Financially Constrained Arbitrage and Cross-Market Contagion

Denis Gromb*
INSEAD and CEPR

Dimitri Vayanos†
LSE, CEPR and NBER

February 2015

Abstract

We propose a continuous time infinite horizon equilibrium model of financial markets in which arbitrageurs have multiple valuable investment opportunities but face financial constraints. The investment opportunities, heterogeneous along different dimensions, are provided by pairs of similar assets trading at different prices in segmented markets. By exploiting these opportunities, arbitrageurs alleviate the segmentation of markets, providing liquidity to other investors by intermediating their trades. We characterize the arbitrageurs' optimal investment policy, and derive implications for market liquidity and asset prices. We show that liquidity is smallest, volatility is largest, correlations between asset pairs with uncorrelated fundamentals are largest, and correlations between asset pairs with highly correlated fundamentals are smallest for intermediate levels of arbitrageur wealth.

Keywords: Financial constraints, arbitrage, liquidity, contagion.

*denis.gromb@insead.edu
†d.vayanos@lse.ac.uk

1 Introduction

The ongoing crisis has highlighted the importance of intermediary capital for the functioning of financial markets. Indeed, the large losses banks incurred in the subprime market has led them to cut their lending across the board, notably their financing of other intermediaries, causing liquidity to dry up in many otherwise unrelated markets. Central banks the world around struggled to deal with a combined banking liquidity and financial market liquidity crisis.

This paper develops a framework to examine the relation between intermediary capital, financial market liquidity and asset prices. The framework itself has three main features.

First, we model arbitrageurs as specialized investors able to exploit profitable trades that other, less sophisticated market participants cannot access directly as easily or quickly. Arbitrageurs are to be understood here as individuals and institutions responsible for providing liquidity in different financial markets. At the same time, arbitrage is assumed to require capital to which arbitrageurs have only limited access, i.e., they face financial constraints. These financial constraints, be they margin requirements, limited access to external capital or barriers to entry of new capital, affect the arbitrageurs' investment capacity.

Second, ours is a dynamic general equilibrium model. On the one hand, arbitrageurs' capital affects their ability to provide liquidity, which is ultimately reflected in asset prices. On the other hand, asset price movements determine arbitrage profits and, therefore, arbitrageurs' capital. This dynamic interaction shapes arbitrageurs' investment policies, asset prices and market liquidity.

Third, in our model, arbitrageurs face multiple arbitrage opportunities with different characteristics, across which they must allocate their scarce capital. This aspect is important to study the cross-sectional properties of arbitrageurs' optimal investment policy, as well as those of market liquidity and asset prices. In particular, it allows us to analyze phenomena of price contagion and liquidity linkages across markets.

We aim to analyze a number of questions relative to arbitrageurs' investment strategy. To start with, what is the optimal investment strategy of an arbitrageur with financial constraints? How is the need for risk management created by financial constraints resolved when there are multiple arbitrage opportunities with different characteristics? How does an arbitrageur's optimal investment policy respond to shocks to their capital?

More importantly, we are interested in questions about asset prices and market liquidity. Financial constraints lead to wealth effects creating price and liquidity linkages across markets. Which asset or trade characteristics make them more sensitive to changes in arbitrageurs' capital? How much time-variation in convergence spreads is explained by contagion vs. fundamentals? Is diversification of arbitrageurs effective despite contagion effects?

Our model's building block is as in Gromb and Vayanos (2002). There are two risky assets paying similar (possibly identical) dividends but traded in segmented markets. The demand by investors on each market for the local risky asset is affected by endowment shocks that covary with the asset's payoff. Since the covariances differ across the two markets, the assets' prices can differ. Said differently, the investors in the segmented markets would benefit from trading with each other to improve risk sharing. However, there is no liquidity due to the assumed segmentation.

This unsatisfied demand for liquidity creates a role for arbitrageurs. We model arbitrageurs as competitive specialists able to invest across markets and thus exploit price discrepancies between the risky assets. Doing so, they facilitate trade between otherwise segmented investors, providing liquidity to them. Arbitrageurs, however, face financial constraints in that their risky asset positions must be collateralized separately with a position in the riskfree asset. Given these constraints, the arbitrageurs' ability to provide market liquidity depends on their wealth. The arbitrageurs' wealth is to be understood as the pool of capital they can access frictionlessly. In that case, there is no distinction between arbitrageurs' internal funds and the "smart capital" they raise externally. If this total pool of capital is insufficient, arbitrageurs may be unable to provide perfect liquidity.

Based on that building block, we develop a continuous time general equilibrium model in which competitive arbitrageurs face at each point in time several arbitrage opportunities, i.e., multiple asset pairs as above. These opportunities are heterogeneous along different dimensions (e.g., volatility, market size, margin requirements). Due to their financial constraints, arbitrageurs face a complex investment problem. On the one hand, they must allocate their scarce capital across opportunities and over time. On the other hand, the performance of these investments affect their investment capacity.

To begin with, we study the case of riskless arbitrage, in which two assets in a pair pay the exact same dividends. In this case, we are able to derive all equilibrium variables in closed form. This allows us to draw many implications. The following are but examples.

Some implications are cross-sectional in nature, i.e., comparing variables across opportunities with different characteristics. For instance, we show that opportunities with higher margin requirements are more illiquid, offer higher excess returns and have higher risk premia. The intuition is that investment opportunities requiring arbitrageurs to tie up more capital as collateral must provide them with a greater reward, i.e., a higher excess return. Risk premia being the present value of future excess returns they must be higher for such opportunities. In our model, risk premia are a measure of the illiquidity arbitrageurs do not eliminate. Therefore, illiquidity is higher for opportunities with higher margin requirements.

Other implications involve comparative statics, in particular with respect to arbitrageur wealth. One way to interpret these results is as the effect of an unanticipated exogenous shock to arbitrageur wealth. For instance, we show that illiquidity and risk premia are more sensitive to arbitrageur wealth for opportunities with higher margin requirements. Intuitively, changes in arbitrageur wealth affect the excess return (current or future) per unit of collateral, and therefore impact more opportunities with higher collateral requirements.

Next, we analyze the case of risky arbitrage. There are two sources of arbitrage risk: fundamental risk and supply risk. Fundamental risk means that the assets in a given pair may not pay the exact same dividends. Fundamental shocks affect arbitrage profits, and therefore arbitrageur wealth and ultimately assets prices and liquidity. Supply risk means that the demand for liquidity may not be predictable. Shocks to the demand for liquidity affect risk premia both directly, i.e., holding arbitrageur wealth constant, and indirectly because changes in premia affect arbitrageur's profit and therefore arbitrageur wealth.

First, we show that the arbitrageurs' financial constraints create a linkage across otherwise independent assets, i.e., fundamental and supply shocks to one opportunity affect all opportunities' risk premia. The linkage goes through arbitrageur wealth. Indeed, a fundamental shock to one opportunity affects the dividend arbitrageurs derive from that opportunity, and hence arbitrageur's wealth. Similarly, a supply shock to one opportunity affects that opportunity's risk premium and hence the capital gains arbitrageurs realize from their investment in that opportunity. In both cases, a change in arbitrageur wealth affects all other opportunities' risk premia.

To derive further implications, we consider the effect of small shocks, i.e., we study equilibrium variables around the riskfree arbitrage equilibrium. We characterize the effect of different shocks on arbitrageurs' portfolios, market liquidity, the volatility of asset prices as well as the correlation between the prices of different assets.

We show that liquidity, volatility and correlations are generally non-monotonic in arbitrageur wealth. (Here an asset's liquidity is defined as the inverse of the impact a supply shock would have on the asset return.) If arbitrageur wealth is high, as arbitrageur wealth increases, liquidity increases, while price volatility decreases for all assets. As for correlations, they decrease for assets whose fundamentals are uncorrelated, and increase for assets whose fundamentals are correlated. Maybe more surprisingly, these relationships are reversed when arbitrage wealth is low. In that case, as arbitrageur wealth increases, liquidity decreases, while price volatility increases for all assets. At the same time, correlations increase for assets whose fundamentals are uncorrelated, and decrease for assets whose fundamentals are correlated.

The reason for this reversal is that for high levels of wealth, arbitrageurs are unlikely to be constrained and therefore their positions are not very sensitive to wealth. In that case, a drop in wealth does not reduce much the positions arbitrageurs take and therefore does not affect much the amount of wealth arbitrageurs put at risk. The opposite holds for low levels of wealth. In that case, a drop in wealth leads to a large reduction in arbitrageurs' positions, reducing the amount of wealth arbitrageurs put at risk.

Relation to the Literature (Incomplete)

Our analysis builds on the recent literature on the limits to arbitrage and, more particularly, on financially constrained arbitrage.[1,2] Gromb and Vayanos (2002) introduce a model of arbitrageurs providing liquidity across two segmented markets but facing collateral-based financial constraints. Their setting is dynamic, i.e., they consider explicitly the link between arbitrageurs' past performance and their ability to provide market liquidity, and how arbitrageurs take this link into account in their investment decision. They also conduct a welfare analysis. This paper extends the analysis by considering multiple investment opportunities.[3]

The analysis of Gromb and Vayanos (2002) is extended to multiple investment opportunities in a static setting by Brunnermeier and Pedersen (2009) who show how financial constraints imply that shocks propagate and liquidity co-moves across markets. In contrast, ours is a dynamic setting. Kyle and Xiong (2001) obtain similar financial contagion effects driven by the wealth of arbitrageurs. These arise not from financial constraints but from arbitrageurs' logarithmic utility implying that their demand for risky assets is increasing in wealth.

The paper proceeds as follows. Section 2 presents the model. Sections 3 and 4 study riskless and risky arbitrage respectively. Section 5 concludes. The Appendix contains mathematical proofs.

2 The Model

2.1 Assets

Time t is continuous and goes from zero to infinity. There is a set $\mathcal{I}$ of risky assets and a riskless asset. The risky assets come in pairs, and we denote by $-i$ the other asset in asset i's pair. The payoff $dD_{i,t}$ of asset i between time t and $t+dt$ is

$$dD_{i,t} = D_i dt + \sigma_i dB_{i,t} + \sigma_i^f dB_{i,t}^f, \tag{1}$$

[1] Alternative theories of the limits to arbitrage are generally based on incentive problems in delegated portfolio management or bounded rationality of investors.

[2] Here, we discuss the relation of our paper to only the closest literature. It is however connected to a broader set of contributions which, given binding (time) constraints, we intend to discuss in future versions. Gromb and Vayanos (2010) survey the theoretical literature on the limits of arbitrage.

[3] Also, the model is cast in an infinite horizon, rather than a finite horizon.

where $(D_i, \sigma_i, \sigma_i^f)$ are constants and $(B_{i,t}, B_{i,t}^f)$ are Brownian motions. The constant D_i is asset i's instantaneous expected payoff. The Brownian motions $(B_{i,t}, B_{i,t}^f)$ are common to assets i and $-i$, i.e., $B_{i,t} = B_{-i,t}$ and $B_{i,t}^f = B_{-i,t}^f$. The effect of $B_{i,t}$ is the same on the two assets, i.e., $\sigma_i = \sigma_{-i}$, while that of $B_{i,t}^f$ is opposite, i.e., $\sigma_i^f = -\sigma_{-i}^f$. Therefore, assets i and $-i$ have identical payoffs if $\sigma_i^f = 0$, and have correlated but different payoffs if $\sigma_i^f \neq 0$. In both cases, we refer to the asset pair as an arbitrage opportunity because in equilibrium arbitrageurs engage in a "spread trade," holding a long position in one asset and an equal short position in the other. Arbitrage opportunity i, corresponding to the asset pair $(i, -i)$, has fundamental risk if assets i and $-i$ have different payoffs ($\sigma_i^f \neq 0$).

We allow for a general correlation structure between the fundamental risks of different arbitrage opportunities, and denote by $\rho_{i,j}^f$ the instantaneous correlation between $(B_{i,t}^f, B_{j,t}^f)$ for assets (i, j) not in the same pair. The correlation between $(B_{i,t}, B_{j,t})$ does not matter for our analysis. The correlation between $(B_{i,t}, B_{j,t}^f)$ is also unimportant, and we set it to zero for all assets (i, j).

We assume that all risky assets are in zero net supply. This assumption simplifies our analysis because it helps ensure that arbitrageurs hold opposite positions in two assets in a pair. We treat the return of the riskless asset as exogenous, and denote by r the continuously compounded riskless return. This assumption is also for simplicity and because our focus is on the price discrepancies between risky assets. The price of asset i is endogenously determined in equilibrium, and we denote it by $p_{i,t}$.

2.2 Outside Investors

Assets can be traded by outside investors and arbitrageurs. Outside investors face market segmentation and can invest only in a subset of the risky assets. For simplicity, we assume that segmentation takes an extreme form, whereby each outside investor can only invest in one specific risky asset and in the riskless asset.[4] We refer to the outside investors who can invest in asset i as i-investors. We take market segmentation as given, i.e., assume that i-investors face prohibitively large transaction costs for investing in any risky asset other than asset i. These costs can be due to unmodelled physical factors (e.g., distance), information asymmetries or institutional constraints.

We assume that i-investors are competitive and form a continuum with measure μ_i. They maximize expected utility of intertemporal consumption, with utility being negative exponential,

[4] More generally, we can assume that outside investors can invest in multiple risky assets, provided that none of these assets belong to the same pair.

i.e.,

$$E_0\left[\int_0^\infty e^{-a_i c_{i,t}-\beta_i t}dt\right]. \tag{2}$$

Negative exponential utility simplifies our analysis by ruling out wealth effects.

Investors i and $-i$ are identical in terms of their measure ($\mu_i = \mu_{-i}$), risk-aversion coefficient ($a_i = a_{-i}$) and discount rate ($\beta_i = \beta_{-i}$). They differ, however, in their propensity to hold their respective assets because of an endowment that they receive. The covariance between this endowment and asset payoffs differs across investors, and this creates the different propensity to hold the assets. Moreover, the covariance can vary over time, and this can cause time variation in the price discrepancy between the assets.

The endowment of i-investors between time t and $t + dt$ is $u_{i,t}dD_{i,t}$. If the coefficient $u_{i,t}$ is positive, then the endowment covaries positively with the payoff of asset i, and this reduces the i-investors' willingness to hold the asset. If instead $u_{i,t}$ is negative, then the covariance is negative, and i-investors view asset i as a valuable hedge. We refer to $u_{i,t}$ as the i-investors' supply parameter: an increase in $u_{i,t}$ renders i-investors more willing to supply the asset.

Investors i and $-i$ have a different propensity to hold their respective assets if the supply parameters $(u_{i,t}, u_{-i,t})$ differ. We assume that these parameters are opposites, i.e.,

$$u_{-i,t} = -u_{i,t}. \tag{3}$$

This assumption is for simplicity: together with the zero-net-supply assumption, it ensures that arbitrageurs hold opposite positions in two assets in a pair.

To reduce the number of cases, we assume that the supply parameter of one asset in a pair is always positive (and hence the other is always negative). We denote by $\mathcal{A}$ the set of assets with positive supply parameters, i.e., $\mathcal{A} \equiv \{i \in \mathcal{I} : u_{i,t} > 0\}$. We assume that the supply parameter of an asset in $\mathcal{A}$ follows the process

$$du_{i,t} = \kappa_i^u(u_i - u_{i,t})dt + \sigma_i^u f(u_{i,t})dB_{i,t}^u, \tag{4}$$

where $\kappa_i^u > 0$, $u_i > 0$ and σ_i^u are constants, $f(u_{i,t})$ is a function such that $u_{i,t}$ is always positive (e.g., $f(u_{i,t}) = \sqrt{u_{i,t}}$), and $B_{i,t}^u$ is a Brownian motion. Arbitrage opportunity i has supply risk if $u_{i,t}$ is stochastic, i.e., $\sigma_i^u \neq 0$.

We allow for a general correlation structure between the supply risks of different arbitrage opportunities, and denote by $\rho_{i,j}^u$ the instantaneous correlation between $(B_{i,t}^u, B_{j,t}^u)$ for assets (i, j) not in the same pair. The correlation between $(B_{i,t}, B_{j,t}^u)$ is unimportant for our analysis, and we set it to zero for all assets (i, j). For simplicity, we also set the correlation between $(B^f_{i,t}, B^u_{j,t})$ to zero for all (i, j).

The optimization problem $\mathcal{P}_i$ of an i-investor is to choose a position $y_{i,t}$ in asset i to maximize (2) subject to the dynamic budget constraint

$$dw_{i,t} = r(w_{i,t} - y_{i,t}p_{i,t})dt + y_{i,t}(dD_{i,t} + dp_{i,t}) + u_{i,t}dD_{i,t}. \tag{5}$$

The first term in (5) is the return from the riskless asset, which receives dollar investment $w_{i,t} - y_{i,t}p_{i,t}$, the second term is the return from the risky asset, and the third term is the endowment.

2.3 Arbitrageurs

Outside investors have different propensities to hold their respective assets, but cannot realize the potential gains from trade due to market segmentation. This unsatisfied demand for liquidity creates a role for arbitrageurs, who can invest across markets and exploit price discrepancies between assets. We assume that arbitrageurs are competitive and form a continuum with measure one.[5] Unlike outside investors, they can invest in all risky assets and in the riskless asset. They maximize expected utility of intertemporal consumption, with utility being logarithmic, i.e.,

$$E_0 \left[\int_0^\infty \log(c_t) e^{-\beta t} dt \right]. \tag{6}$$

Logarithmic utility simplifies our analysis by ensuring that arbitrageurs' consumption is a constant fraction of their wealth regardless of the return on their opportunities, which is time-varying. The constant fraction of wealth consumed by arbitrageurs can be interpreted not only as consumption per se, but also as a proportional cost of running an arbitrage business.[6]

The optimization problem $\mathcal{P}$ of an arbitrageur is to choose positions $\{x_{i,t}\}_{i\in\mathcal{I}}$ in the risky assets to maximize (6) subject to a dynamic budget constraint and a financial constraint. The dynamic budget constraint is

$$dW_t = r\left(W_t - \sum_{i\in\mathcal{I}} x_{i,t}p_{i,t}\right)dt + \sum_{i\in\mathcal{I}} x_{i,t}(dD_{i,t} + dp_{i,t}) - c_t dt, \tag{7}$$

where W_t denotes the arbitrageur's wealth. The first term in (7) is the return from the riskless

[5] By fixing the measure of the arbitrageurs, we rule out entry into the arbitrage industry. This seems a reasonable assumption at least for understanding short-run market behavior.

[6] Logarithmic utility introduces wealth effects, in the form of wealth-dependent risk aversion. These effects are not present in the case of riskless arbitrage (Section 3), but arbitrageur wealth still matters because of the financial constraint (8). In the case of risky arbitrage (Section 4), the wealth effects introduced by logarithmic utility coexist with those introduced by the financial constraint.

asset, which receives dollar investment $W_t - \sum_{i \in \mathcal{I}} x_{i,t} p_{i,t}$, the second term is the return from the risky assets, and the third term is consumption.

We derive a financial constraint from the requirement that arbitrageurs must post riskless collateral to establish a position in each risky asset. We assume that a long or short position of $x_{i,t}$ shares of asset i requires collateral $m_i |x_{i,t}|$, where $m_i > 0$ is an exogenous margin. Since the total collateral that an arbitrageur must post cannot exceed his wealth W_t, the arbitrageur faces the financial constraint

$$W_t \geq \sum_{i \in \mathcal{I}} m_i |x_{i,t}|. \tag{8}$$

To keep the model symmetric, we assume that $m_i = m_{-i}$.

An endogenous derivation of the margin m_i can be found, for example, in Gromb and Vayanos (2002), who show (8) in the case of one asset pair. The margin is derived from the requirement that the position in each risky asset must be collateralized fully and separately from the positions in other assets. The margin satisfies $m_i = m_{-i}$ because of the symmetry of the model (zero net supply and opposite supply parameters), and is an increasing function of asset volatility. Gromb and Vayanos (2002) relate the assumption that arbitrageurs must collateralize their position in each asset separately to that of market segmentation.

The financial constraint (8) limits the arbitrageurs' investment capacity as a function of their wealth. The arbitrageurs' wealth is to be understood not only is their personal wealth, but also as the pool of capital they can access frictionlessly.

Arbitrageurs in our model act as intermediaries, exploiting price discrepancies between assets and providing liquidity to the other investors. Suppose, for example, that i-investors experience an increase in their supply parameter, in which case $-i$-investors experience a decrease. Then arbitrageurs buy asset i from i-investors and sell asset $-i$ to investors $-i$. Through this transaction arbitrageurs make a profit, while also providing liquidity to the other investors.

2.4 Equilibrium

Definition 1 *A competitive equilibrium consists of prices $\{p_{i,t}\}_{i \in \mathcal{I}}$, positions $y_{i,t}$ of the i-investors for all $i \in \mathcal{I}$, and positions $\{x_{i,t}\}_{i \in \mathcal{I}}$ of the arbitrageurs, such that:*

- *Given $\{p_{i,t}\}_{i \in \mathcal{I}}$, $y_{i,t}$ solves problem $\mathcal{P}_i$ for all $i \in \mathcal{I}$, and $\{x_{i,t}\}_{i \in \mathcal{I}}$ solve problem $\mathcal{P}$.*

- *The markets for all risky assets clear, i.e.,*

$$\mu_i y_{i,t} + x_{i,t} = 0 \quad \text{for all } i \in \mathcal{I}. \tag{9}$$

We define the risk premium $\phi_{i,t}$ of asset i as the difference between the present value of the asset's expected payoffs and price, i.e.,

$$\phi_{i,t} \equiv E_t \left[\int_t^{\infty} e^{-r(s-t)} dD_{i,s} \right] - p_{i,t} = \frac{D_i}{r} - p_{i,t}. \tag{10}$$

Definition 2 *A competitive equilibrium is symmetric if for each asset pair $(i, -i)$ the risk premia are opposites ($\phi_{-i,t} = -\phi_{i,t}$), the arbitrageurs' positions in the two assets are opposites ($x_{-i,t} = -x_{i,t}$), and so are the outside investors' positions ($y_{-i,t} = -y_{i,t}$).*

In the following sections we show that a symmetric competitive equilibrium exists. Existence follows because of our model's symmetry. Intuitively, the risk premia of assets i and $-i$ are opposites because the assets are in zero net supply and the supply shocks of investors i and $-i$ are opposites. The arbitrageurs' positions in the two assets are opposites because the risk premia are opposites. The outside investors' positions are also opposites because markets must clear. Note that symmetry and (10) imply that asset i's risk premium is one-half of the price wedge between assets $-i$ and i, i.e.,

$$\phi_{i,t} = \frac{p_{-i,t} - p_{i,t}}{2}. \tag{11}$$

We denote by

$$dR_{i,t} \equiv dD_{i,t} + dp_{i,t} - rp_{i,t}dt \tag{12}$$

the instantaneous return per share of asset i at time t in excess of the riskless asset, and refer to it simply as asset i's return. Using (1) and (10), we can write this return as

$$dR_{i,t} = r\phi_{i,t}dt + \sigma_i dB_{i,t} + \sigma_i^f dB_{i,t}^f - d\phi_{i,t}. \tag{13}$$

We refer to arbitrage opportunities by the asset i in the pair that is in $\mathcal{A}$, and by arbitrageurs' investment in an arbitrage opportunity by the investment $x_{i,t}$ in that asset.

3 Riskless Arbitrage

In this section we study the case where arbitrage opportunities have no fundamental risk and no supply risk. No fundamental risk means that the assets in each pair $(i, -i)$ have identical payoffs

($\sigma_i^f = 0$). No supply risk means that each supply parameter $u_{i,t}$ is deterministic ($\sigma_i^u = 0$). For simplicity, we also assume that $u_{i,t}$ is constant over time. Eq. (4) implies that a constant $u_{i,t}$ must equal to its long-run mean u_i.

In the absence of fundamental and supply risk, arbitrageurs earn a riskless return from spread trades. Indeed, no fundamental risk implies that spread trades are not affected by shocks to asset payoffs. Moreover, no supply risk implies that there are no shocks to supply parameters than can affect asset risk premia. Since arbitrageurs earn a riskless return, their wealth W_t is deterministic. Hence, the arbitrageurs' positions $x_{i,t}$, the outside investors' positions $y_{i,t}$, and asset risk premia $\phi_{i,t}$ are also deterministic. We confirm below that a symmetric equilibrium with deterministic $(W_t, x_{i,t}, y_{i,t}, \phi_{i,t})$ exists.

The case of riskless arbitrage yields non-trivial dynamics for arbitrageur wealth and asset risk premia. Indeed, wealth increases faster when risk premia are high, but an increase in wealth triggers a reduction in risk premia. We compute the dynamics of wealth and risk premia in closed form, and determine how premia depend on wealth, supply parameters and margin requirements. These results are useful not only for the analysis of riskless but also of risky arbitrage. Indeed, in Section 4 we use the closed-form solutions to derive properties of arbitrageurs' portfolios and asset prices when arbitrage risk is small.

3.1 Optimal Investment Policies

We first derive the optimal investment policies of outside investors and arbitrageurs. When the risk premium $\phi_{i,t}$ is deterministic, it only has a drift term that we denote by $\nu_{i,t}^{\phi}$:

$$d\phi_{i,t} \equiv \nu_{i,t}^{\phi} dt. \tag{14}$$

Eqs. (13), (14) and $\sigma_i^f = 0$ imply that asset i's return is

$$dR_{i,t} = \Phi_{i,t} dt + \sigma_i dB_{i,t}, \tag{15}$$

where

$$\Phi_{i,t} \equiv r\phi_{i,t} - \nu_{i,t}^{\phi}, \tag{16}$$

denotes the asset's expected return.

3.1.1 Outside Investors

Using (1), (12), (15), $\sigma_i^f = 0$ and $u_{i,t} = u_i$, we can simplify the budget constraint (5) of an i-investor to

$$dw_{i,t} = (rw_{i,t} + y_{i,t}\Phi_{i,t} + u_i D_i)\, dt + (y_{i,t} + u_i)\, \sigma_i dB_{i,t}. \tag{17}$$

We conjecture that the investor's value function is negative exponential as the expected utility, but with a different risk-aversion coefficient, i.e.,

$$V(w_{i,t}) = -e^{-A_i w_{i,t} - g_{i,t}}, \tag{18}$$

where $g_{i,t}$ is a deterministic function.

Lemma 1 *The value function of an i-investor has the form (18) with $A_i = ra_i$. The investor's first-order condition with respect to the position $y_{i,t}$ in asset i is*

$$\Phi_{i,t} = A_i \sigma_i^2 (y_{i,t} + u_i). \tag{19}$$

The first-order condition (19) equates asset i's expected return $\Phi_{i,t}$ to the marginal cost of bearing asset i's risk. This marginal cost is proportional to the investor's total exposure to asset i's risk, which is the sum of the position $y_{i,t}$ and the supply parameter u_i. The proportionality coefficient is the product of the risk-aversion coefficient A_i times the variance σ_i^2 of asset i's payoff. Eq. (19) can be viewed as the i-investor's demand function, determining the investor's position $y_{i,t}$ as a function of asset i's expected return $\Phi_{i,t}$.

If (19) holds for an i-investor, it also holds for an investor $-i$ in a symmetric equilibrium. Indeed, since the risk premia of assets i and $-i$ are opposites, the same is true for the assets' expected returns, i.e., $\Phi_{-i,t} = -\Phi_{i,t}$. Therefore, if (19) holds for an i-investor, it also holds for an investor $-i$ if $y_{-i,t} + u_{-i} = -(y_{i,t} + u_i)$. The latter condition is met because the supply parameters u_i and u_{-i} are opposites, and so are the positions $y_{i,t}$ and $y_{-i,t}$.

3.1.2 Arbitrageurs

Using (1), (12) and (15), and assuming that the risk premia of the assets in each pair are opposites, we can write the budget constraint (7) of an arbitrageur as

$$dW_t = \left[rW_t + \sum_{i \in \mathcal{A}} (x_{i,t} - x_{-i,t})\Phi_{i,t} - c_t \right] dt + \sum_{i \in \mathcal{A}} (x_{i,t} + x_{-i,t})\sigma_i dB_{i,t}. \tag{20}$$

The arbitrageur's expected return from investing in arbitrage opportunity i depends on the difference $x_{i,t} - x_{-i,t}$ between his positions in assets i and $-i$. This is because the risk premia of the two assets are opposites and hence the assets' expected returns are also opposites. The arbitrageur's risk from investing in the same arbitrage opportunity depends instead on the sum $x_{i,t} + x_{-i,t}$ of his positions in the two assets because this determines his total exposure to the Brownian motion $B_{i,t}$ that characterizes the assets' risk.

Eq. (20) implies, and Proposition 1 confirms, that the arbitrageur's optimal positions in assets i and $-i$ are opposites. Indeed, if the positions were not opposites, the arbitrageur could modify them by the same amount until their sum becomes zero. According to (20), the expected return from investing in arbitrage opportunity i would not change, but the risk would become zero. Setting $x_{-i,t} = -x_{i,t}$, we can simplify (20) to

$$dW_t = \left(rW_t + 2 \sum_{i \in \mathcal{A}} x_{i,t} \Phi_{i,t} - c_t \right) dt. \tag{21}$$

The first term in (21) is the arbitrageur's return from investing in the riskless asset, and the second term is the return from investing in the arbitrage opportunities. The latter return is riskless because the arbitrageur holds opposite positions in the assets in each pair. Hence, the arbitrageur can achieve a riskless return superior to that available to the other investors.

The financial constraint (8) limits the arbitrageur's ability to realize his excess return. We can simplify (8) by noting that the arbitrageur's spread trades involve long positions in the assets with positive supply parameters, i.e., assets $i \in \mathcal{A}$, because in equilibrium these assets offer positive expected returns. Using $x_{i,t} = -x_{-i,t} > 0$ for $i \in \mathcal{A}$, we can simplify (8) to

$$W_t \geq 2 \sum_{i \in \mathcal{A}} m_i x_{i,t}. \tag{22}$$

Proposition 1 *Suppose that the risk premia of the assets in each pair are opposites. An arbitrageur*

- *Consumes a constant fraction β of his wealth ($c_t = \beta W_t$).*
- *Holds opposite positions in assets $i \in \mathcal{A}$ and $-i$, with the long position being in asset i.*
- *Invests only in opportunities i yielding the maximum excess return per unit of collateral ($i \in \arg\max_{j \in \mathcal{A}} \frac{\Phi_{j,t}}{m_j}$) and is indifferent between any of them.*
- *Invests up to the financial constraint if the maximum excess return per unit of collateral is positive ($\max_{j \in \mathcal{A}} \frac{\Phi_{j,t}}{m_j} > 0$).*

The intuition for the last two results of the proposition is as follows. The arbitrageur chooses positions $\{x_{i,t}\}_{i\in\mathcal{I}}$ to maximize his excess return $2\sum_{i\in\mathcal{A}} x_{i,t}\Phi_{i,t}$ subject to the financial constraint (8). The solution to this problem is simple: the arbitrageur focuses on the opportunities yielding the highest excess return per unit of collateral, and is indifferent between any of them. The excess return per unit of collateral associated to opportunity i is $\frac{\Phi_{i,t}}{m_i}$: buying one share of asset i and shorting one share of asset $-i$ yields excess return $2\Phi_{i,t}$ but requires collateral $2m_i$. If the maximum excess return per unit of collateral is positive, then the arbitrageur's excess return is positive. Hence, the arbitrageur invests as much as possible, "maxing out" his financial constraint.

3.2 Equilibrium

Proposition 1 characterizes optimal investment policies given asset prices, but can be restated in terms of the prices implied by these policies.

Corollary 1 *There exists* $\Pi_t \geq 0$ *such that*

- *All opportunities* i *in which arbitrageurs invest* ($x_{i,t} > 0$) *offer the same excess return* Π_t *per unit of collateral, i.e.,*

$$\frac{\Phi_{i,t}}{m_i} = \Pi_t, \tag{23}$$

while the remaining opportunities offer return lower than Π_t.

- *Arbitrageurs invest only in opportunities* i *such that*

$$\frac{A_i\sigma_i^2 u_i}{m_i} > \Pi_t. \tag{24}$$

The equalization of excess return per unit of collateral across all opportunities in which arbitrageurs invest is a consequence of equilibrium: if returns differed, then arbitrageurs would focus on the opportunities with the highest returns (Proposition 1), which would be a contradiction. The common return Π_t of opportunities in which arbitrageurs invest can be viewed as a threshold: arbitrageurs invest in opportunity i if its excess return per unit of collateral in their absence exceeds Π_t. The corresponding inequality is (24): the left-hand side is the excess return per unit of collateral in the absence of arbitrageurs since (19) implies that when $y_{i,t} = 0$, asset i's expected return is $\Phi_{i,t} = A_i\sigma_i^2 u_i$. Eq. (24) implies that opportunities i in which arbitrageurs invest are associated with low margin requirements m_i and with high hedging motives by outside investors because of high risk aversion A_i, payoff variance σ_i^2, and supply parameter u_i.

An increase in the excess return Π_t per unit of collateral raises the right-hand side of (24) and can reduce the set of opportunities in which arbitrageurs invest. An increase in Π_t can be triggered by a reduction in arbitrageur wealth W_t. Indeed, when arbitrageurs are less wealthy, they are less able to exploit price discrepancies and to offer liquidity to outside investors. Therefore, price discrepancies are larger, and so is the return from arbitrage activity. To compute Π_t as function of W_t, we denote by N the number of elements of $\mathcal{A}$, by $\hat{\Pi}_n$ the n'th largest value of $\frac{A_i \sigma_i^2 u_i}{m_i}$ for $i \in \mathcal{A}$, by i_n the element of $\mathcal{A}$ corresponding to that value, and by $\mathcal{A}_n$ the set $\{i_m\}_{m=1,..,n}$, for $n \in \{1,..,N\}$. The value $\hat{\Pi}_n$ represents the excess return per unit of collateral above which arbitrageurs cease to invest in opportunity i_n. We also set

$$B_n \equiv \frac{1}{2\sum_{i \in \mathcal{A}_n} \frac{m_i^2 \mu_i}{A_i \sigma_i^2}},$$

$$C_n \equiv 2B_n \sum_{i \in \mathcal{A}_n} m_i \mu_i u_i + r - \beta,$$

$$\hat{W}_n \equiv 2 \sum_{i \in \mathcal{A}_n} m_i \mu_i u_i - \frac{\hat{\Pi}_{n+1}}{B_n},$$

for $n \in \{1,..,N\}$, and $B_{N+1} \equiv 0$, $C_{N+1} \equiv r - \beta$, $\hat{\Pi}_{N+1} \equiv 0$, $\hat{W}_0 \equiv 0$ and $\hat{W}_{N+1} = \infty$.

Proposition 2 *The arbitrageurs' excess return Π_t per unit of collateral is a decreasing, convex and piece-wise linear function of their wealth W_t.*

- *If $W_t \geq W_N$, then the financial constraint is slack and arbitrageurs earn zero excess return ($\Pi_t = 0$). They invest $x_{i,t} = \mu_i u_i$ in opportunity i, thus eliminating the price discrepancy between assets i and $-i$, and providing perfect liquidity to outside investors.*

- *If $W_t < W_N$, then the financial constraint is binding and arbitrageurs earn the positive excess return*

$$\Pi_t = B_n \left(2 \sum_{i \in \mathcal{A}_{n(W_t)}} m_i \mu_i u_i - W_t \right) \tag{25}$$

per unit of collateral, where $n(W) \in \{1,..,N\}$ is such that $\hat{W}_{n(W)-1} < W \leq \hat{W}_{n(W)}$. They invest $0 \leq x_{i,t} < \mu_i u_i$ in opportunity i, thus not eliminating the price discrepancy between assets i and $-i$, and not providing perfect liquidity to outside investors. Their investment is positive ($x_{i,t} > 0$) only in opportunities $i \in \mathcal{A}_n$.

If arbitrageurs are sufficiently wealthy, then they compete their excess return down to zero and eliminate price discrepancies between asset pairs. Eliminating the discrepancy between assets

i and $-i$ requires investing $x_{i,t} = \mu_i u_i$ in opportunity i and posting collateral $2m_i\mu_i u_i$. Therefore, arbitrageurs can eliminate all discrepancies if their wealth W_t exceeds the total required collateral $2\sum_{i\in\mathcal{A}} m_i\mu_i u_i = \hat{W}_N$. If instead $W_t < \hat{W}_N$, then arbitrageurs' excess return is positive and the financial constraint binds. A reduction in W_t in the constrained region raises Π_t.

The arbitrageurs' excess return Π_t per unit of collateral is not only decreasing in their wealth W_t but is also convex. Indeed, a decrease in W_t has no effect on $\Pi_t = 0$ in the unconstrained region $W_t \geq \hat{W}_N$, but raises Π_t in the constrained region $W_t < \hat{W}_N$. Moreover, the increase in Π_t in the constrained region occurs at an increasing rate. This is because as Π_t increases, arbitrageurs withdraw completely from the less profitable opportunities. Therefore, their collateral is spread across a small number of opportunities, and a reduction in their total collateral W_t causes a large reduction in collateral allocated to each opportunity. This results in a large reduction in their investment in each opportunity and a large increase in the opportunity's excess return.

The relationship between the arbitrageurs' excess return Π_t and their wealth W_t goes in both directions. At a given point in time, Π_t is fully determined by W_t according to Proposition 2. Conversely, Π_t determines the dynamics of W_t: if, for example, Π_t is large, arbitrageur wealth earns a high return and grows faster. Combining Corollary 1 with the arbitrageurs' budget constraint (21), we find that arbitrageur wealth earns the riskless return $r + \Pi_t$.

Lemma 2 *Arbitrageur wealth evolves according to*

$$dW_t = (r + \Pi_t - \beta)\, W_t dt. \tag{26}$$

Determining arbitrageur wealth W_t requires solving the differential equation (26), for Π_t determined by Proposition 2. To rule out trivial cases, we make the following assumption

Assumption 1 *Arbitrageurs' discount factor β satisfies*

$$r < \beta < r + \hat{\Pi}_1. \tag{27}$$

If $\beta < r$, then arbitrageurs save more than they consume even when they earn the riskless rate r on their savings. Hence, their wealth converges to infinity. If $r + \hat{\Pi}_1 < \beta$, then arbitrageurs consume more than they save even when they earn the riskless rate $r + \hat{\Pi}_1$, which is the highest possible return from arbitrage activity. Hence, their wealth converges to zero. Under Assumption 1 instead, arbitrageur wealth converges to a steady-state value $W^* > 0$. Critical for convergence, and for the steady state's uniqueness, is that the arbitrageurs' return Π_t is decreasing in their wealth W_t. For example, when W_t is high, Π_t is low, and hence W_t decreases (because $r + \Pi_t < \beta$). This raises Π_t, and eventually W_t stops decreasing and reaches its steady-state value. Conversely, when W_t is low, Π_t is high, and hence W_t increases. This lowers Π_t, and eventually W_t stops increasing and reaches its steady-state value. Thus, the dynamics of arbitrageur wealth are self-correcting.

To determine arbitrageur wealth, we define the function $F(W,u,n)$ by

$$F(W,u,n) \equiv \frac{We^{C_n u}}{\frac{B_n}{C_n}W\left[e^{C_n u}-1\right]+1}, \tag{28}$$

and the times u_n by

$$\hat{W}_{n-1} \equiv F(\hat{W}_n, u_n, n) \tag{29}$$

for $n = n(W^*)+1, .., N$,

$$\hat{W}_n \equiv F(\hat{W}_{n-1}, u_n, n) \tag{30}$$

for $n = 2, .., n(W^*)-1$, and $u_n \equiv \infty$ for $n = n(W^*)$. The function (28) describes the dynamics of arbitrageur wealth in the interval $(\hat{W}_{n-1}, \hat{W}_n]$, where they invest in the n most profitable opportunities $i \in \mathcal{A}_n$. Wealth is decreasing in that interval if $n \geq n(W^*)+1$, i.e., if the interval is above the steady-state value W^*, and is increasing if $n \leq n(W^*)-1$, i.e., if the interval is below W^*. The time u_n in (29) measures how long it takes for wealth to decrease from $\hat{W}_n$ to $\hat{W}_{n-1} > W^*$, while the time u_n in (30) measures how long it takes for wealth to increase from $\hat{W}_{n-1}$ to $\hat{W}_n < W^*$.

Proposition 3 *Arbitrageur wealth converges monotonically to its steady state value W^*.*

- *Starting from $W_t > W^*$, wealth decreases to W^* as follows*

$$W_{t+u} = F(W_t, u, n(W_t)) \qquad \text{for } u \in [0, u(W_t)) \text{ where } W_{t+u(W_t)} \equiv \hat{W}_{n(W_t)-1}, \tag{31}$$

$$W_{t+u(W_t)+\sum_{m=n+1}^{n(W_t)-1} u_m + u} = F(\hat{W}_n, u, n) \qquad \text{for } u \in [0, u_n), \tag{32}$$

where $n = n(W^*), .., n(W_t)-1$.

- *Starting from $W_t < W^*$, wealth increases to W^* as follows*

$$W_{t+u} = F(W_t, u, n(W_t)) \qquad \text{for } u \in [0, u(W_t)) \text{ where } W_{t+u(W_t)} \equiv \hat{W}_{n(W_t)}, \tag{33}$$

$$W_{t+u(W_t)+\sum_{m=n(W_t)+1}^{n-1} u_m + u} = F(\hat{W}_{n-1}, u, n) \qquad \text{for } u \in [0, u_n), \tag{34}$$

where $n = n(W_t)+1, .., n(W^*)$.

Lemma 3 *Arbitrageurs' investment in opportunity i is*

$$x_{i,t} = \max\left\{\mu_i\left(u_i - \frac{m_i\Pi_t}{a_i\sigma_i^2}\right), 0\right\}. \tag{35}$$

In steady state, arbitrageurs earn excess return excess $\Pi = \beta - r$ per unit of collateral, as can be seen by setting $dW_t = 0$ in (26). This excess return is larger if arbitrageurs are more impatient (large β). Indeed, since arbitrageurs consume a larger fraction of wealth, they have less wealth to use as collateral. Therefore price discrepancies are larger, and so is the return from arbitrage activity.

We now turn to risk premia which, from (16) equal the present value of future instantaneous expected excess returns:

$$\phi_{i,t} = \int_t^\infty \Phi_{i,s}e^{-r(s-t)}ds. \tag{36}$$

From $\Phi_{i,s} = m_i\Pi_s = m_iB\max\{0, W_c - W_s\}$, we can derive the risk premia dynamics from that of arbitrageur wealth:

$$\phi_{i,t} = m_iB\int_t^\infty \max\{0, W_c - W_s\}e^{-r(s-t)}ds. \tag{37}$$

Proposition 4 *The risk premium of asset $i \in \mathcal{A}$ at time t is*

- *If $W_t < W_c$,*

$$\phi_{i,t} = m_iB\int_0^\infty \left[W_c - \frac{W_te^{As}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}\right]e^{-rs}ds. \tag{38}$$

- *If $W_t \geq W_c$,*

$$\phi_{i,t} = m_iB\left(\frac{W_c}{W_t}\right)^{\frac{r}{\beta-r}}\int_0^\infty \left[W_c - \frac{W_ce^{As}}{\frac{B}{A}W_c\left(e^{As}-1\right)+1}\right]e^{-rs}ds. \tag{39}$$

- *As t increases, the risk premium $\phi_{i,t}$ converges monotonically towards its steady state value*

$$\phi_i = m_i\left(\frac{\beta - r}{r}\right).$$

Finally, we can derive the arbitrageurs' equilibrium positions (Lemma 3).

Proposition 5 *The arbitrageurs' position in asset $i \in \mathcal{A}$ at time t is as follows:*

$$x_{i,t} = \mu_i \left(u_i - \frac{m_i}{a_i \sigma_i^2} B \max \{W_c - W_t; 0\} \right). \tag{40}$$

3.3 Properties

Having derived all equilibrium variables in closed form, we can draw many implications. Some of these are cross-sectional in nature, i.e., comparing variables across opportunities with different characteristics. Others involve comparative statics with respect to arbitrageur wealth. These can be considered in two ways. First, because arbitrageur wealth varies over time out of steady state, the comparative statics results can be translated into time series predictions while the equilibrium is off the steady state. Alternatively, they can be interpreted as the effect of an unanticipated exogenous shock to arbitrageur wealth. These are also useful for the analysis of risky arbitrage.

Note that due to the model's symmetry, optimal risk-sharing, which would result from unconstrained trading, would imply $\phi_{i,t} = 0$.

Definition 3 *The risk premium $\phi_{i,t}$ is a measure of the illiquidity (i,t)- and $(-i,t)$-investors face.*

Corollary 2 *The risk premia are decreasing and convex in arbitrageur wealth, i.e., for all $i \in \mathcal{A}$*

$$\frac{\partial \phi_{i,t}}{\partial W_t} < 0 \qquad \textit{and} \qquad \frac{\partial^2 \phi_{i,t}}{\partial W_t^2} > 0.$$

Consider a drop in arbitrageur wealth. Intuitively, the risk premia should increase because arbitrageurs being poorer, they reduce their liquidity provision and allow prices to diverge. Moreover, when arbitrageur wealth is smaller, the return on arbitrageurs' wealth is larger and therefore a drop in arbitrageur wealth has a larger impact on future arbitrage wealth and thus on risk premia.

Corollary 3 *An asset's risk premium is increasing in its supply, and more so the lower arbitrageur wealth is, i.e., for all $i \in \mathcal{A}$*

$$\frac{\partial \phi_{i,t}}{\partial u_i} > 0 \qquad \textit{and} \qquad \frac{\partial^2 \phi_{i,t}}{\partial u_i \partial W_t} < 0.$$

Intuitively, $\phi_{i,t}$ increases with u_i since the discrepancy between the valuations of (i,t)- and $(-i,t)$-investors is larger. There is a mitigating effect. Indeed, the higher u_i, the higher the arbitrageurs' return and their future wealth. This tends to reduce future excess returns, and therefore the current risk premium. For low levels of W_t the mitigating effect is small, and therefore u_i has a large effect on $\phi_{i,t}$.

Corollary 4 *Illiquidity is higher for opportunities with higher margin requirements. These opportunities offer higher instantaneous excess returns and have higher risk premia, i.e., for all* $(i,j) \in \mathcal{A}^2$

$$m_i > m_j \qquad \Rightarrow \qquad \Phi_{i,t} > \Phi_{j,t} \qquad \textit{and} \qquad \phi_{i,t} > \phi_{j,t}. \tag{41}$$

Intuitively, investment opportunities requiring arbitrageurs to tie up more capital as collateral must provide them with a greater reward, i.e., a higher excess return. Risk premia being the present value of future excess returns, they must be higher for such opportunities.

Corollary 5 *Illiquidity and risk premia are more sensitive to arbitrageur wealth for opportunities with higher margin requirements, i.e., for all* $(i,j) \in \mathcal{A}^2$

$$m_i > m_j \qquad \Rightarrow \qquad \frac{\partial \phi_{i,t}}{\partial W_t} < \frac{\partial \phi_{j,t}}{\partial W_t} < 0. \tag{42}$$

Intuitively, changes in arbitrageur wealth affect the excess return (current or future) per unit of collateral, and therefore impact more strongly opportunities with higher collateral requirements.

Corollary 6 *Illiquidity and risk premia are more sensitive to the supply of other assets for opportunities with higher margin requirements, i.e., for all* $(i,j,k) \in \mathcal{A}^2$

$$m_i > m_j \qquad \Rightarrow \qquad \frac{\partial \phi_{i,t}}{\partial u_k} > \frac{\partial \phi_{j,t}}{\partial u_k} > 0$$

Intuitively, changes in supply affect the excess return (current or future) per unit of collateral, and therefore impact more strongly opportunities with higher collateral requirements.

Corollary 7 *Suppose* $W_t < W_c$*. Changes in arbitrageur wealth impact more strongly their position in opportunity* $(i,-i)$ *than in* $(j,-j)$ *if* $\frac{a_i \sigma_i^2}{m_i \mu_i} < \frac{a_j \sigma_j^2}{m_j \mu_j}$.

When arbitrageurs are unconstrained ($W_t > W_c$), they invest $x_{i,t} = \mu_i u_i$ in opportunity $(i,-i)$, independently of their wealth. Instead, when they are constrained ($W_t < W_c$), their wealth affect their positions. For example, following a drop in wealth, arbitrageurs are more constrained and reduce their investment in all opportunities. Investment is more wealth-sensitive for opportunities with higher collateral requirements because the excess returns that arbitrageurs require to invest in those opportunities are more affected by wealth changes (Corollary 4). Investment is less wealth-sensitive for opportunities where outside investors are more risk-averse or assets are riskier because outside investors for those opportunities have a more inelastic demand for insurance.

4 Risky Arbitrage

We now consider the possibility of arbitrage risk which in our model, stems from two sources: fundamental risk and supply risk. Fundamental risk means that assets i and $-i$ in a pair need not pay identical dividends, i.e., $\sigma_i^f \neq 0$. We assume $\sigma_i^f > 0$ for $i \in \mathcal{A}$. Supply risk means that asset i's supply $u_{i,t}$ is stochastic, i.e., $\sigma_i^u \neq 0$. We assume $\sigma_i^u > 0$ for $i \in \mathcal{A}$.[7]

We derive equilibrium conditions in Section 4.1, derive general properties of the equilibrium in Section 4.2, and characterize the equilibrium more fully for small arbitrage risk in Section 4.3.

4.1 Optimal Investment Policies

As we will see, an asset's risk premium is affected by the fundamental shocks and the supply shocks to *all* assets. Hence, for $i \in \mathcal{A}$, we denote the dynamics of the risk premium $\phi_{i,t}$ by

$$d\phi_{i,t} \equiv \nu_{i,t}^{\phi} dt + \sum_{j \in \mathcal{A}} \sigma_{i,j,t}^{f\phi} dB_{j,t}^f + \sum_{j \in \mathcal{A}} \sigma_{i,j,t}^{u\phi} dB_{j,t}^u. \tag{43}$$

Similarly, the instantaneous return of asset i is affected by *all* shocks to *all* assets because they affect the asset's risk premium. From Eqs. (13), (16) and (43), we have

$$dR_{i,t} = \Phi_{i,t} dt + \sigma_i dB_{i,t} + \sigma_i^f dB_{i,t}^f - \sum_{j \in \mathcal{A}} \sigma_{i,j,t}^{f\phi} dB_{j,t}^f - \sum_{j \in \mathcal{A}} \sigma_{i,j,t}^{u\phi} dB_{j,t}^u. \tag{44}$$

[7] Assuming $\sigma_i^f > 0$ and $\sigma_i^u > 0$ is without loss of generality as we can replace $B_{i,t}^f$ and $B_{i,t}^u$ with their opposites.

4.1.1 Outside Investors

We first characterize the optimal investment policies of outside investors. Using (1), (10), (16) and (43), we can write the (i,t)-investors' dynamic budget constraint (5) as

$$dw_{i,t} = \left[rw_{i,t} + u_{i,t}\left(D_i - \nu^{\phi}_{i,t}\right) + y_{i,t}\Phi_{i,t}\right] dt$$

$$+ (y_{i,t} + u_{i,t}) \left[\sigma_i dB_{i,t} + \sigma^f_i dB^f_{i,t} - \sum_{j\in\mathcal{A}} \sigma^{f\phi}_{i,j,t} dB^f_{j,t} - \sum_{j\in\mathcal{A}} \sigma^{u\phi}_{i,j,t} dB^u_{j,t}\right]. \quad (45)$$

The drift is the same as for riskfree arbitrage. The diffusion term captures the risky part of the return of asset i (Eq. (44)). Given this, the (i,t)-investors' objective is

$$\max_{y_{i,t}} \left[y_{i,t}\Phi_{i,t} - \frac{a_i}{2}(y_{i,t} + u_{i,t})^2 \left(\sigma^R_{i,t}\right)^2\right]. \quad (46)$$

The first term is the expected excess return (i,t)-investors derive from their holding in asset i. The second term is a cost of bearing risk. It depends on asset i's instantaneous volatility computed as

$$\left(\sigma^R_{i,t}\right)^2 \equiv \frac{Var_t(dR_{i,t})}{dt} = \sigma_i^2 + \left(\sigma^f_i - \sigma^{f\phi}_{i,i,t}\right)^2 + \sum_{j\in\mathcal{A}/\{i\}} \left(\sigma^{f\phi}_{i,j,t}\right)^2 + \sum_{j\in\mathcal{A}} \left(\sigma^{u\phi}_{i,j,t}\right)^2. \quad (47)$$

At the optimum, the expected excess return of asset i to the marginal cost of risk-bearing, i.e.,

$$\Phi_{i,t} = a_i \left(\sigma^R_{i,t}\right)^2 (y_{i,t} + u_{i,t}). \quad (48)$$

Their first order condition determines the (i,t)-investors' demand $y_{i,t}$. By symmetry, problem $\mathcal{P}_{-i,t}$ yields the same first-order condition as $\mathcal{P}_{i,t}$.

4.1.2 Arbitrageurs

We characterize an arbitrageur's optimal consumption and investment policy under the restriction that the prices of assets in the same pair are driven by symmetric processes, i.e., $\phi_{-i,t} = -\phi_{i,t}$. Using (1), (10), (16), (43) and symmetry, the arbitrageurs' dynamic budget constraint (7) is

$$dW_t = \left(rW_t + 2\sum_{i\in\mathcal{A}} x_{i,t}\Phi_{i,t} - c_t\right) dt + 2\sum_{i\in\mathcal{A}} x_{i,t} \left(\sigma^f_i dB^f_{i,t} - \sum_{j\in\mathcal{A}} \sigma^{f\phi}_{i,j,t} dB^f_{j,t} - \sum_{j\in\mathcal{A}} \sigma^{u\phi}_{i,j,t} dB^u_{j,t}\right). \quad (49)$$

The drift is the same as for riskless arbitrage, i.e., arbitrageur wealth increases by the risk free return plus the expected excess returns provided by all opportunities net of the arbitrageurs' consumption.

Now however, there are also diffusion terms because arbitrage is risky. Denote the respective diffusion coefficients for the fundamental shock $dB^f_{j,t}$ and the supply shock $dB^u_{j,t}$ as

$$\sigma^{fW}_{j,t} \equiv 2x_{j,t}\sigma^f_j - 2\sum_{i\in\mathcal{A}} x_{i,t}\sigma^{f\phi}_{i,j,t} \qquad \text{and} \qquad \sigma^{uW}_{j,t} \equiv -2\sum_{i\in\mathcal{A}} x_{i,t}\sigma^{u\phi}_{i,j,t}. \tag{50}$$

A fundamental shock to opportunity $(j,-j)$ means that assets j and $-j$ do not pay the exact same dividend. The "net dividend" affects arbitrageurs' profit and hence their wealth. This direct effect is captured by $2x_{j,t}\sigma^f_j$. At the same time, the shock affects all opportunities' risk premia, and hence the arbitrageurs' capital gains from their investments in these opportunities. This effect is captured by $-2\sum_{i\in\mathcal{A}} x_{i,t}\sigma^{f\phi}_{i,j,t}$. A supply shock to opportunity $(j,-j)$ means that (j,t)- and $(-j,t)$-investors are more eager to trade. Such a shock does not affect arbitrageur wealth directly but indirectly through its effect on the risk premia of all opportunities, which in turn affects arbitrageurs' capital gains and ultimately their wealth. This effect is captured by $-2\sum_{i\in\mathcal{A}} x_{i,t}\sigma^{u\phi}_{i,j,t}$.

For $i\in\mathcal{A}$, denote $2\hat{\Phi}_{i,t}$ the arbitrageurs' risk-adjusted return from opportunity $(i,-i)$. Indeed, their expected excess return from opportunity $(i,-i)$, $2\Phi_{i,t}$, must be adjusted for the fundamental and supply risk the opportunity entails. This is done by multiplying the arbitrageurs' coefficient of absolute risk aversion, equal to $1/W_t$ due to logarithmic utility, with the covariance of the return of opportunity $(i,-i)$ and that of the arbitrageurs' portfolio, i.e.,

$$2\hat{\Phi}_{i,t} \equiv 2\Phi_{i,t} - \frac{Cov_t(dR_{i,t} - dR_{-i,t}, dW_t)}{W_t dt}. \tag{51}$$

The covariance is obtained by summing over all fundamental and supply shocks the loading of $(i,-i)$'s return on each shock times the arbitrageurs' portfolio loading on that shock, i.e.,[8]

$$\hat{\Phi}_{i,t} = \Phi_{i,t} - \frac{1}{W_t}\left[(\sigma^f_i - \sigma^{f\phi}_{i,i,t})\sigma^{fW}_{i,t} - \sum_{j\in\mathcal{A}/\{i\}} \sigma^{f\phi}_{i,j,t}\sigma^{fW}_{j,t} - \sum_{j\in\mathcal{A}} \sigma^{u\phi}_{i,j,t}\sigma^{uW}_{j,t}\right].$$

We can now derive the arbitrageurs' optimal policy.

Proposition 6 *Denote* $\Pi_t \equiv \max_{i\in\mathcal{A}}\left|\hat{\Phi}_{i,t}/m_i\right|$. *Each arbitrageurs consumes a fraction* β *of his wealth, i.e.,* $c_t = \beta W_t$, *and his investment policy satisfies one of the following conditions.*

- *The financial constraint (8) is slack and* $\hat{\Phi}_{i,t} = 0$ *for all* i.

[8] The arbitrageurs' logarithmic utility simplifies the analysis because risk is measured by the covariance with the arbitrageurs' portfolio and not with other state variables.

- *The financial constraint (8) is binding and for all $i \in \mathcal{A}$,*

$$x_{i,t} > 0 \Rightarrow \frac{\hat{\Phi}_{i,t}}{m_i} = \Pi_t \qquad \text{and} \qquad x_{i,t} < 0 \Rightarrow \frac{\hat{\Phi}_{i,t}}{m_i} = -\Pi_t. \tag{52}$$

Proposition 6 is Proposition 1's counterpart for risky arbitrage. When the financial constraint is slack, arbitrageurs close all opportunities. When the financial constraint is binding, arbitrageurs invest only in opportunities yielding the maximum return on collateral. There are however two differences with Proposition 1. First, the relevant return from opportunity $(i, -i)$ is the risk-adjusted return $\hat{\Phi}_{i,t}$, which depends both on prices and arbitrageur positions. Second, arbitrageurs can "short" some opportunities, i.e., long the pricier asset and short the cheaper one ($x_{i,t} < 0$ for $i \in \mathcal{A}$). This can be optimal for arbitrageurs for hedging their long positions in other opportunities.

4.2 Amplification and Contagion: Direct and Indirect Effects

Equilibrium prices and positions solve the first-order condition of outside investors (Eq. (48)) and arbitrageurs (Proposition 6). This system of equations is complex. Here we derive general properties of equilibrium.

Assumption 2 *Define $W_{c,t} \equiv 2\sum_{i\in\mathcal{A}} m_i \mu_i u_{i,t}$. We assume*

$$0 < \beta - r < BW_{c,t} \qquad \text{and} \qquad \min_{i\in\mathcal{A}} \frac{a_i \sigma_i^2 u_{i,t}}{m_i} > \beta - r.$$

In equilibrium, the risk premium $\phi_{i,t}$ is a function of arbitrageur wealth W_t and the supply parameters $\{u_{j,t}\}_{j\in\mathcal{A}}$. Eq. (49) and Ito's Lemma imply

$$\sigma^{f\phi}_{i,j,t} = \frac{\partial \phi_{i,t}}{\partial W_t} \sigma^{fW}_{j,t} \qquad \text{and} \qquad \sigma^{u\phi}_{i,j,t} = \frac{\partial \phi_{i,t}}{\partial W_t} \sigma^{uW}_{j,t} + \frac{\partial \phi_{i,t}}{\partial u_{j,t}} \sigma^{u}_{j}. \tag{53}$$

As for riskfree arbitrage, arbitrageur wealth creates a linkage between the different opportunities even though their fundamentals are independent.

Lemma 4 *Fundamental and supply shocks to one opportunity affect arbitrageur wealth and the risk premia of all opportunities. More precisely, for $(i, j) \in \mathcal{A}^2$, the effect of a fundamental shock $dB^f_{j,t}$*

to opportunity $(j,-j)$ on arbitrageur wealth and on asset i's the risk premium $\phi_{i,t}$ are respectively

$$\sigma^{fW}_{j,t} = \frac{2x_{j,t}\sigma^f_j}{1+2\sum_{k\in\mathcal{A}} x_{k,t}\frac{\partial\phi_{k,t}}{\partial W_t}}, \tag{54}$$

$$\sigma^{f\phi}_{i,j,t} = \frac{\partial\phi_{i,t}}{\partial W_t}\frac{2x_{j,t}\sigma^f_j}{1+2\sum_{k\in\mathcal{A}} x_{k,t}\frac{\partial\phi_{k,t}}{\partial W_t}}. \tag{55}$$

The effect of the supply shock $dB^u_{j,t}$ on the same variables are respectively

$$\sigma^{uW}_{j,t} = -\frac{2\sum_{k\in\mathcal{A}} x_{k,t}\frac{\partial\phi_{k,t}}{\partial u_{j,t}}\sigma^u_j}{1+2\sum_{k\in\mathcal{A}} x_{k,t}\frac{\partial\phi_{k,t}}{\partial W_t}}, \tag{56}$$

$$\sigma^{u\phi}_{i,j,t} = \frac{\partial\phi_{i,t}}{\partial u_{j,t}}\sigma^u_j - \frac{\partial\phi_{i,t}}{\partial W_t}\frac{2\sum_{k\in\mathcal{A}} x_{k,t}\frac{\partial\phi_{k,t}}{\partial u_{j,t}}\sigma^u_j}{1+2\sum_{k\in\mathcal{A}} x_{k,t}\frac{\partial\phi_{k,t}}{\partial W_t}}. \tag{57}$$

To develop an intuition, assume that (as for riskless arbitrage) arbitrageurs long all opportunities ($x_{i,t} > 0$ for all $i \in \mathcal{A}$), and risk premia decrease with arbitrageur wealth.

After a positive fundamental shock $dB^f_{j,t}$ to opportunity $(j,-j)$, asset j's dividend exceeds asset $-j$'s, and arbitrageurs receive the "net dividend" $2x_{j,t}\sigma^f_j dB^f_{j,t}$. This direct effect on wealth corresponds to the numerator in (54). Moreover, arbitrageurs being richer, risk premia decrease and the arbitrageurs realize capital gains $2\sum_{k\in\mathcal{A}} x_{k,t}\frac{\partial\phi_{k,t}}{\partial W_t}$. This indirect effect on wealth corresponds to the denominator in (54). Since $x_{k,t} > 0$ and $\frac{\partial\phi_{k,t}}{\partial W_t} < 0$, this indirect effect amplifies the direct effect. The indirect effect on the risk premium $\phi_{i,t}$ is (55).

A positive supply shock $dB^u_{j,t}$ to opportunity $(j,-j)$ means that (j,t)- and $(-j,t)$-investors are more eager to trade. Holding wealth constant, such a shock has the direct effect of increasing asset i's risk premium by $\frac{\partial\phi_{i,t}}{\partial u_{j,t}}\sigma^u_j dB^u_{j,t}$, the first term in (57). Due to the increase in risk premia, arbitrageurs realize a capital loss $2\sum_{k\in\mathcal{A}} x_{k,t}\frac{\partial\phi_{k,t}}{\partial u_{j,t}}\sigma^u_j$. Moreover, arbitrageurs being poorer, risk premia increase and the arbitrageurs' loss is amplified. The indirect effect on wealth is (56) and on the risk premium $\phi_{i,t}$ is the second term in Eq. (57).

4.3 Small Arbitrage Risk

In this section, we characterize the solution more fully when arbitrage risk is small ($\sigma^f_i \simeq 0$ and $\sigma^u_i \simeq 0$), and supply parameters are slowly mean-reverting ($\kappa^u_i \simeq 0$). Specifically, we study how an asset's liquidity, volatility and correlation with other assets depend on arbitrageur wealth. We also study arbitrageurs' positions.

4.3.1 Liquidity

Eq. (44) implies that the impact of a supply shock $dB^u_{i,t}$ to asset i at time t on the asset return $dR_{i,t}$ is $|\sigma^{u\phi}_{i,i,t}|$. Hence we define asset i's liquidity as

$$\lambda_{i,t} \equiv \frac{1}{|\sigma^{u\phi}_{i,i,t}|}. \tag{58}$$

All markets are less liquid than absent constraints. Indeed, arbitrageurs cannot absorb as much of the supply shocks as they otherwise would. Since the extent to which financial constraints bind depends on arbitrageur wealth, it is clear that liquidity should depend on arbitrageur wealth. We show that while it does indeed, more arbitrageur wealth does not always yield more liquid markets.

Proposition 7 *There exists $\epsilon > 0$ going to zero when $\{\sigma^f_j, \sigma^u_j, \kappa^u_j\}_{j\in\mathcal{A}}$ go to zero such that*

- *If $W_t > W_{c,t} + \epsilon$, asset i's liquidity $\lambda_{i,t}$ increases with arbitrageur wealth W_t.*
- *If $\underline{W} \leq W_t < W_{c,t} - \epsilon$, asset i's liquidity $\lambda_{i,t}$ decreases with arbitrageur wealth W_t.*

The intuition is as follows. Supply shocks affect risk premia directly but also indirectly through arbitrageur wealth (Lemma 4). The direct effect is weaker when arbitrageur wealth is high (Corollary 3).[9] The indirect effect, however, is a hump-shaped function of wealth. Indeed, at low levels of wealth, the financial constraint is binding and an increase in wealth triggers a sharp increase in arbitrageurs' positions. When positions are larger, arbitrageur wealth is more sensitive to changes in risk premia, and therefore the indirect effect is stronger. Instead, at high values of wealth, arbitrageurs' positions are less sensitive to wealth. The main effect of an increase in wealth is to render risk premia less sensitive to wealth (Corollary 2), implying a weaker indirect effect. The hump-shaped indirect effect drives the U-shaped pattern of liquidity.

4.3.2 Volatility

We next examine how arbitrageur wealth affects the volatility of assets. The volatility of asset i is given by Eq. (47). All assets are more volatile than absent financial constraints. And again, it

[9] Corollary 3 is for the case $\sigma^f_j = \sigma^u_j = \kappa^u_j = 0$, but by continuity, the result extends to small values of $(\sigma^f_j, \sigma^u_j, \kappa^u_j)$.

is intuitive that volatility should depend on arbitrageur wealth. Indeed asset i's volatility depends on factors affecting the asset's dividend, σ_i and σ_i^f, but also on factors affecting the supply of and demand for the asset, $\sigma_{i,j,t}^{f\phi}$ and $\sigma_{i,j,t}^{u\phi}$. Unlike the former, the latter do depend on arbitrageur wealth, so that asset volatilities do too. We show however that they do so in a non-trivial and generally non-monotonic way.

Proposition 8 *There exists $\epsilon > 0$ going to zero when $\{\sigma_j^f, \sigma_j^u, \kappa_j^u\}_{j\in\mathcal{A}}$ go to zero such that:*

- *If $W_t > W_{c,t} + \epsilon$, asset i's volatility $\sigma_{i,t}^R$ decreases with arbitrageur wealth W_t.*
- *If $\underline{W} \le W_t < W_{c,t} - \epsilon$, the component of asset i's volatility $\sigma_{i,t}^R$ due to supply shocks $dB_{j,t}^u$ increases in arbitrageur wealth W_t, and that due to the fundamental shock $dB_{j,t}^f$, $j \in \mathcal{A}$, increases if*

$$\frac{a_j \sigma_j^2 u_{j,t}}{m_j} \le \frac{\sum_{k\in\mathcal{A}} m_k \mu_k u_{k,t}}{\sum_{k\in\mathcal{A}} \frac{m_k^2 \mu_k}{a_k \sigma_k^2}}. \tag{59}$$

The intuition is as follows. A supply shock to asset $j \neq i$ has no direct effect on asset i's risk premium. However, as for a supply shock to asset i itself, its indirect effect is a hump-shaped function of wealth. Therefore, absent fundamental shocks, volatility would be hump-shaped.

The fundamental shock $dB_{j,t}^f$ also generates hump-shaped volatility if j satisfies condition (59). This condition is satisfied by a non-empty subset of $\mathcal{A}$, and by all assets in $\mathcal{A}$ if they are homogenous (and in particular if there is only one opportunity). It is not satisfied when $a_j \sigma_j^2 u_{j,t}/m_j$ is large relative to a weighted average of this variable across assets, and in that case the volatility due to $dB_{j,t}^f$ decreases with arbitrageur wealth. The intuition is that an increase in wealth leads to an increase in arbitrageur positions (implying larger volatility), but to a reduction in the wealth-sensitivity of risk premia (implying smaller volatility). When $u_{j,t}$ is large, arbitrageurs are invested heavily in opportunity $(j, -j)$, and the second effect dominates because the shock $dB_{j,t}$ has a large impact on wealth.[10]

[10]Condition (59) is not needed for supply shocks because the direct effect of $dB_{j,t}^f$ is through arbitrageurs' position in opportunity $(j, -j)$, while that of $dB_{j,t}^u$ concerns all opportunities.

4.3.3 Correlations

We now turn to asset correlations which again differ from the unconstrained case. First, some assets have uncorrelated fundamentals, i.e., dividends and supply. In our model, these are assets not in the same pair. Absent constraints or segmentation, these assets' returns would be uncorrelated. With constraints however they are correlated because arbitrageur wealth is a common factor affecting all asset returns. Second, assets in the same pair have correlated fundamentals but their returns' correlation is below that absent constraints. We show that correlations depend on arbitrageur wealth in a non-trivial way.

Proposition 9 *Consider $(i, i') \in \mathcal{A}^2$, $i \neq i'$. There exists $\epsilon > 0$ going to zero when $\{\sigma_j^f, \sigma_j^u, \kappa_j^u\}_{j \in \mathcal{A}}$ go to zero such that*

- *If $W_t > W_{c,t} + \epsilon$, the correlation between assets i and i' decreases with arbitrageur wealth W_t.*
- *If $W \leq W_t < W_{c,t} - \epsilon$, the component of the correlation between assets i and i' due to supply shocks $dB_{j,t}^f$ increases with arbitrageur wealth W_t, and that due to the fundamental shock $dB_{j,t}^f$, $j \in \mathcal{A}$, increases if (59) is satisfied.*
- *The opposite holds for the correlation between assets i and $-i$, and for that between assets i and $-i'$.*

The intuition is as follows. Assume that (as for riskless arbitrage) arbitrageurs long all opportunities ($x_{i,t} > 0$ for all $i \in \mathcal{A}$), and risk premia decrease with arbitrageur wealth.

Consider first two assets $i \neq i'$ that arbitrageurs long. For such assets, correlation is positive despite their fundamentals' independence. For high levels of wealth, as wealth increases, their correlation converges to that absent constraint, i.e., zero. Things are different for low levels of wealth. Indeed, a given increase in arbitrageur wealth translates into a larger increase in arbitrageurs' positions, and hence in their exposure to supply shocks. Since arbitrageur wealth is a factor common to all assets, this increases the correlation between i and i'. Hence, the correlation between i and i' tends to be hump-shaped or decreasing in wealth.

Consider now assets i and $-i'$. For such assets, correlation is negative despite their fundamentals' independence. Because fundamental and supply shocks have opposite effects on assets i' and $-i'$, the correlation between assets i and $-i'$ tends to be inverse U-shaped or increasing.

Finally, consider assets i and $-i$. These assets tend to be less correlated than absent constraints. Because fundamental and supply shocks have opposite effects on assets i and $-i$, the correlation between assets i and $-i$ tends to be inverse hump-shaped or increasing.

One interesting aspect of these results is that the effect of a change in arbitrageur wealth on correlations is not uniform across asset pairs or across wealth levels. For instance, a reduction in arbitrage capital (e.g., as during a financial crisis) does not necessary lead to an increase in correlations across all assets, a phenomenon often viewed as contagion, and this for two distinct reasons. First, arbitrageurs' activity tends to bring the prices of assets with correlated fundamentals (e.g., i and $-i$) in line with each other. When they are poorer, they may be able to perform that role, and the correlation between such assets decreases. Second, for low levels of arbitrage wealth, arbitrageurs hold small positions and this weakens the transmission of shocks through arbitrageur wealth, reducing the correlation between assets with uncorrelated fundamentals.

4.3.4 Arbitrage Positions

We next examine how arbitrageur positions depend on their wealth and on the risk of investment opportunities. When arbitrageurs long all opportunities (i.e., $x_{i,t} > 0$ for $i \in \mathcal{A}$), Eq. (51) and Proposition 6 imply that for all $i \in \mathcal{A}$, the expected excess return from opportunity $(i, -i)$ is

$$\Phi_{i,t} = m_i \Pi_t + \frac{Cov_t(dR_{i,t} - dR_{-i,t}, dW_t)}{2W_t dt}. \tag{60}$$

The first term is a compensation for tying up capital as collateral. The risk-adjusted return on collateral Π_t is positive when the financial constraint binds and zero when it is slack. The second term is a compensation for risk. It is positive because both fundamental and supply shocks induce positive correlation between the return on opportunity $(i, -i)$ and arbitrageur wealth. Indeed, a positive fundamental shock $dB^f_{j,t}$ to $j \in \mathcal{A}$ raises arbitrageur wealth, leading to lower risk premia and higher returns from all opportunities. A positive supply shock $dB^u_{j,t}$ to $j \in \mathcal{A}$ raises premia, leading to lower arbitrageur wealth and lower returns from all opportunities.

Lemma 5 *The financial constraint becomes slack at a lower level of wealth than under riskless arbitrage. More precisely, if $\{\sigma^f_j, \sigma^u_j, \kappa^u_j\}_{j \in \mathcal{A}}$ are small, financial constraint (8) holds as an equality if and only if $W_t \leq W_{c,t} - \epsilon$, where $\epsilon > 0$.*

The intuition is as follows. Contrary to the riskfree arbitrage case, aggregate risk is not zero. Hence optimal risk sharing does not involve full insurance for outside investors. Said differently, because arbitrageurs require positive compensation for risk from each opportunity, they do not drive expected excess returns down to zero even when they have enough wealth to do so.

Proposition 10 *For $\{\sigma_j^f, \sigma_j^u, \kappa_j^u\}_{j\in\mathcal{A}}$ small, consider $(i,i') \in \mathcal{A}^2$ such that $(\sigma_i, a_i, \mu_i, u_{i,t}) = (\sigma_{i'}, a_{i'}, \mu_{i'}, u_{i',t})$.*

- *If $m_i > m_{i'}$ and $\sigma_i^f = \sigma_{i'}^f$, changes in arbitrageur wealth impact more strongly their position in opportunity $(i, -i)$ than $(i', -i')$.*
- *If $m_i = m_{i'}$ and $\sigma_i^f > \sigma_{i'}^f$, changes in arbitrageur wealth impact more strongly their position in opportunity $(i, -i)$ than $(i', -i')$ when $W_t > W_{c,t} + \epsilon$ for $\epsilon > 0$ that converges to zero when $\{\sigma_j^f, \sigma_j^u, \kappa_j^u\}_{j\in\mathcal{A}}$ go to zero.*

When arbitrageurs are unconstrained, their positions are limited only by risk aversion. If arbitrageur wealth decreases within that region, risk-aversion increases (the coefficient of absolute risk aversion is $1/W_t$) and returns become more volatile (Proposition 8). These mutually-reinforcing effects induce arbitrageurs to scale down their positions, especially in opportunities that involve more risk. These are the opportunities with high collateral requirements (high m_i) and high fundamental risk (high σ_i^f). Note that opportunities with high collateral requirements are more affected not because the opportunity cost of collateral increases, but because their returns are more volatile.

Consider next the region where the financial constraint binds. Under riskless arbitrage, arbitrageurs scale down more their positions in opportunities with high collateral requirements (Corollary 7). Under risky arbitrage, arbitrageurs are also concerned about the risk of each opportunity, but the variation of this effect with wealth is ambiguous. On the one hand, when wealth decreases, arbitrageurs become more risk-averse. On the other hand, return volatility can decrease (Proposition 8). As a consequence, arbitrageurs can scale down their positions less in riskier opportunities. For small arbitrage risk, the effect of m_i is unambiguous (same as under riskless arbitrage), but the effect of σ_i^f is ambiguous.

5 Conclusion

This paper develops a framework to examine the relationship between intermediary capital, financial market liquidity and asset prices. Its main features are as follows. First, arbitrageurs are sophisticated investors with better investment opportunities than other investors, but they face financial constraints. Second, ours is a dynamic general equilibrium model capturing the dynamic interaction between asset prices and arbitrageur capital. Third, arbitrageurs face multiple arbitrage opportunities with different characteristics, across which they must allocate their scarce capital.

We compute asset prices and arbitrageur positions in closed form when arbitrage is riskless. Using these closed-form solutions, we also compute asset prices and arbitrageur positions when arbitrage risk is small. We show that liquidity, volatility and correlations are non-monotonic in arbitrageur wealth: liquidity is smallest, volatility is largest, correlations between asset pairs with uncorrelated fundamentals are largest, and correlations between asset pairs with highly correlated fundamentals are smallest for intermediate levels of arbitrageur wealth.

Our analysis has left aside a number of important questions which we intend to address in future research. Market segmentation and financial constraints have been imposed exogenously; deriving these from more primitive frictions is an important question. Likewise, our analysis precludes capital flows in or out of the arbitrage industry, as well as imperfect competition between arbitrageurs.

References (Incomplete)

Brunnermeier, Markus, and Lasse Pedersen (2009), "Market Liquidity and Funding Liquidity." *Review of Financial Studies*, 22, 2201-2238.

Gromb, Denis, and Dimitri Vayanos (2002), "Equilibrium and Welfare in Markets with Financially Constrained Arbitrageurs," *Journal of Financial Economics*, 66, 361-407.

Gromb, Denis, and Dimitri Vayanos (2010), "Limits of Arbitrage," *Annual Review of Financial Economics*, 2, 251-275.

Kondor, Péter (2009), "Risk in Dynamic Arbitrage: Price Effects of Convergence Trading," *Journal of Finance*, forthcoming.

Kyle, Albert S., and Wei Xiong (2001), "Contagion as a Wealth Effect of Financial Intermediaries," *Journal of Finance*, 56, 1401-1440.

Xiong, Wei (2001), "Convergence Trading with Wealth Effects" *Journal of Financial Economics*, 62, 247-292.

Appendix (Incomplete)

A Riskless Arbitrage

Proof of Proposition 1 and Corollary 1: We solve a given arbitrageur A's problem $\mathcal{P}_t$ using dynamic programming. We distinguish between arbitrageur A's wealth $\hat{W}_t$, and the arbitrageurs' total wealth W_t. In equilibrium $\hat{W}_t = W_t$, but distinguishing $\hat{W}_t$ from W_t is important as W_t influences prices while arbitrageur A can affect only $\hat{W}_t$. We denote $\{\hat{x}_{i,t}\}_{i\in\mathcal{A}}$ arbitrageur A's positions and $\hat{c}_t$ his consumption to distinguish them from the arbitrageurs' total positions $\{x_{i,t}\}_{i\in\mathcal{A}}$ and consumption c_t. We conjecture the value function

$$V(\hat{W}_t, W_t) = \frac{\log(\hat{W}_t)}{\beta} + v(W_t). \tag{A.1}$$

For riskless arbitrage, $(\hat{W}_t, W_t)$ are deterministic, and the Bellman equation is

$$\max_{\hat{x}_{i,t},\hat{c}_t} \left[\log(\hat{c}_t) + V_{\hat{W}}\left(r\hat{W}_t + 2\sum_{i\in\mathcal{A}} \hat{x}_{i,t}\Phi_{i,t} - \hat{c}_t\right) + V_W \mu_t^W - \beta V\right] = 0, \tag{A.2}$$

where μ_t^W denotes the drift of W_t. The first-order condition with respect to $\hat{c}_t$ yields $\hat{c}_t = \beta\hat{W}_t$. Optimizing over $\{\hat{x}_{i,t}\}_{i\in\mathcal{A}}$ amounts to maximizing $\sum_{i\in\mathcal{A}} \hat{x}_{i,t}\Phi_{i,t}$ subject to financial constraint (22). Since $\Phi_{i,t} \geq 0$ for $i \in \mathcal{A}$, the first-order condition yields the policy in the proposition. The maximum value of $2\sum_{i\in\mathcal{A}} \hat{x}_{i,t}\Phi_{i,t}$ is $\hat{W}_t \max_{j\in\mathcal{A}}\left(\frac{\Phi_{j,t}}{m_j}\right)$. Substituting into (A.2), the terms in $\hat{W}_t$ cancel out. Setting the remaining terms to zero determines the function $v(W_t)$. ■

Proof of Lemma 2: From Corollary 1, $\forall i \in \mathcal{I}$, $x_{i,t} = 0$ or $\Phi_{i,t} = m_i\Pi_t$, which implies $x_{i,t}\Phi_{i,t} = x_{i,t}m_i\Pi_t$. Substituting together with $c_t = \beta W_t$ into (21) and yields

$$dW_t = \left[(r-\beta)W_t + 2\Pi_t \sum_{i\in\mathcal{A}} m_i x_{i,t}\right] dt. \tag{A.3}$$

Eq. (26) follows from (22) and (A.3) by noting that when (22) is slack, $\Pi_t = 0$.

For (22) to be slack, arbitrageurs must be able to hold $x_{i,t} = \mu_i u_i$ for all $i \in \mathcal{I}$, which requires $W_t \geq 2\sum_{i\in\mathcal{A}} m_i\mu_i u_i \equiv W_c$. This also implies $y_{i,t} = -u_i$ for all $i \in \mathcal{I}$, and therefore $\Phi_{i,t} = 0$ from (21). $x_{i,t} > 0$ for all $i \in \mathcal{A}$ and $\Phi_{i,t} = 0$ implies $\Pi_t = 0$ (Corollary 1).

If $W_t < W_c$, $\exists i \in \mathcal{A}$ such that $x_{i,t} < \mu_i u_i$, which implies $y_{i,t} > u_i$ and $\Phi_{i,t} > 0$ (from Eq. (21)). This implies $\Pi_t > 0$ (Corollary 1). Moreover if arbitrageurs invest in all opportunities, their position $x_{i,t}$ in each of them is given by (35). Substituting into (22) (which holds as an equality) yields

$$W_t = 2\sum_{i\in\mathcal{A}} m_i\mu_i\left(u_i - \frac{m_i\Pi_t}{a_i\sigma_i^2}\right) = W_c - \frac{\Pi_t}{B}, \tag{A.4}$$

which implies (25). ■

Proof of Lemma 3: Eqs. (9) and (19) imply that

$$\Phi_{i,t} = a_i\sigma_i^2\left(u_i - \frac{x_{i,t}}{\mu_i}\right). \tag{A.5}$$

For $\Phi_{i,t}/m_i = \Pi_t$ and $x_{i,t} > 0$, (A.5) implies $a_i\sigma_i^2 u_i/m_i > \Pi_t$. Solving (A.5) for $x_{i,t}$ yields (35). ■

Proof of Proposition 3: We first determine W_s for $s \geq t$ such that $W_t < W_c$. Using (25), we can write (26) as

$$dW_t = (A - BW_t)W_t dt. \tag{A.6}$$

To integrate (A.6), we note that

$$\frac{d}{dt}\left(\frac{W_t e^{-At}}{A - BW_t}\right) = \frac{Ae^{-At}}{(A - BW_t)^2}\left[\frac{dW_t}{dt} - (A - BW_t)W_t\right] = 0,$$

where the second step follows from (A.6). Therefore, for $s > t$,

$$\frac{W_s e^{-As}}{A - BW_s} = \frac{W_t e^{-At}}{A - BW_t}.$$

Solving for W_s, we find the equations in the proposition. ■

Proof of Proposition 5: From Lemma 3 and Lemma 2. ■

Proof of Corollary 2: Eq. (38) implies that in the constrained region is

$$\frac{\partial\phi_{i,t}}{\partial W_t} = -m_i B \int_0^\infty \frac{e^{As}}{\left[\frac{B}{A}W_t\left(e^{As} - 1\right) + 1\right]^2} e^{-rs} ds. \tag{A.7}$$

Eq. (39) implies that in the unconstrained region is

$$\frac{\partial\phi_{i,t}}{\partial W_t} = -\frac{rm_i B}{(\beta - r)W_t}\left(\frac{W_c}{W_t}\right)^{\frac{r}{\beta-r}} \int_0^\infty \left[W_c - \frac{W_c e^{As}}{\frac{B}{A}W_c\left(e^{As} - 1\right) + 1}\right] e^{-rs} ds. \tag{A.8}$$

In both cases, $\partial\phi_{i,t}/\partial W_t$ is negative and increases as W_t increases. To show strict convexity, we also need to show that $\partial\phi_{i,t}/\partial W_t$ is continuous at $W_t = W_c$. Integrating (A.7) by parts, we find

$$\begin{aligned}
\frac{\partial\phi_{i,t}}{\partial W_t} &= \left[\frac{m_i}{W_t}\frac{e^{-rs}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}\right]_0^\infty + \frac{m_i}{W_t}\int_0^\infty \frac{re^{-rs}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}ds \\
&= -\frac{m_i}{W_t}\left(1-\int_0^\infty \frac{re^{-rs}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}ds\right), \qquad \text{(A.9)}
\end{aligned}$$

and therefore,

$$\left.\frac{\partial\phi_{i,t}}{\partial W_t}\right|_{W_t=W_c^-} = -\frac{m_i}{W_c} + \frac{m_i}{W_c}\int_0^\infty \frac{re^{-rs}}{\frac{B}{A}W_c\left(e^{As}-1\right)+1}ds. \qquad \text{(A.10)}$$

Moreover, (A.8) implies that

$$\begin{aligned}
\left.\frac{\partial\phi_{i,t}}{\partial W_t}\right|_{W_t=W_c^+} &= -\frac{rm_iB}{\beta-r}\int_0^\infty \left[1-\frac{e^{As}}{\frac{B}{A}W_c\left(e^{As}-1\right)+1}\right]e^{-rs}ds \\
&= -\frac{rm_iB}{\beta-r}\int_0^\infty \left[1-\frac{A}{BW_c}-\frac{1-\frac{A}{BW_c}}{\frac{B}{A}W_c\left(e^{As}-1\right)+1}\right]e^{-rs}ds. \qquad \text{(A.11)}
\end{aligned}$$

Using the definition of A (Eq. (27)), we find that (A.11) coincides with (A.10). ■

Proof of Corollary 3: The variable u_j affects $\phi_{i,t}$ through W_c and $A \equiv r-\beta+BW_c$. Since $\partial W_c/\partial u_j$ is a positive constant, it suffices to show the corollary for W_c rather than u_j.

To determine the sign of the cross-effect, we examine how the effect of W_t on $\phi_{i,t}$ depends on W_c. Consider first the constrained region. Since A is increasing in W_c, (A.9) implies that $\partial\phi_{i,t}/\partial W_t$ is decreasing in W_c, i.e., $\partial^2\phi_{i,t}/\partial W_c\partial W_t < 0$. Consider next the unconstrained region. Eq. (A.8) implies that $\partial^2\phi_{i,t}/\partial W_c\partial W_t < 0$ if

$$\frac{\partial}{\partial W_c}\left[W_c-\frac{W_ce^{As}}{\frac{B}{A}W_c\left(e^{As}-1\right)+1}\right] > 0. \qquad \text{(A.12)}$$

For a general value of W_t,

$$W_c-\frac{W_te^{As}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1} = W_c-\frac{A}{B}-\frac{W_t-\frac{A}{B}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1} = \frac{\beta-r}{B}-\frac{W_t-\frac{A}{B}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}. \qquad \text{(A.13)}$$

Therefore, for $W_t = W_c$,

$$W_c-\frac{W_ce^{As}}{\frac{B}{A}W_c\left(e^{As}-1\right)+1} = \frac{\beta-r}{B}-\frac{\beta-r}{B}\frac{1}{\frac{B}{A}W_c\left(e^{As}-1\right)+1}.$$

This expression is increasing in W_c since A is increasing in W_c. Thus, in both the constrained and unconstrained regions, $\partial^2\phi_{i,t}/\partial W_c\partial W_t = \partial^2\phi_{i,t}/\partial W_t\partial W_c < 0$. To conclude that the effect of W_c on $\phi_{i,t}$ is more negative the larger W_t is, we also need to show that $\partial\phi_{i,t}/\partial W_c$ is continuous at $W_t = W_c$. Eq. (38) implies that in the constrained region

$$\frac{\partial\phi_{i,t}}{\partial W_c} = m_i B \int_0^\infty \frac{\partial}{\partial W_c}\left[W_c - \frac{W_t e^{As}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}\right] e^{-rs} ds. \tag{A.14}$$

Eq. (39) implies that in the unconstrained region

$$\begin{aligned}\frac{\partial\phi_{i,t}}{\partial W_c} =& \frac{rm_iB}{(\beta-r)W_c}\left(\frac{W_c}{W_t}\right)^{\frac{r}{\beta-r}}\int_0^\infty\left[W_c - \frac{W_c e^{As}}{\frac{B}{A}W_c\left(e^{As}-1\right)+1}\right]e^{-rs}ds \\ &+ m_iB\left(\frac{W_c}{W_t}\right)^{\frac{r}{\beta-r}}\int_0^\infty\frac{\partial}{\partial W_c}\left[W_c - \frac{W_c e^{As}}{\frac{B}{A}W_c\left(e^{As}-1\right)+1}\right]e^{-rs}ds. \end{aligned}\tag{A.15}$$

Eqs. (A.14) and (A.15) imply that

$$\begin{aligned}&\left.\frac{\partial\phi_{i,t}}{\partial W_c}\right|_{W_t=W_c^-} = \left.\frac{\partial\phi_{i,t}}{\partial W_c}\right|_{W_t=W_c^+} \\ &\Leftrightarrow 0 = \frac{rm_iB}{\beta-r}\int_0^\infty\left[1-\frac{e^{As}}{\frac{B}{A}W_c\left(e^{As}-1\right)+1}\right]e^{-rs}ds - m_iB\int_0^\infty\frac{e^{As}}{\left[\frac{B}{A}W_c\left(e^{As}-1\right)+1\right]^2}e^{-rs}ds \\ &\Leftrightarrow \left.\frac{\partial\phi_{i,t}}{\partial W_t}\right|_{W_t=W_c^-} = \left.\frac{\partial\phi_{i,t}}{\partial W_t}\right|_{W_t=W_c^+},\end{aligned}$$

which holds.

We next show that $\partial\phi_{i,t}/\partial W_c > 0$. Eq. (A.13) implies that $\partial\phi_{i,t}/\partial W_c > 0$ in the constrained region if the function

$$G : A \longrightarrow \frac{W_t - \frac{A}{B}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}$$

is decreasing in A. Since the denominator is increasing in A, $G(A)$ is decreasing if $W_t > A/B$. Since $\partial^2\phi_{i,t}/\partial W_t\partial W_c < 0$, inequality $\partial\phi_{i,t}/\partial W_c > 0$ holds also if $W_t < A/B$. Finally, (A.12) and (A.15) imply $\partial\phi_{i,t}/\partial W_c > 0$ in the unconstrained region. ■

Proof of Corollaries 4, 5 and 6: The first result follows from (38) and (39) by observing that the only asset-specific term in each equation is m_i. Using the same observation, we can derive the second result from (A.7) and (A.8), and the third result from (A.14) and (A.15). ■

Proof of Corollary 7: Follows from Proposition 5. ■

B Risky Arbitrage

Proof of Proposition 6: We proceed as in the proof of Proposition 1, conjecturing the value function (A.1). The Bellman equation is

$$\max_{\hat{x}_{i,t},\hat{c}_t} \Bigg\{ \log(\hat{c}_t) + V_{\hat{W}} \left(r\hat{W}_t + 2\sum_{i\in\mathcal{A}} \hat{x}_{i,t}\Phi_{i,t} - \hat{c}_t \right) + \frac{1}{2}V_{\hat{W}\hat{W}} \left[\sum_{j\in\mathcal{A}} \left(\hat{x}_{j,t}\sigma_j^f - \sum_{i\in\mathcal{A}} \hat{x}_{i,t}\sigma_{i,j,t}^{f\phi} \right)^2 + \sum_{j\in\mathcal{A}} \left(\sum_{i\in\mathcal{A}} \hat{x}_{i,t}\sigma_{i,j,t}^{u\phi} \right)^2 \right] + V_W\mu_t^W + \frac{1}{2}V_{WW} \left[\sum_{j\in\mathcal{A}} \left(\sigma_{j,t}^{fW}\right)^2 + \sum_{j\in\mathcal{A}} \left(\sigma_{j,t}^{uW}\right)^2 \right] - \beta V \Bigg\} = 0, \tag{B.1}$$

where μ_t^W denotes the drift of W_t. The first-order condition with respect to $\hat{c}_t$ yields $\hat{c}_t = \beta\hat{W}_t$. Optimization over $\{\hat{x}_{i,t}\}_{i\in\mathcal{A}}$ amounts to maximizing

$$\sum_{i\in\mathcal{A}} x_{i,t}\Phi_{i,t} - \frac{1}{2\hat{W}_t} \left[\sum_{j\in\mathcal{A}} \left(\hat{x}_{j,t}\sigma_j^f - \sum_{i\in\mathcal{A}} \hat{x}_{i,t}\sigma_{i,j,t}^{f\phi} \right)^2 + \sum_{j\in\mathcal{A}} \left(\sum_{i\in\mathcal{A}} \hat{x}_{i,t}\sigma_{i,j,t}^{u\phi} \right)^2 \right] \tag{B.2}$$

subject to the financial constraint (8). The first-order condition yields the policy in the proposition. The policy $\{\hat{x}_{i,t}\}_{i\in\mathcal{A}}$ and the maximum value of (B.2) are linear in $\hat{W}_t$. Substituting into (B.1), the terms in $\hat{W}_t$ cancel. Setting the remaining terms to zero, determines the function $v(W_t)$. ■

Proof of Lemma 4: Substituting $\sigma_{i,j,t}^{f\phi}$ from (53) into (50) and solving for $\sigma_{j,t}^{fW}$, we find (54). Substituting $\sigma_{i,j,t}^{u\phi}$ from (53) into (50) and solving for $\sigma_{j,t}^{uW}$, we find (57). ■

Proof of Proposition 7: When $\{\sigma_j^f, \sigma_j^u, \kappa_j^u\}_{j\in\mathcal{A}}$ are small, the highest-order term in (57) is

$$\sigma_{i,j,t}^{u\phi 0} \equiv -\frac{\partial\phi_{i,t}^0}{\partial W_t} \frac{2\sum_{k\in\mathcal{A}} x_{k,t}^0 \frac{\partial\phi_{k,t}^0}{\partial u_{j,t}}\sigma_j^u}{1 + 2\sum_{k\in\mathcal{A}} x_{k,t}^0 \frac{\partial\phi_{k,t}^0}{\partial W_t}} + \frac{\partial\phi_{i,t}^0}{\partial u_{j,t}}\sigma_j^u, \tag{B.3}$$

where $(\phi_{k,t}^0, x_{k,t}^0)$ denote the functions $(\phi_{k,t}, x_{k,t})$ evaluated under riskless arbitrage at the point $(W_t, \{u_{k,t}\}_{k\in\mathcal{A}})$. The proposition will follow if we show that $\sigma_{i,j,t}^{u\phi 0}$ is positive, decreasing in W_t for $W_t > W_{c,t}$, and increasing in W_t for $A/B \le W_t < W_c$.

When $W_t > W_{c,t}$, $x^0_{k,t} = \mu_k u_{k,t}$, and the denominator in (B.3) is equal to

$$1 + 2\sum_{k\in\mathcal{A}} \mu_k u_{k,t} \frac{\partial \phi^0_{k,t}}{\partial W_t} = 1 + \frac{2\sum_{k\in\mathcal{A}} m_k \mu_k u_{k,t}}{W_t}\left(-1 + \int_0^\infty \frac{re^{-rs}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}ds\right)$$

$$= 1 - \frac{W_{c,t}}{W_t} + \frac{W_{c,t}}{W_t}\int_0^\infty \frac{re^{-rs}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}ds > 0, \tag{B.4}$$

where the first step follows from (A.9), and the second from $A \equiv BW_{c,t} - (\beta - r) > 0$. The variable $\sigma^{u\phi 0}_{i,j,t}$ is positive and decreasing in W_t because (B.4) is positive, $\frac{\partial \phi^0_{k,t}}{\partial u_{j,t}}$ is positive and decreasing in W_t (Corollary 3), and $\frac{\partial \phi^0_{k,t}}{\partial W_t}$ is negative and increasing in W_t (Corollary 2).

When $W_t < W_{c,t}$, (A.9) implies that

$$\frac{\partial \phi^0_{i,t}}{\partial W_t} = -\frac{m_i}{W_t} + \frac{m_i}{W_t}\int_0^\infty \frac{re^{-rs}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}ds, \tag{B.5}$$

and (A.13) and (A.14) imply that

$$\frac{\partial \phi^0_{i,t}}{\partial u_{j,t}} = -m_i m_j \mu_j B \int_0^\infty \frac{\partial}{\partial W_{c,t}}\left[\frac{W_t - \frac{A}{B}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}\right] e^{-rs}ds. \tag{B.6}$$

Moreover, since $\{x^0_{k,t}\}_{k\in\mathcal{A}}$ satisfy the financial constraint (22), (A.9) implies that

$$1 + 2\sum_{k\in\mathcal{A}} x^0_{k,t} \frac{\partial \phi^0_{k,t}}{\partial W_t} = 1 + \frac{2\sum_{k\in\mathcal{A}} m_k x^0_{k,t}}{W_t}\left(-1 + \int_0^\infty \frac{re^{-rs}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}ds\right)$$

$$= \int_0^\infty \frac{re^{-rs}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}ds, \tag{B.7}$$

and (A.13) and (A.14) imply that

$$2\sum_{k\in\mathcal{A}} x^0_{k,t} \frac{\partial \phi^0_{k,t}}{\partial u_{j,t}} = \left(2\sum_{k\in\mathcal{A}} m_k x^0_{k,t}\right) m_j \mu_j B \int_0^\infty \frac{\partial}{\partial W_{c,t}}\left[W_{c,t} - \frac{W_t e^{As}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}\right] e^{-rs}ds.$$

$$= -W_t m_j \mu_j B \int_0^\infty \frac{\partial}{\partial W_{c,t}}\left[\frac{W_t - \frac{A}{B}}{\frac{B}{A}W_t\left(e^{As}-1\right)+1}\right] e^{-rs}ds. \tag{B.8}$$

Substituting (B.5)-(B.8) into (B.3), we find

$$\sigma^{u\phi 0}_{i,j,t} = -m_i m_j \mu_j B \sigma^u_j \frac{\int_0^\infty \frac{\partial}{\partial W_{c,t}}\left[\frac{W_t - \frac{A}{B}}{\frac{B}{A}W_t(e^{As}-1)+1}\right] e^{-rs}ds}{\int_0^\infty \frac{re^{-rs}}{\frac{B}{A}W_t(e^{As}-1)+1}ds}$$

$$= m_i m_j \mu_j B \sigma^u_j \left\{\frac{1}{r} + \frac{\int_0^\infty \frac{\left(W_t - \frac{A}{B}\right)\frac{B}{A^2}W_t\left(Ase^{As} - e^{As} + 1\right)}{\left[\frac{B}{A}W_t(e^{As}-1)+1\right]^2} e^{-rs}ds}{\int_0^\infty \frac{re^{-rs}}{\frac{B}{A}W_t(e^{As}-1)+1}ds}\right\}, \tag{B.9}$$

where the second step follows from $A \equiv r - \beta + BW_{c,t}$. Since the function

$$W_t \longrightarrow \frac{\left(W_t - \frac{A}{B}\right) \frac{B}{A^2} W_t \left(Ase^{As} - e^{As} + 1\right)}{\left[\frac{B}{A} W_t \left(e^{As} - 1\right) + 1\right]^2}$$

is positive for $W_t > A/B$ and increasing in W_t for $W_t > A/(2B)$, and the function

$$W_t \longrightarrow \int_0^\infty \frac{re^{-rs}}{\frac{B}{A} W_t \left(e^{As} - 1\right) + 1} ds$$

is positive and decreasing in W_t, $\sigma^{u\phi 0}_{i,j,t}$ is positive and increasing in W_t for $A/B \leq W_t < W_c$. ■

Proof of Proposition 8: The loading $\sigma^{f\phi}_{i,j,t}$ for $(i,j) \in \mathcal{A}$ is given by (55). When $\{\sigma^f_j, \sigma^u_j, \kappa^u_j\}_{j\in\mathcal{A}}$ are small, the highest-order term in (55) is

$$\sigma^{f\phi 0}_{i,j,t} \equiv \frac{\partial \phi^0_{i,t}}{\partial W_t} \frac{2x^0_{j,t}\sigma^f_j}{1 + 2\sum_{k\in\mathcal{A}} x^0_{k,t} \frac{\partial \phi^0_{k,t}}{\partial W_t}}. \tag{B.10}$$

The proposition will follow from (47) and the properties of $\sigma^{u\phi 0}_{i,j,t}$ shown in the proof of Proposition 7, if we show that $\sigma^{f\phi 0}_{i,j,t}$ is negative, increasing in W_t for $W_t > W_{c,t}$, and decreasing in W_t for $W_t < W_{c,t}$ and j satisfying (59).

When $W_t > W_{c,t}$, $x^0_{k,t} = \mu_k u_{k,t}$ and the denominator in (B.10) is equal to (B.4). Since σ^f_j, $x^0_{k,t}$, and (B.4) are positive, and $\frac{\partial \phi^0_{k,t}}{\partial W_t}$ is negative and increasing in W_t, $\sigma^{f\phi 0}_{i,j,t}$ is negative and increasing in W_t. When $W_t < W_{c,t}$,

$$x^0_{j,t} = \mu_j \left(u_{j,t} - \frac{m_j \Pi^0_t}{a_j \sigma^2_j} \right) = \mu_j \left[u_{j,t} - \frac{m_j B(W_{c,t} - W_t)}{a_j \sigma^2_j} \right], \tag{B.11}$$

where the first step follows from (35) and the second from (25). Substituting (B.5), (B.7) and (B.11) into (B.10), we find

$$\sigma^{f\phi 0}_{i,j,t} = -2m_i \mu_j \sigma^f_j \left[\frac{u_{j,t} - \frac{m_j B W_{c,t}}{a_j \sigma^2_j}}{W_t} + \frac{m_j B}{a_j \sigma^2_j} \right] \left[\frac{1}{\int_0^\infty \frac{re^{-rs}}{\frac{B}{A} W_t (e^{As} - 1) + 1} ds} - 1 \right]. \tag{B.12}$$

The first square bracket is positive since $x^0_{j,t} > 0$, and increases with W_t for j satisfying (59). Since the second square bracket is positive and increasing in W_t, $\sigma^{f\phi 0}_{i,j,t}$ is negative and decreasing in W_t.
■

Proof of Proposition 9: The correlation between assets $(i, i') \in \mathcal{A}^2$ is

$$\rho_{i,i',t} \equiv \frac{\left(\sigma^{f\phi}_{i,i,t} - \sigma^f_i\right)\sigma^{f\phi}_{i',i,t} + \sigma^{f\phi}_{i,i',t}\left(\sigma^{f\phi}_{i',i',t} - \sigma^f_{i'}\right) + \sum_{j\in\mathcal{A}/\{i,i'\}}\sigma^{f\phi}_{i,j,t}\sigma^{f\phi}_{i',j,t} + \sum_{j\in\mathcal{A}}\sigma^{u\phi}_{i,j,t}\sigma^{u\phi}_{i',j,t}}{\sigma^R_{i,t}\sigma^R_{i',t}}. \quad \text{(B.13)}$$

When $\{\sigma^f_j, \sigma^u_j, \kappa^u_j\}_{j\in\mathcal{A}}$ are small, the highest-order term in (B.13) is

$$\frac{\left(\sigma^{f\phi0}_{i,i,t} - \sigma^f_i\right)\sigma^{f\phi0}_{i',i,t} + \sigma^{f\phi0}_{i,i',t}\left(\sigma^{f\phi0}_{i',i',t} - \sigma^f_{i'}\right) + \sum_{j\in\mathcal{A}/\{i,i'\}}\sigma^{f\phi0}_{i,j,t}\sigma^{f\phi0}_{i',j,t} + \sum_{j\in\mathcal{A}}\sigma^{u\phi0}_{i,j,t}\sigma^{u\phi0}_{i',j,t}}{\sigma_i\sigma_{i'}}. \quad \text{(B.14)}$$

The properties of $\rho_{i,i',t}$ follow from (B.14) and the properties of $(\sigma^{u\phi0}_{i,j,t}, \sigma^{f\phi0}_{i,j,t})$ shown in the proofs of Propositions 7 and 8. The properties of $\rho_{i,-i',t}$ follow from $\rho_{i,-i',t} = -\rho_{i,i',t}$, which is implied from symmetry. To show the properties of $\rho_{i,-i,t}$, we note that symmetry implies that

$$\rho_{i,-i,t} = \frac{\sigma_i^2 - \left(\sigma^f_i - \sigma^{f\phi}_{i,i,t}\right)^2 - \sum_{j\in\mathcal{A}/\{i\}}\left(\sigma^{f\phi}_{i,j,t}\right)^2 - \sum_{j\in\mathcal{A}}\left(\sigma^{u\phi}_{i,j,t}\right)^2}{\left(\sigma^R_{i,t}\right)^2}.$$

When $\{\sigma^f_j, \sigma^u_j, \kappa^u_j\}_{j\in\mathcal{A}}$ are small, $\rho_{i,-i,t}$ is close to one. Using (47), we find that the highest-order term in $1 - \rho_{i,-i,t}$ is

$$\frac{2\left[\left(\sigma^f_i - \sigma^{f\phi0}_{i,i,t}\right)^2 + \sum_{j\in\mathcal{A}/\{i\}}\left(\sigma^{f\phi0}_{i,j,t}\right)^2 + \sum_{j\in\mathcal{A}}\left(\sigma^{u\phi0}_{i,j,t}\right)^2\right]}{\sigma_i^2}. \quad \text{(B.15)}$$

The comparative statics of (B.15) are the same as for (B.14). Therefore, the properties of $\rho_{i,-i,t}$ are opposite to those of $\rho_{i,i',t}$. ■

Proof of Lemma 5: When $\{\sigma^f_j, \sigma^u_j, \kappa^u_j\}_{j\in\mathcal{A}}$ are small, arbitrageurs long all opportunities and their first-order condition is (60). Combining with (48) and using (51), we find

$$x_{i,t} = \mu_i\left[u_{i,t} - \frac{\Psi_{i,t} + m_i\Pi_t}{a_i\left(\sigma^R_{i,t}\right)^2}\right], \quad \text{(B.16)}$$

where

$$\Psi_{i,t} \equiv \frac{1}{W_t}\left[(\sigma^f_i - \sigma^{f\phi}_{i,i,t})\sigma^{fW}_{i,t} - \sum_{j\in\mathcal{A}/\{i\}}\sigma^{f\phi}_{i,j,t}\sigma^{fW}_{j,t} - \sum_{j\in\mathcal{A}}\sigma^{u\phi}_{i,j,t}\sigma^{uW}_{j,t}\right].$$

Multiplying (B.16) by m_i and summing over $i \in \mathcal{A}$, we find that the financial constraint (8) holds as an equality if and only if

$$W_t = W_{c,t} - 2\sum_{i\in\mathcal{A}} \frac{m_i\mu_i\left(\Psi_{i,t} + m_i\Pi_t\right)}{a_i\left(\sigma^R_{i,t}\right)^2}. \tag{B.17}$$

For small $\{\sigma^f_j, \sigma^u_j, \kappa^u_j\}_{j\in\mathcal{A}}$, the highest-order term in $\Psi_{i,t}$ is

$$\Psi^0_{i,t} \equiv \frac{1}{W_t}\left[(\sigma^f_i - \sigma^{f\phi 0}_{i,i,t})\sigma^{fW0}_{i,t} - \sum_{j\in\mathcal{A}/\{i\}} \sigma^{f\phi 0}_{i,j,t}\sigma^{fW0}_{j,t} - \sum_{j\in\mathcal{A}} \sigma^{u\phi 0}_{i,j,t}\sigma^{uW0}_{j,t}\right], \tag{B.18}$$

where

$$\sigma^{fW0}_{j,t} \equiv \frac{2x^0_{j,t}\sigma^f_j}{1 + 2\sum_{i\in\mathcal{A}} x^0_{k,t}\frac{\partial\phi^0_{k,t}}{\partial W_t}} \quad \text{and} \quad \sigma^{uW0}_{j,t} \equiv -\frac{2\sum_{k\in\mathcal{A}} x^0_{k,t}\frac{\partial\phi^0_{k,t}}{\partial u_{j,t}}\sigma^u_j}{1 + 2\sum_{k\in\mathcal{A}} x^0_{k,t}\frac{\partial\phi^0_{k,t}}{\partial W_t}} \tag{B.19}$$

are the highest-order terms in $(\sigma^{fW}_{j,t}, \sigma^{uW}_{j,t})$, respectively. Since $\sigma^{f\phi 0}_{i,j,t} < 0$, $\sigma^{u\phi 0}_{i,j,t} > 0$, $\sigma^{fW0}_{j,t} > 0$ and $\sigma^{uW0}_{j,t} < 0$, $\Psi^0_{i,t} > 0$. Lemma 5 follows from (B.17), $\Pi_t \geq 0$ and $\Psi^0_{i,t} > 0$. ■

Proof of Proposition 10: In the region where arbitrageurs are unconstrained, $\Pi_t = 0$. Eq. (B.16) implies that when $\{\sigma^f_j, \sigma^u_j, \kappa^u_j\}_{j\in\mathcal{A}}$ are small, the two highest-order terms in $x_{i,t}$ are

$$x^1_{i,t} \equiv \mu_i\left(u_{i,t} - \frac{\Psi^0_{i,t}}{a_i\sigma^2_i}\right). \tag{B.20}$$

Since $\sigma^{f\phi 0}_{i,j,t}$ and $\sigma^{uW0}_{j,t}$ are negative and increasing in W_t, and $\sigma^{u\phi 0}_{i,j,t}$ and $\sigma^{fW0}_{j,t}$ are positive and decreasing in W_t, (B.18) implies that $\Psi^0_{i,t}$ is decreasing in W_t, and therefore, $x^1_{i,t}$ is increasing. Eq. (B.20) implies that

$$x^1_{i,t} - x^1_{i',t} = \frac{\mu}{a\sigma^2}\left(\Psi^0_{i',t} - \Psi^0_{i,t}\right), \tag{B.21}$$

where $(\sigma, a, \mu, u_t) \equiv (\sigma_i, a_i, \mu_i, u_{i,t}) = (\sigma_{i'}, a_{i'}, \mu_{i'}, u_{i',t})$. Noting that $\phi^0_{i,t}/m_i = \phi^0_{i',t}/m_{i'}$ and $x^0_{i,t} = \mu_i u_{i,t}$, and using (B.3), (B.10), (B.18), (B.19) and (B.21), we find

$$x^1_{i,t} - x^1_{i',t} = \frac{\mu}{a\sigma^2 W_t}\left\{\frac{2\mu u_t\left[\left(\sigma^f_{i'}\right)^2 - \left(\sigma^f_i\right)^2\right]}{1 + 2\sum_{i\in\mathcal{A}}\mu u_{k,t}\frac{\partial\phi^0_{k,t}}{\partial W_t}} + \left(1 - \frac{m_{i'}}{m_i}\right)\left(\sum_{j\in\mathcal{A}} \sigma^{f\phi 0}_{i,j,t}\sigma^{fW0}_{j,t} + \sum_{j\in\mathcal{A}} \sigma^{u\phi 0}_{i,j,t}\sigma^{uW0}_{j,t}\right)\right\}. \tag{B.22}$$

If $m_i > m_{i'}$ and $\sigma_i^f = \sigma_{i'}^f$, the first term in the curly bracket is zero. Since the second term is negative and increasing in W_t, $x_{i,t}^1 - x_{i',t}^1$ is increasing in W_t. If $m_i = m_{i'}$ and $\sigma_i^f > \sigma_{i'}^f$, the second term in the curly bracket is zero. Since the first term is negative and increasing in W_t, $x_{i,t}^1 - x_{i',t}^1$ is increasing in W_t. In both cases, $\partial x_{i,t}^1/\partial W_t > \partial x_{i',t}^1/\partial W_t > 0$, i.e., changes in arbitrageur wealth impact more strongly their position in opportunity $(i, -i)$ than $(i', -i')$.

In the region where arbitrageurs are constrained, $\Pi_t > 0$. Eq. (B.16) implies that when $\{\sigma_j^f, \sigma_j^u, \kappa_j^u\}_{j \in \mathcal{A}}$ are small, the two highest-order terms in $x_{i,t}$ are

$$x_{i,t}^1 = \mu_i \left\{ u_{i,t} - \frac{\Psi_{i,t} + m_i \Pi_t}{a_i \sigma_i^2} \left[1 + \frac{\left(\sigma_{i,t}^R\right)^2 - \sigma_i^2}{\sigma_i^2} \right] \right\}. \tag{B.23}$$

The comparative statics with respect to m_i follow by considering the highest-order term

$$x_{i,t}^0 = \mu_i \left(u_{i,t} - \frac{m_i \Pi_t}{a_i \sigma_i^2} \right),$$

i.e., as in the case of riskless arbitrage. (The ambiguous comparative statics with respect to σ_i^f follow by considering the term in the next order.) ■