the same portfolio share in the risky asset. Because the riskless bond is in zero net supply, market clearing implies that $\alpha_t = 1$.

Second, when $mW_t < W_t^h$, investing the household's entire wealth into the intermediary $T_t^h = W_t^h$ violates the capital constraint. Now we are at the constrained region, and in equilibrium the intermediaries have a total capital of W_t plus the household's capital investment of mW_t . Since the risky asset must be held by intermediaries, using (10) we find the portfolio share in the risky asset to be,

$$\alpha_t = \frac{P_t}{W_t + mW_t} = \frac{1 + (\rho^h - \rho) w_t}{(1 + m) \rho^h w_t}. \quad (11)$$

Finally, since $W_t + W_t^h = P_t$, the critical point w^c where the capital constraint is binding ($mW_t = W_t^h$) can be easily derived as

$$w^c = \frac{1}{m\rho^h + \rho}. \quad (12)$$

When the scaled specialist's wealth $w \geq w^c$, the economy is unconstrained; while the economy is constrained when $w < w^c$. The following proposition summarizes our result.

Proposition 4 *Let $w^c = \frac{1}{m\rho^h + \rho}$. We have:*

1. *The economy is in the unconstrained region when $w_t \geq \frac{1}{m\rho^h + \rho}$. In this region, $mW_t \geq W_t^h$, and the specialist's portfolio share $\alpha_t = 1$.*
2. *The economy is in the constrained region when $w_t < \frac{1}{m\rho^h + \rho}$. In this region, $W_t < W_t^h$, and specialist's portfolio share $\alpha_t = \frac{1 + (\rho^h - \rho) w_t}{(1 + m) \rho^h w_t}$.*

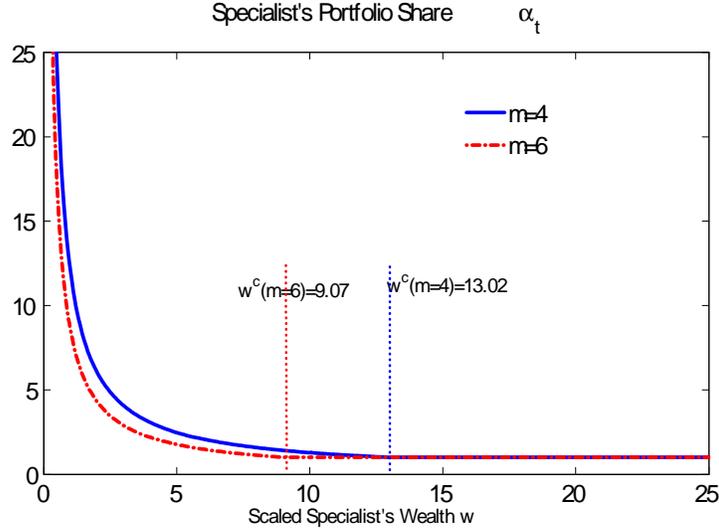


Figure 1: The specialist’s portfolio share α_t in the risky asset is graphed against the scaled specialist wealth w for $m = 4$ and 6 . The constrained (unconstrained) region is on the left (right) of the threshold w^c . Other parameters are $g = 1.84\%$, $\sigma = 12\%$, $\rho = 1\%$, and $\rho^h = 1.67\%$ (see Table 1).

In Figure 1 we plot the specialist’s portfolio share α_t in the risky asset against the scaled specialist’s wealth, the only relevant state variable in our model. The specialist’s portfolio holding in the risky asset rises above 100% once the economy is capital constrained, and rises even higher when the specialist’ wealth falls further.

Two Effects on m : Constraint Effect and Sensitivity Effect Figure 1 illustrates the comparative static results for the cases of $m = 4$ and $m = 6$. There are two effects of the intermediation multiplier m . The first is a “constraint effect.” The intermediation multiplier m captures the maximum amount of households’ (outside) capital that can be raised per specialist’s (insider’s) capital, thus giving an inverse measure of the severity of agency problems in our model. Increasing m reduces the agency problem and thereby loosens intermediaries’ capital constraint for a given wealth distribution. From (12), it is immediate to see that $w^c(m = 6)$ is smaller than $w^c(m = 4)$, and therefore the unconstrained region (where $w < w^c$) is larger when $m = 6$.¹²

Additionally, Figure 1 shows that in the constrained region, the specialist’s portfolio share α_t invested in the risky asset, through market clearing, rises as the capital constraint tightens. When m is lower, the capital constraint binds for smaller values of w . This in turn means that for a given

¹²In the limit, if we raise m towards infinity, households participate fully in the risky asset market regardless of the specialist wealth, and the constrained region vanishes.

value of w , the lower the m , the higher the specialist's holding in the risky asset.

There is a second, more subtle, "sensitivity effect" of m , when we consider the economic impact of a marginal change in the specialist's wealth, given some tightness of constraints. This sensitivity effect is rooted in the nature of the capital constraint. When in the constrained region, a \$1 drop in the specialist's capital reduces the households' equity participation in the intermediary by \$ m . A higher m makes the economy more sensitive to the changes in the underlying state, and therefore magnifies capital shocks.

It is possible, although not readily apparent, to see the sensitivity effect in Figure 1. For the $m = 6$ case, α_t rises faster in the constrained region than for the $m = 4$ case. To analytically show this point, we calculate the derivative of portfolio share α_t with respect to w_t using (11), and evaluate this derivative (in its absolute value) across the same level of α_t :

$$\left| \frac{d\alpha_t}{dw_t} \right| = \frac{1}{(1+m)\rho^h} \frac{1}{w_t^2} = \frac{[\alpha_t(1+m)\rho^h - (\rho^h - \rho)]^2}{(1+m)\rho^h}.$$

Differentiating this expression with respect to m , we find that,

$$\frac{d}{dm} \left| \frac{d\alpha_t}{dw_t} \right| = \rho^h \left((1+m)^2 \alpha_t^2 - (1 - \rho/\rho^h)^2 \right),$$

which is positive for all relevant parameters (recall that $\alpha_t \geq 1$ and that $\rho^h > \rho$). In other words, when m is higher, a change in specialist wealth leads to a larger change in α_t . While we do not go through the computations in the next sections, this sensitivity effect arises in most of the asset pricing measures that we consider.

The two effects of m shed light on crises episodes. If consider that an economy like the U.S. has institutions with higher m s, then our model can help explain why crisis episodes are unusual (constraint effect), but on incidence, are often dramatic (sensitivity effect).

In Figure 1, the observation that the specialist's holding becomes higher in the constrained region is critical in understanding our asset pricing results throughout the paper. Recall that in our model, the specialist, not the household, is in charge of the intermediaries investment decisions. Thus, asset prices have to adjust to make the higher risk share optimal. The next sections detail the asset pricing implications of our model.

3.3 Volatility of Specialist Wealth

We may write the equilibrium evolution of the specialist's wealth W_t as

$$\frac{dW_t}{W_t} = \mu_W dt + \sigma_W dZ_t, \tag{13}$$

where the drift μ_W and the volatility σ_W are to be determined in equilibrium. By matching the diffusion term in (13) with the specialist's budget equation (7), it is straightforward to see that,

$$\sigma_W = \alpha_t \sigma_R. \quad (14)$$

The volatility of the specialist's wealth is equal to the volatility of the risky asset return, modulated by the position of the risky asset held by the specialist.

Given (10), the diffusion term on the risky asset price is,

$$\sigma(dP_t) = \sigma \frac{D_t}{\rho^h} + \left(1 - \frac{\rho}{\rho^h}\right) W_t \sigma_W.$$

Then,

$$\sigma_R = \frac{1}{P_t} \left(\sigma \frac{D_t}{\rho^h} + \left(1 - \frac{\rho}{\rho^h}\right) W_t \sigma_W \right). \quad (15)$$

Combining (14) and (15) we solve for σ_W :

$$\sigma_W = \frac{\sigma}{\frac{\rho^h}{\alpha_t} \frac{P_t}{D_t} - (\rho^h - \rho) w_t}.$$

Now based on the equilibrium portfolio share α_t derived in Proposition 3, we can solve for the volatility of the specialist's wealth.

Proposition 5 *In the unconstrained region,*

$$\sigma_W = \sigma.$$

In the constrained region,

$$\sigma_W = \frac{\sigma}{w_t(m\rho^h + \rho)}.$$

Not surprisingly, Figure 2 shows that the volatility of the specialist's wealth displays a similar pattern as that of α_t . In the unconstrained region, the volatility of the specialist's wealth is constant. In the constrained region, the volatility of wealth rises as the specialist's wealth falls, and the specialist bears disproportionately more risk in the economy. The two effects—constrained effect and sensitivity effects—are also visible from the figure.

3.4 Risky Asset Volatility

Now we are ready to solve for the volatility of risky asset σ_R , as $\sigma_R = \frac{\sigma_W}{\alpha_t}$ according to (14).

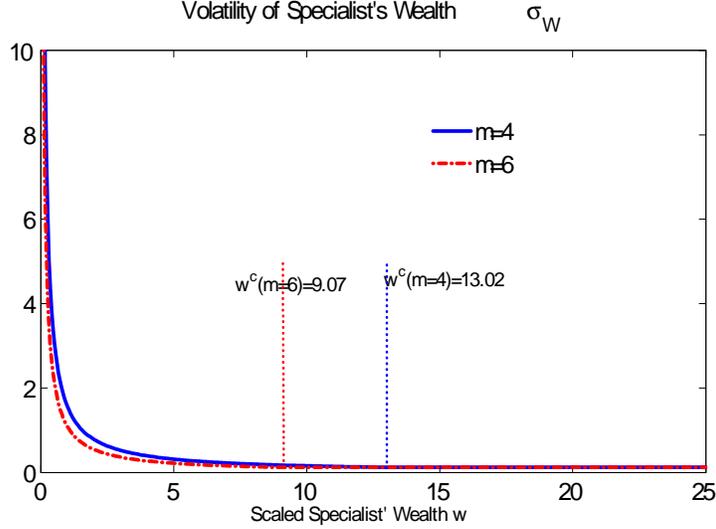


Figure 2: The volatility of the specialist's wealth σ_W is graphed against the scaled specialist wealth w for $m = 4$ and 6 . The constrained (unconstrained) region is on the left (right) of the threshold w^c . Other parameters are $g = 1.84\%$, $\sigma = 12\%$, $\rho = 1\%$, and $\rho^h = 1.67\%$ (see Table 1).

Proposition 6 *In the unconstrained region, we have,*

$$\sigma_R = \sigma.$$

In the constrained region, we have,

$$\sigma_R = \sigma \left(\frac{(1+m)\rho^h}{m\rho^h + \rho} \right) \left(\frac{1}{1 + (\rho^h - \rho)w_t} \right).$$

As Figure 3 shows, the volatility of risky asset is constant in the unconstrained region, which is just the dividend volatility σ . The volatility rises in the constrained region, as the constraint tightens (i.e. W_t falls). To see this, equation (15) implies that

$$\sigma_R = \frac{1}{P_t/D_t} \left(\sigma \frac{1}{\rho^h} + \left(1 - \frac{\rho}{\rho^h} \right) w_t \sigma_W \right).$$

We have seen that in Proposition 5, $w_t \sigma_W$ is a constant in the constrained region. Therefore, for smaller scaled specialist wealth w 's, σ_R increases because the price/dividend ratio P_t/D_t falls, a phenomenon consistent with the fire-sale discount of the intermediated assets.

The model can help explain the rise in volatility that accompanies period of financial turmoil where intermediary capital is low. It can also help to explain the rise in the VIX index during these periods, and why the VIX has come to be called a “fear” index. We will next show that the periods of low intermediary capital also lead to high expected returns. Taking these results

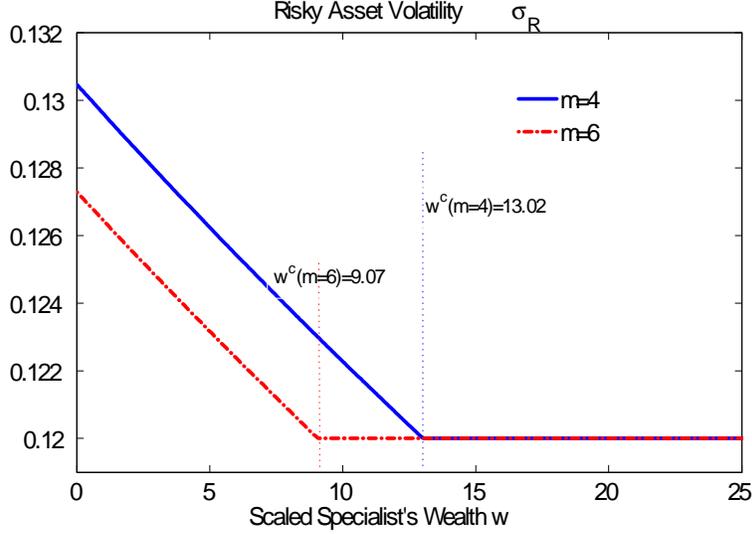


Figure 3: The risky asset volatility σ_R is graphed against the scaled specialist wealth w for $m = 4$ and 6 . The constrained (unconstrained) region is on the left (right) of the threshold w^c . Other parameters are $g = 1.84\%$, $\sigma = 12\%$, $\rho = 1\%$, and $\rho^h = 1.67\%$ (see Table 1).

together, we provide one possible explanation for recent empirical observations relating the VIX index and risk premia on intermediated assets. Bondarenko (2004) documents that the VIX index helps explain the returns to many different types of hedge funds. Berndt, et al. (2004) note that the VIX index is highly correlated with the risk premia embedded in default swaps. In both cases, the assets involved are specialized and intermediated assets that match those of our model. Our model suggests that, as intermediaries hit their capital constraints, the intermediation capital—which is the wealth of marginal investors (as specialists in this model)—becomes more volatile, and this translates to rising market volatilities and rising VIX index. At the same time, as we see in the next section, increased volatility gives rise to higher risk premia on the assets that they are trading.

3.5 Risk Premium

The key observation regarding our model is that the specialist is in charge of the investment decisions into the risky asset. This means that asset prices have to be such that it is optimal for him to buy the market clearing amount of α_t . For the households on the other hand, their (indirect) investment in the risky asset may be constrained and affected by the intermediation frictions.

The specialist's Euler equation for pricing risky asset return dR_t is,

$$m f_t dt - \rho dt + E_t \left[\frac{dc_t}{c_t} \right] + Var_t \left[\frac{dc_t}{c_t} \right] + E_t [dR_t] = Cov_t \left[\frac{dc_t}{c_t}, dR_t \right]. \quad (16)$$

This expression looks like the standard consumption Euler equation, except for the first term $mf_t dt$, which is the total fee that the specialist earns per unit of his wealth. Note that this expression encompasses both regions, as $mf_t = 0$ when the economy is constrained.

To understand this additional term due to the intermediation fee, consider a specialist who decreases consumption today by δ and uses the δ to increase his investment in the intermediary. As in the usual argument, this strategy has a consumption cost today and a gain tomorrow when the proceeds of this investment are consumed. Relative to the usual argument there is a twist in our case, because the increased investment, δ , attracts further households investment on which the specialist gets a fee. The additional fee amounts to $mf_t \delta$ that the specialist can immediately consume. This explains the first term in the Euler equation.

We can easily verify that consumption policy of $c_t = \rho W_t$ satisfies the Euler equation (16). Then, dc_t/c_t is equal to dW_t/W_t . Applying the Euler equation to risky asset return dR_t and to a riskless bond, we find,

$$\pi_R dt = E_t[dR_t - r_t dt] = \sigma_R \sigma_W dt.$$

This is the familiar CAPM pricing result. Since the specialist has log preferences, a CAPM holds with the market portfolio defined as the return on the specialist's wealth.

Proposition 7 *In the unconstrained region, we have,*

$$\pi_R = \sigma^2.$$

In the constrained region, we have,

$$\pi_R = \frac{\sigma^2}{w_t(m\rho^h + \rho)} \left(\frac{(1+m)\rho^h}{(1+m)\rho^h - (\rho^h - \rho)} \right) \left(\frac{1}{1 + (\rho^h - \rho)w_t} \right).$$

Since both σ_R and σ_W rise as W_t falls, the risk premium on the risky asset rises through the constrained region, as shown in Figure 4. It is easy to show that this pattern also prevails for the Sharpe ratio.

An interesting point of comparison for our results is to the literature on state-dependent risk premia, notably, Campbell and Cochrane (1999), Barberis, Huang, and Santos (2001), and Kyle and Xiong (2001). In these models, as in ours, the risk premium is increasing in the adversity of the state. In Campbell and Cochrane, the state dependence arises because marginal utility is dependent on the agent's consumption relative to his habit stock. In Barberis, Huang, and Santos, the state dependence comes about because risk aversion is modeled directly as a function of the

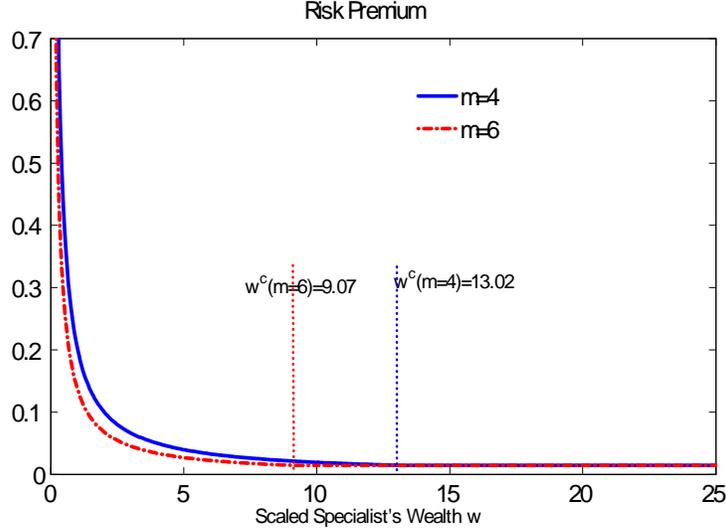


Figure 4: Risk premium π_R is graphed against the scaled specialist wealth w for $m = 4$ and 6 . The constrained (unconstrained) region is on the left (right) of the threshold w^c . Other parameters are $g = 1.84\%$, $\sigma = 12\%$, $\rho = 1\%$, and $\rho^h = 1.67\%$ (see Table 1).

previous period’s gains and losses. Relative to these two models, we work with a standard CRRA utility function, but generate state dependence endogenously as a function of the frictions in the economy.

For empirical work, our approach suggests that measures of intermediary capital/capacity will help to explain risk premia. In this regard, our model is closer in spirit to Kyle and Xiong who generate a risk premium that is a function of “arbitrageur” wealth. The main theoretical difference between Kyle and Xiong and our model is that the wealth effect in their model comes from assuming that the arbitrageur has log utility, while in our model it comes because the intermediation constraint arises endogenously as a function of intermediary capital under an explicitly modeled delegation friction. One clear difference across these models is revealed in the sharp asymmetry of our model’s risk premia: no dependence on capital in the unconstrained region and a strong dependence in the constrained region. In Kyle and Xiong, the log utility assumption delivers a risk premium that is a much smoother function of arbitrageur wealth. Plausibly, to explain a crisis episode, one needs the type of asymmetry delivered by our model.

3.6 Intermediation fee

We now turn to the intermediation market to determine the equilibrium intermediation fee. In Section 2.5 we have shown that in the unconstrained region, the excess supply of intermediation

service implies that households pay zero fee ($f_t = 0$) when purchasing equity shares in the intermediaries. In the constrained region, when households purchase equity claims of intermediaries, the total supply of outside equity is binding due to the capital constraint—or more fundamentally, due to the agency problem in the delegated asset management. Then, the competitive intermediation market gives rise to a positive intermediation fee $f_t > 0$ for the scarce intermediation service.

Let us solve for f_t in the constrained region. With log preference, the specialist's portfolio choice is myopic and mean-variance efficient. Specifically, given his wealth W_t , the specialist's optimal (dollar) exposure in the risky asset is

$$\mathcal{E}^* = \frac{\pi_R}{\sigma_R^2} W_t,$$

where π_R (σ_R) is the risky asset's risk premium (volatility).¹³ Note that the specialist's optimal exposure is not affected by the intermediation fee f_t .

The households' demand for exposure, on the other hand, is decreasing with the intermediation fee. The equilibrium fee f_t then equates the demand with the (inelastic) supply. To find the equilibrium fee f_t , it is easier to derive the equilibrium price of risk exposure k_t first. Log preferences implies that the household is myopic and mean-variance efficient. However, as discussed in Section 2.5.2, the household's effective risk premium from obtaining exposure in the risky asset is the risk premium π_R , minus the delegation cost k_t per unit of exposure. Therefore the household's demand for risk exposure is,

$$\mathcal{E}^{h*}(k_t) = \frac{\pi_R - k_t}{\sigma_R^2} W_t^h.$$

In the constrained equilibrium, $m\mathcal{E}^* = \mathcal{E}^{h*}$ and $\mathcal{E}^{h*} + \mathcal{E}^* = P = W_t^h + W_t$. These relations imply that

$$k_t = \frac{P_t - (1 + m)W_t}{P_t - W_t} \pi_R.$$

Finally, since we can express the total delegation transfer K_t as either $k_t m \mathcal{E}^*$ or $f_t T_t^h = f_t m W_t$, the equilibrium per-unit wealth intermediation fee is

$$f_t = k_t \frac{\mathcal{E}^*}{W_t} = k_t \frac{\pi_R}{\sigma_R^2}.$$

Plugging in the previous results, we have the following result:

¹³Since the total delegation rents that the specialist earns is proportional to \mathcal{E}^* , and his optimal exposure \mathcal{E}^* is proportional to W , the specialist earns a rent proportional to W . Therefore, effectively, in this economy the specialist obtains an extra return per unit of his wealth in addition to any regular investment return (from bonds and/or stocks). This gives a formal justification for the claim that the specialist's total fee through intermediation is linear in his wealth.

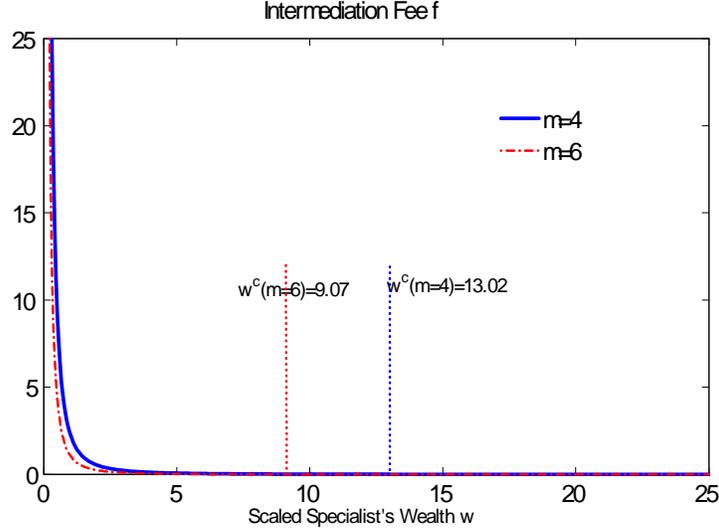


Figure 5: Intermediation fee f per unit of delegated wealth is graphed against the scaled specialist wealth w for $m = 4$ and 6 . The constrained (unconstrained) region is on the left (right) of the threshold w^c . Other parameters are $g = 1.84\%$, $\sigma = 12\%$, $\rho = 1\%$, and $\rho^h = 1.67\%$ (see Table 1).

Proposition 8 *In the unconstrained region, the intermediation fee is*

$$f_t = 0.$$

In the constrained region, the intermediation fee is

$$f_t = \frac{\sigma^2}{(\rho + m\rho^h)^2} \frac{1 - (\rho + m\rho^h) w_t}{1 - \rho w_t} \left(\frac{1}{w_t}\right)^2 > 0.$$

In Figure 5, the intermediation fee displays a similar pattern as the risk premium in Figure 4. This is intuitive: the higher risk premium in the constrained region implies a higher household demand for investment in intermediaries to gain access to the higher risk premium. Because the supply is fixed at mW_t , to clear the intermediation market, the equilibrium fee rises.

The higher intermediation fee is the logical result of our model of scarce supply of intermediation. However, it seems counterfactual that during a crisis period in which agency concerns may be widespread, that specialists can demand a higher fee from their investors. One resolution of this anomalous result is to assume that households, lacking the knowledge of the risky asset market, are also not aware of time variation in the risk premium on the risky asset. That is, they are not aware of the current state of the economy. For example, one can explore a model in which households hold static beliefs over the mean-variance ratio of the payoffs delivered by intermediaries. This

model may deliver the result that fees are state independent, thereby resolving the counterfactual result on fees. We do not pursue this extension here.

On the other hand, the positive intermediation fee can also be seen as a reflection of the scarcity of the specialists' capital. This take on the fee reflects one of the key points of our model: intermediation capital becomes increasingly valuable during the liquidity event when the intermediary sector suffers more losses. Thus the intermediation fee f_t also measures the shadow price of the scarce intermediary capital in this economy. The following example illustrate this point.

Example: Lending Spreads and Market Liquidity

During periods of financial turmoil in the intermediary sector, the terms of credit for new loans get worse. That is, lending spreads rise, even on relatively safe borrowers. In our model, we can interpret this rise as reflecting the scarcity of intermediary capital.

We interpret the intermediary now as not just a purchaser of secondary market assets, but also a lender in the primary market. That is, the intermediaries are commercial banks or investment banks. Suppose that a borrower (infinitesimal) asks the intermediary for a loan at date t to be repaid at date $t + dt$, with zero default risk. We denote the interest rate on this loan as R_t , and ask what R_t lenders will require.

Suppose that making the loan uses up capital. That is to say, if a specialist makes a loan of size δ , he has less wealth ($W_t - \delta$) available for coinvestment with the household in the intermediary. In particular, if in the constrained region, the lender is able to attract $m\delta$ less funds from the households.¹⁴

If $mW_t > W_t^h$, intermediation capital is not scarce and thus $R_t = r_t$. However, if intermediation capital is scarce, then using intermediation capital on the loan reduces the size of the intermediary. A lender could have used the δ in the intermediary to purchase the riskless bond yielding r_t and received a fee from households of $mf_t\delta$. Since both investments are similarly riskless, we must have that,

$$R_t = r_t + mf_t.$$

We have seen that falling into the constrained region causes the intermediation fee f_t to rise, and

¹⁴To develop this example in terms of the primitive incentive constraint, we need to assume that households only observe the specialist's wealth net of the loan, and do not observe the actual loan. Also, households' beliefs are that every specialist will contribute his entire wealth into the intermediary when the delegation fee is positive, a belief that is consistent with the current equilibrium. In this case, observing wealth of $W_t - \delta$ leads households to believe that the risk exposure delivered by that specialist is reduced proportionately, which in turn tightens the intermediation capacity constraint.

so does the lending spread mf_t .

In this example, even a no-default-risk borrower is charged the extra spread of mf_t . The key reason is that the specialist-intermediary is marginal in pricing the loan to the new borrower, so that the opportunity cost of specialist capital is reflected in the lending spread. If we had assumed that households could also have made such a loan, then we will find that $R_t = r_t$. Of course a business loan, which requires expertise and knowledge of borrowers, is the prime example of an intermediated investment.¹⁵

3.7 Interest Rate and Flight to Quality

We can derive the equilibrium interest rate r_t from the household's Euler equation, which is

$$r_t dt = \rho^h dt + E_t \left[\frac{dc_t^h}{c_t^h} \right] - \text{Var}_t \left[\frac{dc_t^h}{c_t^h} \right].$$

The equilibrium condition gives us,

$$\frac{dc_t^h}{c_t^h} = \frac{d(\rho^h W_t^h)}{\rho^h W_t^h} = \frac{d(P_t - W_t)}{P_t - W_t}.$$

Recall that the specialist's budget equation,

$$dW_t/W_t = \alpha_t (dR_t - r_t dt) + r_t dt - \rho dt + mf_t dt.$$

Using the expressions for α_t , σ_R , and f_t that have been derived previously, we have the following result:

Proposition 9 *In the unconstrained region, the interest rate is*

$$r_t = \rho^h + g + \rho (\rho - \rho^h) w_t - \sigma^2.$$

In the constrained region, the interest rate is

$$r_t = \rho^h + g + \rho (\rho - \rho^h) w_t - \sigma^2 \frac{\left[\rho \left((1+m) \left(\frac{1}{w_t} - \rho \right) - m^2 \rho^h \right) + (m\rho^h)^2 \right]}{(1 - \rho w_t) (\rho + m\rho^h)^2}.$$

We observe that in the unconstrained region, the interest rate is decreasing in the scaled specialist's wealth w (recall that $\rho < \rho^h$). This just reflects the divergence in both parties' discount rates. In the limiting case where $W_t = \frac{D_t}{\rho}$, the economy only consists of specialists. Then, consistent

¹⁵The results illustrated in this example are also present in the Holmstrom and Tirole (1997) model, although the connection to secondary market activity is not apparent in their model.

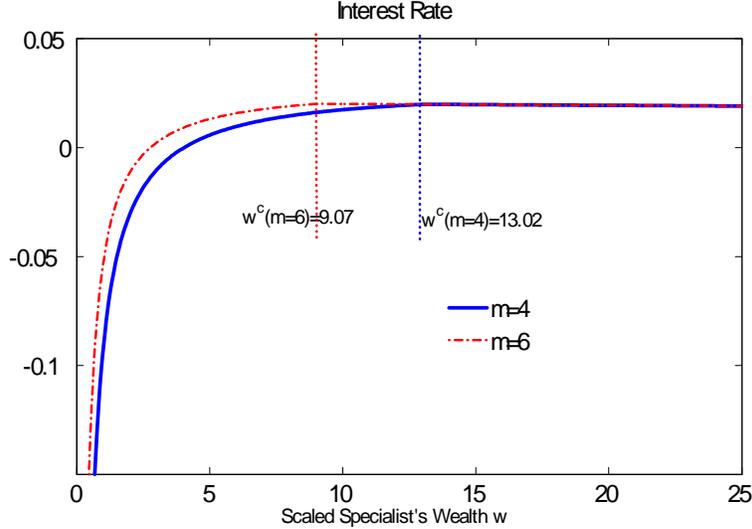


Figure 6: Interest rate r is graphed against the scaled specialist wealth w for $m = 4$ and 6 . The constrained (unconstrained) region is on the left (right) of the threshold w^c . Other parameters are $g = 1.84\%$, $\sigma = 12\%$, $\rho = 1\%$, and $\rho^h = 1.67\%$ (see Table 1).

with the familiar result of an economy with specialists as representative log-investors, the interest rate converges to $\rho + g - \sigma^2$. For a smaller w , where households play a larger part of the economy, the bond's return also reflects the households' discount rate ρ^h , and the equilibrium interest rate is higher.

In the constrained region, the pattern is reversed: The smaller the specialist's wealth, the lower the interest rate. This is because the capital constraint brings about two larger effects that reinforce each other. First, when the capital constraint is binding, the result in Proposition 3 implies that the specialists bear disproportionately greater risk in this economy: The specialist's wealth volatility increases dramatically, and more so when the specialist's wealth further shrinks. As a result, the volatility of the specialist's consumption growth rises with the tightness of the intermediation constraint, and the precautionary savings effect increases his demand for the riskless bond. Second, as specialist wealth falls, households withdraw equity from intermediaries and channel these funds into the riskless bond. The extra demand for bonds from both specialist and households lowers the equilibrium interest rate.

The pattern of decreasing interest rate presented in Figure 6 is consistent with a “flight to quality.” Households withdraw funds from intermediaries and increase their investment in bonds in response to negative price shocks. This disintermediation leaves the intermediaries more vulnerable

to the fundamental asset shocks.

3.8 Illiquidity and Correlation

In the capital constrained region, an individual specialist who may want to sell some risky asset faces buyers with reduced capital. Additionally, since households reduce their (indirect) participation in the risky asset market, the set of buyers of the risky asset effectively shrinks in the constrained region. In this sense, the market for the risky asset “dries up.” On the other hand, if a specialist wished to sell some bonds, then the potential buyers include both specialists as well as households. Thus the bond is more liquid than the risky asset.

There are further connections we can draw between low intermediary capital and aggregate illiquidity periods. As we have already seen, a negative shock in the constrained region leads to a rise in risk premia, volatility, and fall in interest rate. In this subsection, we show that our model also generates increasing comovement of assets that many papers have documented as an empirical regularity during periods of low aggregate liquidity (see, e.g., Chordia, Roll, and Subrahmanyam, 2000). We illustrate this point through two examples.

Example 1: Orthogonal Dividend Process

We introduce a second asset held by the intermediaries.¹⁶ The asset is in infinitesimal supply so that the endowment process and the equilibrium wealth process for specialists is unchanged. We assume that the dividend on this second asset is:

$$\frac{d\hat{D}_t}{\hat{D}_t} = gdt + \sigma dZ_t + \hat{\sigma} d\hat{Z}_t = \frac{dD_t}{D_t} + \hat{\sigma} d\hat{Z}_t.$$

Here, Z_t is the common factor modeled earlier; and \hat{Z}_t is a second Brownian motion, orthogonal to Z_t , which captures the asset’s idiosyncratic variation. Put differently, this second asset is a noisy version of the market asset.

¹⁶If the asset was traded by both households and specialists then its introduction will have an effect on equilibrium, since the market is incomplete. However, introducing an intermediated asset will not alter the equilibrium.

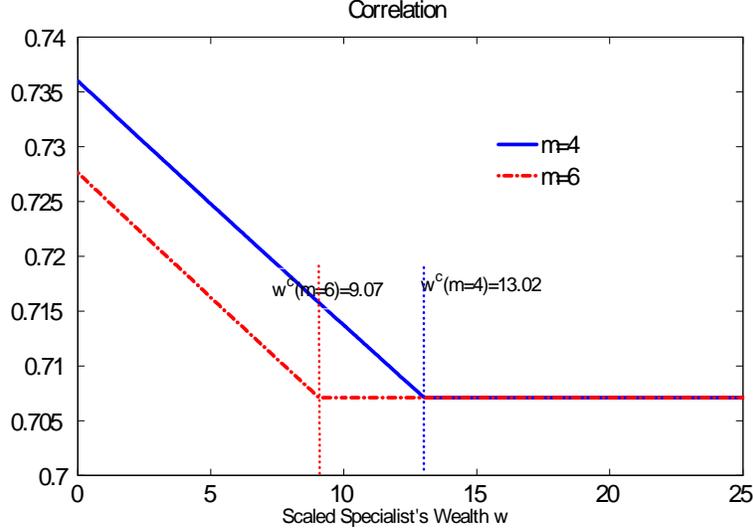


Figure 7: The correlation between the market return and the return on an individual asset, $\text{corr}(dR_t, \hat{dR}_t)$, is graphed against the scaled specialist wealth w for $m = 4$ and 6 . The constrained (unconstrained) region is on the left (right) of the threshold w^c . Other parameters are $g = 1.84\%$, $\sigma = 12\%$, $\rho = 1\%$, $\rho^h = 1.67\%$ (see Table 1), and $\hat{\sigma} = 12\%$.

We can show that the price of this second asset is given by,¹⁷

$$\hat{P}_t = \hat{D}_t \frac{P_t}{D_t} = \hat{D}_t \left[\frac{1}{\rho^h} + \left(1 - \frac{\rho}{\rho^h}\right) w_t \right]. \quad (17)$$

Consider the correlation between dR_t and the return \hat{dR}_t on the second asset:

$$\text{corr}(dR_t, \hat{dR}_t) = \frac{1}{\sqrt{1 + (\hat{\sigma}/\sigma_R)^2}}.$$

In the unconstrained region, since σ_R is constant, the correlation is constant. But, in the constrained region, as σ_R rises, the common component of returns on the two assets becomes magnified, causing the assets to become more correlated. We graph this state-dependent correlation in Figure 7, where we simply take $\hat{\sigma} = \sigma$.

Example 2: Liquidation-sensitive Asset

¹⁷The steps to show this result are as follows. For the market asset, the Euler equation for the specialist is,

$$(mf_t - \rho)dt - E_t \left[\frac{dW_t}{W_t} \right] + \text{Var}_t \left[\frac{dW_t}{W_t} \right] + E_t [dR_t] = \text{Cov}_t \left[\frac{dW_t}{W_t}, dR_t \right].$$

Since, $\frac{d\hat{P}_t}{\hat{P}_t} = \frac{dP_t}{P_t} + \frac{d(\hat{D}_t/D_t)}{D_t/D_t} = \frac{dP_t}{P_t} + \hat{\sigma}d\hat{Z}_t$, we have $\hat{dR}_t = \frac{\hat{D}_t}{P_t} + \frac{d\hat{P}_t}{\hat{P}_t} = dR_t + \hat{\sigma}d\hat{Z}_t$. Substituting this expression into the Euler equation, and using the result that $E_t [\hat{\sigma}d\hat{Z}_t] = 0$, we find that the Euler equation is satisfied for the pricing function in (17).

The preceding example illustrates how the risk-price of a common dividend rises during crises periods and causes increased comovement in asset prices. Another mechanism for comovement that is often emphasized by observers centers on forced liquidations by constrained intermediaries. The following example illustrates this case.

Normalize the initial date as time 0 with the state pair $(W_0, D_0 = 1)$. Consider an (infinitesimal) asset that pays off X_T at the maturity date T , where the dividend is state-contingent, i.e., $X_T = X(W_T, D_T)$. We are interested in how the economy-wide shocks drive the asset price, when the asset is subject to forced liquidation. A simple way to explore this idea is to assume that this dividend $X(W_T, D_T)$ is received only if the economy-wide intermediary capital W_T at the maturity date is above a minimum threshold \underline{W} ; below this threshold the asset pays less, which we normalize to zero for simplicity.

To this end, we study a liquidation-sensitive zero-coupon bond, with the state-contingent payoff as

$$X(W_T, D_T) = \begin{cases} 1 & \text{if } W_T > \underline{W}; \\ 0 & \text{otherwise.} \end{cases}$$

For example, this asset reflects an investment-grade corporate bond or a mortgage backed-security that is at low risk during normal times. However, during a period of low intermediation capital, the asset value is determined by an exogenous fire-sale value, which we have normalized to be zero. Denote the time-0 price of this liquidation-sensitive asset as $Q_0(W, D) = Q_0(W_0, 1)$, which is simply the time-0 present value of $X(W_T, D_T)$ under the pricing kernel in this economy. We focus on the constrained region to illustrate the interesting dynamics in this example, and perform the computations numerically.

The value of this liquidation-sensitive zero-coupon bond $Q_0(W_0, 1)$ varies with the state of the economy. Interestingly, the sign of the correlation switches depending on the state. Consider a negative shock to this economy causing intermediary capital W to fall. A lower W leads to a lower interest rate in the constrained region, which in turn leads to a higher bond price. This *interest rate effect* generates a negative correlation between the returns of our (intermediated) market risky asset and the liquidation sensitive asset. The pattern prevails for high levels of intermediary capital.

When the intermediary capital W_0 is sufficiently low, i.e., in the vicinity of the liquidation boundary \underline{W} , an opposite *liquidation effect* kicks in. Under this effect, a negative shock makes forced liquidation more likely, and the price of the liquidation-sensitive asset falls. As a result, there is positive correlation between the market return and the asset return.

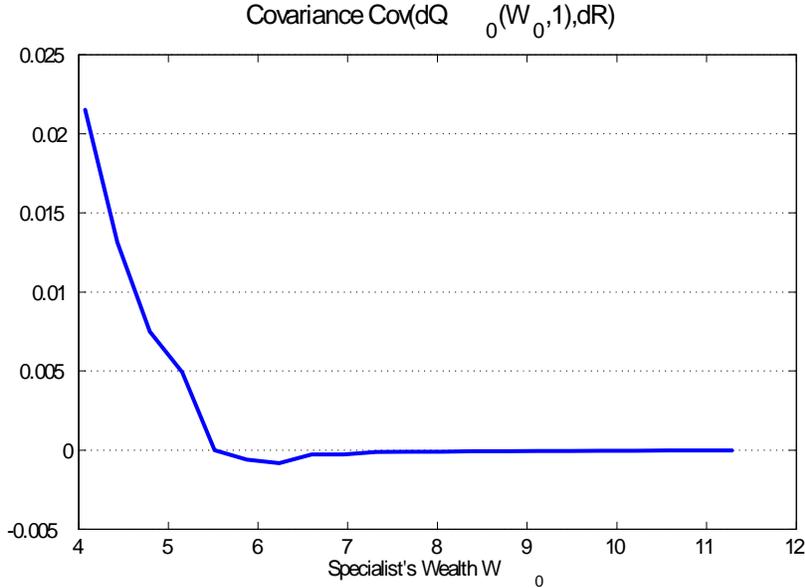


Figure 8: The instantaneous covariance between the returns of intermediated market asset and the liquidation-sensitive asset, i.e., $\text{cov}(dR, dQ_0(W_0, 1))$. The x -horizontal is the time-0 specialist's wealth $w = W_0$, as we normalize $D_0 = 1$. We take $m = 4$, so the capital constraint binds at $w^c = 13$. The liquidation threshold is $\underline{W} = 3.57$. Other parameters are $g = 1.84\%$, $\sigma = 12\%$, $\rho = 1\%$, and $\rho^h = 1.67\%$ (see Table 1).

Figure 8 graphs these two effects by considering the instantaneous covariance between $dQ_0(W_0, 1)$, and the market return dR_t . When the scaled specialist's wealth is high, the correlation is negative, although close to zero for the parameters in our example. The covariance becomes more negative as W_0 shrinks due to the interest rate effect. Finally, when W_0 falls around \underline{W} (which is 3.57 in our example), the liquidation effect dominates, and the liquidation-sensitive asset comoves with the intermediated market asset.

3.9 Liquidity factor

A number of recent papers have provided evidence that asset returns are driven by both a market factor and a liquidity factor (see Amihud, 2002, Acharya and Pedersen, 2003, Pastor and Staambaugh, 2003, and Sadka, 2003). These empirical papers suggest that the marginal investor is particularly averse to holding assets with low payoffs in episodes of low aggregate liquidity, consistent with the logic of our model in which a capital-constrained intermediary is the marginal investor. We next show that our model can indeed rationalize a liquidity factor.

To show this result, we extend the model to include a second source of shocks. Since the

model is driven by a single source of uncertainty (the one-dimensional Brownian motion governing dividends), changes in both the risky asset price and intermediary capital are perfectly correlated. This makes it difficult to clarify the role of a liquidity factor for asset returns separate from the market factor. In our model, the intermediated-market factor is itself affected by liquidity.

We consider the following thought experiment. We perturb our model by adding a second shock process that is orthogonal to dividends, but directly affects intermediary capital. We then trace the effects of this second factor on asset returns. This exercise gives us some understanding of the separate, additive, role of intermediary capital risk, without working out a full-blown two-factor model.

We imagine that nature randomly redistributes a small amount of wealth between specialists and households (we can also think about this as tax/transfer by the government). The redistribution to specialists is $\sigma_l D_t dZ_{l,t}$, where $Z_{l,t}$ is orthogonal to Z_t and $\sigma_l/\sigma \rightarrow 0$. Thus this second shock process is small compared to the primary dividend process. Without loss of generality we assume $\sigma_l > 0$. The shock is scaled by D_t for ease of comparison.

The transfer adds a random disturbance to the wealth evolution equations for specialists and households. But since $\sigma_l \rightarrow 0$, both agents' consumption policies of eating a fraction of wealth is close to optimal (up to an error in the order of σ_l^2).¹⁸ Then, the same equilibrium argument as offered earlier implies that the risky asset price is $P_t = \frac{D_t}{\rho^h} + \left(1 - \frac{\rho}{\rho^h}\right) W_t$.

We write the diffusion terms on the specialist's percentage wealth evolution dW/W as,

$$\sigma_W dZ_t + \sigma_{W,l} dZ_{l,t}. \quad (18)$$

We focus on the diffusion terms since we are interested only in understanding how the new factor affects risk premia. Given the risky asset price P_t , we can write the diffusion terms on dR_t as,

$$\begin{aligned} & \frac{1}{P_t} \left(\sigma \frac{D_t}{\rho_t} + \left(1 - \frac{\rho}{\rho^h}\right) W \sigma_W \right) dZ_t + \frac{1}{P_t} \left(1 - \frac{\rho}{\rho^h}\right) W \sigma_{W,l} dZ_{l,t} \\ & = \sigma_R dZ_t + \sigma_{R,l} dZ_{l,t}. \end{aligned}$$

¹⁸Consider the households. Their wealth evolution is $\frac{dW_t^h}{W_t^h} = -\rho^h dt + \widetilde{dR}_t^h - \sigma_l \frac{D_t}{W_t^h} dZ_{l,t}$, where \widetilde{dR}_t^h is the return on their optimally chosen asset portfolio. The Euler equation for the household is,

$$-\rho^h dt + E_t \left[\frac{dc_t^h}{c_t^h} \right] + Var_t \left[\frac{dc_t^h}{c_t^h} \right] + E_t[\widetilde{dR}_t^h] - Cov_t \left[\frac{dc_t^h}{c_t^h}, dR_t^h \right] = 0.$$

If we substitute the consumption policy of $c_t^h = \rho^h W_t^h$ into this equation, we find that the equation is satisfied up to an error of $\sigma_l^2 \left(\frac{D_t}{W_t^h}\right)^2$. In this sense, as $\sigma_l \rightarrow 0$, the consumption policy is near optimal. A similar argument applies to the specialists, so that we take their consumption policy to be unchanged as well.

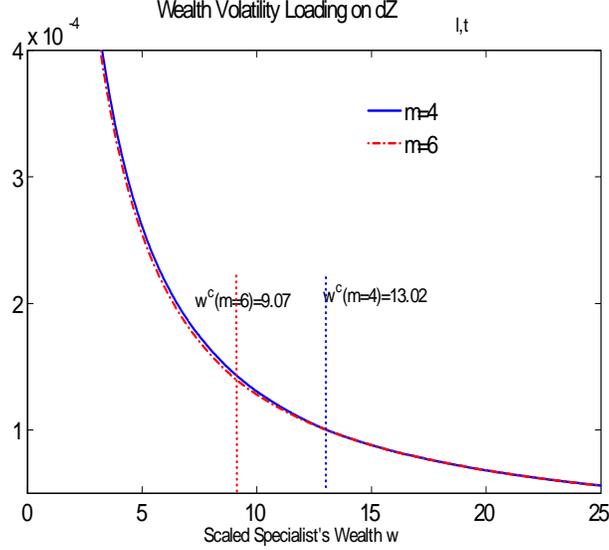


Figure 9: The volatility loading of the specialist's wealth on the second shock $dZ_{l,t}$ is graphed against the scaled specialist wealth w for $m = 4$ and 6 . The constrained (unconstrained) region is on the left (right) of the threshold w^c . Other parameters are $g = 1.84\%$, $\sigma = 12\%$, $\rho = 1\%$, $\rho^h = 1.67\%$ (see Table 1), and $\sigma_l = \frac{\sigma}{100} = 0.12\%$.

There are two contributing sources in the diffusion term of the specialist's wealth growth dW/W : first, the specialist has a position of α_t in the risky asset; second, the specialist's wealth has a percentage shock of $\frac{\sigma_l D_t}{W_t} dZ_{l,t} = \frac{\sigma_l}{w_t} dZ_{l,t}$ due to wealth redistribution. Then we can go back through the steps of matching coefficients on the diffusion terms in (18), and find that σ_W is the same as previously derived (see Proposition 4 in Section 3.3), while

$$\sigma_{W,l} = \begin{cases} \sigma_l \left(\frac{1}{w_t} + \rho^h - \rho \right) & \text{in the unconstrained region,} \\ \frac{(1+m)\rho^h}{\rho+m\rho^h} \frac{\sigma_l}{w_t} & \text{in the constrained region.} \end{cases}$$

Figure 9 graphs $\sigma_{W,l}$ as a function of w_t . The key point to note is that $\sigma_{W,l}$ rises faster in the constrained region, even though the exogenous shock process is itself not a function of the tightness of constraints. The exogenous shock is amplified as constraints tighten.

The risk premium on the risky asset now reflects the new shock process,¹⁹

$$E_t [dR_t] - r_t dt = \frac{\sigma_W^2}{\alpha_t} dt + \frac{\sigma_{W,l} - \frac{\sigma_l}{w_t}}{\alpha_t} \sigma_{W,l} dt. \quad (19)$$

¹⁹The risk premium on the risk source of $dZ_{l,t}$ is just $\sigma_{R,l}\sigma_{W,l}$, where $\sigma_{R,l}$ is the loading of dR_t on $dZ_{l,t}$. Now we have $\sigma_{W,l} = \alpha_t \sigma_{R,l} + \frac{\sigma_l}{w_t}$, i.e., the percentage change of specialist's wealth due to $dZ_{l,t}$ equals the sum of two contributions from holding risky asset and the wealth redistribution. Therefore $\sigma_{R,l} = \frac{\sigma_{W,l} - \frac{\sigma_l}{w_t}}{\alpha_t}$, and the result follows.

The last term on the right-hand side is new relative to our previous expressions. This term reflects the effect of shocks to intermediary capital on the market's expected return. The shock $dZ_{l,t}$ has no direct effect on dividends and hence on prices of risky assets; rather, it affects prices through a liquidity channel. Endogenously the risk aversion of the specialist is affected by changes in w_t , and hence the asset prices and consumption are affected by $dZ_{l,t}$.

We next consider the pricing of a second asset in order to derive a cross-sectional asset pricing model. We introduce an infinitesimal amount of asset- j whose return process dR_t^j has diffusion terms,

$$\sigma^j dZ_t + \sigma_l^j dZ_{l,t}.$$

We can rearrange this expression, and write the above diffusion as the sum of the loading on the market return, and the loading on the new liquidity risk $dZ_{l,t}$ (recall the $\sigma_{R,1}$ expression derived in footnote 19):

$$\frac{\sigma^j}{\sigma_R} dR_t + \left(\sigma_l^j - \sigma^j \frac{\sigma_{W,l} - \frac{\sigma_l}{w_t}}{\sigma_W} \right) dZ_{l,t}.$$

We then arrive at the main result of this section.

Proposition 10 *The risk premium of asset- j satisfies a two-factor asset pricing model:*

$$E_t [dR^j] - r_t dt = \overbrace{\beta^{m,j} (E_t [dR_t] - r_t dt)}^{\text{Market}} + \overbrace{\beta^{liq,j} (\sigma_{W,l} dt)}^{\text{Liquidity}}. \quad (20)$$

Here, $E_t [dR_t] - r_t dt$ is the risk premium for the market factor, and $\sigma_{W,l} dt$ is the risk premium for an asset with unit of loading on the liquidity factor $dZ_{l,t}$.

The first term on the right hand side of (20) is the return on asset- j for bearing market risk, with its sensitivity β as

$$\beta^{m,j} = \frac{\sigma^j}{\sigma_R}.$$

For example, an asset with a higher σ^j loads more heavily on the market return, and therefore has a higher market β .

The second term on the right hand side of (20) is the additional return for bearing liquidity risk, with its sensitivity β as

$$\beta^{liq,j} = \sigma_l^j - \sigma^j \frac{\sigma_{W,l} - \frac{\sigma_l}{w_t}}{\sigma_W}.$$

We say ‘‘additional’’ return, because the market return already reflects a premium for bearing liquidity risk. Thus, the additional return comes from the added liquidity risk of asset- j . Assets whose returns load more heavily on the $dZ_{l,t}$ shock (i.e. higher σ_l^j) have a higher liquidity β .

4 Parameter Choices

Table 1 lists the parameter choices that are used in the graphs presented in the previous section. The parameter choices are based on matching a hedge fund crisis episode such as the 1998 crisis. We choose parameters so that the intermediaries of the model resemble a hedge fund, and the asset of the model reflects those of a hedge fund. It is worth stressing that the parameter choices should be viewed not as a precise calibration but rather as a plausible representation of a hedge fund crisis scenario.

Table 1: Parameters

Panel A: Intermediation		
m	Intermediation multiplier	4, 6
Panel B: Cashflows and Preferences		
g	Dividend growth	1.84%
σ	Dividend volatility	12%
ρ^h	Time discount rate of household	1.67%
ρ	Time discount rate of specialist	1%

The multiplier m parameterizes the intermediation constraint in our model. We choose m based on standard hedge fund contracts. We note that m measures the share of returns that specialists need to receive in order to satisfy the incentive compatibility constraints. Hedge fund contracts typically pay the manager 20% of the fund’s return in excess of a benchmark, plus 1 – 2% of funds under management (Fung and Hsieh, 2006). A value of $m = 4$ implies that the specialist’s inside stake is $1/5 = 20\%$. The 20% is an option contract so it is not a full equity stake. The percentage fee 1% is on funds under management and therefore grows as the fund is successful and garners more inflows. Thus, a 20% stake is in the range of parameters that may reasonably capture a hedge fund manager’s inside stake. We also present an $m = 6$ case to provide a sense as to the sensitivity of the results to the choice of m .

The risky asset cashflows are governed by growth rate g and volatility σ . Hedge funds invest in a variety of complex investment strategies each with their own cashflow characteristics. We apply the model to fit an amalgam of these strategies, rather than any single type of hedge fund. As a benchmark for such an amalgamate strategy, we use the aggregate stock market and set $\sigma = 12\%$ and $g = 1.84\%$. As another benchmark, Chan, et al. (2005) report the volatility of returns on different categories of hedge funds, finding standard deviations ranging between 3% to 17%. They also note that these numbers underestimate the true volatility of returns, because the underlying

assets of hedge funds are illiquid and there is evidence that hedge funds smooth reported returns.

We set ρ and ρ^h to match a riskless interest rate in the unconstrained region around 1%. The ratio of ρ to ρ^h measures the ratio of the lowest value of P/D to the highest value of P/D (reached at the two endpoints where $W^h = 0$ and $W = 0$). We set this ratio to be 60%. This choice is based on thinking about a “liquidation” value for the risky asset, and loosely, from considering the Warren Buffett/AIG/Goldman Sachs bid for the LTCM portfolio. This bid was reported to be \$4 billion for a 90% equity stake, suggesting a liquidation value of \$4.44 billion for LTCM’s assets. LTCM was said to have lost close to \$3 billion of capital at the time of this bid, suggesting that LTCM lost 40% of its value to arrive at the liquidation price of \$4.44 billion. Our calculation here is clearly rough.

5 Conclusion

We have presented a model to study the effects of capital constraints in the intermediary sector on asset prices. Capital effects arise because (1) households lack the knowledge to participate in the risky asset; and, (2) intermediary capital determines the endogenous amount of exposure that households can achieve to the risky asset. The model builds on an explicit microeconomic foundation for intermediation. The model is also cast within a dynamic economy in which one can articulate the dynamic effects of capital constraints on asset prices. We show that the model can help to explain the behavior of asset markets during aggregate liquidity events and can rationalize a liquidity factor for asset returns.

There are a number of interesting directions to take this research. First, the model we have presented has a degenerate steady-state distribution, which means that we cannot meaningfully simulate the model. For typical parameter values, the specialist will eventually end up with all of the wealth. This aspect of the model is well-known and arises in many two-agent models (see Dumas, 1989, for further discussion). He and Krishnamurthy (2008) analyze a closely related model, which has a non-degenerate steady-state distribution. That model is sufficiently complex that it does not allow for the simple closed-form solutions of this paper. There, we solve the model numerically and simulate to compute a number of asset pricing moments.

A second avenue of research is to expand the number of traded assets. Currently the only non-intermediated asset in the model is the riskless bond. However, in practice, even unsophisticated households have the knowledge to invest in many risky assets directly, or to invest in low

intermediation-intensive assets such as an S&P500 index fund. It will be interesting to introduce a second asset in positive supply in which households can directly invest, and study the differential asset pricing effects across these different asset classes. This exercise seems particularly relevant in light of the evidence in the fall of 1998 that it was primarily the asset classes invested in by hedge funds that were affected during the crises. We intend to investigate these issues more fully in future work.

References

- [1] Acharya, Viral and Lasse Pedersen, 2005, Asset Pricing with Liquidity Risk, *Journal of Financial Economics*, 77, 375-410.
- [2] Allen, Franklin and Douglas Gale, 1994, Limited Market Participation and Volatility of Asset Prices, *American Economic Review* 84, 933-955.
- [3] Allen, Franklin and Douglas Gale, 2005, Financial Intermediaries and Markets, *Econometrica*, 72(4), 1023-1061.
- [4] Allen, Franklin and Gary Gorton, 1993, Churning Bubbles, *Review of Economic Studies*, 60, 813-836.
- [5] Amihud, Yacov, 2002, Illiquidity and Stock Returns: Cross-Section and Time-Series Effects, *Journal of Financial Markets*, 5, 31-56.
- [6] Basak, Suleyman and Domenico Cuoco, 1998, An Equilibrium Model with Restricted Stock Market Participation, *Review of Financial Studies*, 11, 309-341.
- [7] Bates, David, 2003, Empirical Option Pricing: A Retrospection, *Journal of Econometrics*, 116, 387-404.
- [8] Berndt, Antje, Rohan Douglas, Darrell Duffie, Mark Ferguson, and David Schranz, 2004, Measuring Default Risk Premia from Default Swap Rates and EDFs, mimeo, Stanford University.
- [9] Bondarenko, Oleg, 2004, Market Price of Variance Risk and Performance of Hedge Funds, mimeo, University of Illinois-Chicago.
- [10] Brennan, Michael, 1993, Agency and Asset Pricing, working paper, UCLA.
- [11] Brunnermeier, Markus and Lasse Pedersen, 2008, Market Liquidity and Funding Liquidity, forthcoming, *Review of Financial Studies*.
- [12] Chan, Nicholas, Mila Getmansky, Shane Haas, and Andrew Lo, 2005, Systemic Risk and Hedge Funds, with Nicholas Chan, to appear in M. Carey and R. Stulz, eds., *The Risks of Financial Institutions and the Financial Sector*. Chicago, IL: University of Chicago Press.

- [13] Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam, 2000, Commonality in liquidity, *Journal of Financial Economics*, 56, 3-28.
- [14] Collin-Dufresne, Pierre, Robert Goldstein, and J. Spencer Martin, 2001, The Determinants of Credit Spread Changes, *The Journal of Finance*, 66, 2177-2207.
- [15] Dasgupta, Amil, Andrea Prat, and Michela Verardo, 2005, The Price Impact of Institutional Herding, working paper, LSE.
- [16] Dow, James and Gary Gorton, 1994, Arbitrage Chains, *Journal of Finance* 49, 819-849.
- [17] Duffie, Darrell, 2007, Princeton Lectures in Finance: Capital Mobility and Asset Pricing, lecture notes on web.
- [18] Dumas, Bernard, 1989, Two-Person Dynamic Equilibrium in the Capital Market, *The Review of Financial Studies*, 2(2), 157-188.
- [19] Froot, Kenneth, 2001. The Market for Catastrophe Risk: A Clinical Examination, *Journal of Financial Economics* 60, 529-571.
- [20] Froot, Kenneth A., and Paul G. O'Connell, 1999, The Pricing of US Catastrophe Reinsurance in *The Financing of Catastrophe Risk*, edited by Kenneth Froot, University of Chicago Press.
- [21] Fung, William and David Hsieh, 2006, Hedge Funds: An Industry in Its Adolescence, *Federal Reserve Bank of Atlanta Economic Review*, (Fourth Quarter), 91, 1-33.
- [22] Gabaix, Xavier, Arvind Krishnamurthy and Olivier Vigneron, 2007, Limits of Arbitrage: Theory and Evidence from the MBS market, *Journal of Finance*, 62(2), 557-596.
- [23] Garleanu, Nicolae, Lasse Heje Pedersen, and Allen M. Poteshman, 2005, Demand-Based Option Pricing, Wharton mimeo.
- [24] Gromb, Denis and Dimitri Vayanos, 2002, Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs, *Journal of Financial Economics* 66, 361-407.
- [25] Grossman, Sanford J. and Zongquan Zhou, 1996, Equilibrium Analysis of Portfolio Insurance, *Journal of Finance* 51, 1379-1403.
- [26] Guerrieri, Veronica and Peter Kondor, 2008, Fund Managers and Defaultable Debt, working paper, Chicago GSB.

- [27] He, Zhiguo and Arvind Krishnamurthy, 2008, Intermediated Asset Prices, working paper, Chicago GSB and Northwestern University.
- [28] Holmstrom, Bengt and Jean Tirole, 1997, Financial Intermediation, Loanable Funds, and the Real Sector, *Quarterly Journal of Economics* 112, 663-691.
- [29] Krishnamurthy, Arvind, 2003, Collateral Constraints and the Amplification Mechanism, *Journal of Economic Theory*, 111(2), 277-292.
- [30] Kyle, Albert S. and Wei Xiong, 2001, Contagion as a Wealth Effect, *Journal of Finance* 56, 1401-1440.
- [31] Liu, Jun and Francis Longstaff, 2004, "Losing Money on Arbitrages: Optimal Dynamic Portfolio Choice in Markets with Arbitrage Opportunities" , *Review of Financial Studies* 17(3), 611-641.
- [32] Lucas, Robert E., Jr, 1978, Asset Prices in an Exchange Economy, *Econometrica*, 46, 1429-1445.
- [33] Mankiw, Gregory N, and Stephen Zeldes, 1991, The Consumption of Stockholders and Non-stockholders, *Journal of Financial Economics*, 29, 97-112.
- [34] Mas-Colell, A., M. Whinston, and J. Green. *Microeconomic Theory*. Oxford University Press, 1995.
- [35] Mitchell, Mark, Lasse Pedersen, and Todd Pulvino, 2007, Slow Moving Capital, *American Economic Review Papers and Proceedings*, forthcoming.
- [36] Pastor, Lubos and Robert Stambaugh, 2003, Liquidity Risk and Expected Stock Returns, *Journal of Political Economy*, 111, 642-685.
- [37] Pavlova, Anna and Roberto Rigobon, 2008, The Role of Portfolio Constraints in the International Propagation of Shocks, *Review of Economic Studies*, forthcoming.
- [38] Sadka, Ronnie, 2006, Momentum and Post-Earnings-Announcement Drift Anomalies: The Role of Liquidity Risk, *Journal of Financial Economics* 80, 309-349.
- [39] Shleifer, Andrei and Robert W. Vishny, 1997, The Limits of Arbitrage, *Journal of Finance* 36, 35-55.

- [40] Vayanos, Dimitri, 2005, Flight to Quality, Flight to Liquidity, and the Pricing of Risk, working paper, LSE.
- [41] Vissing-Jorgensen, Annette, 2002, Limited asset market participation and the elasticity of intertemporal substitution, *Journal of Political Economy*, 110(4), 825-853.
- [42] Warther, Vincent, 1995, Aggregate mutual fund flows and security returns, *Journal of Financial Economics*, 39, 209-235.
- [43] Xiong, Wei, 2001, Convergence Trading with Wealth Effects: An Amplification Mechanism in Financial Markets, *Journal of Financial Economics* 62, 247-292.