Capital Mobility and Asset Pricing *

Darrell Duffie
Stanford University

Bruno Strulovici
Northwestern University

September 14, 2009

Abstract

We present a model for the equilibrium movement of capital between asset markets that are distinguished only by the levels of capital invested in each. Investment in that market with the greatest amount of capital earns the lowest risk premium. 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. Those fees also depend on the bargaining power of the investor, in light of potential alternative intermediaries. In equilibrium, the speeds of adjustment of mean returns and of capital between the two markets 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, The London School of Economics, London Business School, Yale, and especially for comments from Bruno Biais, Eddie Dekel, Julien Hugonnier, Glen Weyl, Jean Tirole, Jean-Charles Rochet, Larry Samuelson, and Dimitri Vayanos. We are thankful for the research assistance of Sergey Lobanov.

1 Introduction

We present a model for the equilibrium movement of capital between markets. Equilibrium conditional mean rates of return vary across markets according to 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, intermediaries find 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. An intermediary charges suppliers of capital a fee that is based on their gain from the move, and based on the degree of competition in the market for intermediation.

This paper is motivated by extensive empirical evidence, some of which is reviewed in the last section, 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 conditional mean asset returns, in a way that reflects delays in the adjustments of investors' portfolios. We are particularly interested in how those adjustments are affected by the endogenous behavior of intermediaries.

In our equilibrium model, the greater the relative difference in capital levels across the markets, the more intensive are intermediaries' 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 toward a common level.

An example is the 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."

We are particularly interested in the impact of competition among intermediaries on the equilibrium degree of capital mobility. The impact of competition among intermediaries is through two channels. First, intermediaries do not internalize the entire impact of their search activity on the market because each gets only a fraction of the aggregate intermediation fees. This prompts intermediaries to search more as the number of intermediaries increases. Competition has a second and potentially offsetting effect on capital mobility through the impact of fee bargaining on lower incentives to intermediate. In the simplest setting that we analyze, the second effect dominates: Increasing the number of intermediaries reduces capital mobility.

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 substantial 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 inefficient (relative to a market without intermediation 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. Exceptions include the recent work of Coval and Stafford (2007) and Lou (2009), who note that the conditional mean returns of an equity tend to be lower due to price pressure when mutual funds owning that equity are experiencing liquidation-motivated outflows, and that the conditional mean returns recover as price pressure abates.

A significant body of theory examines the implications of search frictions for asset pricing. For example, differences in search frictions across different asset markets are treated by Weill (2008) and Vayanos and Wang (2007). Duffie, Gârleanu, and Pedersen (2005) study the implications of search frictions in a single asset market with marketmaking. In the context of a single asset market, Duffie, Gârleanu, and Pedersen (2007) and Lagos, Rocheteau, and Weil (2008) model recoveries in mean returns, after a shock to the preferences of investors, corresponding to a gradual re-allocation of the asset to more suitable investors, rather than by cross-market capital dynamics as here. Earlier search-based models of intermediation include Rubinstein and Wolinsky (1987), Bhattacharya and Hagerty (1987), Moresi (1991), Gehrig (1993), and Yavaş (1996).

Related 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 some objectives and model features, the study by Gromb and Vayanos (2007) is closely related to ours. Our respective approaches were developed independently. Common to our models, local hedgers are 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. Our respective approaches, however, are quite different. Our model focuses on capital dynamics and their impact on risk premia.

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) investors, who provide risk-bearing to hedgers in each of two local markets; and (iii) intermediaries (or asset managers) who provide the fee-based service to investors of moving capital from one market to the other. In equilibrium, investors move their capital, subject to intermediation frictions, into that market with the higher premium for the same risk.

We fix a probability space $(\Omega, \mathcal{F}, P)$ 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 (a non-atomic measure space) of local risk-averse agents own short-lived risky assets that they are willing to sell at or above their respective reservation prices. Equivalently, they are willing to buy insurance contracts against the risks to which they are exposed. 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. A continuum of investors that supply capital have access to cross-market trading, subject to intermediation frictions to be described. These suppliers of capital are risk-neutral, offering to bear the risk that hedgers desire to shed in return for any strictly positive risk premium. In an insurance context, one might think of these suppliers of capital as stylized versions of the "Names" that supply risk bearing capacity to the insurance market known as "Lloyd's of London."

The total 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

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

for sale 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 double auctions at which the supply and demand for the asset in market i are matched at each point in time. As the amount of capital available to invest in the asset is increased, the market-clearing reset rate declines. In Appendix A, 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 that 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 promises $1 + d\rho_{it} + \pi(X_{it})\,dt$ 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 in that market until some time $\tau > t$ accumulates to W_τ units of capital, according to the stochastic differential equation $dW_t = W_{t-}\,d\rho_{it}$, and in the meantime generates cash flows at the rate $W_{t-}\,\pi(X_{it})$. (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 illustrative case of an insurance market, we can take ρ_i to be a compound Poisson process 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 instantaneous 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 a loss-size distribution ν with mean -ν, then the mean loss rate is η-ν. In this case, as the amount x of capital gets large, the market clearing payout rate $\pi(x)$ cannot go below $\eta\overline{\nu} + r$, where $r > 0$ is the time preference rate of the investors. Investors optimally supply all of their local capital inelastically, so long as the mean rate of return $\pi(x) - \eta\overline{\nu}$ is strictly larger than their time preference rate r.

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

We assume that $\rho_a = \epsilon_a + \epsilon_c$ and $\rho_b = \epsilon_b + \epsilon_c$, where the market-specific processes ϵ_a and ϵ_b as well as the common component ϵ_c are independent Lévy processes. We assume that ϵ_a and ϵ_b have the same distribution, so that the two markets have identically and symmetrically distributed risks. This symmetry 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.

If there were no capital-market frictions, investors would move capital between the markets so as to obtain the higher reset rate, and in doing so would equate the payout rates $\pi(X_{at})$ and $\pi(X_{bt})$, and thereby equate X_{at} and X_{bt} at all times. Indeed, given the symmetrically distributed returns of the two markets, investors 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.

An investor chooses 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 a particular investor from market a into market b through time t, the investor's 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 have introduced a model for intermediation of capital movements. A investor is infinitely-lived, and has a utility of

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

where $|C|_t$ denotes the total variation of C up to time t. A minor alteration of the model that allows for randomly timed exit and entrance of investors would be equally tractable.[2] For simplicity, we have assumed that transactions costs are paid directly by investors, and not deducted from the capital moved from market to market.

Each investor 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, investors form correct conjectures regarding the dynamics of (X_a, X_b, K).

Intermediaries contact investors in order to profit from fees for moving their capital from one market to another. In equilibrium, at any time, only investors in that market with greater capital agree to have any of their capital moved to the other market. Because a investor 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.[3]

We let $W_{ij}(t)$ denote the level of capital in market i of investor j at time t. Conditional on the intermediation contact intensity process λ, investors are contacted, pairwise independently at the conditional mean rate λ_t. In a manner similar to that of Weill (2007), the law of large numbers allows us to calculate the aggregate rate of movement of capital. Letting $m(\,\cdot\,)$ denote the non-atomic measure over the space of investors, the total rate at which capital is moved from market a to market b is almost surely[4]

$$\begin{aligned} \int \lambda_t 1_{\{X_{at} > X_{bt}\}} W_{aj}(t)\, dm(j) &= \lambda_t 1_{\{X_{at} > X_{bt}\}} \int W_{aj}(t)\, dm(j) \\ &= \lambda_t 1_{\{X_{at} > X_{bt}\}} X_{at}. \end{aligned}$$

Likewise, the rate at which capital moves from market b to market a is $\lambda_t 1_{\{X_{bt} > X_{at}\}} X_{bt}$.

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

[3]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 an intermediary.

[4]One can apply the results of Sun (2006), relying on a particular style of product measure space for states of the world and investors.

Given the 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 one additional unit of capital in market i at time t is

$$\theta_{it}^{\lambda} = E\left(\int_t^{\infty} e^{-r(s-t)} \left[W_s \pi \left(X_{D(s),s}^{\lambda} \right) \, ds - W_{s-} \, K_{s-} \, dN_s \right] \, \middle| \, \mathcal{F}_t \right), \tag{1}$$

where 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 accumulated capital W_s is currently located in market a or b. This 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 $X_{it}^{\lambda} > X_{jt}^{\lambda}$. The marginal value of moving capital is thus

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

At each time t, intermediaries charge investors some fraction $q \in [0, 1]$ of the gain ϕ_t^{λ} from switching each unit of capital. That is, the proportional intermediation fee is $K_t^{\lambda} = q\phi_t^{\lambda}$. One can view q as the bargaining power of an intermediary. We later treat the effect of intermediary competition on q.

To assume that an investor 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 investors is reasonable in some circumstances.

Our model can also be generalized by supposing that each investor has an alternative technology by which opportunities to move capital to the other market arrive at random times, independent across investors, with a constant mean arrival rate. This would cause only minor modifications to the structure and solution of our model. We avoid it for simplicity. Increasing the mean arrival rates of these alternative capital-shifting opportunities reduces the average degree of imbalance of capital and the difference in risk premia between the two markets, and thus reduces the profitability of intermediation.

An intermediary's rate of cost for applying contact intensity λ_t is assumed to be $c\lambda_t$, for some technological cost coefficient $c \geq 0$. For example, doubling the expected rate at which investors are contacted costs twice as much.[5] We restrict λ to be a progressively-measurable process so that, at each time, the intermediary chooses a contact intensity that depends only on information that is currently available. The maximum feasible contact intensity of the market is some constant $\overline{\lambda} > 0$.

3 Equilibrium with Monopolistic Intermediation

We first take the monopolistic case, $n = 1$. In the next section, the solution of the monopolistic case leads immediately to a solution for the oligopolistic case via a simple equivalence result.

3.1 The Monopolist's Problem

A monopolistic intermediary's total rate of fee revenue is $\lambda_t \max(X_{at}, X_{bt}) q \Phi_t$, where $\Phi_t = \phi_t^\lambda$ is the gain from switching capital under policy λ. 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 impose this consistency property as part of the definition of an equilibrium.

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

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

[5] This can be viewed as 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 assume that the parameters are such that this utility is finite, which is the case in the equilibria that we analyze. We restrict attention to intermediation policies that depend on the available information only through 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.

The inability to commit to an intermediation strategy may in principle be overcome by sophisticated punishment threats, as in Ausubel and Deneckere (1989) and Mailath and Samuelson (2006). In such equilibria, if the intermediary deviates, investors would update their beliefs about the intermediary's strategy in a way that harms the intermediary. Such equilibria are based on sophisticated off-equilibrium-path investor beliefs, which are not in the spirit of our assumption that investors are less sophisticated than intermediaries. Another possible justification for our focus on Markov equilibria is the fact that more sophisticated equilibria unravel in finite-horizon models where (possibly state-dependent) stage games have a unique Nash equilibrium.

Given the symmetry of the two markets, it suffices to characterize equilibrium behavior in terms of

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

The payoff processes to investments in the "larger" and "smaller" markets are, respectively,

$$d\rho_t^X = 1_{\{X_{at} > X_{bt}\}}\, d\rho_{at} + 1_{\{X_{at} \le X_{bt}\}}\, d\rho_{bt}$$

$$d\rho_t^Y = 1_{\{X_{at} \le X_{bt}\}}\, d\rho_{at} + 1_{\{X_{at} > X_{bt}\}}\, d\rho_{bt}.$$

From the Lévy property, (ρ^X, ρ^Y) has the same joint distribution as the primitive payoff processes (ρ_a, ρ_b).

Because we restrict attention to an intermediation intensity process λ that depends only on current capital levels, and because of symmetry, we can suppose that $\lambda_t = \Lambda(X_t, Y_t)$ for some measurable policy function $\Lambda : \mathbb{R}_+^2 \to [0, \overline{\lambda}]$ with the property that there is a solution to the associated stochastic differential equation

$$dX_t = -\Lambda(X_t, Y_t) X_t\, dt + X_t\, d\rho_t^X \tag{2}$$

$$dY_t = \Lambda(X_t, Y_t) X_t \, dt + Y_t \, d\rho_t^Y. \tag{3}$$

Letting $\mathcal{L}$ denote the space of intermediation intensity processes of this form, and given an assumed gain-from-switching process Φ, the intermediary solves the problem

$$\sup_{\lambda \in \mathcal{L}} U(x, y, \Phi, \lambda). \tag{4}$$

An equilibrium is a pair (λ, Φ) consisting of an intermediation intensity process λ that attains the supremum (4) given Φ, and such that $\Phi = \phi_t^\lambda$. This definition includes consistency with the optimality for investors to move their capital, in exchange for the marginal fee determined by Φ, when contacted by the intermediary, and includes consistency between the conjectured and actual dynamics for capital movements and search intensity. In the some of the cases that we analyze, we show that there exists a unique equilibrium.

3.2 Homogeneous Case

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 hedgers' dis-utilities 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.

Also without loss of generality, we can take $k = 1$ by re-scaling. That is, the equilibrium behavior for (k, c) is the same as that 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, we can further assume the independence of ρ_a and ρ_b without loss of generality because any common Lévy component would have no effect on the ratio of X to Y. (The sole exception is a 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.)

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. The proportional payoff processes 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 L. 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 ever reach zero at the same time is 0. The partial-recovery case that we later consider 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 investors of the marginal future cash flows per unit of capital held at time t in the over-capitalized and under-capitalized markets, respectively, as defined by (1). Subject to the usual smoothness and integrability conditions, Itô's formula implies that these functions satisfy the coupled equations

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

where subscripts denote partial derivatives. The first of these equations states that the time-preference effect $rG(x,y)$ is equal to the expected rate of gain per unit of capital to an investor currently in the "large market," that with greater capital. This rate of gain first includes dividend payout rate $\pi(x)$. The next two terms capture the rate of change of $G(X_t, Y_t)$ due to intermediated flows of capital out of the large market and into the small market, respectively. The expected rate of gain in value to those in the large market also includes the expected rate of gain $(1-q)\Lambda(x,y)(H(x,y) - G(x,y))$, net of intermediation fees, associated with switching to the higher-premium market. The final two terms reflect the expected rate of impact of loss events. As there is no recovery value to large-market investors of a loss event in the large market, the first of these expected loss rates is $\eta G(x,y)$. A loss event in the small market replaces the value $G(x,y)$ with $G(x,0)$, which explains the final term. The equation for small-market investors is similarly explained.

The marginal gain from switching capital is then

$$\phi_t^{\lambda} = F^{\Lambda}(X_t, Y_t) \equiv H(X_t, Y_t) - G(X_t, Y_t). \tag{5}$$

3.3 The Bellman Equation

Given the marginal value $F^{\Gamma}(x,y)$ for moving capital that is associated with an assumed policy function Γ, the intermediary's value function is

$$V(x,y) = \sup_{\Lambda}\; E\left(\int_0^{\infty} e^{-rt}\Lambda(X_t,Y_t)(X_t q F^{\Gamma}(X_t,Y_t) - c)\,dt\right). \tag{6}$$

We assume that $V(x,y)$ is finite, which is the case in the equilibria that we analyze. The associated Hamilton-Jacobi-Bellman (HJB) equation is

$$0 = \sup_{\ell \in [0,\overline{\lambda}]} \{-rV(x,y) + \mathcal{U}(V,x,y,\ell,\Gamma)\}, \tag{7}$$

where, by Itô's formula,

$$\mathcal{U}(V,x,y,\ell,\Gamma) = -V_x(x,y)\ell x + V_y(x,y)\ell x + \eta[V(y,0) + V(x,0) - 2V(x,y)] + \ell(xqF^{\Gamma}(x,y) - c).$$

Proposition 1 *Given an assumed intermediation policy Γ, suppose that $\hat{V}$ is a bounded differentiable function satisfying the HJB equation (7). Then $\hat{V}$ is the value function V of the optimization problem 6 and any policy $(x,y) \mapsto \Lambda(x,y)$ which, for each (x,y), attains the supremum (7) is an optimal policy given the switching-gain function F^{Γ} associated* *with the policy* Γ. *If, moreover,* $\Lambda = \Gamma$, *then* (λ^*, Φ) *is an equilibrium where, for all* t, $\lambda^*_t = \Lambda(X_t, Y_t)$ *and* $\Phi_t = F^\Lambda(X_t, Y_t)$.

The proof is by a traditional martingale verification argument given in Appendix B. We will show that the assumption that the candidate value function $\hat{V}$ is bounded and differentiable is satisfied by the candidate we calculate in our main parametric example with $\gamma = 1$. Thus, the proposition implies that the HJB equation characterizes optimality in this setting.

The homogeneity of the payout-rate function 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 $f(z) = F^\Lambda(z, 1)$ and $L(z) = \Lambda(z, 1)$. Homogeneity of F^Λ implies that f solves the ordinary differential equation

$$0 = -rf(z) + (1 - z^{-\gamma}) - zL(z)f'(z) + (-\gamma f(z) - zf'(z))L(z)z \tag{8}$$

$$-(1 - q)f(z)L(z) + \eta[g_0(1 - z^{-\gamma}) - 2f(z)].$$

The relevant boundary condition is $f(1) = 0$, corresponding to no gain from switching when the two markets have the same capital levels. Using (8), Appendix C contains a proof of the following result that the switching gain $f(z)$ is strictly positive when capital levels are unequal.

Proposition 2 *Given any intermediation policy* Λ, $f(z)$ *is strictly positive for* $z > 1$. *That is, given* Λ, *investors 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.*

3.4 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 take $a = 0$ and $k = 1$ without loss of generality. In this case, for any admissible policy Λ, the switching gain function F^Λ is homogeneous of degree -1. Thus, taking F^Λ as given, the optimal present value V of intermediation profits is homogeneous of degree 0, that is, $V(x, y) = V(x/y, 1)$ for $y > 0$. In particular, given any reduced switching-gain function f, the policy Λ achieving the supremum of the HJB equation (7) must also be homogeneous of degree 0; that is,

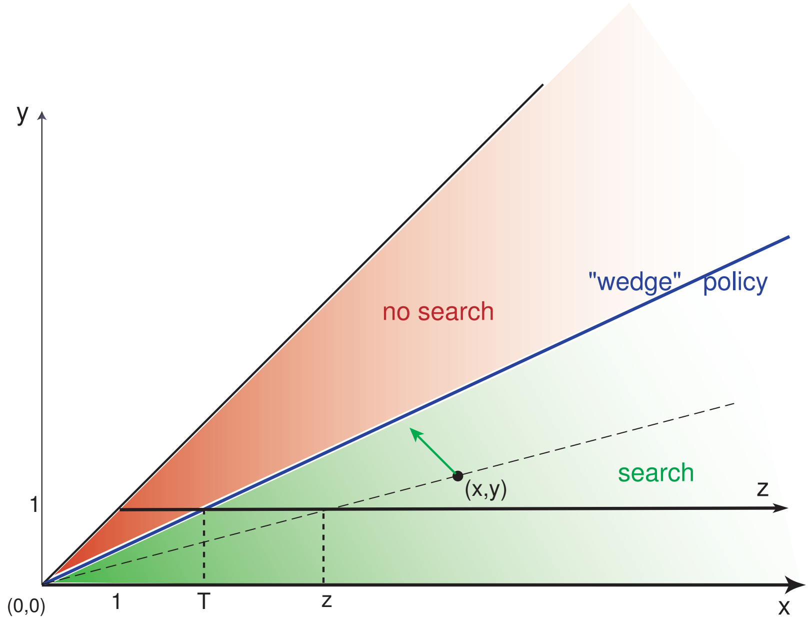

Figure 1: Optimal search policy and solution homogeneity.

$\Lambda(x, y) = L(x/y)$ for some $L(\cdot)$. Because the switching-gain function f depends on the policy function L, we have a fixed-point problem: Find a pair (f, L) such that: (i) given f, the policy L is optimal, and (ii) given L, the marginal gain function f is that determined by L through (8).

In Appendix F (Proposition 9), we show that any equilibrium must be of the "bang-bang" form $\Lambda(x, y) = 0$ for $x < Ty$ and $\Lambda(x, y) = \overline{\lambda}$ for $x \geq Ty$, 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 is intuitive. Because the HJB equation is linear, we anticipate the optimality of switching from minimal to maximal intensity whenever there is sufficient marginal gain from moving capital from one market to the other. This occurs when the levels of capital in the two markets are sufficiently different. Such a trigger policy is illustrated in Figure 1. Our problem is reduced to finding the optimal trigger ratio T, which then completely determines equilibrium behavior.

In order to identify the constant g_0, we use a conservation equation: the sum of the value functions of all investors and of the intermediary must be equal to the present value of all cash dividend payments of the hedgers net of the search costs incurred by the intermediary. After calculations shown in Appendix D, this yields

$$g_0 = \frac{2}{r} - \frac{c\bar{\lambda}}{r}\left(1 - e^{-(2\eta+r)a(T)}\right) - V(1,0), \tag{9}$$

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

The differential equation (8) for f thus reduces to

$$(r + 2\eta + \overline{\lambda}[(1-q) + z])f(z) + \overline{\lambda}(1+z)zf'(z) = (1 + \eta g_0)\left(1 - \frac{1}{z}\right), \quad z \geq T, \tag{10}$$

and

$$(r + 2\eta)f(z) = (1 + \eta g_0)\left(1 - \frac{1}{z}\right), \quad z \in [1, T]. \tag{11}$$

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

$$f(z) = \frac{1 + \eta g_0}{r + 2\eta}\left(1 - \frac{1}{z}\right). \tag{12}$$

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 (10) as

$$(a + z)f(z) + z(1 + z)f'(z) = \left(1 - \frac{1}{z}\right)b, \quad z \geq T, \tag{13}$$

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

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

$$0 = \sup_{\ell \in [0,\overline{\lambda}]} \{-rv(z) - \ell z v'(z) - \ell z^2 v'(z) + 2\eta[v_0 - v(z)] + (qzf(z) - c)\ell\}, \tag{14}$$

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

$$v(z) = v_1, \quad z \in [1, T], \tag{15}$$

where

$$v_1 = \frac{2\eta}{r + 2\eta}v_0 < v_0, \tag{16}$$

and

$$\kappa v(z) + v'(z)z(1 + z) = d + qzf(z), \quad z \geq T, \tag{17}$$

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

Appendix E contains a proof of the following monotonicity and regularity of $v(\,\cdot\,)$. Monotonicity of the value $v(z)$ in the capital heterogeneity measure z is not an obvious result, in particular because the switching gain $f(z)$ is not in general monotonic.

Proposition 3 (Value Function Monotonicity) *For any trigger capital ratio T, the solution v of (14)-(17) is bounded, increasing, and strictly increasing on $[T, \infty)$.*

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

$$T = 1 + \frac{c(r + 2\eta)}{(1 + \eta g_0)q}. \tag{18}$$

A proof of the following result guaranteeing the existence and uniqueness of a trigger strategy is found in Appendices G (existence) and H (uniqueness).

Proposition 4 (Existence and Uniqueness) *There exists a unique trigger capitalization ratio T satisfying* (9), (10), (11), *and* (18).

This analysis leads to the following characterization of equilibrium, 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 $\overline{\lambda}$ on contact intensity.

Proposition 5 *Suppose that the payout-rate function π is of the form $\pi(x) = a + k/x$. Then there exists a unique equilibrium. The equilibrium intermediation process is inactive ($\lambda_t = 0$) whenever the ratio of capital levels in the two markets is between $1/T$ and T, for some capital-ratio trigger T, and is otherwise at full capacity ($\lambda_t = \overline{\lambda}$). The capital ratio trigger T is given by (18), where the constant g_0 is given by (9). If there is no intermediation cost ($c = 0$), then the intermediary always works at full capacity (that is, $T = 1$).*

Relation (18) also provides an upper bound on the equilibrium capital-ratio trigger level:

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

This bound is useful for computing numerical solutions to the optimization problem. An algorithm for computing the constant g_0, and thus T, is given in Appendix I.

3.5 Partial Recovery

We now allow the fraction W recovered after a loss to be randomly distributed on $(0, 1)$. This will be the basis for our numerical illustration of the model.

Subject to the usual smoothness and integrability conditions, Itô's formula and the definition (1) of the value of a unit of capital held in market i imply that the value $G(x,y)$ of a unit of capital in the over-capitalized market satisfies:

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

A similar equation for $H(x,y)$ is found in Appendix J. Unlike the zero-recovery case, these equations cannot be reduced to a single equation for the gain F from switching.

Using the homogeneity of π as before, one can solve the HJB equation for the intermediary's value $V(x,y)$ in the form $v(x/y) = V(x,y)$, as a function of the capital ratio $z = x/y$, in the form

$$\begin{aligned} 0 &= \sup_{\ell \in [0,\bar{\lambda}]} \left\{ -rv(z) - \ell z(1+z)v'(z) + \eta \left(E\left[v\left(\frac{\max(zW,1)}{\min(zW,1)} \right) \right] - v(z) \right) \right. \\ &\quad \left. + \eta \left(E\left[v\left(\frac{z}{W} \right) \right] - v(z) \right) + \ell\,(qzf(z) - c) \right\}. \end{aligned}$$

In this setting, the intermediary's value function cannot be computed by solving a differential equation because $v'(z)$ depends on $v(z')$ for all other z'. We have the same issue to overcome in order to solve for $G(x,y)$ and $H(x,y)$. Exploiting the linear structure of the problem, however, Appendix J provides a numerical algorithm for solving the corresponding integro-differential equations. The associated smooth-fit condition is

$$qTf(T) - c = T(1+T)v'(T). \tag{19}$$

3.6 Numerical Illustration

We provide an illustrative example of equilibrium for the case of partial recovery. We take the parameters $r = 0.04$, $\eta = 1.5$, $c = 0.04$, $\overline{\lambda} = 0.1$, $q = 1/30$. We assume beta-distributed recovery (one minus proportion lost) 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

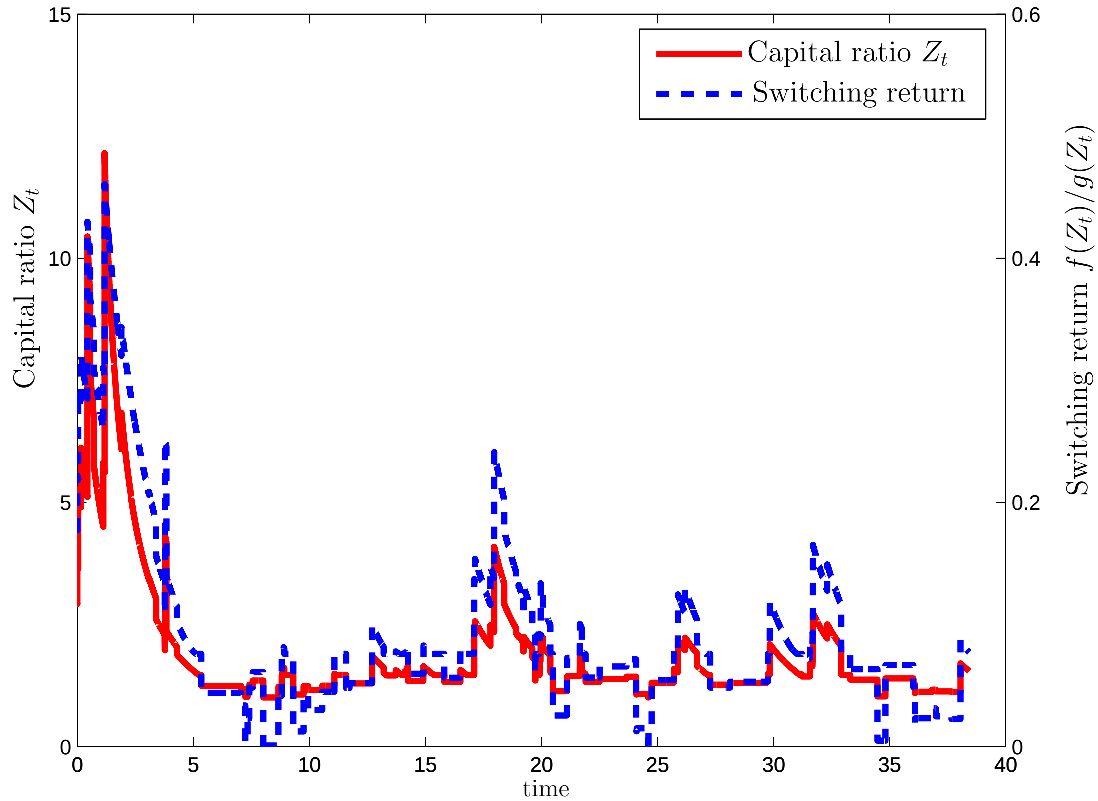

Figure 2: Simulated sample paths of the capitalization ratio, $Z_t = X_t/Y_t$, and the return from switching, $f(Z_t)/g(Z_t)$.

with switching capital into the under-capitalized market. Figure 3 shows the present values, 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 $z = x/y$ of the level of capital x in the over-capitalized market to the level y of capital in the under-capitalized market.

4 Intermediary Competition

We now provide solutions for equilibria with oligopolistic or perfectly competitive markets for intermediation.

There are two channels through which the intermediary competition might affect the

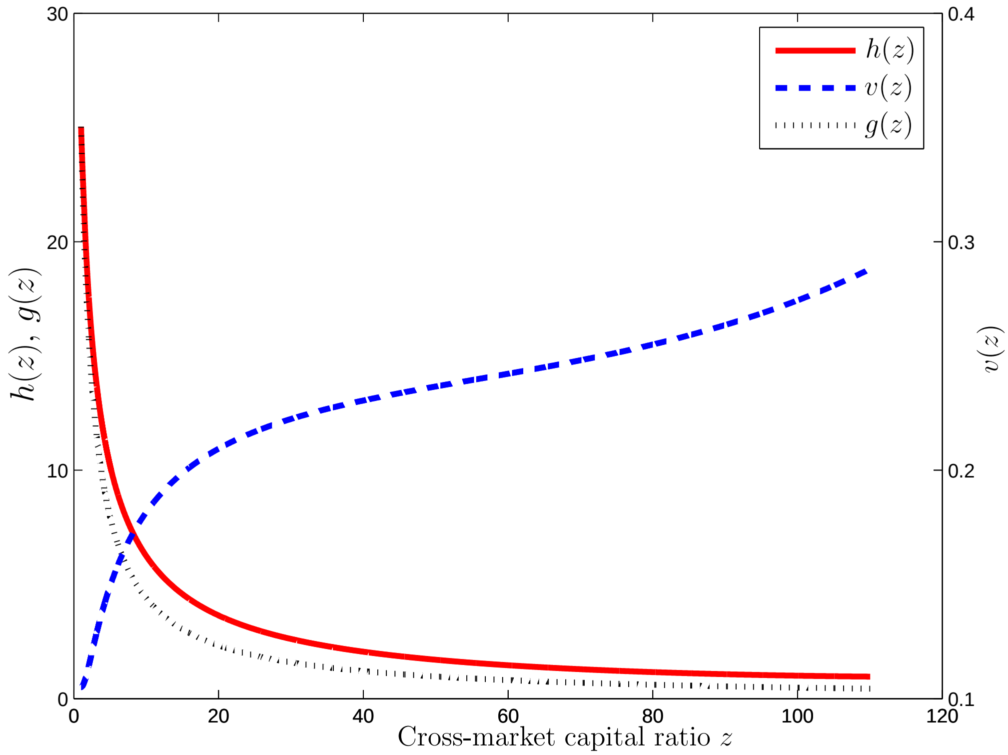

Figure 3: Value function $v(z)$ of the intermediary and the marginal values $g(z)$ and $h(z)$ of capital held in the over-capitalized and undercapitalized markets, respectively.

equilibrium level of intermediation offered by the market. First, a large intermediary internalizes the impact of intermediation intensity on the heterogeneity of capital levels across the two markets, and thus the degree to which there are gains from trade to outside investors. The more intensive is the intermediation policy, the lower are the potential future gains from trade to be split with a investor moving capital. Second, when in contact with a investor, an intermediary considers the ability of the investor to compare the intermediation fee offered with the fees offered by other intermediaries. This plays a role in determining the effective bargaining power of the intermediary, and through that channel, the impact on the profitability of intermediation. We will examine the effects of both channels, and start by taking bargaining power as fixed.

4.1 Intermediary Competition At Fixed Bargaining Power

For a given bargaining power q, equilibrium trigger policies for the oligopolistic case can be translated directly from the case of monopolistic intermediation, by a change of variables.

For the oligopolistic case, we take n identical intermediaries, each with an upper bound $\overline{\lambda}/n$ 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 treat "$n = \infty$" by considering a non-atomic measure space of intermediaries of total mass 1. Each intermediary in this continuum has maximal intermediation intensity $\overline{\lambda}$, again providing a market-wide total intermediation capacity of $\overline{\lambda}$. Thus, all cases have the same feasible market dynamics and costs.

We again consider only Markov equilibria. Equilibrium incorporates the degree to which intermediaries internalize the impact of their intermediation intensity on the heterogeneity of capital levels across markets. We first analyze the case of zero recovery, then briefly comment on the case of partial recovery.

For an oligopolistic 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

$$\begin{aligned} 0 \quad = \quad & \sup_{\ell \in [0, \overline{\lambda}/n]} \left\{ - rv(z) + \left(-\frac{(n-1)}{n}\overline{\lambda} 1_{\{z \geq T\}} - \ell \right) zv'(z) \right. \\ & \left. - \left(\frac{(n-1)}{n}\overline{\lambda} z 1_{\{z \geq T\}} + \ell z \right) zv'(z) + \, 2\eta[v_0 - v(z)] + (qzf(z) - c)\ell \right\}, \end{aligned} \tag{20}$$

reflecting the presumption by the given intermediary that the $n - 1$ other intermediaries have adopted a specific trigger capital ratio T. The equilibrium condition is that the same trigger policy is optimal for the given intermediary. Verification of the HJB solution as the value function is as for the monopolistic case.

Thus, an 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 $\overline{\lambda}/n$ whenever $Z_t \geq T$, for a trigger capital ratio T. We will show that optimality implies that there is no intermediation at or below the capital ratio T satisfying the smooth pasting condition $v'(T) = 0$. This, along with (20), implies that

$$qTf(T) - c = 0. \tag{21}$$

From (21), we see that an intermediary's optimization problem in a setting with n intermediaries is equivalent to that of a monopolistic intermediary with the same maximum intermediation intensity $\overline{\lambda}/n$. Indeed, for a given threshold T, the monopolistic and oligopolistic cases yield the same function f determining proportional intermediation fees, and hence the same smooth-pasting condition (21). In fact, this is actually the unique equilibrium, even allowing for the possibility of non-trigger strategies! To see this, consider any Markov equilibrium, not necessarily of the trigger-ratio form, and let f denote the function determining the associated gain from switching. An intermediary's HJB equation is of the form (20), except that *(i)* the aggregate of other intermediaries' contact intensities may be almost arbitrary, and *(ii)* the value functions may vary across intermediaries. Owing, however, to the form of the HJB equation, the indifference condition is nevertheless given by (21), and thus is the same for all intermediaries. This shows that any Markov equilibrium must be symmetric and of the trigger form.[6] In fact, repeating arguments from the monopolistic case leads to the following proposition.

PROPOSITION 6 *Consider the case of zero recovery and n intermediaries. There exists a unique Markov equilibrium. This equilibrium is symmetric and determined by a trigger capital ratio equal to that of a monopolistic intermediary with the oligopolistic maximal contact intensity $\overline{\lambda}/n$.*

We now informally discuss the case of partial recovery. Recall from (19) the smooth-pasting condition for the monopolistic case:

$$qTf(T) - c = T(1+T)v'(T). \tag{22}$$

One can see that the trigger capital ratio T is determined not only by the function f determining the marginal gain from moving capital, but also by the derivative $v'(T)$ of the intermediary's value function. In order to understand the impact of oligopolistic intermediation, suppose that intermediaries were to use, instead of the optimal trigger ratio T, the equilibrium trigger ratio of a monopolist with the same aggregate capacity for intermediation. In that case, f would be unchanged. Each intermediary, however, would receive only a fraction $1/n$ of the total intermediation fees. The righthand side of (22) is thus lowered, implying that intermediaries prefers to continue intermediating after the capital ratio exceeds the monopolistic trigger. This is the first channel through which oligopolistic competition matters: Because an oligopolistic intermediary does not

[6] The trigger form comes from showing, as in the monopoly case (Lemma 2), that the function $z \mapsto zf(z)$ is increasing.

internalize the full impact of his search on intermediation fees, he has a greater incentive to intermediate. More precisely, an intermediary does not work for opportunities to move capital when the immediate net marginal benefit of doing so, $qzf(z) - c$, is below the marginal value $z(1+z)v'(z)$ associated with future capital heterogeneity. For a given trigger ratio T, an intermediary's value function v declines in direct proportion to the number n of intermediaries, and, hence, so does the derivative v'. This implies that the term $z(1+z)v'(z)$ diminishes with n, while the immediate marginal benefit $qzf(z) - c$ is unchanged, keeping T constant. Thus, as n increases, the incentive to intermediate at the given trigger ratio T becomes strictly positive, prompting intermediaries to search more.[7]

As n goes to infinity, an intermediary's value function goes to zero (because the size of the pie to be shared among intermediaries is uniformly bounded above by $2/r$), and the derivative $v'(T)$ also goes to 0. The limit as n diverges is the competitive equilibrium, in which the trigger capital ratio T is determined by

$$qTf(T) - c = 0.$$

With perfect competition, an intermediary has no impact on aggregate search activity, and thus cares only about the immediate net benefit from switching.

Competition for intermediation, does, however, play a role through the sharing of gains from trade when in contact with an investor. So far, we have taken the fraction q of gains that are allocated to intermediaries to be fixed. We next consider the implications of market structure for the determination of q.

4.2 Endogenous Bargaining Power

With $n > 1$ intermediaries, we suppose that some fraction ψ_n of investors are "well connected," meaning that as they prepare to switch capital from one market to another, they are in simultaneous contact with more than one intermediary. The number of intermediaries with whom a given investor is in contact could also be random, exploiting the law of large numbers, in which case ψ_n can be taken to be the probability that a investor, when contacted, is in contact with more than one intermediary. Intuitively, a well-connected

[7] When there is zero recovery from a loss event, the after-event heterogeneity (which is infinite) does not depend of the pre-event heterogeneity. In that case, intermediaries already ignore the impact of their search activity on heterogeneity and the monopolistic solution coincides with the competitive one.

investor has more bargaining power than a "captive" investor, one who is in contact with only one intermediary.

When modeling this intuition with a bargaining game, an issue is whether the contacted intermediary is assumed to know whether the investor is in contact with other intermediaries. We will take this case.[8] Another modeling approach is a multilateral bargaining game with complete information, as in Stole and Zwiebel (1996). The Shapley value from such a bargaining game is identical to that of the solution below.[9] We consider a bargaining procedure à la Rubinstein (1982), in which the investor and a particular intermediary alternate offers. In our continuous-time setting, the times between offer rounds can be treated as arbitrarily small, so the inter-round discount factor can be taken to be 1. In that case, the investor and intermediary agree immediately to split the surplus according to the Nash bargaining solution. The investor's share depends on his outside option. If the investor is captive, his outside option is simply $g(z)$, the value of remaining in the over-capitalized market. Thus, the Nash product associated with a proportional fee of s is

$$[v(z) + s - v(z)][h(z) - s - g(z)],$$

which is maximal at $s = f(z)/2$, corresponding to $q = 1/2$, meaning an equal splitting of the gains with the intermediary.

For a well-connected investor, the Nash product is

$$[v(z) + s - v(z)][h(z) - s - g(z) - (1 - q_0)f(z)],$$

where q_0 is the conjectured proportion of the gain from trade that the investor would pay to another intermediary if this first round of bargaining were to break down. The Nash product is maximized at $s = 0$, for a proportional intermediary share of $q = 0$, corresponding to the extraction of all surplus by the well-connected investor.

If the number of intermediaries in contact with the investor is known only by the

[8]It would be possible to allow for one-sided information. The fees derived could be obtained as equilibrium outcomes of a bargaining process, although there may be additional equilibria. See, for example, Sutton (1986). For an alternative approach to treating uncertainty about the degree to which an intermediary's customer is in contact with other intermediaries, see Green (2007).

[9]In that case, the payoff of an intermediary is zero whenever at least two intermediaries take part in the bargaining, since the surplus that can be achieved from any coalition is independent of the number of intermediaries, provided that number is nonzero.

investor, then q is similarly obtained, and depends on the probability that the investor is captive.

The average of an intermediary's share of gains across the population of investors is

$$q(n) = 0 \times \psi_n + \frac{1}{2}(1 - \psi_n) = \frac{1 - \psi_n}{2}.$$

In particular, $q(n)$ is decreasing in n if ψ_n is increasing in n. Obviously, $\psi_2 \geq \psi_1$. Going beyond the case of $n = 2$, it is somewhat intuitive that an investor is more likely to be well connected as the number of intermediaries increases. Appendix M briefly outlines a model with this natural feature.

Lowering q reduces an intermediary's incentive to search, all else equal, because, for given capital dynamics, lowering q reduces intermediation profits, and therefore the marginal benefit of raising intermediation intensity. We will next illustrate the second channel through which oligopolistic intermediation affects capital mobility: By reducing each intermediary's bargaining power, the incentive to intermediate is lowered.

Endogenous bargaining leads to complex dynamics, in which the number of intermediaries *actively* searching for capital varies over time. In order to see this, consider a candidate equilibrium in which n intermediaries search at full capacity whenever $z > T$, and no intermediary searches when $z \leq T$. If a single intermediary deviates by searching for capital when z is in a left neighborhood of T, then his fee per unit of capital switched is that of a monopolist, not that of the n-intermediary case. This increases the value of this deviation. Despite this added complexity, we now show that oligopolistic intermediation may reduce capital mobility. We focus on the case of zero recovery.

4.3 Reduced Capital Mobility With More Intermediaries

A Markov strategy profile for n intermediaries consists of functions $L_1, L_2, \ldots, L_n$ on $[1, \infty)$ into $[0, \bar{\lambda}/n]$. Here, $L_i(z)$ denotes the search intensity of intermediary i when the heterogeneity of capital across the two markets is $z = x/y$. The aggregate capital mobility is

$$L(z) = \sum_{i=1}^{n} L_i(z).$$

Let f^L denote the marginal gain to a investor from switching to the market with less capital, given aggregate intensity policy L. In order to exploit the fee share $q(n)$ derived above, we focus on simple strategies, for which $L_i(z)$ is either 0 or $\bar{\lambda}/n$. With this restriction,[10] we can associate with any strategy profile an increasing sequence $T_0, T_1, \ldots, T_K$ of capital-ratio thresholds with the property that, whenever the capital ratio Z_t is in $[T_k, T_{k+1})$, a particular set N_k of intermediaries is active. We let $n_k = |N_k|$ denote the number of intermediaries in N_k.

Using our previous analysis of the oligopolistic case with fixed bargaining power, we say that a profile of simple strategies is a *Markov equilibrium* if, for all k and $z \in [T_k, T_{k+1})$,

$$q(n_k) z f^L(z) - c \geq 0, \quad i \in N_k, \tag{23}$$

and

$$q(n_k + 1) z f^L(z) - c \leq 0, \quad i \notin N_k. \tag{24}$$

The first inequality means that any intermediary searching at level z does so optimally, given equilibrium fee share $q(n_k)$. The second equation states that any intermediary not searching at level z does so optimally, given the equilibrium fee share $q(n_k + 1)$ that he would get if he searched. We let $T_n = \inf\{\bar{z} : L(z) = \bar{\lambda},\ z \geq \bar{z}\}$, the smallest level of capital heterogeneity above which intermediaries search at full capacity. Thus, T_1 denotes the monopolistic threshold. For the result to follow, recall that η is the mean arrival rate of loss events and that T_n depends, through L, on the particular Markov equilibrium under consideration. The following result applies for all equilibria.

PROPOSITION 7 *For any $n \geq 2$, there exists some $\bar{\eta} > 0$ such that for any $\eta \in (0, \bar{\eta})$ and any Markov equilibrium with n players associated with mean shock intensity η, we have $T_1 < T_n$.*

In words, the reduced bargaining power caused by oligopolistic competition reduces the domain of maximal capital mobility relative to that of the monopolistic case. A proof of this proposition may be found in Appendix N. Proposition 7 shows that oligopolistic competition results in less intermediation than achieved by a monopolist, for some range of market heterogeneity. This does not, however, rule out intermediation by oligopolists at capital ratios below the monopolistic trigger level. The next result shows that, provided that loss events are sufficiently rare, oligopolistic and monopolistic settings lead to a cessation of intermediation at approximately the same levels of market heterogeneity.

[10] Extending the analysis to general Markov strategies would be possible if one computes, for any possible strategy, the expected fee for each intermediary as a function of his search intensity and of the aggregate search intensity.

For any n-intermediary Markov equilibrium with aggregate strategy L, let

$$S_n = \inf\{z : L(z) > 0\},$$

the smallest heterogeneity level above which capital mobility is nonzero. A proof of the next proposition may be found in Appendix N.

PROPOSITION 8 *For any $\varepsilon > 0$, there exists a strictly positive $\bar{\eta}$ such that for any $\eta \in (0, \bar{\eta})$ and any Markov equilibrium with n players and mean loss-event arrival rate η, we have $S_n \geq T_1 - \varepsilon$.*

Propositions 7 and 8 together show that capital mobility is lower, at any levels of capital, with oligopolistic intermediation than with monopolistic intermediation, provided that loss events are sufficiently rare.

5 Concluding Remarks

We have examined a simple setting in which, absent trading frictions, investors would adjust their portfolios so as to achieve the highest possible mean return for a given risk, thereby equating mean returns across assets. Because of trading frictions, however, investors cannot instantaneously adjust their portfolios. Over time, investors make portfolio adjustments that cause mean returns across markets to revert toward each other. In our analysis, capital is mobilized through optimal intermediation. Although other market microstructures may lead to similar patterns of adjustment of capital and mean returns, we are particularly focused on the endogenous role of intermediaries.

For example, in corporate bond markets, which are not traded on a central exchange, one observes large price drops and delayed price recovery in connection with major downgrades or defaults, as described by Hradsky and Long (1989) and Chen, Lookman, and Schürhoff (2008), when certain classes of investors have an incentive or a contractual requirement to sell their holdings. 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 was the LTCM collapse in 1998. Newman and Rierson (2003) show that large issuances 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.

Examples of slow price adjustments to supply shocks in equity markets include those of Holthausen and Mayers (1990), Scholes (1972), Coval and Stafford (2007), Andrade, Chang, and Seasholes (2008), Kulak (2008), 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).

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). Periods of high re-insurance rates 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 and capitalize a viable new provider of catastrophe risk insurance.

In these examples, the time pattern of returns or prices after a supply or demand shock reveals that the friction at work is not merely a transaction cost for trade. If that were the nature of the friction, then all investors would immediately adjust their portfolios, or not, optimally. The new market price and expected return would be immediately established, and remain constant until the next change 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 reverts in mean 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 quickly 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 investors. This is only an abstraction, which can also proxy for other forms of delay, including time to educate investors about assets with which they have limited familiarity (awareness), time for contracting, and time for investors to dispose of their current positions, which could involve similar delays and price shocks, as suggested by Chaiserote (2008). 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.

We have assumed that the markets segmented by intermediation frictions are symmetric in all respects other than the level of capital in each. Thus, differences in mean returns, and the value of moving capital from one market to another, are entirely due to the nature of intermediation and differences in capital levels. We could, however, extend the model so as to treat asymmetric markets. Provided that the dividend functions satisfy similar homogeneity assumptions, intermediation would be characterized by two distinct thresholds of capital ratios, one for movement of capital from market a to market b, and another for the reverse movement. For example, if returns in market a are riskier than those in market b, then, all else equal, capital will be less mobile toward market a than toward market b. Asymmetry, for example, would allow a consideration of capital mobility from a low-risk "money market" into a high-risk market such as that for catastrophe risk or private equity. Many of the qualitative features of our symmetric model, such as the dynamics of capital mobility and the impact of intermediation competition, are anticipated to carry over to asymmetric settings, at least under regularity conditions.

Another natural extension concerns the case of three or more markets. Consider, for example, three symmetric markets differing only in their capital levels, and satisfying our homogeneity conditions. We conjecture that capital will flow exclusively to the highest premium market, with more mobility from the the lowest-premium market than from the mid-premium market one. Moreover, when these two flows are strictly positive, the gains from switching will be equalized across the two flows, with the lower mobility from the mid-premium market just offsetting its lower immediate premium differential, illustrating the idea that future capital mobility affects today's gain from switching.

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 $\zeta_1, \zeta_2, \ldots$ 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 precisely that time.[11] 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

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

[11]The appropriate measurability restriction is "predictability."

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 $\overline{p}$, the total demand $\chi(\overline{p}) = M\left(\{\alpha : p_\alpha \leq \overline{p}\}\right)$ for insurance, where $M(\,\cdot\,)$ is the measure on the space A of consumers in the market.[12] Associated with the strictly decreasing demand function χ, assuming continuity, is a strictly decreasing and continuous inverse demand function $\pi(\,\cdot\,)$. The expected loss rate is $\eta E(\zeta_i)$, so the risk premium is $\pi(x) - \eta E(\zeta_i)$. 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 proportional 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 allowed merely for generality.

The total cumulative proportional accumulation process for capital in market i, before

[12] In order for the premium rate $\pi(x)$ to be strictly decreasing in the capital level x, for simplicity we can take the total measure $M(A)$ of buyers of protection to be infinite.

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 $(q, c, \overline{\lambda})$ of the intermediation of capital between the two markets, the primitives $(\pi, \rho_a, \rho_b, r, q, c, \overline{\lambda})$ of our basic model are fixed.

B Verification of Optimality of HJB Solution

This appendix provides a proof that the HJB equation (14) characterizes optimality. For this, given an arbitrary intensity process λ, let

$$S_t = e^{-rt}\hat{V}(X_t^\lambda, Y_t^\lambda) + \int_0^t e^{-rs}\lambda_s[X_s^\lambda q F^\Gamma(X_s^\lambda, Y_s^\lambda) - c]\, ds.$$

By Itô's Formula, a local martingale is defined by

$$\hat{V}(X_t^\lambda, Y_t^\lambda) - \int_0^t \left(-\hat{V}_x(X_s^\lambda, Y_s^\lambda)\lambda_s X_s^\lambda + \hat{V}_y(X_s^\lambda, Y_s^\lambda)\lambda_s X_s^\lambda + \eta[\hat{V}(X_s^\lambda, 0) + \hat{V}(X_s^\lambda, 0) - 2\hat{V}(X_s^\lambda, Y_s^\lambda)]\right) ds.$$

Because λ and $\hat{V}$ are bounded, this local martingale is in fact a martingale. From this and the implication of the HJB equation that

$$-r\hat{V}(X_t^\lambda, Y_t^\lambda) - \mathcal{U}(\hat{V}, X_t^\lambda, Y_t^\lambda, \lambda_t, \Gamma) \leq 0,$$

another application of Itô's formula implies that S is the sum of a decreasing process and a martingale. Thus, S is a supermartingale. Because $\hat{V}$ is bounded, we have the "transversality" condition that for any intermediation intensity process λ,

$$\lim_{t\to\infty} E[e^{-rt}\hat{V}(X_t^\lambda, Y_t^\lambda)] = 0. \tag{25}$$

Thus, for any intermediation intensity process λ,

$$\hat{V}(x, y) \geq \mathcal{V}(x, y, \lambda, \Gamma) \equiv E\left(\int_0^\infty e^{-rt}\lambda_t[X_t^\lambda q F^\Gamma(X_t^\lambda, Y_t^\lambda) - c]\, dt\right). \tag{26}$$

Let Λ be a policy such that, for each (x, y), $\Lambda(x, y)$ attains the supremum (7). For each t, let $\lambda_t^* = \Lambda(X_t, Y_t)$. Then, the fact that

$$-r\hat{V}(X_t, Y_t) - \mathcal{U}(\hat{V}, X_t, Y_t, \lambda_t^*, \Gamma) = 0$$

implies that S is a martingale. Thus

$$\hat{V}(x, y) = \mathcal{V}(x, y, \lambda^*, \Gamma). \tag{27}$$

Thus, for any intermediation intensity process λ,

$$\mathcal{V}(x, y, \lambda^*, \Gamma) \geq \mathcal{V}(x, y, \lambda, \Gamma),$$

proving the result.

C Nonnegativity of the Gain From Switching f

In order to prove Proposition 2, we rewrite (8) as

$$(r + 2\eta + L(z)(\gamma z + (1 - q)))\, f(z) + z(1 + z)L(z)f'(z) = (1 + \eta g_0)(1 - z^{-\gamma}). \tag{28}$$

Because the righthand side is strictly positive, f or f' must be strictly positive. This implies that f cannot cross 0 from above. Hence, f must be strictly positive on some interval of the form (z, ∞), and is non-positive on $[1, z]$ for some level z. It remains to show that $z = 1$. Because $f(1) = 0$, the intermediary does not search when the markets have equal levels of capital, given that $c > 0$. That is, $L(z)$ vanishes on a neighborhood of 1. From (28), this implies that f is positive on that neighborhood, which concludes the proof.

D Valuation of Search Costs

The conservation equation is

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

where $R(x, y)$ is the present value 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 investors, and $P_T(x, y)$ is the the intermediary's expected discounted search costs over the infinite horizon, given a trigger T.

Because π is homogeneous of degree -1, we have $R(x, y) = 2/r$. The search-cost present value $P_T(1, 0)$ solves

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

where p is present value of search costs from time zero to the exponentially distributed time τ of the next loss event. We now show that, for the case of no recovery at loss event,

$$P_T(1,0) = \frac{c\bar{\lambda}}{r}\left(1 - e^{-(2\eta+r)a(T)}\right), \tag{29}$$

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

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

$$P_T(1,0) = p + E[e^{-r\tau}P_T(1,0)], \tag{30}$$

where 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 (30) yields the result (29).

E Proof of Proposition 3

That v is bounded follows from the fact that it is dominated by $2/r$. The monotonicity result is based on two intermediate lemmas.

First, given the function f determining intermediation fees, let

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

The first term of $\kappa(z)$ is the present value of switching capital to the under-capitalized market if the intermediary arrests intermediation efforts from the point at which the capital ratio Z_t is at z until the next loss event occurs, given g_0. Suppose in particular, a given reduced policy $L(z) = \Lambda(z, 1)$, and a particular z at which $L(z) = 0$. Then $\kappa(z) = 0$. As a special case, $\kappa(1) = 0$ (which can also be checked directly from the definition of κ and the fact that $f(1) = 0$). We note that, since the first term defining κ is strictly increasing in z, $\kappa'(z)$ must be positive whenever $f'(z)$ is negative. Given a policy L, we will show that κ is nonnegative. In order to see this, we observe that for $z \geq 1$, (28) can be re-written as

$$L(z)\left[((1 - q) + z\gamma)f + z(1 + z)f'\right] = (r + 2\eta)\kappa(z). \tag{31}$$

We already know that $\kappa(1) = 0$. Since f is positive from Proposition 2, this implies that $f'(z)$ is negative whenever $\kappa(z) \leq 0$, and hence that $\kappa' > 0$ whenever $\kappa \leq 0$. Therefore, κ cannot cross 0 from above, which proves our first lemma.

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

This result is intuitive: other things equal, the expected gain from moving one's capital is larger if the intermediary immediately stops switching capital after that last movement, since the difference between capital levels, and hence between premia, is larger in that case. Lemma 1 has a crucial consequence for the case $\gamma = 1$: the rate at which fees are paid to the intermediary when he searches is strictly increasing in z. 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 strictly increasing in z. We can re-write (31) when $\gamma = 1$ as

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

Since f is positive and κ is nonnegative, this implies that $f(z)+zf'(z)$ is positive whenever $L(z) > 0$, hence that $zf(z)$ is strictly 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 strictly increasing, and, a fortiori, so is $zf(z)$.

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

We can now show monotonicity of v for any trigger policy. From (16), 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. The function $z \mapsto [qf(z)z - c]1_{z>S}$ is nondecreasing in z from Lemma 2, and strictly 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 depend on z or z'). This proves Proposition 3.

F Optimality of a Trigger Policy

This appendix shows that for any equilibrium pair (f, L), the reduced policy function L must be a trigger policy. In fact, we will show that for any switching-gain function f that can arise as the result of an admissible intermediation policy, equilibrium or otherwise, the optimal policy must be of the trigger form.

From Appendix B, we know that, for a given f, any bounded solution of the HJB equation yields an optimal policy. We also know that f is continuous (and, in fact, differentiable) from (13). From Lemma 2, we also know that for any admissible policy, $zf(z)$ must be increasing. Finally f must be such that the value function v is bounded by $2/r$. These conditions define what we call "admissibility" of f. In particular, these conditions must be satisfied in any equilibrium.

We first show that there exists a solution to the HJB equation that is achieved by a trigger policy. Then we verify that any policy that achieves the value function that solves the HJB equation must be of the trigger form.

For any equilibrium, the function f is bounded, because

$$f(z) = |h(z) - g(z)| \leq h(z) + g(z) \leq h(z) + zg(z) \leq \frac{2}{r}.$$

Therefore, given any candidates for the capital trigger ratio T and the constant v_1, one can integrate the HJB equation (17) on $[T, \infty)$. The smooth-pasting condition is satisfied if $v'(T) = 0$, and this is equivalent to the condition that $qTf(T) = c$. (For this, see (14).) Given f, this uniquely determines T, because $Tf(T)$ is strictly increasing in T by Lemma 2. The only difficulty is to show the consistency condition $v_1 = (2\eta/2\eta + r)v_0$ (see (16)), where $v_0 = \lim_{z\to\infty} v(z)$, noting that v_0 enters as a coefficient of ODE 17 (in the constant d). In order to show this, we exploit the linearity of the ODE (17). Making the change of variables $u(z) = v(z) - v_1$, we have $u(T) = 0$. The dynamics of u do not depend on v_0, in that

$$u(z) + \alpha z(1+z)u'(z) = \beta(z), \tag{32}$$

where $\beta(z) = \bar{\lambda}(qzf(z) - c)/(r + 2\eta)$ and $\alpha = \bar{\lambda}/(r + 2\eta) > 0$ is positive on (T, ∞). Moreover, the limit u_∞ is by construction equal to $v_0 - v_1$. This allows us to re-express the consistency condition as $u_\infty = (r/2\eta + r)v_0$. Therefore, having integrated u over $[T, \infty)$, one may simply read off the values v_0 and v_1. The resulting function $v(t) = u(t) + v_1$ solves the initial HJB equation with a v_0-dependent coefficient, and also satisfies the smooth pasting condition.

Thus, for any admissible f, there is an optimal policy of the trigger form. To conclude, we will show that there are no policies solving the HJB equation that are not of the trigger form. This follows from the linearity in ℓ of the HJB equation, implying a bang-bang solution, which is strict because indifference is characterized by the equation $qzf(z) = c$, which has a unique solution by Lemma 2. This analysis is summarized as follows.

Proposition 9 *Suppose that the payout-rate function π is of the form $\pi(x) = a + k/x$. Then any equilibrium intermediation policy Λ corresponds to a trigger capital ratio T. That is $\Lambda(x, y) = \bar{\lambda}1_{\{x/y>T\}}$.*

G Existence of Equilibrium

So far, we have shown that any equilibrium must be of the trigger form. In this appendix we show that there exists such an equilibrium. Appendix H shows uniqueness of such equilibria.

For any candidate trigger capital ratio T, let $f(z \mid T)$ be the net expected gain from switching capital across markets under the policy with trigger T, given current market heterogeneity z. We need to show that there exists some T such that $qTf(T \mid T) = c$, that is, such that the intermediary ceases intermediation, given the switching gain function $f(\,\cdot\,) = f(\,\cdot \mid T)$, exactly when $z = T$. It suffices to show that $Tf(T \mid T)$ takes all values between 0 and ∞ as T varies from 1 to ∞.

Because $zf(z)$ is increasing, equation (12) implies that

$$f(z \mid T) \geq \frac{(T-1)}{T(r+2\eta)}, \quad z \geq T.$$

This implies that $Tf(T \mid T) \geq (T-1)/(r+2\eta)$. We note that the lower bound grows linearly with T. Because $Tf(T \mid T) = 0$ for $T = 1$, we know that $T \mapsto Tf(T \mid T)$ goes from 0 to ∞ as T goes from 0 to ∞. This function is continuous, so there exists some T^* such that $T^* f(T^* \mid T^*) = c/q$.

Proposition 10 *Suppose that the payout-rate function π is of the form $\pi(x) = a + k/x$. Then, there exists an equilibrium with a trigger policy.*

H Proof of Uniqueness of Trigger

Proof of Proposition 4. Suppose that trigger levels S and T, $S < T$, both satisfy the equations of the proposition. Let $\phi(z) = f^T(z) - f^S(z)$ denote the difference between the gains from switching capital under policies S and T, as a function of z. (Throughout, we use superscripts to denote dependence on S or T.) From (18), $S < T$ implies that $g_0^S > g_0^T$. Optimality of S (respectively T) with respect to f^S (respectively, f^T) implies that, for any z in $(S, T]$,

$$qzf^S(z) - c - z(1+z)(v^S)'(z) > 0$$

and

$$qzf^T(z) - c - z(1+z)(v^T)'(z) \leq 0.$$

Because $(v^T)'(z) = 0$ for z in this interval $(S < T]$, while $(v^S)'(z) \geq 0$ by Proposition 3, we know that $\phi(T) < 0$. Subtracting the version of equation (10) for T from the version of the same equation for S yields

$$(a+z)\phi + z(1+z)\phi' = \alpha \left(1 - \frac{1}{z}\right), \quad z > T, \tag{33}$$

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 $\phi(T) < 0$, this[13] implies that $\phi < 0$ for $z > T$, so that ϕ is everywhere negative.

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

$$g_0 = \frac{2}{r} - \Phi_0, \tag{34}$$

where Φ_0 is the expected discounted value of all future fees that investors will pay to the intermediary. (Recall that $2/r$ is the expected discounted stream of dividends paid on both market; see the "conservation equation" (9).) We have seen that $\phi < 0$, that is, $f^S(z) > f^T(z)$ for all $z > T$. This means that investors pay, for any z, more fees with S than with T for $z > T$. Moreover, for $z \in [S, T]$, investors pay fees (which are positive, from Proposition 2) for trigger S, whereas they pay nothing for trigger T. Therefore, $\Phi_0^S > \Phi_0^T$, which implies from (34) that $g_0^S < g_0^T$, a contradiction. ■

I Algorithm for Trigger Calculation

In general, (18) provides the following fixed-point algorithm for computing the equilibrium trigger capital ratio T.

Combining (17), the equation obtained by differentiating (17), as well as the equation (13) for f, yields the second-order linear ordinary differential equation for v:

$$\alpha v(z) + (\beta + 2z)z(1 + z)v'(z) + z^2(1 + z)^2 v''(z) = \omega + \delta z, \quad z \geq T, \tag{35}$$

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, $v_0 = V(1, 0)$.

1. Start with some candidate value for v_0, which we call v^0. From (9) and (18) 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, (16) provides a corresponding value for $v(T^0)$.

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

2. Starting with the initial conditions $v(T^0)$ and $v'(T^0) = 0$, evaluate a candidate for $v(\infty) = \lim_{z\to\infty} v(z)$ by integration of the differential equation (35) 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.[14]

J HJB Analysis with Partial Recovery

In this appendix, we analyze the Hamilton-Jacobi-Bellman equation for the case of general proportional losses. The equation for G is given in Section 3.5. The equation for H is

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

We let $\Phi(\,\cdot\,)$ denote the cumulative recovery-rate distribution function associated with the fractional event loss measure ν. That is, $\Phi(u) = 1 - \nu([0,u])$. We let $g(z) = G(z,1)$ and $h(z) = H(z,1)$, obtaining the coupled equations

$$\begin{aligned} (r + 2\eta + \Lambda(z,1)z)g(z) + \Lambda(z,1)(1+z)zg'(z) = {} & \frac{1}{z} + \Lambda(z,1)(1-q)(h(z) - g(z)) \\ & + \eta\left[\int_{1/z}^{1} ug(uz)\,d\Phi_u + \int_0^{1/z} \frac{1}{z}h\left(\frac{1}{uz}\right)\,d\Phi_u + \int_0^1 \frac{1}{u}g\left(\frac{z}{u}\right)\,d\Phi_u\right] \quad (36) \end{aligned}$$

[14] One can prove that $v' \sim \log(z)/z^2$ as z goes to ∞. Unfortunately, this convergence rate is not particularly fast. A possible improvement is to integrate v numerically up to some value $\hat{z}$ above which non-dominant terms in (35) 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

$$(r+2\eta+\Lambda(z,1)z)h(z)+\Lambda(z,1)(1+z)zh'(z)$$
$$= 1 \,+\, \eta\left[\int_{1/z}^{1} h(uz)\,d\Phi_u + \int_0^{1/z} \frac{1}{uz} g\left(\frac{1}{uz}\right)\,d\Phi_u + \int_0^1 h\left(\frac{z}{u}\right)\,d\Phi_u\right]. \quad (37)$$

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

$$0 = \sup_{\ell\in[0,\overline{\lambda}]} \Bigg\{ -rv(z) - \ell z v'(z) - \ell z^2 v'(z) + \ell(qzf(z)-c)$$
$$+\eta\left[\int_{1/z}^{1} v(uz)\,d\Phi_u + \int_0^{1/z} v\left(\frac{1}{uz}\right)\,d\Phi_u + \int_0^1 v\left(\frac{z}{u}\right)\,d\Phi_u - 2v(z)\right]\Bigg\}. \quad (38)$$

The equation reduces to

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

and

$$(r+2\eta)v(z)+\overline{\lambda}(1+z)zv(z)' = [qzf(z)-c]\overline{\lambda}$$
$$+\eta\left[\int_{1/z}^{1} v(uz)\,d\Phi_u + \int_0^{1/z} v\left(\frac{1}{uz}\right)\,d\Phi_u + \int_0^1 v\left(\frac{z}{u}\right)\,d\Phi_u\right], \quad z\geq T. \quad (39)$$

The smooth-pasting condition is

$$(1+T)Tv'(T) = qTf(T) - c. \quad (40)$$

K 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 integro-differential equations for g, h, and v, which arise thanks to the special structure of our problem.

K.1 Primitives

The parameters are r, η, $\overline{\lambda}$, q, c, and the recover rate distribution function $\Phi : [0,1] \to [0,1]$, a beta distribution with given parameters. The algorithm will determine the trigger level T for intermediation and the value functions g, h, and v.

K.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 (40) 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.

K.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 $z_i = \delta^i$, $i \in \{0, \ldots, n\}$, 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, $g = (g_0, \ldots, g_n)$ and $h = (h_0, \ldots, h_n)$ 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 $\lambda(z,T) = \overline{\lambda} 1_{z>T}$. Thus, $\lambda = \overline{\lambda}$ if $z > T$ and 0 otherwise.

K.4 Discretization conventions

For any $0 \leq u < u' \leq 1$, we let $K(u,u') = \Phi(u') - \Phi(u)$ 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 $h_{n+1} = h_n$.

K.5 Discretized Equations

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

$$
\begin{aligned}
g_i[r + 2\eta + \lambda_i(z_i + (1-q))] &+ g_{i+1}\frac{\lambda_i z_i(1+z_i)}{z_{i+1} - z_{i-1}} + g_{i-1}\frac{-\lambda_i z_i(1+z_i)}{z_{i+1} - z_{i-1}} + h_i(q-1)\lambda_i \\
&- \eta \sum_{j=0}^{i} g_j \frac{z_j}{z_i} K\left(\frac{z_{j-1} + z_j}{2z_i}, \frac{z_j + \min\{z_{j+1}, z_i\}}{2z_i}\right) \\
&- \eta \sum_{j=0}^{n} \frac{h_j}{z_i z_j} K\left(1_{j<n}\left(\frac{1}{2z_i z_j} + \frac{1}{2z_i z_{j+1}}\right), \frac{1}{2z_i z_j} + \frac{1}{2z_i z_{j-1}}\right) \\
&- \eta \sum_{j=i}^{n} g_j \frac{z_j}{z_i} K\left(1_{j<n}\left(\frac{z_i}{2z_j} + \frac{z_i}{2z_{j+1}}\right), \frac{z_i}{2z_j} + \frac{z_i}{2\max\{z_{j-1}, z_i\}}\right) = \frac{1}{z_i}. \quad (41)
\end{aligned}
$$

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

$$
\begin{aligned}
h_i[r + 2\eta + \lambda_i z_i] &+ h_{i+1}\frac{\lambda_i z_i(1+z_i)}{z_{i+1} - z_{i-1}} + h_{i-1}\frac{-\lambda_i z_i(1+z_i)}{z_{i+1} - z_{i-1}} \\
&- \eta \sum_{j=0}^{i} h_j K\left(\frac{z_{j-1} + z_j}{2z_i}, \frac{z_j + \min\{z_{j+1}, z_i\}}{2z_i}\right) \\
&- \eta \sum_{j=0}^{n} g_j z_j K\left(1_{j<n}\left(\frac{1}{2z_i z_j} + \frac{1}{2z_i z_{j+1}}\right), \frac{1}{2z_i z_j} + \frac{1}{2z_i z_{j-1}}\right) \\
&- \eta \sum_{j=i}^{n} h_j K\left(1_{j<n}\left(\frac{z_i}{2z_j} + \frac{z_i}{2z_{j+1}}\right), \frac{z_i}{2z_j} + \frac{z_i}{2\max\{z_{j-1}, z_i\}}\right) = 1. \quad (42)
\end{aligned}
$$

K.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.

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

$$
\begin{aligned}
v_i[r + 2\eta] + v_{i+1}\frac{\lambda_i z_i(1+z_i)}{z_{i+1} - z_{i-1}} + v_{i-1}\frac{-\lambda_i z_i(1+z_i)}{z_{i+1} - z_{i-1}} \\
- \eta \sum_{j=0}^{i} v_j K\left(\frac{z_{j-1} + z_j}{2z_i}, \frac{z_j + \min\{z_{j+1}, z_i\}}{2z_i}\right) \\
- \eta \sum_{j=0}^{n} v_j K\left(1_{j<n}\left(\frac{1}{2z_i z_j} + \frac{1}{2z_i z_{j+1}}\right), \frac{1}{2z_i z_j} + \frac{1}{2z_i z_{j-1}}\right) \\
- \eta \sum_{j=i}^{n} v_j K\left(1_{j<n}\left(\frac{z_i}{2z_j} + \frac{z_i}{2z_{j+1}}\right), \frac{z_i}{2z_j} + \frac{z_i}{2\max\{z_{j-1}, z_i\}}\right) = \lambda_i[q z_i(h_i - g_i) - c]. \quad (43)
\end{aligned}
$$

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

K.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 (40) yields the condition, for $T = z_t$

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

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.

L 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, $\tilde{H}(x,y,\alpha) = H(x,y)$, 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

$$
\begin{aligned}
&- r\tilde{G}(x,y,\alpha) + \alpha\pi(x) - \tilde{G}_x(x,y,\alpha)x\Lambda(x,y) + \tilde{G}_y(x,y,\alpha)x\Lambda(x,y) \\
&\quad + (1-q)\eta(\tilde{H}(x,y,\alpha) - \tilde{G}(x,y,\alpha)) - \eta\tilde{G}(x,y,\alpha) + \eta(\tilde{G}(x,0,\alpha) - \tilde{G}(x,y,\alpha)) \\
&\qquad + \frac{1}{2}\sigma^2[\tilde{G}_{xx}(x,y,\alpha)x^2 + \tilde{G}_{yy}(x,y,\alpha)y^2 + \tilde{G}_{\alpha\alpha}(x,y,\alpha)\alpha^2] \\
&\qquad\quad + \sigma^2[xyR\tilde{G}_{xy}(x,y,\alpha) + x\alpha\tilde{G}_{x\alpha}(x,y,\alpha) + y\alpha R\tilde{G}_{y\alpha}(x,y,\alpha)] = 0
\end{aligned} \tag{44}
$$

and

$$
\begin{aligned}
&- r\tilde{H}(x,y,\alpha) + \alpha\pi(y) - \tilde{H}_x(x,y,\alpha)x\Lambda(x,y) + \tilde{H}_y(x,y,\alpha)x\Lambda(x,y) \\
&\qquad + \eta(\tilde{G}(y,0,\alpha) - \tilde{H}(x,y,\alpha)) - \tilde{\eta}H(x,y,\alpha) \\
&\quad + \frac{1}{2}\sigma^2[\tilde{H}_{xx}(x,y,\alpha)x^2 + \tilde{H}_{yy}(x,y,\alpha)y^2 + \tilde{H}_{\alpha\alpha}(x,y,\alpha)\alpha^2] \\
&\qquad + \sigma^2[xyR\tilde{H}_{xy}(x,y,\alpha) + x\alpha R\tilde{H}_{x\alpha}(x,y,\alpha) + y\alpha\tilde{H}_{y\alpha}(x,y,\alpha)] = 0,
\end{aligned} \tag{45}
$$

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):

$$\begin{aligned}
&- rG(x,y) + \pi(x) - G_x(x,y)x\Lambda(x,y) + G_y(x,y)x\Lambda(x,y) \\
&\qquad + (1-q)\eta(H(x,y) - G(x,y)) - \eta G(x,y) + \eta(G(x,0) - G(x,y)) \\
&+ \frac{1}{2}\sigma^2[G_{xx}(x,y)x^2 + G_{yy}(x,y)y^2] + \sigma^2[xyRG_{xy}(x,y) + xG_x(x,y) + yRG_y(x,y)] = 0
\end{aligned} \tag{46}$$

and

$$\begin{aligned}
&- rH(x,y) + \pi(y) - H_x(x,y)x\Lambda(x,y) + H_y(x,y)x\Lambda(x,y) \\
&\qquad + \eta(G(y,0) - H(x,y)) - \eta H(x,y) + \frac{1}{2}\sigma^2[H_{xx}(x,y)x^2 + H_{yy}(x,y)y^2] \\
&\qquad\qquad + \sigma^2[xyRH_{xy}(x,y) + xRH_x(x,y) + yH_y(x,y)] = 0.
\end{aligned} \tag{47}$$

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 $(x,y) = (z,1)$,

$$\begin{aligned}
&\frac{1}{2}\sigma^2[F_{xx}(x,y)x^2 + F_{yy}(x,y)y^2 + 2zF_{xy}(x,y)] \\
&\qquad = \sigma^2\left[\frac{\gamma}{2}(\gamma+1)f(z) + (\gamma+1)(1-R)f'(z) + (1-R)z^2f''(z)\right].
\end{aligned} \tag{48}$$

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

$$\begin{aligned}
&\frac{1}{2}\sigma^2[F_{xx}(x,y)x^2 + F_{yy}(x,y)y^2 + 2xyF_{xy}(x,y)] \\
&\qquad = \sigma^2[f(z) + 2(1-R)zf'(z) + (1-\rho)z^2f''(z)].
\end{aligned} \tag{49}$$

M Connectedness

In this appendix, we outline a model with the natural feature that an investor is increasingly likely to be in contact with multiple intermediaries at the point of bargaining as the total number of intermediaries is increased.

Suppose that there is an advertising medium handling intermediary ads. An intermediary's effort corresponds to the probability p that its advertisement will place the intermediary in contact with an investor at the time at which the investor checks the medium. We assume that p is bounded by some capacity constraint $\bar{p} < 1$. Each investor, pairwise independently across investors, has some exogenous intensity χ for the times of monitoring his capital and observing the advertising medium.[15] This is consistent with the framework of our main model: The intensity of times at which an investor is contacted by least one intermediary is $\chi\phi_n(p)$, where

$$\phi_n(p) = 1 - (1-p)^n.$$

Then, $\overline{\lambda} = \chi\phi_n(\bar{p})$ is the intermediation capacity parameter of the basic model. Assuming that a well-connected investor initiates bargaining with a randomly selected intermediary from among those contacted, each intermediary has maximal contact intensity $\overline{\lambda}/n$. The probability that, when in contact with an intermediary, an investor is in contact with at least two intermediaries is

$$\psi_n(p) = 1 - (1-p)^n - np(1-p)^{n-1}.$$

For a fixed $\overline{\phi} \in [0,1]$, let $\bar{p}_n$ solve $\phi_n(\bar{p}_n) = \overline{\phi}$, so that $\bar{\lambda}$ is independent of n, as in our basic model. One may easily check that $\bar{p}_n$ is decreasing in n. Moreover, using that

$$\psi_n(p_n) = \bar{\phi} - np_n(1-p_n)^{n-1} = \bar{\phi} - \frac{(1-\bar{\phi})np_n}{1-p_n},$$

one can show that $\psi_n(p_n)$ is increasing in n.[16] Therefore, keeping constant the flow of investors being contacted at any given time, the average number of intermediaries in contact with any given investor is increasing in n. As the number of intermediaries goes to infinity, the probability that investor is well connected is:

$$\lim_{n\to\infty} \psi_n(p_n) = \bar{\phi} + (1-\bar{\phi})\log(1-\bar{\phi}).$$

The second term is negative. This specification can be generalized to an arbitrary number of media, with the same result that ψ_n is increasing in n.

[15] At such times, the investor observes the medium and plays a bargaining game with advertised intermediaries. If bargaining breaks down, the investor leaves his capital in the large market, until the next monitoring time.

[16] In order to verify this, one is to show that $np_n/(1-p_n)$ is decreasing. Expressing p_n in terms of $\alpha = (1-\bar{\phi})^{-1} > 1$ and letting $x = 1/n$, this is equivalent to showing that $(\alpha^x - 1)/x$ is increasing in x. This is easily done by checking the positivity of the derivative, whose numerator is increasing in $u = \alpha^x$ and vanishes for $u = 1$.

N Proofs of Section 4.3

Proof of Proposition 7. As before, we let $g_0^L = G(1,0)$, under strategy L. For any equilibrium with aggregate mobility $z \mapsto L(z)$ and fee $z \mapsto q(z)$, one can easily modify the proof of Lemmas 1 and 2 to show that

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

is nonnegative and that $zf^L(z)$ is increasing in z. If $T_n \leq T_1$, we have

$$f^1(T_n) = \left(\frac{1+\eta g_0^1}{r+2\eta}\right)\left(1-\frac{1}{T_n}\right),$$

where f^1 and g_0^1 denote the corresponding quantities for the monopolistic case, since the intermediary does not search at T_n. Further,

$$f^L(T_n) \leq \left(\frac{1+\eta g_0^L}{r+2\eta}\right)\left(1-\frac{1}{T_n}\right),$$

from the nonnegativity of $\kappa^L(T_n)$. Therefore,

$$T_n\left(f^1(T_n) - f^L(T_n)\right) \geq \eta\left(\frac{g_0^1 - g_0^L}{r+2\eta}\right)(T_n - 1). \tag{50}$$

Since $g_0^L \leq 2/r$ for any policy, there exists some $\bar{\eta}$ such that for all $\eta < \bar{\eta}$, the righthand side of (50) is bounded in norm by ε whenever $T_n \leq T_1$, since we have an upper bound on T_1 from (18). Choosing ε below $(1/q(n) - 1/q(1))c$ and setting $\bar{\eta}$ accordingly, we have for any $T_n \leq T_1$,

$$q(1)T_n f^1(T_n) \geq \frac{q(1)}{q(n)}\left(q(n)T_n f^L(T_n) - q(n)\varepsilon\right) \geq \frac{q(1)}{q(n)}(c - q(n)\varepsilon) > c, \tag{51}$$

which shows that it is strictly optimal for the monopolist to search at T_n, contradicting the assumption that $T_n \leq T_1$.

Proof of Proposition 8. At S_n, it cannot be strictly profitable for an intermediary to deviate by continuing to search and receive the net payoff $q(1)S_n f^L(S_n) - c$ per unit of effort, but it was profitable to some intermediaries to search at a capital heterogeneity just above S_n. This implies that S_n must satisfy the equation

$$q(1)S_n f^L(S_n) = c.$$

We recall that in the monopolistic case, T_1 satisfies the equation

$$q(1)T_1 f^1(T_1) = c.$$

Therefore, it suffices to show that the roots of these two equations are arbitrarily close if η is arbitrarily small. We have

$$f^L(S_n) = \left(\frac{1+\eta g_0^L}{r+2\eta}\right)\left(1-\frac{1}{S_n}\right)$$

and

$$f^1(T_1) = \left(\frac{1+\eta g_0^1}{r+2\eta}\right)\left(1-\frac{1}{T_1}\right).$$

Therefore, S_n and T_1 must satisfy

$$\left(\frac{1+\eta g_0^L}{r+2\eta}\right)(S_n-1) - \left(\frac{1+\eta g_0^1}{r+2\eta}\right)(T_1-1) = 0,$$

which may be rewritten as

$$\left(\frac{1+\eta g_0^L}{r+2\eta}\right)(S_n-T_1) = \eta\left(\frac{g_0^L-g_0^1}{r+2\eta}\right)(1+T_1).$$

Since T_1 is uniformly bounded from (18) and the g_0's are uniformly bounded by $2/r$, the righthand side is less than ε if η is chosen small enough. The first factor of the lefthand side is equivalent to $1/r$ when η is small enough. Combining these observations shows that $|S_n - T_1| \leq \varepsilon$ for any arbitrary $\varepsilon > 0$, provided that η is small enough.

References

Andrade, S., C. Chang, and M. Seasholes, 2008, "Trading Imbalances, Predictable Reversals, and Cross-Stock Price Pressure," *Journal of Financial Economics*, 88, 406–423.

Ausubel, L. M., and R. J. Deneckere, 1989, "Reputation in Bargaining and Durable Goods Monopoly," *Econometrica*, 57, 511–531.

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.

Chaiserote, S., 2008, "Liquidity Provision by International Institutional Investors: Evidence from Brazilian Privatization," Stanford University, Working Paper.

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.

Chen, Z., A. Lookman, and N. Schürhoff, 2008, "Why Ratings Matter: Evidence from Lehman's Index Inclusion Rule Change," University of Lausanne, Working Paper.

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.

——— , 2007, "Valuation in Over-the-Counter Markets," *Review of Financial Studies*, 20, 1865–1900.

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.

Green, R., 2007, "Issuers, Underwriter Syndicates, and Aftermarket Transparency," *Journal of Finance*, 62, 1529–1550.

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, 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 Curces for Stocks do Slope Down: New Evidence from an Index Weights Adjustment," *Journal of Finance*, 55, 893–912.

Kulak, J. P., 2008, "Price-Pressures And Foreign Investor Biases Revealed Through Firms: Geographical Placement Choices," University of Lausanne, Working Paper.

Lagos, R., G. Rocheteau, and P.-O. Weil, 2008, "Crashes and Recoveries in Illiquid Markets," Working Paper, New York University.

Lou, D., 2009, "A Flow-Based Explanation for Return Predictability," Working Paper, Yale University.

Mailath, G., and L. Samuelson, 2006, *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press.

Mitchell, M., L. H. Pedersen, and T. Pulvino, 2007, "Slow Moving Capital," *American Economic Review, Papers and Proceedings*, 97, 215–220.

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., 1982, "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50, 97–108.

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

Scholes, M., 1972, "The Market for Securities: Substitution verus Price Pressure and the Effects of Information on Share Prices," *Journal of Business*, 45, 179211.

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

Stole, L., and J. Zwiebel, 1996, "Intra-Firm Bargaining under Non-Binding Contracts," *Review of Economic Studies*, 63, 375–410.

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

Sutton, J., 1986, "Non-Cooperative Bargaining Theory: An Introduction," *Review of Economic Studies*, 53, 709–724.

Vayanos, D., and T. Wang, 2007, "Search and Endogenous Concentration of Liquidity in Asset Markets," *Journal of Economic Theory*, 136, 66–104.

Weill, P.-O., 2007, "Leaning Against the Wind," *Review of Economic Studies*, 74, 1329–1354.

——— , 2008, "The Liquidity Premium in a Dynamic Bargaining Market," *Journal of Economic Theory*, pp. 66–96.

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