Market Structure and the Limits of Arbitrage*

Vincent Fardeau†
Federal Reserve Board
Job Market Paper

November 2012

Abstract

Financial markets are intermediated by large arbitrageurs that may benefit from their market power but face margin constraints. This paper studies how margin constraints affect liquidity provision and welfare under different structures of the arbitrage industry. In competitive markets, margin requirements may impair the ability of arbitrageurs to provide liquidity, which reduces other investors' welfare. Instead, imposing margins on a monopolistic arbitrageur can in some cases make the market more liquid and increase social welfare. Further, for a given level of capital in the arbitrage industry, markets are not necessarily less liquid under a monopolistic structure. A monopolistic arbitrageur captures rents that can be used as collateral, which relaxes her margin constraints and may result in a more liquid market. A delicate interaction between margin constraints and market power thus determines how efficiently financial institutions deploy available capital.

JEL classification: G12, D41, D60

Keywords: Limits of arbitrage, liquidity, strategic arbitrage, market structure, price impact, margin requirements, Dodd-Frank

*I benefited from comments from seminar and conference participants at the Federal Reserve Board and the 2012 Econometric Society European Meetings. I am grateful to Denis Gromb, Xuewen Liu, Anna Obizhaeva and Dimitri Vayanos for detailed suggestions.

†Risk Analysis Section, Division of Research and Statistics, Federal Reserve Board, Mail Stop 91, Washington, DC 20551, USA. The views presented here are solely those of the author and do not necessarily represent those of the Federal Reserve Board or its staff. Email: vincent.fardeau@frb.gov.

1 Introduction

Financial markets are dominated by large intermediaries who may benefit from their market power. As the 2007-2009 crisis illustrated, the financial constraints faced by these arbitrageurs play an important role in their ability to provide market liquidity (see e.g. Garleânu and Pedersen (2011)). How these constraints interact with arbitrageurs' market power and affect the intermediation process and the provision of liquidity, however, remains unclear. This question has gained importance as the crisis – and the regulatory response to it – defined a new landscape characterized by even greater concentration and a likely increase in the collateral needs of arbitrageurs. Major players collapsed or merged during the peak of the financial turmoil in 2008. The remaining investment banks and intermediaries are now facing new rules under Dodd-Frank and Basel III such as new margin requirements on the trading of derivatives, or the new liquidity coverage ratio, which requires institutions to hold enough high quality liquid assets to cover potential outflows in stressed market conditions.

Whether regulators should aim at increasing competition among intermediaries, or at relaxing margin and capital requirements (or some combination of both) to smooth the provision of liquidity has been the source of intense debates since the crisis. A starting point to answer these questions consists in understanding how the interactions between the market structure and margin constraints affects liquidity provision by arbitrageurs. This paper offers a framework to do so. For a given level of capital, I compare the properties of two polar structures: a monopolistic arbitrageur on one hand, and a competitive arbitrage sector on the other hand.

One might expect that market power reduces an arbitrageur's incentives to provide liquidity. I show that this intuition applies when capital is abundant. However, when capital is less abundant, margin requirements may curb some of the negative effects of market power on liquidity and efficiency. Thus imposing margins or limiting arbitrageurs' capital may improve the allocative efficiency of financial markets when they are dominated by a few large players. Instead, when the arbitrage sector is competitive, binding margin constraints always decrease the arbitrageurs' ability to enforce the law of one price, which hurts other investors in the market. Further, for a given level of capital, the market is not always more liquid under a competitive structure. A monopolistic arbitrageur captures rents that can be used as collateral, relaxing her financial constraints. As a result, when capital is relatively scarce, the monopolist may be less constrained than competitive arbitrageurs, and provide more market liquidity at some point in the life of the arbitrage opportunity. In other words, for a given level of capital, the cost of short-term debt funding should be lower for intermediaries in concentrated markets, leading to a different pattern of liquidity provision over time.

I consider a model where a monopolistic arbitrageur exploits price differences between two identical risky assets traded in segmented markets. In each market, some local investors receive endowment shocks over time, which affects their valuation for the asset. I assume that endowment shocks offset each other. In one market, local investors become more exposed to the risky asset and become natural sellers, while it is the opposite in the other market. Thus, local investors could perfectly share risks if they could trade directly with each other. However, market segmentation prevents this from happening, which creates a wedge between the prices of the risky assets. When the assets mature, they pay off identical liquidating dividends, which eliminates the price wedge. This generally captures the idea that market power may be temporary, and that sufficient capital may eventually enter the market. Prior to maturity, the arbitrageur can exploit the price wedge by intermediating trades between local investors (i.e. selling high and buying low), and so doing, provides liquidity to the market. Financial intermediaries use their capital to leverage their investments (e.g. through repo transactions). To capture this important aspect of the intermediation process, I assume that the arbitrageur must collateralize separately each leg of the arbitrage. The margins on her positions are set to cover the maximum potential losses over the next period. This generates limits on the size of her positions as a function of her wealth. As a benchmark, I consider an identical model where the arbitrageur stands for a continuum of identical price-taking traders, who as a group hold the same amount of capital, and face the same constraints. The price wedge is the main measure of liquidity.

In a competitive economy, well-capitalized arbitrageurs fully eliminate the price wedge, which drives their profits to zero. A shortage of capital, however, can impair their ability to enforce the law of one price. Constrained arbitrageurs are unable to fully absorb the demand from local investors, which leaves the market imperfectly liquid. The positive price wedge, however, generates capital gains for arbitrageurs, which progressively relaxes their constraints, and allows them to provide further liquidity.

In a monopolistic economy, market power alters the interaction between market liquidity and funding liquidity (the arbitrageur's ability to finance her positions). The arbitrageur recognizes that her trades decrease the price wedge and thus the profitability of the arbitrage. Thus she takes limited positions, even if capital is abundant. As a result, prices do not fully converge until the assets pay off, and the arbitrageur makes positive profits in equilibrium. This implies that the arbitrageur remains unconstrained with a lower level of capital than competitive arbitrageurs. The first mechanical effect is that, as the arbitrageur takes smaller positions than competitive arbitrageurs (considered as a whole), she requires a lower amount of capital. The second effect is that rents can be used as collateral, decreasing the margin per unit. This effect is present even at the onset, since the financiers setting margins anticipate that the arbitrageur will make equilibrium profits, which lowers margins.

The lesser severity of the constraint on the monopolistic arbitrageur drives the result that she in some cases provides more liquidity than competitive arbitrageurs. When capital is relatively low, there exists a region where the monopolistic arbitrageur is still unconstrained, while for the same level of capital, competitive arbitrageurs would be constrained. Thanks to the additional financial flexibility, the monopolistic arbitrageur captures larger capital gains and provides more liquidity at later dates.

I show more generally that margin constraints can have sharply different effects under different market structures. When arbitrageurs are competitive, margins may limit their ability to provide liquidity, which hurts local investors. When the arbitrageur is a monopolist, imposing margins may be Pareto-improving and increase liquidity relative to a no-constraint economy. The result relies on the analogy between a monopolistic arbitrageur and a large durable good provider, as analyzed by Coase (1972). To maximize profits, the arbitrageur breaks up her trades into small orders. This implies a decreasing price wedge over time. Local investors can anticipate this price pattern, which erodes the arbitrageur's market power. Consider for instance the market where local investors are natural sellers. In this market, the arbitrageur seeks to buy progressively to limit her price impact. In equilibrium, the price increase over the next period compensates local investors to hold some of the asset for an additional period. Thus, if the arbitrageur could commit not to improve the price in the future, local investors would be demanding more liquidity today, or willing to accept larger price concessions. This is precisely what the financial constraint achieves. By limiting the arbitrageur's ability to provide additional liquidity in the future, the margin constraint operates as a commitment device for the arbitrageur. As a result, the arbitrageur faces a larger price wedge today, and buys more aggressively, which increases her profit. This also benefits local investors, who, in the model, value receiving liquidity early.

This equilibrium arises only for an intermediate level of capital. On one hand, the arbitrageur must hold enough capital to be able to exploit the larger price wedge early on. On the other hand, the arbitrageur's opportunity cost of being constrained in the future is larger when she has a larger amount of capital. Further, a large amount of capital would allow the arbitrageur to re-optimize. Local investors being rational, they can foresee this behaviour, and the commitment power of the financial constraint unravels. Interestingly, the arbitrageur in some case holds enough capital to remain unconstrained throughout, but chooses (in a dynamically-consistent way) to trade more aggressively early on in order to "consume" her capital more quickly, and be constrained in the future. In this sense, margin constraints may bind as a result of deliberate actions by the monopolistic arbitrageur, rather than a shortage of capital.

My analysis has implications for the debates about the new architecture of financial markets. First, the analysis nuances some of the negative effects of market power or collusion among broker-dealers, or other institutions.[1] Policies fostering competition (e.g. by breaking up financial

[1]Christie and Shultz (1994) find evidence consistent with collusive behaviour of dealers on the NASDAQ. Further, Ellis, Michaely and O'Hara (2002) show that the dealers who were lead underwriters in a stock are likely to become the main, if not sole, dealer in this stock on the NASDAQ.

institutions, opening specialist market-making businesses to entry, or through other regulations preventing collusion such as tick rules) benefit liquidity consumers (local investors), but decrease arbitrageur's profits. I show that the overall effect can be negative, as aggregate welfare may decrease. Further, liquidity does not necessarily increase at all dates. This effect is related to the ability of a monopolistic arbitrageur to secure cheaper funding. It is often argued that large financial institutions can fund themselves more cheaply than others because of an implicit government put.[2] The model suggests a complementary rationale: investors with market power have a larger "pledgeable income" since they extract rents from the market. This effect echoes some ideas developed in the banking literature. For instance, Keeley (1990) shows that banks with market power are more likely to act prudently with regard to risk-taking, because they risk losing valuable bank charters.

Second, the model implies that regulators should not consider the issues of margin regulation independently of the market structure. This point is relevant for the new organization of derivatives markets. A concern among regulators and major broker-dealers is that newly imposed margins could impair arbitrageurs' ability to make these markets, with ripple effects on other assets.[3] As derivatives markets are often dominated by a few large dealers, the model shows that these concerns may be justified only when capital is very scarce. When capital is intermediate, the model predicts that margin requirements may improve liquidity and social welfare. When capital is abundant, liquidity provision is insensitive to margins.

This point relates more broadly to the long-standing debates about the benefits of a monopolistic market-maker, such as the NYSE specialists, as internal risk-management practices (e.g. VaR) are likely to generate margin-like constraints on the market-maker.[4] The paper shows that the specialist's obligation to maintain a continuous presence on the market is *in itself* a factor that erodes her market power (independently of actual participation costs). Margin constraints can in some cases *increase* a specialist's market power, by limiting her ability to participate in the market in the future. Surprisingly, when capital is intermediate, this does not make the market less liquid than if there were no constraint at all.[5]

The paper belongs to the literature on the limits of arbitrage. My setting is closely related to Gromb and Vayanos (2002, 2010) and Brunnermeier and Pedersen (2009). As in Gromb and

[2]See e.g. Acharya, Cooley, Richardson and Walter (2010).

[3]See e.g. Bullock, Braithwaite and McCrum, "Dealer and investors talks over liquidity fears", *Financial Times*, June 1, 2012.

[4]Specialists enjoy a privileged position on the NYSE, but face some competition from limit orders and off-exchange investors. See Seppi (1997) for a discussion.

[5]Glosten (1989) shows that arbitrageur's market power can have benefits in a model of asymmetric information. When arbitrageurs (in Glosten's context, market-makers) are competitive, the market may break down when the adverse selection problem becomes extreme. A monopolistic market-maker (e.g. a specialist) can average profits over time, which reduces the likelihood of a market break-down. In my model, information is perfect, and the benefit of market power comes from the fact that the monopolist can fund itself more cheaply.

Vayanos (2002), the focus of this paper is on welfare. My contribution is to consider the interaction between financial constraints and the market structure. The literature on the limits of arbitrage relies extensively on the assumption of price competitiveness (See Gromb and Vayanos (2010) for a survey). My results show that the effects of financial constraints on liquidity may be different and sometimes opposite when an arbitrageur has price impact.

The arbitrageur's price impact stems from market power. In this respect, this paper can be seen as bridging the gap between two groups of papers in the literature on asset pricing with imperfectly competitive investors. The first group, which includes Basak (1997), Vayanos (1999, 2001), Kihlstrom (2000), Pritsker (2009) and DeMarzo and Urosevic (2007), models all investors as rational and emphasizes the parallel with the durable goods problem studied by Coase (1972) (in particular in Vayanos (1999, 2001), Kihlstrom (2000) and DeMarzo and Urosevic (2007)). However, this literature does not model financial constraints. The second group, which includes Attari and Mello (2006) and Oehmke (2010), considers the effects of financial constraints but models only large traders, assuming that local investors have an exogenous downward-sloping demand curves. Attari and Mello (2006), in particular, study the trading strategy of a financially constrained monopolistic arbitrageur. Relative to their analysis, the modeling of local investors as rational agents has two important consequences: i) it allows me to carry out a welfare analysis under different market structures; ii) it generates an endogenous market depth, which is determined by the complex interaction between the arbitrageur's market power and the margin constraints. This mechanism is at the heart of my results. Therefore, to the best of my knowledge, this paper is the first to solve the dynamic problem of a monopolistic investor under realistic financial constraints when all investors are rational. My contribution in this context is to show that financial constraints may alleviate the arbitrageur's commitment problem by providing a commitment device without impairing market liquidity.[6]

The model predicts that with a monopolistic arbitrageur, arbitrage opportunities should converge only gradually, whether margin constraints bind or not, while the competitive arbitrage model predicts this pattern only when there is a shortage of capital. There are additional interesting price and liquidity effects caused by imperfect competition in the unconstrained region, but these are outside of the scope of this paper. I present them in a slightly more general setting allowing for an oligopolistic structure and time-varying risk-sharing needs of local investors in a companion paper (Fardeau, 2012).

The paper proceeds as follows. I present the model in the next section. In section 3, I review the competitive equilibrium and its properties. In Section 4, I study the monopolistic equilibrium. I compare liquidity across market structures in Section 5. Section 6 concludes. The appendix contains the proofs.

[6] There is also a loose analogy between this effect and the use of leverage as a strategic bargaining tool by shareholders against unions (Perotti and Spier, 1993).

2 Model

Assets and timeline. The model has three periods, indexed by $t = 0, 1, 2$. The financial market is open at time 0 and time 1, and consumption takes place at time 2. There are two identical risky assets, A and B, and a risk-free asset with return r_f normalized to 0. Assets A and B are in zero net supply and pay a dividend D_2 at time 2, with $D_2 = D + \epsilon_1 + \epsilon_2$, where ϵ_t is a random variable with a symmetric bounded support $[-\bar{e}, \bar{e}]$, a mean of 0 and volatility σ. The distribution does not need to be further specified, but to facilitate the interpretation of the results, I will sometimes use a particular distribution described below. The information ϵ_t is revealed to all investors at time t before trading. The price of asset k at time t is denoted p_t^k. Each asset k is traded on its own, segmented market.

Local investors. In each market, there are risk-averse local investors with mean-variance preferences: for $k = A, B$, $U\left(W_2^k\right) = \mathbb{E}\left(W_2^k\right) - \frac{a}{2}\mathbb{V}\left(W_2^k\right)$. Local investors experience endowment shocks $s\epsilon_t$ that are correlated with the dividend of the risky asset. That is, at time $t = 0, 1$, local investors in market A receive a shock $s\epsilon_{t+1}$, where the magnitude of the shock, $s > 0$, is deterministic. B-investors receive opposite shocks, $-s\epsilon_{t+1}$.[7] Since k-investors have only access to asset k (by market segmentation) and the risk-free security, they cannot share risk with the other group, although they could perfectly insure each other. The shocks and market segmentation imply potential price differences between assets A and B, although their cash-flows are identical. In particular, A-investors have a low valuation for the asset, and B-investors a high valuation.

At time 2, local investors consume their wealth W_2^k. Let E_t^k and Y_t^k denote their end-of-period positions in the risk-free and risky asset k, respectively. Then we can write local investors' final wealth as follows:

$$\text{for } k = A, B,\ W_2^k = E_1^k + Y_1^k D_2,$$

The dynamic budget constraint follows from the dynamics of asset holdings: $Y_t^k = Y_{t-1}^k + y_t^k$ and $E_t^k = E_{t-1}^k - y_t^k p_t^k + s\epsilon_{t+1}$, where y_t^k denotes the time-t trade of investors k.

Arbitrageur(s). There is an additional investor, the arbitrageur, who can participate in all markets. The arbitrageur is also endowed with mean-variance preferences over wealth: $u\left(W_2\right) = \mathbb{E}\left(W_2\right) - \frac{b}{2}\mathbb{V}\left(W_2\right)$, albeit with a potentially different risk-aversion b. Given that she has access to all securities, the arbitrageur's final wealth is

$$W_2 = \sum_{k=A,B} X_1^k D_2 + B_1$$

[7]The results do not depend on the mean-variance framework. I use these preferences because they offer greater tractability when endowment shocks are stochastic, an extension that I am planning to consider in future work.

with for each asset k, $X_t^k = X_{t-1}^k + x_t^k$ denotes the end-of-period position at time t in asset k, x_t^k the corresponding trade, and $B_t = B_{t-1} - \sum_{k=A,B} x_t^k p_t^k$, the arbitrageur's risk-free asset holdings at the end of period t. I assume that the arbitrageur has no endowment in the risky assets, $X_{-1}^k = 0$, $k = A, B$, and starts with an initial wealth $W_{-1} = B_{-1}$. Apart from Section 3, where I consider the benchmark case where the arbitrageur stands for a continuum of competitive investors, I assume that the arbitrageur is a price-setter in both A and B markets. Specifically, I assume that the arbitrageur chooses positions, knowing the local investors' demand in each market, and imposing market-clearing.

Financial constraints. Whether the arbitrageur is price-taker or price-setter, she needs capital to trade the risky assets. I model the financial constraint in the same fashion as Gromb and Vayanos (2002, 2010) and Brunnnermeier and Pedersen (2008). Arbitrageurs have a margin account V_t^k in each market, and their positions must be fully collateralized. That is, the arbitrageur's wealth in this account must cover the maximum possible loss on the position over the next period:

$$V_{t-1}^k \geq \max_{p_{t+1}^k} X_t^k \left(p_t^k - p_{t+1}^k\right)$$

Hence, in total, the arbitrageur's wealth must cover the total maximum loss on each account:[8]

$$W_{t-1} \geq \sum_{k=A,B} \max_{p_{t+1}^k} X_t^k \left(p_t^k - p_{t+1}^k\right) \tag{1}$$

The presence of the financial constraint implies that arbitrageurs may not be able to fully eliminate the price differences between A and B assets. The modeling of the constraint also implies that asset A cannot be used as collateral for asset B (and vice-versa). In other words, cross-collateralization is not allowed, which can be viewed as a consequence of the assumption of market segmentation. In practice, cross-collateralization is often limited by financiers who are concerned about imperfect correlation between assets (although this would not be an issue here). Sometimes traders also voluntarily avoid cross-collateralization in order to avoid revealing their trading strategies.[9]

The financial constraint corresponds to a one-period VaR constraint at the 100 percent level (as implied by the assumption of full collateralization). The 100 percent level is for simplicity only, as it rules out default in equilibrium and thus makes welfare comparisons simpler[10], but the constraint

[8] I define W_{t-1} as the end-of-period wealth, while Gromb and Vayanos (2002) use $\tilde{W}_t$ as the beginning-of-period wealth. Given this difference in notation: $W_{t-1} = \tilde{W}_t$. The same applies to the definition of margin accounts.

[9] For instance, Pérold (1999) reports: "LTCM inernalized most of the back-office functions associated with contractual arrangements, due to the complexity and and advanced nature of many of the firm's trades. This also helped maintain the confidentiality of its positions. LTCM chose Bear Stearns as a clearing agent partly because Bear Stearns was committed to customer business rather than being focused on proprietary trading, and thus there were fewer conflicts of interest."

[10] There is no need to compute the welfare of financiers on the other side of the constraint.

is motivated by real-world margin setting.[11] An important feature of the constraint is that it is forward-looking, in the sense that it is based on both current and future prices.

Following Gromb and Vayanos (2002), I will focus on equilibria in which the arbitrageur holds opposite positions in both assets, i.e. $X_t^A = -X_t^B = X_t$. Given that the arbitrageur starts with no endowment in the risky assets, this implies that $x_t^A = -x_t^B = x_t$, for $t = 0, 1$. Using this assumption, we can rewrite the arbitrageur's budget constraint as follows:

$$W_2 = W_{-1} + \sum_{t=0,1} x_t \Delta_t, \text{ with } \Delta_t = p_t^B - p_t^A$$

The equation shows that by setting up opposite position in each leg of the arbitrage, the arbitrageur eliminates all fundamental risk and derives all her profits from exploiting the price difference Δ between the two markets. This assumption also simplifies the financial constraint, because it implies that the risk premia on asset A and B are opposite. That is, $\phi_t^A = D_t - p_t^A = \frac{\Delta_t}{2} = -\phi_t^B$, where D_t is the conditional expected value of the asset at time t: $D_t = D_{t-1} + \epsilon_t$. This implies that $p_t^k - p_{t+1}^k = \phi_{t+1}^k - \phi_t^k - \epsilon_{t+1} = \frac{\Delta_{t+1} - \Delta_t}{2} - \epsilon_{t+1}$. As a result, we can rewrite the financial constraint (1) as follows:

$$\begin{aligned} W_{t-1} &\geq \sum_{k=A,B} \max_{p_{t+1}^k} X_t^k \left(p_t^k - p_{t+1}^k \right) \\ &\geq \max_{\epsilon_{t+1}} X_t \left(\frac{\Delta_{t+1} - \Delta_t}{2} - \epsilon_{t+1} \right) + \max_{\epsilon_{t+1}} -X_t \left(-\frac{\Delta_{t+1} - \Delta_t}{2} - \epsilon_{t+1} \right) \\ &\geq 2X_t \left(\frac{\Delta_{t+1} - \Delta_t}{2} \right) + \max_{\epsilon_{t+1}} X_t \left(-\epsilon_{t+1} \right) + \max_{\epsilon_{t+1}} -X_t \left(-\epsilon_{t+1} \right) \\ &\geq 2|X_t| \bar{e} - X_t \left(\Delta_t - \Delta_{t+1} \right) \end{aligned} \tag{2}$$

The last step follows from the symmetric support of the distribution.

Properties of margins. Suppose that the arbitrageur holds a long position, $X_t \geq 0$. We can rewrite the right-hand side of inequality (2) as $2m_t X_t$, where the margin m_t is

$$m_t = \bar{e} - \frac{1}{2} \left(\Delta_t - \Delta_{t+1} \right) \text{ if } X_t \geq 0 \tag{3}$$

The properties of the margin are key for the dynamics of the model. Clearly, margins increase with the "dispersion" (and consequently, volatility) of the fundamental $\bar{e}$. A more volatile asset leads to a larger potential loss on the position, which induces financiers to ask for more collateral. Margins also depend on the mispricing between asset A and B. More specifically, they depend on the *change* in the mispricing, $\Delta_t - \Delta_{t+1}$. If financiers expect market liquidity to improve, i.e. $\Delta_{t+1} \leq \Delta_t$, they reduce current margins. Hence, the financiers' behaviour assumed here leads to countercyclical

[11] See Brunnermeier and Pedersen (2009), Appendix A, for additional institutional details.

margins relative to mispricings (illiquidity). Said differently, margins play a stabilizing role for asset prices. If a drop in liquidity (i.e. large Δ_t) is temporary, then financiers do not necessarily ask for more capital. Brunnermeier and Pedersen (2009) consider a similar constraint in their benchmark case. They also consider a situation financiers are assumed to be uninformed. They show that in this case, uncertainty about whether the mispricing will decrease or not in the future can lead to procyclical, destabilizing margins.[12]

Remark: relation between volatility, tail risk and support boundary. I show in section 4 that the ratio $\frac{\bar{e}}{\sigma^2}$ plays an important role for the equilibrium. Since the dispersion of the fundamental $\bar{e}$ and its volatility are related, it is useful to specify a simple distribution of fundamental shocks. The following coarse four-point symmetric distribution is enough to gain intuition:

Lemma 1 *Let $(\mu, p) \in \left]0, \infty\right[\times \left]0, 1\right[$, and $\epsilon_t \sim \mathcal{E}\left[-\bar{e}, \bar{e}\right]$, where the random variable $\mathcal{E}$ takes the following values:*

$$\mathcal{E} = \begin{cases} -\bar{e} \text{ with probability } \frac{1}{2} - p \\ -\frac{\bar{e}}{\mu} \text{ with probability } p \\ \frac{\bar{e}}{\mu} \text{ with probability } p \\ \bar{e} \text{ with probability } \frac{1}{2} - p \end{cases}$$

Then $\mathbb{E}\left(\mathcal{E}\right) = 0$ and $\sigma^2 = \mathbb{V}\left(\mathcal{E}\right) = \bar{e}^2 \left[1 + 2p\left(\frac{1}{\mu^2} - 1\right)\right]$.

This example shows how the variance of fundamental shocks relates to the support boundary or "dispersion" $\bar{e}$. Although this distribution is just meant to fix ideas, the relation between σ^2 and $\bar{e}^2$ is more general. Further, this example can help us clarify how the volatility in a symmetric distribution can relate to the shape of the tails. The parameter μ measures how far the median values are from the mean 0, while p measures the weight of the tails: a small p means that tail events in which ϵ_t takes the extreme values $\bar{e}$ or $-\bar{e}$ are likely. Clearly the variance decreases with μ and with p (since $\frac{1}{\mu^2} - 1 < 0$). Hence, when extreme events are likely (small p), the variance is large. More generally, this example shows that while an increase in the boundaries of the distribution of fundamentals $\bar{e}$ always increases volatility, volatility may also increase because of a change in the shape of the distribution, without changing the dispersion $\bar{e}$.

[12] Brunnermeier and Pedersen show that a margin spiral, in which low liquidity leads to higher margins, which further limits the ability of arbitrageurs to provide liquidity, can result from the uninformed case. This margin spiral complements and amplifies the loss spiral created by the financial constraint ("a decrease in arbitrageurs' capital impairs their ability to provide liquidity and eliminate the mispricing, which in turn reduces their capital"). Under our assumptions, there can be a loss spiral, but no margin spiral.

3 Competitive equilibrium benchmark

In this section, I briefly recall the competitive benchmark derived in Gromb and Vayanos (2002). The model illustrates how liquidity (given by the spread between assets A and B) depends on arbitrageurs' capital.

Proposition 1 *(Gromb and Vayanos, 2002) There exists a unique competitive equilibrium with symmetric liquidity premia given by:*

- *If $W_{-1} \geq \omega^* \equiv 2s\bar{e}$, the financial constraint never binds, the arbitrageurs absorb the liquidity shock s, i.e. $X_t = s$ at $t = 0, 1$, and the spread between assets A and B is always 0: $\Delta_0 = \Delta_1 = \Delta_2 = 0$*

- *If $0 \leq W_{-1} < \omega^*$, the financial constraint binds at $t = 0$ and $t = 1$ and the spread between assets A and B narrows over time and is closed only at $t = 2$, i.e. $\Delta_0 > \Delta_1 > \Delta_2 = 0$. The arbitrageur position in asset A is given by:*

$$x_0 - x_0 \frac{a\sigma^2 (s - x_0)}{\bar{e}} = \frac{W_{-1}}{2\bar{e}} \tag{4}$$

$$X_1 - X_1 \frac{a\sigma^2 s - X_1}{\bar{e}} = x_0 \tag{5}$$

The equilibrium links liquidity (via the spread) to arbitrageurs' initial capital and has a simple form: if arbitrageurs' capital is large enough, then the market is perfectly liquid, as reflected by the absence of spread between assets A and B; if instead arbitrageurs start with less capital, then the financial constraints are binding, and assets A and B trade at a positive spread, which decreases over time. An increase in the liquidity shock s affecting local investors or in the dispersion of the fundamental (increase in $\bar{e}$) tightens (proportionately) the financial constraint: the financiers anticipate that the price divergence between assets A and B is potentially larger and demand more collateral.

To facilitate comparison with the monopolistic case and gain further insight, I derive the equilibrium positions and spread as a function of arbitrageurs' capital in the constrained region:

Corollary 1 *If $0 \leq W_{-1} < \omega^*$:*

- *The arbitrageurs' positions in asset A are:*

$$x_0 = \frac{a\sigma^2 s - \bar{e} + \sqrt{Q}}{2a\sigma^2}; \; X_1 = \frac{a\sigma^2 s - \bar{e} + \sqrt{U}}{2a\sigma^2}$$

  with $Q = \left(\bar{e} - a\sigma^2 s\right)^2 + 2a\sigma^2 W_{-1}$ and $U = \left(\bar{e} - a\sigma^2 s\right)^2 + 4a\sigma^2 x_0 \bar{e}$.

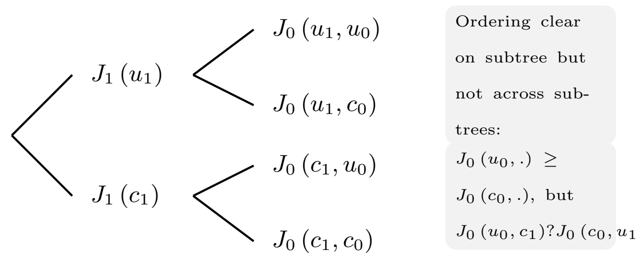

Figure 1: Game tree

- *The equilibrium spreads are:*

$$\Delta_0 = 2\left(a\sigma^2 s + \bar{e}\right) - \sqrt{Q} - \sqrt{U};\ \Delta_1 = a\sigma^2 s + \bar{e} - \sqrt{U}$$

This result shows how the positions and the spreads depend on arbitrageurs' capital: clearly the spread at time 0 and 1 decreases with capital W_{-1} (i.e. $\frac{\partial \Delta_t}{\partial W_{-1}} < 0$, and even more so if capital is low, a non-linear effect (i.e. $\frac{\partial^2 \Delta_t}{\partial W_{-1}^2} < 0$).[13]

4 Equilibrium with a monopolistic arbitrageur

In this section, I derive the trading strategy of a monopolistic arbitrageur and compare it to the competitive case. In the monopoly case, market power allows the arbitrageur to limit market liquidity but also relaxes her margin constraint. The arbitrageur also faces a commitment problem as local investors recognize that, even though the arbitrageur can limit liquidity at the current stage, she always has an interest to provide further liquidity at a later stage. Liquidity provision by a single arbitrageur thus resembles the provision of a durable good by a monopolist and is subject to Coasian dynamics.

4.1 Coasian dynamics and time consistency

I start by introducing some useful notation and presenting the main steps of the analysis. Since the arbitrageur may be constrained (superscript c) or not (superscript u) at each date, there are two payoffs associated with the different combinations at time 1: $J_1^{c_1}$ and $J_1^{u_1}$, and four at time 0: $J_0^{u_1,u_0}$, $J_0^{u_1,c_0}$, $J_0^{c_1,u_0}$, $J_0^{c_1,c_0}$ (Figure 1).

Time 1. The arbitrageur enters time 1 with a position x_0 in asset A (and the opposite in asset B). Local investors in market A have the following demand for the risky asset:

$$Y_1^A = \frac{E_1\left(D_2\right) - p_1^A}{a\sigma^2} - s,$$

[13] See also Brunnermeier and Pedersen (2008).

where the second term s represents the hedging demand stemming from the endowment shocks. In market B, local investors' demand is $\frac{E_1(D_2)-p_1^B}{a\sigma^2}+s$: since shocks are opposite across markets, so is the hedging demand. Market-clearing in markets A and B,

$$Y_1^k + X_1^k = 0,\ k = A, B$$

implies the following mapping between the price difference Δ_1 and quantities:

$$\Delta_1 (X_1) = 2a\sigma^2 (s - X_1),$$

where $X_1 = X_0 + x_1 = x_0 + x_1$. Given that arbitrageurs take opposite positions across markets and that liquidity shocks are known in advance, the arbitrageur faces a risk-free arbitrage opportunity. Thus the variance term in her utility function disappears and her maximization problem boils down to:

$$\begin{aligned} J_1 &= \max_{x_1} B_0 + x_1 \Delta_1 (X_1) \\ &\quad s.t. \quad W_0 \geq 2X_1 \left[\bar{e} - \frac{1}{2} (\Delta_1 - \Delta_2)\right] \text{ if } X_1 \geq 0 \end{aligned}$$

The arbitrageur's problem is to maximize her payoff at time 1, $x_1\Delta_1 (X_1)$, subject to the financial constraint, and for given initial positions in risky and risk-free assets, x_0 and B_0. The constraint is based on the assumption that the arbitrageur holds a long position in the arbitrage at time 1, which will be the case in equilibrium. The dependence of Δ_1 on the arbitrageur's position X_1 shows that the arbitrageur accounts for her price impact. When the constraint is not binding, the arbitrageur trades the quantity $x_1^u = \frac{s-x_0}{2}$. When the constraint is binding, the arbitrageur, who is effectively risk-neutral, "maxes out" her constraint to maximize profits. She chooses a quantity x_1^c which saturates the constraint: $W_0 = 2X_1^c [\bar{e} - (\Delta_1 - \Delta_2)]$, where $X_1^c = x_0 + x_1^c$. Therefore, the two possible payoffs at time 1 are given by:

$$\begin{aligned} J_1^{u_1} (x_0) &= B_0 + 2a\sigma^2 x_1^u (s - X_1^u) = B_0 + \frac{a\sigma^2}{2} (s - x_0)^2 \\ J_1^{c_1} (x_0) &= B_0 + 2a\sigma^2 x_1^c (s - X_1^c), \text{ with } X_1^j = x_0 + x_1^j,\ j = u, c \end{aligned}$$

Of course, by construction, for any initial position x_0, $J_1^{c_1} (x_0) \leq J_1^{u_1} (x_0)$. Similarly, the time-1 equilibrium spread depends on the state and the arbitrageur's beginning-of-period position x_0: $\Delta_1 = \Delta_1^{u_1} (x_0)$ if the constraint is slack, and $\Delta_1 = \Delta_1^{c_1} (x_0)$ otherwise, with $\Delta_1^{c_1} (x_0) \geq \Delta_1^{u_1} (x_0)$.

Time 0 - Coasian dynamics. At time 0, the relation between the four payoffs $J_0^{u_1,u_0}$, $J_0^{u_1,c_0}$, $J_0^{c_1,u_0}$, $J_0^{c_1,c_0}$ is more complicated due to Coasian dynamics. To see why, we can start from local investors' demand at time 0. Since local investors are rational, their demand depends on the current price and their expectation of the future price (which will be correct in equilibrium). For instance, in market A:

$$Y_0^A = \frac{E_0\left(p_1^A\right) - p_0^A}{a\sigma^2} - s \tag{6}$$

It is useful to build intuition to rewrite this equation in terms of liquidity premia $\phi_t^A = D_t - p_t^A$ as

$$Y_0^A = \frac{-\mathbb{E}_0\left(\phi_1^A\right) + \phi_0^A}{a\sigma^2} - s$$

When local investors in market A believe that the time-1 price will be high, i.e. expect a low liquidity premium ϕ_1^A, they hold larger positions at time 0. Since local investors in market A are natural sellers, this means that they demand less liquidity when they anticipate a higher price at time 1. Conversely, an anticipation of a low price at time 1 (i.e. high ϕ_1^A) increases their liquidity demand at time 0.

Further, by providing liquidity early on, the arbitrageur reduces the demand for liquidity later. This is because the liquidity shocks experienced by local investors are correlated with the asset payoff. Hence hedging by trading in their local asset today provides local investors with a "proxy hedge" against the next shock. This is illustrated by the expression of $x_1^u = \frac{s-x_0}{2}$: at time 1, the arbitrageur, if unconstrained, serves a fraction of the liquidity demand of the period, which is $s - x_0$, instead of s, in each market. As a result, providing liquidity is subject to the same Coasian dynamics as providing a durable good. Namely, receiving liquidity "insures" local investors durably (albeit imperfectly) against shocks.

The interaction between these Coasian dynamics and the financial constraints significantly complicates the analysis, as it makes standard backward induction methods inappropriate. When the financial constraint binds at time 1, the prices of assets A and B remain further apart than what the arbitrageur's trading would imply if she had more capital. Since local investors understand it, their liquidity demand is larger at time 0 when they expect the constrained state to occur at time 1. The arbitrageur may thus *benefit* from being constrained at time 1 if the additional benefit coming from increased liquidity demand at time 0 offsets the cost of being constrained at time 1. Thus trading in such a way that the constraint is binding at time 1 may be optimal for the arbitrageur. This contradicts Bellman principle given that being constrained yields a lower payoff in the subgame at time 1.

Maximization problems. *Formally*, the trade-off faced by the arbitrageur is illustrated by the arbitrageur's value functions at time 0. From local investors' demand (6), and market-clearing, $Y_0^k + X_0^k = 0$ ($k = A, B$), we can derive the mapping between the spread and the arbitrageur's position at time 0:

$$\Delta_0(x_0) = \Delta_1(x_0) + 2a\sigma^2(s - x_0) \tag{7}$$

Since Δ_1 depends on the state at time 1, we can define $\Delta_0^{u_1}(x_0)$, and $\Delta_0^{c_1}(x_0)$ the time-0 price schedules implied by the corresponding beliefs about the state at time 1. Depending on the anticipated time-1 state, the arbitrageur's maximization problem is:

$$J_0^{u_1} = \max_{x_0} W_{-1} + x_0\Delta_0^{u_1}(x_0) + 2a\sigma^2 x_1^u(s - X_1^u) \tag{8}$$

$$s.t. \quad W_{-1} \geq 2|x_0|\bar{e} - x_0(\Delta_0 - \Delta_1) \tag{9}$$

$$\text{or } J_0^{c_1} = \max_{x_0} W_{-1} + x_0\Delta_0^{c_1}(x_0) + 2a\sigma^2 x_1^c(s - X_1^c) \tag{10}$$

$$s.t. \quad W_{-1} \geq 2|x_0|\bar{e} - x_0(\Delta_0 - \Delta_1)$$

The last term in the maximand of J_0 is the profit made from the arbitrage at time 1. As noted above, $x_1^u(s - X_1^u) \geq x_1^c(s - X_1^c)$. However, given that the time-0 spread is different depending on local investors' beliefs, it is not guaranteed that being constrained at time 1 always yields a lower payoff than being unconstrained from the point of view of time 0. Indeed, $\Delta_1^{c_1}(x_0) \geq \Delta_1^{u_1}(x_0)$ and equation (7) imply that $\Delta_1^{c_1}(x_0) \geq \Delta_1^{u_1}(x_0)$. For this reason, it is possible in principle for instance that at the optimum, $J_0^{c_1,u_0} \geq J_0^{u_1,u_0}$, where u_0 indicates that the financial constraint does not bind at the optimum at time 0. Given that the arbitrageur is effectively risk-neutral and maxes out her financial constraint in the constrained region, it is only certain ex-ante that on a subtree of the game (i.e. conditional on being in state j at time 1), it is better to be unconstrained, e.g. $J_0^{u_1,u_0} \geq J_0^{u_1,c_0}$ (see Figure 1). Solving for the equilibrium thus requires to compute all four payoffs and comparing them whenever parameters imply that several strategies are possible at the same time. This is in general the case when capital W_{-1} is not too scarce in the economy. For instance, if there is enough capital to be unconstrained at all dates, then there should also be enough to play a strategy where the constraint is binding at some date.

Time consistency. There are some restrictions, however, on which strategies are consistent with local investors' rational expectations. For instance, rational expectations rule out cases where the price schedule $\Delta_0(x_0)$ is based on the belief that the arbitrageur will be constrained at time 1, and where she is not when time 1 comes. Thus *admissible* strategies are determined by

1. *feasibility* at a given date, i.e. whether the constraint is binding or not at this point in time, and

2. along the equilibrium path (*time consistency*)

For instance, my presentation of the maximization problems at time 0 in programmes (8) and (10) is incomplete in that it omits the time consistency condition. For the programmes to make sense, a necessary condition is that in equilibrium, the time-0 trade does satisfy the time-1 constraint if local investors expect the arbitrageur's constraint to be slack, and vice-versa if they expect the constraint to be binding. Thus the maximization problems should be completed with the following condition:

Lemma 2 *(Time consistency) Suppose that in equilibrium the arbitrageur chooses to trade a quantity $x_0 > -s$ at time 0. It is consistent with being unconstrained at time 1 if and only if*

$$k(x_0) \geq 0, \text{ with } k(x_0) \equiv W_{-1} - s\bar{e} + a\sigma^2 \frac{s^2}{2} + \left(2a\sigma^2 s - \bar{e}\right) x_0 - \frac{5}{2} a\sigma^2 x_0^2 \tag{11}$$

The feasibility and time consistency requirements are subsumed under the assumptions that the arbitrageur maximizes expected utility under the financial constraint and that local investors have rational expectations.

Definition 1 *An equilibrium is a collection of arbitrageur's trades $(x_t)_{t=0,1}$ (or equivalently positions x_0, X_1) in asset A and opposite trades (positions) in asset B, such that*

- *given (rationally expected) prices, local investors maximize their expected utility of final consumption,*
- *the arbitrageur maximizes her expected payoff subject to financial constraints, local investors' demands and market-clearing.*

4.2 Equilibrium with a very low level of capital

I start with the case where the arbitrageur's capital is very low. Then the arbitrageur has no financial flexibility and there is only one feasible strategy, so that the equilibrium is easy to determine.

Proposition 2 *Define $\omega^c \equiv \frac{7}{5}s\bar{e} - \frac{9}{10}a\sigma^2 s - \frac{\bar{e}^2}{10a\sigma^2}$. If $0 \leq W_{-1} < \omega^c$, the arbitrageur's constraint binds at $t = 0$ and $t = 1$. Equilibrium trades and spreads are the same as in Corollary 1.*

4.3 Strategies when capital is more abundant

If capital is more abundant ($W_{-1} \geq \omega^c$), the arbitrageur has more financial flexibility and can choose from a larger set of strategies. I now describe the strategies available to the arbitrageur.

Lemma 3 *Strategies with non-binding constraint at time 1. Denote $\omega_0^m = \frac{4}{5}s\bar{e} - \frac{12}{25}a\sigma^2 s^2$, $\omega_1^m = \frac{7}{5}s\bar{e} - \frac{9}{10}a\sigma^2 s^2$, $\omega^p = s\bar{e} - \frac{1}{2}a\sigma^2 s^2$ and $\omega^u = \frac{2}{9}s\bar{e} - \frac{2}{3}a\sigma^2 s^2 + \frac{2\bar{e}^2}{9a\sigma^2}$. Further, denote x_0^0 the long position saturating constraint* (9) *at time 0, and x_0^1 the largest position saturating constraint* (11) *at time 1. The following holds:*

- ***Strategy*** u_1, u_0. *If* $W_{-1} \geq \max(\omega_0^m, \omega_1^m)$, *the arbitrageur can be unconstrained at time 0 and time 1. The unconstrained strategy* (u_1, u_0) *consists of the following trades*

$$x_0^{u_1,u_0} = \frac{2}{5}s, \; x_1^{u_1,u_0} = \frac{3}{10}s$$

and yields a payoff $J_0^{u_1,u_0} = W_{-1} + \frac{9}{10}a\sigma^2 s^2$.

- ***Strategy*** u_1, c_0. *If* $W_{-1} \in [0, \max(\omega_0^m, \omega_1^m)[$, *the arbitrageur may remain unconstrained at time 1 (and trade* $x_1^{u_1,c_0} = \frac{s - x_0^{u_1,c_0}}{2}$*) by reducing her time-0 trade relative to* $x_0^{u_1,u_0}$.

  – *If* $W_{-1} \in [\omega^p, \max(\omega_0^m, \omega_1^m)]$, *or if* $W_{-1} \in [\omega^c, \omega^p[$ *and* $\frac{\bar{e}}{a\sigma^2 s} < 2$, *the arbitrageur can always remain unconstrained at time 1 by taking the following long position at time 0:*

$$x_0^{c_1,u_0} = \begin{cases} x_0^1 \text{ if } \omega_0^m \leq W_{-1} < \omega_1^m \\ x_0^0 \text{ if } \omega_1^m \leq W_{-1} < \omega_0^m \\ \min\left(x_0^0, x_0^1\right) \text{ if } \omega^p \leq W_{-1} < \min(\omega_0^m, \omega_1^m) \end{cases}$$

  – *If* $W_{-1} \in [\omega^c, \omega^p[$ *and* $\frac{\bar{e}}{a\sigma^2 s} \geq 2$, *the arbitrageur can remain unconstrained at time 1 only by taking a short position* $x_0^1 < 0$ *at time 0 and only in the following cases:*

    ∗ *If* $\frac{\bar{e}}{a\sigma^2 s} \in [2, \iota_1[$, *or* $\frac{\bar{e}}{a\sigma^2 s} \in [\iota_1, \iota_2[$ *and* $W_{-1} \in [\omega^u, \omega^p[$.

    ∗ *If* $\alpha W_{-1}^2 + \beta W_{-1} + \gamma < 0$ *when* $\frac{\bar{e}}{a\sigma^2 s} \in [\iota_1, \iota_2[$ *and* $W_{-1} \in [\omega^c, \omega^u[$, *or when* $\frac{\bar{e}}{a\sigma^2 s} > \iota_2$.

  – *If* $0 \leq W_{-1} < \omega^c$, *the arbitrageur cannot remain unconstrained at time 1.*

  – *The payoff of the strategy, as a function of the time 0 trade, is* $J_0^{u_1,c_0}(x_0) = W_{-1} + \frac{a\sigma^2 s^2}{2} + 2a\sigma^2 s x_0 - \frac{5}{2}a\sigma^2 x_0^2$.

With abundant capital, $W_{-1} \geq \max(\omega_0^m, \omega_1^m)$, the arbitrageur can remain unconstrained at all time. The u_1, u_0 strategy is the strategy that the arbitrageur would choose if she was not facing any financial constraints (and could not commit to trade only once). Trading a large quantity at time 0 reduces the price wedge at time 1, and thus the future profitability of the arbitrage. Thus, in the u_1, u_0 strategy, the arbitrageur trades at a pace such that the marginal benefit of buying an additional unit at time 0 equals the marginal cost at time 1.

This strategy can be intensive in capital. When capital is below a certain level (e.g. $W_{-1} < \max(\omega_0^m, \omega_1^m)$), the arbitrageur must alter her trading strategy if she wants to remain unconstrained at time 1. In effect, the arbitrageur must trade less aggressively at time 0, in order to free up some capital for the next period. Surprisingly, when capital is scarcer ($W_{-1} < \omega^p$ and $\frac{\bar{e}}{a\sigma^2 s} \geq 2$), this involves taking a short position in the arbitrage at time 0, whereas intuition suggests that the arbitrageur should always go long the spread. In fact, this short position at time 0 allows the arbitrageur to take a larger long position at time 1. Intuitively, the short position worsens the mispricing, resulting in a potentially larger spread at time 1. This reduces the margin the arbitrageur must pay to set up a long position at time 1. Indeed, from equation (3) at time 1, we can write the time-1 margin as $m_1 = \bar{e} - \frac{\Delta_1}{2}$ which is equal to $\bar{e} - a\sigma^2 (s - x_0)$ in equilibrium. Thus taking a small long position or indeed a short position at time 0 reduces the time-1 margin, which allows the arbitrageur to remain unconstrained. When $W_{-1} < \omega^p$, there may not always be enough capital to set up this position, and the arbitrageur may not be able to remain unconstrained at time 1. With very low capital ($W_{-1} < \omega^c$), the arbitrageur no longer has the ability to stay unconstrained.

Given that in the u_1, u_0, the arbitrageur trades at an optimal pace, it should be suboptimal to buy a larger quantity at time 0. It is indeed the case as long as the increased trading aggressiveness at time 0 does not make the time 1 constraint binding. However, if the arbitrageur's increased aggressiveness leads to a binding constraint at time 1, the properties of the time 0 spread change. In particular, the time 0 spread widens. This is a consequence of equation (7) and the fact that $\Delta_1^{c_1}(x_0) \geq \Delta_1^{u_1}(x_0)$. Hence, as highlighted in Section 4.1, the arbitrageur may benefit from being constrained at time 1. This implies that freeing up capital to remain unconstrained in the next period may not be optimal. However, strategies where the arbitrageur's constraint binds at time 1 are not necessarily time consistent.

Lemma 4 *Strategies with binding constraint at time 1. Denote* $\bar{\omega}^p \equiv \frac{3}{2}s\bar{e} - \frac{7}{8}a\sigma^2 s^2$, $\tilde{\omega} = -6s\bar{e} + 2a\sigma^2 s^2 + \frac{6\bar{e}^2 - \sqrt{\tilde{D}^w}}{a\sigma^2}$, *with* $\tilde{D}^w = 4a^4\sigma^8 s^4 + 36\bar{e}^4 + 88a^2\sigma^4 s^2\bar{e}^2 - 96a\sigma^2 s\bar{e}^3 - 32a^3\sigma^6 s^3\bar{e}$.

- ***Strategy*** c_1, u_0. *In this strategy, the arbitrageur increases her time 0 position relative to the* u_1, u_0 *strategy and her constraint is binding at time 1, but not at time 0. The strategy consists of the following positions:*

$$x_0^{c_1,u_0} = \frac{s}{2},\ X_1^{c_1,u_0} = \frac{s}{2} + \frac{\sqrt{U^m} - \bar{e}}{2a\sigma^2},\ \text{with } U^m = Q + a^2\sigma^4 s^2$$

  The strategy is

  - *feasible if and only if* $W_{-1} \geq \omega^p$,
  - *time consistent if and only if* $W_{-1} < \bar{\omega}^p$.

  Its payoff is $J_0^{c_1,u_0} = \frac{\bar{e}}{a\sigma^2}\left[a\sigma^2 s - \bar{e} + \sqrt{U^m}\right]$.

- ***Strategy*** c_1, c_0. *If* $W_{-1} < \omega^p$, *the arbitrageur has not enough capital to trade* $x_0^{c_1,u_0}$ *at time 0. She thus takes the largest position allowed by her capital, i.e.*

$$x_0^{c_1,c_0} = \frac{a\sigma^2 s - \bar{e} + \sqrt{Q}}{2a\sigma^2}, \text{as in Corollary 1}$$

  When $W_{-1} \in [\max(0, \omega^c), \omega^p[$, *this trade does lead to a binding constraint at time 1, which ensures dynamic consistency, under the following conditions:*

- *If* $\frac{\bar{e}}{a\sigma^2 s} \in \left[0, \frac{1}{3}\right[$, c_1, c_0 *is time consistent if and only if* $W_{-1} \in \left[\min\left(\tilde{\omega}, \omega^p\right), \omega^p\right[$ *when* $\tilde{D}^w > 0$, *and is not time consistent otherwise,*
- *If* $\frac{\bar{e}}{a\sigma^2 s} \left[\frac{1}{3}, \frac{2}{3}\right[$, c_1, c_0 *is not time consistent,*
- *If* $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{2}{3}, 2\right[$, c_1, c_0 *is time consistent if and only if* $W_{-1} \in \left[\max\left(0, \omega^c\right), \min\left(\tilde{\omega}, \omega^p\right)\right[$ *when* $\tilde{D}^w > 0$, *and is not time consistent otherwise.*
- *If* $\frac{\bar{e}}{a\sigma^2 s} \geq 2$, c_1, c_0 *is time consistent.*

The time 1 position $X_1^{c_1,c_0}$ *is given in Corollary 1. The payoff of the strategy is*

$$J_0^{c_1,c_0} = \frac{\bar{e}}{a\sigma^2}\left[a\sigma^2 s - \bar{e} + \sqrt{U}\right]$$

If the arbitrageur has enough capital ($W_{-1} \geq \omega^p$), she increases her time 0 trade up to the point where she is constrained at time 1. Note that the arbitrageur does retrade at time 1. The financial constraint can *limit* the arbitrageur's ability to provide further liquidity, but does not *eliminate* it. Indeed, the arbitrageur captures capital gains between time 0 and time 1. Thus, even if she were to use all her capital at once at time 0, she would still have the ability to retrade thanks to her interim profits. In this sense, the financial constraint is not fully equivalent to a commitment to not retrade.

Strategies with binding constraints at time 1 must satisfy additional dynamic consistency conditions. For instance, in the c_1, u_0 strategy, the arbitrageur's wealth should not be too large, e.g. $W_{-1} < \bar{\omega}^p$, to ensure that an increase in trading aggressiveness at time 0 (indeed $x_0^{c_1,u_0} > x_0^{u_1,u_0}$) does translate into a binding constraint at time 1. A large amount of capital makes it more likely that the arbitrageur will be unconstrained throughout, so that she will be tempted to "re-optimize" at time 1.

Interstingly, in the c_1, c_0 case, the dynamic consistency condition may generate an upper bound or a lower bound on arbitrageur's capital. This non-linear effect is related to the fact that in this strategy, the arbitrageur's constraint binds at time 0 and time 1, and not just at time 1. An increase in capital can have opposite effects on the arbitrageur's interim profit when she is constrained at time 0, and thus on her ability to re-optimize later. When the maximum loss per unit $\bar{e}$ is large relative to liquidity provision profits, determined by as (i.e. high $\frac{\bar{e}}{a\sigma^2 s}$), dynamic consistency imposes an upper bound on arbitrageur's initial capital. In this region, positions are so margin-intensive that it is likely that the arbitrageur will indeed be constrained at time 1, unless she has a lot capital (e.g. $W_{-1} \geq \tilde{\omega}$). When the maximum loss is smaller relative to liquidity provision profits (i.e. low $\frac{\bar{e}}{a\sigma^2 s}$), the arbitrageur may be able to re-optimize when she has a lot of capital ($W_{-1} \geq \omega^p$) or little capital ($W_{-1} < \tilde{\omega}$). This is because a small amount of capital implies a small position but also a large capital gain per unit, which results in a large interim profit, and the ability for the arbitrageur to re-optimize at time 1.

Equilibrium trade-off of a binding constraint at time 1. It is interesting to understand under which conditions the arbitrageur would like to be constrained at time 1, independently of any dynamic consistency consideration. The following result shows that the arbitrageur prefers to be constrained at time 1 when her capital is not too large.

Lemma 5 *Denote* $v_1 \equiv \left(1 + \frac{1}{\sqrt{5}}\right) s\bar{e} - \frac{9}{10} a\sigma^2 s^2$ *and* $v_2 \equiv \left(1 - \frac{1}{\sqrt{5}}\right) s\bar{e} - \frac{9}{10} a\sigma^2 s^2$*. It holds that*

1. *The* u_1, u_0 *strategy is more profitable than the* c_1, u_0 *strategy if and only if* $W_{-1} \leq v_2$ *or* $W_{-1} \geq v_1$*,*
2. *For all parameters,* $v_2 < \omega_1^m < v_1$*.*

This lemma implies that v_2 is not a relevant threshold to assess the benefits of the c_1, u_0 strategy, since the unconstrained strategy is not feasible for $W_{-1} < \omega_1^m$. Thus the arbitrageur prefers to be unconstrained at time 1 when her capital is above the threshold $v_1 > \omega_1^m$. Intuitively, when the arbitrageur is well capitalized, being constrained at time 1 entails a high opportunity cost relative to the additional profit that can be captured at time 0, and is thus better off following the u_1, u_0 strategy.

4.4 Equilibrium when capital is more abundant

4.4.1 Main regions

There are four main regions, corresponding to the different strategies. The main drivers of the equilibrium are the amount of capital W_{-1} and the ratio $\frac{\bar{e}}{a\sigma^2 s}$. Intuitively, this ratio measures the maximum loss per unit, relative to the local investors' hedging demand, which determines the price wedge, and thus the profitability of the arbitrage. In other words, the ratio balances the availability of funding and the potential profits that the arbitrageur can earn from intermediating trades. The ratio determines the order of the thresholds v_1, ω_0^m, ω_1^m, ω^p, etc. which condition feasibility, dynamic consistency, or payoff rankings, generating a large number of parameter cases. The equilibrium is represented in Figures 2 and 3.

In a nutshell, the equilibrium is c_1, c_0 when capital is abundant and/or the ratio $\frac{\bar{e}}{a\sigma^2 s}$ is low enough, i.e. when the trade is not very risky and quite profitable. In the opposite corner of the figures, i.e. if capital is low or $\frac{\bar{e}}{a\sigma^2 s}$ is high, the arbitrageur is constrained (region c_1, c_0 under the red curve, see also Proposition 2). In the intermediate region, the arbitrageur's constraint binds either at time 0 or time 1, or both. When capital is large enough, the arbitrageur's constraint binds at time 1. As seen in Lemma 4, the arbitrageur cannot play the c_1, u_0 strategy unless she has not enough capital to trade more aggressively at time 0 ($W_{-1} \geq \omega^p$). Instead, if there is less capital, the arbitrageur must free up capital by reducing x_0 to remain unconstrained at time 1, and may be constrained at all dates. In the next two sections, I focus on two interesting parameter regions. The full analytic characterization of the equilibrium is available in Appendix B.5.5.

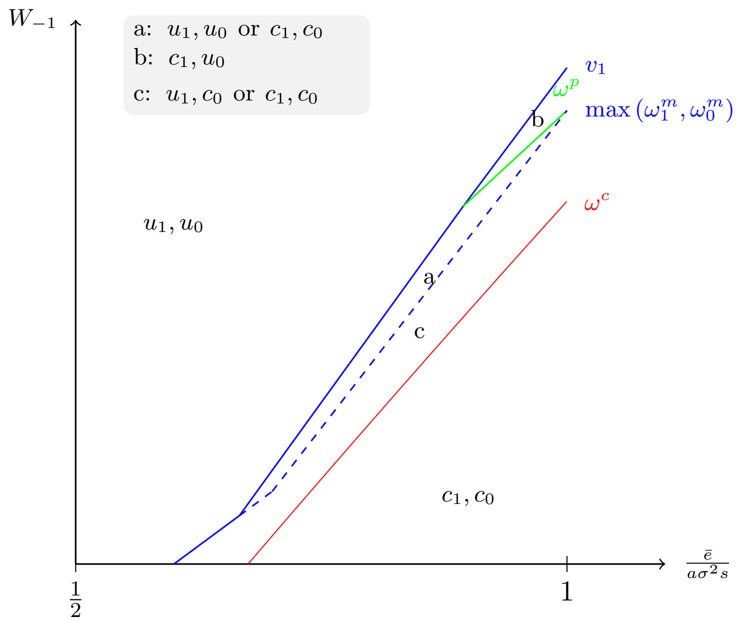

Figure 2: Equilibrium when $\frac{\bar{e}}{a\sigma^2 s} < 1$

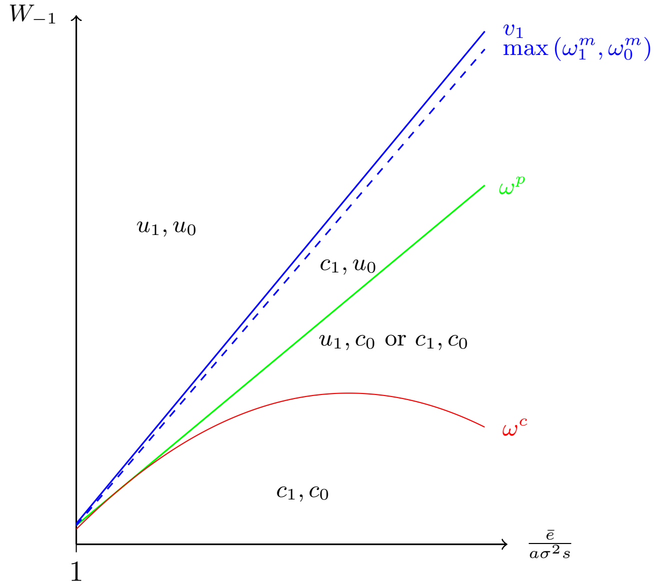

Figure 3: Equilibrium when $\frac{\bar{e}}{a\sigma^2 s} \geq 1$

4.4.2 Equilibrium with abundant capital: unconstrained trading

Proposition 3 *The following holds:*

1. *If* $W_{-1} \geq \bar{\omega}^m = \Lambda s\bar{e} - \Gamma a\sigma^2 s^2$*, there exists a unique equilibrium in which the financial constraint is slack at* $t = 0$ *and* $t = 1$*, the arbitrageur's trades in asset A are* $x_0^{u_1,u_0} = \frac{2}{5}s$ *and* $x_1^{u_1,u_0} = \frac{3}{10}s$ *and the equilibrium spreads are* $\Delta_0^m = \frac{9}{5}a\sigma^2 s$ *and* $\Delta_1^m = \frac{3}{5}a\sigma^2 s$*.*

2. Λ *and* Γ *are such that* $0 < \Lambda < 2$ *and* $0 < \Gamma < 1$*, and depend on the ratio* $\frac{\bar{e}}{a\sigma^2 s}$ *as follows:*

$$\begin{cases} \Lambda = \frac{4}{5} \text{ and } \Gamma = \frac{12}{25} \text{ i.e. } \bar{\omega}^m = \omega_0^m \text{ if } \frac{\bar{e}}{a\sigma^2 s} \in \left[0, \frac{21}{10\left(1+\sqrt{5}\right)}\right[\\ \Lambda = 1 + \frac{1}{\sqrt{5}} \text{ and } \Gamma = \frac{9}{10} \text{ i.e. } \bar{\omega}^m = v_1 \text{ if } \frac{\bar{e}}{a\sigma^2 s} \geq \frac{21}{10\left(1+\sqrt{5}\right)} \end{cases}$$

3. *Relative to the benchmark competitive case:*

   - *Market liquidity is lower:* $\Delta_t^m > \Delta_t^*$, $t = 0, 1$.
   - *The monopolistic arbitrageur remains unconstrained with a lower initial capital:* $\bar{\omega}^m < \omega^*$*, for all* $\bar{e}$, a, σ, s.
   - *Prices converge more slowly:* $\Delta_0^m > \Delta_1^m > \Delta_2^m = 0$*, while* $\Delta_0^* = \Delta_1^* = \Delta_2^* = 0$.

This region is above the thick blue line in Figures 2 and 3. It is comparable to the first region of the competitive case ($W_{-1} \geq \omega^*$). There are three noticeable differences. First, although the arbitrageur is unconstrained, assets A and B trade at a spread, i.e. liquidity is imperfect, and prices converge gradually, as opposed to instantaneously in the competitive economy. This is simply due to the arbitrageur's market power. Given the absence of competition, the arbitrageur limits the amount she buys from local investors with low valuation for the asset and sells to those with high valuation. This keeps the spread open in equilibrium, which allows the arbitrageur to make a profit.[14] Further, to minimize her price impact, the arbitrageur trades "slowly". Given that providing liquidity early on reduces liquidity demand later on, the arbitrageur trades to equalize the marginal benefit of liquidity provision at time 0 and the marginal cost of liquidity reduction at time 1. Instead competitive arbitrageurs trade until the marginal benefit breaks even in each period, ignoring their intertemporal impact on the price. This results in gradual price convergence in the monopolistic economy, and immediate convergence in the competitive economy.

Second, the financial constraint is no longer linear in the dispersion of the fundamental $\bar{e}$ and the liquidity shock s. In fact, it is now quadratic in s[15], so that the following comparative statics obtains:

[14] Indeed in equilibrium the arbitrageur's position is $X_t < s$. (note that $x_0 = X_0$)

[15] It is also quadratic in $\bar{e}$ since σ^2 is a function of $\bar{e}^2$.

Corollary 2 *The threshold $\bar{\omega}^m$ features the following comparative statics:*

- *If $\frac{\bar{e}}{a\sigma^2 s} \in \left[0, \frac{21}{10\left(1+\sqrt{5}\right)}\right[$, i.e. if volatility is high enough, then a small increase in the liquidity shock s loosens the financial constraint,*
- *If $\frac{\bar{e}}{a\sigma^2 s} > \frac{21}{10\left(1+\sqrt{5}\right)}$, i.e. if volatility is low enough, a small increase in s tightens the financial constraint.*

The intuition for this result is simple. On one hand, an increase in the dispersion of the fundamental $\bar{e}$ increases its volatility σ, which in turn increases the magnitude of the potential divergence from fundamental in the next period and makes a default by the arbitrageur more likely from the viewpoint of financiers. This tightens the financial constraint, an effect akin to the competitive case. On the other hand, under our modeling assumptions, the financiers (implicitly) recognize that an increase in volatility is equivalent to an increase in the willingness of local investors to share their risk and to accept large price concessions. This increases the profitability of the arbitrage strategy and allows the arbitrageur to capture larger rents. The arbitrageur reaps larger capital gains, which relaxes the financial constraint. This second effect relies on the assumption that financiers can distinguish between fundamental and liquidity effects, which generates countercyclical (stabilizing) margins.[16]

Bearing this simple trade-off in mind, it is easy to interpret $\bar{\omega}^m$ as the sum of two terms:

$$\bar{\omega}^m = \underbrace{\Lambda s \bar{e}}_{\text{maximum position loss}} - \underbrace{\Gamma a \sigma^2 s^2}_{\text{profit adjustment}}$$

The first term, $\Lambda s \bar{e}$, represents the maximum loss on the position caused by a change in the fundamental and is therefore a multiple of $\bar{e}$, which measures the largest possible change in the fundamental. Note that it depends on the arbitrageur's position , which is less than $2s$. By contrast, in the competitive case, $\omega^* = 2s\bar{e}$, because arbitrageurs fully absorb the liquidity shock affecting market A and B, which is $2s$ in total (corresponding to a position of size s in each leg of the arbitrage). The second term in $\bar{\omega}^m$, $-\Gamma a \sigma^2 s^2$, is an adjustment measuring how much past or future profits due to rent extraction lower the capital requirement. It is thus specific to the monopoly case. In the competitive economy, financiers anticipate that perfect competition drives profits to zero, and therefore there is no profit adjustment.

The third noticeable difference is that there are two different regions for the threshold $\bar{\omega}^m$, while there is a unique threshold ω^* in the competitive benchmark. These regions can be expressed in

[16] Brunnermeier and Pedersen (2008)'s model nests both stabilizing and destabilizing margins. With destabilizing margins, the second effect would remain but would bite less.

terms of low or high volatility. For instance, using the four-point distribution given in Section 2, the ratio $\frac{\bar{e}}{a\sigma^2 s}$ can be expressed as follows:

$$\frac{\bar{e}}{a\sigma^2 s} = \frac{\bar{e}}{a\bar{e}^2\left(1+2p\left(\frac{1}{\mu^2}-1\right)\right)s} = \frac{1}{a\bar{e}\left(1+2p\left(\frac{1}{\mu^2}-1\right)\right)s}$$

For simplicity, I will refer to a situation with large fundamental dispersion $\bar{e}$ as high volatility and small fundamental dispersion as low volatility.[17] Note that in the high volatility region, the time-0 feasibility constraint is the binding constraint, i.e. $\bar{\omega}^m = \omega_0^m$. In the region of low volatility, the binding constraint is the one ensuring that the unconstrained strategy u_1, u_0 dominates the voluntarily constrained strategy c_1, u_0, i.e. $\bar{\omega}^m = v_1$. The interpretation is in terms of which effects of the maximum position loss and the profit adjustment dominates. When volatility is high, the profit adjustment is large and therefore $v_1 < \omega_0^m$, because by definition, v_1 takes into account all expected profits, while ω_0^m reflects only one period expected profits. This can be seen from equation 1, which shows that margins depend on the change in liquidity (and thus capital gains) one period ahead only. It is the opposite in the low volatility region, where the profit adjustment is small, meaning that the feasibility constraint is not the binding constraint.

4.4.3 Equilibrium with intermediate level of capital: voluntarily-constrained trading

When the arbitrageur has less capital, she may credibly choose to be constrained at time 1.

Proposition 4 *If* $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{2\sqrt{5}}{5}, 1\right[$ *and* $W_{-1} \in [\omega^p, v_1[$, *or if* $\frac{\bar{e}}{a\sigma^2 s} > 1$ *and* $W_{-1} \in [\omega_1^m, v_1[$, *then:*

1. *The unconstrained strategy* u_1, u_0 *is feasible but is not played in equilibrium;*

2. *In the unique equilibrium, the financial constraint is slack at* $t = 0$ *and binding at* $t = 1$, *i.e. the arbitrageur is voluntarily constrained at time 1 by playing the* c_1, u_0 *strategy.*

3. *The equilibrium spreads are:*

$$\Delta_0^{c_1,u_0} = 2a\sigma^2 s + \bar{e} - \sqrt{U^m} \text{ and } \Delta_1^{c_1,u_0} = a\sigma^2 s + \bar{e} - \sqrt{U^m}$$

The proposition corresponds to region b in Figure 2 and the region between the full blue curve and the dashed blue curve in Figure 3. The result shows that the arbitrageur voluntarily chooses to be constrained when volatility is low enough and her level of capital is intermediate. The reason why being constrained at time 1 might be optimal is related to the Coasian dynamics of the model. Intuitively, the arbitrageur chooses her trading strategy to keep the spread open as long as possible.

[17] Equivalently, low volatility can stem from low risk in the tail (large μ or low p), so that we could rephrase the analysis in terms of large or small tail risk. Note that because the distribution is symmetric, tail risk equally includes good and bad events.

Since the asset matures only at time 2, local investors have some freedom to chose the date at which they consume liquidity. They rationally anticipate that after providing liquidity at time 0, the arbitrageur will provide further liquidity at time 1, further decreasing the spread. Hence providing liquidity early, at time 0, reduces the profitability of later liquidity provision for the arbitrageur, unless she is able to credibly commit to keep the spread large in the future. To this extent, the financial constraint works as a commitment device for the arbitrageur, who can then extract larger rents at time 0. Indeed, in equilibrium, local investors anticipate that the arbitrageur's constraint binds time 1, which increases their willingness to accept large price concessions at time 0, increasing the potential price gap and thus the arbitrageur's capital gain.

When does this occur in equilibrium? The conditions on capital and volatility given in Proposition 4 have an intuitive interpretation. First, the arbitrageur's capital must be below some threshold $v_1 < \bar{\omega}^p$, which guarantees that i) the opportunity cost of being constrained at time 1 is low enough, and ii) the arbitrageur cannot re-optimize (by Lemma 4). If capital is too low, however, the arbitrageur cannot serve the additional liquidity demand at time 0 and cannot benefit from committing to be constrained. These conditions on arbitrageur's capital are combined with the requirement that the volatility be low enough, i.e. that $\frac{\bar{e}}{a\sigma^2 s}$ is high enough. Intuitively, if this was not the case, the unconstrained strategy would be so profitable that the opportunity cost of being constrained would be too large.

At a deeper level, one may wonder how it is possible for the constraint to be binding only at one date in equilibrium, while in the competitive case, either the constraint binds at all dates or never. This point is related to the trade-off between position funding and profit adjustment. The position funding effect depends on the size of the arbitrageur's position. The profit adjustment depends on expected profits. As time passes, the position increases, and therefore the constraint should tighten. But at the same time, the profit adjustment also increases, so that the constraint at time 1 may be less severe than the constraint at time 0. When arbitrageurs are competitive, they collectively fully absorb the liquidity shock in each period, so that their total position is always $2s$. Further, perfect competition eliminates the arbitrageurs' profits, hence there is no profit adjustment, and the constraint either binds all the time or never.

When the arbitrageur chooses to be constrained at time 1, she trades more aggressively at time 0 than in the unconstrained case: she sets up a trade $x_0 = \frac{s}{2}$ instead of $\frac{2}{5}s$. Conversely, her time-1 trade is lower than if she were unconstrained.

Corollary 3 *In the voluntarily constrained equilibrium given in Proposition 4:*

1. *The arbitrageur trades a larger quantity at time 0 and a smaller at time 1, than if she were using the (feasible) unconstrained strategy, i.e. in the relevant parameter space,*

$$x_0^{c_1,u_0} > x_0^{u_1,u_0} \text{ and } x_1^{c_1,u_0} < x_1^{u_1,u_0}$$

2. *The overall effect is that the arbitrageur builds a larger position in the voluntarily constrained equilibrium than if she chose to remain unconstrained:*

$$X_1^{c_1,u_0} \geq X_1^{u_1,u_0}$$

Implications for market liquidity and empirical predictions. The arbitrageur trades more aggressively in this equilibrium than when she is unconstrained. But her behaviour is motivated by the fact that local investors shift liquidity demand towards the first period, pushing the prices of assets A and B further apart. Given these conflicting effects, it is natural to analyze the overall impact on the equilibrium spread. The following result shows that the increased trading aggressiveness always dominates:

Corollary 4 *The spread is lower at all dates when the arbitrageur chooses to be constrained in equilibrium, i.e.* $\Delta_t^{c_1,u_0} < \Delta_t^m$*,* $t = 0, 1$.

This improvement in liquidity means either that the arbitrageur more than compensates the additional liquidity demand at time 0, or that her trades have a larger price impact than in the unconstrained case. Even more surprising is the result that the liquidity improves at all dates: this is related to the fact that trades have a permanent impact on the price. Moreover, the arbitrageur acquires a larger position than if she was unconstrained.

More generally, this result implies that a drop in arbitrageur's capital may not have a monotonically decreasing effect on market liquidity. If volatility is low enough, a reduction in the arbitrageur's capital may first leave market liquidity unchanged (if $W_{-1} \geq \bar{\omega}^m$), then improve it and decrease it again later. This is in contrast to the competitive case.

This result has implications for the amount of capital monopolistic arbitrageurs or market-makers should hold. NYSE specialists can be seen as real-world counterparts to our monopolistic arbitrageur, although their ability to exert market power may be limited by other institutional features not considered here, e.g. competition from limit orders by off-exchange investors, or from investment banks in the upstairs market.[18] Given this caveat, one can use effective spreads as a proxy for market liquidity as in Comerton-Forde, Jones, Hendershott, Moulton, Seasholes (2010).[19] The model prediction is that for firms with low enough dividend volatility, effective spreads should increase in the amount of capital available to the specialists running the stock, when capital is

[18] Note however that sequential arrival and execution costs can limit competition from limit orders, as in Seppi (1997).

[19] To strengthen the analogy with market-making, consider that market A and B at time t may represent two subperiods t^A and t^B of date t. Thus A-investors may come to the market at time t^A to share the risk of their liquidity shocks, while B-investors with whom there would be gains from trade arrive only later at time t^B. Arbitrageurs fill the gap between the two subperiods, providing immediacy as market-makers smoothing out order imbalances. See Gromb and Vayanos (2010) for further details on this alternative interpretation of the model.

intermediate. For capital low enough, effective spreads should be decreasing in the amount of specialists' capital, and even more so if specialist capital is low, for any level of dividend volatility. For a high level of capital, there should be no significant relation between spreads and the level of capital. Comerton-Forde et al. find evidence for the effects in the constrained region, but base their tests on the assumption that there are two regions for market-maker capital (as in the competitive case), and not three as in the monopolistic case. Instead of the cross-section, the test may be run in the time-series, i.e. by comparing spreads in times where specialists appear to be more constrained than others.

5 Market structure, liquidity provision and welfare

When the competitive economy is unconstrained, because capital is abundant, it is clear that the market is more liquid than in the monopolistic economy. However, given that market power relaxes the arbitrageur's margin constraint in equilibrium, it is natural to ask whether the monopoly can provide more liquidity than a constrained competitive market when capital is relatively scarce. In this section, I show that a monopoly - whether it is unconstrained or voluntarily constrained - may provide more liquidity than a constrained competitive market but only at time 1, just before the asset matures. In addition to these comparisons across market structures, I present comparisons within market structures. These results may shed light on different types of policies: long-term policies aimed at opening markets to competition, and short-term policies aiming to affect margin or capital requirements in a given type of market.

5.1 Liquidity provision across market structures

5.1.1 Constrained perfect competition vs unconstrained monopoly

Given that the thresholds $\bar{\omega}^m$ and ω_1^m associated with the monopoly are lower than the threshold ω^* of the competitive market, there is a parameter region in which the competitive market is constrained but the monopoly is unconstrained in equilibrium. I denote Δ_t^* and Δ_t^m the spreads at time t in the competitive and monopoly cases, respectively.

Proposition 5 *Liquidity under different market structures:*

1. *At time 0: Suppose that $W_{-1} \in [\bar{\omega}^m, \omega^*[$. The constrained competitive market features more liquidity than an unconstrained monopoly at time 0 if $\frac{\bar{e}}{a\sigma^2 s} < \frac{21}{10\left(1+\sqrt{5}\right)}$ or if $\frac{\bar{e}}{a\sigma^2 s} \geq \frac{21}{10\left(1+\sqrt{5}\right)}$ and $W_{-1} \in [\tilde{\omega}, \omega^*[$ with $\tilde{\omega} > \bar{\omega}^m$.*
2. *At time 1: Suppose that $W_{-1} \in [\omega_1^m, \omega^*[$ and $h\left(W_{-1}\right) \geq 0$ or $h\left(W_{-1}\right) < 0$ and $f\left(W_{-1}\right) \geq 0$, then*

- *If* $\frac{\bar{e}}{a\sigma^2 s} \geq 7+q$*, then if* W_{-1} $[\omega_1^m, \hat{\omega}[$*, the unconstrained monopoly provides more liquidity than the constrained competitive market, and less if* $W_{-1} \in [\hat{\omega}, \omega^*[$*,*
- *If* $\frac{\bar{e}}{a\sigma^2 s} < 7+q$*, then constrained competitive market always provide more liquidity than the unconstrained monopoly.*

Since the ratio $\frac{\bar{e}}{a\sigma^2 s}$ can be rewritten as a function of $\bar{e}$, a and s only, we can discuss the result in terms of high and low volatility regions (or dispersion of fundamental $\bar{e}$). To understand the result, note that volatility has different effects in the constrained competitive case and the unconstrained monopoly. For the unconstrained monopoly, volatility has an unequivocal positive effect. It increases local investors' demand for liquidity, making the arbitrage opportunity more profitable. Thus the spread increases with volatility σ^2 (and thus with $\bar{e}$). For the constrained competitive market, volatility has two opposite effects: first, it increases local investors' demand for liquidity, as in the monopoly case, thus pushing asset prices apart. Second, by making the arbitrage more profitable, it can increase the intermediate capital gain and relax the financial constraint.

As a consequence, in the high volatility region (low $\frac{\bar{e}}{a\sigma^2 s}$), the unconstrained monopoly is less liquid than the constrained competitive economy. This is because the constrained competitive economy benefits from the softening effect of volatility on the financial constraint. In the low volatility region, this effect is reduced, and thus there is less liquidity in the monopoly case only if arbitrageurs' capital is large enough at time 0, and there can be more liquidity in the monopoly case at time 1. Intuitively, the intermediate capital gain is small in this case, thus the constrained competitive economy remains severely constrained at time 1. The condition on capital is intuitive, since when competitive arbitrageurs are constrained, the spread decreases in the amount of capital they hold. Note that this result obtains only when considering ω_1^m instead of v_1 as a lower threshold for an unconstrained monopoly. This requires that the conditions $h(W_{-1}) \geq 0$ or $h(W_{-1}) < 0$ and $f(W_{-1}) \geq 0$ be satisfied. In numerical examples, these conditions seem easy to meet.

My results do not rule out the possibility of having more liquidity in the monopoly case also at time 0. However, even if I do not have an analytical proof, I have not been able to generate this case numerically, suggesting that liquidity improvement may occur only at time 1.

5.1.2 Constrained perfect competition vs voluntarily constrained monopoly

When capital is relatively abundant but close to the constrained region, the monopolistic arbitrageur may find it optimal to be constrained at time 1. I showed that this decreases the spread relative to the unconstrained case in Section 4. Hence from Proposition 5, one would expect that the monopoly provides more liquidity than the constrained competitive market at least at time 1. The following result confirms this conjecture.

Proposition 6 *Under the conditions of Propositions 4, the voluntarily-constrained monopolist provides more liquidity at time 1 than the constrained competitive market.*

Given Corollary 4, it is not surprising to see that the condition for $\Delta_1^* \geq \Delta_1^{c_1,u_0}$ is easier to satisfy than in the unconstrained monopoly case. In particular, there is no condition on arbitrageurs' capital, although we are looking at the same region with $W_{-1} \geq \omega_1^m$. The result, however, holds only at time 1. At time 0, I show in the proof that for the monopoly to provide more liquidity at time 0, a non-trivial condition on parameters must be satisfied. In numerical examples, I have always found a larger spread in the monopoly case than in the constrained competitive case. This is confirmed by the following result:

Corollary 5 *Under the conditions of Proposition 4,*

- *the voluntarily constrained monopoly captures the largest possible intermediate capital gain, $x_0\left(\Delta_0 - \Delta_1\right)$, by limiting liquidity more than the constrained competitive market: $x_0^{c_1,u_0} \leq x_0^{c_1,c_0}$.*
- *$\Delta_0 - \Delta_1$ is larger in the voluntarily constrained case than in the constrained competitive case*
- *$\Delta_0^{c_1,u_0} \geq \Delta_0^{c_1,c_0}$*
- *The speed of arbitrage is higher under the voluntarily constrained monopoly than the constrained competitive case at time 0, and lower at time 1.*

It may be surprising that the competitive market yields a tighter spread at time 0 and a larger one at time 1, all the more than the model features permanent price impact, implying that a large spread at time 0 should translate into a large spread at time 1. However, the intuition is simple. The monopoly improves liquidity at time 1 relative to competitive arbitrageurs because she captures a larger intermediate capital gain. The larger capital gain, $x_0\left(\Delta_0 - \Delta_1\right) = 2a\sigma^2 x_0\left(s - x_0\right)$, follows precisely from the fact that the monopoly limits liquidity more at time 0 by buying a smaller amount, which causes the spread to be larger at time 0 than in the competitive case. In particular, I show in the proof of the Corollary that the intermediate capital gain in the monopoly case, U^m, is greater than that of the competitive case, U, and that this implies $\Delta_1^* \geq \Delta_1^{c_1,u_0}$.

5.2 Welfare

5.2.1 Across market structures: monopoly vs constrained competitive market

Given that liquidity may improve at time 1 when the market is monopolistic, it is natural to study whether investors' welfare improves. As a first step, I calculate the expression of local investors' welfare as a function of the spreads.

Lemma 6 *Let χ_0^A denote A-local investors' welfare, and let autarky ($\chi_0^{A,a}$) define a situation without arbitrageur ($n = 0$), i.e. where there is no trade across markets, and full insurance (χ_0^{A*}) the situation where a continuum of unconstrained competitive arbitrageurs trade across markets.*

Then we have:

$$\begin{aligned}\chi_0^A &= \frac{(\Delta_0 - \Delta_1)^2 + \Delta_1^2}{8a\sigma^2} - \frac{s}{2}\Delta_0, \\ \chi_0^{A,a} &= -a\sigma^2 s^2 < \chi_0^{A,m} \leq \chi_0^{A,*} = 0\end{aligned}$$

The arbitrageur's profit is larger in the monopoly case than in autarky or full insurance cases.

As expected, for local investors, autarky and full insurance form two polar cases, and the monopolistic case is somewhere in the middle. In autarky, local investors have no options to hedge, and their certainty equivalent is minimal. When there is a continuum of unconstrained competitive arbitrageurs, local investors can trade the asset at its fair value, and can access a perfect hedge thanks to arbitrageurs' intermediation to market B, resulting in perfect insurance. When there is a monopolistic arbitrageur (whether she is constrained or not), local investors receive some imperfect insurance as the market is imperfectly liquid. To understand how the investors' welfare with an unconstrained monopoly ($\chi_0^{A,m}$) compares to a constrained competitive market ($\chi_0^{A,c}$) when we place ourselves under the conditions of Propositions 5 and 6, we could directly compare welfare. However, it is difficult to derive analytical results. Thus I use an indirect approach based on comparative statics.

Corollary 6 *The following holds:*

- *Local investors' welfare decreases with Δ_0 if and only if $x_0 > 0$, $\frac{\partial \chi_0^A}{\partial \Delta_0} < 0 \Leftrightarrow x_0 > 0$*
- *If $\Delta_1 < \frac{1}{2}\Delta_0$, then local investors' welfare decreases with the time-1 spread and with a decrease in liquidity: $\frac{\partial \chi_0^A}{\partial \Delta_1} < 0$, $\frac{\partial \chi_0^A}{\partial (\Delta_0 - \Delta_1)} < 0$, and $\frac{\partial \chi_0^A}{\partial (\Delta_1 - \Delta_2)} < 0$*

An immediate implication of this result is that, if Δ_1 is small enough relative to Δ_0, the improvement in liquidity at time 1 - because it is due to a larger difference $\Delta_0 - \Delta_1$ as numerical results and analysis suggest - may not be Pareto improving. Put differently, switching from a constrained competitive market to a monopolistic market unambiguously increases arbitrageurs' welfare but may decrease local investors' if the improvement in liquidity at time 1 does not offset the worsening at time 0:

Corollary 7 *If $\Delta_1 < \frac{1}{2}\Delta_0$, and the conditions of Proposition 6 are satisfied, switching from a competitive market to a monopolistic market increases the arbitrageur's welfare but can decrease local investors' welfare. Aggregate welfare may rise.*

Intuitively, the improvement in liquidity at time 1 is associated with a decrease in the improvement of liquidity between time 0 and time 1, measured by $\Delta_0 - \Delta_1$, and a quicker improvement between time 1 and time 2, $\Delta_1 - \Delta_2$.[20] Under the conditions of Corollary 7, the first effect can outweigh the second. There are two reasons why it can be the case. i) Since local investors experience two shocks, they face higher risks at time 0 (conditionally), thus receiving liquidity at time 0 matters more than receiving liquidity at time 1. ii) The improvement in liquidity at time 1 may require a large worsening at time 0, implying that $\frac{d\chi_0^A}{d\Delta} = \frac{\partial \chi_0^A}{\partial \Delta_0} d\Delta_0 + \frac{\partial \chi_0^A}{\partial \Delta_1} d\Delta_1 \leq 0$ (see numerical example below). The condition for the result of Corollary 6 is that $\Delta_1 < \frac{1}{2}\Delta_0$. In practice it seems verified. In many numerical examples, including the one reported below, the spread at time 1 is somewhere between a third and a half of the spread at time 0. I prove the result on aggregate welfare on a numerical example:

Aggregate welfare. It is interesting to study the aggregate effects of the change in market structure. Although some redistribution effects may be negative, aggregate welfare may increase with a change in market structure (This, of course, takes into account both A and B local investors). For instance, assume that ϵ_t follows the example distribution described in Section 2 and consider the following parameter values: $a = 9$, $s = 0.1$, $\bar{e} = 1$, $p = 0.48$, $\mu = 150$, and set the arbitrageur's capital W_{-1} to $\omega_1^m \approx 0.1368$. We are then under the conditions of Proposition 6, and if the market is competitive, the arbitrageurs would be constrained, since $\omega^* = 0.2$. The equilibrium spreads in the monopolistic (voluntarily-constrained) structure are $\Delta_0^{c_1,u_0} \approx 0.058$, and $\Delta_1^{c_1,u_0} \approx 0.0216$. The spreads in the constrained competitive case are $\Delta_0^{c_1,c_0} \approx 0.044$ and $\Delta_1^{c_1,c_0} \approx 0.0217$, implying that the spread at time 1 is less than a half of that of time 0. Comparing the market structure shows that the improvement in liquidity at time 1 is moderate relative to the deterioration at time 0. (observe that we assumed a value of the arbitrageur's capital at the low end of the possible range) This implies that local investors' welfare decreases from -0.0019 to -0.0023. Instead arbitrageur's profit increases 0.1398 to 0.1401. The total effect (taking into account both markets A and B) is positive: 0.013.

5.2.2 Within market structure: unconstrained vs voluntarily constrained monopoly

Last, I highlight some welfare effects of imposing constraints on a monopolistic arbitrageur. When capital is very scarce or particularly abundant, financial constraints have the same effect as in the competitive case. I.e., when capital is abundant, liquidity, although limited because of arbitrageur's market power, is insensitive to the arbitrageur's wealth. Instead, when capital is very scarce, the extra slack caused by financiers' recognition that the monopolist makes profits in equilibrium, is not sufficient and liquidity and welfare are as in the constrained competitive case.

A more interesting situation is when capital is intermediate and volatility bounded above, as

[20] Recall that $\Delta_2 = 0$ and that $\Delta_0 > \Delta_1$.

in Proposition 4. Then the arbitrageur chooses to be constrained, although she has enough capital to remain unconstrained. Since in equilibrium, spreads are tighter when the monopoly chooses to be constrained than if she remained unconstrained, a straightforward implication of Corollary 6 is that local investors' welfare increases when the monopoly chooses to be constrained.

Corollary 8 *Imposing financial constraints when parameters satisfy the conditions in Proposition 4 (voluntarily constrained equilibrium) is Pareto-improving.*

An implication of this result is that imposing margin or capital requirements on arbitrageurs has differential effects depending on the market structure. When the market is competitive, binding financial constraints lead to transfers from local investors to arbitrageurs. When the market is monopolistic, binding constraints lead to similar transfers when capital is very scarce and improve all market participants' welfare when capital is intermediate and asset volatility is not too large.

6 Conclusion

In this paper, I studied a model of financially constrained arbitrage under two polar market structures: competitive and monopolistic. I showed that the interaction between market liquidity and funding liquidity is profoundly affected when the assumption of price-taking behavior is relaxed for the arbitrageur. Market power may to some extent alleviate the arbitrageur's margin constraint. As a result, an arbitrageur with price impact does not always provide less liquidity than many small arbitrageurs holding collectively the same amount of capital. Another important conclusion from the paper is that margin constraints can under stated conditions improve liquidity and welfare under a monopolistic structure, while they decrease liquidity and redistribute wealth at the expenses of liquidity consumers in competitive markets.

The current framework can be seen as first step towards a comprehensive analysis of several important policy issues, such as the optimal capital structure of financial intermediaries, systemic risk and the effects of implicit government protection. My analysis implies that an arbitrageur with market power may want to commit to decrease her capital level in the future, e.g. by distributing dividends. The degree of competition among arbitrageurs should thus result in different capital structures and financial policies.

The modeling of the financial constraints precludes defaults in equilibrium. An interesting extension would thus to allow for the possibility of default. The introduction of a third party, e.g. a government, would allow for a discussion of the effects of systemic risk and bailouts under different market structures. An intermediate step may consist in extending the model to risky arbitrage. I have studied a textbook situation in which the arbitrage is risk-free. In practice, arbitrage strategies such as relative-value and convergence trading entail risk. Gromb and Vayanos (2002) show that in this case competitive arbitrageurs may not take the efficient level of risk, as they fail to internalize the effects of their strategies on others' financial constraints. With imperfect competition, one can expect that arbitrageurs would to some extent internalize the impact of their decisions, even though this would also decrease efficiency. These extensions are left for future research.

Appendix

A Competitive equilibrium

A.1 Proposition 1

Proof. The result is a special case of Proposition 1 in Gromb and Vayanos (2002) with $T = 2$, and $f'(x) = a\sigma^2 x$. ■

A.2 Corollary 1

Proof. I solve the system of equations (4)-(5). Rearranging terms in equation (4) gives:

$$2a\sigma^2 x_0^2 + 2\left(\bar{e} - a\sigma^2 s\right) x_0 - W_{-1} = 0$$

Since $2a\sigma^2 > 0$ and $-W_{-1} \leq 0$, the unique positive root is

$$x_0 = \frac{a\sigma^2 s - \bar{e} + \sqrt{Q}}{2a\sigma^2}, \text{ with } Q = \left(\bar{e} - a\sigma^2 s\right)^2 + 2a\sigma^2 W_{-1}. \tag{12}$$

Similarly, the time-1 position follows from equation (5):

$$a\sigma^2 X_1 + \bar{e} - a\sigma 2s - x_0\bar{e} = 0 \tag{13}$$

$a\sigma^2$ and $-x_0\bar{e}$ have opposite signs, hence the unique positive root is

$$X_1 = \frac{a\sigma^2 s - \bar{e} + \sqrt{U}}{2a\sigma^2}, \text{ with } U = \left(\bar{e} - a\sigma^2 s\right)^2 + 4a\sigma^2 x_0 \bar{e}.$$

Equilibrium spreads are obtained from the first-order conditions of local investors' maximization problems. At time 0, $a\sigma^2\left(x_0^A + s\right) = \mathbb{E}_0\left(p_1^A\right) - p_0^A = \frac{\Delta_0 - \Delta_1}{2}$. Similarly at time 1, $a\sigma^2\left(X_1^A + s\right) = \mathbb{E}_1\left(p_2^A\right) - p_1^A = \frac{\Delta_1 - \Delta_2}{2} = \frac{\Delta_1}{2}$. Market-clearing for asset A requires $y_0^A + x_0 = 0$, and $Y_1^A + X_1 = 0$, hence

$$\Delta_1 = 2a\sigma^2\left(s - X_1\right) \tag{14}$$

$$\frac{\Delta_0}{2} = a\sigma^2\left(s - x_0\right) + a\sigma^2\left(s - X_1\right) \tag{15}$$

Substituting equations (12) and (13) into (14) and (15) yields the equilibrium spreads in the proposition. ■

B Monopoly equilibrium

B.1 Lemma 2

Proof. At time 1, local investors' problem in market A is:

$$\begin{aligned}
\chi_1^A &= \max_{y_1^A} U_1 = \mathbb{E}_1\left(W_2^A\right) - \frac{a}{2}\mathbb{V}_1\left(W_2^A\right) \\
&\text{s.t.} \quad W_2^A = E_1^A + Y_1^A D_2
\end{aligned}$$

Substituting W_2^A into the objective function, and using the law of motion of E_1^A leads to:

$$\chi_1^A = \max_{y_1^A} E_0^A + Y_1^A D_1 - y_1^A p_1^A - \frac{a\sigma^2}{2}\left(Y_1^A + s_1\right)^2$$

From the first-order condition, $a\sigma^2\left(Y_1^A + s_1\right) = D_1 - p_1^A$, and market-clearing, $Y_t^A + X_t = 0$, the price schedule faced by the arbitrageur in market A is $p_1^A\left(X_1\right) = D_1 - a\sigma^2 s_1 + a\sigma^2 X_1$. By symmetry, in market B: $p_1^B\left(X_1\right) = D_1 + a\sigma^2 s_1 - a\sigma^2 X_1$. With $\Delta_1 = p_1^B - p_1^A$, this gives:

$$\Delta_1\left(X_1\right) = 2a\sigma^2\left(s - X_1\right) \tag{16}$$

The arbitrageur takes the price schedule as given in her maximization problem:

$$\begin{aligned}
J_1 &= \max_{x_1} \mathbb{E}_1\left(W_2^i\right) - \frac{b}{2}\mathbb{V}\left(W_2^i\right) \\
s.t.\ W_2 &= W_1 = B_0 - x_1 p_1^A + x_1 p_1^B = B_0 + x_1 \Delta_1\left(X_1\right) \\
W_0 &\geq 2X_1\left[\bar{e} - \frac{\Delta_1 - \Delta_2}{2}\right] \text{ if } X_1 \geq 0 \\
W_0 &\geq 2X_1\left[-\bar{e} - \frac{\Delta_1 - \Delta_2}{2}\right] \text{ if } X_1 < 0 \\
\Delta_1\left(X_1\right) &= 2a\sigma^2\left(s - X_1\right) \\
\Delta_2 &= 0
\end{aligned}$$

Prices of asset A and B coincide at time 2 since both assets pay off a liquidating dividend D_2, hence $\Delta_2 = 0$. The financial constraint differs depending on the sign of the position. In equilibrium, it will be the case that $X_1 \geq 0$, thus for brevity, I will keep only the condition corresponding to $X_1 \geq 0$. I will follow the same rule for the time-0 constraint. Replacing W_2 and Δ_1 in the objective function by their expressions, the programme boils down to:

$$\begin{aligned}
J_1 &= \max_{x_1} B_0 + 2a\sigma^2 x_1\left(s - X_1\right) \qquad (17) \\
s.t.\ W_0 &\geq 2X_1\left[\bar{e} - a\sigma^2 s\left(s - X_1\right)\right]
\end{aligned}$$

The terms related to variance and risk aversion disappear since the arbitrageur takes opposite positions in markets A and B, and thus eliminates all fundamental risk. From the first-order condition, and using $X_1 = x_0 + x_1$, the unconstrained solution is

$$x_1^u = \frac{s - x_0}{2} \tag{18}$$

This trade satisfies the $t = 1$ financial constraint iff

$$W_0 \geq 2X_1^u \left[\bar{e} - a\sigma^2 \left(s - X_1^u\right)\right], \text{ with } X_1^u = x_0 + x_1^u \tag{19}$$

On the left hand side, $W_0 = B_{-1} + x_0 \left(\Delta_0 - \Delta_1\right)$. Thus to express this inequality as a function of the time-0 trade x_0, it is necessary to derive the price schedule Δ_0, which is a function of x_0. To do so, one must solve the local investors' demand at time 0, assuming that inequality (19) is satisfied. At time 0, the local investors choose their holdings (trades) in the risky asset y_0^A. Their final wealth is given by

$$W_2^A = E_{-1}^A - y_0^A p_0^A + s_0 \epsilon_1 + Y_0^a p_1^A + X_1^A \left(D_2 - p_1^A\right) + s_1 \epsilon_2$$

Note that in equilibrium of the subgame, the price p_1^A is the sum of the expected conditional value of the asset at time 1, $D_1 = D + \epsilon_1$ and the liquidity discount, -ϕ_1^A, which is independent of ϵ_1. Hence $D_1 - p_1^A$ is independent of ϵ_1, which implies that $D_2 - p_1$ depends only on ϵ_2. Thus:

$$\begin{aligned} \mathbb{E}_0\left(W_2^A\right) &= E_{-1}^A - y_0^A p_0^A + Y_0^A \mathbb{E}_0\left(p_1^A\right) + X_1^A \left(D_1 - p_1\right) \\ \mathbb{V}_0\left(W_2^A\right) &= \sigma^2 \left(Y_0^A + s_0\right)^2 + \sigma^2 \left(Y_1^A + s_1\right)^2 \end{aligned}$$

Hence the local investors' maximization problem at time 0 is:[21]

$$\begin{aligned} \chi_0^A &= \max_{y_0^A} \mathbb{E}_0\left(W_2^A\right) - \frac{a}{2}\mathbb{V}_0\left(W_2^A\right) \\ &= \max_{y_0^A} E_{-1}^A - y_0^A p_0^A + Y_0^A \mathbb{E}_0\left(p_1^A\right) + Y_1^A \left(D_1 - p_1\right) - \frac{a}{2}\sigma^2 \left(Y_0^A + s_0\right)^2 \\ &- \frac{a}{2}\sigma^2 \left(Y_1^A + s_1\right)^2 \end{aligned} \tag{20}$$

The first-order condition yields:

$$\mathbb{E}_0\left(p_1^A\right) - p_0^A = a\sigma^2 \left(Y_0^A + s_0\right)$$

[21] Note that following Basak and Chabakauri (2009) we could write a recursive representation for the local investors' problem, using the law of the conditional variance. This would yield of course the same solution.

Using the symmetry of the B-market, and market-clearing in both markets, gives:

$$\Delta_0 - \Delta_1 = 2a\sigma^2 (s - x_0) \tag{21}$$

Using this result, equation (18), the fact that $X_1 = x_0 + x_1$, the notation $B_{-1} = W_{-1}$ and the financial constraint (19), the condition under which the constraint is slack is:

$$W_{-1} \geq \bar{e}(s + x_0) - a\sigma^2 (s - x_0) \frac{s + 5x_0}{2}, \tag{22}$$

$$\text{i.e. } W_{-1} - s\bar{e} + a\sigma^2 \frac{s^2}{2} + \left(2a\sigma^2 s - \bar{e}\right) x_0 - \frac{5}{2} a\sigma^2 x_0^2 \geq 0 \tag{23}$$

■

B.2 Proposition 2

Proof. Building on Lemma 2, we know that the unconstrained trade $x_1 = \frac{s - x_0}{2}$ is feasible as long as the left-hand side of inequality (11) has a solution, i.e. as long as the discriminant is positive:

$$R = \left(2a\sigma^2 s - \bar{e}\right)^2 + 10a\sigma^2 \left(W_{-1} - s\bar{e} + a\sigma^2 \frac{s^2}{2}\right) \tag{24}$$

Hence, rearranging terms, I get that $R \geq 0$ if and only if

$$W_{-1} \geq \frac{7}{5} s\bar{e} - \frac{9}{10} a\sigma^2 s^2 - \frac{e^2}{a\sigma^2} \equiv \omega^c \tag{25}$$

Note that the threshold ω^c is in some cases negative:

Lemma 7 *If* $\frac{\bar{e}}{a\sigma^2 s} < 7 - 2\sqrt{10}$ *or* $\frac{\bar{e}}{a\sigma^2 s} > 7 + 2\sqrt{10}$, $\omega^c < 0$.

Proof. Rewriting ω^c as $\frac{14a\sigma^2 s\bar{e} - 9a^2\sigma^4 s^2 - \bar{e}^2}{10a\sigma^2}$, one can view the numerator as a second-order equation in $\bar{e}$ and calculate its discriminant, $\delta = 160a^2\sigma^4 s^2 > 0$. There are two positive roots, $\left(7 - 2\sqrt{10}\right) a\sigma^2 s \approx 0.67a\sigma^2 s$ and $\left(7 + 2\sqrt{10}\right) a\sigma^2 s \approx 13.3a\sigma^2 s$ ■

Let us assume that $W_{-1} < \omega^c$ so that the arbitrageur is necessarily constrained at $t = 1$, i.e. there is no position x_0 such that $x_1 = \frac{s - x_0}{2}$ is feasible.

Using $X_1 = x_0 + x_1$ to rewrite the arbitrageur's problem (17) gives:

$$J_1 = \max_{x_1} \quad B_0 - a\sigma^2 s x_0 + 2a\sigma^2 (s + x_0) X_1 - 2a\sigma^2 X_1^2 \tag{26}$$

$$\text{s.t.} \quad W_0 + 2X_1 \left(a\sigma^2 s - \bar{e}\right) - a\sigma^2 X_1^2 \geq 0 \tag{27}$$

The function (26) has a maximum at $X_1^u = x_0 + x_1^u = \frac{s + x_0}{2}$. Given that W_0 will be positive in equilibrium, the function (27) has one negative root and one positive root, which are smaller than

X_1^u. The constraint is based on the assumption that $X_1 \geq 0$, which will be true in equilibrium. Thus, if the arbitrageur's constraint binds at time 1, her trade must be between 0 and the positive root. Given that the objective function is increasing for $X_1 \leq X_1^u$, the arbitrageur maxes out her financial constraint and chooses the positive root:

$$X_1^c = \frac{a\sigma^2 s - \bar{e} + \sqrt{U^m}}{2a\sigma^2} > 0, \text{ with } U^m = \left(\bar{e} - a\sigma^2 s\right)^2 + 2a\sigma^2 \underbrace{\left(W_{-1} + 2a\sigma^2 x_0 \left(s - x_0\right)\right)}_{W_0} \tag{28}$$

This implies that:

$$2a\sigma^2 X_1^c \left(s - X_1^c\right) = \frac{\bar{e}}{a\sigma^2}\left[a\sigma^2 s - \bar{e} + \sqrt{U^m}\right] - \left(W_{-1} + 2a\sigma^2 x_0 \left(s - x_0\right)\right) \tag{29}$$

Further, substituting (28) into (14) yields $\Delta_1^{c_1} = a\sigma^2 s + \bar{e} - \sqrt{U^m}$. Using (15), this implies that

$$\Delta_0^{c_1}\left(x_0\right) = a\sigma^2 s + \bar{e} - \sqrt{U^m} + 2a\sigma^2 \left(s - x_0\right) \tag{30}$$

I can now solve the arbitrageur's problem at time 0.

$$\begin{aligned} J_0^{c_1} = \quad & \max_{x_0} \quad W_{-1} + x_0 \Delta_0^{c_1}\left(x_0\right) + 2a\sigma^2 X_1^{c_1}\left(s - X_1^{c_1}\right) \\ & \text{s.t.} \qquad W_{-1} \geq 2x_0 \left(\bar{e} - a\sigma^2 \left(s - x_0\right)\right) \end{aligned} \tag{31}$$

The constraint is based on the assumption that the arbitrageur's position will be positive, which will be true in equilibrium. Substituting equations (29) and (30) into the maximand yields:

$$\begin{aligned} J_0^{c_1} = \quad & \max_{x_0} \quad \frac{\bar{e}}{a\sigma^2}\left[\sqrt{U^m} - \bar{e} + a\sigma^2 s\right] \\ & \text{s.t.} \qquad W_{-1} \geq 2x_0 \left(\bar{e} - a\sigma^2 \left(s - x_0\right)\right) \end{aligned}$$

Since $U^m = \left(\bar{e} - a\sigma^2 s\right)^2 + 2a\sigma^2 \left(W_{-1} + 2a\sigma^2 x_0 \left(s - x_0\right)\right)$ is concave in x_0, the maximization problem admits an interior solution (ignoring the constraint for now):

$$\text{FOC:} \quad \frac{\bar{e}}{a\sigma^2}\frac{\partial\sqrt{U^m}}{\partial x_0} = 0$$

Since $\frac{\partial\sqrt{U^m}}{\partial x_0} = \frac{\frac{\partial U^m}{\partial x_0}}{\sqrt{U^m}} = \frac{4a^2\sigma^4(s-2x_0)}{\sqrt{U^m}}$, the unconstrained optimum is

$$x_0^{c_1,u_0} = \frac{s}{2}$$

The time 0 constraint is satisfied if and only if

$$W_{-1} \geq 2x_0^{c_1,u_0}\left(\bar{e} - a\sigma^2\left(s - x_0^{c_1,u_0}\right)\right) = s\bar{e} - \frac{1}{2}a\sigma^2 s^2 \equiv \omega^p$$

The objective function is increasing on $]-\infty, x_0^{c_1,u_0}]$. Thus, if $W_{-1} < \omega^p$, the arbitrageur trades the largest quantity satisfying her time-0 constraint. The trade saturating the financial constraint is given by (Corollary 1)

$$x_0^0 = \frac{a\sigma^2 s - \bar{e} + \sqrt{Q}}{2a\sigma^2} > 0 \tag{32}$$

I now compare the two thresholds ω^c and ω^p:

$$\omega^c \geq \omega^p \Leftrightarrow \frac{2}{5}s\bar{e} - \frac{2}{5}a\sigma^2 s^2 - \frac{\bar{e}^2}{10a\sigma^2} \Leftrightarrow -\frac{\left(\bar{e} - 2a\sigma^2 s\right)^2}{10a\sigma^2} \geq 0$$

Since the last inequality is never satisfied, $\omega^c \leq \omega^p$, implying that for any arbitrageur capital W_{-1} strictly below ω^c, the arbitrageur is constrained at both $t = 0$ and $t = 1$. ■

B.3 Lemma 3

Proof. Suppose that the arbitrageur is unconstrained at time 1. Substituting $x_1^u = \frac{s - x_0}{2}$ into the arbitrageur's objective function (17) yields her value function in the unconstrained state of the world:

$$J_1^{u_1} = B_0 + \frac{a\sigma^2}{2}\left(s - x_0\right)^2 \tag{33}$$

The u_1, u_0 strategy. I first derive the conditions under which the unconstrained strategy is feasible. If local investors believe that the arbitrageur will be unconstrained at time 1, the arbitrageur's problem at time 0 is (assuming her positions are positive at time 0 and time 1):

$$\begin{aligned} J_0^{u_1} = \quad & \max_{x_0} \quad W_{-1} + x_0 \Delta_0^{u_1}(x_0) + \frac{a\sigma^2}{2}\left(s - x_0\right)^2 \\ & \text{s.t.} \quad W_{-1} \geq 2x_0\left[\bar{e} - \frac{1}{2}\left(\Delta_0 - \Delta_1\right)\right] \\ & \qquad\quad W_{-1} - s\bar{e} + a\sigma^2 \frac{s^2}{2} + \left(2a\sigma^2 - \bar{e}\right)x_0 - \frac{5}{2}a\sigma^2 x_0^2 \geq 0 \end{aligned}$$

Using (16) and (21) gives

$$\Delta_0^{u_1}(x_0) = 3a\sigma^2\left(s - x_0\right) \tag{34}$$

Substituting (34), (16) and (21) into the arbitrageur's problem thus yields:

$$J_0^{u_1} = \max_{x_0} \quad W_{-1} + \frac{a\sigma^2 s^2}{2} + 2a\sigma^2 s x_0 - \frac{5}{2}\sigma^2 x_0^2 \tag{35}$$

$$\text{s.t.} \quad W_{-1} + 2\left(a\sigma^2 s - \bar{e}\right) x_0 - 2a\sigma^2 x_0^2 \geq 0 \tag{36}$$

$$W_{-1} + \frac{a\sigma^2 s^2}{2} - s\bar{e} + \left(2a\sigma^2 s - \bar{e}\right) x_0 - \frac{5}{2}a\sigma^2 x_0^2 \geq 0 \tag{37}$$

The first-order condition (ignoring the financial constraints) gives:

$$x_0^{u_1,u_0} = \frac{2}{5}s \tag{38}$$

$$\text{which implies: } x_1^{u_1,u_0} = \frac{3}{10}s \tag{39}$$

Substituting (38) and (39) into (16) and (34) gives the spreads:

$$\Delta_0^{u_1,u_0} = \frac{9}{5}a\sigma^2 s; \; \Delta_1^{u_1,u_0} = \frac{3}{5}a\sigma^2 s \tag{40}$$

Further, from equations (38) and (39), the payoff is:

$$J_0^{u_1,u_0} = W_{-1} + \frac{9}{10}a\sigma^2 s^2 \tag{41}$$

Substituting equations (38)-(39) into the financial constraints gives:

$$\text{At } t=0,\ W_{-1} \geq 2x_0^{u_1,u_0}\left[\bar{e} - a\sigma^2\left(s - x_0^{u_1,u_0}\right)\right] \Leftrightarrow \ W_{-1} \geq \omega_0^m \equiv \frac{4}{5}s\bar{e} - \frac{12}{25}a\sigma^2 s^2 \tag{42}$$

$$\text{At } t=1,\ W_0 \geq 2X_1^{u_1,u_0}\left[\bar{e} - a\sigma^2\left(s - X_1^{u_1,u_0}\right)\right] \Leftrightarrow \ W_{-1} \geq \omega_1^m \equiv \frac{7}{5}s\bar{e} - \frac{9}{10}a\sigma^2 s^2 \tag{43}$$

From the expressions of ω_0^m and ω_1^m, one can see that a slack constraint at $t = 0$ does not necessarily imply the same at $t = 1$, in particular, $\omega_1^m \geq \omega_0^m \Leftrightarrow \bar{e} \geq \frac{7}{10}a\sigma^2 s$. Hence, the arbitrageur can play the strategy u_1, u_0 if and only if $W_{-1} \geq \max\left(\omega_0^m, \omega_1^m\right)$.

The u_1, c_0 strategy. For $\omega^c < W_{-1} < \max\left(\omega_0^m, \omega_1^m\right)$, one or two constraints may be binding if the arbitrageur trades $x_0^{u_1,u_0}$ and $x_1^{u_1,u_0}$. From the proof of Proposition 2, if $W_{-1} \geq \omega^c$, there always exists an x_0 such that the time-1 constraint is not binding. The time-0 trade, however, must also satisfy the financial constraint at time 0.

First consider the time-1 constraint (37) on positive positions. It has two roots, denoted x_0^1 and $x_0^{1'}$:

$$x_0^1 = \frac{2a\sigma^2 s - \bar{e} + \sqrt{R}}{5a\sigma^2}, \; x_0^{1'} = \frac{2a\sigma^2 s - \bar{e} - \sqrt{R}}{5a\sigma^2}$$

By inspecting (37), one can see that if $W_{-1} \geq \omega^p = s\bar{e} - \frac{1}{2}a\sigma^2 s^2$, $x_0^1 > 0$ and $x_0^{1'} < 0$. Instead, if $W_{-1} < \omega^p$, the sign of the roots depends on the sign of the coefficient of the term in x_0: if $\frac{\bar{e}}{a\sigma^2 s} < 2$, then $x_0^1 > x_0^{1'} > 0$, otherwise, $0 > x_0^1 > x_0^{1'}$. Note that $\omega^p > \omega^c$ (Lemma 12). Thus there are three cases:

1. If $W_{-1} \geq \omega^p$, there are two positive upper bounds x_0^0 and x_0^1 on the time-0 trade feasible to satisfy the time-0 constraint, and remain unconstrained at time 1. The constraints being quadratic in x_0, with a negative coefficient of the term in x_0, they are satisfied for any $x_0 \in \mathcal{I}_1 \equiv \left[0, x_0^0\right] \cap \left[0, x_0^1\right]$. The objective function (35) is also quadratic in x_0. It is increasing on $\left]-\infty, x_0^{u_1,u_0}\right]$ and decreasing otherwise. Since $W_{-1} < \max\left(\omega_0^m, \omega_1^m\right)$, either x_0^0 or x_0^1, or both, are smaller than $x_0^{u_1,u_0}$. Thus the arbitrageur's optimal strategy is to choose the largest x_0 belonging to $\mathcal{I}_1$.

   - If $\omega_1^m \leq W_{-1} < \omega_0^m$, only the time-0 constraint is binding, thus the arbitrageur chooses $x_0 = x_0^0$.
   - If $\omega_0^m \leq W_{-1} < \omega_1^m$, only the time-1 constraint is binding, thus the arbitrageur chooses $x_0 = x_0^1$.
   - If $\omega^p \leq W_{-1} < \min\left(\omega_1^m, \omega_0^m\right)$, both the time-0 and time-1 constraints are binding, thus the arbitrageur chooses $x_0 = \min\left(x_0^1, x_0^0\right)$.

2. If $\max\left(0, \omega^c\right) \leq W_{-1} < \omega^p$ and $\frac{\bar{e}}{a\sigma^2 s} < 2$, then $x_0^0 > 0$ and $x_0^1 > x_0^{1'} > 0$. Thus the time-0 constraint still imposes a positive upper bound on the time-0 trade, while the time-1 constraint imposes both a positive lower bound $x_0^{1'}$ and a positive upper bound x_0^1. If $x_0^0 > x_0^{1'}$, the interval $\mathcal{I}_2 \equiv \left[0, x_0^0\right] \cap \left[x_0^{1'}, x_0^1\right] \neq \{\emptyset\}$ then the arbitrageur's optimal strategy is to choose the largest $x_0 \in \mathcal{I}_2$, thus her strategy is the same as in Case 1. Let's study the position of x_0^0 relative to $x_0^{1'}$: the following result shows that $x_0^0 > x_0^{1'}$, so that this case is similar to Case 1.

   Lemma 8 *If* $\frac{\bar{e}}{a\sigma^2 s} < 2$ *and* $W_{-1} \in \left[\max\left(0, \omega^c\right), \omega^p\right[$, $x_0^0 > x_0^{1'}$.

   Proof. $x_0^0 > x_0^{1'} \Leftrightarrow a\sigma^2 s - 3\bar{e} + 5\sqrt{Q} + 2\sqrt{R} > 0$. Thus

   - If $\frac{\bar{e}}{a\sigma^2 s} \in \left[0, \frac{1}{3}\right[$, the condition is satisfied.
   - If $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{1}{3}, 2\right[$, then

   $$x_0^0 > x_0^{1'} \Leftrightarrow 25Q + 4R + 20\sqrt{QR} > 9\bar{e}^2 + a^2\sigma^2 s^2 - 6a\sigma^2 s\bar{e}$$

   Substituting for R (equation (24)) and Q (corollary 1) gives:

   $$\begin{aligned} 25Q + 4R &= 90a\sigma^2 W_{-1} + 29\bar{e}^2 + 61a\sigma^2 s^2 - 106a\sigma^2 s\bar{e} \\ &\Rightarrow \quad x_0^0 > x_0^{1'} \Leftrightarrow 90a\sigma^2 W_{-1} + 20\bar{e}^2 + 60a\sigma^2 s^2 - 100a\sigma^2 s\bar{e} + 20\sqrt{QR} > 0 \end{aligned}$$

A sufficient condition for this inequality to be satisfied is

$$W_{-1} \geq \omega^r \equiv \frac{10}{9}s\bar{e} - \frac{2}{3}a\sigma^2 s^2 - \frac{2\bar{e}^2}{9a\sigma^2} \tag{44}$$

To assess the previous condition, it is useful to study the sign of ω^r and its position relative to ω^c. It is easy to show that if $\omega^r > \omega^c$ if $\frac{\bar{e}}{a\sigma^2 s} \in [0, \kappa_2[$ and $\omega^r \leq \omega^c$ otherwise, where $\kappa_2 \equiv \frac{\sqrt{253}-13}{11} \approx 0.26 < \frac{1}{3}$. Further, $\omega^r \geq 0$ if $\frac{\bar{e}}{a\sigma^2 s} \in [\kappa_1, 2[$, where $\kappa_1 = \frac{10-\sqrt{52}}{4} \approx 0.697$. Thus, there are two cases:

- If $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{1}{3}, \kappa_1\right[$, $\omega^r \leq \omega^c$, and $\omega^r < 0$, thus $W_{-1} \geq \omega^c \Rightarrow W_{-1} \geq \omega^r$, implying that condition 44 is satisfied.
- If $\frac{\bar{e}}{a\sigma^2 s} \in [\kappa_1, 2[$, $0 \leq \omega^r \leq \omega^c$. thus $W_{-1} \geq \omega^c \Rightarrow W_{-1} \geq \omega^r$, implying that condition 44 is satisfied.

This concludes the proof.

■

3. If $\max(0, \omega^c) \leq W_{-1} < \omega^p$ and $\frac{\bar{e}}{a\sigma^2 s} \geq 2$ then $x_0^0 > 0$ and $0 > x_0^1 > x_0^{1'}$. The fact that equation (48) has two negative roots implies that there is no positive trade satisfying both inequalities (36) and (37). Thus, we need to consider the complete optimization problem, by explicitly writing the financial constraints on positive and negative positions:

$$J_0^{u_1} = \max_{x_0} \quad W_{-1} + \frac{a\sigma^2 s^2}{2} + 2a\sigma^2 s x_0 - \frac{5}{2}\sigma^2 x_0^2 \tag{45}$$

$$\text{s.t.} \quad W_{-1} + 2\left(a\sigma^2 s - \bar{e}\right) x_0 - 2a\sigma^2 x_0^2 \geq 0 \text{ if } x_0 \geq 0 \tag{46}$$

$$W_{-1} + 2\left(a\sigma^2 s + \bar{e}\right) x_0 - 2a\sigma^2 x_0^2 \geq 0 \text{ if } x_0 < 0 \tag{47}$$

$$W_{-1} + \frac{a\sigma^2 s^2}{2} - s\bar{e} + \left(2a\sigma^2 s - \bar{e}\right) x_0 - \frac{5}{2}a\sigma^2 x_0^2 \geq 0 \text{ if } \tfrac{s+x_0}{2} \geq 0 \tag{48}$$

$$W_{-1} + \frac{a\sigma^2 s^2}{2} + s\bar{e} + \left(2a\sigma^2 s + \bar{e}\right) x_0 - \frac{5}{2}a\sigma^2 x_0^2 \geq 0 \text{ if } \tfrac{s+x_0}{2} < 0 \tag{49}$$

Equations (47) and (49) are easily obtained from equation (2) and by following the same steps as in Lemma 2. Clearly, inequality (47) imposes a lower bound $x_0^{0-} \equiv \frac{2a\sigma^2 s + \bar{e} - \sqrt{\Delta_0^-}}{2a\sigma^2}$ on short positions, with $\Delta_0^- = \left(\bar{e} + a\sigma^2 s\right) + 2a\sigma^2 W_{-1}$. Since there is no positive trade allowing the arbitrageur to remain unconstrained at time 1, consider the financial constraints (47), (48) and (49) for $x_0 < 0$. The time-1 constraints (48)-(49) switch around the axis $x_0 = -s$, thus one must first determine the position of x_0^{0-} relative to $-s$:

$$\begin{aligned} x_0^{0-} \leq -s \quad &\Leftrightarrow \quad 3a\sigma^2 s + \bar{e} \leq \sqrt{\Delta_0^-} \\ &\Leftrightarrow \quad s\bar{e} + 4a\sigma^2 s^2 \leq W_{-1} \end{aligned} \tag{50}$$

Given that by assumption, $W_{-1} < \omega^p < s\bar{e} + 4a\sigma^2 s^2$, condition (50) is not satisfied. Thus $x_0^{0-} > -s$. Hence there is no time-0 trade implying $X_1 < 0$ (i.e. $x_0 < -s$) that can jointly satisfy inequalities (47) and (49). Thus only trades above $-s$ are possible, so that one must look for $x_0 < 0$ such that both inequalities (47) and (48) are satisfied. The time-1 constraint (48) is satisfied iff $x_0 \in \left[x_0^{1'}, x_0^1\right]$ when $x_0 > -s$. the time-0 constraint being positive for $x_0 \in \left[x_0^{0-}, 0\right[$, to jointly satisfy (47) and (48), it is necessary to have: $x_0^1 > x_0^{0-}$. Using the definitions of x_0^1 and $x_0^{0'}$ given above gives:

$$x_0^{0-} > x_0^1 \;\Leftrightarrow\; a\sigma^2 s + 7\bar{e} > 2\sqrt{R} + 5\sqrt{\Delta_0^-}$$

Thus, powering each side to the square gives:

$$x_0^{0-} > x_0^1 \;\Leftrightarrow\; a^2\sigma^4 s^2 + 49\bar{e}^2 + 14a\sigma^2 s\bar{e} > 4R + 25\Delta_0^- + 20\sqrt{R\Delta_0^-}$$

Given that $4R + 25\Delta_0^- = 90a\sigma^2 W_{-1} - 6a\sigma^2 s\bar{e} + 61a^2\sigma^4 s^2 + 29\bar{e}^2$, we get:

$$x_0^{0-} \;>\; x_0^1 \;\Leftrightarrow\; W_{-1} < \omega^u - \frac{2\sqrt{R\Delta_0^-}}{9a\sigma^2} \tag{51}$$

$$\omega^u \;=\; \frac{2}{9}s\bar{e} - \frac{2}{3}a\sigma^2 s^2 + \frac{2\bar{e}^2}{9a\sigma^2} \tag{52}$$

It is necessary to determine the sign of ω^u and its position relative to the boundaries of the interval under consideration: $[\max(0, \omega^c), \omega^p[$. First, note that $\omega^u > 0$ if $\frac{\bar{e}}{a\sigma^2 s} > \frac{\sqrt{13}-1}{2} \approx 1.3$. Thus $\frac{\bar{e}}{a\sigma^2 s} \geq 2 \;\Rightarrow\; \omega^u > 0$. Further,

$$\omega^p > \omega^u \;\Leftrightarrow\; \frac{7}{9}s\bar{e} + \frac{1}{6}a\sigma^2 s^2 - \frac{2\bar{e}^2}{9a\sigma^2} \;\Leftrightarrow\; \frac{14a\sigma^2 s\bar{e} + 3a^2\sigma^4 s^2 - 4\bar{e}^2}{18a\sigma^2}$$

Thus, $\omega^p > \omega^u \;\Leftrightarrow\; \frac{\bar{e}}{a\sigma^2 s} < \frac{7+\sqrt{61}}{4} \approx 3.7$. Next,

$$\omega^u > \omega^c \;\Leftrightarrow\; \frac{-106a\sigma^2 s\bar{e} + 21a^2\sigma^4 s^2 + 29\bar{e}^2}{90a\sigma^2} > 0$$

Analyzing the numerator as a quadratic function in $\bar{e}$ gives:

$$\omega^u \geq \omega^c \;\Leftrightarrow\; \frac{\bar{e}}{a\sigma^2 s} \in \left[0, \frac{106 - \sqrt{8800}}{58}\right] \cup \left[\frac{106 + \sqrt{8800}}{58}, \infty\right)$$

Denote

$$\iota_1 \equiv \frac{106 + \sqrt{8800}}{58}; \iota_2 \equiv \frac{7 + \sqrt{61}}{4}$$

the position of ω^u relative to the boundaries of the interval is:

- If $\frac{\bar{e}}{a\sigma^2 s} \in [2, \iota_1[$, $\omega^u < \omega^c < \omega^p$.
- If $\frac{\bar{e}}{a\sigma^2 s} \in [\iota_1, \iota_2[$, $\omega^c \leq \omega^u < \omega^p$.
- If $\frac{\bar{e}}{a\sigma^2 s} > \iota_2$, $\omega^c < \omega^p < \omega^u$.

Thus there are three cases:

(a) If $\frac{\bar{e}}{a\sigma^2 s} \in [2, \iota_1[$, $W_{-1} \in [\max(0, \omega^c), \omega^p[\quad \Rightarrow \quad W_{-1} > \omega^u$. Thus, condition 51 is not satisfied, and $x_0^{0-} < x_0^1$.

(b) If $\frac{\bar{e}}{a\sigma^2 s} \in [\iota_1, \iota_2[$, if $W_{-1} \in [\omega^u, \omega^p[$, then as in the previous case, $x_0^{0-} < x_0^1$. If $W_{-1} \in [0 \wedge \omega^c, \omega^u[$, then $W_{-1} < \omega^u$, thus we can rewrite condition 51 as

$$x_0^{0-} > x_0^1 \;\Leftrightarrow\; [-W_{-1} + \omega^u]^2 > \frac{4}{81a^2\sigma^4} R\Delta_0^-$$

Developing terms on each side, and rearranging gives:

$$\begin{aligned} x_0^{0-} &> x_0^1 \;\Leftrightarrow\; \alpha W_{-1}^2 + \beta W_{-1} + \gamma > 0 \\ \alpha &= \frac{1}{81};\; \beta = -2\left[\frac{34}{81} s\bar{e} + \frac{42}{81}\frac{\bar{e}^2}{a\sigma^2} + \frac{2}{81} a\sigma^2 s^2\right] < 0;\; \gamma = (\omega^u)^2 + \frac{40}{81a\sigma^2}\left(a\sigma^2 s + \bar{e}\right)\omega^c \end{aligned} \tag{53}$$

(c) If $\frac{\bar{e}}{a\sigma^2 s} > \iota_2$, $W_{-1} \in [\max(0, \omega^c), \omega^p[\;\; \Rightarrow \;\; W_{-1} < \omega^u$, thus the analysis is similar to the previous case and yields condition (53).

■

B.4 Lemma 4

Proof.

The c_1, u_0 strategy was derived, in part, in the proof of Proposition 2. I recall the main equations here for convenience. The positions are:

$$\begin{aligned} x_0^{c_1,u_0} &= \frac{s}{2} > 0 \\ X_1^{c_1,u_0} &= \frac{a\sigma^2 s - \bar{e} + \sqrt{U^m}}{2a\sigma^2} > 0 \\ \text{where } U^m &= \left(\bar{e} - a\sigma^2 s\right)^2 + 2a\sigma^2\left[W_{-1} + 2a\sigma^2 x_0^{c_1,u_0}\left(s - x_0^{c_1,u_0}\right)\right] \\ &= 2a\sigma^2 W_{-1} + \left(\bar{e} - a\sigma^2 s\right)^2 + a^2\sigma^4 s^2 \end{aligned} \tag{54}$$

The payoff is:

$$J_0^{c_1,u_0} = \frac{\bar{e}}{a\sigma^2}\left[\sqrt{U^m} - \bar{e} + a\sigma^2 s\right] \tag{55}$$

From Proposition 2, the time-0 trade is feasible if and only if $W_{-1} \geq \omega^p$. I now check under which condition the arbitrageur is indeed constrained at time 1 after trading $x_0 = \frac{s}{2}$. It must be that if the arbitrageur re-optimizes at time 1 with an initial position $x_0^{c1,u_0} = \frac{s}{2}$, her financial constraint binds, i.e. $k\left(\frac{s}{2}\right) < 0$, where k is given in Lemma 11.

$$k\left(\frac{s}{2}\right) < 0 \;\Leftrightarrow\; W_{-1} + 2a\sigma^2 x_0\left(s - x_0\right) \geq 2X_1\left[\bar{e} - a\sigma^2\left(s - X_1\right)\right] \Leftrightarrow\; W_{-1} \geq \frac{3}{2}s\bar{e} - \frac{7}{8}a\sigma^2 s^2 \equiv \bar{\omega}^p$$

To sum up: the c_1, u_0 strategy is

- feasible if and only if $W_{-1} \geq \omega^p$,
- time consistent if and only if $W_{-1} < \bar{\omega}^p$.

The c_1, c_0 strategy is given in Corollary 1. Substituting $x_0^{c_1,c_0}$ into the arbitrageur's payoff (31) gives:

$$U^m\left(x_0^{c_1,c_0}\right) = \left(\bar{e} - a\sigma^2 s\right)^2 + 4a\sigma^2 x_0^{c_1,c_0}\bar{e}$$

where I used the fact that the time-0 constraint is binding so that $W_{-1} = 2x_0^{c_1,c_0}\left(\bar{e} - 2a\sigma^2\left(s - x_0^{c_1,c_0}\right)\right)$. Substituting x_0^{c,c_0}, this gives:

$$U^m\left(x_0^{c_1,c_0}\right) = U = a^2\sigma^4 s^2 - \bar{e}^2 + 2\bar{e}\sqrt{Q}$$

Thus the arbitrageur's payoff is

$$J_0^{c_1,c_0} = \frac{\bar{e}}{a\sigma^2}\left[a\sigma^2 s - \bar{e} + \sqrt{U)}\right]$$

The arbitrageur cannot re-optimize at time 1 if inequality (48) is violated when it is evaluated at $x_0^{c_1,c_0} = \frac{a\sigma^2 s - \bar{e} + \sqrt{Q}}{2a\sigma^2}$, i.e. when $k\left(x_0^{c_1,c_0}\right) < 0$: Substituting (32) into (22) yields:

$$\begin{aligned} k\left(x_0^{c_1,c_0}\right) < 0 \quad &\Leftrightarrow \quad W_{-1} < \frac{\left(3a\sigma^2 s - \bar{e} + \sqrt{Q}\right)\bar{e}}{2a\sigma^2} - \frac{\left(a\sigma^2 s + \bar{e} - \sqrt{Q}\right)\left(7a\sigma^2 s - 5\bar{e} + 5\sqrt{Q}\right)}{8a\sigma^2} \\ &\Leftrightarrow \quad W_{-1} < \frac{10a\sigma^2 s\bar{e} - 7a^2\sigma^4 s^2 + \bar{e}^2 + 5Q + \left(2a\sigma^2 s - 6\bar{e}\right)\sqrt{Q} + 10a\sigma^2 W_{-1}}{8a\sigma^2} \end{aligned}$$

Table 1: Subcases

$\frac{\bar{e}}{a\sigma^2 s}$	$\left[0, \frac{1}{3}\right[$	$\left[\frac{1}{3}, \frac{\sqrt{19}-1}{6}\right[$	$\left[\frac{\sqrt{19}-1}{6}, \frac{1}{\sqrt{3}}\right[$	$\left[\frac{1}{\sqrt{3}}, \frac{2\sqrt{150}-7}{29}\right[$	$\left[\frac{2\sqrt{150}-7}{29}, \infty\right[$
	$a\sigma^2 s - 3\bar{e} \geq 0$	$a\sigma^2 s - 3\bar{e} < 0$	$a\sigma^2 s - 3\bar{e} < 0$	$a\sigma^2 s - 3\bar{e} < 0$	$a\sigma^2 s - 3\bar{e} < 0$
	$\omega^c < \omega^p \leq \bar{\omega}^c$	$\omega^c < \omega^p \leq \bar{\omega}^c$	$\omega^c < \bar{\omega}^c \leq \omega^p$	$\omega^c < \bar{\omega}^c \leq \omega^p$	$\bar{\omega}^c < \omega^c < \omega^p$
	$\bar{\omega}^c > 0$	$\bar{\omega}^c > 0$	$\bar{\omega}^c > 0$	$\bar{\omega}^c \leq 0$	$\bar{\omega}^c \leq 0$

Using the definition of Q (corollary 1) and rearranging terms gives:

$$k\left(x_0^{c_1,c_0}\right) < 0 \;\Leftrightarrow\; \underbrace{a\sigma^2 W_{-1} - a^2\sigma^4 s^2 + 3\bar{e}^2 + \left(a\sigma^2 s - 3\bar{e}\right)\sqrt{Q\left(W_{-1}\right)}}_{h(W_{-1})} > 0$$

To analyze the sign of h, we must determine the sign of $a\sigma^2 s - 3\bar{e}$ and of $W_{-1} - a^2\sigma^4 s^2 + 3\bar{e}^2$:

$$a\sigma^2 s - 3\bar{e} \;\geq\; 0 \;\Leftrightarrow\; \frac{\bar{e}}{a\sigma^2 s} < \frac{1}{3} \tag{56}$$

$$W_{-1} - a^2\sigma^4 s^2 + 3\bar{e}^2 \;\geq\; 0 \;\Leftrightarrow\; W_{-1} \geq \bar{\omega}^c \equiv a\sigma^2 s^2 - \frac{3\bar{e}^2}{a\sigma^2} \tag{57}$$

The strategy c_1, c_0 is dominated by c_1, u_0, which is feasible for $W_{-1} \geq \omega^p$. Further, there is no commitment problem if $W_{-1} < \omega^c$. Thus, we can restrict the analysis to $W_{-1} \in \left[\max\left(\omega^c, 0\right), \omega^p\right[$. To this end, I determine the position of $\bar{\omega}^c$ relative to ω^c and ω^p:

$$\omega^c > \bar{\omega^c} \;\Leftrightarrow\; 14a\sigma^2 s\bar{e} - 19a^2\sigma^4 s^2 + 29\bar{e}^2 > 0$$

It is easy to show that this inequality is true for $\frac{\bar{e}}{a\sigma^2 s} > \frac{2\sqrt{150}-7}{29} \approx 0.603$. Similarly,

$$\bar{\omega}^c > \omega^p \;\Leftrightarrow\; 3a^2\sigma^4 s^2 - 2a\sigma^2 s\bar{e} - 6\bar{e}^2 > 0 \Leftrightarrow \frac{\bar{e}}{a\sigma^2 s} \in \left[0, \frac{\sqrt{19}-1}{6}\right[$$

Further, $\bar{\omega}^c > 0 \;\Leftrightarrow\; \frac{\bar{e}}{a\sigma^2 s} < \frac{1}{\sqrt{3}}$. Thus, we have the following cases, summarized in Table 1.

1. If $\frac{\bar{e}}{a\sigma^2 s} \in \left[0, \frac{1}{3}\right[$, $W_{-1} \in \left[\omega^c, \omega^p\right[\;\;\Rightarrow\;\; W_{-1} \leq \bar{\omega}^c$, and $a\sigma^2 s - 3\bar{e} \geq 0$. Thus $k\left(x_0^{c_1,c_0}\right) < 0 \;\Leftrightarrow\; \left(a\sigma^2 s - 3\bar{e}\right)\sqrt{Q} > -a\sigma^2\left(W_{-1} - \bar{\omega}^c\right) \geq 0$. Thus $\left(a\sigma^2 s - 3\bar{e}\right)^2 Q > a^2\sigma^4\left(W_{-1} - \bar{\omega}^c\right)^2$, which gives:

$$k\left(x_0^{c_1,c_0}\right) \;<\; 0 \;\Leftrightarrow\; \kappa + \Gamma W_{-1} - a^2\sigma^4 W_{-1}^2 > 0 \tag{58}$$

$$\kappa \;=\; -a^2\sigma^2\left(\bar{\omega}^c\right)^2 + \left(a\sigma^2 s - 3\bar{e}\right)^2\left(\bar{e} - a\sigma^2 s\right)^2$$

$$\;=\; 4a\sigma^2 s\bar{e}\left(7a\sigma^2 s\bar{e} - 2a^2\sigma^4 s^2 - 6\bar{e}^2\right) \tag{59}$$

$$\Gamma \;=\; 2a\sigma^2\left[\left(a\sigma^2 s - 3\bar{e}\right)^2 + a\sigma^2\bar{\omega}^c\right] = 2a\sigma^2\left(2a^2\sigma^4 s^2 + 6\bar{e}^2 - 6a\sigma^2 s\bar{e}\right) \tag{60}$$

It is easy to show that $\Gamma > 0$ and that $\kappa \geq 0 \;\Leftrightarrow\; \frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{1}{2}, \frac{2}{3}\right]$. Thus if $\frac{\bar{e}}{a\sigma^2 s} \in \left[0, \frac{1}{3}\right[$, $\kappa < 0$, and condition 58 is not trivially satisfied. The discriminant of the left-hand side of condition 58 is $D^w = 4^2\sigma^4\tilde{D}^w$, where $\tilde{D}^w$ is given in the Lemma.

- If $\tilde{D}^w < 0$, condition 58 is not satisfied, and c_1, c_0 is not time-consistent
- If $\tilde{D}^w > 0$ (ignoring the knife-edge case), then 58 is satisfied if $W_{-1} \in \left]p - \frac{\tilde{D}^w}{a\sigma^2}, p + \frac{\tilde{D}^w}{a\sigma^2}\right[$, where $p = -6s\bar{e} + 2a\sigma^2 s^2 + \frac{6\bar{e}^2}{a\sigma^2}$. Given that $p > \omega^p$, we can eliminate the second case. Thus, c_1, c_0 is time consistent if $W_{-1} > \tilde{\omega} \equiv p - \frac{\tilde{D}^w}{a\sigma^2}$.

2. If $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{1}{3}, \frac{\sqrt{19}-1}{6}\right[$, $a\sigma^2 s - 3\bar{e} < 0$ and $W_{-1} \in [\omega^c, \omega^p[\;\Rightarrow\; W_{-1} \leq \bar{\omega}^c$. This implies that $h(W_{-1}) \leq 0$, thus c_1, c_0 is not time consistent when $W_{-1} \in [\max(\omega^c, 0), \omega^p[$.

3. If $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{\sqrt{19}-1}{6}, \frac{1}{\sqrt{3}}\right[$, $a\sigma^2 s - 3\bar{e} < 0$, and $\omega^c < \bar{\omega}^c \leq \omega^p$. Thus if $W_{-1} \in [\omega^c, \bar{\omega}^c[$, $h(W_{-1}) \leq 0$, and c_1, c_0 is not time-consistent. If $W_{-1} \in [\bar{\omega}^c, \omega^p[$, then $h(W_{-1}) > 0 \;\Leftrightarrow\; \kappa + \Gamma W_{-1} - a^2\sigma^4 W_{-1}^2 < 0$. Since $\kappa > 0$ on this interval, this condition imposes an upper bound $p + \frac{\tilde{D}^w}{a\sigma^2}$, where $p > \omega^p$. Thus, c_1, c_0 is not time consistent when $W_{-1} \in [\bar{\omega}^c, \omega^p[$ either.

4. If $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{1}{\sqrt{3}}, \frac{2\sqrt{150}-7}{29}\right[$, the analysis is similar to the previous case. (the only difference is that $\bar{\omega}^c$ is negative so that the subcase $W_{-1} \in [\omega^c, \bar{\omega}^c[$ is not relevant anymore.

5. If $\frac{\bar{e}}{a\sigma^2 s} > \frac{2\sqrt{150}-7}{29}$, then $a\sigma^2 s - 3\bar{e} < 0$ and $W_{-1} \in [\omega^c, \omega^p[\;\Rightarrow\; W_{-1} \geq \bar{\omega}^c$. Thus this case is the mirror case of case 1, which implies that $h(W_{-1}) > 0 \;\Leftrightarrow\; \kappa + \Gamma W_{-1} - a^2\sigma^4 W_{-1}^2 < 0$. The sign of κ changes over the interval, with $\kappa > 0$ if $\frac{\bar{e}}{a\sigma^2 s} < \frac{2}{3}$. Since $\frac{2\sqrt{150}-7}{29} < \frac{2}{3}$, we have two cases:

   - If $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{2\sqrt{150}-7}{29}, \frac{2}{3}\right[$, $\kappa > 0 \;\Rightarrow \tilde{D}^w > 0$, and $h(W_{-1}) > 0$ iff $W_{-1} > p + \frac{\tilde{D}^w}{a\sigma^2} > \omega^p$. Thus $W_{-1} \in [\bar{\omega}^c, \omega^p[$ implies that c_1, c_0 is not time consistent.
   - If $\frac{\bar{e}}{a\sigma^2 s} \geq \frac{2}{3}$, then $\kappa < 0$. If $\tilde{D}^w < 0$, c_1, c_0 is not time consistent. Otherwise, c_1, c_0 is time consistent if $W_{-1} < \tilde{\omega}$ or $W_{-1} > p + \frac{\tilde{D}^w}{a\sigma^2} > \omega^p$. We can eliminate the second case.

We can collect together cases 2, 3 and 4, and 5a, so that there are two cutoffs, $\frac{1}{3}$ and $\frac{2}{3}$.

To conclude the proof, I show that if $\frac{\bar{e}}{a\sigma^2 s} \geq 2$, c_1, c_0 is time consistent. If $W_{-1} < \omega^p$and $\frac{\bar{e}}{a\sigma^2 s} \geq 2$, the time 1 financial constraint, given by the function k (Lemma 2), has two negative roots. In particular, the largest root, $x_0^1 < 0 < x_0^{c_1,c_0}$ (this is shown in the proof of Lemma 3). Since k is decreasing for $x_0 > x_0^1$, $x_0^1 < 0 < x_0^{c_1,c_0}$ implies that $k\left(x_0^1\right) = 0 > k\left(x_0^{c_1,c_0}\right)$, i.e. c_1, c_0 is time consistent when $\frac{\bar{e}}{a\sigma^2 s} \geq 2$ and $W_{-1} < \omega^p$. ■

B.5 Equilibrium

B.5.1 Preliminary results

The following result will be useful to determine the equilibrium:

Corollary 9 *Suppose that $\bar{\omega}^p \leq W_{-1} < \omega^p$, then both the c_1, u_0 strategy and the c_1, c_0 strategy are not time consistent.*

Proof. First note that $\omega^p \geq 0 \;\Leftrightarrow\; \frac{\bar{e}}{a\sigma^2 s} \geq \frac{1}{2}$. Second, note that by lemma 12, $\bar{\omega}^p \leq \omega^p \;\Leftrightarrow\; \frac{\bar{e}}{a\sigma^2 s} < \frac{3}{4}$. Thus the interval of interest is $\left[\frac{1}{2}, \frac{3}{4}\right[$. Let us use $G(x_0)$ as a short-hand for equation (48). c_1, u_0 not time consistent means that $G(x_0^{c_1,u_0}) \geq 0$. By lemma 4, this is equivalent to $W_{-1} \geq \bar{\omega}^p$. Thus, we just need to show that if $W_{-1} \in [\bar{\omega}^p, \omega^p[$, $G(x_0^{c_1,c_0}) \geq 0$.

By Lemma 3, if $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{1}{2}, \frac{3}{4}\right[$, G has two positive roots $x_0^1 > x_0^{1'} > 0$ and G is positive for $x_0 \in \left[x_0^{1'}, x_0^1\right]$. We can show that if $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{1}{2}, \frac{3}{4}\right[$, $x_0^{c_1,c_0} \in \left[x_0^{1'}, x_0^1\right]$. First, note that $W_{-1} < \omega^p \;\Rightarrow\; x_0^{c_1,c_0} < x_0^{c_1,u_0}$ (this is by definition). Then compare $x_0^{c_1,c_0}$ and $x_0^{1'}$. From Corollary 1 and the proof of Lemma 3, we substitute for the expressions of $x_0^{c_1,c_0}$ and $x_0^{1'}$:

$$x_0^{c_1,c_0} \geq x_0^{1'} \;\Leftrightarrow\; a\sigma^2 s - 3\bar{e} + 5\sqrt{Q} + 2\sqrt{R} > 0$$

If $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{1}{2}, \frac{3}{4}\right[$, $a\sigma^2 s - 3\bar{e} < 0$, thus by elevating to the square on each side, $x_0^{c_1,c_0} \geq x_0^{1'}$ is equivalent to:

$$9a\sigma^2 W_{-1} - 10a\sigma^2 s\bar{e} + 6a^2\sigma^4 s^2 + 2\bar{e}^2 + 2\sqrt{QR} \geq 0$$

The first term is positive if $9a\sigma^2 W_{-1} - 10a\sigma^2 s\bar{e} + 6a^2\sigma^4 s^2 + 2\bar{e}^2 \geq 0 \;\Leftrightarrow\; W_{-1} \geq v = \frac{10}{9}s\bar{e} - \frac{2}{3}a\sigma^2 s^2 - \frac{2}{9}\frac{\bar{e}^2}{a\sigma^2}$. Since $v < \bar{\omega}^p$, the assumption $W_{-1} \geq \bar{\omega}^p$ implies $W_{-1} \geq v$ and thus $x_0^{c_1,c_0} \geq x_0^{1'}$. Thus $x_0^{c_1,c_0} \in \left[x_0^{1'}, x_0^1\right]$, which implies that $G(x_0^{c_1,c_0}) \geq 0$ and concludes the proof. ■

The proofs of propositions 3 and 4 rely on several intermediate results. I first compare the payoffs of the different strategies, then derive parameter conditions to order the capital thresholds, and finally determine the equilibrium in each region.

Lemma 5

Proof. The first point of the lemma follows from direct comparison of $J_0^{u_1,u_0}$ and $J_0^{c_1,u_0}$. I recall the expressions of the respective payoffs for convenience: $J_0^{u_1,u_0} = W_{-1} + \frac{9}{10}a\sigma^2 s^2$, and $J_0^{c_1,u_0} = \frac{\bar{e}}{a\sigma^2}\left[\sqrt{U^m} - \left(\bar{e} - a\sigma^2 s\right)\right]$. This gives:

$$J_0^{u_1,u_0} \geq J_0^{c_1,u_0} \;\Leftrightarrow\; a\sigma^2 W_{-1} + \bar{e}^2 + \frac{9}{10}a^2\sigma^4 s^2 - a\sigma^2 s\bar{e} \geq \bar{e}\sqrt{U^m} \tag{61}$$

The u_1, u_0 strategy requires $W_{-1} \geq \max(\omega_0^m, \omega_1^m) \geq \omega_1^m$ (Lemma 3). This implies that $W_{-1} \geq \omega^l = s\bar{e} - \frac{9}{10}a\sigma^2 s^2 - \frac{\bar{e}^2}{a\sigma^2}$. Thus by elevating each side to the square, inequality (61) is equivalent to:

$$\left[a\sigma^2 W_{-1} - a\sigma^2 \omega^l\right]^2 \geq \bar{e}^2 U^m,$$

Using $U^m = \left(\bar{e} - a\sigma^2 s\right)^2 + 2a\sigma^2 W_{-1} + a^2\sigma^4 s^2$ (Lemma 4) and rearranging terms gives :

$$a^2\sigma^4 W_{-1}^2 - 2a\sigma^2\left(a\sigma^2\omega^l + \bar{e}^2\right)W_{-1} + a^2\sigma^4\left(\omega^l\right)^2 - \bar{e}^2\left(\bar{e}^2 - 2a\sigma^2 s\bar{e} + 2a^2\sigma^4 s^2\right) \geq 0 \qquad (62)$$

The left-hand side is a second-order equation in W_{-1}. Its discriminant is:

$$\delta = 4a^2\sigma^4\left[a\sigma^2\omega^l + \bar{e}^2\right]^2 - 4a^2\sigma^4\left[a^2\sigma^4\left(\omega^l\right)^2 - \bar{e}^2\left(\bar{e}^2 - 2a\sigma^2 s\bar{e} + 2a^2\sigma^4 s^2\right)\right]$$

Using the definition of ω^l and regrouping terms yields:

$$\delta = \frac{4}{5}a^4\sigma^8 s^2\bar{e}^2 > 0$$

Thus, there are always two roots $v_1 = \frac{2a\sigma^2\left(a\sigma^2\omega^l + \bar{e}^2\right) + \sqrt{\delta}}{2a^2\sigma^4}$ and $v_2 = \frac{2a\sigma^2\left(a\sigma^2\omega^l - \bar{e}^2\right) + \sqrt{\delta}}{2a^2\sigma^4}$, with $v_2 \leq v_1$. Hence inequality (62) is satisfied for $W_{-1} \leq v_2$ or $W_{-1} \geq v_1$. v_1 and v_2 can be simplified to the expressions given in the lemma by replacing δ and ω^l by their expressions. It is easy to check that the conditions are both necessary and sufficient.

The second result of the lemma is proved in Lemma 12 below. ∎

Lemma 9 *If $W_{-1} \geq \bar{\omega}^m$, the u_1, u_0 strategy weakly dominates the c_1, c_0 strategy if and only if*

$$f\left(W_{-1}\right) = a^2\sigma^4 W_{-1}^2 - 2a^2\sigma^4\bar{\omega}^l W_{-1} + a^2\sigma^4\left(\bar{\omega}^l\right)^2 - 2\bar{e}^3\sqrt{Q} + \bar{e}^4 - a^2\sigma^4 s^2\bar{e}^2 \geq 0 \qquad (63)$$

Proof. Using the expressions of $J_0^{u_1,u_0}$ (see Lemma 3) and $J_0^{c_1,c_0}$ (Lemma 4) gives:

$$J_0^{u_1,u_0} \geq J_0^{c_1,c_0} \;\Leftrightarrow\; a\sigma^2 W_{-1} - a\sigma^2 s\bar{e} + \frac{9}{10}a^2\sigma^4 s^2 + \bar{e}^2 \geq \bar{e}\sqrt{a^2\sigma^4 s^2 - \bar{e}^2 + 2\bar{e}\sqrt{Q}}$$

Comparing u_1, u_0 to c_1, c_0 makes sense only if $W_{-1} \geq \omega_1^m$, which in turn implies that the left-hand side of the inequality is positive. Taking the square on each side gives:

$$J_0^{u_1,u_0} \geq J_0^{c_1,c_0} \;\Leftrightarrow\; a^2\sigma^4\left[W_{-1} - \bar{\omega}^l\right]^2 \geq a^2\sigma^4 s^2\bar{e}^2 - \bar{e}^4 + 2\bar{e}^3\sqrt{Q}, \text{ with } \bar{\omega}^l = s\bar{e} - \tfrac{9}{10}a\sigma^2 s^2 - \tfrac{\bar{e}^2}{a\sigma^2}$$

Developing and rearranging terms yields condition (63). ∎

Lemma 10 *If $x_0^{u_1,c_0} = x_0^1$ (Lemma 3), then:*

- *the u_1, c_0 strategy is always dominated by the c_1, u_0 strategy,*

- *the u_1, c_0 strategy dominates c_1, c_0 if and only if*

$$g\left(W_{-1}\right)=\frac{2}{5}a\sigma^2W_{-1}-\frac{12}{25}a^2\sigma^4s^2+\frac{42}{25}e^2+\frac{2}{25}a\sigma^2s\bar{e}+\frac{4}{25}\left(a\sigma^2s+2\bar{e}\right)\sqrt{R}-2\bar{e}\sqrt{Q}\geq 0 \quad (64)$$

Proof. First I derive the payoff of the u_1, c_0 strategy when $x_0^{c_1,u_0}=x_0^1$. As a preliminary step, rewrite the objective function (45) as

$$J_0^{u_1,c_0}\left(x_0\right)=W_{-1}+a\sigma^2\left(s-x_0\right)\frac{s+5x_0}{2}$$

Substituting x_0^1 into this equation gives

$$J_0^{u_1,c_0}\left(x_0^1\right)=W_{-1}+\frac{3a\sigma^2s+\bar{e}-\sqrt{R}}{5}\frac{3a\sigma^2s-\bar{e}+\sqrt{R}}{2a\sigma^2}=W_{-1}+\frac{9a^2\sigma^4s^2-\left(\bar{e}-\sqrt{R}\right)^2}{10a\sigma^2}$$

Developing the squared term and substituting for the expression of R (Lemma 3), this simplifies into:

$$J_0^{u_1,c_0}\left(x_0^1\right)=\frac{\bar{e}}{a\sigma^2}\left[\frac{7}{5}a\sigma^2s+\frac{\sqrt{R}-\bar{e}}{5}\right] \quad (65)$$

Then, using the expression for $J_0^{c_1,u_0}$ given in Lemma 4 gives:

$$J_0^{u_1,c_0}\left(x_0^1\right)\geq J_0^{c_1,u_0} \;\Leftrightarrow\; \frac{2}{5}\left(a\sigma^2s+2\bar{e}\right)+\frac{\sqrt{R}}{5}\geq\sqrt{U^m}>0$$

Taking the square on each side yields, after developing and rearranging the terms,

$$\left(a\sigma^2s+2\bar{e}\right)\sqrt{R}\geq 10a\sigma^2\left(W_{-1}-\underline{\omega}^c\right),\text{ with }\underline{\omega}^c=\tfrac{13}{10}s\bar{e}-\tfrac{23}{20}a\sigma^2s^2-\tfrac{e^2}{5a\sigma^2}$$

Given that $W_{-1}\geq\omega^c$ by assumption, and that $\omega^c>\underline{\omega}^c$, the right-hand side is positive and taking the square on each side does not change the order. This yields the following condition on a second-order equation in W_{-1}:

$$-100a^2\sigma^4W_{-1}^2+10a\sigma^2\left[20a\sigma^2\underline{\omega}^c+\left(a\sigma^2s+2\bar{e}\right)^2\right]W_{-1}-10a\sigma^2\left[\left(a\sigma^2s+2\bar{e}\right)^2\omega^c+10a\sigma^2\left(\underline{\omega}^c\right)^2\right]\geq 0$$

The discriminant of the equation is:

$$d=100a^2\sigma^4\left[20a\sigma^2\underline{\omega}^c+\left(a\sigma^2s+2\bar{e}\right)^2\right]^2-400a^2\sigma^4\left[10a\sigma^2\left[\left(a\sigma^2s+2\bar{e}\right)^2\omega^c+10a\sigma^2\left(\underline{\omega}^c\right)^2\right]\right]$$

Developing and regrouping the terms gives

$$d = 100a^2\sigma^4\left(a\sigma^2 s + 2\bar{e}\right)^2\left[\left(a\sigma^2 s + 2\bar{e}\right)^2 - 40a\sigma^2\left(\omega^c - \underline{\omega}^c\right)\right]$$

Since $\omega^c - \underline{\omega}^c = \frac{1}{10}s\bar{e} + \frac{1}{4}a\sigma^2 s^2 + \frac{\bar{e}^2}{10a\sigma^2}$, the discriminant boils down to

$$d = -900a^4\sigma^8 s^2\left(a\sigma^2 s + 2\bar{e}\right)^2 < 0$$

Therefore, given that the coefficient of the second-order term is negative, the inequality is never satisfied.

I now turn to the second point:

$$J_0^{u_1,c_0} \geq J_0^{c_1,c_0} \;\Leftrightarrow\; \frac{2}{5}\left(a\sigma^2 s + 2\bar{e}\right) + \frac{\sqrt{R}}{5} \geq \sqrt{a^2\sigma^4 s^2 - \bar{e}^2 + 2\bar{e}\sqrt{Q}}$$

Elevating both sides to the square, substituting for R and simplifying yields condition (64). ■

Lemma 11 *If $x_0^{u_1,c_0} = x_0^0$ (Lemma 3), then*

- *The u_1, c_0 strategy dominates c_1, u_0 if and only if*

$$\tilde{g}_u\left(W_{-1}\right) = -a\sigma^2 W_{-1} + \left(5\bar{e} - a\sigma^2 s\right)\sqrt{Q} + 2a\sigma^2 s\bar{e} - \bar{e}^2 + a^2\sigma^4 s^2 - 4\bar{e}\sqrt{U^m} \geq 0 \quad (66)$$

- *the u_1, c_0 strategy dominates c_1, c_0 if and only if*

$$\tilde{g}_c\left(W_{-1}\right) = -a\sigma^2 W_{-1} + \left(5\bar{e} - a\sigma^2 s\right)\sqrt{Q} + 2a\sigma^2 s\bar{e} - \bar{e}^2 + a^2\sigma^4 s^2 - 4\bar{e}\sqrt{a^2\sigma^4 s^2 - \bar{e}^2 + 2\bar{e}\sqrt{Q}} \geq 0 \quad (67)$$

Proof. Substituting x_0^0 (Corollary 1) into $J_0^{u_1,c_0}\left(x_0\right)$ (equation (65)) gives:

$$J_0^{u_1,c_0}\left(x_0^0\right) = W_{-1} + \frac{a\sigma^2 s + \bar{e} - \sqrt{Q}}{2}\,\frac{7a\sigma^2 s - 5\bar{e} + 5\sqrt{Q}}{4a\sigma^2}$$

Developing and substituting for R, this simplifies into:

$$J_0^{u_1,c_0}\left(x_0^0\right) = -\frac{1}{4}W_{-1} + \frac{a^2\sigma^4 s^2 + 6a\sigma^2 s\bar{e} - 5\bar{e}^2 + \left(5\bar{e} - a\sigma^2 s\right)\sqrt{Q}}{4a\sigma^2}$$

Then, using $J_0^{c_1,c_0}$ (Lemma 4) gives:

$$J_0^{u_1,c_0}\left(x_0^0\right) \geq J_0^{c_1,u_0} \;\Leftrightarrow\; -a\sigma^2 W_{-1} + a^2\sigma^4 s^2 + 2a\sigma^2 s\bar{e} - \bar{e}^2 + \left(5\bar{e} - a\sigma^2 s\right)\sqrt{Q} - 4\bar{e}\sqrt{U^m} \geq 0$$

Similarly, using $J_0^{c_1,c_0}$ (Lemma 4) and comparing to $J_0^{u_1,c_0}\left(x_0^0\right)$ gives the necessary and sufficient condition inequality 67 in the proposition. ■

B.5.2 Order and signs of thresholds

Lemma 12 *The thresholds are ranked in the following order:*

1. *If* $\frac{\bar{e}}{a\sigma^2 s}\in\left[0,\frac{1}{10}\right[$, *then* $\omega^c<\omega_1^m\leq v_1\leq\bar{\omega}^p\leq\omega^p<\omega_0^m$,
2. *If* $\frac{\bar{e}}{a\sigma^2 s}\in\left[\frac{1}{10},\frac{79}{140}\right[$, *then* $\omega^c<\omega_1^m\leq v_1\leq\bar{\omega}^p<\omega_0^m\leq\omega^p$,
3. *If* $\frac{\bar{e}}{a\sigma^2 s}\in\left[\frac{79}{140},\frac{21}{10\left(1+\sqrt{5}\right)}\right[$, *then* $\omega^c<\omega_1^m\leq v_1<\omega_0^m\leq\bar{\omega}^p\leq\omega^p$,
4. *If* $\frac{\bar{e}}{a\sigma^2 s}\in\left[\frac{21}{10\left(1+\sqrt{5}\right)},\frac{7}{10}\right[$, *then* $\omega^c<\omega_1^m<\omega_0^m\leq v_1\leq\bar{\omega}^p\leq\omega^p$,
5. *If* $\frac{\bar{e}}{a\sigma^2 s}\in\left[\frac{7}{10},\frac{3}{4}\right[$, *then* $\omega^c\leq\omega_0^m\leq\omega_1^m\leq v_1\leq\bar{\omega}^p<\omega^p$,
6. *If* $\frac{\bar{e}}{a\sigma^2 s}\in\left[\frac{3}{4},3-2\sqrt{\frac{6}{5}}\right[$, *then* $\omega^c<\omega_0^m\leq\omega_1^m\leq v_1<\omega^p\leq\bar{\omega}^p$,
7. *If* $\frac{\bar{e}}{a\sigma^2 s}\in\left[3-2\sqrt{\frac{6}{5}},\frac{2\sqrt{5}}{5}\right[$, *then* $\omega_0^m\leq\omega^c\leq\omega_1^m\leq v_1<\omega^p\leq\bar{\omega}^p$,
8. *If* $\frac{\bar{e}}{a\sigma^2 s}\in\left[\frac{2\sqrt{5}}{5},1\right[$, *then* $\omega_0^m<\omega^c\leq\omega_1^m<\omega^p\leq v_1\leq\bar{\omega}^p$,
9. *If* $\frac{\bar{e}}{a\sigma^2 s}\in\left[1,3+2\sqrt{\frac{6}{5}}\right[$, *then* $\omega_0^m<\omega^c\leq\omega^p\leq\omega_1^m\leq v_1\leq\bar{\omega}^p$,
10. *If* $\frac{\bar{e}}{a\sigma^2 s}\geq 3+2\sqrt{\frac{6}{5}}$, *then* $\omega^c\leq\omega_0^m\leq\omega^p\leq\omega_1^m\leq v_1\leq\bar{\omega}^p$.

Proof.

First, recall the expressions of the different thresholds:

$$\begin{aligned}
\omega^c &= \frac{7}{5}s\bar{e}-\frac{9}{10}a\sigma^2 s^2-\frac{\bar{e}^2}{10a\sigma^2};\ \omega_0^m=\frac{4}{5}s\bar{e}-\frac{12}{25}a\sigma^2 s^2;\ \omega_1^m=\frac{7}{5}s\bar{e}-\frac{9}{10}a\sigma^2 s^2\\
\omega^p &= s\bar{e}-\frac{1}{2}a\sigma^2 s^2;\ \bar{\omega}^p=\frac{3}{2}s\bar{e}-\frac{7}{8}a\sigma^2 s^2;\ v_1=\left(1+\frac{\sqrt{5}}{5}\right)s\bar{e}-\frac{9}{10}a\sigma^2 s^2\\
v_2 &= \left(1-\frac{\sqrt{5}}{5}\right)s\bar{e}-\frac{9}{10}a\sigma^2 s^2
\end{aligned}$$

Then determine relative positions:

- ω_1^m vs ω^c: $\omega^c=\omega_1^m-\frac{\bar{e}}{10a\sigma^2}<\omega_1^m$.

- ω_0^m vs ω^c:

$$\begin{aligned}
\omega^c \geq \omega_0^m \quad &\Leftrightarrow \quad \frac{7}{5}s\bar{e} - \frac{9}{10}a\sigma^2 s^2 - \frac{\bar{e}^2}{10a\sigma^2} \geq \frac{4}{5}s\bar{e} - \frac{12}{25}a\sigma^2 s^2 \\
&\Leftrightarrow \quad 6a\sigma^2 s\bar{e} > \frac{42}{10}a^2\sigma^4 s^2 + \bar{e}^2 \\
&\Leftrightarrow \quad -\bar{e}^2 + 6a\sigma^2 s\bar{e} - \frac{42}{10}a^2\sigma^4 s^2 > 0 \qquad (68)
\end{aligned}$$

One can view the left-hand side as a second-order equation in $\bar{e}$. The discriminant of the left-hand side is $D = \frac{96}{5}a^2\sigma^4 s^2$, and given that the coefficient of the second-order term and the constant have the same sign, and that the coefficient of the first-order term is positive, there are two positive roots, given by $\left(3 - 2\sqrt{\frac{6}{5}}\right) a\sigma^2 s \approx 0.81a\sigma^2 s$ and $\left(3 + 2\sqrt{\frac{6}{5}}\right) a\sigma^2 s \approx 5.2a\sigma^2 s$. Hence $\omega^c \geq \omega_0^m$ if and only if $\frac{\bar{e}}{a\sigma^2 s} \in \left[3 - 2\sqrt{\frac{6}{5}}, 3 + 2\sqrt{\frac{6}{5}}\right]$.

- ω^p vs ω_1^m and ω_0^m:

$$\begin{aligned}
\omega^p \geq \omega_0^m \quad &\Leftrightarrow \quad \bar{e} \geq \frac{1}{10}a\sigma^2 s \\
\omega^p \geq \omega_1^m \quad &\Leftrightarrow \quad \bar{e} \leq a\sigma^2 s
\end{aligned}$$

- ω^p vs ω^c:

$$\begin{aligned}
\omega^p \leq \omega^c \quad &\Leftrightarrow \quad \frac{2}{5}s\bar{e} - \frac{2}{5}a\sigma^2 s^2 - \frac{\bar{e}^2}{10a\sigma^2} \geq 0 \\
&\Leftrightarrow \quad -\frac{\left(\bar{e} - 2a\sigma^2 s\right)^2}{10a\sigma^2} \geq 0 \\
&\Rightarrow \quad \text{impossible, hence } \omega^p > \omega^c
\end{aligned}$$

- ω^p vs $\bar{\omega}^p$: $\omega^p \leq \bar{\omega}^p \;\Leftrightarrow\; \bar{e} \geq \frac{3}{4}a\sigma^2 s$.

- $\bar{\omega}^p$ vs ω_0^m and ω_1^m:

$$\begin{aligned}
\bar{\omega}^p \geq \omega_0^m \quad &\Leftrightarrow \quad \bar{e} \geq \frac{79}{140}a\sigma^2 s \\
\bar{\omega}^p \geq \omega_1^m \quad &\Leftrightarrow \quad \bar{e} \geq -4a\sigma^2 s \text{ which always holds}
\end{aligned}$$

- ω_0^m vs ω_1^m: $\omega_0^m \leq \omega_1^m \;\Leftrightarrow\; \bar{e} \geq \frac{7}{10}a\sigma^2 s$.

- v_1 vs ω_0^m, ω_1^m, ω^p and $\bar{\omega}^p$:

$$
\begin{aligned}
v_1 &> \omega_1^m \text{ since } 1+\tfrac{\sqrt{5}}{5} > \tfrac{7}{5} \\
v_1 &\geq \omega_0^m \Leftrightarrow \bar{e} \geq \frac{21}{10\left(1+\sqrt{5}\right)} a\sigma^2 s \\
v_1 &\geq \omega^p \Leftrightarrow \bar{e} \geq \frac{2\sqrt{5}}{5} a\sigma^2 s \\
v_1 &< \bar{\omega}^p \text{ since } \tfrac{3}{2} > 1+\tfrac{\sqrt{5}}{5} \text{ and } \tfrac{9}{10} > \tfrac{7}{8}
\end{aligned}
$$

  Note that $v_2 < \omega_1^m$ hence the condition $W_{-1} < v_2$ is not going to bind, and therefore it is not useful to study the relative position of v_2.

- Overall, without condition on the parameters, we have: $\bar{\omega}^p \geq \omega_1^m$, $\omega^c < \omega_1^m$, $\omega^c < \omega^p$, $v_1 > \omega_1^m$, $\bar{\omega}^p > v_1$ and $v_2 < \omega_1^m$. For the other relationships, there are 9 thresholds, in ascending order: $\frac{1}{10}$, $\frac{79}{140}$, $\frac{21}{10\left(1+\sqrt{5}\right)}$, $\frac{7}{10}$, $\frac{3}{4}$, $3-2\sqrt{\frac{6}{5}}$, $\frac{2\sqrt{5}}{5}$, 1, $3+2\sqrt{\frac{6}{5}}$.

■

Lemma 13 *Positive tresholds. The following holds:*

- $\omega^p \geq 0 \Leftrightarrow \frac{\bar{e}}{a\sigma^2 s} \geq \frac{1}{2}$
- $\bar{\omega}^p \geq 0 \Leftrightarrow \frac{\bar{e}}{a\sigma^2 s} \geq \frac{7}{12} \approx 0.58$
- $\omega_0^m \geq 0 \Leftrightarrow \frac{\bar{e}}{a\sigma^2 s} \geq \frac{3}{10}$
- $v_1 \geq 0 \Leftrightarrow \frac{\bar{e}}{a\sigma^2 s} \geq \frac{9}{10\left(1+\frac{\sqrt{5}}{5}\right)} \approx 0.62$
- $\omega_1^m \geq 0 \Leftrightarrow \frac{\bar{e}}{a\sigma^2 s} \geq \frac{9}{14} \approx 0.64$

Proof. Immediate from the expressions of the thresholds given in Lemma 12. ■

B.5.3 Abundant capital

Proposition 3 **Proof.** For u_1, u_0 to be an equilibrium, it must be i) feasible, ii) weakly dominating c_1, u_0 and c_1, c_0. A necessary and sufficient condition for i) to hold is that $W_{-1} \geq \max\left(\omega_0^m, \omega_1^m\right)$ by Lemma 3. A necessary and sufficient condition for ii) to hold is that $W_{-1} \geq v_1$ or $W_{-1} \leq v_2$. Given that $v_2 \leq \omega_1^m$ and $v_1 > \omega_1^m$ for all parameter values, one can eliminate $W_{-1} \leq v_2$ from the equilibrium conditions (Lemma 5). Hence a sufficient condition is $W_{-1} \geq \max\left(v_1, \omega_0^m\right)$. Lemma 12 shows that $v_1 \geq \omega_0^m$ if and only if $\frac{\bar{e}}{a\sigma^2 s} \geq \frac{21}{10\left(1+\sqrt{5}\right)}$, hence the expression of $\bar{\omega}^m$. ■

Note that u_1, u_0 may be an equilibrium for a larger parameter interval, because Proposition 3 merely states a sufficient condition. There are cases (see Lemmata 19-22 below) where c_1, u_0 dominates u_1, u_0 but is either not time consistent or not feasible. This is the case in particular if $v_1 \leq \omega^p$ and / or $v_1 \geq \bar{\omega}^p$. If at the same time, u_1, u_0 dominates c_1, c_0, and / or c_1, c_0 is not time consistent, then u_1, u_0 is an equilibrium, but these cases are not accounted for by the proposition.

Corollary 2 Proof. Building on Proposition 3, if $\frac{\bar{e}}{a\sigma^2 s} \leq \frac{9}{10\left(1+\sqrt{5}\right)} \approx 0.65$, the equilibrium condition is $W_{-1} \geq \omega_0^m = \frac{4}{5}s\bar{e} - \frac{12}{25}a\sigma^2 s^2$. Since ω_0^m is highest for $\bar{e} \geq \frac{6}{5}a\sigma^2 s > \frac{21}{10\left(1+\sqrt{5}\right)}$, an increase in s decreases ω_0^m if $\frac{\bar{e}}{a\sigma^2 s} \leq \frac{21}{10\left(1+\sqrt{5}\right)}$. Conversely, if $\frac{\bar{e}}{a\sigma^2 s} > \frac{21}{10\left(1+\sqrt{5}\right)}$, the threshold is $v_1 = \left(1 + \frac{1}{\sqrt{5}}\right) s\bar{e} - \frac{9}{10}a\sigma^2 s^2$. Taking the first-order condition shows that v_1 is increasing in s if $\bar{e} \geq \frac{9}{10\left(1+\frac{1}{\sqrt{5}}\right)} \approx 0.62$. Given that $\frac{9}{10\left(1+\frac{1}{\sqrt{5}}\right)} < \frac{21}{10\left(1+\sqrt{5}\right)}$, an increase in s tightens the constraint if $\frac{\bar{e}}{a\sigma^2 s} \geq \frac{21}{10\left(1+\sqrt{5}\right)}$. ■

B.5.4 Intermediate capital

Proposition 4 Proof. The result requires that the u_1, u_0 and c_1, u_0 are both feasible and that c_1, u_0 is the dominant strategy. The first point requires $W_{-1} \geq \max\left(\omega_0^m, \omega_1^m\right)$ and $W_{-1} \in [\omega^p, \bar{\omega}^p[$ (by Lemmata 3 and 4). The second point requires $W_{-1} < v_1$ (Lemma 5). Hence we must have $W_{-1} \geq \max\left(\omega_0^m, \omega_1^m, \omega^p\right)$ and $W_{-1} < \min\left(v_1, \bar{\omega}^p\right)$. From Lemma 12, this interval is non-empty only in the two instances stated in the result. ■

Corollary 3 Proof. Recall the expressions of $x_0^{c_1,u_0}$ and $X_1^{c_1,u_0}$ from Lemma 4:

$$x_0^{c_1,u_0} = \frac{s}{2};\ X_1^{c_1,u_0} = \frac{a\sigma^2 s - \bar{e} + \sqrt{U^m}}{2a\sigma^2};$$

This implies that:

$$x_1^{c_1,u_0} = X_1^{c_1,u_0} - x_0^{c_1,u_0} = \frac{\sqrt{U^m} - \bar{e}}{2a\sigma^2}$$

- Clearly, $x_0^{c_1,u_0} > x_0^{u_1,u_0} = \frac{2}{5}s$ (Lemma 3).
- Consider the time-1 trades:

$$x_1^{c_1,u_0} - x_1^{u_1,u_0} = \frac{\sqrt{U^m} - \bar{e}}{2a\sigma^2} - \frac{3}{10}s \leq 0 \ \Leftrightarrow\ \sqrt{U^m} \geq \bar{e} + \frac{3}{5}a\sigma^2 s$$

  Taking squares on both sides and rearranging terms yields:

$$x_1^{c_1,u_0} - x_1^{u_1,u_0} \leq 0 \ \Leftrightarrow\ W_{-1} \leq \frac{8}{5}s\bar{e} - \frac{41}{5}a\sigma^2 s$$

By assumption, $W_{-1} < v_1$. Thus we must compare the threshold to v_1: $v_1 \leq \frac{8}{5}s\bar{e} - \frac{41}{5}a\sigma^2 s \Leftrightarrow \frac{\sqrt{5}-3}{5}\bar{e} \leq \frac{2}{25}a\sigma^2 s$, which is always true since $\sqrt{5} - 3 < 0$. Thus the condition $W_{-1} < v_1$ implies that $x_1^{c_1,u_0} \leq x_1^{u_1,u_0}$.

- Turning to the total position: $X_1^{u_1,u_0} = x_0^{u_1,u_0} + x_1^{u_1,u_0} = \frac{2}{5}s + \frac{3}{10}s = \frac{7}{10}s$ (see Lemma 3). Thus:

$$X_1^{c_1,u_0} \geq X_1^{u_1,u_0} \Leftrightarrow 5\sqrt{U^m} \geq 2a\sigma^2 s + 5\bar{e} \Leftrightarrow 25U^m \geq \left(2a\sigma^2 s + 5\bar{e}\right)^2$$

Substituting for U^m (Lemma 4) and simplifying gives:

$$X_1^{c_1,u_0} \geq X_1^{u_1,u_0} \Leftrightarrow W_{-1} \geq v' = \frac{7}{5}s\bar{e} - \frac{23}{25}a\sigma^2 s^2$$

We must compare v' to the lower bound of the interval $[\max(\omega^p, \omega_1^m), v_1]$, with $v_1 = \left(1 + \frac{1}{\sqrt{5}}\right)s\bar{e} - \frac{9}{10}a\sigma^2 s^2$, $\omega^p = s\bar{e} - \frac{1}{2}a\sigma^2 s^2$ and $\omega_1^m = \frac{7}{5}s\bar{e} - \frac{9}{10}a\sigma^2 s^2$. Clearly, $v' < \omega_1^m$. Further, $v' - \omega^p = \frac{2}{5}s\bar{e} - \frac{21}{50}a\sigma^2 s^2 \Leftrightarrow \frac{\bar{e}}{a\sigma^2 s} < \frac{21}{20}$. Since ω^p is the lower bound of the interval if $\frac{\bar{e}}{a\sigma^2 s} < 1$, this condition is satisfied. This implies that $v' < \omega^p$. Thus under the conditions of Proposition 4, $X_1^{c_1,u_0} \geq X_1^{u_1,u_0}$.

■

Corollary 4 Proof.

- Given that $\Delta_1 = 2a\sigma^2(s - X_1)$ (shown in the text), $X_1^{c_1,u_0} \geq X_1^{u_1,u_0}$ (Corollary 3) implies that $\Delta_1^{c_1,u_0} \leq \Delta_1^{u_1,u_0}$.

- At time 0, the equilibrium spreads are $\Delta_0^{u_1,u_0} = \frac{9}{5}a\sigma^2 s$ (Proposition 3) and $\Delta_0^{c_1,u_0} = 2a\sigma^2 s + \bar{e} - \sqrt{U^m}$ (Proposition 4). Thus:

$$\begin{aligned}
\Delta_0^{u_1,u_0} - \Delta_0^{c_1,u_0} < 0 \quad &\Leftrightarrow \quad 0 < \frac{1}{5}a\sigma^2 s + \bar{e} < \sqrt{U^m} \\
&\Leftrightarrow \quad \left(\frac{1}{5}a\sigma^2 s + \bar{e}\right)^2 < U^m = \left(\bar{e} - a\sigma s^2 s\right)^2 + 2a\sigma^2 W_{-1} + a^2\sigma^4 s^2 \\
&\Leftrightarrow \quad W_{-1} > \frac{6}{5}s\bar{e} - \frac{49}{50}a\sigma^2 s^2 = \omega'
\end{aligned}$$

One must compare the position of ω' relative to ω_1^m and ω^p:

  - Clearly, $\omega' < \omega_1^m = \frac{7}{5}s\bar{e} - \frac{9}{10}a\sigma^2 s^2$.
  - Consider the position of ω' relative to ω^p:

$$\omega' - \omega^p = \frac{1}{5}s\bar{e} - \frac{12}{25}a\sigma^2 s^2 < 0 \Leftrightarrow \frac{\bar{e}}{a\sigma^2 s} < \frac{12}{5}$$

Note that ω^p is the lower bound of the interval if $\frac{\bar{e}}{a\sigma^2 s} < 1$, thus the previous condition is satisfied. Hence $\Delta_0^{c_1,u_0} \leq \Delta_0^{u_1,u_0}$.

■

B.5.5 Full characterization of the equilibrium

For the sake of exhaustiveness, I now derive the equilibrium for each of the parameter regions of Lemma 12. Since the algorithm for the equilibrium determination is always the same, I write the proof only in the first case. I first recall some notation.

Notation 7 *The following notation is used as a shorthand:*

- *Weakly dominant strategy:*
  - u_1, u_0 *dominates* c_1, c_0 *if and only if* $f(W_{-1}) \geq 0$ *(Lemma 9),*
  - *If* $x_0^{u_1,c_0} = x_0^1$*, the strategy* u_1, c_0 *dominates* c_1, c_0 *if and only if* $g(W_{-1}) \geq 0$ *(Lemma 10),*
  - *If* $x_0^{u_1,c_0} = x_0^0$*, the strategy* u_1, c_0 *dominates* c_1, u_0 *if and only if* $\tilde{g}_u(W_{-1}) \geq 0$ *(Lemma 11),*
  - *If* $x_0^{u_1,c_0} = x_0^0$*, the strategy* u_1, c_0 *dominates* c_1, c_0 *if and only if* $\tilde{g}_c(W_{-1}) \geq 0$ *(Lemma 11).*
- *Time consistency condition:*
  - c_1, c_0 *is time consistent if and only if* $h(W_{-1}) > 0$*, where* h *is defined in Lemma 3.*

I now present the equilibrium for each parameter region:

Lemma 14 ***Case A:*** *If* $\frac{\bar{e}}{a\sigma^2 s} \in \left[0, \frac{1}{10}\right[$*, then* $\omega^c < \omega_1^m \leq v_1 \leq \bar{\omega}^p \leq \omega^p < \omega_0^m < 0$*,*

1. *If* $W_{-1} \geq 0$*, the equilibrium is* u_1, u_0*,*

Proof. If $W_{-1} \geq 0$, u_1, u_0 is feasible, and u_1, c_0 is feasible ($W_{-1} \geq \omega^p$) but not time consistent ($W_{-1} \geq \bar{\omega}^p$, see Lemma 4). Given that $W_{-1} \geq v_1$, u_1, u_0 would dominate c_1, u_0, hence dominates c_1, c_0. Hence in equilibrium, the arbitrageur chooses strategy u_1, u_0. ■

Lemma 15 ***Case B:*** *If* $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{1}{10}, \frac{79}{140}\right[$*, then* $\omega_0^m < 0$ *and* $\omega^c < \omega_1^m \leq v_1 \leq \bar{\omega}^p < \omega_0^m \leq \omega^p$*.*

1. *If* $W_{-1} \geq \omega^p$ *or* $W_{-1} \in [0, \omega^p[$*, the equilibrium is* u_1, u_0*.*

Proof. Same as A.1. ■

Lemma 16 ***Case C:*** *If* $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{79}{140}, \frac{21}{10(1+\sqrt{5})}\right[$, *then* $\omega^c < 0$ *and all parameters are positive on part of the interval. They are ranked in the following order:* $\omega^c \leq \omega_1^m \leq v_1 \leq \omega_0^m \leq \bar{\omega}^p \leq \omega^p$.

1. *If* $W_{-1} \geq \omega_0^m$, *the equilibrium is* u_1, u_0.

2. *If* $W_{-1} \in [0, \omega_0^m[$, *the equilibrium is* u_1, c_0, *with* $x_0^{u_1,c_0} = x_0^0$ *if* $W_{-1} \geq \omega_1^m$, *and* $\min\left(x_0^0, x_0^1\right)$ *otherwise.*

Proof. C.1: u_1, u_0 is feasible and dominates c_1, u_0, whether it is feasible or not.

C.2: u_1, u_0 is not feasible anymore ($W_{-1} < \omega_0^m$). $W_{-1} < \omega^p$ implies that c_1, u_0 is not feasible (Lemma 4). Thus, one must compare c_1, c_0 and u_1, c_0. From Lemma 4, the strategy c_1, c_0 is not time consistent, as the interval $\left[\frac{79}{140}, \frac{21}{10(1+\sqrt{5})}\right[$ is included in $\left[\frac{1}{3}, \frac{2}{3}\right]$. Thus the arbitrageur's strategy in equilibrium is u_1, c_0, and $x_0^{u_1,c_0}$ depends on which constraints binds most severely, as in Lemma 3. ■

Lemma 17 ***Case D:*** *If* $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{21}{10(1+\sqrt{5})}, \frac{2}{3}\right[$, *then* $\omega^c < \omega_1^m < \omega_0^m \leq v_1 \leq \bar{\omega}^p < \omega^p$,

1. *If* $W_{-1} \geq v_1$ *or* $W_{-1} \in [\omega_0^m, v_1[$, *the equilibrium is* u_1, u_0.

2. *If* $W_{-1} \in [\omega_1^m, \omega_0^m[$, *the equilibrium is* u_1, c_0 *with* $x_0^{u_1,c_0} = x_0^0$

3. *If* $W_{-1} \in [\omega^c \wedge 0, \omega_0^m[$, *the equilibrium is* u_1, c_0, *with* $x_0^{u_1,c_0} = \min\left(x_0^0, x_0^1\right)$.

Proof. See case D'. ■

Lemma 18 ***Case D':*** *If* $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{2}{3}, \frac{7}{10}\right[$, *then* $\omega^c < \omega_1^m < \omega_0^m \leq v_1 \leq \bar{\omega}^p < \omega^p$ *(same order as in case D),*

1. *If* $W_{-1} \geq v_1$, *the equilibrium is* u_1, u_0.

2. *If* $W_{-1} \in [\omega_0^m, v_1[$, *then the equilibrium is* u_1, u_0 *when* $\tilde{D}^w < 0$, *or if* $W_{-1} \geq \tilde{\omega}$ *when* $\tilde{D}^w > 0$, *or if* $f(W_{-1}) \geq 0$ *when* $\tilde{D}^w > 0$ *and* $W_{-1} < \tilde{\omega}$. *Otherwise, i.e. if* $\tilde{D}^w > 0$ *and* $W_{-1} < \tilde{\omega}$ *and* $f(W_{-1}) < 0$, *the equilibrium is* c_1, c_0.

3. *If* $W_{-1} \in [\omega_1^m, \omega_0^m[$, *then the equilibrium is* u_1, c_0 *with* $x_0^{u_1,c_0} = x_0^0$ *if* $\tilde{D}^w < 0$, *or if* $W_{-1} \geq \tilde{\omega}$ *when* $\tilde{D}^w > 0$, *or if* $\tilde{g}_u(W_{-1}) \geq 0$ *when* $\tilde{D}^w > 0$ *and* $W_{-1} < \tilde{\omega}$. *Otherwise, the equilibrium is* c_1, c_0.

4. *If* $W_{-1} \in [\omega^c \wedge 0, \omega_1^m[$, *the equilibrium is* u_1, c_0, *with* $x_0^{u_1,c_0} = \min\left(x_0^0, x_0^1\right)$ *if* $\tilde{D}^w < 0$, *or if* $W_{-1} \geq \tilde{\omega}$ *when* $\tilde{D}^w > 0$, *or if* $g(W_{-1}) \geq 0$ *when* $\tilde{D}^w > 0$ *and* $W_{-1} < \tilde{\omega}$. *Otherwise, the equilibrium is* c_1, c_0.

Proof. For D and D'. D'.1: $W_{-1} \geq v_1$. Same as C.1.

D.1. and D'.2: $W_{-1} < v_1$ implies that the u_1, u_0 strategy is dominated by the c_1, u_0 strategy (Lemma 5), but c_1, u_0 is not feasible ($W_{-1} < \omega^p$, Lemma 4). Thus one must compare u_1, u_0 to c_1, c_0, if c_1, c_0 is time-consistent. By Lemma 4, c_1, c_0 is not time consistent if $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{21}{10\left(1+\sqrt{5}\right)}, \frac{2}{3}\right[$. If $\left[\frac{2}{3}, \frac{7}{10}\right[$, c_1, c_0 is time-consistent if and only if $\tilde{D}^w > 0$ and $W_{-1} < \tilde{\omega}$. In this case, c_1, c_0 dominates u_1, u_0 if and only if $f\left(W_{-1}\right) < 0$ (Lemma 9). If c_1, c_0 is not time consistent, the equilibrium is u_1, u_0, as it dominates u_1, c_0 by definition.

D.2 and D'3: $W_{-1} < \omega_0^m$ implies that u_1, u_0 is not feasible anymore, thus the arbitrageur chooses between u_1, c_0 and c_1, c_0. The result follows from applying Lemma 4 and Lemma 10 along the same lines as in D.2.

D.3 and D'.4: $W_{-1} < \omega_0^m \leq \omega_1^m$ implies that $x_0^{u_1,c_0} = \min\left(x_0^0, x_0^1\right)$. The rest follows from applying a similar reasoning as D.2. ■

Lemma 19 ***Case E:*** *If* $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{7}{10}, \frac{3}{4}\right[$, *then* $\omega^c < \omega_0^m \leq \omega_1^m \leq v_1 \leq \bar{\omega}^p < \omega^p$,

1. *If* $W_{-1} \geq v_1$, *the equilibrium is* u_1, u_0.
2. *If* $W_{-1} \in [\omega_1^m, v_1[$, *the equilibrium is as D'.2.*
3. *If* $W_{-1} \in [\omega_0^m, \omega_1^m[$, *then the equilibrium as in D'.4.*
4. *If* $W_{-1} \in [\omega^c, \omega_0^m[$, *then the equilibrium as in D'.3.*

Lemma 20 ***Case F:*** *If* $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{3}{4}, 3 - 2\sqrt{\frac{6}{5}}\right[$, *then* $\omega^c < \omega_0^m \leq \omega_1^m \leq v_1 < \omega^p \leq \bar{\omega}^p$,

1. *If* $W_{-1} \geq v_1$, *then the equilibrium is* u_1, u_0,
2. *If* $W_{-1} \in [\omega_1^m, v_1[$, *then the equilibrium is as D'.2.*
3. *If* $W_{-1} \in [\omega_0^m, \omega_1^m[$, *then the equilibrium as in D'.4.*
4. *If* $W_{-1} \in [\omega^c, \omega_0^m[$, *then the equilibrium as in D'.3.*

Lemma 21 ***Case G:*** *If* $\frac{\bar{e}}{a\sigma^2 s} \in \left[3 - 2\sqrt{\frac{6}{5}}, \frac{2\sqrt{5}}{5}\right[$, *then* $\omega_0^m \leq \omega^c \leq \omega_1^m \leq v_1 \leq \omega^p \leq \bar{\omega}^p$,

1. *If* $W_{-1} \geq v_1$, *then the equilibrium is* u_1, u_0,
2. *If* $W_{-1} \in [\omega_1^m, v_1[$, *then the equilibrium is as D'.2.*
3. *If* $W_{-1} \in [\omega_0^m, \omega_1^m[$, *then the equilibrium as in D'.4.*
4. *If* $W_{-1} \in [\omega^c, \omega_0^m[$, *then the equilibrium as in D'.3.*

Lemma 22 ***Case H:*** *If* $\frac{\bar{e}}{a\sigma^2 s} \in \left[\frac{2\sqrt{5}}{5}, 1\right[$, *then* $\omega_0^m \le \omega^c \le \omega_1^m \le \omega^p \le v_1 \le \bar{\omega^p}$,

1. *If* $W_{-1} \ge v_1$, *then the equilibrium is* u_1, u_0,
2. *If* $W_{-1} \in [\omega^p, v_1[$, *then the equilibrium is* c_1, u_0,
3. *If* $W_{-1} \in [\omega_1^m, \omega^p[$, *then the equilibrium is as D'.2.*
4. *If* $W_{-1} \in [\omega^c, \omega_1^m[$, *the equilibrium is as in D'.4.*

Proof. H.1: same as D'.1.

H.2: Since $W_{-1} \ge \max(\omega_0^m, \omega_1^m)$, u_1, u_0 is feasible. However, $W_{-1} < v_1$ implies that c_1, u_0 yields a higher payoff (Lemma 5). This implies that c_1, u_0 also yields a higher payoff than any strategy u_1, c_0 (since it is dominated by u_1, u_0). Further, $W_{-1} \in [\omega^p, v_1[$ implies that $W_{-1} \in [\omega^p, \bar{\omega}^p[$, thus c_1, u_0 is feasible and time-consistent (Lemma 4). Thus, in equilibrium, the arbitrageur plays c_1, u_0. ∎

Lemma 23 ***Case I:*** *If* $\frac{\bar{e}}{a\sigma^2 s} \in \left[1, 3 + 2\sqrt{\frac{6}{5}}\right[$, *then* $\omega_0^m < \omega^c \le \omega^p \le \omega_1^m \le v_1 \le \bar{\omega^p}$,

1. *If* $W_{-1} \ge v_1$, *the equilibrium is* u_1, u_0,
2. *If* $W_{-1} \in [\omega^p, v_1[$, *the equilibrium is* c_1, u_0,
3. *If* $W_{-1} \in [\omega^c, \omega^p[$, *then*
   - *If* $\frac{\bar{e}}{a\sigma^2 s} \in [1, 2[$, *the equilibrium is* c_1, c_0,
   - *If* $\frac{\bar{e}}{a\sigma^2 s} \in [2, \iota_1[$ *and* $g(W_{-1}) \ge 0$, *or if* $\frac{\bar{e}}{a\sigma^2 s} \in \left[2, 3 + 2\sqrt{\frac{6}{5}}\right[$, *the equilibrium is* u_1, c_0 *if* $g(W_{-1}) \ge 0$ *and*
     - *If* $\frac{\bar{e}}{a\sigma^2 s} \in [1, 2[$, *the equilibrium is as in D'.4,*
     - *If* $\frac{\bar{e}}{a\sigma^2 s} \in \left[2, 3 + 2\sqrt{\frac{6}{5}}\right[$, *the equilibrium is* u_1, c_0 *with* $x_0^{u_1,c_0} = x_0^1$ *if* $g(W_{-1}) \ge 0$ *and* u_1, c_0 *is feasible as given by Lemma 3. Otherwise, the equilibrium is* c_1, c_0.

Proof. I.1: same as F.1.

I.2: same as H.2.

II.3: $W_{-1} < \omega^p < \max(\omega_0^m, \omega_1^m)$ implies that the arbitrage has not enough capital for the u_1, u_0 and c_1, u_0 strategies.

- If $\frac{\bar{e}}{a\sigma^2 s} \in [1, 2[$, the proof is similar to D'.4.
- If $\frac{\bar{e}}{a\sigma^2 s} \in \left[2, 3 + 2\sqrt{\frac{6}{5}}\right[$, then the proof follows by combining Lemmata 3, 4, and 10 ($\iota_2 < 3 + 2\sqrt{\frac{6}{5}}$).

■

Lemma 24 ***Case J:*** *If $\frac{\bar{e}}{a\sigma^2 s} > 3 + 2\sqrt{\frac{6}{5}}$, then $\omega^c \leq \omega_0^m \leq \omega^p \leq \omega_1^m \leq v_1 \leq \bar{\omega^p}$,*

1. *If $W_{-1} \geq v_1$, the equilibrium is u_1, u_0,*
2. *If $W_{-1} \in [\omega^p, v_1[$, the equilibrium is c_1, u_0,*
3. *If $W_{-1} \in [\omega^c, \omega^p[$, the equilibrium is u_1, c_0 if $\alpha W_{-1}^2 + \beta W_{-1} + \gamma < 0$ and $g(W_{-1}) \geq 0$, and c_1, c_0 otherwise.*

Proof. J.1: same as F.1.

J.2: same as H.2.

J.3: For this level of capital, the arbitrageur cannot play the u_1, u_0 and the c_1, u_0 strategies. Thus the arbitrageur chooses between u_1, c_0 if it is feasible and c_1, c_0. Since $\frac{\bar{e}}{a\sigma^2 s} > \iota_2$, we are in the same case as I.3. second bullet point. ■

C Liquidity and welfare comparisons

C.1 Proposition 5

Proof. First I recall the expressions of the spread at time 0 in both cases:

$$\Delta_0^m = \frac{9}{5}a\sigma^2 s;\ \Delta_0^* = 2\left(a\sigma^2 s + \bar{e}\right) - \sqrt{Q} - \sqrt{U}$$

with $Q = \left(\bar{e} - a\sigma^2 s\right)^2 + 2a\sigma^2 W_{-1}$ and $U = \left(\bar{e} - a\sigma^2 s\right)^2 + 4a\sigma^2 x_0 \bar{e}$. It is convenient to rewrite U by plugging the expression for the time-0 constrained trade given in Corollary 1:

$$\begin{aligned} U &= \left(\bar{e} - a\sigma^2 s\right)^2 + 2\bar{e}\left(a\sigma^2 s - \bar{e}\right) + 2\bar{e}\sqrt{Q} \\ &= a^2\sigma^4 s^2 - \bar{e}^2 + 2\bar{e}\sqrt{Q} \end{aligned}$$

$$\Delta_0^* \geq \Delta_0^m \ \Leftrightarrow\ \frac{1}{5}a\sigma^2 s + 2\bar{e} \geq \sqrt{Q} + \sqrt{U}$$

This implies that

$$\frac{1}{5}a\sigma^2 s + 2\bar{e} \geq Q + U + 2\sqrt{QU}$$

Since $Q + U = 2a^2\sigma^4 s^2 - 2a\sigma^2 s\bar{e} + 2\bar{e}\sqrt{Q} + 2a\sigma^2 W_{-1}$, the condition becomes, after regrouping terms:

$$-2a\sigma^2 W_{-1} + \frac{14}{5}a\sigma^2 s\bar{e} - \frac{49}{25}a^2\sigma^4 s^2 + 4\bar{e}^2 \geq 2\sqrt{Q}\left(\bar{e} + \sqrt{U}\right)$$

A necessary condition for this inequality to be satisfied is that the LHS is positive, i.e.

$$W_{-1} \leq \tilde{\omega} = \frac{7}{5}s\bar{e} - \frac{49}{50}a\sigma^2 s^2 + \frac{2\bar{e}^2}{a\sigma^2} \tag{69}$$

The interval of interest to compare $\tilde{\omega}$ to is $[\omega_0^m, \omega^*[$ if $\frac{\bar{e}}{a\sigma^2 s} < \frac{21}{10(1+\sqrt{5})}$, $[v_1, \omega^*[$ if $\frac{\bar{e}}{a\sigma^2 s} \geq \frac{21}{10(1+\sqrt{5})}$ and $[v_1, \omega^*[$ if $\frac{\bar{e}}{a\sigma^2 s} \geq \frac{7}{10}$ and $h(W_{-1}) \geq 0$ or $h(W_{-1}) < 0$ and $f(W_{-1}) \geq 0$. Note that the latter case is not given in Proposition 3, which gives only a necessary condition and not a sufficient one. I compare $\tilde{\omega}$ to the different thresholds:

$$v_1 \geq \tilde{\omega} \;\Leftrightarrow\; \frac{5\left(\sqrt{5}-2\right)a\sigma^2 s\bar{e} - 2a^2\sigma^4 s^2 - 50\bar{e}^2}{25a\sigma^2} \geq 0$$

We can consider the numerator of the LHS as a second-order equation in $\bar{e}$. Calculating the discriminant $d = \left(25\left(\sqrt{5}-2\right)^2 - 400\right)a^2\sigma^4 s^2 < 0$ shows that the LHS is always negative (since the second-order term has a negative coefficient), i.e. $v_1 < \tilde{\omega}$.

Next, I compare $\tilde{\omega}$ and ω_0^m:

$$\omega_0^m \leq \tilde{\omega} \;\Leftrightarrow\; \frac{-6a\sigma^2 s\bar{e} + 5a^2\sigma^4 s^2 - 20\bar{e}^2}{10a\sigma^2}$$

The LHS in the numerator can be seen as a second-order equation in $\bar{e}$, I calculate its discriminant: $d = 436a^2\sigma^4 s^2 > 0$. Since the constant and the second-order term have opposite signs, there is a positive and a negative root. The positive root is equal to $\frac{3+\sqrt{109}}{20}a\sigma^2 s \approx 0.68a\sigma^2 s > \frac{21}{10(1+\sqrt{5})}a\sigma^2 s \approx 0.65a\sigma^2 s$. Hence if $\frac{\bar{e}}{a\sigma^2 s} < \frac{21}{10(1+\sqrt{5})}$, (69) does not hold.

To complete the time-0 case, I assume that $h(W_{-1}) > 0$ or $h(W_{-1}) \leq 0$ and $f(W_{-1}) \geq 0$. The relevant threshold for the monopoly is then ω_1^m:

$$\omega_1^m \leq \tilde{\omega} \;\Leftrightarrow\; \frac{a^2\sigma^4 s^2}{25} \leq \bar{e}^2 \;\Leftrightarrow \bar{e} \geq \frac{a\sigma^2 s}{5} \;\; (\bar{e} > 0)$$

Next, I compare the time-1 spreads, $\Delta_1^* = a\sigma^2 s + \bar{e} - \sqrt{U}$ and $\Delta_1^m = \frac{3}{5}a\sigma^2 s$:

$$\Delta_1^* \geq \Delta_1^m \;\Leftrightarrow\; \frac{4}{5}a\sigma^2 s\bar{e} + 2\bar{e}^2 - \frac{21}{25}a^2\sigma^2 s^4 \geq 2\bar{e}\sqrt{Q} \tag{70}$$

I study the sign of the LHS, taking it as a second-order equation in $\bar{e}$. The discriminant is $d = \frac{184}{25}a^2\sigma^4 s^2$ and given that the constant and the second-order term have opposite signs, there is a positive and a negative root. The positive root is $\frac{\sqrt{46}-2}{10}a\sigma^2 s \approx 0.48a\sigma^2 s$. Hence for $\frac{\bar{e}}{a\sigma^2 s} \leq \frac{\sqrt{46}-2}{10}$, $\Delta_1^* \leq \Delta_1^m$. Otherwise, one can take the square in each side of inequality (70), which gives, after a few lines of algebra:

$$\Delta_1^* \geq \Delta_1^m \;\Rightarrow\; W_{-1} \leq \hat{\omega} \equiv \frac{7}{5}s\bar{e} - \frac{21}{25}a\sigma^2 s^2 + \frac{441}{5000}\frac{a^3\sigma^6 s^4}{\bar{e}^2} - \frac{21}{125}\frac{a^2\sigma^4 s^3}{\bar{e}}$$

To assess the existence of this case, I compare $\hat{\omega}$ to the thresholds v_1, ω_0^m and ω_1^m.

$$\hat{\omega} \geq v_1 \;\Leftrightarrow\; \frac{2-\sqrt{5}}{5}s\bar{e} + \frac{3}{50}a\sigma^2 s^2 + \frac{21}{125}\frac{a^2\sigma^4 s^2}{\bar{e}}\left[\frac{21}{40}\frac{a\sigma^2 s}{\bar{e}} - 1\right] \geq 0$$

It is not possible to derive the roots of this equation analytically. However, one can look at sufficient conditions. There are two options and both are not satisfied. i) It is enough that:

$$\begin{cases} \frac{21}{40}\frac{a\sigma^2 s}{\bar{e}} - 1 \geq 0 \\ \frac{2-\sqrt{5}}{5}s\bar{e} + \frac{3}{50}a\sigma^2 s^2 \geq 0 \end{cases} \Leftrightarrow \begin{cases} \frac{\bar{e}}{a\sigma^2 s} \leq 0.525 \\ \frac{\bar{e}}{a\sigma^2 s} \leq \frac{3}{10\left(\sqrt{5}-2\right)} \approx 1.27 \end{cases} \Leftrightarrow \frac{\bar{e}}{a\sigma^2 s} \leq 0.525$$

This is not compatible with the initial assumption that $\frac{\bar{e}}{a\sigma^2 s} \geq \frac{21}{10\left(1+\sqrt{5}\right)} \approx 0.65$. ii) Another sufficient condition is:

$$\begin{cases} \frac{3}{50}a\sigma^2 s^2 \geq \frac{21}{125}\frac{a^2\sigma^4 s^3}{\bar{e}} \\ \frac{441}{5000\bar{e}^2}a^3\sigma^6 s^4 \geq \frac{\sqrt{5}-2}{5}s\bar{e} \end{cases} \Leftrightarrow \begin{cases} \frac{\bar{e}}{a\sigma^2 s} \geq \frac{14}{5} \\ \frac{\bar{e}}{a\sigma^2 s} \leq \frac{(441)^{\frac{1}{3}}}{10\left(\sqrt{5}-2\right)^{\frac{1}{3}}} \approx 1.23 \end{cases}$$

These two conditions are contradictory.

Now compare $\hat{\omega}$ and ω_0^m:

$$\hat{\omega} \geq \omega_0^m \;\Leftrightarrow\; \frac{3}{5}s\bar{e} - \frac{9}{25}a\sigma^2 s^2 + \frac{441}{5000}\frac{a^3\sigma^6 s^3}{\bar{e}^2} - \frac{21}{125}\frac{a^2\sigma^4 s^3}{\bar{e}} \geq 0$$

Again, there are two types of sufficient conditions and both are not satisfied. i) It is enough that:

$$\begin{cases} \frac{3}{5}s\bar{e} \geq \frac{9}{25}a\sigma^2 s^2 \\ \frac{441}{5000}\frac{a^3\sigma^6 s^3}{\bar{e}^2} \geq \frac{21}{125}\frac{a^2\sigma^4 s^3}{\bar{e}} \end{cases} \Leftrightarrow \begin{cases} \frac{\bar{e}}{a\sigma^2 s} \geq \frac{3}{5} \\ \frac{\bar{e}}{a\sigma^2 s} \leq \frac{21}{40} \end{cases}$$

These two conditions contradict each other. ii) Another sufficient set of sufficient conditions is:

$$\begin{cases} \frac{3}{5}s\bar{e} \geq \frac{21}{125}\frac{a^2\sigma^4 s^3}{\bar{e}} \\ \frac{441}{5000}\frac{a^3\sigma^6 s^3}{\bar{e}^2} \geq \frac{9}{25}a\sigma^2 s^2 \end{cases} \Leftrightarrow \begin{cases} \frac{\bar{e}}{a\sigma^2 s} \geq \frac{\sqrt{7}}{5} \approx 0.53 \\ \frac{\bar{e}}{a\sigma^2 s} \leq \frac{21}{15\sqrt{8}} \approx 0.49 \end{cases}$$

Again these two conditions contradict each other. Last I compare ω_1^m and $\hat{\omega}$:

$$\hat{\omega} \geq \omega_1^m \;\Leftrightarrow\; \frac{3}{50}a\sigma^2 s^2 + \frac{441}{5000}\frac{a^3\sigma^6 s^3}{\bar{e}^2} - \frac{21}{125}\frac{a^2\sigma^4 s^3}{\bar{e}} \geq 0$$

We can rewrite the LHS as $\frac{a\sigma^2 s\left(300\bar{e}^2+441a^2\sigma^4 s^2-4200a\sigma^2 s\bar{e}\right)}{5000\bar{e}^2}$, which can be seen as a second-order equation in $\bar{e}$. The discriminant of the numerator is $d=\left(4200^2-1200.441\right)a^2\sigma^4 s^2 \equiv \bar{q}^2 a^2\sigma^4 s^2>0$. The constant and the second-order term has the same sign, and the first-order term is negative, thus there are two positive roots. The smallest root is equal to $(7-q)\,a\sigma^2 s \approx 0.1a\sigma^2 s$, with $q=\frac{\bar{q}}{600}$, which is lower than the threshold $\frac{21}{10\left(1+\sqrt{5}\right)}a\sigma^2 s$, and the largest root is $(7+q)\,a\sigma^2 s \approx 14a\sigma^2 s > \frac{21}{10\left(1+\sqrt{5}\right)}a\sigma^2 s$. Hence if $\frac{\bar{e}}{a\sigma^2 s}\geq 7+q$, $\Delta_1^*\geq\Delta_1^m$ if $W_{-1}\in[\omega_1^m,\hat{\omega}]$. ■

C.2 Proposition 6

Proof. The equilibrium spreads in the constrained competitive and voluntarily constrained cases are given in Corollary 1 and Proposition 4. I consider the time-1 spreads:

$$\Delta_1^*\geq\Delta_1^{c_1,u_0} \;\Leftrightarrow\; U^m\geq U \;\Leftrightarrow\; 2\bar{e}^2+a^2\sigma^4 s^2-2a\sigma^2 s\bar{e}+2a\sigma^2 W_{-1}\geq 2\bar{e}\sqrt{Q}$$

I first check the sign of the LHS. It is positive if and only if $W_{-1}\geq s\bar{e}-\frac{1}{2}a\sigma^2 s^2-\frac{\bar{e}^2}{a\sigma^2}=\omega^p-\frac{\bar{e}^2}{a\sigma^2}$. This inequality is always satisfied since by definition $W_{-1}\geq\omega^p$ in the c_1,u_0 equilibrium. I rewrite the LHS as $2a\sigma^2\left(W_{-1}-\left(\omega^p-\frac{\bar{e}^2}{a\sigma^2}\right)\right)$. Thus:

$$\Delta_1^*\geq\Delta_1^{c_1,u_0} \;\Leftrightarrow\; 4a^2\sigma^4\left(W_{-1}-\left(\omega^p-\frac{\bar{e}^2}{a\sigma^2}\right)\right)^2\geq 4\bar{e}^2 Q=4\bar{e}^2\left(\bar{e}-a\sigma^2 s\right)^2+8a\sigma^2\bar{e}^2 W_{-1}$$

Developing each side and skipping some lines of algebra, I find that the previous inequality is equivalent to:

$$W_{-1}^2-2\omega^p W_{-1}+s^2\left(\bar{e}-\frac{1}{2}a\sigma^2 s\right)^2\geq 0$$

Since $s^2\left(\bar{e}-\frac{1}{2}a\sigma^2 s\right)^2=\left(s\bar{e}-\frac{1}{2}a\sigma^2 s^2\right)^2=(\omega^p)^2$, the LHS is equal to $(W_{-1}-\omega^p)^2$ which is always positive. This proves the result about time-1 spreads.

■

C.3 Corollary 5

Proof. I showed in the proof of Proposition 6 that the capital gain is larger in the monopoly case by showing that $U^m\geq U$. Further, recalling that $x_0^{c_1,c0}=\frac{s}{2}-\frac{\bar{e}-\sqrt{Q}}{2a\sigma^2}$ and $x_0^{c_1,u_0}=\frac{s}{2}$, I get:

$$x_0^{c_1,u_0}\leq x_0^{c_1,c_0} \;\Leftrightarrow\; \sqrt{Q}\geq\bar{e}\Rightarrow W_{-1}\geq s\bar{e}-\frac{1}{2}a\sigma^2 s^2=\omega^p$$

This is always true for the c_1,u_0 equilibrium under consideration.

Clearly, i) $x_0^{c_1,c_0}\geq x_0^{c_1,u_0}$ and ii) the fact that $x_0\left(\Delta_0-\Delta_1\right)$ is larger in the voluntarily constrained case imply that $\Delta_0 - \Delta_1$ is larger in the voluntarily constrained case.

Since by Proposition 6, $\Delta_1^* \geq \Delta_1^{c_1,u_0}$, the previous result implies that $\Delta^* \leq \Delta_0^{c_1,u_0}$.

This in turn implies that $\frac{\Delta_0}{\Delta_1}$ is larger in the voluntarily constrained case than in the constrained competitive case. Since Δ_1 is smaller in this case, this means that the arbitrage converges more quickly in the voluntarily constrained case at time 0, and more slowly at time 1.

■

C.4 Lemma 6

Proof. Starting from equation (20), and using the expression of local investors' demand, $\mathbb{E}_0(p_1^a) - p_0^A = a\sigma^2\left(Y_0^A + s_0\right)$, we can rewrite local investors' equilibrium utility as

$$\begin{aligned}\chi_0^A &= \frac{\left(\mathbb{E}_0\left(p_1^A\right)-p_0^A\right)^2+\left(D_1-p_1\right)^2}{a\sigma^2} - s\left(\mathbb{E}_0\left(p_1^A\right)-p_0^A+D_1-p_1\right) - \frac{a\sigma^2}{2}\left[\left(Y_0^A+s\right)^2+\left(Y_1^A+s\right)^2\right] \\ &= \frac{\left(\mathbb{E}_0\left(p_1^A\right)-p_0^A\right)^2+\left(D_1-p_1\right)^2}{2a\sigma^2} - s\left(\mathbb{E}_0\left(p_1^A\right)-p_0^A+D_1-p_1\right) \qquad (71)\end{aligned}$$

Given that risk premia are symmetric, we have: $\phi_0^A = D - p_0^A = \frac{\Delta_0}{2}$ and $\mathbb{E}_0\left(-\phi_1^A\right) = \mathbb{E}_0\left(p_1^A - D_1\right) = -\frac{\Delta_1}{2}$. Further, $\mathbb{E}_0\left(\Delta_1\right) = \Delta_1$ since the spread, unlike the individual price, does not depend on ϵ_1. Hence local investors' welfare is

$$\chi_0^A = \frac{\left(\Delta_0-\Delta_1\right)^2+\Delta_1^2}{8a\sigma^2} - \frac{s}{2}\Delta_0 \qquad (72)$$

When unconstrained competitive arbitrageurs are present, all spreads are 0, as shown in Proposition 1, hence $\chi_0^{A,*} = 0$. In the autarky situation, local investors are constrained to hold their local asset in equilibrium. Hence using market-clearing and investors' demand functions, we have $Y_1^A = 0$, hence $Y_1^A + s = s$, which implies from investors' demand that $p_1^A = D_1 - a\sigma^2 s$. At time 0, by market-clearing, $Y_0^A = 0$, hence $Y_0^A + s = \frac{\mathbb{E}_0\left(p_1^A\right)-p_0^A}{a\sigma^2} = s$. Hence $p_0^A = \mathbb{E}_0\left(p_1^A\right) - a\sigma^2 s = D - 2a\sigma^2 s$. The prices in market B are opposite, by construction. Plugging this prices into the local investors's welfare function gives $\chi_0^{A,a} = -a\sigma^2 s$.

To rank $\chi_0^{A,m}$ relative to $\chi_0^{A,*}$ and $\chi_0^{A,a}$, note that

$$\begin{cases} 0 \leq \mathbb{E}_0\left(p_1^A\right) - p_0^A \leq 2a\sigma^2 s \\ 0 \leq D_1 - p_1^A \leq 2a\sigma^2 s \end{cases}$$

Using these four inequalities and equation (71) yields the result.

Finally, it is clear that in the full insurance and autarky cases, arbitrageurs do not make any profit, while they do in any of the monopolistic cases. ■

C.5 Corollaries 6 and 7

Proof. The comparative statics in Corollary 6 obtain by differentiation from equation (72). The first part of Corollary 7 follows immediately. The second part about aggregate welfare is proved on an example in the text. ■

C.6 Corollary 8

Proof. The effect of imposing financial constraints when parameters are such that conditions in Proposition 4 are satisfied is to increase the arbitrageur's payoff from $J_0^{u_1,u_0}$ to $J_0^{c_1,u_0}$. The effect on local investors' welfare (e.g. in market A) is $\frac{\partial \chi_0^A}{\partial \Delta_0} d\Delta_0 + \frac{\partial \chi_0^A}{\partial \Delta_1} d\Delta_1$. From Corollary 4, we know that $d\Delta_0 = \Delta_0^{c_1,u_0} - \Delta_0^{u_1,u_0} < 0$ and $d\Delta_1 = \Delta_1^{c_1,u_0} - \Delta_1^{u_1,u_0} < 0$ when imposing the constraint. Using Corollary 6, we get that $\frac{\partial \chi_0^A}{\partial \Delta_0} < 0$, and $\frac{\partial \chi_0^A}{\partial \Delta_1} < 0$ if $\Delta_1 < \frac{1}{2}\Delta_0$. I now show that this condition is verified.

From Proposition 4, we have $\frac{\Delta_1^{c_1,u_0}}{\Delta_0^{c_1,u_0}} = 1 - \frac{a\sigma^2 s}{\Delta_0^{c_1,u_0}}$. Thus $\Delta_1 < \frac{1}{2}\Delta_0 \Leftrightarrow \frac{\bar{e} - \sqrt{U^m}}{a\sigma^2 s} < 0 \Leftrightarrow \bar{e} < \sqrt{U^m}$. Taking each side to the square and developing yields

$$2a\sigma^2 W_{-1} - 2a\sigma^2 s\bar{e} + 2a^2\sigma^4 s^2 > 0, \ \ i.e. \ W_{-1} > s\bar{e} - a\sigma^2 s^2$$

This condition is always satisfied since the c_1, u_0 strategy requires $W_{-1} > \omega^p = s\bar{e} - \frac{1}{2}a\sigma^2 s^2$ to be feasible. Hence $\Delta_1 < \frac{1}{2}\Delta_0$. ■

References

ACHARYA, V., T. COOLEY, M. RICHARDSON, AND I. WALTER (2010): *Regulating Wall Street: The Dodd-Frank Act and the New Architecture of Global Finance",*. John Wiley and Sons.

ATTARI, M., AND A. MELLO (2006): "Financially Constrained Arbitrage in Illiquid Markets," *Journal of Economic Dynamics and Control*, 30, 2793–2822.

BASAK, S. (1997): "Consumption Choice and Asset Pricing with Non-Price Taking Agents," *Economic Theory*, 10, 437–462.

BRUNNERMEIER, M., AND L. PEDERSEN (2009): "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22(6), 2201–2238.

BULOW, J. (1982): "Durable Goods Monopolists," *Journal of Political Economy*, 90, 314–332.

CHAN, L., AND J. LAKONISHOK (1993): "Institutional trades and intraday stock price behavior," *Journal of Financial Economics*, 33(2), 173–199.

——— (1995): "The Behavior of Stock Prices Around Institutional Trades," *Journal of Finance*, 50(4), 1147–1174.

——— (1997): "Institutional Equity Trading Costs: NYSE Versus Nasdaq," *Journal of Finance*, 52(2), 713–735.

CHEN, Z., W. STANZL, AND M. WATANABE (2002): "Price Impact Costs and the Limits of Arbitrage," Yale University Working Paper.

COASE, R. (1972): "Durability and Monopoly," *Journal of Law and Economics*, 15, 143–149.

COMERTON-FORDE, C., T. HENDERSHOTT, C. JONES, P. MOULTON, AND M. SEASHOLES (2010): "Time Variation in Liquidity: The Role of Market-Maker Inventories and Revenues," *Journal of Finance*, 65(1), 295–331.

FARDEAU, V. (2011): "Dynamic Strategic Arbitrage," LSE Working Paper.

GABAIX, X., A. KRISHNAMURTHY, AND O. VIGNERON (2007): "Limits of Arbitrage: Theory and Evidence from the Mortgage-Backed Securities Market," *Journal of Finance*, 62(2), 557–595.

GARLEANU, N., AND L. H. PEDERSEN (2011): "Margin-Based Asset Pricing and the Law of One Price," *Review of Financial Studies*, 24, 1980–2022.

GLOSTEN, L. (1989): "Insider Trading, Liquidity, and the Role of the Monopolist Specialist," *Journal of Business*, 62(2), 211–235.

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

——— (2010): "Limits of Arbitrage: The State of the Theory," *Annual Review of Financial Economics*, 2, 251–275.

Grossman, S., and M. Miller (1988): "Liquidity and Market Structure," *Journal of Finance*, 38, 617–633.

Keeley, M. (1990): "Deposit Insurance, Risk, and Market Power in Banking," *American Economic Review*, 80(5), 1183–1200.

Kihlstrom, R. (2000): "Monopoly Power in Dynamic Securities Markets," Wharton Working Paper.

Liu, X., and A. Mello (2011): "The Fragile Capital Structure of Hedge Funds and the Limits to Arbitrage," *Journal of Financial Economics*, 102(3), 491–506.

Oehmke, M. (2010a): "Gradual Arbitrage," Columbia University, Working Paper.

——— (2010b): "Liquidating Illiquid Collateral," Columbia Business School Working Paper.

Pérold, A. (1999): *Long-Term Capital Management, L.P.* Harvard Business Case.

Perotti, E., and K. Spier (1993): "Capital Structure as a Bargaining Tool: The Role of Leverage in Contract Renegotiation," *American Economic Review*, 83(5), 1131–1141.

Pritsker, M. (2009): "Large Investors: Implications for Equilibrium Asset Returns, Shock Absorption, and Liquidity," Finance and Economics Discussion Series, Federal Reserve Board, Washington, D.C.

Rostek, M., and M. Weretka (2011): "Dynamic Thin Markets," University of Wisconsin-Madison working paper.

Seppi, D. (1997): "Liquidity Provision with Limit Orders and a Strategic Specialist," *Review of Financial Studies*, 10(1), 103–150.

Vayanos, D. (1999): "Strategic Trading and Welfare in a Dynamic Market," *Review of Economic Studies*, 66, 219–254.