

National Centre of Competence in Research
Financial Valuation and Risk Management

Working Paper No. 482

Capital Mobility and Asset Pricing

Darrell Duffie

Bruno Strulovici

First version: May 2008
Current version: May 2008

This research has been carried out within the NCCR FINRISK project on
“Equilibrium Asset Pricing”

Capital Mobility and Asset Pricing*

Darrell Duffie
Stanford University

Bruno Strulovici
Oxford University

Preliminary and Incomplete

May 1, 2008

Abstract

We present a model for the equilibrium movement of capital between markets. Two markets with symmetrically distributed risks are distinguished only by the levels of capital invested in the two markets. That market with the greater amount of capital earns lower conditional mean returns. Intermediaries optimally trade off the costs of intermediation against fees that depend on the gain they can offer to investors for moving their capital to the market with the higher mean return. In equilibrium, the speed of adjustment of returns and of capital are increasing in the degree to which capital is imbalanced between the two markets.

*We are grateful for reactions at Oxford University, the 2007 Gerzensee European Summer Symposium in Financial Markets, the University of Toulouse, and especially for comments from Bruno Biais, Glen Weyl, Jean Tirole, and Jean-Charles Rochet. We are especially thankful for the research assistance of Sergey Lobanov.

1 Introduction

We present a model for the equilibrium movement of capital between markets. The equilibrium mean rates of return across the markets are distinguished only by the levels of capital invested in the respective markets. As a matter of supply and demand within each market, that market with the greater amount of capital earns lower conditional mean returns. Given a sufficient disparity in the capital levels in the markets, an intermediary finds it optimal to search for investors in the market with “surplus” capital and offer them the opportunity to move their capital to the other market. The intermediary charges suppliers of capital a fee that is based on their gain from the move. In equilibrium, the greater the relative difference in capital levels across the markets, the more intensive are the intermediary’s efforts to re-balance the distribution of capital across the markers, and the greater the rate of convergence of the two mean rates of return. Absent search frictions, capital is equally distributed across the markets. A shock to the relative levels of capital to the two markets causes a relative improvement in the mean rate of return offered to investors in the market whose capitalization has been reduced.

Our general objective, closely related to that of Gromb and Vayanos (2002) and Gromb and Vayanos (2007), is to provide a framework that can be used to analyze equilibrium asset pricing in a setting of imperfect capital mobility. In order to achieve analytical tractability, our model is highly stylized in several regards. Foremost among these, as in Gromb and Vayanos (2007), some investors in each market are assumed to have no mobility.

Our work is motivated in part by the example of limited mobility of capital into reinsurance markets, documented by Froot and O’Connell (1997), who write: “Our results suggest that capital market imperfections are more important than shifts in actuarial valuation for understanding catastrophe reinsurance pricing. Supply, rather than demand, shifts seem to explain most features of the market in the aftermath of a loss.” In subsequent work, Froot (2000) continues: “We ... find the most compelling (evidence) to be supply restrictions associated with capital market imperfections and market power exerted by traditional reinsurers.”

2 The Market Setting

This section presents a stylized model for the endogenous adjustment of capital and risk premia across markets. There are three types of agents: (i) local hedgers; (ii) speculators, who provide risk-bearing to hedgers in each of two local markets; and (iii) intermediaries (or asset managers) who provide the fee-based service of moving speculators' capital from one market to the other. In equilibrium, speculators move their capital, subject to intermediation frictions, into that market with the higher premium for the same risk. We provide an illustrative example based on insurance markets.

We fix a probability space (Ω, \mathcal{F}, P) and a common information filtration $\{\mathcal{F}_t : t \geq 0\}$ satisfying the usual conditions.¹

In each of two financial markets, labeled a and b , a continuum of local risk-averse agents own short-lived risky assets that they are willing to sell at or above their respective reservations prices. Equivalently, they are willing to buy insurance contracts against the risks to which they are exposed from these assets. These “hedgers” are not mobile across markets. They can be viewed in this respect as relatively unsophisticated in the use of cross-capital-market transactions, or as having high transactions costs for trading outside of their local markets. Suppliers of capital, however, have access to cross-market trading, subject to intermediation frictions to be described. These suppliers of capital are risk-neutral, and can be viewed as speculators offering to bear the risk that hedgers desire to shed, in return for a risk premium. Because the total potential demand for risk bearing exceeds the available supply of capital to bear it, there are strictly positive equilibrium risk premia. One can imagine the suppliers of capital to be institutional or individual investors. In an insurance context, one might think of these speculators as stylized versions of the “Names” that supply risk bearing capacity to the reinsurance market that is known as “Lloyd’s of London.”

The levels of capital available in the two markets at time t are X_{at} and X_{bt} , respectively. Capital can be reinvested continually at the discretion of each provider of capital, that is, “rolled over” in the short-lived assets that are continually made available for sale

¹See, for example, Protter (2004) for the usual conditions and for other standard properties of stochastic processes to which we refer.

by hedgers. Each unit of capital that is currently invested in market i at time t is paid cash dividends at the going market “reset rate” $\pi(X_{it})$, where $\pi(\cdot)$ is a strictly decreasing continuous function. The payout rate $\pi(X_{it})$ is continually reset in auctions in which the supply and demand for the asset in market i are matched at each point in time. As the amount x of capital available to invest in the asset is increased, the reset rate declines. In the appendix, we provide an example in which $\pi(x)$ is the equilibrium insurance premium in a market with x units of insurance capital.

In return for the payout rate $\pi(X_{it})$, the provider of each unit of capital in market i agrees to absorb the risky increments of a payoff process ρ_i , which is Lévy, that is, has independently and identically distributed increments over non-overlapping time periods of the same length. (Examples include Brownian motions, Poisson processes, compound Poisson processes, and linear combinations of these.) The idea is that the short-lived risky asset pays $1 + d\rho_{it} + \pi(X_{it}) dt$ at time $t + dt$ per unit of capital invested at time t , in the instantaneous sense. More precisely, each unit of capital invested in market i at any time s , and rolled over continually, accumulates to W_t units of capital by time t , where $dW_t = W_{t-} d\rho_{it}$, and in the meantime generates cash flows at the rate $\pi(X_{it})W_t$. (The notation “ W_{t-} ” means the left limit of the path of W at time t , that is, the level just before any jump at time t .)

In the case of an insurance contract, for example, we can take ρ_i to be a compound Poisson that jumps down at the arrival times of loss events, and is otherwise constant. In this case, one unit of capital invested at time t pays the supplier of capital $1 + \pi(X_{it}) dt$ at time $t + dt$ (in the above sense) if there is no loss event, and if there is a loss event, has a recovery value of $1 + \Delta\rho_{it}$, where $\Delta\rho_{it}$ is the jump size. The jumps of ρ_i are bounded below by -1 , preserving limited liability. If the loss events have mean arrival rate η and the loss size distribution ν has mean $\bar{\nu}$, then the mean loss rate is $\eta\bar{\nu}$. In this case, as the amount x of capital gets large, the market clearing payout rate $\pi(x)$ will never go below $\eta\bar{\nu} + r$, where r is the time preference rate of the speculators. In any case, suppliers of capital optimally supply all of their local capital inelastically, so long as the mean rate of return is at least as large as their common time preference rate r .

In order to solve for simple examples of equilibrium capital dynamics, we will later pick parametric examples for $\pi(x)$, constructing the population of hedgers so as to support these examples.

As with typical asset-management contracts used by hedge funds and private equity partnerships, cash dividends are not re-invested into the capital pool. For us, this is merely a modeling convenience.

We assume that $\rho_i = \epsilon_i + \epsilon_c$, where the market-specific processes ϵ_a and ϵ_b , and the common component ϵ_c , are independent Lévy processes. We assume throughout that ϵ_a and ϵ_b have the same distribution, so that the two markets have identically and symmetrically distributed capital-gain processes. This simplifies the calculation of an equilibrium and has the further illustrative advantage that any differences in the conditional expected returns in the two markets are due solely to differences in the capital levels of the markets.

Without capital-market frictions, suppliers of capital would move capital between the markets so as to obtain the higher reset rate, and by virtue of their common capital movements would equate $\pi(X_{at})$ and $\pi(X_{bt})$ in equilibrium, and thereby equate X_{at} and X_{bt} at all times. Indeed, given the symmetrically distributed returns of the two markets, the suppliers of capital would do so even if they were risk-averse, provided that they have no other hedging motives.

Frictions in the movement of capital may, however, lead to unequal levels of capital in the two markets. If, for example, $X_{at} < X_{bt}$, then the conditional excess mean rate of return of the risky asset in market a exceeds that in market b by $\pi(X_{at}) - \pi(X_{bt})$, despite the identical idiosyncratic and systematic risks of the two assets. Whichever market has “too much capital” receives the lower risk premium.

A supplier of capital decides only how to deploy re-invested capital between the two markets, subject to the available trading technology. Letting C_t denote the net cumulative amount of capital moved by that supplier of capital from market a into market b through time t , the supplier’s levels of capital, W_{at}^C in market a and W_{bt}^C in market b , jointly satisfy

$$dW_{at}^C = W_{at-}^C d\rho_{at} - dC_t$$

and

$$dW_{bt}^C = W_{bt-}^C d\rho_{bt} + dC_t.$$

Capital can be moved only through the services of an intermediary, and at the times of contact with an intermediary, as will be explained. A model for a proportional

transactions-fee process K will be determined in equilibrium, once we introduce a model for intermediation of capital movements. A supplier of capital is infinitely-lived and discounts net cash flow at a given rate r , so has a utility of

$$E \left(\int_0^\infty e^{-rt} ([W_{at}^C \pi(X_{at}) + W_{bt}^C \pi(X_{bt})] dt - K_{t-} dC_t) \right).$$

A minor alteration of the model for randomly timed exit and entrance of capital suppliers would be equally tractable.² For simplicity, we have assumed that transactions costs are paid directly by suppliers of capital, and not deducted from the capital moved from market to market.

There is a continuum (a non-atomic measure space) of suppliers of capital. Each takes as given the total capital processes X_a and X_b of the respective markets as well as the proportional transactions-cost process K . Among other equilibrium consistency conditions, suppliers of capital form conjectures regarding (X_a, X_b, K) that are consistent in equilibrium.

Intermediaries contacts suppliers of capital in order to profit from fees for moving their capital from one market to another. In equilibrium, at any time, only suppliers of capital in that market with greater capital agree to have any of their capital moved to the other market. Because a supplier of capital has linear preferences and takes (X_a, X_b, K) as given, it is optimal when contacted to move either no capital or to move all capital to the other market. If he or she has any capital in the market with more total capital, then all of this investor's capital will be moved, provided the proportional transaction-costs process K is not too large, and this is the case in any equilibrium for our model, as we shall see once the model is completely specified. Thus, although we allow that a given supplier of capital may initially have non-zero capital in both markets, all of his or her invested capital will optimally be held in just one of the two markets at any time after the first time of contact with the intermediary.

We let $W_{ij}(t)$ denote the level of capital in market i of investor j at time t . Conditional on the intensity process λ with which suppliers of capital are individually contacted by intermediaries, it is assumed that suppliers are contacted independently of each

²For this, suppliers of capital would exit at independent exponentially distributed times, and consume their capital at exit. New suppliers of capital would appear in proportion to the current levels of capital. Any difference between exit and entrance rates would be subtracted from the proportional drifts of the capital accumulation processes X_a and X_b .

other. In a manner similar to that of Weill (2007), we appeal formally to the law of large numbers to calculate the aggregate rate of movement of capital. The total rate at which capital is moved from market a to market b is almost surely³

$$\int \lambda_t 1_{\{X_{at} > X_{bt}\}} W_{aj}(t) dm(j) = \lambda_t 1_{\{X_{at} > X_{bt}\}} \int W_{aj}(t) dm(j) \quad (1)$$

$$= \lambda_t 1_{\{X_{at} > X_{bt}\}} X_{at}, \quad (2)$$

where $m(\cdot)$ is the measure over the space of suppliers of capital. Likewise, the rate at which capital moves from market b to market a is $\lambda_t 1_{\{X_{bt} > X_{at}\}} X_{bt}$.

Given the total market intermediation contact intensity process λ and initial conditions for capital in each market, we let X_{it}^λ denote the total capital in market i at time t . Given an associated transaction-cost process K , the marginal value to a supplier of capital of an additional unit of capital in market i at time t is

$$\theta_{it}^\lambda = E \left(\int_t^\infty e^{-r(s-t)} \left[W_s \pi \left(X_{D(s),i}^\lambda \right) ds - W_{s-} K_{s-} dN_s \right] \middle| \mathcal{F}_t \right), \quad (3)$$

where, for each s , N_s is the cumulative number of switches back and forth between the two markets through time s by the holder of this unit of capital, and the market indicator $D(s)$ is a or b , depending on whether, at time s , the associated accumulated capital W_s is currently located⁴ in market a or b . The unit of capital thus accumulates according to

$$dW_s = W_{s-} d\rho_{D(s-)}(s),$$

with initial condition $W_t = 1$. The market-indicator process D is a marked point process with an initial condition at time t of $D(t) = i$, and with an intensity of jumping from market i to market j at time s of $\lambda_s 1_{\{X_{is}^\lambda > X_{js}^\lambda\}}$. In the equilibrium that we shall describe, the value of switching from market i to market j is strictly positive if and only if $X_{it}^\lambda > X_{jt}^\lambda$. In general, the switching value is

$$\phi_t^\lambda = \max(\theta_{at}^\lambda, \theta_{bt}^\lambda) - \min(\theta_{at}^\lambda, \theta_{bt}^\lambda).$$

We assume that, at each time t , intermediaries charge suppliers of capital some fraction $q \in [0, 1]$ of the gain ϕ_t^λ from switching each unit of capital. That is, $K_t^\lambda = q\phi_t^\lambda$. One

³ In spirit, we apply the results of Sun (2006), relying on pairwise independence and a particular style of product measure space for states of the world and suppliers of capital. We do not provide formal conditions for the same result in our particular setting.

⁴In equilibrium, investors always have all their capital in a single market, see Proposition 1.

can view q as the bargaining power of an intermediary. We do not endogenize an intermediary's bargaining power q , a weakness of our model. If we were to adopt the Nash bargaining solution, linear preferences would lead to the bargaining power $q = 1/2$.

To assume that a supplier of capital can move capital from one market to another only through intermediation is tantamount to an assumption that the alternative technologies for moving capital are prohibitively expensive. For this, it would be enough, in equilibrium, that for any alternative trading technology such as directly contacting and negotiating with hedgers, the proportional cost of moving capital exceeds the marginal value ϕ_t^λ of switching. This is a strong assumption that simplifies the model and its solution. We will calculate the marginal switching value in examples and show that it can be arbitrarily small depending on the parameters of the intermediation technology. So, the assumption that alternative capital movement technologies would not be used by suppliers of capital is reasonable in some circumstances.

Our model can also be generalized by supposing that each supplier of capital has an alternative technology by which opportunities to move capital to the other market arrive at random times, independent across suppliers of capital, with a constant mean arrival rate. This would cause only minor modifications to the structure of our model. We avoid it for simplicity. Obviously, the higher the mean arrival rates of these alternative capital-shifting opportunities, the less the average degree of imbalance of capital and mean asset returns between the two markets, and the smaller the role of the intermediaries.

An intermediary's rate of cost for applying contact intensity λ_t is assumed to be $c\lambda_t$, for some technological coefficient $c \geq 0$. This means, for example, that doubling the expected rate at which a unit mass of capital suppliers is contacted costs twice as much.⁵ We restrict λ to be a progressively-measurable process so that, at each time, the contact intensity can depend only on currently available information. The maximum feasible contact intensity of the market is assumed to be some $\bar{\lambda} \in (0, \infty)$.

⁵If the intermediary can direct search efforts to only those suppliers of capital currently in a given local capital market, then this search cost model can be viewed consistent with a contact technology in which the intermediary adjusts a "broadcast" intensity, for example adjusting the rate of purchase of advertisements or other forms of market-wide intermediation efforts. This differs from a model in which, for example, contacting twice as many individuals at a given intensity costs twice as much.

We will solve the model separately for the case a monopolistic intermediary, or a finite number n of identical oligopolistic intermediaries, or perfectly competitive intermediaries. Given an inverse-demand function $\pi(\cdot)$, which as explained in Appendix A can be engineered from primitive assumptions on the risk preferences of hedgers, the model is defined by $(\pi, \rho_a, \rho_b, r, q, c, \bar{\lambda}, n)$.

3 Equilibrium with Monopolistic Intermediation

We focus first on the monopolistic case, and generalize in the next section. The solution of the monopolistic case leads immediately to a solution for the oligopolistic case, via a simple equivalence result.

A monopolistic intermediary's total rate of fee revenue is $\lambda_t X_t^\lambda q \phi_t^\lambda$. This assumes that the intermediary and supplier of capital both correctly anticipate that the intermediary's future contact intensity is indeed given by the process λ . We will later impose this consistency property as part of the definition of an equilibrium.

Given the initial conditions $X_{a0}^\lambda = x_a$ and $X_{b0}^\lambda = x_b$, the intermediary's utility for a contact intensity process λ is

$$U(x_a, x_b, \lambda) = E \left(\int_0^\infty e^{-rt} \lambda_t [q \phi_t^\lambda \max(X_{at}^\lambda, X_{bt}^\lambda) - c] dt \right).$$

We restrict attention to intermediation policies that depend only on the current capital levels (X_{at}, X_{bt}) . The intermediary might otherwise prefer to commit once and for all time to a path-dependent intensity policy that could, at some future time, be dominated by another policy available at that time, given the current capital market conditions at that time. Our restriction rules this out. In work to appear, we show technical conditions under which, absent the ability to commit, intermediation intensity policies that depend only on current capital levels are indeed optimal.

That is, we consider only a search intensity process λ that is "Markov," in the sense that there is some measurable function $\Lambda : \mathbb{R}_+^2 \rightarrow [0, \bar{\lambda}]$ with the property that $\lambda_t = \Lambda(X_{at}^\lambda, X_{bt}^\lambda)$. Letting \mathcal{L} denote the space of policies of this form, the intermediary's problem is

$$\sup_{\lambda \in \mathcal{L}} U(x_a, x_b, \lambda). \tag{4}$$

An equilibrium is defined to be an intermediation intensity process λ^* that attains the supremum (4). This definition includes consistency with the optimality for suppliers of capital to move their capital for the stipulated fee when contacted by the intermediary, and includes consistency of the conjectured and actual dynamics of capital movements and of search intensity. A given equilibrium uniquely determines the dynamics of capital levels and risk premia in the two markets.

Under restrictive parametric assumptions, we will characterize equilibria and solve for equilibria in illustrative examples.

3.1 Problem Setup

Given the intermediary's contact intensity process λ , we let

$$\begin{aligned} X_t &= \max(X_{at}, X_{bt}) \\ Y_t &= \min(X_{at}, X_{bt}), \end{aligned}$$

and

$$\begin{aligned} d\rho_t^X &= 1_{\{X_{at} > X_{bt}\}} d\rho_{at} + 1_{\{X_{at} \leq X_{bt}\}} d\rho_{bt} \\ d\rho_t^Y &= 1_{\{X_{at} \leq X_{bt}\}} d\rho_{at} + 1_{\{X_{at} > X_{bt}\}} d\rho_{bt}, \end{aligned}$$

The Lévy processes (ρ^X, ρ^Y) have the same joint distribution as primitive processes (ρ_a, ρ_b) . Fixing a Markov intermediation intensity process λ of the form $\lambda_t = \Lambda(X_t, Y_t)$, the associated state process (X, Y) is Markov, satisfying the stochastic differential equation

$$\begin{aligned} dX_t &= -\Lambda(X_t, Y_t)X_t dt + X_t d\rho_t^X \\ dY_t &= \Lambda(X_t, Y_t)X_t dt + Y_t d\rho_t^Y. \end{aligned}$$

Consistent with the insurance example, we suppose that ρ_a and ρ_b are of the form $\rho_{it} = \mu t + \epsilon_{it}$, where μ is a constant and ϵ_a and ϵ_b are independent compound Poisson processes with common jump intensity η and a given jump-size probability distribution ν .

In order to obtain the simplification associated with homogeneity, we suppose that the inverse demand function $\pi(\cdot)$ is of the form $a + kx^{-\gamma}$ for positive constants a , k , and γ .

As explained in the insurance setting of Appendix A, this can be arranged by suitable assumptions on the cross-sectional distribution of dis-utility for insurance premia and losses. Because the constant a is common to the two markets, it has no effect on benefits to switching capital, and can be taken to be zero without loss of generality.

Without loss of generality, we can take $k = 1$, by re-scaling. That is, the solution of the intermediary for coefficients (k, c) is the same as the solution for $(1, c/k)$. Because the intermediary has linear time-additive preferences and because of the homogeneity of π , and therefore of ϕ^λ , the ratio $Z = X/Y$ of total capital in the over-capitalized market to total capital in the under-capitalized market determines the optimal intermediation intensity. Thus, the assumption of independence of ϵ_a and ϵ_b is without loss of generality, for common jumps would have no effect on the ratio of X to Y . (The sole exception is for the case of common jumps with a jump-size distribution that supports -1 , in which case there is a non-zero probability that X_t and Y_t can be zero simultaneously. We rule out this exception.)

The proportional payoff processes ρ_a and ρ_b could also be given a common Brownian component without affecting our analysis, for this also has no effect on the relative proportions of capital in the two markets. Cases with market-specific Brownian components are analyzed in Appendix J. Likewise, the constant drift rate μ plays no role in the analysis of optimal intermediation, and can be taken to be zero without loss of generality for purposes of determining equilibrium intermediation policies. The effect of non-zero μ on actual capital levels can be reintroduced later with the scaling by $e^{\mu t}$ of both X_t and Y_t .

We begin our analysis with the simple case in which the jump-size distribution ν places all mass at -1 , meaning complete loss of invested capital at an event. We later relax this to random partial recovery, for which we offer an illustrative numerical example. For the zero-recovery case, a loss event in the market with less capital would cause the capital ratio X_t/Y_t to jump to $+\infty$. While we allow this formally, the analysis can be done similarly in terms of the ratio Y_t/X_t , which remains in $[0, 1)$ almost surely, and our results apply with only notational changes. Provided the initial conditions include a strictly positive amount of capital in at least one market, the probability that X_t and Y_t reach zero at the same time is 0. The partial-recovery case that we consider, for which the jump size distribution places zero probability on -1 , has strictly positive capital levels in both markets at all times after time zero, given a strictly positive level

of capital in at least one of the markets at time zero.

Let $G(X_t, Y_t)$ and $H(X_t, Y_t)$ denote the present values to suppliers of capital of the marginal future cash flows per unit of capital held at time t in the over-capitalized and under-capitalized markets, respectively. Subject to the usual smoothness and integrability conditions, Itô's formula and the definition (3) of the value of a unit of capital held in market i imply that these functions satisfy the coupled equations

$$\begin{aligned} 0 = & -rG(x, y) + \pi(x) - G_x(x, y)x\Lambda(x, y) + G_y(x, y)x\Lambda(x, y) \\ & + (1 - q)\Lambda(x, y)(H(x, y) - G(x, y)) - \eta G(x, y) + \eta(G(x, 0) - G(x, y)) \end{aligned}$$

and

$$\begin{aligned} 0 = & -rH(x, y) + \pi(y) - H_x(x, y)x\Lambda(x, y) + H_y(x, y)x\Lambda(x, y) \\ & + \eta(G(y, 0) - H(x, y)) - \eta H(x, y). \end{aligned}$$

When moving a unit of capital, the intermediary gets a fraction q of the gain in present value to the supplier of capital, which is

$$\phi_t^\lambda = F^\Lambda(X_t, Y_t) \equiv H(X_t, Y_t) - G(X_t, Y_t).$$

The intermediary's value function is thus

$$V(x, y) = \sup_{\Lambda} E \left(\int_0^\infty e^{-rt} \Lambda(X_t, Y_t) (X_t q F^\Lambda(X_t, Y_t) - c) dt \right).$$

3.2 The Bellman Equation

Taking $F^\Lambda(x, y)$ as given, the associated Hamilton-Jacobi-Bellman (HJB) equation is

$$\begin{aligned} 0 = \sup_{\ell \in [0, \bar{\lambda}]} & \{-rV(x, y) - V_x(x, y)\ell x + V_y(x, y)\ell x \\ & + \eta[V(y, 0) + V(x, 0) - 2V(x, y)] + \ell(xqF^\Lambda(x, y) - c)\}. \end{aligned} \quad (5)$$

This is not a traditional stochastic control problem, in that the running reward to the intermediary includes the effect of $F^\Lambda(x, y)$, which is determined by the intermediation policy function Λ itself. This leads to an additional optimality consistency condition that we will show our solution satisfies.

The homogeneity of π implies that H and G are homogeneous of degree $-\gamma$. As a result, $G(z, 0) = g_0 z^{-\gamma}$ for some positive constant g_0 to be determined. Let $f(z) = F^\Lambda(x, 1)$. Using the calculations

$$\begin{aligned} F_x^\Lambda(x, y) &= y^{-\gamma-1} f' \left(\frac{x}{y} \right) \\ F_y^\Lambda(x, y) &= (-\gamma) y^{-\gamma-1} f \left(\frac{x}{y} \right) - x y^{-\gamma-2} f' \left(\frac{x}{y} \right), \end{aligned}$$

we can recover F^Λ from the solution f of the ordinary differential equation

$$\begin{aligned} -r f(z) + (1 - z^{-\gamma}) - z \Lambda(z, 1) f'(z) + (-\gamma f(z) - z f'(z)) \Lambda(z, 1) z \\ -(1 - q) f(z) \Lambda(z, 1) + \eta (g_0 (1 - z^{-\gamma}) - 2f(z)) = 0. \end{aligned} \quad (6)$$

The boundary condition is $f(1) = 0$, since this corresponds to the case $x = y$.

We now check that f is nonnegative. We can rewrite (6) as

$$(r + 2\eta + \Lambda(z, 1)(\gamma z + (1 - q))) f + z(1 + z) \Lambda(z, 1) f' = (1 + \eta g_0)(1 - z^{-\gamma}) \quad (7)$$

Because the righthand side is positive, f or f' must be positive. This implies that f cannot cross 0 from above. Hence, f must be positive on some interval of the form (z, ∞) and non-positive on $[1, z]$ for some level z . It remains to show that $z = 1$. Because $f(1) = 0$, this implies that the intermediary does not search at this level (since $c > 0$). That is, $\Lambda(z, 1)$ vanishes on a neighborhood of 1. From (7), this implies that f is positive on that neighborhood, which concludes a proof of the following result.

PROPOSITION 1 *Given any intermediation policy Λ , $f(z)$ is positive for $z > 1$. That is, suppliers of capital in the over-capitalized market optimally accept the offer to move all of their capital out of the over-capitalized market whenever given the opportunity.*

3.3 Trigger Intermediation Solution

We now solve for the equilibrium intermediation policy for the special case in which $\pi(x) = a + k/x$. As we have explained, we can without loss of generality for purposes of solving for equilibrium intermediation take $a = 0$ and, by scaling, take $k = 1$. In this case, F is homogeneous of degree -1 and therefore V is homogeneous of degree 0, that is, $V(x, y) = v(x/y)$ for some function v . In particular, given the function

f determining intermediation fees, the policy Λ achieving the supremum of the HJB equation (5) must be also be homogeneous of degree 0; that is, $\Lambda(x, y) = L(x/y)$ for some function L . Because the fee function f depends on the intermediary's policy L , we have a fixed point problem, to find a pair (f, L) such that: i) given f , L is the intermediary's optimal policy, and ii) given the policy L , f is the fee function determined by (6). Moreover, among pairs (f, L) satisfying these two conditions, L must also lead to maximal intermediation utility.

In the appendix (Proposition 5), we show that any such equilibrium must be of the “bang-bang” form

$$\Lambda(x, y) = 0, \quad x < Ty, \quad (8)$$

$$\Lambda(x, y) = \bar{\lambda}, \quad x \geq Ty, \quad (9)$$

for some trigger ratio $T \geq 1$ of the capital level in the over-capitalized market to the capital level in the under-capitalized market. This trigger policy is illustrated in Figure 1. A standard martingale-based verification argument implies that, given a function f , any pair (v, λ) that satisfies the homogenized version (16) of the HJB equation (5), is indeed a best response (optimal policy) given f . Our objective is to solve for the optimal trigger ratio T , which then determines the complete equilibrium behavior of the two markets.

In order to determine the constant g_0 , we note that, for any initial capital levels x and y ,

$$V(x, y) + xG(x, y) + yH(x, y)$$

is the present value $R(x, y)$ of the total future cash flows at rate $X_t\pi(X_t) + Y_t\pi(Y_t)$, to be divided among the intermediaries and the suppliers of capital, net of the present value $P_T(x, y)$ of the intermediary's expected discounted search costs over the infinite horizon, given trigger T . That is,

$$V(x, y) + xG(x, y) + yH(x, y) = R(x, y) - P_T(x, y).$$

Because of the homogeneity of π of degree -1 , we have $R(x, y) = 2/r$. The search-cost present value $P_T(1, 0)$ solves the equation $P_T(1, 0) = p + E[e^{-r\tau} P_T(1, 0)]$, where p is present value of search costs from time zero to the exponentially distributed time τ of the next loss event. From a calculation given in the appendix,

$$P_T(1, 0) = \frac{c\bar{\lambda}}{r} \left(1 - e^{-(2\eta+r)a(T)} \right), \quad (10)$$

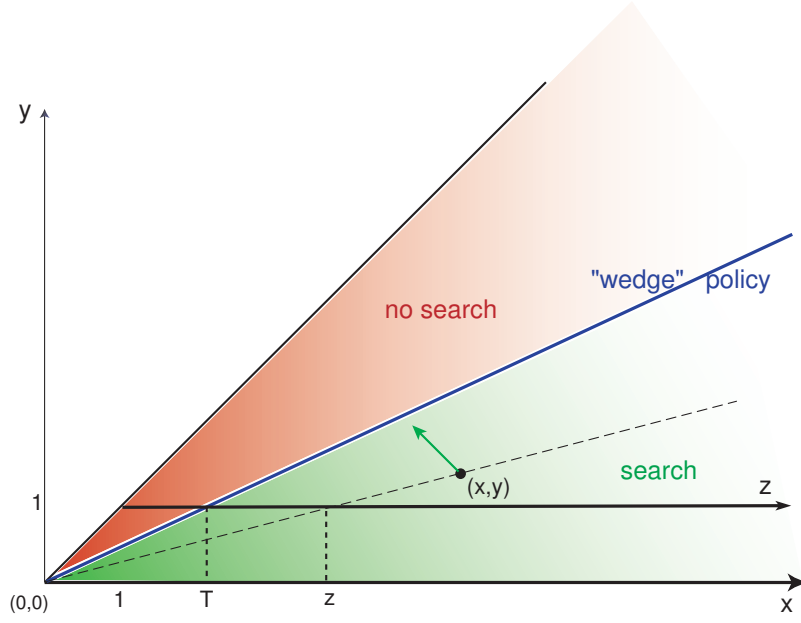


Figure 1: Optimal search policy and solution homogeneity.

where $a(T) = \log(1 + 1/T)/\bar{\lambda}$. We therefore have

$$g_0 = \frac{2}{r} - \frac{c\bar{\lambda}}{r} \left(1 - e^{-(2\eta+r)a(T)}\right) - V(1, 0). \quad (11)$$

The equation for f reduces to

$$(r + 2\eta + \bar{\lambda}[(1 - q) + z])f(z) + \bar{\lambda}(1 + z)zf'(z) = (1 + \eta g_0) \left(1 - \frac{1}{z}\right), \quad z \geq T, \quad (12)$$

and

$$(r + 2\eta)f(z) = (1 + \eta g_0) \left(1 - \frac{1}{z}\right), \quad z \in [1, T]. \quad (13)$$

For $z \in [1, T]$, the solution is trivial:

$$f(z) = \frac{1 + \eta g_0}{r + 2\eta} \left(1 - \frac{1}{z}\right). \quad (14)$$

In particular, we verify that $f(1) = 0$, consistent with the observation that the net present value of moving capital from one market to the other is 0 when the levels of capital in the two markets are the same.

We can re-write (12) as

$$(a + z)f(z) + z(1 + z)f'(z) = \left(1 - \frac{1}{z}\right)b, \quad z \geq T, \quad (15)$$

where $a = (r + 2\eta + (1 - q)\bar{\lambda})/\bar{\lambda}$ and $b = (1 + \eta g_0)/\bar{\lambda}$.

Letting $v(z) = V(z, 1)$, the HJB equation reduces to

$$0 = \sup_{\ell \in [0, \bar{\lambda}]} \{-rv(z) - \ell z v'(z) - \ell z^2 v''(z) + 2\eta[v_0 - v(z)] + (qzf(z) - c)\ell\}, \quad (16)$$

where $v_0 = V(y, 0) = V(x, 0)$. Therefore,

$$v(z) = v_1, \quad z \in [1, T],$$

where

$$v_1 = \frac{2\eta}{r + 2\eta}v_0 < v_0, \quad (17)$$

and

$$\kappa v(z) + v'(z)z(1 + z) = d + qzf(z), \quad z \geq T, \quad (18)$$

where $\kappa = (r + 2\eta)/\bar{\lambda}$ and $d = (2\eta v_0 - c\bar{\lambda})/\bar{\lambda}$.

In the appendix, we prove the following result.

PROPOSITION 2 *For any trigger policy, the value function v is everywhere nondecreasing, and increasing on $[S, \infty)$, where S is the trigger level.*

Combining (18), the equation obtained by differentiating (18), as well as the equation (15) for f , yields the second-order linear ordinary differential equation for v :

$$\alpha v(z) + (\beta + 2z)z(1 + z)v'(z) + z^2(1 + z)^2v''(z) = \omega + \delta z, \quad z \geq T, \quad (19)$$

where $\alpha = (a - 1)\kappa$, $\beta = (a + \kappa)$, $\omega = d(a - 1) - qb$, and $\delta = qb$. We bear in mind that some of the coefficients of this equation depend on a constant to be determined, $v_0 = V(1, 0)$.

The smooth-pasting condition $v'(T) = 0$ implies the trigger capital ratio

$$T = 1 + \frac{c(r + 2\eta)}{(1 + \eta g_0)q}. \quad (20)$$

A proof of the following result guaranteeing unique solutions for T is found in the appendix.

PROPOSITION 3 (UNIQUENESS) *There is at most one threshold T satisfying (11), (12), (13), and (20).*

This analysis leads to the following characterization of optimality, which includes the result that in the absence of search costs, the intermediary does not exploit his position to restrict movement of capital, but rather provides maximal intermediation, nevertheless generating fee income from his or her imperfect ability to instantaneously move capital from one market to the other due to the upper bound $\bar{\lambda}$ on contact intensity.

PROPOSITION 4 *Suppose that the payout rate function $\pi(\cdot)$ is a constant plus a function that is homogeneous of degree -1 , and that optimal intermediation is characterized by the HJB equation and smooth pasting. Then the intermediary has a unique optimal intermediation policy λ . This policy is inactive ($\lambda_t = 0$) whenever the ratio of capital levels in the two markets is between $1/T$ and T , and is otherwise at full capacity ($\lambda_t = \bar{\lambda}$). The uniquely determined trigger ratio T is given by (20), where the constant g_0 is given by (11). If there is no intermediation cost ($c = 0$), then the intermediary always works at full capacity (that is, $T = 1$).*

Relation (20) also provides an upper bound on the trigger level T , in that

$$T \leq 1 + \frac{c(r + 2\eta)}{q}.$$

An algorithm for computing the constant g_0 , and thus T , is given in the appendix.

3.4 Partial Recovery

We now allow the jump-size distribution ν of the gain process ρ_i to be distributed on $(0, 1)$, the case of partial recovery from a loss event.

Subject to the usual smoothness and integrability conditions, Itô's formula and the definition (3) of the value of a unit of capital held in market i imply that the value functions G and H satisfy the following coupled equations, taking $1 - W$ to be a random

variable with the proportional event-loss probability distribution ν :

$$\begin{aligned}
0 &= -rG(x, y) + \pi(x) + G_x(x, y) - x\Lambda(x, y) + G_y(x, y)x\Lambda(x, y) \\
&\quad + \eta P(Wx < y)[E(H(y, xW) | Wx < y) - G(x, y)] \\
&\quad + \eta P(Wx \geq y)[E(G(Wx, y) | Wx \geq y) - G(x, y)] \\
&\quad + (1 - q)\eta(H(x, y) - G(x, y)) + \eta E(G(x, Wy) - G(x, y))
\end{aligned}$$

and

$$\begin{aligned}
0 &= -rH(x, y) + \pi(y) - H_x(x, y)x\Lambda(x, y) + H_y(x, y)x\Lambda(x, y) \\
&\quad + \eta P(Wx < y)[E(G(y, Wx) | Wx < y) - H(x, y)] \\
&\quad + \eta P(Wx \geq y)[E(H(Wx, y) | Wx \geq y) - H(x, y)] \\
&\quad + \eta E[(H(x, Wy) - H(x, y))].
\end{aligned}$$

In the appendix, our prior analysis of the HJB equation and smooth-fit condition is extended from the case of zero recovery to this partial-recovery setting, and a solution algorithm is given. The solution algorithm exploits the linearity of the differential equations for g , h , and v , which arise thanks to the special structure of our problem.

3.5 Numerical Illustration

This subsection presents an illustrative example for the case of partial recovery. We take the parameters $r = 0.04$, $\eta = 1.5$, $c = 0.04$, $\bar{\lambda} = 0.1$, $q = 1/30$, and beta distributed recovery on $(0, 1)$ with parameters $(5, 1)$. The equilibrium intermediation trigger ratio T of capital in the over-capitalized market to capital in the under-capitalized market is found numerically to be 1.465.

Figure 2 shows simulated sample paths of the capitalization ratio $Z_t = X_t/Y_t$ and the immediate return $f(Z_t)/g(Z_t)$ to a supplier of capital, before transactions fees, associated with switching. Figure 3 shows the present value, with one unit of capital in the under-capitalized market, of future cash flows to a provider of one unit capital in the over-capitalized market (net of fees), to a provider of one unit of capital in the under-capitalized market (net of fees), and to the intermediary (in the form of fees net of search costs). These are, respectively, $g(z)$, $h(z)$, and $v(z)$, and depend on the ratio

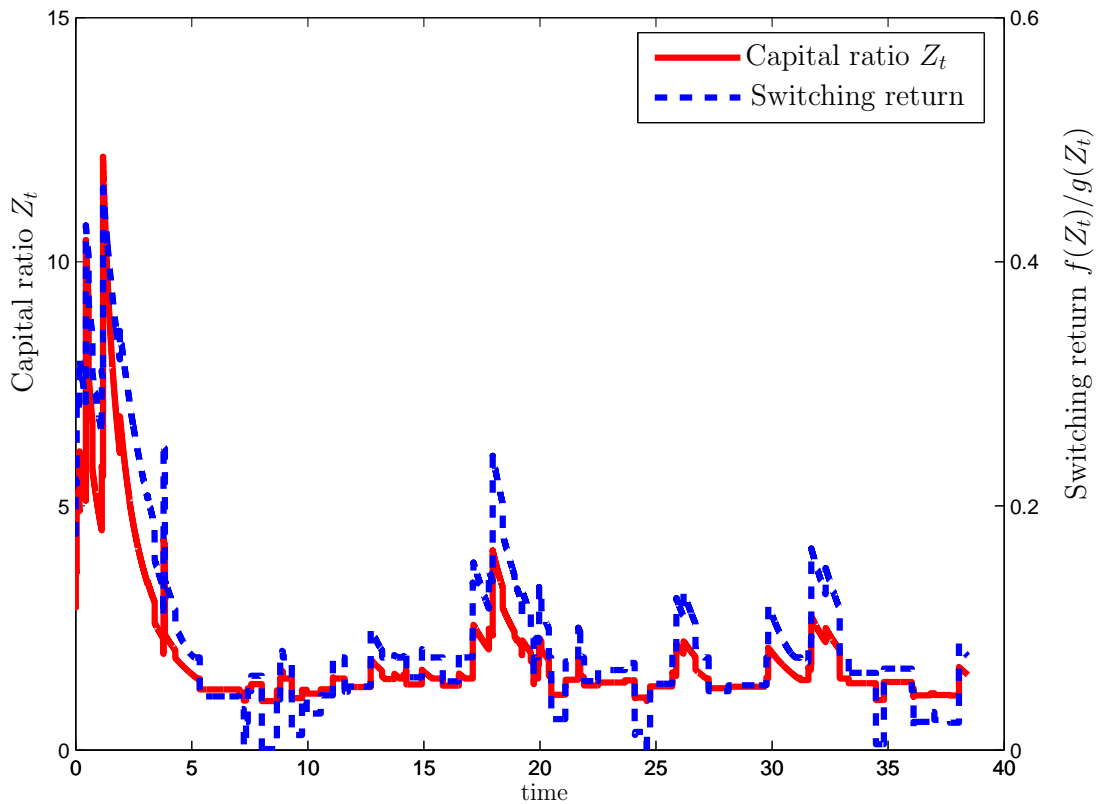


Figure 2: Simulated sample paths of the capitalization ratio $Z_t = X_t/Y_t$ and the return from switching $f(Z_t)/g(Z_t)$.

$z = x/y$ of the level of capital x in the over-capitalized market to the level y of capital in the under-capitalized market.

4 Oligopolistic and Competitive Intermediation

We now provide solutions for equilibria with oligopolistic or perfectly competitive markets for intermediation. Equilibrium trigger policies for the oligopolistic case can be translated directly from the case of monopolistic intermediation, by a change of variables. The solution is even simpler with perfectly competitive intermediaries, because they do not internalize any impact of their own intermediation intensity on the market dynamics.

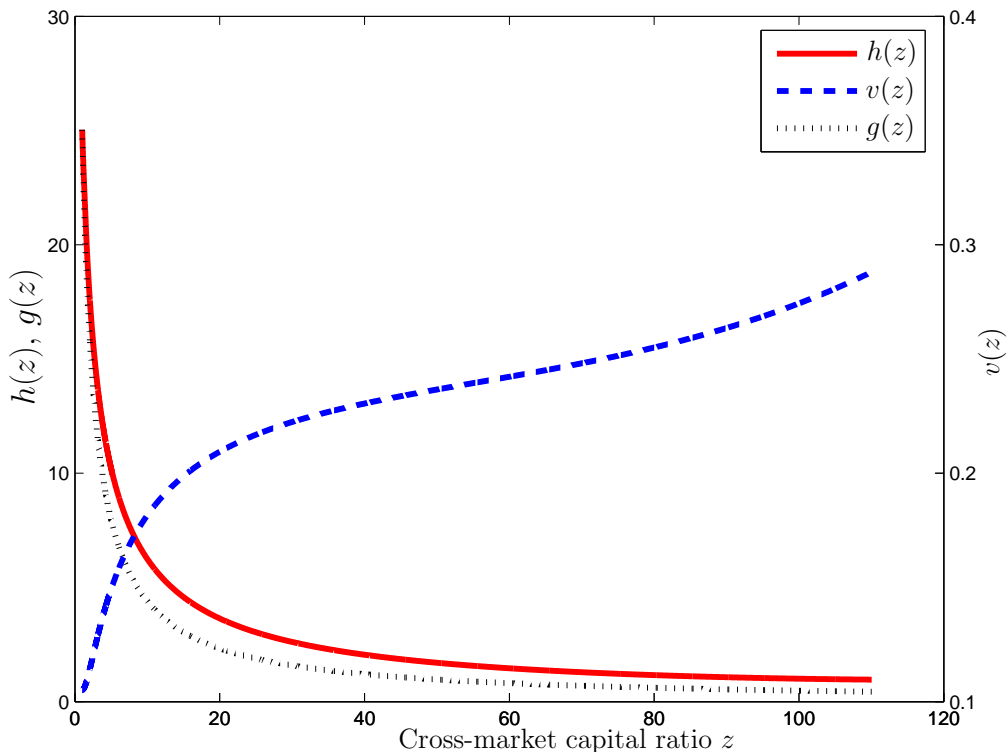


Figure 3: Value function $v(z)$ of the intermediary and the marginal values $g(z)$ and $h(z)$ of capital held in the over-capitalized and undercapitalized markets, respectively.

For the oligopolistic case, we take n identical intermediaries, each with an upper bound $\bar{\lambda}/n$ on intermediation intensity, and with the same proportional cost c of intermediation. The monopolistic case ($n = 1$) is the special case considered in the previous section. For the case of perfectly competitive intermediation, we take “ $n = \infty$ ” by considering a non-atomic measure space of intermediaries of total mass 1. Each intermediary has maximal intermediation intensity $\bar{\lambda}$, again providing for a market-wide total intermediation capacity of $\bar{\lambda}$. Thus, all cases have the same feasible market dynamics and costs.

We again consider only Markov equilibria. As opposed to the monopolistic case, however, these need not be the only form of no-commitment equilibria. There might exist equilibria with “punishment phases” involving path-dependent variation of intermediation intensities. We focus on equilibria where such cartel-like behavior is disallowed. Equilibrium incorporates the degree to which intermediaries internalize the impact of

their intermediation intensity on the heterogeneity of capital levels across markets.

For the oligopolistic case, for a Markov equilibrium in trigger strategies, each of the n intermediaries has a reduced value function v , with $v(z) = V(z, 1)$, solving the reduced HJB equation

$$0 = \sup_{\ell \in [0, \bar{\lambda}/n]} \left\{ -rv(z) + \left(\mu - \frac{(n-1)\bar{\lambda}1_{\{z \geq T\}}}{n} - \ell \right) zv'(z) \right. \\ \left. - \left(\mu + \frac{(n-1)\bar{\lambda}z1_{\{z \geq T\}}}{n} + \ell z \right) zv'(z) \right. \\ \left. + 2\eta[v_0 - v(z)] + (qzf(z) - c)\ell \right\},$$

where $v_0 = V(y, 0) = V(x, 0)$. This HJB equation reflects the presumption by the given intermediary that the $n - 1$ other intermediaries have adopted a specific trigger level T . An equilibrium condition is that the same trigger policy is optimal for the given intermediary.

The HJB equation is solved by

$$v(z) = v_1, \quad z \in [1, T],$$

where

$$v_1 = \frac{2\eta}{r + 2\eta} v_0 < v_0, \quad (21)$$

and

$$\kappa v(z) + v'(z)z(1+z) = d_n + \frac{qzf(z)}{n}, \quad z \geq T, \quad (22)$$

where

$$d_n = \frac{2\eta v_0}{\bar{\lambda}} - \frac{c}{n}.$$

From this reduced HJB equation, we see that that the dynamics of capital movements for an oligopolistic model with parameters (n, c, k) , for any number n of intermediaries, and intermediation effort cost coefficient c , and any payout rate scaling k , are equivalent to those for a market with corresponding parameters $(m, cn/m, kn/m)$, for any positive integer m . For example, the effect of doubling the number of intermediaries on improving capital mobility is precisely offset by halving the cost of intermediation and halving the asset payout rate to suppliers of capital. In particular, this implies that no further analysis is required, for we can translate the solution that we have already obtained for the monopolistic case $n = 1$, to the case of any number of intermediaries.

Thus, the equilibrium for the n -intermediary problem is again given by bang-bang control for all intermediaries, each exerting no effort when $Z_t < T$ and maximal intermediation intensity $\bar{\lambda}/n$ whenever $Z_t \geq T$, where the smooth-pasting condition $v'(T) = 0$ implies the trigger capital ratio

$$T = 1 + \frac{c(r + 2\eta)}{n(1 + \eta g_{0,n})q}, \quad (23)$$

where $g_{0,n} = G(1, 0; n)$ denotes the solution corresponding to the monopolistic case, for cost parameter c/n and asset payout rate kx/n .

The fraction q of gains from mobility that are allocated to intermediaries is taken to be fixed in this analysis, but would be better treated as a property of the entire market structure. Increasing the market power of intermediaries does not, in this setting, discourage suppliers of capital from moving their capital; they will do so in equilibrium so long as there are any gains from trade, that is, whenever given the chance to move out of the market with too much capital. Reducing the number of intermediaries while maintaining the total market intermediation capacity, represented by the technology coefficients $(c, \bar{\lambda})$, has an adverse impact on the outside option of a supplier of capital when bargaining with an intermediary. For example, if there are only two intermediaries, then at a contact with one of them, the contact intensity with the other one is $\bar{\lambda}/2$, whereas when there are 10 intermediaries, the contact intensities with the others total $9\bar{\lambda}/10$. In the sense of Rubinstein and Wolinsky (1987), this would dramatically improve the effective bargaining power of a supplier of capital.

For the extreme case of perfect competition (a continuum of identical intermediaries with total mass 1), each intermediary has no concern with respect to the impact of his or her efforts on the future levels of capital across markets, and merely solves the myopic intermediation problem

$$\sup_{\ell \in [0, \bar{\lambda}]} (qzf(z) - c)\ell. \quad (24)$$

It follows that the perfectly competitive equilibrium intermediation policy is $\lambda_t = 1_{\{Z_t \geq T\}} \bar{\lambda}/n$, where T solves $qTf(T) = c$. Because the solution of the equation for f involves T , this leaves a fixed-point calculation of T .

5 Discussion

This paper is motivated by empirical evidence that supply or demand shocks in asset markets, in addition to causing an immediate price response, also lead to adjustments over time in the distribution of capital across markets and adjustments over time in relative asset returns, in a way that reflects delays in the adjustments of investors' portfolios.

With trading frictions that delay portfolio adjustments, there can be periods of time over which assets with identical risks have different mean returns. More generally, there can be differences in mean returns across assets that are due not only to cross-sectional differences in “fundamental” cash-flow risks, but are also due to the degree to which the distribution of asset holdings across investors is not efficient (absent frictions). Empirical “factor” models of asset returns do not often account for factors related to the distribution of ownership of assets, or related to likely changes in the distribution of ownership.

We have examined a simple setting in which, absent trading frictions, risk-neutral investors would adjust their portfolios so as to achieve the highest possible mean return, thereby equating mean returns across assets. Because of trading frictions, however, investors cannot instantaneously adjust their portfolios. Over time, investors do make portfolio adjustments that cause mean returns across the markets to revert toward each other, until the next supply shock occurs. In our analysis, capital is mobilized through optimal intermediation. Other market microstructures would, however, lead to similar patterns of adjustment of capital and mean returns.

For example, in corporate bond markets, one observes large price drops and delayed recovery in connection with major downgrades or defaults, as described by Hradsky and Long (1989), when certain classes of investors have an incentive or a contractual requirement to sell their holdings. Dramatic examples of slow price adjustments to supply shocks in equity markets include those discussed by Coval and Stafford (2007), Andrade, Chang, and Seasholes (2005), and, with respect to index recomposition events, Shleifer (1986), Harris and Gurel (1986), Kaul, Mehrotra, and Morck (2000), Chen, Noronha, and Singhal (2004), and Greenwood (2005). Early work on the price impact of supply shocks include Holthausen and Mayers (1990) and Scholes (1972), who

consider the supply shocks to equity markets associated with block transactions.

Mitchell, Pedersen, and Pulvino (2007) document the effect on convertible bond hedge funds of large capital redemptions in 2005. Convertible bond prices dropped and rebounded over several months. A similar drop-and-rebound pattern was observed in connection with the LTCM collapse in 1998. Newman and Riersen (2003) show that large issues of credit-risky bonds temporarily raise credit spreads throughout the issuer's sector, because providers of liquidity such as underwriters and hedge funds bear extra risk as they search for long-term investors. They provide empirical evidence of temporary bulges in credit spreads across the European Telecom debt market during 1999-2002 in response to large issues by individual firms in this sector.

Our introduction uses the example of the market for catastrophe risk reinsurance. Sudden price surges, then multi-year price declines, follow sudden large aggregate claims against providers of insurance at times of major natural disasters, as explained by Froot and O'Connell (1999). The periods of price decline are typically accompanied by new entrants to the market, including hedge funds and other new re-insurers, whose capital has been mobilized by the price discrepancy, but not immediately. It takes time to set up a viable new provider of catastrophe risk insurance.

In all of these examples, the time pattern of returns or prices after a supply or demand shock reveals that the friction at work is not a simple transaction cost for trade. If that were the only friction, then all investors would immediately adjust their portfolios, or not, optimally, and the new market price and expected return would be immediately established, and remain constant until the next supply shock, absent changes in fundamentals. In all of the above examples, however, after the immediate price response, whose magnitude reflects the size of the shock and the degree of short-term price elasticity, there is a relatively lengthy period of time over which the price recovers somewhat (up after an adverse shock, down after a positive shock) reverting toward its new fundamental level. In the meantime, of course, additional shocks can occur, with overlapping consequences. The typical pattern suggests that the initial price response is larger than would occur with perfect capital mobility, and reflects the demand curve of the limited pool of investors that are immediately available to absorb the shock. The speed of adjustment after the initial price response is a reflection of the time that it takes more investors to realign their portfolios in light of the new market conditions, or for the initially responding investors to gather more capital.

In our model, delays in portfolio adjustments are due to the time that it takes for intermediaries to locate suitable suppliers of capital. This is only an abstraction, of course, but it may proxy for other delays, including time to educate investors about assets with which they have limited familiarity, time for contracting, and time for investors to dispose of their current positions (which involves similar delays). Some of the delays in practice could be due to time for information about investment opportunities to percolate through the population of suitable investors. Incorporating informational differences in our model would, however, involve substantial complications.

There is already a significant body of theory dealing with the implications of search frictions for asset pricing. For example, the implications of differences in search frictions across different asset markets are treated by Weill (2002) and Vayanos and Wang (2007). Duffie, Gârleanu, and Pedersen (2005) study the implications of search frictions in a single asset market with marketmaking. Earlier search-based models of intermediation include Rubinstein and Wolinsky (1987), Bhattacharya and Hagerty (1987), Moresi (1991), Gehrig (1993), and Yavaş (1996).

Related prior work on the implications of capital market frictions for asset pricing dynamics includes Basak and Croitoru (2000) and He and Krishnamurthy (2007).

In terms of general objectives and some model features, the study by Gromb and Vayanos (2007) is closely related to ours. Our respective approaches were developed independently. Common to the models in these two studies, certain local classes of investors are completely immobile, while arbitrageurs can work across markets driving returns toward fundamental levels, subject to frictions that prevent them from perfectly equating returns in the two markets. In Gromb and Vayanos (2007), the shock that initially drives a wedge between the prices of the assets in the two markets is a change in the direction of hedging demands of local investors. In our model, the shock is a loss in capital in one of the markets, relative to the other, due to an adverse realized asset return in one of the markets. The friction in Gromb and Vayanos (2007) is that arbitrageurs' positions must be collateralized, combined with a rule that prohibits the use of the asset from one of the markets as collateral against losses in the asset of other market. In our study, the friction is costly intermediation. We focus on the implications of optimal intermediation, under rational expectations by arbitrageurs and intermediaries of the implications of future optimal intermediation efforts for current intermediation effort, for current and future asset returns, and for intermediation pric-

ing. In our model, investors who contemplate moving their capital to the other market consider the likelihood that future shocks will at some point cause them to wish to move capital from the other market back to their current market, and the likelihood of delays in doing so, thereby lowering the value of switching to the other market in the first place.

Appendices

A An Insurance Example

We illustrate the model with an example motivated by catastrophe insurance contracts.

In a particular market, at each of the event times of a Poisson process J with a constant intensity η , a catastrophe occurs that causes losses throughout a population of consumers who are potential buyers of protection. Each of a continuum of consumers in the given insurance market has a property that experiences a loss at each catastrophe event. The losses of the consumers at a given event are identically and symmetrically distributed. The distribution of consumer losses at each catastrophe has the property that if a quantity x of the consumers have bought insurance at the time of the i -th catastrophe, then total claims of $x\zeta_i$ are paid by sellers of protection, where ζ_1, ζ_2, \dots is a sequence of independent random variables, identically distributed on $[0, 1]$, and independent of J . For this, it need not be the case that the damage of a particular consumer at the i -th event is equal to the average damage rate ζ_i , but we will assume so for notational simplicity only.

Each consumer chooses to be insured, or not, at each point in time, based on information available up to that time, but of course not including the information about loss events at that time. Whenever insured, the consumer pays premiums at the current rate p_t in his or her market, and is covered against damages in the event of a loss. Consumer α in a particular market has an insurance purchase policy process δ , valued in $\{0, 1\}$, providing total expected dis-utility of

$$E \left[\int_0^\infty e^{-\beta t} u_{1\alpha}(\delta_t p_t) dt + \sum_{i=1}^\infty e^{-\beta \tau_i} u_{2\alpha}((1 - \delta_t) \zeta_i) \right],$$

where τ_i is the time of the i -th catastrophe, β is a discount rate, and $u_{1\alpha}(\cdot)$ and $u_{2\alpha}(\cdot)$ are strictly decreasing dis-utility functions.

Given the additive nature of this utility, the insurance purchase policy δ minimizes total lifetime dis-utility if and only if, almost everywhere, δ_t solves, time by time, the

insurance purchase decision

$$\min_{\bar{\delta} \in \{0,1\}} u_{1\alpha}(\bar{\delta}p_t) + \gamma E[u_{2\alpha}((1 - \bar{\delta})\zeta_i)].$$

This problem is solved by 0 or 1 depending on whether p_t is greater or less than some reservation price p_α . We can therefore calculate, for each premium level \bar{p} , the total demand $\chi(\bar{p}) = M(\{\alpha : p_\alpha \leq \bar{p}\})$ for insurance, where $M(\cdot)$ is the measure on the space A of consumers in the market.⁶ Associated with the strictly decreasing demand function χ , assuming continuity, is a strictly decreasing and continuous inverse demand function, which we can express as $\pi(\cdot) + \eta E(\zeta_i)$, where $\pi(x)$ is the premium above expected loss. Alternative approaches, for example partial coverage, could be used to model the inverse demand function. In the end, to achieve a tractable solution of the intermediary's problem, we will make parametric assumptions for $\pi(\cdot)$ that can be justified by suitable construction of $u_{1\alpha}, u_{2\alpha}$, and the measure M .

The cumulative insurance claims process L for a quantity of one unit of insurance sold at all times is the compound Poisson is defined by $L_t = \sum_{i=1}^{J(t)} \zeta_i$. In order to offer one unit of insurance in a particular market, a seller of protection is required to commit one unit of capital. This is natural if one requires (say, as a regulatory matter) that insurance is default free, under the assumption that the essential supremum of the fractional event loss ζ_i is 1, which is the case in our illustrative numerical examples. (In any case, this supremum loss can be taken to be 1 without loss of generality by normalization of the definition of one unit of capital and of the associated construction of returns per unit of capital.) Thus, in a given market with x units of available insurance capital, the demand for insurance is $\chi(\pi(x) + \eta E(\zeta_i)) = x$, because the risk premium $\pi(x)$ is positive and providers of insurance capital have no better use for their capital at that moment in time.

Markets a and b are assumed to have identically distributed preferences among their respective pools of buyers of protection, and thus have the same inverse-demand function $\pi(\cdot)$. Their cumulative proportional claims processes L_a and L_b are identically distributed, but need not be independent. For example, some of the loss events could strike both markets.

While capital is deployed in insurance market i , it is subject to the cumulative pro-

⁶Because, for tractability, the supply x of capital is unbounded and the premium rate $\pi(x)$ is strictly decreasing in x , the total measure $M(A)$ of buyers of protection is taken to be infinite.

portional loss process L_i and is re-invested over time in a financial asset with Lévy cumulative return process R_i . Investment in this additional local asset is merely for generality, and plays no crucial role in the characterization of equilibrium, although it does have an impact on the specific equilibrium solution.

The total cumulative proportional accumulation process for capital in market i , before considering the movement of capital between the markets, is thus $\rho_i = -L_i + R_i$, where ρ_a and ρ_b have the joint distribution described earlier for the general model. Given the characteristics $(q, c, \bar{\lambda})$ of the intermediation of capital between the two markets, the primitives $(\pi, \rho_a, \rho_b, r, q, c, \bar{\lambda})$ of our basic model are fixed.

B Valuation of Search Fees

This appendix demonstrates the calculation (10) of the present value $P_T(1, 0)$ of search fees that was used in the text, in the case of no recovery at loss events.

Homogeneity implies that this present value returns to the same level at each loss event, so

$$P_T(1, 0) = p + E[e^{-r\tau} P_T(1, 0)], \quad (25)$$

where p is the present value of search costs until the next catastrophe and τ is the time until the first to arrive of the loss events in the two markets, which is exponentially distributed with parameter 2η . Starting with $X_0 = 1$ and $Y_0 = 0$, we have

$$dX_t = -\bar{\lambda}X_t 1_{\{Z_t > T\}} dt$$

and

$$dY_t = \bar{\lambda}X_t 1_{\{Z_t > T\}} dt.$$

This yields $X_t = e^{-\bar{\lambda}t}$ and $Y_t = 1 - e^{-\bar{\lambda}t}$, for $t < \tau$. The intermediary will stop searching at that time $a(T)$ at which $Z_{a(T)} = T$, so

$$\frac{e^{-\bar{\lambda}a(T)}}{1 - e^{-\bar{\lambda}a(T)}} = T.$$

This yields

$$a(T) = \frac{1}{\bar{\lambda}} \log \left(\frac{1+T}{T} \right).$$

The present value of search costs until the next loss event is

$$p = E \left[\int_0^{\min(a(T), \tau)} e^{-rt} \bar{\lambda} c dt \right] = \frac{\bar{\lambda} c}{r} \left(1 - E[e^{-r \min(a(T), \tau)}] \right).$$

Because τ is exponentially distributed with parameter 2η ,

$$E(e^{-r\tau}) = \frac{2\eta}{2\eta + r}$$

and

$$E[e^{-r \min(a(T), \tau)}] = \frac{2\eta}{2\eta + r} [1 - e^{-r(2\eta+r)a(T)}].$$

Substitution of these into (25) yields the result (10).

C Proof of Monotonicity of Value Function

The proof is based on two intermediate lemmas.

First, given the function f determining intermediation fees, let

$$\kappa(z) = \left(1 - \frac{1}{z} \right) \left(1 + \frac{\eta g_0}{r} + 2\eta \right) - f(z).$$

The first term of $\kappa(z)$ is the present value of arresting intermediation efforts from the point at which $Z_t = z$ until the next loss event occurs, given g_0 . In particular, suppose for a given z and a given policy L that $L(z) = 0$. Then $\kappa(z) = 0$. In particular, $\kappa(1) = 0$ (which can also be checked directly from the definition of κ and the fact that $f = 1$). Because the first term defining κ is increasing in z , we have $\kappa'(z) > 0$ provided that $f'(z) < 0$. Given a policy L , we will show that κ is nonnegative. In order to see this, we observe that for $z \geq 1$, (7) can be re-written as

$$L(z) [((1 - q) + z\gamma)f + z(1 + z)f'] = (r + 2\eta)\kappa(z). \quad (26)$$

We already know that $\kappa(1) = 0$. Since f is positive from Proposition 1, this implies that f' is negative, hence that $\kappa' > 0$ whenever κ is nonnegative. Therefore, κ cannot cross 0 from above, which shows our first lemma.

LEMMA 1 *For any policy, κ is everywhere nonnegative.*

This result is intuitive: The gain from moving capital, at any relative capital level z , given a trigger policy with trigger level S , is less than the gain from switching at the same level z if the intermediary immediately stops searching. Lemma 1 has an important consequence for the case $\gamma = 1$, namely that the rate at which fees are paid to the intermediary when he searches is increasing in z . That is, the more heterogeneous the markets, the higher is the intermediary's immediate profit from switching. Since this rate of fee payment, net of search costs, is $qzf(z) - c$, we must show that $zf(z)$ is increasing in z . We can re-write (26) when $\gamma = 1$ as

$$L(z)(1+z)(f+zf') = (r+2\eta)\kappa(z) + qf.$$

Because f is positive and κ is nonnegative, this implies that $f+zf'$ is positive whenever $L(z) > 0$, hence that $zf(z)$ is increasing in z . On any interval on which $L(z) = 0$, we have $f(z) = (1+\eta g_0)/(r+2\eta)(1-1/z)$, so f is increasing, and, a fortiori, so is $zf(z)$.

LEMMA 2 *For $\gamma = 1$ and any policy, the flow $zf(z)$ is increasing.*

We can now show monotonicity of v for any trigger policy. From (21), v is constant for $z \leq S$. Starting with some capital ratio $Z_0 = z > S$,

$$v(z) = E \left[\int_0^\tau e^{-rt} [qf(Z_t)Z_t - c] 1_{\{Z_t > S\}} dt + e^{-r\tau} v_0 \right],$$

where τ is the time of the next loss event and z_t denote heterogeneity at time $t < \tau$ starting at $z_0 = z$. The function $z \mapsto [qf(z)z - c] 1_{z > S}$ is nondecreasing in z from Lemma 2, and increasing for $z > S$. For $S < z < z'$, this implies that $v(z) < v(z')$ (because the event time τ has a distribution that does not on z or z'). This proves Proposition 2.

D Proof of Uniqueness of Trigger

PROOF OF PROPOSITION 3 Suppose $S < T$ both satisfy the equations of the proposition. From (20), this implies that $g_0^S > g_0^T$, using superscripts here and throughout to denote dependence on S and T . Optimality of S (resp.) T with respect to f^S (resp. f^T) implies that $qzf^S(z) - c - z(1+z)(v^S)'(z) > 0$ and $qzf^T(z) - c - z(1+z)(v^T)'(z) \leq 0$

for z in $(S, T]$. Since $(v^T)'(z) = 0$ for z on this interval, while $(v^S)'(z) \geq 0$ by Proposition 2, this implies that $\phi(T) < 0$. Subtracting equation (12) for T from the same equation for S for $z > T$ yields

$$(a + z)\phi + z(1 + z)\phi' = \alpha \left(1 - \frac{1}{z}\right), \quad z \in [S, T] \quad (27)$$

where

$$a = \frac{r + 2\eta}{\bar{\lambda}} + (1 - q) > 0$$

and

$$\alpha = \frac{\eta(g_0^T - g_0^S)}{\bar{\lambda}} < 0.$$

Because $\phi(T) < 0$, this⁷ implies that $\phi < 0$ for $z > T$, so that ϕ is everywhere negative.

By definition, g_0 is the present value per unit of capital of investors in the overcapitalized market when $x = 1$ and $y = 0$ (that is, when no investor is initially present in the small market). Therefore,

$$g_0 = \frac{2}{r} - \Phi_0, \quad (28)$$

where Φ_0 is the expected discounted stream of fees that investors will pay to the intermediary over the whole horizon (and $2/r$ is the expected discounted stream of dividends paid on both market, see conservation equation (11)). We have seen that $\phi < 0$, that is, $f^S(z) > f^T(z)$ for all z . This means that investors pay, for any z , more fees with S than with T for $z > T$. Moreover, investors pay fees (which are positive, from Proposition 1) for $z \in [S, T]$, for trigger S , whereas they pay nothing for trigger T . Therefore, $\Phi_0^S > \Phi_0^T$, which implies from (28) that $g_0^S < g_0^T$, a contradiction. ■

E Optimality of Trigger Policy

This appendix shows that for any equilibrium pair (f, μ) , μ must be a trigger policy. Suppose on the contrary that there exist levels $z < z' < z''$ such that $\mu(z) = \mu(z'') = 0$ but $\mu(z') > 0$. Then there must exist $S \in [z, z')$ and $T \in (z', z'']$ such that $\mu(S) = \mu(T) = 0$ and such that the intermediary is indifferent between several actions at those levels of market heterogeneity. Since the intermediary completely stops search at S and T this implies that $v(S) = v(T) = 2\eta/(2\eta + r) = v_1$, since the intermediary gets

⁷Indeed, $\phi(z) = 0$ implies $\phi'(z) < 0$, so ϕ cannot cross zero from below.

nothing until a catastrophe occurs (see (21)). Note that v_1 is the lowest possible level of the value function achievable by the intermediary, since he can always stop searching until the next catastrophe. Therefore, the value function reaches a minimum at S and T . This implies that $v'(S) = v'(T) = 0$. Equation (16) implies that $qf(z)z - c - v'(z)$ at any level z where the intermediary is indifferent. From above, this implies that $f(S)S = f(T)T$. Since the intermediary completely stops search at S and T , $f(S) = \alpha(1 - 1/S)$ and $f(T) = \alpha(1 - 1/T)$, where $\alpha = (1 + \eta g_0)/(r + 2\eta)$. Indeed, this is an investor's gain from switching from the overcapitalized market to the undercapitalized one, given that the intermediary stops searching until the next catastrophe occurs (see (14)). Therefore, $f(S)S \neq f(T)T$, a contradiction. This shows that the domain over which the intermediary stops searching is convex: it is of the form $[1, T]$ (since the intermediary does not search for $z = 1$). Over that interval, the intermediary either searches at full capacity or is indifferent between several search levels. Indifference however would imply that $v(z) = v_1$, i.e. that the intermediary reaches a global minimum at some level $z > T$ (since this is the value he would reach if he stopped searching until the next catastrophe, an action that is part of his indifference set, by (16)). This implies that $qzf(z) = c$ at such indifference level. From Lemma 2, we know however that $zf(z)$ is increasing. Since $qTf(T) = c$, this implies that $qzf(z) - c$ is positive for $z > T$, hence that the intermediary strictly prefers to search at full capacity for any level $z > T$. The analysis is summarized in the following statement.

PROPOSITION 5 *Any equilibrium policy is determined by a trigger level T such that $\lambda(z) = \bar{\lambda}1_{z>T}$.*

F Verification of Optimality of HJB Solution

In this section, we show that, given a fee function f , any bounded solution of the HJB equation (16) is optimal, that is, a best response to f .

G Algorithm for Trigger Calculation

In general, (20) provides the following fixed-point algorithm to compute the unknown trigger ratio T .

1. Start with some initial value for v_0 , which we call v^0 . From (11) and (20) we can then determine values for g_0 and T (it is easy to show that such values always exist). Call T^0 the corresponding trigger level. Furthermore (21), provides a corresponding value for $v(T^0)$.
2. Starting with the initial conditions $v(T^0)$ and $v'(T^0) = 0$, evaluate a candidate for $v(\infty) = \lim_{z \rightarrow \infty} v(z)$ by integration of the differential equation (19) on $[T^0, \infty)$.
3. The limit $v(\infty)$ corresponds to a new value for v_0 (since $v(\infty) = V(1, 0) = v_0$), which we call v^1 .
4. These steps are iterated until a fixed point is reached.

We have considered methods for speeding up the computation.⁸

H HJB Analysis with Partial Recovery

In this appendix, we analyze the Hamilton-Jacobi-Bellman equation for the case of general proportional losses.

Letting $K(\cdot)$ denote the cumulative distribution function of the event loss-fraction measure ν , $g(z) = G(z, 1)$ and $h(z) = H(z, 1)$ satisfy the coupled equations

$$\begin{aligned} (r + 2\eta + \Lambda(z, 1)z)g(z) + \Lambda(z, 1)(1 + z)zg'(z) &= \frac{1}{z} + \Lambda(z, 1)(1 - q)(h(z) - g(z)) \\ &+ \eta \left[\int_{1/z}^1 ug(uz) dK_u + \int_0^{1/z} \frac{1}{z} h\left(\frac{1}{uz}\right) dK_u + \int_0^1 \frac{1}{u} g\left(\frac{z}{u}\right) dK_u \right] \end{aligned} \quad (29)$$

⁸One can prove that $v' \sim \log(z)/z^2$ as z goes to ∞ . Unfortunately, this convergence rate is not particularly fast. A possibility is to integrate v numerically up to some value \hat{z} above which non-dominant terms in (19) are neglected. Above \hat{z} , the simplified equation becomes $2z^2v'(z) + z^3v''(z) = \delta$, which implies that $v'(z) = v'(\hat{z}) + \log(z)/z^2 - \log(\hat{z})/(\hat{z})^2$, which can be integrated to yield $v(z) - v(\hat{z})$ in closed form (up to the simplification of the equation).

and

$$\begin{aligned}
& (r + 2\eta + \Lambda(z, 1)z)h(z) + \Lambda(z, 1)(1 + z)zh'(z) \\
&= 1 + \eta \left[\int_{1/z}^1 h(uz) dK_u + \int_0^{1/z} \frac{1}{uz} g\left(\frac{1}{uz}\right) dK_u + \int_0^1 h\left(\frac{z}{u}\right) dK_u \right]. \quad (30)
\end{aligned}$$

As opposed to the case of total loss, these equations cannot be combined to yield a single equation for $f = h - g$, because of differing integrands.

Letting $v(z) = V(z, 1)$, the 0-homogeneity of V implies that the value after a loss event is $v(uz)$ if $ux \geq y$, $v(1/uz)$ if $ux \leq y$, and $v(z/u)$ if the loss occurs on the smaller market. The HJB equation is thus

$$\begin{aligned}
0 = \sup_{\ell \in [0, \bar{\lambda}]} & \left\{ -rv(z) - \ell zv'(z) - \ell z^2 v'(z) + \ell(qzf(z) - c) \right. \\
& \left. + \eta \left[\int_{1/z}^1 v(uz) dK_u + \int_0^{1/z} v\left(\frac{1}{uz}\right) dK_u + \int_0^1 v\left(\frac{z}{u}\right) dK_u - 2v(z) \right] \right\}. \quad (31)
\end{aligned}$$

The equation reduces to

$$(r + 2\eta)v(z) = \eta \left[\int_{1/z}^1 v(uz) dK_u + \int_0^{1/z} v\left(\frac{1}{uz}\right) dK_u + \int_0^1 v\left(\frac{z}{u}\right) dK_u \right], \quad z \in [1, T],$$

and

$$\begin{aligned}
& (r + 2\eta)v(z) + \bar{\lambda}(1 + z)zv(z)' = [qzf(z) - c]\bar{\lambda} \\
& + \eta \left[\int_{1/z}^1 v(uz) dK_u + \int_0^{1/z} v\left(\frac{1}{uz}\right) dK_u + \int_0^1 v\left(\frac{z}{u}\right) dK_u \right], \quad z \geq T. \quad (32)
\end{aligned}$$

The smooth-pasting condition, as for the case of no recovery at loss events, is

$$(1 + T)Tv'(T) = qTf(T) - c. \quad (33)$$

I Algorithm for Partial Recovery Model

This appendix includes an algorithm for solving the partial-recovery equations of the previous appendix. The algorithm exploits the linearity of the differential equations for g , h , and v , which arise thanks to the special structure of our problem.

I.1 Primitives

The parameters are $r, \eta, \bar{\lambda}, q, c$, and the distribution function $K : [0, 1] \rightarrow [0, 1]$, a beta distribution with parameters ν and ω . The algorithm will determine the trigger level T for intermediation and the value functions g, h , and v .

I.2 Strategy

We use the following fixed-point algorithm. Start with a value of T , then iterate the following steps:

1. Numerically evaluate g and h (which are independent of the rest of the system, given T).
2. Numerically evaluate v (which depends on T, g and h).
3. Use (33) to obtain a new value of T .
4. Stop if the last iteration is such that the new value of T is close enough to the value of T at the beginning of the loop. Otherwise, return to the first step.

Separate analysis shows that the solution T lies in $1 \leq T \leq 1 + c(r + 2\eta)/q$ which bounds the starting value.

The remaining subsections provide guidelines for the realization of each step. Except for the last subsection, the value of T is fixed.

I.3 A system of equations for g and h

We first discretize the equations for g and h to obtain a linear system of equations of the form

$$Ax = b.$$

The variable $z \in [1, \infty)$ is discretized: we use a grid \mathcal{G} with $n + 1$ points such that $z_i = \delta^i, i \in \{0, \dots, n\}$, where $\delta > 1$ is fixed. Such a grid is finer near 1, where T is more likely to be found. Considering other grids does not affect the equations below.

To each z_i corresponds two rows of the matrix A , which is $(2n + 2) \times (2n + 2)$. The vector $x = [g, h]$ corresponds to the discretized values of the unknown functions g and h . In what follows, $g = (g_0, \dots, g_n)$ and $h = (h_0, \dots, h_n)$ are vectors approximating the functions, and x is the concatenation of these vectors.

For any condition C let 1_C denote the function equal to 1 if C is true and 0 otherwise.

For z and T in \mathcal{G} , we let $\lambda(z, T) = \bar{\lambda}1_{z > T}$. Thus, $\lambda = \bar{\lambda}$ if $z > T$ and 0 otherwise.

I.4 Discretization conventions

For any $0 \leq u < u' \leq 1$, we let $K(u, u')$ denote the probability that the recovery rate is between u and u' , according to the stipulated beta distribution. For each i , let $\lambda_i = \lambda(z_i, T)$

In the computations to follow, we let $z_{-1} = 1$, $z_{n+1} = z_n$, $g_{-1} = g_0$, $g_{n+1} = g_n$, $h_{-1} = h_0$, and $h_{n+1} = h_n$.

I.5 Discretized Equations

The discretized equation for g yields, for $i \in \{0, \dots, n\}$,

$$\begin{aligned}
& g_i[r + 2\eta + \lambda_i(z_i + (1 - q))] + g_{i+1} \frac{\lambda_i z_i (1 + z_i)}{z_{i+1} - z_{i-1}} + g_{i-1} \frac{-\lambda_i z_i (1 + z_i)}{z_{i+1} - z_{i-1}} + h_i(q - 1)\lambda_i \\
& - \eta \sum_{j=0}^i g_j \frac{z_j}{z_i} K\left(\frac{z_{j-1} + z_j}{2z_i}, \frac{z_j + \min\{z_{j+1}, z_i\}}{2z_i}\right) \\
& - \eta \sum_{j=0}^n \frac{h_j}{z_i z_j} K\left(1_{j < n} \left(\frac{1}{2z_i z_j} + \frac{1}{2z_i z_{j+1}}\right), \frac{1}{2z_i z_j} + \frac{1}{2z_i z_{j-1}}\right) \\
& - \eta \sum_{j=i}^n g_j \frac{z_j}{z_i} K\left(1_{j < n} \left(\frac{z_i}{2z_j} + \frac{z_i}{2z_{j+1}}\right), \frac{z_i}{2z_j} + \frac{z_i}{2 \max\{z_{j-1}, z_i\}}\right) = \frac{1}{z_i}. \quad (34)
\end{aligned}$$

The discretized equation for h yields for $i \in \{0, \dots, n\}$

$$\begin{aligned}
& h_i[r + 2\eta + \lambda_i z_i] + h_{i+1} \frac{\lambda_i z_i (1 + z_i)}{z_{i+1} - z_{i-1}} + h_{i-1} \frac{-\lambda_i z_i (1 + z_i)}{z_{i+1} - z_{i-1}} \\
& \quad - \eta \sum_{j=0}^i h_j K \left(\frac{z_{j-1} + z_j}{2z_i}, \frac{z_j + \min\{z_{j+1}, z_i\}}{2z_i} \right) \\
& \quad - \eta \sum_{j=0}^n g_j z_j K \left(1_{j < n} \left(\frac{1}{2z_i z_j} + \frac{1}{2z_i z_{j+1}} \right), \frac{1}{2z_i z_j} + \frac{1}{2z_i z_{j-1}} \right) \\
& \quad - \eta \sum_{j=i}^n h_j K \left(1_{j < n} \left(\frac{z_i}{2z_j} + \frac{z_i}{2z_{j+1}} \right), \frac{z_i}{2z_j} + \frac{z_i}{2 \max\{z_{j-1}, z_i\}} \right) = 1. \quad (35)
\end{aligned}$$

I.6 Linear system

We index from 0 to $2n + 1$ the rows and columns of A as well as the rows of b . Indices from 0 to n correspond to equations or variables related to g , while indices from $n + 1$ to $2n + 1$ correspond to equations or variables related to h . The above discretized equations determine the coefficients of A and b . First, $b_i = 1/z_i$ for $i \leq n$ and $b_i = 1$ for $i > n$, as is clear from the above. We can decompose A into four $(n + 1) \times (n + 1)$ submatrices as

$$A = \begin{bmatrix} B & C \\ D & E \end{bmatrix}.$$

The coefficients of these submatrices are determined by the previous discretized equations. We have

$$B_{ii} = r + 2\eta + \lambda_i(z_i + (1 - q)) - \eta \left[K \left(\frac{z_{i-1} + z_i}{2z_i}, 1 \right) + K \left(\frac{1}{2} + \frac{z_i}{2z_{i+1}}, 1 \right) \right].$$

For $i < n$,

$$B_{i(i+1)} = \frac{\lambda_i z_i (1 + z_i)}{z_{i+1} - z_{i-1}} - \eta \frac{z_{i+1}}{z_i} K \left(1_{i+1 < n} \left(\frac{z_i}{2z_{i+1}} + \frac{z_i}{2z_{i+2}} \right), \frac{z_i}{2z_{i+1}} + \frac{1}{2} \right).$$

For $i > 0$,

$$B_{i(i-1)} = \frac{-\lambda_i z_i (1 + z_i)}{z_{i+1} - z_{i-1}} - \eta \frac{z_{i-1}}{z_i} K \left(\frac{z_{i-2} + z_{i-1}}{2z_i}, \frac{z_{i-1} + z_i}{2z_i} \right).$$

For all i and $j > i + 1$,

$$B_{ij} = -\eta \frac{z_j}{z_i} K \left(1_{j < n} \left(\frac{z_i}{2z_j} + \frac{z_i}{2z_{j+1}} \right), \frac{z_i}{2z_j} + \frac{z_i}{2 \max\{z_{j-1}, z_i\}} \right).$$

For all i and $j < i - 1$,

$$B_{ij} = -\eta \frac{z_j}{z_i} K \left(\frac{z_{j-1} + z_j}{2z_i}, \frac{z_j + \min\{z_{j+1}, z_i\}}{2z_i} \right).$$

The coefficients of the matrices C, D , and E can be obtained similarly.

Once A is computed, we solve the system $A[g; h] = b$. This yields the vector of candidate values for g and h that is needed in the next step of the algorithm.

For $n = 100$, the system can easily be solved by any reasonable computation package, as long as A is invertible. Usual algorithms proceed by factorization of A and direct computation of the solution by pivot methods, which are faster and more robust than inversion of A .

I.7 Computation of v

We discretize the equation for v similarly, using the candidate values of g and h obtained in the previous step. The goal of this subsection is to determine the coefficients of the matrix F and a vector d defining the system $Fv = d$, where $v \in \mathbb{R}_+^{n+1}$ is the discretization vector of the function v , F is a $(n+1) \times (n+1)$ square matrix, and d is an $(n+1)$ -dimensional vector.

The discretized equation for $v = (v_0, \dots, v_n)$ yields for $i \in \{0, \dots, n\}$, keeping the same notational scheme used before and letting $v_{-1} = v_0$ and $v_{n+1} = v_n$,

$$\begin{aligned} v_i[r + 2\eta] + v_{i+1} \frac{\lambda_i z_i (1 + z_i)}{z_{i+1} - z_{i-1}} + v_{i-1} \frac{-\lambda_i z_i (1 + z_i)}{z_{i+1} - z_{i-1}} \\ - \eta \sum_{j=0}^i v_j K \left(\frac{z_{j-1} + z_j}{2z_i}, \frac{z_j + \min\{z_{j+1}, z_i\}}{2z_i} \right) \\ - \eta \sum_{j=0}^n v_j K \left(1_{j < n} \left(\frac{1}{2z_i z_j} + \frac{1}{2z_i z_{j+1}} \right), \frac{1}{2z_i z_j} + \frac{1}{2z_i z_{j-1}} \right) \\ - \eta \sum_{j=i}^n v_j K \left(1_{j < n} \left(\frac{z_i}{2z_j} + \frac{z_i}{2z_{j+1}} \right), \frac{z_i}{2z_j} + \frac{z_i}{2 \max\{z_{j-1}, z_i\}} \right) = \lambda_i [qz_i (h_i - g_i) - c]. \end{aligned} \tag{36}$$

Therefore, the right-hand side of the linear system is $d_i = \lambda_i [qz_i (h_i - g_i) - c]$. The coefficients F are determined as were those of A .

I.8 New Value of T

The last step of the loop of the fixed-point algorithm is the determination of a new candidate trigger level of T . Discretizing (33) yields the condition, for $T = z_t$

$$(1 + z_t)z_t \frac{v_{t+1} - v_{t-1}}{z_{t+1} - z_{t-1}} = qz_t(h_t - g_t) - c.$$

The new candidate value of T is thus the element of the grid \mathcal{G} whose corresponding index t is the closest to satisfying the above equation.

J Diffusion Risk

In this appendix, we allow invested capital to be exposed to diffusive reinvestment risk. Specifically, we suppose that the Lévy process ρ_i driving proportional capital changes in market i is the sum of a Brownian motion ζ_i and an independent compound Poisson process. The value function retains the same degree of homogeneity found in the main text.

With perfect correlation between the Brownian sources of risk in the two markets, ζ_a and ζ_b , the analysis is identical to that shown in the main text.

More generally, suppose that the Brownian motions ζ_a and ζ_b have volatility parameter σ and correlation parameter R . In the remainder of this appendix, we derive the characterizing equations for G and H , then g and h .

To clarify computations with diffusion terms, we temporarily consider investor wealth. Let $\tilde{G}(x, y, \alpha)$ and $\tilde{H}(x, y, \alpha)$ denote the present value of having α units of capital initially in the large and small markets, respectively. Of course, $\tilde{G}(x, y, \alpha) = \alpha G(x, y)$, where $G(x, y) = \tilde{G}(x, y, 1)$. Similarly, $\tilde{H}(x, y, \alpha) = \alpha H(x, y)$, where $H(x, y) = \tilde{H}(x, y, 1)$. We first provide equations for \tilde{G} and \tilde{H} , and then use those to derive equations for G and H .

We assume, to begin, zero recovery. As before, we can take the drift rate μ to be zero

without loss of generality. We have

$$\begin{aligned}
& -r\tilde{G}(x, y, \alpha) + \alpha\pi(x) - \tilde{G}_x(x, y, \alpha)x\Lambda(x, y) + \tilde{G}_y(x, y, \alpha)x\Lambda(x, y) \\
& + (1 - q)\eta(\tilde{H}(x, y, \alpha) - \tilde{G}(x, y, \alpha)) - \eta\tilde{G}(x, y, \alpha) + \eta(\tilde{G}(x, 0, \alpha) - \tilde{G}(x, y, \alpha)) \\
& + \frac{1}{2}\sigma^2[\tilde{G}_{xx}(x, y, \alpha)x^2 + \tilde{G}_{yy}(x, y, \alpha)y^2 + \tilde{G}_{\alpha\alpha}(x, y, \alpha)\alpha^2] \\
& + \sigma^2[xyR\tilde{G}_{xy}(x, y, \alpha) + x\alpha\tilde{G}_{x\alpha}(x, y, \alpha) + y\alpha R\tilde{G}_{y\alpha}(x, y, \alpha)] = 0 \quad (37)
\end{aligned}$$

and

$$\begin{aligned}
& -r\tilde{H}(x, y, \alpha) + \alpha\pi(y) - \tilde{H}_x(x, y, \alpha)x\Lambda(x, y) + \tilde{H}_y(x, y, \alpha)x\Lambda(x, y) \\
& + \eta(\tilde{G}(y, 0, \alpha) - \tilde{H}(x, y, \alpha)) - \tilde{\eta}H(x, y, \alpha) \\
& + \frac{1}{2}\sigma^2[\tilde{H}_{xx}(x, y, \alpha)x^2 + \tilde{H}_{yy}(x, y, \alpha)y^2 + \tilde{H}_{\alpha\alpha}(x, y, \alpha)\alpha^2] \\
& + \sigma^2[xyR\tilde{H}_{xy}(x, y, \alpha) + x\alpha R\tilde{H}_{x\alpha}(x, y, \alpha) + y\alpha\tilde{H}_{y\alpha}(x, y, \alpha)] = 0, \quad (38)
\end{aligned}$$

where we used the fact that, when the investor is in market x , the correlation between x and α is 1, and the correlation between y and α is R . The symmetric correlations apply when the investor is in market y .

Using the fact that $\tilde{G}_\alpha(x, y, 1) = G(x, y)$, $\tilde{G}_{\alpha\alpha}(x, y, 1) = 0$, $\tilde{G}_{x\alpha}(x, y, 1) = G_x(x, y)$, and $\tilde{G}_{y\alpha}(x, y, 1) = G_y(x, y)$, with identical relations between \tilde{H} , H , and their derivatives, we get the following equations for G and H (letting $\alpha = 1$ in the previous equations):

$$\begin{aligned}
& -rG(x, y) + \pi(x) - G_x(x, y)x\Lambda(x, y) + G_y(x, y)x\Lambda(x, y) \\
& + (1 - q)\eta(H(x, y) - G(x, y)) - \eta G(x, y) + \eta(G(x, 0) - G(x, y)) \\
& + \frac{1}{2}\sigma^2[G_{xx}(x, y)x^2 + G_{yy}(x, y)y^2] + \sigma^2[xyRG_{xy}(x, y) + xG_x(x, y) + yRG_y(x, y)] = 0 \quad (39)
\end{aligned}$$

and

$$\begin{aligned}
& -rH(x, y) + \pi(y) - H_x(x, y)x\Lambda(x, y) + H_y(x, y)x\Lambda(x, y) \\
& + \eta(G(y, 0) - H(x, y)) - \eta H(x, y) + \frac{1}{2}\sigma^2[H_{xx}(x, y)x^2 + H_{yy}(x, y)y^2] \\
& + \sigma^2[xyRH_{xy}(x, y) + xRH_x(x, y) + yH_y(x, y)] = 0. \quad (40)
\end{aligned}$$

If π is homogeneous of degree $-\gamma$, then so is F . In this case, letting $f(z) = F(z, 1)$, we have $F_{xx}(x, y) = y^{-\gamma-2} f''\left(\frac{x}{y}\right)$,

$$F_{xy}(x, y) = -(\gamma + 1)y^{-\gamma-2} f'\left(\frac{x}{y}\right) - xy^{-\gamma-3} f''\left(\frac{x}{y}\right),$$

and

$$F_{yy}(x, y) = \gamma(\gamma + 1)y^{-\gamma-2} f\left(\frac{x}{y}\right) + 2(\gamma + 1)xy^{-\gamma-3} f'\left(\frac{x}{y}\right) + x^2y^{-\gamma-4} f''\left(\frac{x}{y}\right).$$

This implies that, at $(x, y) = (z, 1)$,

$$\begin{aligned} & \frac{1}{2}\sigma^2[F_{xx}(x, y)x^2 + F_{yy}(x, y)y^2 + 2zF_{xy}(x, y)] \\ &= \sigma^2\left[\frac{\gamma}{2}(\gamma + 1)f(z) + (\gamma + 1)(1 - R)f'(z) + (1 - R)z^2f''(z)\right]. \end{aligned} \quad (41)$$

With $\gamma = 1$, this reduces at $(x, y) = (z, 1)$, to

$$\begin{aligned} & \frac{1}{2}\sigma^2[F_{xx}(x, y)x^2 + F_{yy}(x, y)y^2 + 2xyF_{xy}(x, y)] \\ &= \sigma^2[f(z) + 2(1 - R)zf'(z) + (1 - \rho)z^2f''(z)]. \end{aligned} \quad (42)$$

References

- Andrade, G., C. Chang, and M. Seasholes, 2005, “Predictable Reversals, Cross-Stock Effects, and the Limits of Arbitrage,” Working Paper, Haas School of Business, University of California.
- Basak, S., and B. Croitoru, 2000, “Equilibrium Mispricing in a Capital Market with Portfolio Constraints,” *Review of Financial Studies*, 13, 715.
- Bhattacharya, S., and K. M. Hagerty, 1987, “Dealerships, Trading Externalities, and General Equilibrium,” in *Contractual Arrangements for Intertemporal Trade*, ed. by E. Prescott, and N. Wallace. Minneapolis: University of Minnesota Press, chap. Minnesota Studies in Macroeconomics Series, vol. 1, pp. 81–104.
- Chen, H., G. Noronha, and V. Singhal, 2004, “The Price Response to S&P 500 Index Additions and Deletions: Evidence of Asymmetry and a New Explanation,” *Journal of Finance*, 59, 1901–1929.
- Coval, J., and E. Stafford, 2007, “Asset Fire Sales (and Purchases) in Equity Markets,” *Journal of Financial Economics*, forthcoming.
- Duffie, D., N. Gârleanu, and L. H. Pedersen, 2005, “Over-the-Counter Markets,” *Econometrica*, 73, 1815–1847.
- Froot, K., and P. O’Connell, 1999, “The Pricing of US Catastrophe Reinsurance,” in *The Financing of Catastrophe Risk*, ed. by K. Froot. University of Chicago Press, pp. 195–232.
- Gehrig, T., 1993, “Intermediation in Search Markets,” *Journal of Economics and Management Strategy*, 2, 97–120.
- Greenwood, R., 2005, “Short and Long Term Demand Curves for Stocks: Theory and Evidence,” *Journal of Financial Economics*, 75, 607–650.
- Gromb, D., and D. Vayanos, 2002, “Equilibrium and Welfare in Markets with Constrained Arbitrageurs,” *Journal of Financial Economics*, 66, 361–407.
- , 2007, “Financially Constrained Arbitrage and the Cross Section of Market Liquidity,” Working Paper, London School of Economics.
- Harris, L., and E. Gurel, 1986, “Price and Volume Effects Associated with Changes in the S&P 500: New Evidence of Price Pressures,” *Journal of Finance*, 41, 815–830.
- He, Z., and A. Krishnamurthy, 2007, “Intermediation, Capital Immobility, and Asset Prices,” Northwestern University Working Paper.

Holthausen, R. W., R. L., and D. Mayers, 1990, “Large-Block Transactions, the Speed of Response and Temporary and Permanent Stock-Price Effects,” *Journal of Financial Economics*, 26, 7195.

Hradsky, G. T., and R. D. Long, 1989, “High-Yield Default Losses and the Return Performance of Bankrupt Debt,” *Financial Analysts Journal*, (July-August 1989), 38–49.

Kaul, A., V. Mehrotra, and R. Morck, 2000, “Demand Curves for Stocks do Slope Down: New Evidence from an Index Weights Adjustment,” *Journal of Finance*, 55, 893–912.

Mitchell, M., L. H. Pedersen, and T. Pulvino, 2007, “Slow Moving Capital,” *American Economic Review, Papers and Proceedings*, forthcoming.

Moresi, S., 1991, “Three Essays in Economic Theory,” Ph.D. Thesis, MIT.

Newman, Y., and M. Rierson, 2003, “Illiquidity Spillovers: Theory and Evidence from European Telecom Bond Issuance,” Working Paper, Graduate School of Business, Stanford University.

Protter, P., 2004, *Stochastic Integration and Differential Equations, Second Edition*. New York: Springer-Verlag.

Rubinstein, A., and A. Wolinsky, 1987, “Middlemen,” *Quarterly Journal of Economics*, 102, 581–594.

Scholes, M., 1972, “The Market for Securities: Substitution versus Price Pressure and the Effects of Information on Share Prices,” *Journal of Business*, 45, 179–211.

Shleifer, A., 1986, “Do Demand Curves for Stocks Slope Down?,” *Journal of Finance*, 41, 579–590.

Sun, Y., 2006, “The Exact Law of Large Numbers via Fubini Extension and Characterization of Insurable Risks,” *Journal of Economic Theory*, 126, 31–69.

Vayanos, D., and T. Wang, 2007, “Search and Endogenous Concentration of Liquidity in Asset Markets,” *Journal of Economic Theory*, forthcoming.

Weill, P.-O., 2002, “The Liquidity Premium in a Dynamic Bargaining Market,” Working Paper, Stanford University.

———, 2007, “Leaning Against the Wind,” *Review of Economic Studies*, forthcoming.

Yavaş, A., 1996, "Search and Trading in Intermediated Markets," *Journal of Economics and Management Strategy*, 5, 195–216.