Leverage and Liquidity Dry-ups: A Framework and Policy Implications *

Denis Gromb

London Business School
London School of Economics
and CEPR

Dimitri Vayanos

London School of Economics
CEPR and NBER

First draft: May 18, 2008

This draft: June 19, 2008

Preliminary and incomplete draft

Abstract

We model financial market liquidity as provided by financially constrained arbitrageurs. Market liquidity increases with the level of arbitrage capital, i.e., internal and external capital arbitrageurs can access frictionlessly. We show that liquidity dry-ups follow periods of low returns of arbitrageurs' risky investment opportunities, and that liquidity is correlated across markets. A welfare analysis reveals that arbitrageurs may fail take socially optimal positions in their investment opportunities, adversely affecting their ability to provide market liquidity. Finally, we discuss possible policy responses.

Keywords: Financial constraints, arbitrage, liquidity, contagion, welfare, public policy, central bank.

*Correspondence: Dimitri Vayanos, D.Vayanos@lse.ac.uk and Denis Gromb, dgromb@london.edu. We thank Alan Morrison and seminar participants at the Imperial College "Liquidity: Pricing and Risk Management" conference at the Bank of England for comments. All remaining errors are ours.

1 Introduction

The recent subprime crisis has highlighted (yet again!) the importance of intermediary capital for the liquidity of financial markets. As banks incurred large losses in the subprime market, they curtailed their funding of other activities, notably that 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.

Standard financial markets models are however ill-suited to analyze liquidity problems and related public policy issues. Indeed, the textbook counterparts of liquidity providers are arbitrageurs, meant to represent not only professional arbitrageurs such as hedge funds and proprietary trading desks, but more generally real world financial intermediaries such as dealers or investment banks. Arbitrageurs provide liquidity to other investors and bring prices closer to fundamentals. In standard models, they do such a good job of it that liquidity is perfect and prices coincide with fundamental values. This Absence of Arbitrage Opportunities result is the cornerstone of both the modern theory of asset pricing and its practice in industry. For those with a public policy interest however, it delivers a somewhat frustratingly simple and optimistic message: The equilibrium in financial markets is efficient, the standard welfare theorems apply and public intervention can at best consist only in redistributing a cake, not increasing its size. In other words, market liquidity is as good as it gets.

This paper presents a framework in which the welfare implications of financial market liquidity dry-ups and related policy issues can be discussed in a meaningful way. Building on the recent literature on the limits to arbitrage, it starts from the premise that arbitrageurs face financial constraints. Again, arbitrageurs are to be understood as individuals and institutions responsible for providing liquidity in different financial markets. The arbitrageurs' financial constraints, be they margin requirements, limited access to external capital or barriers to entry of new capital, affect the arbitrageurs' investment capacity, which in turn has consequences for asset prices and liquidity. Our focus in on these financial constraints' welfare effects. Indeed, we show that competitive arbitrageurs can fail to follow socially optimal investment strategies. We use this framework to discuss public policy issues.

We begin by modelling financial markets needing liquidity as in Gromb and Vayanos (2002, 2008). We consider two risky assets, assets A and B, with identical payoffs but traded in segmented markets. The demand by investors on each market, A- and B-investors respectively, for the corresponding 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 because of 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 positions in any risky asset 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, which we call "arbitrage capital", is insufficient, arbitrageurs may be unable to provide perfect liquidity to markets.

Given this building block, we consider a second investment opportunity that arbitrageurs can exploit before they face the arbitrage opportunity described above. This may be, for instance, an alternative arbitrage opportunity for a hedge fund, another stock for which a dealer is a market maker, or subprime mortgages financed by commercial banks. We model this opportunity as a risky asset, asset C, that some investors, C-investors are eager to sell to arbitrageurs for hedging reasons. The important point is that the investment opportunity is risky and its performance draws from and contributes to the same pool of arbitrage capital needed to ensure market liquidity. For instance, losses in the alternative investment opportunity will reduce the capital available to arbitrageurs for liquidity provision, resulting in liquidity dry-ups in otherwise unrelated financial markets. More generally, liquidity will tend to comove across markets which ultimately rely on the same pool of arbitrage capital for their liquidity.

We next move to the welfare analysis. Understanding the welfare implications of arbitrageurs' financial constraints is important as they underlie many policy debates. An example is the debate on systemic risk, i.e., on whether a worsening of the financial condition of some market participants can propagate into the financial system with harmful effects. One question our model allows us to study is whether market liquidity providers take an appropriate level of risk?[1]

In our model, arbitrage activity benefits all investors because arbitrageurs supply market liquidity. We show, however, that they can fail to do so in a (constrained) socially optimal fashion, given their financial constraints. We consider social welfare from an ex-ante perspective, defining it as the sum of all agents' certainty equivalents. We show that an exogenous change (away from its equilibrium value) in the arbitrageurs' initial position in asset C can yield a social welfare increase. While changing quantities has only second-order welfare effects because they are optimal given prices, this is not so for price changes. Indeed, price changes redistribute wealth among agents. In complete markets, wealth transfers would affect individual agents' welfare but not total welfare. Here, however, wealth transfers can and do affect total welfare because of financial constraints and imperfect risk-sharing. Our analysis highlights the channels through which this occurs.

First, wealth is transferred between C-investors and arbitrageurs. Consider, for instance, a decrease in the arbitrageurs' initial holding of asset C. This in effect, rations insurance to C-

[1] For example, during the 1998 crisis, it was feared that the positions of Long-Term Capital Management (LTCM), a major hedge fund and one of the worst hit, were so large that their forced liquidation would depress prices. This could disrupt markets and possibly jeopardize the financial system, with consequences reaching far beyond LTCM's investors. Such concerns were behind the Fed's controversial decision to orchestrate LTCM's rescue.

investors raising the price of insurance, i.e., allowing arbitrageurs to buy asset C at a lower price. Hence arbitrageurs are better off and C-investors worse off. This transfer, however, increases welfare because arbitrageurs have better investment opportunities than C-investors. Indeed, they can provide the liquidity A- and B-investors crave, while C-investors cannot.

Second, wealth is transferred between arbitrageurs and A- and B-investors. Again, consider a decrease in the arbitrageurs' initial holding of asset C. Since arbitrageurs take less risk, arbitrageurs' wealth is less volatile and so is therefore their ability to provide liquidity to A- and B-investors. This benefits A- and B-investors in bad states where arbitrage capital and liquidity are low, and benefits arbitrageurs in good states where they are high. These state-contingent transfers increase welfare if A- and B-investors suffer more than arbitrageurs in bad states. This is the case, for example, if they are more risk-averse than arbitrageurs.

In summary, central to our welfare analysis is the result that competitive arbitrageurs do not generally follow a socially optimal risk-management policy. This arises from their failing to internalize the price effects of their investment decisions. We use these insights from our welfare analysis to discuss a number of policy options. Indeed, policy responses must curb arbitrageurs' risk-management policy, directly or indirectly. Suppose for instance that arbitrageurs take too much risk.

A first possible route may be to increase the cost of risk-taking, by setting a tax or, perhaps more realistically, a capital requirement. Importantly, a capital requirement would be effective even if it applied not to the arbitrageurs themselves but only to their providers of frictionless external capital, e.g., banks financing hedge funds. Indeed, the aim of the policy would be to reduce the amount of arbitrage capital available for risk-taking. Note that much that while much of the regulation of financial institutions is concerned with default risk, this is not so in our model as there is no default. Instead, regulation would have the broader objective of ensuring that arbitrageurs have capital when it matters most for society.

Another related policy option might be to affect directly the arbitrageurs' risk-management policy by forcing them to save capital for bad times, when it is needed most. Such a policy may even be budget-neutral in that it would amount to a tax in good times and a positive transfer in bad times. This may again be implemented by tightening and loosening capital requirements of suppliers of market liquidity and/or their close financiers.

In our model, arbitrageurs fail to internalize the price effects of their investment decisions because they are perfectly competitive. Imperfect competition among arbitrageurs might lead them to internalize some of the price effects, possibly leading them to adopt investment policies closer to the social optimum.

Our analysis builds on the recent literature on the limits to arbitrage and, more particularly, on financially constrained arbitrage.[2,3] Gromb and Vayanos (2002) introduce a model of arbi-

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

[3]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

trageurs 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. The present paper uses a simplified version of the same model, but extends the analysis by considering alternative investment opportunities. These can increase the volatility of arbitrage capital, and ultimately that of market liquidity.

The Gromb and Vayanos (2002) model is extended to multiple investment opportunities in a static setting by Brunnermeier and Pedersen (2008) and a dynamic setting by Gromb and Vayanos (2008). Both show how financial constraints imply that shocks propagate and liquidity co-moves across markets. 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. None of these three papers does however conduct a welfare analysis.

The paper proceeds as follows. Section 2 presents a model in which market liquidity is determined by arbitrage capital. This serves as the building block for the rest of the analysis. Section 3 introduces an alternative investment opportunity for arbitrageurs and solves for the resulting equilibrium. Section 4 presents a welfare analysis showing that arbitrage capital may not be used in a socially (constrained) efficient manner. Section 5 discusses some policy implications of our analysis. Section 6 concludes. The Appendix contains mathematical proofs.

2 A Model of Market Liquidity Based on Arbitrage Capital

We consider the role of limited arbitrage and intermediary capital for financial market liquidity. We model an arbitrage opportunity as a situation of unsatisfied demand for liquidity, and arbitrageurs as specialist individuals or institutions uniquely able to providing liquidity. However, financial constraints can prevent them from closing the arbitrage opportunity. Consequently, market liquidity is imperfect and depends on arbitrage capital, i.e., the pool of internal or external capital arbitrageurs can access frictionlessly.

2.1 Arbitrage Opportunity

There are two dates $t = 1, 2$. Consider two segmented markets, A and B, in each of which investors can trade a different risky asset and the riskfree asset, with an exogenous return of 0. The risky assets yield identical payoffs. The only difference across markets is the investors' demand for the risky asset. Hence the risky assets' prices can differ, reflecting gains from trade between investors in the two markets unrealized due to market illiquidity, i.e., segmentation.

Segmented markets. In market $i = A, B$, investors are competitive and form a measure μ_i

Vayanos (2002, 2008) contain more detailed descriptions of the relevant literature.

continuum. At $t = 1$, these investors, whom we call i-investors, can invest in a single risky asset, asset i, and in the riskfree asset. For instance, A-investors cannot take positions in asset B. Market segmentation is taken as given, i.e., i-investors are assumed to face prohibitively large transaction costs for investing in any other risky asset than asset i. These costs can be due to physical factors (e.g., distance), information asymmetries or institutional constraints.

Payoffs. Asset i is in zero net supply and pays a single random dividend at $t = 2$

$$\delta_i = \overline{\delta} + \epsilon_{i,2}, \tag{1}$$

where $\overline{\delta}$ is a constant, and $\epsilon_{i,t}$ is a zero-mean random variable revealed at $t = 2$, and symmetrically distributed around zero over the finite support $[-\overline{\epsilon}_{i,2}, +\overline{\epsilon}_{i,2}]$.[4] We assume that assets A and B pay the same dividend, i.e., $\delta_A = \delta_B$.

Utility. i-investors have initial wealth $w_{i,1}$ at $t = 1$ and exponential utility over their wealth $w_{i,2}$ at $t = 2$, which is restricted to be non-negative, i.e.,

$$-\exp\left(-\alpha_i w_{i,2}\right) \qquad \text{with } \alpha_i > 0 \quad \text{and} \quad w_{i,2} \geq 0. \tag{2}$$

At $t = 2$, each i-investor receives an endowment correlated with asset i's dividend

$$u_i \cdot \epsilon_{i,2}. \tag{3}$$

The coefficient u_i measures the extent to which the endowment covaries with δ_i. If u_i is large and positive (negative), the shock and the dividend are highly positively (negatively) correlated, and thus the i-investors' willingness to hold asset i is low (high). Hence, we refer to u_i as i-investors' supply shock to reflect that their demand for asset i decreases with u_i.[5]

We assume that A- and B-investors are identical (i.e., $\mu_A = \mu_B$ and $\alpha_A = \alpha_B$) but for the fact that they incur opposite supply shocks, i.e.,

$$u_A = -u_B > 0. \tag{4}$$

Because A- and B-investors incur different shocks at $t = 2$, they have different demands for the risky asset at $t = 1$: A-investors want to sell asset A and B-investors buy asset B. However, they cannot realize the gains from trade due to the market illiquidity implied by segmentation.

2.2 Arbitrageurs

A- and B-investors' unsatisfied demand for liquidity creates a role for individuals or institutions able to provide such liquidity, which we call arbitrageurs. These represent not only professional arbitrageurs such as hedge funds and prop desks, but more generally financial intermediaries

[4] The bounded support assumption plays a role for the financial constraint (see below). Less importantly, it implies that assuming $\overline{\delta}$ large enough ensures that dividends are positive.

[5] To be consistent with the zero net supply assumption, the endowments can be interpreted as positions in a different but correlated asset, e.g., labor income. This specification of endowments is quite standard in the market microstructure literature (see O'Hara, 1995).

such as dealers or investment banks. For instance, arbitrageurs may be hedge funds trading bond vs. futures, or market makers for particular stocks. To study the role of the short supply of smart money for market liquidity, we assume that arbitrageurs face financial constraints in that the pool of capital, internal or external, they can access frictionlessly is limited. These constraints can prevent them from providing perfect market liquidity.

Utility. Arbitrageurs are competitive and form a measure 1 continuum. They have wealth W_1 at $t = 1$ and have exponential utility over wealth W_2 at $t = 2$, which is restricted to be non-negative, i.e., [6]

$$-\exp(-\alpha W_2) \qquad \text{with } \alpha > 0 \quad \text{and} \quad W_2 \geq 0. \tag{5}$$

Better investment opportunities. Arbitrageurs are specialists who, unlike other investors, are able to invest in all risky assets as well as in the riskless asset. Since assets A and B pay the same dividend, a price discrepancy between them constitutes an arbitrage opportunity, which arbitrageurs can exploit. So doing, arbitrageurs provide liquidity in otherwise illiquid markets.

Financial constraints. We assume arbitrageurs face financial constraints in that they must collateralize their positions in each asset separately. To buy or short $x_{i,1}$ shares of asset $i = A, B$ at $t = 1$, they must post $mx_{i,1}$ units of riskfree asset as collateral until $t = 2$. We treat m as exogenous. The case of no financial constraint corresponds to $m = 0$. That where arbitrageurs must fully collateralize positions in asset $i = A, B$ corresponds to $m = \left(\overline{\epsilon}_{A,2} - \phi_{A,1}\right)$ (Gromb and Vayanos (2002, 2008)).[7] While we refer to W_t as the arbitrageurs' wealth, it should be understood as the pool of capital, internal or external, which arbitrageurs can access frictionlessly. Hence our model captures the effect of arbitrage capital being in limited supply.

2.3 Equilibrium

At date t, asset i's price is denoted by $p_{i,t}$, and its risk premium defined and denoted by:

$$\phi_{i,t} \equiv E_t[\delta_i] - p_{i,t}. \tag{6}$$

Definition 1 *A competitive equilibrium consists of prices $p_{i,t}$, asset holdings of the i-investors, $y_{i,t}$, and of the arbitrageurs, $x_{i,t}$ at date t, such that given the prices, $y_{i,t}$ is optimal for each i-investor and $x_{i,t}$ for each arbitrageur, and the market for each risky asset clears:*

$$\mu_i y_{i,t} + x_{i,t} = 0. \tag{7}$$

[6] By fixing the measure and wealth of the arbitrageurs, we rule out entry in the arbitrage industry. This seems a reasonable assumption at least for understanding short-run market behavior. However, an alternative interpretation is that after $t = 2$, enough new arbitrageurs or new arbitrage capital enter and eliminate the arbitrage opportunity.

[7] We do not consider financial constraints for outside investors. That is, we assume that either these investors do not face constraints or that their constraints are not binding. The latter situation will arise if the outside investors' initial wealth is large enough. Indeed, with exponential utility, optimal holdings of the risky asset are independent of wealth, and so are capital gains. Moreover, since asset payoffs and supply shocks have bounded support, capital gains are also bounded. Therefore, for large enough initial wealth, capital losses are always smaller than wealth, and the financial constraint is not binding. Note that the initial wealth of the outside investors need not exceed that of the arbitrageurs. Indeed, if the measures μ_i of the outside investors are large enough, the arbitrageurs' positions are much larger than those of the outside investors, and thus require more collateral.

Because of our model's symmetry, we can show the existence of a competitive equilibrium that is AB-symmetric in the following sense.

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

Intuitively, the risk premia are opposites because assets are in zero net supply and the supply shocks of the A- and B-investors are opposites. The arbitrageurs' positions are opposites because the risk premia are opposites. Note that arbitrageurs act as intermediaries. Suppose A-investors receive a positive supply shock, in which case B-investors' shock is negative. Arbitrageurs buy asset A from the A-investors, and sell asset B to the B-investors, thereby making a profit, while at the same time providing liquidity to A- and B-investors. Finally, A- and B-investors' positions must be opposites for markets to clear. Note that by symmetry, we have:

$$\frac{p_{A,t} + p_{B,t}}{2} = E_t[\delta_i]. \tag{8}$$

Hence asset A's risk premium at date t equals half the price wedge between assets A and B, i.e.,

$$\phi_{A,t} = \frac{p_{B,t} - p_{A,t}}{2}. \tag{9}$$

It is therefore also an inverse measure of market liquidity, with perfect liquidity corresponding to markets A and B being integrated, and therefore to $\phi_{A,t} = 0$.

2.3.1 Outside Investors

At $t = 1$, each i-investor chooses $y_{i,1}$, her holding of asset i, to maximize their expected utility:

$$\max_{y_{i,1}} -E_t \exp(-\alpha_i w_{i,2}), \tag{10}$$

subject to their budget constraint which takes the following form. At $t = 1$, each i-investor invests $y_{i,1}p_{i,1}$ in asset i and the rest of his wealth, $(w_{i,1} - y_{i,1}p_{i,1})$, in the riskfree asset. By $t = 2$, the investor has received an endowment $u_i\delta_i$. Hence an i-investor's budget constraint is

$$w_{i,2} = y_{i,1}p_{i,2} + (w_{i,1} - y_{i,1}p_{i,1}) + u_i\epsilon_{i,2} = w_{i,1} + y_{i,1}\phi_{i,1} + (y_{i,1} + u_i)\epsilon_{i,2}. \tag{11}$$

The term $w_{i,1}$ is the investor's wealth at $t = 1$. The term $y_{i,1}\phi_{i,1}$ is the investor's expected profit between $t = 1$ and $t = 2$. The term $(y_{i,1} + u_i)\epsilon_{i,2}$ is the risk i-investors bear. It is the sum of the uncertain part of the dividend and of the uncertain endowment at $t = 2$.

Lemma 1 *A function $f_{i,1}$ positive and strictly convex exists such that each i-investor's problem at $t = 1$ can be written as:*

$$\max_{y_{i,1}} \quad w_{i,1} + y_{i,1}\phi_{i,1} - f_{i,1}(y_{i,1} + u_i). \tag{12}$$

This expression reflects the risk-return trade-off facing i-investors. The second term is their expected profit. The third is the cost of bearing risk given the uncertainty of dividends and of the supply shock, i.e., $f_{i,1}$ is a cost function. The first-order condition can be written as

$$f'_{i,1}(y_{i,1} + u_i) = \phi_{i,1}. \tag{13}$$

2.3.2 Arbitrageurs

Arbitrageurs can invest in all risky assets as well as in the riskless asset. Hence, an arbitrageur's budget constraint when he enters AB spread trades (i.e., $x_{A,1} = -x_{B,1}$) is

$$W_2 = W_1 + x_{A,1}(-p_{A,1} + p_{A,2}) + x_{B,1}(-p_{B,1} + p_{B,2}) = W_1 + 2x_{A,1}\phi_{A,1}. \tag{14}$$

In other words, arbitrageurs can eliminate the dividend risk $\epsilon_{A,2}$ by taking opposite positions in assets A and B, and exploit any price discrepancy between these assets (i.e., if $\phi_{A,1} > 0$).

Each arbitrageur maximizes this objective subject to his financial constraint

$$2mx_{A,1} \leq W_1. \tag{15}$$

2.3.3 Equilibrium

Clearly, an arbitrageur must invest as much as possible in the AB spread trade as long as $\phi_{A,1} > 0$, i.e., they "max out" their financial constraint. At the same time, the arbitrageurs' aggregate spread trade reduces the price wedge $\phi_{A,1}$. The equilibrium outcome depends on whether arbitrageurs close the price wedge before their financial constraint binds.

Proposition 1 *The equilibrium is as follows.*

- *If $W_1 < 2m\mu_A u_A$, the arbitrageurs' financial constraint binds and markets are illiquid: Assets A and B trade at different prices ($\phi_{A,1} > 0$). Market liquidity increases with arbitrage capital (i.e., $\partial\phi_{A,1}/\partial W_1 < 0$).*
- *If $W_1 \geq 2m\mu_A u_A$, the arbitrageurs' financial constraint is slack and market liquidity is perfect: The price wedge between assets A and B is closed ($\phi_{A,1} = 0$).*

This result highlights the role of arbitrage capital for market liquidity. First, absent financial constraints (i.e., $m = 0$), markets are perfectly liquid ($\phi_{A,1} = 0$) and therefore arbitrage capital plays no role ($\partial\phi_{A,1}/\partial W_1 = 0$). With financial constraints however, market liquidity depends positively on the level of arbitrage capital ($\partial\phi_{A,1}/\partial W_1 < 0$). In particular, liquidity dry-ups correspond to situations in which arbitrage capital is low relative to the need for liquidity, i.e., W_1 small and $\mu_A u_A$ large. In reality, arbitrage capital is, in part, the result of the performance of past investments. In that case, our model predicts that liquidity dry-ups follow periods of low returns for arbitrageurs and/or their financiers.

3 Alternative Investment Opportunity

We now study the effect of arbitrageurs' alternative investment opportunities on their ability to provide market liquidity. Alternative opportunities can be interpreted as assets in the portfolios of arbitrageurs, e.g., hedge funds. Under the broader interpretation of arbitrageur wealth as the pool of capital that they can access without frictions, alternative opportunities can also be viewed as assets available to the providers of arbitrage capital. For instance, they could represent mortgages extended by banks that are also financing hedge funds or market makers.

3.1 The Model

We model the alternative opportunities as asset C, traded at $t = 0, 1, 2$. At $t = 0$, some investors, C-investors, are keen on selling asset C for hedging reasons, and arbitrageurs can profit by absorbing some of that risk. Doing so, however, they must consider the effect of trading profits and losses on their ability to exploit the AB arbitrage opportunity at $t = 1$.

Segmented markets. In market C, investors are competitive and form a measure μ_C continuum. At $t = 0, 1$, C-investors can invest in a single risky asset, asset C, and in the riskfree asset which has an exogenous return of 0. Market C is segmented from markets A and B.

Payoff. Asset C is in zero net supply and pays a single random dividend at $t = 2$

$$\delta_C = \overline{\delta} + \epsilon_{C,1} + \epsilon_{C,2}, \tag{16}$$

where $\epsilon_{C,t}$ is a zero-mean random variable revealed at date t, and symmetrically distributed around zero over the finite support $[-\overline{\epsilon}_{C,t}, +\overline{\epsilon}_{C,t}]$. We assume that $\epsilon_{C,1}$, $\epsilon_{C,2}$ and $\epsilon_{A,2}$ are independently distributed. We define $p_{C,t}$ and $\phi_{C,t}$ as before.

Utility. C-investors have initial wealth $w_{C,0}$ at $t = 0$ and exponential utility over $w_{C,2}$, which is restricted to be non-negative, i.e., $-\exp(-\alpha_C w_{C,2})$ with $\alpha_C > 0$ and $w_{C,2} \geq 0$. At $t = 2$, each C-investor receives an endowment correlated with asset C's dividend

$$u_C \cdot (\epsilon_{C,1} + \epsilon_{C,2}). \tag{17}$$

Without loss of generality, we assume that $u_C > 0$, i.e., C-investors want to sell asset C. For instance, C-investors may be individuals selling mortgages to banks.

Arbitrageurs. Arbitrageurs have initial wealth W_0 at $t = 0$, and can trade asset C at $t = 0, 1$. However, they must ensure that the value of any position $x_{C,t}$ in asset C at date t is non-negative at $t+1$ (irrespective of positions in assets A and B) by adding enough riskfree asset as collateral:

$$-\min x_{C,t}(-p_{C,t} + p_{C,t+1}) = -\min x_{C,t}\left(\phi_{C,t} + \epsilon_{C,t+1}\right) = x_{C,t}\left(\overline{\epsilon}_{C,t+1} - \phi_{C,t}\right). \tag{18}$$

Equilibrium. Definitions 1 and 2 are extended in the obvious way.

3.2 Equilibrium at $t = 1$

We begin by characterizing the equilibrium at $t = 1$, which is similar to that without asset C.

3.2.1 Outside Investors

Since A- and B-investors cannot take positions in asset C, their portfolio choice problem is unaffected by its existence. Therefore, their first order condition remains (13).

C-investors' problem is similar. Indeed, their budget constraint is similar to A- and B-investors', except that some dividend uncertainty is resolved at $t = 1$:

$$
\begin{aligned}
w_{C,2} &= y_{C,1}p_{C,2} + (w_{C,1} - y_{C,1}p_{C,1}) + u_i(\epsilon_{C,1} + \epsilon_{C,2}) && (19)\\
&= y_{C,1}\left(\overline{\delta} + \epsilon_{C,1} + \epsilon_{C,2}\right) + \left(w_{C,1} - y_{C,1}\left(\overline{\delta} + \epsilon_{C,1} - \phi_{C,1}\right)\right) + u_C(\epsilon_{C,1} + \epsilon_{C,2}) && (20)\\
&= (w_{C,1} + u_C\epsilon_{C,1}) + y_{C,1}\phi_{C,1} + (y_{C,1} + u_C)\epsilon_{C,2}. && (21)
\end{aligned}
$$

The term $(w_{C,1} + u_C\epsilon_{C,1})$ would be the investor's wealth at $t = 2$ if there was no dividend risk at $t = 2$ ($\overline{\epsilon}_{C,2} = 0$). The two other terms, familiar from (13), represent expected return and risk.

Since $(w_{C,1} + u_C\epsilon_{C,1})$ is certain as of $t = 1$, C-investors' problem is formally identical to A- and B-investors' and Lemma 1 applies, i.e., a function $f_{C,1}$ positive and strictly convex exists such that each C-investor's problem at $t = 1$ can be written as:

$$\max_{y_{C,1}} \quad (w_{C,1} + u_C\epsilon_{C,1}) + y_{C,1}\phi_{C,1} - f_{C,1}(y_{C,1} + u_C)\,. \tag{22}$$

As a result, C-investors first-order condition can also be written as (13).

3.2.2 Arbitrageurs

As arbitrageurs can invest in all assets, their budget constraint between $t = 1$ and $t = 2$ is as before (condition (14)) but for the $t = 2$ return on their $t = 1$ position in asset C:

$$W_2 = W_1 + 2x_{A,1}\phi_{A,1} + x_{C,1}(\phi_{C,1} + \epsilon_{i,2}). \tag{23}$$

Unlike the AB spread trade, arbitrageurs' position in asset C involves dividend risk. For simplicity, we assume from now on that arbitrageurs and C-investors have the same risk-aversion coefficient, i.e., $\alpha = \alpha_C$. Applying Lemma 1, we can write an arbitrageur's objective at $t = 1$ as

$$\max_{x_{A,1},x_{C,1}} \quad W_1 + 2x_{A,1}\phi_{A,1} + x_{C,1}\phi_{C,1} - f_{C,2}(x_{C,1}). \tag{24}$$

The cost of bearing dividend risk is the same for arbitrageurs and C-investors because utility functions are assumed the same.

Each arbitrageur maximizes his objective subject to the financial constraint

$$2mx_{A,1} + x_{C,1}\left(\overline{\epsilon}_{C,2} - \phi_{C,1}\right) \leq W_1. \tag{25}$$

At $t = 1$, each arbitrageur must trade off investing in the AB spread trade vs. longing asset C as both require scarce collateral. This trade off is illustrated by substituting $x_{A,1}$ from the financial constraint (25) into the objective (24), which can be rewritten as[8]

$$W_1 \left(1 + \frac{\phi_{A,1}}{m}\right) + x_{C,1} \left[\phi_{C,1} - \frac{\left(\overline{\epsilon}_{C,2} - \phi_{C,1}\right) \phi_{A,1}}{m}\right] - f_{C,2}(x_{C,1}). \tag{26}$$

The derivative of this expression has the same sign as

$$\frac{\phi_{C,1} - f'_{C,2}(x_{C,1})}{\overline{\epsilon}_{C,2} - \phi_{C,1}} - \frac{\phi_{A,1}}{m}. \tag{27}$$

The first term is the return per unit of collateral of investing in asset C, the second term is that of investing in the AB spread trade. Indeed, the marginal unit position in asset C yields an expected return $\phi_{C,1}$ but forces the arbitrageur to bear some risk at a cost that depends on his current holdings of asset C, $f'_{C,2}(x_{C,1})$. It also requires $\left(\overline{\epsilon}_{C,2} - \phi_{C,1}\right)$ units of collateral. Similarly, the AB spread trade yields a safe return $\phi_{A,1}$ per m units of collateral.

As $x_{C,1}$ increases, so does $f'_{C,2}(x_{C,1})$ since the arbitrageur is increasingly reluctant to bear asset C's dividend risk. An interior solution obtains eventually if expression (27) equals zero, i.e., when the return on collateral is equalized across opportunities (Gromb and Vayanos (2008)).

3.2.3 Equilibrium

To make the problem interesting, we assume that in equilibrium arbitrageurs always find it optimal to invest in the AB spread trade. Using the i-investors' first order condition (13) and the market clearing conditions (7), expression (27) can be rewritten as

$$\frac{f'_{C,2}(u_C - \frac{x_{C,1}}{\mu_C}) - f'_{C,2}(x_{C,1})}{\overline{\epsilon}_{C,2} - f'_{C,2}(u_C - \frac{x_{C,1}}{\mu_C})} - \frac{f'_{A,2}(u_A - \frac{x_{A,1}}{\mu_A})}{m}. \tag{28}$$

To ensure that arbitrageurs invest in the spread trade, we assume that this condition never holds for $x_{A,1} = 0$. The LHS being maximum for $x_{C,1} = 0$, we make the following assumption.

Assumption 1 *We assume*

$$\frac{f'_{C,2}(u_C)}{\overline{\epsilon}_{C,2} - f'_{C,2}(u_C)} - \frac{f'_{A,2}(u_A)}{m} < 0. \tag{29}$$

This condition holds for m small enough. We can now characterize the equilibrium at $t = 1$.

Proposition 2 *Thresholds $W_1^* > W_1^{**} > 0$ exist such that the equilibrium at $t = 1$ is as follows.*

[8] Expression (26) is valid whether the financial constraint is binding or not because if the constraint is not binding then $\phi_{A,1} = 0$.

- *If $W_1 < W_1^*$, the arbitrageurs' financial constraint binds and markets are illiquid ($\phi_{A,1} > 0$). Market liquidity increases with arbitrage capital (i.e., $\partial\phi_{A,1}/\partial W_1 < 0$). Arbitrageurs share risk less than optimally with C-investors.*
  - *If $W \leq W_1^{**}$ arbitrageurs do not hold asset C (i.e., $x_{C,1} = 0$) and $\partial\phi_{C,1}/\partial W_1 = 0$.*
  - *If $W > W_1^{**}$ arbitrageurs invest $x_{C,1} < \mu_C u_C/(1+\mu_C)$ in asset C and $\partial\phi_{C,1}/\partial W_1 < 0$.*
- *If $W_1 \geq W_1^*$, the arbitrageurs' financial constraint is slack and market liquidity is perfect: The price wedge between assets A and B is closed ($\phi_{A,1} = 0$). Arbitrageurs share risk optimally with C-investors, i.e., invest $x_{C,1} = \mu_C u_C/(1 + \mu_C)$ in asset C.*

Since arbitrageurs face financial constraints, their wealth W_1 determines their capacity to invest in all assets at $t = 1$, i.e., their ability to provide liquidity in all markets. Hence, the financial constraint creates a linkage between all markets despite their fundamentals being uncorrelated (Brunnermeier and Pedersen (2008), Gromb and Vayanos (2008)).

Corollary 1 *Uncertainty over arbitrage capital W_1 creates correlation in market liquidity across markets, i.e., $Cov\left(\phi_{A,1}, \phi_{C,1}\right) > 0$.*

3.2.4 Effect of $\epsilon_{C,1}$

In our model, the source of uncertainty over W_1 is asset C's dividend risk $\epsilon_{C,1}$. If, for example, $\epsilon_{C,1}$ is low, asset C is expected to pay a small dividend at $t = 2$, and its price at $t = 1$ drops. The uncertainty over $\epsilon_{C,1}$ generates knock-on effects through arbitrageurs' wealth. Indeed, if arbitrageurs hold a long position in asset C at $t = 0$ ($x_{C,0} > 0$), a low $\epsilon_{C,1}$ reduces their wealth W_1. This reduces their ability to provide liquidity at $t = 1$. As a result, the risk premium $\phi_{C,1}$ of asset C increases, amplifying the effect of $\epsilon_{C,1}$. Moreover, the risk premium $\phi_{A,1}$ of asset A increases, and so does the price wedge between assets A and B. Note that assets A and B are effected by asset C even though their fundamentals are unrelated with C.

Consider the effect of $\epsilon_{C,1}$ on arbitrage capital W_1. Arbitrageurs' wealth at $t = 1$ is their initial wealth plus the capital gain in their position in asset C. Hence, the arbitrageurs' budget constraint between $t = 0$ and $t = 1$ is

$$W_1 = W_0 + x_{C,0}(p_{C,1} - p_{C,0}) = W_0 + x_{C,0}\left(\phi_{C,0} - \phi_{C,1} + \epsilon_{C,1}\right), \tag{30}$$

which implies

$$\frac{\partial W_1}{\partial \epsilon_{C,1}} = x_{C,0}\left(1 - \frac{\partial \phi_{C,1}}{\partial \epsilon_{C,1}}\right). \tag{31}$$

The first term is the direct effect that asset C's dividend news have on arbitrageur wealth. The second term is the indirect amplification effect: because arbitrageur wealth changes, the risk premium $\phi_{C,1}$ of asset C changes and this further changes arbitrageur wealth.

Now let z_1 be any of the variables at $t = 1$, i.e., $x_{i,1}$ and $\phi_{i,1}$ for $i = A, B, C$. We have

$$\frac{\partial z_1}{\partial \epsilon_{C,1}} = \left(\frac{\partial W_1}{\partial \epsilon_{C,1}}\right)\left(\frac{\partial z_1}{\partial W_1}\right) = x_{C,0}\left(1 - \frac{\partial \phi_{C,1}}{\partial \epsilon_{C,1}}\right)\left(\frac{\partial z_1}{\partial W_1}\right). \tag{32}$$

Setting $z_1 = \phi_{C,1}$ in this expression, we find

$$\frac{\partial \phi_{C,1}}{\partial \epsilon_{C,1}} = \frac{x_{C,0}\left(\frac{\partial \phi_{C,1}}{\partial W_1}\right)}{1 + x_{C,0}\frac{\partial \phi_{C,1}}{\partial W_1}}. \tag{33}$$

Therefore expression (32) can be rewritten as

$$\frac{\partial z_1}{\partial \epsilon_{C,1}} = \frac{x_{C,0}}{1 + x_{C,0}\frac{\partial \phi_{C,1}}{\partial W_1}}\left(\frac{\partial z_1}{\partial W_1}\right). \tag{34}$$

This expression illustrates how the effect of $\epsilon_{C,1}$ varies with $x_{C,0}$, the arbitrageurs' position in asset C. Note that $\partial z_1/\partial W_1$ is independent of $x_{C,0}$, and that $\frac{\partial \phi_{C,1}}{\partial W_1} < 0$. Hence all else equal, the effect of $\epsilon_{C,1}$ on any variable increases with $x_{C,0}$. The direct effect of $x_{C,0}$ is through the numerator, while its indirect effect is through the denominator.

3.3 Equilibrium at $t = 0$

At $t = 0$, only asset C and the riskless asset are traded, and only C-investors and arbitrageurs can trade them.

3.3.1 C-investors

The C-investors' budget constraint between $t = 0$ and $t = 1$ is

$$w_{C,1} = w_{C,0} + y_{C,0}(p_{C,1} - p_{C,0}) = w_{C,0} + y_{C,0}\left(\phi_{C,0} - \phi_{C,1} + \epsilon_{C,1}\right). \tag{35}$$

Indeed, C-investors' wealth at $t = 1$ is their initial wealth plus the capital gain in their position in asset C. Substituting into (21), we can write the C-investors' objective at $t = 0$ as

$$-E_0 \exp(-\alpha M_C), \tag{36}$$

where

$$M_C \equiv w_{C,0} + y_{C,0}\left(\phi_{C,0} - \phi_{C,1} + \epsilon_{C,1}\right) + u_C\epsilon_{C,1} + y_{C,1}\phi_{C,1} - f_{C,2}(y_{C,1} + u_C). \tag{37}$$

The first-order condition is

$$E_0\left[\exp(-\alpha M_C)\left(\phi_{C,0} - \phi_{C,1} + \epsilon_{C,1}\right)\right] = 0. \tag{38}$$

The term $\left(\phi_{C,0} - \phi_{C,1} + \epsilon_{C,1}\right)$ is the capital gain between $t = 0$ and $t = 1$. Furthermore, $\exp(-\alpha M_C)$ is the C-investors' expected marginal utility of wealth at $t = 2$, conditional on $\epsilon_{C,1}$. The optimality condition consists in setting the expectation with respect to $\epsilon_{C,1}$ of the product of these two terms to zero.

3.3.2 Arbitrageurs

The arbitrageurs' budget constraint between $t = 0$ and $t = 1$ is condition (30). Substituting into (26), we can write the arbitrageurs' objective at $t = 0$ as

$$-E\exp(-\alpha M), \tag{39}$$

where

$$M \equiv \left[W_0 + x_{C,0}\left(\phi_{C,0} - \phi_{C,1} + \epsilon_{C,1}\right)\right]\left(1 + \frac{\phi_{A,1}}{m}\right) + x_{C,1}\left[\phi_{C,1} - \frac{\left(\bar{\epsilon}_{C,2} - \phi_{C,1}\right)\phi_{A,1}}{m}\right] - f_{C,2}(x_{C,1}). \tag{40}$$

The first-order condition is

$$E\left[\exp(-\alpha M)\left(1 + \frac{\phi_{A,1}}{m}\right)\left(\phi_{C,0} - \phi_{C,1} + \epsilon_{C,1}\right)\right] = 0. \tag{41}$$

The difference with the first-order condition of C-investors is in the term $(1 + \phi_{A,1}/m)$. This term reflects the superior return that arbitrageurs can achieve by investing in the AB arbitrage opportunity. Arbitrageurs' marginal utility of wealth at $t = 1$ is the product of the return $(1 + \phi_{A,1}/m)$, times their expected marginal utility $\exp(-\alpha M)$ of wealth at $t = 2$. For C-investors, the riskless return between $t = 1$ and $t = 2$ is one, and the marginal utility of wealth at $t = 1$ coincides with the expected marginal utility $(-\alpha M_C)$ of wealth at $t = 2$. Notice that the arbitrageurs' return $(1 + \phi_{A,1}/m)$ is higher when they lose money at $t = 1$ because the arbitrage opportunity then widens. This can induce arbitrageurs to invest cautiously as to limit their losses at $t = 1$ and better take advantage of the high returns between $t = 1$ and $t = 2$.

3.3.3 Equilibrium

The variables $(x_{C,0}, y_{C,0}, \phi_{C,0})$ are the solution to (7), (38) and (41).

4 Welfare

We now turn to the welfare analysis. Understanding the welfare implications of investors' financial constraints is important as they underlie many policy debates. An example is the debate on systemic risk, i.e., whether a worsening of the financial condition of some market participants can propagate into the financial system with harmful effects. One important issue in this debate is whether market participants take an appropriate level of risk, given that their potential losses can affect others. Our model provides a framework for studying this question.

Since we evaluate ex-ante decisions, we consider agents' expected utilities at $t = 0$. We convert expected utilities into certainty equivalents, i.e., the monetary value that would leave agents indifferent between participating in and staying out of the market. The $t = 0$ certainty equivalent of arbitrageurs is defined by

$$CEQ_{arb} \equiv -\frac{1}{\alpha}\log E_0 \exp(-\alpha M). \tag{42}$$

The $t = 0$ certainty equivalent of i-investors, $i = A, B, C$, is defined by

$$CEQ_i \equiv -\frac{1}{\alpha_i} \log E_0 \exp(-\alpha_i M_i), \tag{43}$$

where for $i = A, B$ we set

$$M_i \equiv w_{i,1} + y_{i,1}\phi_{i,1} - f_{i,2}(y_{i,1} + u_i). \tag{44}$$

We define welfare as the sum of the certainty equivalents of each agent group, weighted by the group's measure, i.e.,

$$\mathcal{W} \equiv \sum_{i=A,B,C} \mu_i CEQ_i + CEQ_{arb}. \tag{45}$$

We consider only total welfare and not the utility of each agent group separately.[9] If a change in $x_{C,0}$ raises total welfare, there exists a set of side-transfers that raises the certainty equivalent of each agent group. These side-transfers have no effect on total welfare if they are deterministic and take place at $t = 2$. Indeed, deterministic transfers do not affect risksharing among agents, and because of exponential utility they do not affect agents' risk attitudes. Moreover, if the transfers take place at $t = 2$, they have no effect on the financial constraint at $t = 1$.

We start our analysis with a benchmark result. The presence of arbitrageurs raises welfare, while the financial constraint imposes a welfare cost.

Proposition 3 *In equilibrium, total welfare is larger than if arbitrageurs were not allowed to trade ($x_{i,t} = 0$ for $i = A, B, C$ and $t = 0, 1$) and smaller than if they faced no financial constraint ($m = 0$).*

The result that welfare is larger in the presence of arbitrageurs does not necessarily imply that arbitrageurs take an appropriate level of risk. To examine this issue, we consider the following thought experiment. Suppose that a social planner changes the arbitrageurs' position in asset C at $t = 0$ away from its equilibrium value $x_{C,0}$. The social planner affects only that position, and lets the market determine all other positions and prices. In particular, the social planner faces the same financial constraint as arbitrageurs at $t = 1$. Imposing the same constraint on the social planner as on individual agents is the relevant benchmark for financial regulation: it amounts to assuming that regulators are subject to the same informational frictions as individual agents.

To implement the thought experiment formally, we treat $x_{C,0}$, the arbitrageurs' position in asset C at $t = 0$, as an exogenous parameter. For each value of $x_{C,0}$, we define an "$x_{C,0}$ equilibrium" by requiring that agents' positions, except for $x_{C,0}$, are optimal given prices, markets clear, and financial constraints are met. We evaluate welfare in this $x_{C,0}$ equilibrium, and compute the derivative at the value of $x_{C,0}$ that corresponds to the original equilibrium.

[9] Gromb and Vayanos (2002) consider distributional effects and show the possibility of Pareto improvements in a setting where arbitrageurs can only trade with A- and B-investors.

We evaluate the change in welfare in the simple case where there are only two states of nature at $t = 1$, i.e., $\epsilon_{C,1}$ takes only two values. In this case, risksharing between arbitrageurs and C-investors at $t = 0$ is optimal because markets are complete from these agents' viewpoint. (Arbitrageurs and C-investors can equate their marginal rates of substitution across the two states by trading asset C and the riskfree asset.) Allowing more than two states would expand the range of possible market inefficiencies.

We refer to the state where $\epsilon_{C,1}$ is low as the bad state. In that state, arbitrageur wealth is low, the risk premia $\phi_{A,1}$ and $\phi_{C,1}$ are high, and market liquidity is low. Conversely, in the good state where $\epsilon_{C,1}$ is high, arbitrageurs are wealthy and liquidity is high.

Proposition 4 *Suppose that $\epsilon_{C,1}$ takes only two values. A change in $x_{C,0}$ away from its equilibrium value changes total welfare by*

$$
\begin{aligned}
\frac{\partial \mathcal{W}}{\partial x_{C,0}} =& \frac{E\left\{\left[\frac{\partial(\phi_{C,0}-\phi_{C,1})}{\partial x_{C,0}} x_{C,0} + \frac{\partial \phi_{C,1}}{\partial x_{C,0}} x_{C,1}\right] \exp(-\alpha M)\left(1+\frac{\phi_{A,1}}{m}\right)\right\} E\left[\exp(-\alpha M)\frac{\phi_{A,1}}{m}\right]}{E\left[\exp(-\alpha M)\left(1+\frac{\phi_{A,1}}{m}\right)\right] E\exp(-\alpha M)} \\
&+ 2E\left\{\frac{\partial \phi_{A,1}}{\partial x_{C,0}} x_1 \left[\frac{\exp(-\alpha M)}{E\exp(-\alpha M)} - \frac{\exp(-\alpha_A M_A)}{E\exp(-\alpha_A M_A)}\right]\right\}. \qquad (46)
\end{aligned}
$$

A change in $x_{C,0}$ affects equilibrium prices and quantities. Changes in quantities have only second-order effects on welfare because quantities are chosen optimally given prices. Changes in prices, however, can have first-order effects. When prices change, wealth is redistributed among agents. Such redistributions affect the utility of each agent, but more surprisingly can also affect total welfare. This is because of financial constraints and incomplete markets.

A first type of redistribution is between C-investors and arbitrageurs. This redistribution occurs through the changes in the risk premia $\phi_{C,0}$ and $\phi_{C,1}$ of asset C. Its effect on welfare corresponds to the first term in expression (46). Total welfare changes because arbitrageurs have better investment opportunities than C-investors. Consider, for example, a decrease in $x_{C,0}$, i.e., arbitrageurs buy a smaller quantity of asset C, thus offering less insurance to C-investors at $t = 0$. Because C-investors are more eager to sell asset C, the asset's expected capital gain $E_0\left(\phi_{C,0} - \phi_{C,1} + \epsilon_{C,1}\right)$ between $t = 0$ and $t = 1$ increases. This makes the arbitrageurs better off at the expense of C-investors. Effectively, the arbitrageurs move prices to their advantage by restricting quantity. The transfer to arbitrageurs raises welfare because arbitrageurs' return on wealth at $t = 1$ exceeds that of C-investors ($\phi_{A,1} \geq 0$).

A second type of redistribution is between A- and B-investors on one side and arbitrageurs on the other. This redistribution occurs through the change in the risk premium $\phi_{A,1}$ of asset A (and $\phi_{B,1}$ of asset B). Its effect on welfare corresponds to the second term in expression (46). Total welfare changes because the redistribution is stochastic, and because the inability of A- and B-investors to access the market at $t = 0$ leaves unexploited insurance opportunities between them and arbitrageurs. Consider, for example, a decrease in $x_{C,0}$. Since arbitrageurs take less risk at $t = 0$, they realize smaller losses in the bad state where $\epsilon_{C,1}$ is low. Therefore, market liquidity does not drop as much in the bad state, i.e., the increase in $\phi_{A,1}$ is smaller. This makes A- and B-investors better off in the bad state at the expense of arbitrageurs (with a reverse transfer taking place in the good state). Total welfare increases if A- and B-investors suffer more than arbitrageurs in the bad state, i.e., if their marginal rate of substitution between the bad and the good state is higher than arbitrageurs'. This is the case, for example, if the risk aversion coefficient of A- and B-investors is large relative to that of arbitrageurs.

While Proposition 4 highlights the channels through which changes in $x_{C,0}$ affect total welfare, it does not determine conditions under which total welfare increases. To establish such conditions, we can proceed analytically or numerically. On the analytical front, we can solve for equilibrium in closed form when the uncertainty between $t = 0$ and $t = 1$ is small ($\epsilon_{C,1}$ small). We can also solve for large uncertainty numerically, and examine how the change in welfare depends on exogenous parameters. We leave this for the next revision of this paper, and use a simple numerical example to illustrate the basic effects.

We set the risk aversion coefficient of all agents to 0.4, i.e., $\alpha = \alpha_A = \alpha_C = 0.4$. We assume that all shocks are drawn from a symmetric two-point distribution with support $\{1, -1\}$ for $\epsilon_{i,2}$, $i = A, B, C$, and $\{1.2, -1.2\}$ for $\epsilon_{C,1}$. We set all supply shocks and agents' measures to one, i.e., $u_i = \mu_i = 1$ for $i = A, B, C$. Finally, we set the parameter m characterizing the tightness of the financial constraint to 0.2, and the arbitrageurs' initial wealth W_0 to 0.5.

Tables 1 and **??** show the equilibrium prices and quantities. Quantities are reported in terms of arbitrageur positions and prices are reported in terms of risk premia.

Parameters	Notation	Value
Asset C, $t = 0$	$x_{C,0}$	0.26
Asset C, $t = 1$, good state	$x_{C,1}$	0.50
Asset C, $t = 1$, bad state	$x_{C,1}$	0.00
Asset A, $t = 1$, good state	$x_{A,1}$	1.00
Asset A, $t = 1$, bad state	$x_{A,1}$	0.69

Table 1: Arbitrageur Positions.

Table 1 shows that arbitrageurs increase their position in asset C in the good state ($x_{C,1} > x_{C,0}$), while reducing it in the bad state. Moreover, they absorb the entire supply shock of A- and B-investors in the good state, and only 69% in the bad state. Consistent with these findings, Table **??** shows that the risk premium of asset A is zero (perfect liquidity) in the good state, and positive in the bad state. Moreover, the risk premium of asset C is larger in the bad than in the good state.

We next reduce $x_{C,0}$ below its equilibrium value of 0.26 and recalculate quantities and prices. We find that for small reductions in $x_{C,0}$, total welfare increases. A- and B-investors become better off because arbitrageurs absorb a larger fraction of their supply shock in the bad state.

Parameters	Notation	Value
Asset C, $t = 0$	$\phi_{C,0}$	0.73
Asset C, $t = 1$, good state	$\phi_{C,1}$	0.20
Asset C, $t = 1$, bad state	$\phi_{C,1}$	0.38
Asset A, $t = 1$, good state	$\phi_{A,1}$	0.00
Asset A, $t = 1$, bad state	$\phi_{A,1}$	0.12

Table 2: Risk Premia.

C-investors become worse off because they receive less insurance from arbitrageurs at $t = 0$. Arbitrageurs benefit from the depressed price of asset C at $t = 0$ but lose because the price wedge between assets A and B decreases. Overall, their utility increases, as does total welfare.

5 Public Policy

We build on our analysis to discuss some policy options. At this stage, these are preliminary. Their main value, if any, is to illustrate how a careful welfare analysis of market liquidity provision can inform policy debates. Therefore, we also point at directions for future research.

Central to our welfare analysis is the result that competitive arbitrageurs do not generally follow a socially optimal risk-management policy. This, again, arises from their failing to internalize the price effects of their investment decisions. Policy responses must therefore curb arbitrageurs' risk-management policy, directly or indirectly.

5.1 Tightening or Loosening Constraints

In our model, social welfare would be maximized if arbitrageurs faced no financial constraints, i.e. for $m = 0$. Eliminating these constraints may therefore constitute an obvious policy goal. Suppose however that this route is not possible or involves prohibitively high costs. Can instead tightening constraints improve welfare? We conjecture that this can indeed be the case.

Conjecture 1 *Policies constraining the arbitrageurs' investment in asset C can be welfare-improving.*

Suppose that in equilibrium, arbitrageurs overinvest in asset C at $t = 0$ relative to the social optimum. Consider now the possibility for a regulator to increase the arbitrageurs' cost of investing in asset C. For now, assume simply that positions in asset C are taxed. Below, we discuss alternative implementations. For simplicity, assume for now that the policy must be the same at both $t = 0$ and $t = 1$. Such a policy would reduce the arbitrageurs' position in asset C at $t = 0$, moving it towards the social optimum. As a result, arbitrage capital would be more abundant in bad states (when $\epsilon_{C,1}$ is small), when it is needed most. This would surely increase the arbitrageurs' ability to provide liquidity to A- and B-investors in bad states, making these investors better off. As for C-investors, two countervailing effects would affect their welfare at $t = 1$. On the one hand, the increased cost of investing in asset C at $t = 1$ reduces the arbitrageurs' ability to provide risk-sharing per unit of arbitrage capital. On the other hand, arbitrage capital is more abundant in the bad state, when risk-sharing is more valuable. The direction of the net effect remains to be investigated. However, even if the net effect were a cost for C-investors, it might still be outweighed by the benefit to A- and B-investors and possibly to the arbitrageurs.

What form would such a policy take? One possibility would be to tax investment in asset C. Another, perhaps more realistic one, would be to set capital requirements for arbitrageurs investing in asset C. This might also apply to their financiers.

5.2 Arbitrage Capital Transfers

Another related policy option might be to affect directly the arbitrageurs' risk-management policy by forcing them to transfer capital to certain states of the world.

Conjecture 2 *Welfare-improving balanced budget policies reducing arbitrage capital at $t = 0$ and increasing it at $t = 1$ are possible.*

Suppose that in equilibrium, arbitrageurs overinvest in asset C at $t = 0$ relative to the social optimum. Consider now the possibility for a social planner to tax some of the arbitrageurs' wealth at $t = 0$, and release it back to them at $t = 1$. Such a policy would, in effect, force arbitrageurs' risk-management policy towards the socially optimum.

Note that the anticipation of policies tightening or loosening constraints would affect the investment policy of arbitrageurs and of their suppliers of capital. For instance, an anticipated loosening of constraints might lead to moral hazard, i.e., arbitrage capital being over-utilized. These questions call for a full-fledged dynamic analysis.

5.3 Industrial Organization

One policy route might be to increase the amount of arbitrage capital available by facilitating entry in the liquidity provision industry. This might be achieved by lowering entry barriers and/or relaxing regulation. Instead, suppose now that the amount of arbitrage capital is fixed, at least in the short run, and consider the effect of reducing competition among arbitrageurs.

Conjecture 3 *Policies reducing the intensity of competition between arbitrageurs can be welfare-improving.*

As we have emphasized, our welfare results arise from the fact that arbitrageurs fail to internalize the price effects of their investment decisions. This is because we assumed them to be perfectly competitive. Imperfect competition among arbitrageurs would lead them to internalize some of the price effects, possibly leading them to adopt investment policies closer to the social optimum. This positive effect would however have to be weighted against the adverse welfare effects of market power. Indeed, imperfectly competitive arbitrageurs may be tempted to ration liquidity to increase its price.

The industrial organization of arbitrage and market liquidity provision is a fascinating and under-researched question (Duffie and Strulovici (2008)). For instance, one might ask whether certain industry structures are more likely to emerge in particular markets, e.g., whether certain types of arbitrage will be conducted by a few large institutions, and others by many boutique-type firms. The next step would be to consider whether and when the emerging industrial structure is optimal, and when it ought to be corrected or regulated.

5.4 Remarks

We conclude this section with a few remarks. First, note that much of the regulation of financial institutions is concerned with default risk. In our model, however, there is no default (in equilibrium and out of equilibrium). Here, public policy would have the broader objective of ensuring that arbitrageurs have capital when it matters most for society (see Cerasi and Rochet (2008)). Second, the discussion suggests it is important to be able to distinguish situations in which arbitrageurs might over-invest rather from those in which they might under-invest, as these are likely to call for different policy responses. Both types of situations are likely to arise in reality. At times, arbitrageurs may be overleveraged (as hedge funds during the 1998 crisis). At other times, the providers of arbitrage capital may withdraw funds and hoard liquidity, leading to liquidity drying up in financial markets (as during the recent crisis).

6 Conclusion

We model financial market liquidity as provided by financially constrained arbitrageurs. Market liquidity is shown to depend positively on the level of arbitrage capital defined as the capital, internal or external, arbitrageurs can access frictionlessly. In particular, liquidity dry-ups correspond to situations in which arbitrage capital is low relative to the demand for liquidity. Therefore, liquidity dry-ups should follow periods of low returns for arbitrageurs and/or their financiers. In this context, we conduct a welfare analysis of the dynamic allocation of scarce arbitrage capital across different opportunities. We view such a welfare analysis as a prerequisite to any meaningful discussion of public policy in relation to market liquidity. We show that due to price effects, arbitrage capital may not be allocated in a socially optimal fashion. We use the insight from our welfare analysis to discuss a number of possible policy levers.

Our analysis of policy remains preliminary in several respects and much work remains to be done. On the theory front, an important agenda is to move from exogenous to endogenous financial constraints. Also, our model assumes perfectly competitive arbitrageurs. While this is a useful and standard benchmark, episodes in which some individual players can have a price impact have been documented, particularly during crisis situations. The LTCM debacle is an example in which a single hedge fund's liquidations moved prices. In other cases, a few liquidity providers may command substantial market power (Acharya, Gromb and Yorulmazer (2008)). More generally, the Industrial Organization of liquidity provision, its implications and its potential regulation seem important topics. On the applications front, our model ought to be extended to incorporate and, possibly, evaluate the instruments and policies available to parties concerned with the liquidity of financial markets, be they regulators, central banks or financial markets themselves. In particular, it might incorporate a role for monetary policy and distinguish its use in financial crises and during bubbles. We hope our framework can clarify some ideas in this exciting area of future research.

References

Acharya, Viral, Denis Gromb and Tanju Yorulmazer (2008), "Imperfect Competition in the Interbank Market for Liquidity as a Rationale for Central Banking," working paper, London Business School, London School of Economics, and Federal Reserve Bank of New York.

Anshuman, Ravi, and V. Viswanathan (2007), "Collateral and Market Liquidity," working paper, Duke University.

Basak, Suleyman, and Benjamin Croitoru (2000), "Equilibrium Mispricing in a Capital Market with Portfolio Constraints," *Review of Financial Studies*, 13, 715-748.

Basak, Suleyman, and Benjamin Croitoru (2006), "On the Role of Arbitrageurs in Rational Markets," *Journal of Financial Economics*, 81, 143-173.

Brunnermeier, Markus, and Lasse Pedersen (2007), "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, forthcoming.

Caballero, Ricardo J., and Arvind Krishnamurthy (2008), "Collective Risk Management in a Flight to Quality Episode," *Journal of Finance*, forthcoming.

Cerasi, Vittoria, and Jean-Charles Rochet (2008), "Solvency Regulation and Credit Risk Transfer," working paper Bicocca University and Toulouse University.

Duffie, Darrell, and Bruno Strulovici (2008), "Capital Mobility and Asset Pricing," working paper Stanford University and Oxford University.

Ewerhart, Christian, and Natacha Valla (2007), "Financial Market Liquidity and the Lender of Last Resort," working paper, Banque de France.

Fishman, Tal, Harrison Hong, and Jeffrey D. Shubik (2007), "Do Arbitrageurs Amplify Economic Shocks?" working paper, Princeton University and Syracuse University.

Geanakoplos, John (2003), "Liquidity, Default and Crashes: Endogenous Contracts in General Equilibrium," *Advances in Economics and Econometrics: Theory and Applications, Eighth World Conference, Volume II,* Econometric Society Monographs, 170-205.

Goldstein, Itay, and Ady Pauzner (2004), "Contagion of Self-Fulfilling Financial Crises Due to Diversification of Investment Portfolios," *Journal of Economic Theory*, 119, 151-183.

Gorton, G., Ping He, and Lixin Huang (2006), "Asset Prices When Agents Are Marked-to-Market," working paper, University of Pennsylvania.

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 (2008), "Financially Constrained Arbitrage and the Cross-Section of Market Liquidity," working paper, London Business School and London School of Economics.

He, Zhiguo, and Arvind Krishnamurthy (2006), "Intermediation, Capital Immobility, and Asset Prices," working paper, Northwestern University.

He, Zhiguo, and Arvind Krishnamurthy (2008), "Intermediary Asset Pricing," working paper Northwestern University.

Hendershott, Terrence, Pamela C. Moulton, and Mark S. Seasholes (2007), "Market Maker Inventories and Liquidity," working paper, University of California at Berkeley.

Ko, Jeremy (2007), "Breakup of Relative-Value Trades," working paper, Penn State University.

Kondor, Péter (2007), "Career Concerns and Dynamic Arbitrage," working paper, University of Chicago.

Kondor, Péter (2008), "Risk in Dynamic Arbitrage: Price Effects of Convergence Trading," *Journal of Finance*, forthcoming.

Krishnamurthy, Arvind, and Zhiguo He (2007), "Intermediation, Capital Immobility, and Asset Prices," working paper, Northwestern University.

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

Liu, Jun, and Francis Longstaff (2004), "Losing Money on Arbitrages: Optimal Dynamic Portfolio Choice in Markets with Arbitrage Opportunities," *Review of Financial Studies*, 17, 611-641.

Lustig, Hanno, and Stijn van Nieuwerburgh (2006), "Can Housing Explain Long-Run Swings in Asset Returns," working paper, NYU.

O'Hara, Maureen (1995), Market Microstructure Theory, Blackwell Publishers, Cambridge, MA.

Pavlova, Anna, and Roberto Rigobon (2008), "The Role of Portfolio Constraints in the International Propagation of Shocks." *Review of Economic Studies*, forthcoming.

Shleifer, Andrei, and Robert W. Vishny (1997), "The Limits of Arbitrage," *Journal of Finance*, 52, 35-55.

van Nieuwerburgh, Stijn, and Laura Veldkamp (2008), "Information Acquisition and Under-Diversification," working paper, NYU.

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

Appendix

Proof of Lemma 1: Define the function $f_{i,1}$ by $E_1\left[\exp(-\alpha_i y \epsilon_{i,2})\right] \equiv \exp\left[\alpha_i f_{i,1}(y)\right]$. Some properties of $f_{i,1}$ will prove useful. First note that $f_{A,0} = f_{B,0} = 0$ and $f_{A,1} = f_{B,1}$. We can also prove the following.

Lemma 2 *$f_{i,1}$ is positive and strictly convex. Moreover, it satisfies $f_{i,1}(y) = f_{i,1}(-y)$, $f'_{i,1}(y) \in (-1,1)$ and $\lim_{y\to\infty} f'_{i,1}(y) = 1$.*

To show that $f_{i,1}$ is positive, we use Jensen's inequality, which is strict since $\epsilon_{i,2}$ is stochastic, and the fact that $E\left(\epsilon_{i,2}\right) = 0$:

$$\exp(\alpha_i f_{i,1}(y)) = E\exp(-\alpha_i y\epsilon_{i,2}) > \exp\left[E(-\alpha_i y \epsilon_{i,2})\right] = \exp\left[-\alpha_i y E(\epsilon_{i,2})\right] = 1. \tag{47}$$

To show that $f_{i,1}$ is strictly convex, we compute its second derivative. We have

$$f_{i,1}(y) = \frac{1}{\alpha_i}\log\left[E\exp(-\alpha_i y \epsilon_{i,2})\right]. \tag{48}$$

Therefore,

$$f'_{i,1}(y) = -\frac{E\left(\epsilon_{i,2}\exp(-\alpha_i y\epsilon_{i,2})\right)}{E\exp(-\alpha_i y \epsilon_{i,2})}, \tag{49}$$

and

$$f''_{i,1}(y) = \alpha_i \frac{E\left(\epsilon_{i,2}^2\exp(-\alpha_i y\epsilon_{i,2})\right)E\exp(-\alpha_i y\epsilon_{i,2}) - \left[E\left(\epsilon_{i,2}\exp(-\alpha_i y \epsilon_{i,2})\right)\right]^2}{\left[E\exp(-\alpha_i y\epsilon_{i,2})\right]^2}. \tag{50}$$

That $f''_{i,1}(y) > 0$ follows from the Cauchy-Schwarz inequality

$$E(GH)^2 \le E(G^2)E(H^2), \tag{51}$$

for the functions $G = \epsilon_{i,2}\exp(-\alpha_i y \epsilon_{i,2}/2)$ and $H = \exp(-\alpha_i y\epsilon_{i,2}/2)$. The Cauchy-Schwarz inequality is strict since $\epsilon_{i,2}$ is stochastic, and thus G and H are not proportional.

To show that $f_{i,1}(y) = f_{i,1}(-y)$, we use the symmetry of $\epsilon_{i,2}$'s probability distribution around zero:

$$\exp(\alpha_i f_{i,1}(y)) = E\exp(-\alpha_i y\epsilon_{i,2}) = E\exp(\alpha_i y \epsilon_{i,2}) = \exp(\alpha_i f_{i,1}(-y)). \tag{52}$$

Finally, to show that $\lim_{y\to\infty} f'_{i,1}(y) = 1$, we note that

$$|f'_{i,1}(y) - 1| = \left| - \frac{E\left[\epsilon_{i,2}\exp(-\alpha_i y \epsilon_{i,2})\right]}{E\exp(-\alpha_i y \epsilon_{i,2})} - 1 \right| = \left| \frac{E\left[(\epsilon_{i,2}+1)\exp(-\alpha_i y \epsilon_{i,2})\right]}{E\exp(-\alpha_i y \epsilon_{i,2})} \right|. \tag{53}$$

To show that the last term goes to zero when y goes to ∞, we fix $\eta > 0$. We have

$$\left| \frac{E\left[(\epsilon_{i,2}+1)\exp(-\alpha_i y \epsilon_{i,2})\mathbf{1}_{\epsilon_{i,2}\le -1+\eta}\right]}{E\exp(-\alpha_i y \epsilon_{i,2})} \right| \le \eta \left| \frac{E\left[\exp(-\alpha_i y \epsilon_{i,2})\mathbf{1}_{\epsilon_{i,2}\le -1+\eta}\right]}{E\exp(-\alpha_i y \epsilon_{i,2})} \right| \le \eta. \tag{54}$$

Moreover, for y large enough,

$$\left| \frac{E\left[(\epsilon_{i,2}+1)\exp(-\alpha_i y \epsilon_{i,2})\mathbf{1}_{\epsilon_{i,2}> -1+\eta}\right]}{E\exp(-\alpha_i y \epsilon_{i,2})} \right| \le \eta. \tag{55}$$

This completes the technical lemma's proof. With this in mind, the gain $\phi_{i,1}$ being deterministic, we have:

$$E_1\left[-\alpha_i w_{i,2}\right] = E_1 \exp\left[-\alpha_i \left(\left(w_{i,1} + u_i\left(\overline{\delta} + \epsilon_{i,1}\right)\right) + y_{i,1}\phi_{i,1} - f_{i,1}\left(y_{i,1}+u_i\right)\right)\right]. \tag{56}$$

Since $w_{i,1}$ is a constant as of $t = 1$, an i-investor's problem is therefore as in expression (12). ■

Proof of Propositions 1 and 2: Let us prove the more general Proposition 2. (Proposition 1 obtains as a special case of Proposition 2 when asset C is irrelevant to arbitrageurs, i.e., when $\overline{\epsilon}_{C,2} = 0$ which implies $f_{C,2} = 0$.) Define the thresholds

$$W_1^* \equiv 2m\mu_A u_A + \frac{\mu_C u_C}{1+\mu_C}\left[\overline{\epsilon}_{C,2} - f'_{C,2}\left(\frac{\mu_C u_C}{1+\mu_C}\right)\right] \tag{57}$$

and

$$W_1^{**} \equiv 2m\left(\left(f'_{A,2}\right)^{-1}\left(\frac{f'_{C,2}(u_C)}{\overline{\epsilon}_{C,2} - f'_{C,2}(u_C)}\right) - u_A\right). \tag{58}$$

The values $y_{A,1} = x_{A,1}/\mu_A = u_A$ and $\phi_{A,1} = 0$ satisfy the market clearing condition (7), and A- and B-investors' FOCs (13). The corresponding conditions for C-investors are satisfied for $y_{C,1} = x_{C,1}/\mu_C = u_C/(1+\mu_C)$ and $\phi_{C,1} = f'_{C,2}[\mu_C u_C/(1+\mu_C)]$. These values also satisfy the arbitrageurs' financial constraint (25) iff $W_1 \ge W_1^*$, in which case they form an equilibrium. If instead $W_1 < W_1^*$, the financial constraint (25) binds. As long as expression (28) is negative, which by Assumption 1 it is for $x_{A,1} = x_{C,1} = 0$, arbitrageurs do not hold asset C. In that case, the financial constraint (25) being binding, we have $x_{A,1} = W_1/2m$. W_1^{**} is defined by expression (28) being equal to zero for $x_{A,1} = W_1^{**}/2m$ and $x_{C,1} = 0$. Hence, for $W_1 \le W_1^{**}$, $x_{A,1} = W_1/2m$ and $\phi_{A,1} = f'_{A,2}\left(u_A + W_1/2m\mu_A\right)$ which decreases with W_1. At the same time, $x_{C,1} = 0$ and $\phi_{C,1} = f'_{C,2}\left(u_C\right)$. If instead, $W_1 \in (W_1^{**}, W_1^*)$, expression (28) equals zero. As W_1 increases, so do $x_{A,1}$ and $x_{C,1}$ so that both $\phi_{A,1}$ and $\phi_{C,1}$ decrease. ■

Proof of Proposition 4: When there are only two states at $t = 1$, arbitrageurs and C-investors equate their marginal rates of substitution. The marginal value of \$1 at $t = 1$ is

$\exp(-\alpha M)(1+\phi_1/m)$ for arbitrageurs and $\exp(-\alpha M_C)$ for C-investors. Since the ratios of these marginal values are equal across states, we can write the change in social welfare as

$$\frac{\partial \mathcal{W}}{\partial x_{C,0}} = \frac{\partial \phi_{C,0}}{\partial x_{C,0}} \frac{x_{C,0} E\left[\exp(-\alpha M)\frac{\phi_1}{m}\right]}{E\exp(-\alpha M)} + 2E\left\{\frac{\partial \phi_1}{\partial x_{C,0}} x_1 \left[\frac{\exp(-\alpha M)}{E\exp(-\alpha M)} - \frac{\exp(-\alpha M_a)}{E\exp(-\alpha M_a)}\right]\right\}$$
$$\frac{E\left[\frac{\partial \phi_{C,1}}{\partial x_{C,0}}(x_{C,1}-x_{C,0})\exp(-\alpha M)\left(1+\frac{\phi_1}{m}\right)\right] E\left[\exp(-\alpha M)\frac{\phi_1}{m}\right]}{E\left[\exp(-\alpha M)\left(1+\frac{\phi_1}{m}\right)\right] E\exp(-\alpha M)}. \tag{59}$$

The third term in (46) is analogous to the first term. The first term corresponds to the Period 0 cash flow $x_{C,0}\partial\phi_{C,0}/\partial x_{C,0}$. The third term corresponds to the stochastic cash flow at $t=1$ $(x_{C,1}-x_{C,0})\partial\phi_{C,1}/\partial x_{C,0}$ that is equivalent at $t=0$ to

$$\frac{E\left[\frac{\partial \phi_{C,1}}{\partial x_{C,0}}(x_{C,1}-x_{C,0})\exp(-\alpha M)\left(1+\frac{\phi_1}{m}\right)\right]}{E\left[\exp(-\alpha M)\left(1+\frac{\phi_1}{m}\right)\right]}.$$

■