Arbitrage with Financial Constraints and Market Power

Vincent Fardeau*
Higher School of Economics

December 2020

Abstract

I study how financial constraints affect liquidity provision and welfare under different structures of the arbitrage industry. In competitive markets, financial constraints may impair arbitrageurs' ability to provide liquidity, thereby reducing other investors' welfare. Instead, in imperfectly competitive markets, I characterize situations in which imposing constraints on arbitrageurs leads to a Pareto improvement relative to a no-constraint case. Further, unlike the competitive case, a drop in arbitrage capital does not always lead to a reduction in market liquidity. A subtle interaction between financial constraints and arbitrageurs' market power generates these Pareto improvement and novel comparative statics.

JEL classification: G12, D41, D60

Keywords: Limits of arbitrage, liquidity, strategic arbitrage, market structure, price impact, margin requirements, VaR

*International College of Economics and Finance (ICEF), National Research University Higher School of Economics, Pokrovsky Boulevard 11, Moscow, Russia. Email: vfardeau@hse.ru. This article was prepared within the framework of the HSE University Basic Research Program and funded by the Russian Academic Excellence Project '5-100'. This paper was previously circulated "Market Structure and Limits to Arbitrage". I benefited from comments from conference and seminar participants at 2020 Econometric Society World Congress, the 2020 Paris December Finance Meetings, EIEF, Frankfurt School of Finance and Management, HEC Lausanne, Higher School of Economics, University of Naples, University of Vienna, University of Zurich. I am grateful to Sylvain Carré, Denis Gromb, Michel Habib, Ron Kaniel, Xuewen Liu, Dmitri Livdan, Anna Obizhaeva, Christine Parlour, Norman Schürhoff, Günter Strobl, and Dimitri Vayanos for helpful suggestions.

1 Introduction

While the LTCM and 2007-2009 crises highlighted the interactions between funding and market liquidity, our understanding of these interactions remains limited to competitive settings.[1] The reality of financial markets, however, is often closer to imperfect competition. For instance, LTCM, which initially motivated the literature, was nicknamed the "central bank of volatility", due to its dominant position in derivatives markets. LTCM is not an isolated case: many other hedge funds or banks (e.g. Amaranth, the "London whale", etc.) have been under the spotlights for becoming the dominant traders in some markets; more generally, there has been a noted increase in market power in financial markets.[2] Consistent with this evidence, transaction-level data shows that large intermediaries recognize their price impact and use optimal execution techniques to rebalance portfolios.[3] Further, financial constraints due to regulations, internal risk management, or margins imposed by financiers (e.g. brokers, repo market participants, etc.) are likely to limit the funding liquidity of these large traders, and have probably become tighter after 2007-2009.[4]

In this paper, I study the effects of imposing financial constraints on imperfectly competitive arbitrageurs. I show that these constraints affect market liquidity and social welfare in different and sometimes opposite ways when arbitrageurs have market power. In competitive markets, binding financial constraints (i.e. a decrease in funding liquidity) impair arbitrageurs' ability to exploit profitable trading opportunities, thereby reducing market liquidity and hurting the investors who are on the other side of their trades. Instead, imposing financial constraints on imperfectly competitive arbitrageurs may in some cases *improve* both market liquidity and social welfare.

Arbitrageurs with market power may benefit from the constraints because they face a

[1]For models of financially constrained arbitrage in competitive markets, see, among others, Shleifer and Vishny (1997), Gromb and Vayanos (2002, 2010, 2018), Kondor (2008), Brunnermeier and Pedersen (2009). Attari and Mello (2006) study numerically how a monopolistic arbitrageur subject to financial constraints trades, but their setting does not lend itself to welfare analysis.

[2]See for instance, De Loecker, Eeckhout, and Unger (2019), for general evidence about market power. Many financial markets are dominated by a few large players. For instance, five banks represent 90% of the notional amount of derivative contracts (OCC, 2018). Wallen (2020b) finds evidence of dealer market power in the FX market. Wallen (2020a) shows that both shocks to capital and shocks to the *concentration* of capital matter for asset prices.

[3]See, e.g., Gabaix et al., (2006), Ben-David et al. (2015), Chan and Lakonishok (1995), and van Kervel and Menkveld (2019). Given the mixed evidence about the performance of large traders, a large fraction of this rebalancing must stem from reasons other than superior information, e.g. risk-sharing (Vayanos, 1999).

[4]Basel III has tightened capital requirements for banks, and introduced a leverage ratio and several liquidity ratios, depleting some of the capital devoted to chasing arbitrage opportunities. Further, these new or tighter constraints on banks have been passed through to hedge funds and other market players through a reduction in funding (Boyarchenko et al, 2018).

commitment problem, as durable good producers: once they have provided some liquidity, they cannot refrain from providing more, and reduce the arbitrage opportunity further. However, a binding constraint in the future limits their ability to provide further liquidity, mitigating their commitment problem. Surprisingly, the investors on the other side of their trades may also benefit, for two reasons. First, the only way for the arbitrageur to make the constraint binding in the future is to pledge more capital to the arbitrage early on, which speeds up risk-sharing. Second, the constraint is an imperfect commitment device: it does not prevent the arbitrageur entirely from retrading, as early capital gains generate additional collateral for later trading rounds.

Model. I introduce imperfect (Cournot) competition among arbitrageurs in an otherwise standard model of financially constrained arbitrage. The model has two types of investors: *hedgers* and *arbitrageurs*, which trade for two rounds and then consume. There is a risky asset traded in two segmented markets (A and B) and a risk-free asset. In each segmented market, the risk-averse, competitive hedgers receive endowment shocks correlated with the asset payoff. For simplicity, these shocks are symmetric: hedgers in market A are overexposed to the risky asset and would like to hedge by selling, and vice versa in market B. Market segmentation prevents welfare-improving trades between the two groups. As a result, the risky asset trades at a discount in A and at a premium in B.

While hedgers are restricted to trade in their respective market, arbitrageurs can trade across markets. As the risky asset gives claims to the same cash-flows in both markets, prices converge in the final period. Thus, the *spread* between prices in markets A and B creates a textbook arbitrage opportunity for arbitrageurs, who face a relative value trade with a fixed convergence date. Arbitrageurs, however, must separately collateralize each leg of the arbitrage. Arbitrageurs' wealth serves as collateral, for both long and short positions, and must remain sufficiently large over time to absorb adverse price movements.[5] This funding constraint is akin to real VaR constraints imposed by regulators or used by risk managers or financiers (e.g. prime brokers) to set margins. The stressed VaR for capital requirements in Basel 2.5 and the margin requirements for non-centrally cleared derivatives are some examples of newly-imposed constraints based on VaR.[6]

[5] I use wealth and capital as synonyms. To avoid dealing with default in equilibrium, I assume that the worst change in fundamental is bounded above and below, and that arbitrageurs must fully collateralize all potential losses. This is akin to a 100% VaR constraint. Gromb and Vayanos (2002) follow a similar strategy, while Brunnermeier and Pedersen consider an α% VaR, but do not study welfare.

[6] For the former, see https://www.bis.org/publ/bcbs148.pdf, section 4, p. 11. For the latter, see https://www.bis.org/bcbs/publ/d475.pdf, paragraph 3.1, p. 12. The exercise of this paper thus has a positive flavour: given realistic imperfection in competition, I determine the welfare and price effects of widely-used financial constraints. Note that Expected Shortfall generates exactly the same type of con-

Results. Suppose first that arbitrageurs face financial constraints but are competitive. In equilibrium, arbitrageurs eliminate the arbitrage opportunity when capital is abundant. When capital is scarce, arbitrageurs hit their funding constraint, which prevents them from building large enough positions to eliminate the spread. As time passes, arbitrageurs earn capital gains, which increase their wealth and relax the financial constraint, so that the spread decreases. In this competitive setting, imposing financial constraints has either no effect on the equilibrium (when arbitrageurs' capital is large), or prevents arbitrageurs from intermediating trades between markets, which reduces hedgers' welfare.

Imposing financial constraints on arbitrageurs with market power can have sharply different effects. I show that when capital is intermediate and the risk to benefit ratio of the trade is sufficiently high, imposing constraints on arbitrageurs leads to a Pareto improvement. In all other situations, imposing constraints has either no effect, because they never bind, or limits the liquidity provision by arbitrageurs, which reduces the welfare of at least one type of investors.[7]

The Pareto improvement follows from a subtle interaction between market power and financial constraints. Consider first a monopoly without constraints. The arbitrageur is akin to a durable good producer; he faces a commitment problem: having provided some market liquidity, the arbitrageur faces a residual demand for liquidity and cannot refrain from providing further liquidity, thereby reducing the spread further. Hedgers anticipate this behaviour, which erodes the arbitrageur's market power ex-ante.[8]

Because of these Coasian dynamics, the arbitrageur would be better off if he could commit to a trading strategy, i.e. decide ex-ante how much liquidity to provide over time and stick to it. In this case, he would trade only once at the beginning. Doing so, he would earn static monopoly profits, as is well-known in IO; this hurts hedgers relative to the no-commitment case.

With a financial constraint, the arbitrageur chooses trades sequentially. However, in states or at dates where the constraint binds, the choice is purely mechanical: max out the

straint. Brunnermeier and Pedersen, 2009, Appendix A, describe in further detail the mapping of this kind of constraint to various institutional settings.

[7]In the model, the numerator of the risk benefit ratio is the worst case scenario for the fundamental. In practice, it would correspond to a quantile of the return distribution, e.g. the 99% quantile.

[8]As hedgers have reduced their positions at time 0, they are less exposed to the endowment shock at time 1. To this extent, receiving liquidity / sharing risk/ buying insurance (all synonyms in this model) by trading the risky, imperfectly liquid asset against cash (the liquid asset) is akin to buying a durable good. However, in contrast to the literature on the Coase conjecture, the horizon is finite. In IO, it is typically assumed that the good is infinitely durable. Classic papers on the Coase conjecture include Stockey (1981), Bulow, (1982), Gul, Sonnenschein, and Wilson (1986).

constraint. To this extent, a binding constraint in the future works as a commitment device. Yet, this commitment is imperfect. Financial constraints may limit retrading, but do not eliminate it: they merely prevent the arbitrageur from reaching his preferred position. The reason is simple: when providing liquidity, a monopoly *always* earns capital gains, which generate additional collateral and allows the arbitrageur to retrade next period.[9] Besides, since the arbitrageur internalizes his price impact, he has more leeway than price-takers in choosing strategically ex-ante *whether* or at least *when* to make constraints bind. For this reason, financial constraints may bind at some date and not at other.

The welfare and liquidity improvement occur only in cases where the constraint is slack at time 0 and binding at time 1. The binding constraint at time 1 mitigates the arbitrageur's problem ex-ante; the slack constraint at time 0 allows the arbitrageur to reap the benefit of it. Indeed, the prospect of a binding constraint induces hedgers to shift some of their liquidity demand earlier. The slack constraint allows the arbitrageur to exploit this extra demand. The arbitrageur's welfare increases relative to the no commitment, no constraint case, but does not reach the perfect commitment level, because the retrading maintains some Coasian dynamics. Thus, the arbitrageur is not able to charge static monopoly prices.

Hedgers are also better off when the constraint binds at time 1 but not at time 0 than without constraint. In this case, the constraint induces faster risk-sharing but does not eliminate retrading. Instead, perfect commitment speeds up risk-sharing but eliminates retrading.[10] Receiving liquidity early matters to hedgers: the asset is conditionally riskier at time 0, because dividends news accrue every trading round.[11]

Risk-sharing speeds up endogenously: to make the financial constraint binding next round, the arbitrageur *must* trade more aggressively early on than without constraints. By taking larger positions, the arbitrageur pledges more capital to the trade now and makes it more likely to have a binding constraint next period. But doing so, he provides more

[9] By contrast, competitive arbitrageurs earn capital gains only when their constraints bind. With financial constraints, it is essential to generate more collateral for later rounds that the rents the monopolistic arbitrageur extracts in the first round cannot be diverted or repledged elsewhere. Without such ability, it is ex-post *optimal* for the arbitrageur to keep pledging capital to the arbitrage, for lack of a better investment opportunity. Indeed, the only alternative investment is the risk-free asset, which offers lower returns than the arbitrage. Note that in Dow, Han, and Sangiorgi (2019), a financially constrained arbitrageur also prefers to stick to his existing position, not for a lack of better opportunities, but because it is costly to exit.

[10] Under perfect commitment, hedgers anticipate that the arbitrageur will not retrade and shift their demand to the initial trading round. The arbitrageur provides more liquidity initially to exploit this extra demand.

[11] Uncertainty about prices at time 0 gives hedgers a preference to trade early and plays the same role as a discount factor. It also creates a role for arbitrageurs in both periods. Without it, hedgers' demand would not remain downward-sloping at time 0, so that there would be no demand for liquidity. Hedgers would face a temporary risk-free asset and would flatten out demand, becoming arbitrageurs themselves.

liquidity and reduces the spread. The financial constraint does not eliminate retrading: in fact, the larger gains earned ex-ante increase the maximum position the arbitrageur can afford given the constraint.[12] As a result, the arbitrageur accumulates larger positions than in the no commitment, no constraint case. The combination of faster risk-sharing (more liquidity early) and higher total amount of liquidity is necessary for an improvement in hedgers' welfare with a monopoly, but not with an oligopoly. In the latter, more early liquidity and enough retrading suffice.[13]

The conditions for the Pareto improvement are that capital is intermediate and the risk to benefit ratio of the trade is sufficiently high. Intuitively, if capital is very low, the arbitrageur cannot take larger positions early on without violating the constraint. But if capital is very large, he will never be financially constrained ex-post. A high risk to benefit ratio implies that positions are sufficiently capital-intensive relative to potential profits. Hence, even taking into account the capital gains from providing early liquidity, the arbitrageur's constraint does bind next period.

The risk benefit ratio also determines the comparative statics of spreads with respect to initial capital. In markets with a low risk benefit ratio, comparative statics are qualitatively similar to the competitive case: less capital may lead to higher spreads. In markets with a high risk to benefit ratio, however, a drop in capital first *reduces* the spread and then increases it. As capital drops, hedgers anticipate that the arbitrageur will trade more aggressively early on and face a binding constraint later, and eventually provide more liquidity, which reduces spreads. As capital drops further, however, the arbitrageur will no longer be able to trade aggressively early on, and spreads will increase, as in the competitive case. To the best of my knowledge, the empirical literature has not used the risk benefit ratio as a conditioning variable, and has not tested this unique prediction of the model.[14][15] Note that positions and VaR feature similar comparative statics: when the risk benefit ratio is low, a drop in capital leads to smaller positions and VaR, as in the competitive case. When it is high, the arbitrageur's positions and VaR first increase and then decrease.

My results should be of interest to policy-makers regulating markets with large traders.

[12] Without constraint, this position would be even larger.

[13] Competition among arbitrageurs erodes capital gains, so arbitrageurs' financial constraints relax less, and arbitrageurs may not provide more liquidity in total. Still, hedgers may be better off. Hence, what is essential is that the arbitrageur provides more liquidity early on and retrades, not that he provides more liquidity in total.

[14] The welfare result also holds in the oligopolistic case, but spreads may not be lower at all dates.

[15] Under the US Market Risk Capital Rule, large banks must report daily VaR and P&L (see, e.g. Falato et al. (2019). This newly available regulatory data could be used to test the empirical predictions of the paper and for policy analysis.

The debate around the regulations set in place after the 2007-2009 crisis has focused on the possibly negative effects on liquidity provision of imposing new capital and margin requirements. I show, however, that in some situations there are welfare *gains* from imposing financial constraints on arbitrageurs with market power.[16] My analysis shows that this improvement occurs when margins are price-based and forward-looking. Other types of constraints such as fixed margins or mere position limits do not generally have such effects. Indeed, unlike fixed margins, price-based, forward-looking margins reflect the fact that arbitrageurs make profit, which provides "cushion" from the point of view of financiers. This allows arbitrageurs to tackle the arbitrage more aggressively early on. As discussed in the literature review, position limits delay risk-sharing (see Table 1).[17]

My mechanism provides a rationale for (in specific cases) imposing financial constraints on imperfectly competitive arbitrageurs, not in favor of market power itself. A standard argument in favor of market power is that pecuniary externalities arise from the fact that agents do not internalize their price effects. Thus, *given constraints*, it may be beneficial to give traders market power to curb externalities.[18] Here the competitive equilibrium is constrained efficient, so this mechanism does not arise. The mechanism of this paper is different, and to the best of my knowledge novel: *given market power*, which is a feature of many financial markets, imposing financial constraints may improve liquidity and welfare, because of the way constraints interact with arbitrageurs' market power.

Two additional policy implications arise: first, the model may explain why large financial institutions fund themselves more cheaply. It is often argued that this funding cost advantage results from an implicit government put.[19] In my model, large arbitrageurs are always less severely constrained than competitive arbitrageurs, because their profits lead to a larger

[16]One aspect that I abstract from in this paper is that binding constraints may lead to firesales and even default in equilibrium. Thus, the benefits of being constrained should be weighted against the costs generated by firesales and default. However, to the best of my knowledge, such benefits have not been highlighted before.

[17]The model with fixed margins can only be analysed numerically, so this claim is based on numerical analysis. The model with position limits (capacity constraints) is analytically tractable. Fixed margins and position limits are imposed by CCPs or exchanges in some derivatives markets. Interestingly, for non-centrally cleared derivatives, regulators expect sophisticated investors to use VaR models, but allow unsophisticated ones to use a schedule of fixed margins, see https://www.bis.org/bcbs/publ/d475.pdf, Appendix A.

[18]Eisenbach and Whelan (2019) show that this argument may not always go through when imperfectly competitive traders differ in their trading needs. On a different note, 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.

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

"pledgeable income".[20] Second, the model highlights an unintended consequence of capital requirements regulation. They may increase traders' market power by limiting their ability to trade in some states.[21]

Related literature. To the best of my knowledge, this paper is the first to solve the dynamic problem of imperfectly competitive traders under realistic financial constraints when all investors are rational. As a result, the paper contributes to three strands of the literature. My first contribution is to extend the literature on the limits of arbitrage, which is cast into the competitive framework, to imperfect competition. I build on Gromb and Vayanos (2002, 2010) and Brunnermeier and Pedersen (2009). Like the former, I carry out a welfare analysis.[22]

Second, I contribute to the literature on imperfect competition in financial markets by analyzing the interaction between market power and financial constraints. Several papers in this active literature model all investors as rational and emphasize the parallel with the durable goods problem studied by Coase (1972), but not study the effects of financial constraints.[23] Instead, Attari and Mello (2006) do study the effects of financial constraints on a monopolistic arbitrageur, but do not model all agents as rational, assuming that the arbitrageur faces exogenous demand curves, i.e. that hedgers do not optimize. This assumption rules out a welfare analysis and eliminates the Coasian dynamics, which are central to my results.

Finally, the paper contributes more broadly to the literature on durable goods monopoly. In IO, commitment devices lead to a reduction in consumer surplus. However, here financial constraints may lead to a Pareto improvement relative to a no-constraint case, even though they may provide some form of commitment. In this regard, it is instructive to contrast my results to McAfee and Wiseman (2008), where the monopolist pays a small cost to set up ex-ante a maximum capacity per period (in the context of financial markets, capacity constraints are similar to position limits, which are used in some derivatives markets). With such capacity constraints, I find that hedgers are worse off than in all other cases. The

[20]Related ideas have been 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.

[21]In the absence of constraints, the mere continuous presence on the market of a trader or his tendency to break up trades to execute block orders are factors that erode her market power. Indeed, rational investors can anticipate better prices for liquidity in the future and shift their demand.

[22]Other recent papers on financially constrained arbitrage include Kondor and Vayanos (2018) and Dávila and Korinek (2017), among others.

[23]See, e.g., Basak (1997), Vayanos (1999), Kihlstrom (2000), Pritsker (2009), DeMarzo and Urosevic (2007), Marinovic and Varas (2019).

reason is that capacity constraints do not eliminate retrading but delay risk sharing. Indeed, to avoid unused capacity at time 1, the arbitrageur chooses a "small" capacity and trades at maximum capacity over time. This decreases liquidity at time 0 relative to the no-commitment case. Further, if the arbitrageur can set up a time-dependent capacity, he will simply choose one inducing the perfect commitment outcome. By contrast, VaR-based constraints induce both faster risk sharing and sufficient retrading. This is possible because their tighthness depends endogenously on the trading process. It is this endogeneity that overturns their welfare effects as commitment devices.[24]

Commitment/Constraint	Risk sharing	Retrading	Hedgers	Arbitrageurs
Perfect commitment	Earlier	No	Worse off	Better off
Capacity constraints (position limits) MacAfee and Wiseman (2008)	Delayed	Yes	Worse off	(Quasi) Indifferent
(price-based) financial constraints	Earlier	Yes	Better off (sometimes)	Better off (sometimes)

Table 1: Welare effects relative to the no constraint, no commitment equilibrium (u_0, u_1)

2 Model

I consider a standard model of financially constrained arbitrage, where arbitrageurs exploit price differences between two identical assets over time, while facing realistic capital constraints.

Assets and timeline. The model has three periods, indexed by $t = 0, 1, 2$. Financial markets are 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. The risky assets trade in segmented markets at price p_t^k, $k \in \{A, B\}$. They are both in zero net supply and pay the same dividend D_2 at time 2, with $D_2 = D + \epsilon_1 + \epsilon_2$. The dividend news ϵ_t are *iid* random variables with a symmetric bounded support $[-\bar{e}, \bar{e}]$, a mean of 0 and variance σ^2. The news ϵ_t is revealed to all investors at time t before trading. There are two types of investors: hedgers and arbitrageurs.

[24] Because of this endogeneity, there is also a feedback loop between financial constraints and trading strategies, which may lead to multiple equilibria. Equilibria may coexist, because arbitrageurs choose quantities, but not hedgers' expectations. However, hedgers' expectations about whether constraints bind or not in the future determine market depth today, and therefore the arbitrageur's incentives to trade in a way that makes the constraint binding or not.

Hedgers. In each market, there is a continuum mass one of risk-averse competitive hedgers with mean-variance preferences: $U\left(w_2^k\right) = \mathbb{E}\left(w_2^k\right) - \frac{a}{2}\mathbb{V}\left(w_2^k\right)$.[25] Every period, hedgers receive endowment shocks $s^k\epsilon_t$ that are correlated with the dividend of the risky asset, and will therefore affect their demand for the risky asset. At time t, hedgers' weath at is

$$w_t^k = w_{t-1}^k + s^k\epsilon_t + Y_{t-1}^k(p_t^k - p_{t-1}^k), \tag{1}$$

i.e. hedgers' wealth changes because of capital gains on the risky asset (third term) and the endowment shocks. For simplicity, the magnitude of the shock, s^k, is deterministic, constant over time, and symmetric across markets.[26] That is, at time $t = 1, 2$, hedgers in market A receive a shock $s^A\epsilon_t = s\epsilon_t$, while hedgers in market B receive opposite shocks, $s^B\epsilon_t = -s\epsilon_t$. As a result, A-investors have a low valuation for the risky asset, and B-investors a high valuation. Market segmentation prevents hedgers from sharing risk across markets, although they could perfectly insure each other. Therefore, assets A and B may trade at different prices in their respective market, even though their cash-flows are identical. Since the endowment shock will shift hedgers' demand up or down by $|s|$, it is convenient to think of s as the net supply (in absolute value) in each market (see Section 3.2 for additional details).

Arbitrageur(s). There are $n \geq 1$ identical arbitrageurs, who can participate in all markets, but face financial constraints, described below. Arbitrageurs have no initial holdings of the risky asset but own initial wealth (capital) $W_0^i = \frac{W_0}{n}$, where W_0 is the total capital in the arbitrage industry. I focus on the comparison between the monopolistic ($n = 1$) and competitive cases in the text and consider the more general oligopolistic case in the Internet Appendix.

Arbitrageurs also strictly increasing utility over final wealth $u\left(W_2\right)$. Because they have access to all markets, arbitrageur's final wealth is $W_2^i = \sum_{k\in\{A,B\}} X_1^{i,k} D_2 + B_1^i$, where $X_t^{i,k} =$

[25] Mean-variance preferences are also used for tractability reasons in, e.g., Banerjee and Green (2015). With mean-variance preferences, I consider time-consistent trading strategies.

[26] The assumption of constant shock magnitude s can be relaxed at the cost of increased complexity but is not essential for the main results. Similarly, Coasian dynamics would remain in the presence of stochastic shocks. Relaxing the assumption of deterministic shocks, however, would require a separate analysis, as additional effects would arise. As shown in Gromb and Vayanos (2002), stochastic shocks lead to pecuniary externalities, so that the competitive equilibrium is not constrained efficient anymore. Market power would lead arbitrageurs to internalize some of the pecuniary externalities, opening an interaction that is not present with deterministic shocks, where the competitive equilibrium is constrained efficient. Eisenbach and Phelan (2019) start from a constrained inefficient equilibrium and study the effect of giving liquidity suppliers market power, albeit in a static setting. I discuss the bearing of the assumption of deterministic shocks for the results further in Section 7.

$X_{t-1}^{i,k} + x_t^{i,k}$ denotes the end-of-period position at time t in asset k of arbitrageur i, $x_t^{i,k}$ the corresponding trade, and $B_t^i = B_{t-1}^i - \sum_{k\in\{A,B\}} x_t^{i,k} p_t^k$, the arbitrageur's risk-free asset holdings at the end of period t. Wealth consists initially only of cash, $W_0^i = B_{-1}^i$.

When n is finite, arbitrageurs internalize their own price impact in both A and B markets. Hedgers' inverted demand and market clearing define a price schedule, derived below, that links the arbitrageur's trade to the equilibrium price in each market. An arbitrageur chooses trades $x_t^{i,k}$ given these price schedules and other arbitrageurs' trades $x_t^{-i,k}$ (Cournot competition).

Financial constraints. Arbitrageurs need 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 (2009). Arbitrageurs must fully collateralize their positions in each market. The maximum possible loss on the position over the next period in market $k \in \{A, B\}$ is $\max_{p_{t+1}^k} X_t^k \left(p_t^k - p_{t+1}^k\right)$. The arbitrageur's wealth must cover the total maximum loss on each account:

$$W_t \geq \sum_{k=A,B} \max_{p_{t+1}^k} X_t^k \left(p_t^k - p_{t+1}^k\right) \tag{2}$$

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.[27]

Given the symmetry assumptions, and in line with the literature, it is natural to focus on equilibria in which the arbitrageur holds opposite positions in both assets, i.e. $X_t^{i,A} = -X_t^{i,B} = X_t^i$. Given that arbitrageurs start with no endowment in the risky assets, this implies that trades are symmetric $x_t^{i,A} = -x_t^{i,B} = x_t^i$, for $t = 0, 1$. Thus, we can rewrite the

[27] 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."

arbitrageur's budget constraint as follows:

$$W_2^i = W_0^i + \sum_{t=0,1} x_t^i \Delta_t, \text{ with } \Delta_t = p_t^B - p_t^A \tag{3}$$

The equation shows that by setting up opposite position in each leg of the arbitrage, arbitrageurs eliminate all fundamental risk and derive all their profits from exploiting the spread Δ between the prices of the two assets.[28] Symmetry assumptions also lead to opposite risk premia on assets A and B, which simplifies the financial constraint. Let D_t denote the conditional expected value of the asset at time t, i.e. $D_t = \mathbb{E}_t(D_2) = D_{t-1} + \epsilon_t$. Then the risk premium in market A is $D_t - p_t^A = \frac{\Delta_t}{2}$ and opposite in B. Thus, we can write (minus) the price change in each market as $p_t^k - p_{t+1}^k = \frac{\Delta_{t+1} - \Delta_t}{2} - \epsilon_{t+1}$. Then we can rewrite the financial constraint (2) as follows:

$$\begin{aligned} W_t^i &\geq \sum_{k=A,B} \max_{p_{t+1}^k} X_t^{i,k} \left(p_t^k - p_{t+1}^k\right) \\ &\geq \max_{\epsilon_{t+1}} X_t^i \left(\frac{\Delta_{t+1} - \Delta_t}{2} - \epsilon_{t+1}\right) + \max_{\epsilon_{t+1}} -X_t^i \left(-\frac{\Delta_{t+1} - \Delta_t}{2} - \epsilon_{t+1}\right) \\ &\geq 2X_t^i \left(\frac{\Delta_{t+1} - \Delta_t}{2}\right) + \max_{\epsilon_{t+1}} X_t^i \left(-\epsilon_{t+1}\right) + \max_{\epsilon_{t+1}} -X_t^i \left(-\epsilon_{t+1}\right) \\ &\geq 2|X_t^i|\bar{e} - X_t^i \left(\Delta_t - \Delta_{t+1}\right) \end{aligned} \tag{4}$$

The last step follows from the symmetric support of the distribution. Note that the constraint may bind upwards or downwards. An upward-binding constraint generates an upper bound on how much the arbitrageur can hold, e.g. for a long position, $X_t^i \leq \frac{W_t^i}{2\bar{e} - (\Delta_t - \Delta_{t+1})}$. Instead, a downward-binding constraint generates a lower bound on the arbitrageur's position, e.g. for short positions, the arbitrageur needs to hold at least $X_t^i \geq \frac{W_t^i}{-(2\bar{e} + \Delta_t - \Delta_{t+1})}$.

It is convenient to rewrite constraint (4) as

$$f_t^+(X_t^i)\mathbb{1}_{X_t^i \geq 0} + f_t^-(X_t^i)\mathbb{1}_{X_t^i < 0} \geq 0 \tag{5}$$

where $f_t^+(X_t^i) = W_t^i - 2X_t^i\bar{e} + X_t^i \left(\Delta_t - \Delta_{t+1}\right)$ and $f_t^-(X_t^i) = W_t^i + 2X_t^i\bar{e} + X_t^i \left(\Delta_t - \Delta_{t+1}\right)$. The spread change $\Delta_t - \Delta_{t+1}$ will itself depend on the arbitrageurs' positions X_t^i, so the constraint is non-linear in X_t^i. I denote $\bar{X}_t^i$ the largest long position satisfying arbitrageur's i constraint at time t given other arbitrageurs' positions, i.e. $f_t^+(\bar{X}_t^i) = 0$. Since arbitrageurs have no endowments, $\bar{X}_0^i = \bar{x}_0^i$. Because wealth depends on initial positions, and $X_t^i = X_{t-1}^i + x_t^i$, $\bar{X}_1$

[28] I assume that $B_{-1} = W_0$, i.e. the initial wealth is just the endowment in cash.

is a function of the previous position x_0.

VaR / margins. 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[29], but the constraint is motivated by real-world regulations and risk management practices of financiers.[30]

The VaR constraint can also be written as a margin constraint. Suppose that the arbitrageur holds a long position, $X_t^i \geq 0$. We can rewrite the right-hand side of inequality (4) as $2m_t^+ X_t^i$, where $m_t^+ = \bar{e} - \frac{1}{2}(\Delta_t - \Delta_{t+1})$ denotes the margin required on the position. Margins increase with fundamental risk $\bar{e}$ and decrease with the expected price change $\mathbb{E}_t(p_{t+1}^k) - p_t^k = \frac{\Delta_t - \Delta_{t+1}}{2}$. A riskier asset leads to a larger potential loss on the position, which induces financiers to ask for more collateral. Instead, a higher expected return increases the "pledgeable income", reducing the margin requirement. Financiers reduce margins when they expect the spread to decrease, i.e. $\Delta_{t+1} \leq \Delta_t$. To this extent, margins play a stabilizing role for asset prices. [31]

Terminology. In the literature, market liquidity refers to the price spread $\Delta_t \equiv p_t^B - p_t^A$, which resembles a bid-ask spread. However, market liquidity is a multifaceted concept. One measure of liquidity in the model is market depth, given by the slope of hedgers' inverted demand. Thus, to avoid ambiguity, I use spread instead of market liquidity for Δ_t. Further, I use the expression "provide liquidity" as synonym to "provide insurance/risk-sharing". Funding liquidity relates to the tightness of the arbitrageur's financial constraint.

In sum, the model is close to Gromb and Vayanos (2002), but imposes three simplifications: deterministic and constant hedging needs, two trading rounds, and mean-variance preferences. I discuss further the bearing of these assumptions in Section 7, after presenting the main result.

[29] In particular, there is no need to compute the welfare of financiers on the other side of the constraint.

[30] For instance, Brunnermeier and Pedersen (2009), Appendix A, provide additional institutional details to motivate the analysis of this type of constraint. Bruche and Kuong (2019) obtain a similar constraint in a static setting when deriving optimal contract between financiers and arbitrageurs.

[31] Brunnermeier and Pedersen (2009) obtain a similar constraint in their benchmark case with informed financiers. They also consider a situation in wihch financiers are assumed to be uninformed. In this case, uncertainty about whether the mispricing will decrease or not in the future can lead to procyclical, destabilizing margins. 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 presence of uninformed financiers. 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 the assumptions of this paper, there can be a loss spiral, but no margin spiral.

3 Benchmarks

To highlight the novel interaction betwen market power and financial constraints of the model, I first review the effect of each ingredient separately.

3.1 Financial constraints without market power

In a competitive economy, arbitrageurs are aggregated into a representative competitive arbitrageur endowed with aggregate capital W_0. A competitive equilibrium is a collection of prices and trades such that (i) hedgers' holdings are optimal given prices, (ii) the arbitrageur's holdings are optimal given prices and financial constraints, and (iii) markets clear. I denote X_t the position of the representative arbitrageur; a * denotes the competitive outcome in the paper.

Proposition 1 (Gromb and Vayanos, 2002) *There exists a unique competitive equilibrium:*

- *If $W_0 \geq \omega^* \equiv 2s\bar{e}$, the financial constraint never binds, the representative arbitrageur absorbs the supply 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 $0 \leq W_0 < \omega^*$, 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 representative arbitrageur' positions in asset A, $\bar{x}_0$ and $\bar{X}_1$, are the largest (long) positions allowed by the financial constraints, and satisfy $f^+(\bar{X}_t) = 0$.*

The equilibrium links the spread Δ to arbitrageurs' initial capital W_0, and takes a simple form: if arbitrageurs' capital is large enough, then arbitrageurs eliminate the arbitrage opportunity; if instead arbitrageurs start with lower 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 supply s or in the fundamental risk (increase in $\bar{e}$) tightens proportionately the financial constraint. This is because the worst possible loss increases, so arbitrageurs need to post more collateral.

A drop in arbitrage capital has the following consequences:

Corollary 1 (Comparative Statics in the Competitive Benchmark) *Suppose that competitive arbitrageurs are constrained, i.e. $0 \leq W_0 < \omega^*$. A decrease in capital increases the* *spread, even more so if capital was initially low:*

$$\frac{\partial \Delta_t^*}{\partial W_0} < 0, \quad \frac{\partial^2 \Delta_t^*}{\partial W_0^2} < 0$$

Instead, if the constraint is slack, equilibrium spreads and positions are independent of capital.

Proof. The comparative statics follow from Proposition 1 and Corollary 4 (Appendix A.1). ■

3.2 Market power (monopoly) without financial constraints

Definition 1 *The price schedule is a function $p_t^k(X_t) : \mathbb{R} \to \mathbb{R}$ mapping the arbitrageur's position X_t to the equilibrium price in market k. A monopolistic equilibrium consists in arbitrageur's and hedgers' trades $(x_t^k, y_t^k)_{t=0,1}^{k\in\{A,B\}}$ and spreads Δ_t, such that (i) hedgers' demand is optimal given the equilibrium price path in each market, and (ii) the arbitrageur's trades maximize expected utility given the price schedule.*[32]

In Lemma 2 in the Appendix, I show that at $t = 0, 1$, hedgers' demand in market A is $Y_t = \frac{\mathbb{E}(p_{t+1})-p_t}{a\sigma^2} - s$. Inverting the demand and imposing market-clearing gives the price schedule faced by the arbitrageur:

Lemma 1 (Price Schedules) *Suppose that $n = 1$. At $t = 0, 1$, the price schedule faced the monopolist in market A is*

$$p_t(X_t) = \mathbb{E}(p_{t+1}) - a\sigma^2(s - X_t) \tag{6}$$

Proof. See Appendix A.2. ■

This equation shows that the equilibrium price today is increasing in the anticipated price next period. A similar dynamic relationship between prices arises in textbook presentations of Coasian dynamics, see e.g. Tirole (1988), p. 81. To see the intuition, consider hedgers in market A, who are natural sellers of the asset. When they anticipate a high price tomorrow, their willingness to accept a low price to trade today is reduced.

Thus, while the arbitrageur does not face competition from other traders, he competes with himself over time: hedgers understand at time 0 that the arbitrageur will retrade and provide additional liquidity at time 1.

[32]Recall that $x_t = x_t^A = -x_t^B$. As usual in the IO and finance literature, I assume that deviations by a zero mass of hedgers do not affect the course of the game.

The liquidity received at time 0 is durable from hedgers' point of view, because they remain exposed to the same source of risk over time. For instance, hedgers in market A (who have some willingness to sell) will suffer less from the endowment shock at time 1 if they have already reduced their positions at time 0.[33] Risk-sharing is thus akin to a durable "insurance", and is subject to Coasian dynamics. As is well-known from IO, a monopoly can evade Coasian dynamics when he has commitment power. I now compare the equilibrium under no commitment and perfect commitment.

Proposition 2 (No Commitment Equilibrium) *The equilibrium has the following properties:*

1. *The arbitrageur buys less than the supply in each market and increases his total position over time:* $x_0^{u_0,u_1} = \frac{2}{5}s$*,* $X_1^{u_0,u_1} = \frac{7}{10}s$*.*
2. *The spread decreases over time:* $\Delta_0^{u_0,u_1} = \frac{9}{5}a\sigma^2 s$*,* $\Delta_1^{u_0,u_1} = \frac{3}{5}a\sigma^2 s$*.*
3. *The arbitrageur earns strictly positive trading profits:* $\Omega_0^{u_0,u_1} = W_0 + \frac{9}{10}a\sigma^2 s^2$*.*

Proof. This is a special case of Proposition 4. ■

As the arbitrageur has market power, he does not fully integrate markets, even though there are no financial constraints. Further, the arbitrageur splits trades to control his price impact. Because the asset is conditionnally riskier at time 0, the arbitrageur trades more aggressively in the first trading round. Indeed, hedgers are more desperate to share risk at time 0 than at time 1.[34] This higher willingness to trade implies that price impact is higher at time 0 than at time 1.[35] Since markets are not fully integrated, prices do not converge and the arbitrageur realizes trading profits by earning the spread.

Proposition 3 (Perfect Commitment Equilibrium) *Suppose that the arbitrageur can commit to a trading strategy. Then:*

1. *The arbitrageur trades more than in the no-commitment case at time 0 and does not trade at time 1:* $x_0^{pc} = X_1^{pc} = \frac{s}{2}$*.*

[33] Note that this effect is maximal under our assumptions, because the exposure does not change sign.

[34] The uncertainty about the fundamental between 0 and 1 makes the capital gain uncertain, which ensures that hedgers' demand is downward-sloping at time 0. Hedgers have no discount factor in the model, but since hedgers are risk averse the higher risk at time 0 plays a similar role as a discount factor.

[35] This feature of the model is broadly consistent with the empirical evidence. See for instance Zarinelli et al. (2015): "for a given execution size, earlier transactions of the metaorder change the price more than later transactions", p.5. Hence, although stylized, the model captures salient features of real markets.

2. *Equilibrium spreads are larger at time 0 and time 1:* $\Delta_0^{pc} = 2a\sigma^2 s$ *and* $\Delta_1^{pc} = a\sigma^2 s$.

3. *The arbitrageur is better off, and hedgers are worse off:* $\Omega_0^{pc} = W_0 + a\sigma^2 s^2 > \Omega_0^{u_0,u_1} = W_0 + \frac{9}{10}a\sigma^2 s^2$ *and* $U_0^{pc} = -\frac{3}{4}a\sigma^2 s^2 < U_0^{u_0,u_1} = -\frac{27}{40}a\sigma^2 s^2$.

Proof. See Internet Appendix. ■

If the arbitrageur can commit to a trading strategy, he will trade only at time 0, eliminating competition with himself over time. Note that if the arbitrageur were to trade only at time 1, he would forego the benefit of the extra demand for risk-sharing at time 0. With perfect commitment ability, not surprisingly the arbitrageur limits further the amount of risk-sharing, hurting hedgers' welfare, increasing the spreads and his own trading profits relative to the no-commitment case.

4 Financially constrained monopoly

The definition of equilibrium remains the same as in Section 3.2, with the extra requirement that in each period the monopoly's positions must satisfy the financial constraints.

Equilibrium drivers. Inspecting the arbitrageur's financial constraints (4) and the price schedule (6) shows that the equilibrium will be determined by two key variables: the arbitrageur's capital W_0 and the risk benefit ratio ρ, defined as follows.

Notation 1 (Risk Benefit Ratio) *Let* $\rho \equiv \frac{\bar{e}}{a\sigma^2 s}$ *denote the risk benefit ratio.*

The risk benefit ratio is a cost benefit ratio of the trade from the arbitrageur's point of view. Note that risk appears both in the nominator and the denominator, in different forms. In the numerator, $\bar{e}$ measures by how much the fundamental can go up or down relative to the conditional mean, thus it measures the largest potential gain or loss due to the fundamental. In the denominator, the product $a\sigma^2 s$ measures the (maximum) profitability of the arbitrageur's trade in a given market. It is determined by the amount of hedging needs from hedgers. Hedgers are more desperate to share risk if the asset is riskier (larger σ^2), if their endowment shock is larger (larger s), or if they are more risk averse (larger a).[36] Because $\bar{e}$ represents the "tail" risk of the fundamental and σ^2 is its variance, ρ is larger when the distribution of the fundamental has less mass in the tails.

[36]Alternatively, one can think of the inverse of the risk benefit ratio as the maximum profit per unit of maximum risk, for each leg of the arbitrage.

Equilibrium multiplicity. Lemma 1 shows that the price schedule at time 0 depends on the expected price at time 1, which itself depends on whether the arbitrageur's constraint is binding at time 1. But the tightness of the constraint itself depends on prices. Therefore, in the presence of financial constraints, equilibria can be self-fulfilling and multiple equilibria coexist. Indeed, the arbitrageur chooses a position, but does not control hedgers' expectations. The anticipation of a binding constraint next period affects hedgers' demand today, and the price at which they are ready to trade.[37] But this price matters for the tightness of the financial constraints.

4.1 Equilibria with a slack constraint at time 1

Suppose that at time 0 hedgers anticipate that the arbitrageur is unconstrained at time 1. I conjecture that the arbitrageur chooses a position x_0 such that his time-1 constraint is slack, and verify under which conditions this holds. That is, given hedgers' anticipations u_1 (for *u*nconstrained at time 1), I determine under which conditions the arbitrageur chooses a position x_0 leading to state $l = \{u_1, \bar{c}_1, \underline{c}_1\}$ at time 1, where $\bar{c}_1$ denotes an upward-binding constraint and $\underline{c}_1$ a downward-binding constraint. Denoting $\Omega_0^{u_1,l}$ the value function associated with hedgers' expectations u_1 and state l, the arbitrageur chooses x_0 such that his time 0 expected utility Ω_0^u is $\Omega_0^u = \max\ (\Omega_0^{u_1,u_1}, \Omega_0^{u_1,\bar{c}_1}, \Omega_0^{u_1,\underline{c}_1})$. There exists an equilibrium with a slack constraint at time 1 iff $\Omega_0^u = \Omega_0^{u_1,u_1}$.

Given that the arbitrageur takes offsetting positions across markets, his wealth is risk-free, and only the trading profits in each period enter his value functions. Each value function $\Omega_0^{u_1,l}$ is thus defined as follows:

$$
\begin{aligned}
\Omega_0^{u_1,l} = \max_{x_0 \in \mathcal{F}_0^0} \quad & W_0 + x_0 \Delta_0^{u_1}(x_0) + \Omega_1^l(x_0) \\
\text{s.t.} \quad & \Delta_0^{u_1}(x_0) = 2a\sigma^2(s - x_0) + \Delta_1^{u_1}(x_0) \\
& + \text{additional consistency conditions}
\end{aligned}
$$

where $\mathcal{F}_0^0$ is the set of time-0 positions satisfying the financial constraint (4) at time 0, $\Delta_0^{u_1}(\cdot) \equiv p_t^B(\cdot) - p_t^A(\cdot)$ is the time-0 spread schedule, i.e. the difference between the price schedule in each market when hedgers assume that the arbitrageur's constraint is slack at time 1, and Ω_1^l is the continuation value at time 1, given state $l = \{u_1, \bar{c}_1, \underline{c}_1\}$. The spread

[37] Hegders anticipate prices, but since the model is of complete information and given the price schedule (6), there is an equivalence between the price next period and the arbitrageur's position. Thus, when I use the expressions "hedgers anticipate binding constraints", I mean the price induced by binding constraints.

schedule is fixed, in the sense that I keep hedgers' anticipation u_1 fixed. However, given this schedule, the arbitrageur internalizes that his trade leads to a binding or slack constraint and the associated continuation value at time 1 (i.e. Ω_1^l depends on l, not u_1). The associated consistency conditions, provided in the Appendix, ensure that the arbitrageur's actions are time-consistent, e.g. on the equilibrium path, he has indeed sufficient wealth at time 1 to hold his preferred position.

Proposition 4 (Equilibria with slack time-1 constraint) *There exist thresholds ω_0^u, ω_1^u, and ω^f, that define four regions in terms of initial arbitrage capital:*

1. *In the first region, $W_0 \geq \max(\omega_0^u, \omega_1^u)$, arbitrage capital is abundant, both constraints are slack in equilibrium, and the arbitrageur holds his preferred positions $x_0^{u_0,u_1}$ and $X_1^{u_0,u_1}$ given in Proposition 2 ($\mathbf{u_0}, \mathbf{u_1}$ equilibrium).*

2. *In second region ($\max(\omega^f, \min(\omega_0^u, \omega_1^u)) \leq W_0 < \max(\omega_0^u, \omega_1^u)$) and third regions ($\max(0, \omega^f) \leq W_0 < \min(\omega_0^u, \omega_1^u)$), either there is no equilibrium with a slack constraint at time 1 (**no** $\mathbf{u_1}$), or there exists an equilibrium in which the constraint binds at time 0 but not at time 1 ($\mathbf{c_0}, \mathbf{u_1}$ equilibrium).*
   In the latter, the arbitrageur holds a smaller position at time 0, allowing him to hold his preferred position $X_1^u(x_0)$ at time 1 without violating the constraint: $x_0^{c_0,u_1} < x_0^{u_0,u_1}$, $X_1^{c_0,u_1} = X_1^u(x_0^{c_0,u_1}) = \frac{s+x_0^{c_0,u_1}}{2}$. This equilibrium arises in particular in the second region when the risk benefit ratio is sufficiently low ($\rho < \frac{7}{10}$).

3. *In the fourth region, $0 \leq W_0 < \max(0, \omega^f)$, there is little arbitrage capital, and thus there is no equilibrium with a slack constraint at time 1 (**no** $\mathbf{u_1}$).*

Proof. See Appendix C. ■

The equilibrium takes a simple, intuitive form. When arbitrage capital is sufficiently abundant, constraints never bind, and the arbitrageur holds his preferred positions. In the opposite case, where capital is particularly scarce, the arbitrageur cannot trade in such a way that the constraint remains slack at time 1. In between these two regions, either we are in the former case, or in an intermediate case, where the arbitrageur reduces positions at time 0 to ensure that he can hold his preferred position at time 1. In other words, in such equilibria, the arbitrageur decides to save capital at time 0 to ensure that, *given* the time 0 position, he can trade his preferred quantity at time 1. The different cases are represented in Figure 1. To vary the risk benefit ratio, I hold hedgers' risk aversion a, the fundamental $\bar{e}$ and the variance σ^2 fixed, and vary the supply s.

In the second and third regions, the equilibrium can be determined analytically, but is not very tractable.[38] Thus, in Figure 2, I solve numerically for the equilibrium. The parameters chosen here show a typical case. Panel a shows in red the equilibrium that prevails in the regions where two cases may arise. Panel b eliminates the redundant information of the picture. The resulting equilibrium representation is strikingly simple. For $\rho \geq \frac{7}{10}$ (left-hand side of the picture), either capital is large enough such that no constraint binds and the arbitrageur holds his preferred positions, or the arbitrageur has not enough capital to keep the constraint slack at time 1. When the risk benefit ratio is lower, an intermediate case arises, where the arbitrageur's capital is relatively low, but sufficient to allow him to keep his constraint slack at time 1. Doing so, however, requires to trade less at time 0. Intuitively, the arbitrage is profitable enough relative to the risk of the position to relax the constraint at time 1. This is because wealth at time 1 increases sufficiently thanks to the capital gains made by the arbitrageur at time 0.

There may be no equilibrium with a slack constraint at time 1 for two reasons. Either there is not enough capital, so that it is impossible for the arbitrageur to keep the constraint slack and hold his preferred position at time 1 (this is so if $W_0 < \omega^f$, where the superscript f stands for floor). Or it is possible but not optimal for the arbitrageur to do so. When capital is not very abundant, and the risk benefit ratio is high enough, keeping enough dry powder at time 0 to trade his preferred position at time 1 is costly for the arbitrageur. It requires to reduce the time 0 trade away from his preferred level. In that case, the arbitrageur may thus deviate from the conjectured strategy and make the constraint binding at time 1. The loss from being constrained at time 1 may be more than offset by the benefit of trading his preferred position at time 0 (which is not necessarily the same amount as if the constraint is slack at time 1). This trade-off yields an endogenous threshold (represented by a dotted line in panels a and b of Figure 2) under which there is no equilibrium with a slack time 1 constraint.

[38] The reason is that given hegders' anticipations, and thus the price schedule, one must check that it is indeed optimal for the arbitrageur to follow the conjectured strategy. However, deviations involve making the constraint binding at time 1. Such binding constraint implies that the effect of the position x_0 on the time 1 profit is no longer quadratic, so that first-order conditions become highly non-linear. The solution can be written in closed-form, but is not very tractable. Proposition 13 in the Appendix provides somewhat more detailed equilibrium conditions.

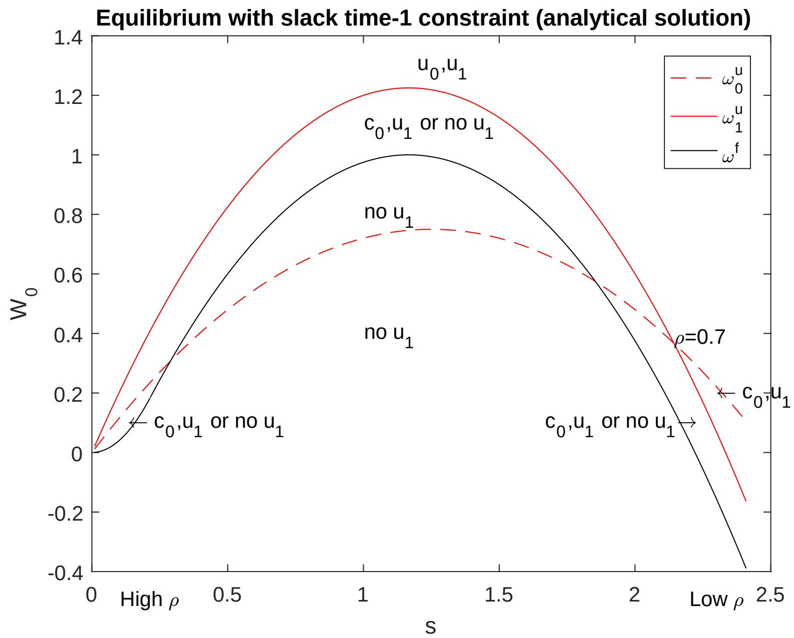


Figure 1: Equilibria with slack time-1 constraint. The parameters are $a = \sigma = 1$, $\bar{e} = 1.5$.

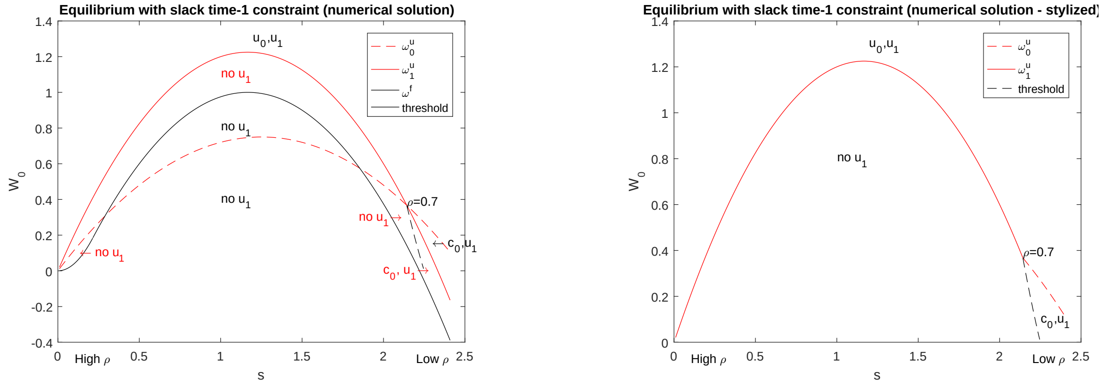
4.2 Equilibria with binding time-1 constraint

There exist equilibria, in which the constraint binds upwards at time 1. However, I show in the Appendix that there is no equilibrium in which the constraint binds downwards. It is intuitive: since arbitrageurs naturally want to go long the spread, the main issue arising from limited capital is that they cannot go long as much as they wish, i.e. that the constraint binds upwards. I proceed as in the previous section: I conjecture an equilibrium strategy and determine under which conditions it holds in equilibrium.

Proposition 5 (Equilibria with binding time-1 constraint) *Let ω_0^p and ω_1^p denote two thresholds. There are equilibria in which the arbitrageur's constraint binds upwards at time 1, as follows.*

1. *If $0 \leq \rho < \frac{3}{4}$, then $\omega_1^p < \omega_0^p$, and there are three regions in terms of arbitrage capital:*

   (a) *In the first region, with $0 \leq W_0 < \max(0, \omega^f)$, the arbitrageur's constraint binds upwards at time 0 and time 1 in equilibrium ($\mathbf{c_0}, \mathbf{c_1}$ equilibrium). This equilibrium is the same as in the constrained competitive case, for a given level of capital. The arbitrageur holds the largest (long) positions satisfying the financial constraints at each date: $x_0^{c_0,c_1} = \bar{x}_0$, and $X_1^{c_0,c_1} = \bar{X}_1(\bar{x}_0)$.*



(a) (b)

Figure 2: Equilibria with slack time-1 constraint. The parameters are $a = \sigma = 1$, $\bar{e} = 1.5$.

(b) *In the second region, with* $\max(0, \omega^f) \leq W_0 < \max(0, \omega_1^p)$*, there are two cases: there is either no equilibrium in which the arbitrageur's constraint binds upwards at time 1 (**no** $\mathbf{c_1}$), or an equilibrium where both constraints bind as in (a).*

(c) *In the third region, with* $\max(0, \omega_1^p) \leq W_0 < \omega_0^p$ *or* $\omega_0^p \leq W_0$*, there is no equilibrium in which the arbitrageur's constraint binds upwards at time 1 (**no** $\mathbf{c_1}$).*

2. *If* $\rho \geq \frac{3}{4}$*, then* $\omega_1^p > \omega_0^p$*, and there are four regions in terms of arbitrage capital:*

   (a) *In the first region, with* $0 \leq W_0 < \omega^f$*, the equilibrium is* $\mathbf{c_0}, \mathbf{c_1}$*, as in case 1a.*

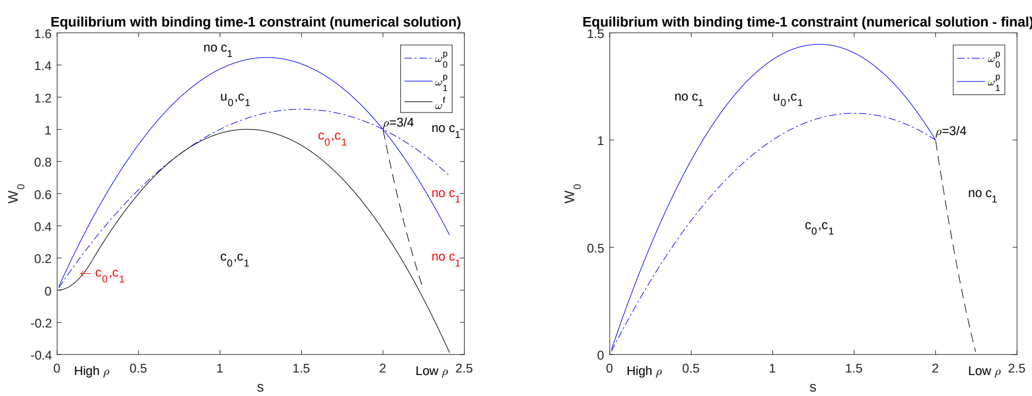
   (b) *In the second region, with* $\omega^f \leq W_0 < \omega_0^p$*, the equilibrium is the same as in 1b.*

   (c) *In the third region, with* $\omega_0^p \leq W_0 < \omega_1^p$*, there may exist an equilibrium in which the abritrageur's constraint binds upwards at time 1 and is slack at time 0 (*$\mathbf{u_0}, \mathbf{c_1}$ *equilibrium). In this equilibrium, the arbitrageur holds the same amount as in the perfect commitment case at time 0,* $x_0^{u_0,c_1} = x_0^{pc}$*, and the largest position satisfying the constraint at time 1:* $X_1^{u_0,c_1} = \bar{X}_1(x_0^{u_0,c_1})$*.*

   (d) *In the fourth region, with* $\omega_1^p \leq W_0$*, there is no equilibrium in which the arbitrageur's constraint binds upwards at time 1 (**no** $\mathbf{c_1}$), as in 1c.*

For the same reason as with the u_1 equilibrium, some cases have an analytical albeit rather intractable solution. So I proceed as before and illustrate Proposition 5 graphically.I determine the numerical solution in Figure 3. The graph is a typical case, and does not critically depend on the choice of parameters. Panel b of Figure 3 shows the equilibrium regions. The intuition for the form of the equilibrium is simple. With abundant capital, there is no equilibrium in which the arbitrageur's constraint binds at time 1 (no c_1). With little capital, constraints are likely to bind not only at time 1, but also at time 0 (c_0, c_1). With intermediate capital, the most interesting case arises when the risk benefit ratio is large enough ($\rho \geq \frac{3}{4}$) and arbitrage capital is intermediate ($\omega_0^p \leq W_0 < \omega_1^p$). In this case, the arbitrageur's constraint binds at time 1, but not at time 0. The conditions on arbitrage capital for this equilibrium are intuitive. The tresholds ω_t^p for this $p$artly constrained equilibrium are associated with the time t constraints. There must be enough capital for the arbitrageur to be able to hold $x_0^{u_0,c_1}$ at time 0 ($W_0 \geq \omega_0^p$). However, capital cannot be large enough ($W_0 < \omega_1^p$), for otherwise, the constraint would not bind at time 1.

Holding a larger position at time 0 is a necessary condition for the constraint to bind at time 1. Doing so, the arbitrageur pledges more capital at time 0, increasing the chance to be constrained at time 1 (of course, not only the position but also the depth of the market, and therefore the arbitrageurs' trading profits, are different across the two equilibria). This equilibrium arises only if the risk benefit ratio is large enough. It makes sense: for the constraint to bind at time 1, it must be that the position is sufficiently risky relative to profits. Otherwise, trading is not very capital intensive, and the arbitrageur will be free to re-optimize when time 1 comes.



(a) (b)

Figure 3: Equilibria with binding time-1 constraint. The parameters are $a = \sigma = 1$, $\bar{e} = 1.5$.

Having a sufficiently low level of capital does not only ensure time consistency. A low W_0 also ensures that deviating from being constrained at time 1 is not attractive. With sufficiently low capital, the arbitrageur must take a small position at time 0 to ensure that his time-1 constraint is slack. This reduces his time-0 profit, and thus prevents the arbitrageur from fully benefiting from the deviation.

4.3 Coexistence

Proposition 6 (Equilibria with Slack and Binding Time-1 Constraints May Coexist)

- *There is a unique equilibrium when arbitrage capital is either sufficiently low or sufficiently high:*
  - *If* $0 \leq W_0 < \max(0, \omega^f)$*, the unique equilibrium is* $\mathbf{c_0}, \mathbf{c_1}$*.*
  - *If* $W_0 \geq \max(\omega_0^u, \omega_1^u, \omega_1^p)$*, the unique equilibrium is* $\mathbf{u_0}, \mathbf{u_1}$*.*
- *When capital is intermediate, i.e. if* $\max(0, \omega^f) \leq W_0 < \max(\omega_0^u, \omega_1^u, \omega_1^p)$*, multiple equilibria may coexist depending on the level of* ρ*, as detailed in Proposition 18.*

Figure 4 superimposes the results of the numerical solutions in panel b of Figures 2 and 3. There are four possible equilibria: u_0, u_1, c_0, u_1, u_0, c_1, and c_0, c_1. Equilibria coexist in two regions characterized by intermediate capital and a large enough risk benefit ratio. In all other regions, the equilibrium is unique. With abundant capital, the unconstrained equilibrium u_0, u_1 is unique, while with scarce capital, the fully constrained equilibrium c_0, c_1 prevails. Further, when the risk benefit ratio and capital are low, the unique prediction of the model is c_0, u_1, i.e. the arbitrageur reduces his time-0 position to remain unconstrained. When the risk benefit ratio is larger, the fully constrained and unconstrained equilibria overlap. If we increase further the risk benefit ratio, the unconstrained equilibrium coexists with the u_0, c_1 equilibrium.

Multiple equilibria arise even though the abitrageur has market power and chooses how much to trade. Intuitively, the arbitrageur does choose quantities, but cannot pick hedgers' expectations. Since hedgers' expectations affect market depth (through the price schedule), they affect the arbitrageur's incentives to trade in one way or another, leading to self-fulfilling equilibria. It is not surprising that equilibria coexist for intermediate amount of capital and moderately capital intensive positions. In such case, the arbitrageur's constraint is close to be binding, and thus hedgers' anticipations, by affecting market depth, may tip the equilibrium outcome in one way or another. Market depth is not as determinant when capital is very scarce or abundant.

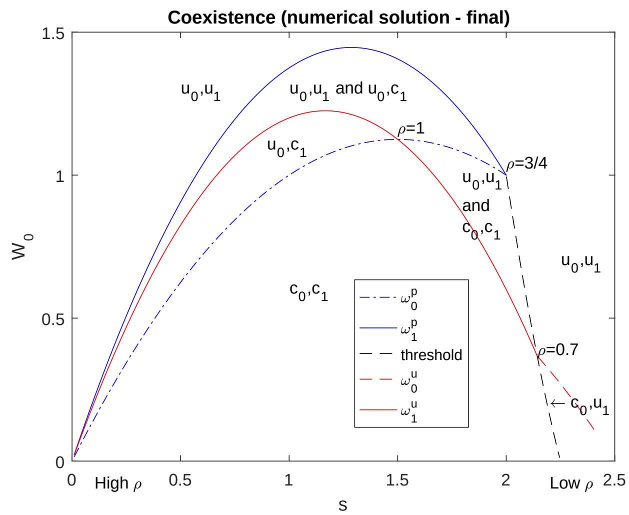


Figure 4: Coexistence. The parameters are $a = \sigma = 1$, $\bar{e} = 1.5$.

5 Welfare

In this section, I show that (i) With a monopolistic arbitrageur, the u_0, c_1 equilibrium Pareto-dominates the u_0, u_1 equilibrium. (ii) Imposing constraints on a competitive arbitrageur reduces welfare. However, imposing the same constraint on a monopoly with the same amount of initial capital may increase welfare.

Proposition 7 **(u_0, c_1 Pareto-dominates u_0, u_1)** *Suppose that a partly constrained equilibrium (u_0, c_1) exists and that it coexists with the unconstrained equilibrium (u_0, u_1). Then in the partly constrained equilibrium:*

1. *Spreads are smaller:* $\Delta_t^{u_0,c_1} < \Delta_t^{u_0,u_1}$, $t = 0, 1$;
2. *The arbitrageur holds larger positions:* $X_t^{u_0,c_1} > X_t^{u_0,u_1}$, $t = 0, 1$;
3. *Hedgers are better off;*
4. *The arbitrageur is better off if and only if* $W_0 \in [\max(\omega_1^u, \omega_0^p), \omega^a)$, *where* $\omega^a < \omega_1^p$. *This interval is non-empty if* $\rho \geq \frac{2\sqrt{5}}{5} > \frac{3}{4}$, *i.e. there exists a non-empty set of parameters such that the arbitrageur is better off.*

Proof. See Appendix F. ■

A key consequence of this result is that imposing financial constraints on a monopolistic arbitrageur improves social welfare in certain markets:

Corollary 2 (Constraints on a Monopoly may be Welfare-improving)

1. *In markets with a large amount of capital, imposing financial constraints on a monopolistic arbitrageur has no effect, because the constraint never binds.*

2. *In markets with intermediate capital ($\omega_0^p \leq W_0 < w^a$) and sufficiently high risk-benefit ratios ($\rho > \frac{2\sqrt{5}}{5}$), imposing financial constraints on a monopolistic arbitrageur*

   - *improves social welfare if $\omega_0^p \leq W_0 < \omega_1^u$ and*
   - *may improve it if $\omega_1^u \leq W_0 < \omega^a$.*

   If the risk-benefit ratio is lower ($\frac{3}{4} \leq \rho < \frac{2\sqrt{5}}{5}$) and / or capital is larger ($\omega^a < W_0 \leq \omega_1^p$), the arbitrageur is worse off.

3. *In markets with sufficiently small capital, imposing constraints either has no effect or leads to a reduction in liquidity in at least one date, with at least one type of investors being worse off.*

Proof. See Appendix F. ■

This result provides conditions under which imposing financial constraints on a large arbitrageur may increase or decrease social welfare. Thus, this result may cast light on the debates about the effects of the tightening of capital requirements that have occurred in the aftermath of the 2007-2009 crisis. These debates have mostly focused on the negative implications of the tightening of existing financial constraints or the introduction of new constraints (see e.g. Boyarchenko et al., 2018). Here instead, I show that there are benefits to regulating an arbitrageur with market power through price-based financial constraints. These benefits occur only in markets with high risk benefit ratios and where the arbitrageur is neither too well nor too poorly capitalized. In other situations, either the constraint has no effect, or hurts at least one type of market participant. This result is at odds with the competitive case, where financial constraints are either irrelevant or reduce hedgers' welfare.

In general, adding a friction on top of another one may bring the economy either closer or further away from the first-best. Here financial constraints interact with the arbitrageur's commitment problem. The arbitrageur faces a Coasian problem of competition with one-self over time, and benefits from being able to commit to trade only once. Without commitment, the arbitrageur cannot help retrading. However, the financial constraint does not rule out retrading. It serves as an endogenous and imperfect commitment device to trade *less* at time 1. Indeed, we have $x_0^{u_0,u_1} < x_0^{u_0,c_1}$, but $x_1^{u_0,u_1} > x_1^{u_0,c_1}$ (recall that $x_1 = X_1 - x_0$). However, the arbitrageur holds larger positions *at all dates* in the u_0, c_1 equilibrium, i.e. $x_0^{u_0,u_1} > x_0^{u_0,u_1}$ and $X_1^{u_0,c_1} = x_0^{u_0,c_1} + x_1^{u_0,c_1} > x_0^{u_0,u_1} + x_1^{u_0,u_1} = X_1^{u_0,u_1}$.

A larger position at time 0 is inherent to the u_0, c_1 equilibrium: by pledging more capital early on, the arbitrageur ensures that the constraint is indeed binding at time 1. However, doing so, the arbitrageur also makes larger gains, which mechanically increase the position he can afford at time 1. The larger positions lead to smaller spreads, although the price impact is different across equilibria: in the u_0, c_1 equilibrium, hedgers demand more liquidity at time 0 in anticipation of the binding constraint at time 1, potentially increasing the spread.

Both the arbitrageur and hedgers are better off in the u_0, c_1 equilibrium, but note that the conditions are stricter for the arbitrageur to be better off, because the retrading maintains some Coasian dynamics. Arbitrageurs would be better off if they were able to divert capital gains entirely between 0 and 1 and keep the same level of capital after trading at time 0. This would allow them to earn full commitment profits. For hedgers, the situation is opposite: they benefit from the fact that the arbitrageur provides more liquidity at time 0, and that he evenutally also holds a larger position at time 1, despite the binding constraint at time 1. In fact, both the increase in early liquidity and the increase in the final position are necessary to obtain hedgers' welfare gain. To see this, I compute hedgers' welfare under two counterfactual allocations.

Proposition 8 (Why hedgers are better off) *Suppose the u_0, u_1 and u_0, c_1 equilibria coexist.*

- *Counterfactual 1: If the time-1 position increases to $X_1^{u_0,c_1}$ without an increase in the time 0 position ($x_0^{cf1} = x_0^{u_0,u_1}$), hedgers are better off in the u_0, c_1 equilibrium than in the counterfactual.*

- *Counterfactual 2: If the time-1 position remains the same $X_1^{cf2} = X_1^{u_0,u_1}$ but the time 0 position increases ($x_0^{cf1} = x_0^{u_0,c_1}$), hedgers are better off in the u_0, u_1 equilibrium than in the counterfactual.*

Thus, hedgers are better off because both the time-1 and time-0 positions increase in the u_0, c_1 equilibrium.

Proof. See Appendix F. ■

This result shows that varying both the extensive margin (the total position, counterfactual 1) and the intensive margin (the amount of liquidity at time 0, counterfactual 2) is necessary to improve hedgers' welfare. It is intuitive that the increase in the final position matters: hedgers have shared more risk in the market, getting closer to the first-best.

The reason why the intensive margin – given a total amount of liquidity, how much liquidity is provided at time 0 – matters is that the asset is conditionally riskier at time 0 than at time 1, making it valuable for hedgers to share risk early. The binding constraint at time 1, which is required to reduce the Coasian dynamics, requires the arbitrageur to trade more aggressively at time 0. This increase in liquidity at time 0 is essential.

Note that the uncertainty about the fundamental between 0 and 1 is a key ingredient of the model: it implies that hedger's demand remains downward-sloping at time 0. Without it, there is no demand for liquidity at time 0. Indeed, the risky asset would be temporarily risk-free, so that hedgers would flatten out the demand – in other words, hedgers would become arbitrageurs themselves. The price would simply equal the expected price next period and arbitrageurs would have no incentive to trade (for any trade would push the price away from the expected time 1 price, and hedgers would step in to correct this distortion).

6 Empirical implications

The model also delivers new, and to the best of my knowledge, untested empirical predictions: (i) A drop in arbitrage capital may increase spreads when the risk benefit ratio is low enough, and first decrease and then increase them otherwise. Similar effects occur for positions, margins, and VaR. (ii) Merging constrained competitive arbitrageurs into a single arbitrageur softens the financial constraint and may reduce spreads at time 1.

6.1 Price effects of a drop in capital

Figures 5a to 8b show the competitive and imperfectly competitive spreads as a function of arbitrage capital at time 0 and 1.[39] For $\rho \leq \frac{3}{4}$, there is no qualitative difference between the competitive and monopolistic cases: a drop in capital leads to an increase in spread when the economy enters the region where constraints bind. For $\rho > \frac{3}{4}$, the comparative statics are qualitatively different: a drop in capital first *reduces* the spread. A further drop leads to an increase in spreads. The reduction is due to the fact that as capital drops, hedgers rationally anticipate that constraints will bind at time 1 (u_0, c_1 becomes the unique equilibrium as capital drops), leading to a reduction in spreads relative to the previous situation, in which u_0, u_1 is the equilibrium (possibly coexisting with u_0, c_1).

[39]For each type of equilibrium, it is possible to write the comparative statics of the spreads with respect to capital in analytical form (see Corollaries 6 and 8 for the expression of the spreads in the different cases). However, it is difficult to do so across equilibria, hence the largely numerical approach.

I am not aware of any other setting delivering such empirical implication, nor of any test of this implication in the literature. The model emphasizes the role of the risk benefit ratio as a conditioning variable for empirical tests of the relation between spreads and arbitrage capital. Markets characterized by a high risk benefit ratio are those in which risk is large relative to profitability. Empirically, the risk could be proxied by a certain quantile of the return distribution (e.g. 99% quantile), and profitability by the P&L of trading desks involved in a given arbitrage.

6.2 Positions, margins and VaR

The non-monotonic effect of a drop in capital on prices carries through to positions, margins, and VaRs.

Corollary 3 *When competitive arbitrageurs hit their funding constraints (e.g. following a drop in capital), positions, margins and the VaR at $t = 0, 1$ decrease. With a monopoly, the same occurs for a sufficiently low risk benefit ratio ($\rho < \frac{3}{4}$). If the ratio is high instead, a drop in capital first increases and then decreases time-0 positions, margins and the VaR, and may do so at time 1.*

Proof. See Appendix F. ■

For an outside observer, an increases in positions and VaRs following a drop in arbitrage capital may look like a gamble for resurrection. Here, however, the effect is solely driven by the interaction between market power and the financial constraint.

6.3 Liquidity fragility

The multiplicity of equilibria that occurs for intermediate capital and sufficiently high risk benefit ratio shows that market liquidity may be "fragile": it may jump following a change in the market's (hedgers') expectations without a large shock, if any, to arbitrage capital. With competitive arbitrageurs, such fragility occurs only when margins are procyclical (Brunnermeier and Pedersen, 2009). Instead, when the arbitrageur is monopolistic, fragility occurs even though margins are countercyclical. Hence, markets with a dominant arbitrageur may be more prone to sudden jumps in market liquidity.

6.4 Effects of a merger

Merging competitive arbitrageurs into a monopoly affects both equilibrium prices, positions, and capital requirements. Note first that

Proposition 9 *The wealth thresholds can be written as the sum of two terms:*

$$\omega_t^l = \underbrace{\Lambda_t^l \, s\bar{e}}_{\textit{maximum position loss}} - \underbrace{\Gamma_t^l \, a\sigma^2 s^2}_{\textit{profit adjustment}}, \quad l \in \{u, p\}$$

Similary, $\omega^* = \Lambda^* s\bar{e} - \Gamma^* a\sigma^2 s^2$, with $\Lambda_0^l < \Lambda_1^l < \Lambda^* = 2$, and $\Gamma^* = 0 < \Gamma_0^l < \Gamma_1^l$ (no profit adjustment in the competitive case). Thus, the monopoly's constraint is always slacker than the competitive arbitrageurs' constraints: $\max(\omega_0^l, \omega_1^l) < \omega^*$*, $l \in \{u, p\}$.*

Proof. Follows from the definitions of ω_t^u, ω_t^p and ω^*. ■

The form of the ω_t^l thresholds is intuitive. The first term, $\Lambda_t^l s\bar{e}$, represents the maximum potential loss caused by fundamentals. It is the product of the worst possible change in fundamental $\bar{e}$ and the arbitrageur's total exposure at time t, $\Lambda_t^l s$. In a competitive market, arbitrageurs fully integrate markets A and B, and their total position is $s - (-s) = 2s$, as each leg of the arbitrage is of size $|s|$.[40] In a monopolistic market, the arbitrageur acquires a smaller position than competitive arbitrageurs and split orders to limit his price impact, so $\Lambda_0^l < \Lambda_1^l$ and $\Lambda_t^l < 2$, for $l = u, p$. The second term in ω_t^l, $-\Gamma_t^l a\sigma^2 s^2$, is an adjustment measuring how much accumulated profits due to market power relax the capital requirement. This term is zero in a competitive market, since profits are competed away. The monopoly always earns capital gains in equilibrium, which provides "cushion" from the financiers' point of view.[41]

[40] E.g. $\Lambda_t^u s = 2X_t$.

[41] The trade-off between the risk of the position and the profit adjustment also means that i) the wealth threshold may be negative and thus non-binding as long as the arbitrageur starts with positive wealth (i.e. ω_t^u may be negative); this occurs if ρ is small enough; and ii) that the arbitrageur may be constrained at time 0, but not at time 1, or vice versa. Intuitively, at time 0, the position is smaller, but so is the profit. At time 1, both the position and the profit increase (i.e. both Λ_t^l and Γ_t^l increase with time). If the profit increases faster than the position, the arbitrageur's constraint relaxes at time 1, even if the constraint was binding at time 0. In a competitive market, being intially unconstrained implies that the price gap between the two assets is closed. Thus wealth does not increase over time. Further, at time 1, when hedgers receive a new shock, the previous shock has been fully hedged. This implies that if arbitrageurs had enough wealth to close the price gap at time 0, they have enough wealth to do so at time 1 as well. Hence the condition boils down to a single threshold. Conversely, in a competitive market, if the constraint binds at time 0, it also binds at time 1. It is not necessarily the case, however, when the arbitrageur has market power. The reason is simply that competitive arbitrageurs do not internalize their price impact, while the monopoly does. Because he takes into account his price impact, the monopoly takes a smaller position at time 1.

Merging arbitrageurs thus softens their funding constraint. As a result, when there is limited arbitrage capital in aggregate, a monopoly may be unconstrained or partly constrained, while competitive arbitrageurs are constrained. Hence, while competitive arbitrageurs' constraints are slack, the spread is always smaller than with a monopoly, but it is no longer always the case when their constraints bind.

Proposition 10 (Merging Arbitrageurs May Reduce Time-1 Spreads) *Suppose that competitive arbitrageurs are constrained, i.e.* $W_0 < \omega^*$. *Merging all arbitrageurs into a monopoly, holding capital constant, leads to a decrease in the time 1 spread (and equivalently, to an increase in time-1 position)*

- *If* $W_0 \in [\omega_0^p, \omega_1^p[$ *when the equilibrium is* u_0, c_1,
- *If* $W_0 \in [\omega_1^u, \omega^m[$ *and* $\rho \geq \frac{21}{10}$ *when the equilibrium is* u_0, u_1 $(\omega_1^u < \omega^m)$.

From this result, we see that merging arbitrageurs always reduces the time-1 spread if the monopoly is partly constrained in equilibrium, and may reduce it if the monopoly is unconstrained. Figure 5 illustrates the case with a partly constrained arbitrageur, although the difference between the competitive and monopolistic spreads is very small. At time 0, it is possible to write conditions under which there is no such decrease in the spread. It is difficult to rule out such improvement analytically. However, in all numerical examples I considered, time-0 spreads were larger under a monopoly than constrained competitive arbitrageurs. In any event, the reduction in spread at time 1 does not lead to a welfare improvement, as the competitive equilibrium is constrained efficient (Gromb and Vayanos, 2002). Thus, in this set up, merging arbitrageurs may lead to smaller spreads, but not to a Pareto improvement.

7 Discussion and extensions

The oligopolistic case. With a monopolistic arbitrageur, imposing VaR-based constraints may be socially desirable in certain markets. Is it a special case? The short answer is no. In the Internet Appendix, I derive the analogs of the u_0, u_1 and u_0, c_1 equilibria with n equally capitalized arbitrageurs. These equilibria may coexist, as in the monopolistic case. When they do coexist, the u_0, c_1 equilibrium no longer Pareto-dominates. However, if we start from a no-constraint oligopolistic economy and impose constraints, then we still obtain a

The profits from time 0 may be large enough to finance this smaller position, but are never large enough to finance a position that eliminates the spread.

Pareto-improvement under similar conditions as the monopolistic case: intermediate capital and sufficiently high risk benefit ratio.

The fact that u_0, c_1 no longer Pareto-dominates relates to the equilibrium tightness of the constraint: capital must be sufficiently large to ensure that the u_0, u_1 exists, and as n increases, this condition first tightens (unless ρ is very large). When we compare the expected utilities between u_0, u_1 and u_0, c_1, instead, we obtain an upper bound on capital. This is because expected utility is concave in W_0 in the u_0, c_1 equilibrium, while it is linear in the u_0, u_1 case. I show that for $n > 1$, these two conditions cannot be satisfied simultaneously. However, if we compare a no-constraint case to a case with constraint, there is no need to take into account the lower bound for the u_0, u_1 equilibrium: without constraints, it is always the prevailing equilibrium.

In the oligopolistic case, I provide examples in which hedgers are better off even without an increase in the time-1 position (relative to the u_0, u_1 equilibrium). It suffices that the time-0 position increases. That the final position does not increase is a consequence of competition: more competition leads to lower capital gains, so that arbitrageurs' constraint at time 1 is tighter than under a monopoly. This fact implies that spreads may be reduced only at time 0 relative to a no-constraint case.

Assumptions. I now discuss the bearing of my assumptions for the results. First, the fact that there are only two trading rounds is for simplicity only. Given that constraints may bind only occasionally, adding trading rounds will simply multiply the already large number of potential cases without affecting the economics at play. Further, the assumption of two trading rounds entails a loss of generality when the arbitrage is risky (see Gromb and Vayanos, 2002), but not when the arbitrage is risk-free. Second, the assumption of mean-variance preferences allows to invert hedgers' demand and keep it linear. The model with linear demand is widely used in applications in economic theory and IO. Working with a more general expected utility would make the model intractable, ruling out a sharp characterisation of the equilibrium. Note that I solve for time-consistent strategies for hedgers. Given that arbitrageurs face a riskless arbitrage opportunity, their preferences have no bearing on the result.

The third simplifying assumption is precisly that hedgers' hedging needs are constant and deterministic, implying that the arbitrage is risk-free. The fact that hedging needs are constant is for simplicity and does not affect primarily the results. For instance, if hedging needs are known to decrease over time, the demand for liquidity to the arbitrageur would decrease over time for exogenous reasons, worsening the effects of competition with oneself.

The fact that hedging needs are deterministic implies that arbitrageurs take the appropriate level of risk; the equilibrium is constrained-efficient (Gromb and Vayanos, 2002).

With risky arbitrage, Coasian dynamics would remain, but additional effects would arise. On the one hand, hedgers would be even more eager to share risk early. On the other hand, arbitrageurs would be less strategic, because they also would also be willing to share risk. However, they would still compete with themselves over time: when the arbitrage is risky, it is possible to derive a generalization of the price schedule (6), which is at the heart of the Coasian dynamics.

There is no reason to believe that the welfare improvement induced by a binding constraint at time 1 would not remain in this more general setting. Indeed, the risk-free arbitrage case is simply the limiting case of the risky arbitrage case when arbitrage risk goes to zero. In fact, hedgers would benefit even more from more liquidity at time 0, since they would be more risk averse with risky arbitrage (to the extent that there would be two sources of risk regarding their endowment shock, both s and ϵ being random). Thus, the main result of this paper is likely to remain after introducing some arbitrage risk, in particular if arbitrageurs are not too risk averse. But as arbitrageurs may fail to take the efficient level of risk, and in particular fail to internalize their pecuniary externalities, other interesting topics would arise. Investigating these topics, however, is beyond the scope of this paper.

Form of the financial constraint. In my model, the tightness of the constraint (4) is endogenous at all dates, to the extent that wealth (on the left-hand side) is endogenous at time 1, and that margins (on the right-hand side) are endogenous at time 0 and time 1. Further, the constraint is forward-looking, as margins depend on expected price changes.

I analyse a model with the same constraint but exogenous margins m_t in the Internet Appendix. This case is also relevant in practice, as regulators or exchanges may impose fixed margins.[42] Perhaps surprisingly, this model is less tractable than the model with endogenous margins, so I used numerical analysis. In this model, I could not find parameter combinations such that a u_0, c_1 equilibrium exists when margins are constant over time, i.e. when $m_0 = m_1 = m$. In this case, when the time-1 constraint is sufficiently tight to prevent deviations, the time-0 constraint is also tight and prevents the arbitrageur from taking his preferred position at time 0. In other words, the equilibrium is always c_0, c_1, not u_0, c_1. When margins are allowed to vary over time, it is of course possible to pick m_0 such that the constraint is slack at time 0. However, this model is clearly inferior to the one where

[42] This is the case for instance for VIX futures on CBOE. See also my remarks in footnote 17 in the introduction.

margins endogenously change over time, in particular because it allows one extra degree of freedom.

Note that with fixed margins, the time-1 constraint is endogenous, because wealth at time 1 is endogenous, but not the time-0 constraint; thus it is the fact that the constraint with endogenous margins is forward-looking that generates the u_0, c_1 equilibrium. Margins are forward-looking, because they are set using VaR.

In the Internet Appendix, I also derived margins in the case where they are set to cover an Expected Shortfall of level α. Altough parametrically different, margins keep exacly the same functional form $m_t = \zeta_\alpha - \beta_\alpha \frac{\Delta_t - \Delta_{t+1}}{2}$, and thus retain the same properties.

Next, I consider another type of constraint used in the IO literature.

Comparison to IO. The result of this paper may be puzzling from the point of view of the IO literature, since generally giving arbitrageurs some commitment power reduces hedgers' (consumers') welfare. The key reason for the difference is that the tightness of the constraint is endogenous, because it depends on equilibrium prices.

The IO literature, instead, has considered the effects of *exogenous* capacity constraints (McAfee and Wiseman, 2008). To facilitate comparisons, I now study the effects of such constraints. Capacity constraints resemble position limits, which are in use in some derivatives markets (i.e. a maximum number of contracts in a given derivative).

The arbitrageur trades sequentially; however, at time 0, before the first trading round, the arbitrageur (or its risk manager) chooses the maximum number of shares k he may trade per period. A capacity k costs $c(k)$, e.g. $c(k) = ck$. I look at vanishly small costs, $c \to 0$. The tightness of the constraint is set ex-ante and is independent of prices, so the equilibrium remains unique:

Proposition 11 (Capacity Constraints) *Suppose that the per unit capacity cost c is small, but strictly positive. The arbitrageur chooses optimal capacity* $k = \frac{3}{10}s = x_1^{u_0,u_1}$.

1. *The arbitrageur trades less than in the non commitment, no constraint case, and less than the perfect commtiment:* $x_0^{cc} < x_0^{u_0,u_1} < x_0^{pc}$. *At time 1, he holds a larger position than with perfect commitment and a smaller one than without commitment:* $X_1^{pc} < X_1^{cc} < X_1^{u_0,u_1}$.

2. *Equilibrium spreads at time 0 are larger than in the other cases:* $\Delta_0^{u_0,u_1} < \Delta_0^{pc} < \Delta_0^{cc} = \frac{11}{5}a\sigma^2 s$. *At time 1, spreads increase relative to the no commitment case and decrease relative to perfect commitment: and* $\Delta_1^{u_0,u_1} < \Delta_1^{cc} = \frac{4}{5}a\sigma^2 s < \Delta_1^{pc}$.

3. *The arbitrageur is worse off than if he could fully commit and almost as well-off as without commitment:* $\Omega_0^{u_0,u_1} \approx \Omega_0^{cc} < \Omega_0^{pc}$

4. *Hedgers ar worse off than in the other cases:* $U_0^{cc} = -\frac{31}{40}a\sigma^2 s^2 < U_0^{pc} < U_0^{u_0,u_1}$. *Hence, hedgers are also worse-off than in the* u_0, c_1 *equilibrium.*

Proof. See Internet Appendix. ■

The proof shows that the arbitrageur is indifferent between two capacities: $k = x_1^{u_0,u_1}$ and $k = x_0^{u_0,u_1}$. However, with a small but positive cost, the latter is more expensive. Because of the small capacity, the arbitrageur restricts liquidity at time 0 more than in the other cases, which explains why the time-0 spread is the largest. The final position, however, is intermediate, and thus so is the time-1 spread. While the arbitrageur provide less liquidity in total than in the no commitment case, he benefits from the larger spreads, and so achieves almost the same welfare (up to the cost). Hedgers, however, suffer from the lack of liquidity at time 0; they would be better off with higher liquidity at time 0 and a lower final position, as in the perfect commitment case. In sum, capacity constraints *delay* risk-sharing, hurting hedgers.

8 Conclusion

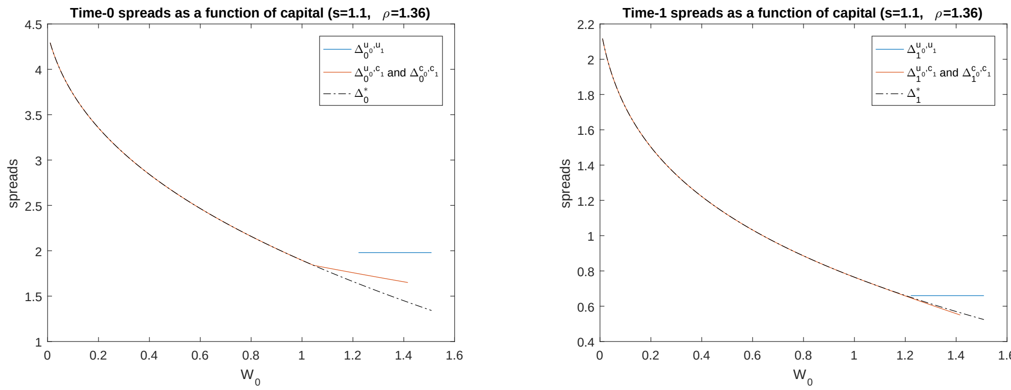
In this paper, I consider the effects of imperfect competition among arbitrageurs subject to financial constraints. I characterize markets in which imposing these constraints may benefit both arbitrageurs and their trading counterparties (hedgers) and improve market liquidity. This analysis reveals novel and subtle mechanisms through which constraints affects both types of investors in the presence of arbitrageurs' market power. On the one hand, a binding constraint can mitigate the commitment problem of an imperfectly competitive arbitrageur. On the other hand, the constraint is endogenous to the arbitrageur's trading strategy: to make a constraint binding in the future, an arbitrageur must trade more aggressively today, which speeds up the arbitrage and risk-sharing. This strategy yields capital gains, leading to some amount of retrading. The increase in early liquidity combined with sufficient retrading makes hedgers better off. These mechanisms are specific to imperfect competition: in a competitive economy, arbitrageurs take prices as given and do not recognize the commitment problem.

The analysis also delivers new empirical predictions about spreads, VaR, positions, and margins as a function of capital and the risk benefit ratio of the arbitrage. Spreads are

U-shaped and positions, margins, and VaR are hump-shaped in arbitrage capital for a large enough risk benefit ratio. But they are respectively increasing and decreasing in capital for low risk benefit ratio, or when arbitrageurs are competitive. Overall, imposing financial constraints on arbitrageurs may have diametrically different effects for different structures of the arbitrage industry.

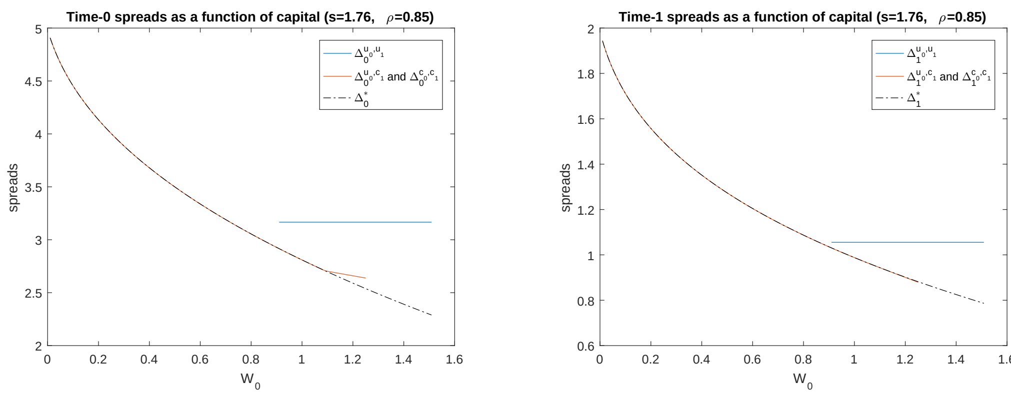
The model may be extended to consider internal allocation of capital across trading desks or systemic risk. In my framework, an arbitrageur with market power would benefit from being able to commit to decrease her capital level in the future. This may be achieved by pledging capital gains to new trades, for instance by reallocating capital across different trading desks over time. Such effect would not arise with competitive arbitrageurs. Imperfect competition among arbitrageurs should thus result in different internal capital allocations.

I compare the equilibrium impact of arbitrageurs under two structures of the arbitrage industry, given a specific financial constraint. This exercise makes sense, given that the type of constraint I consider is so widely used by practitioners and regulators and has been microfounded in a static setting (see, e.g., Bruche and Kuong, 2019). However, a related and important exercise would consist in deriving the optimal constraint from first principles for each type of structure in a dynamic setting. These extensions are left for future research.



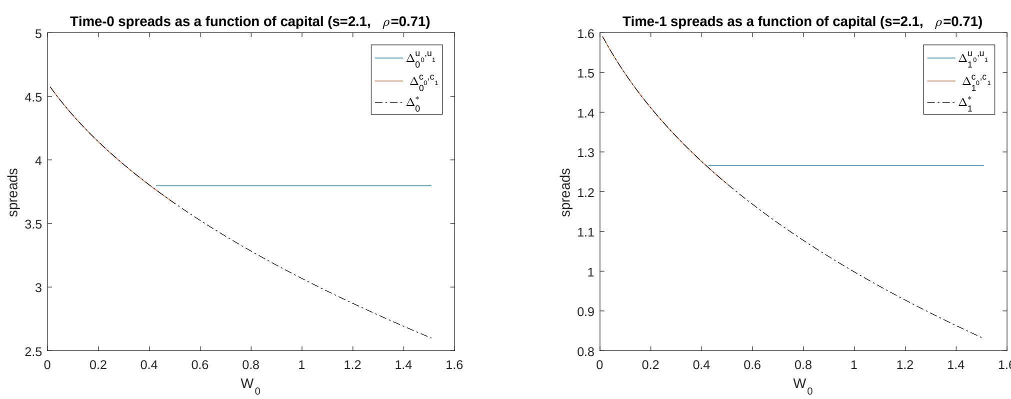
(a) Time 0 (b) Time 1

Figure 5: Spreads as a function of arbitrage capital for $\rho > 1$. The parameters are $a = \bar{e} = \sigma = 1$ and $s = 1.1$.



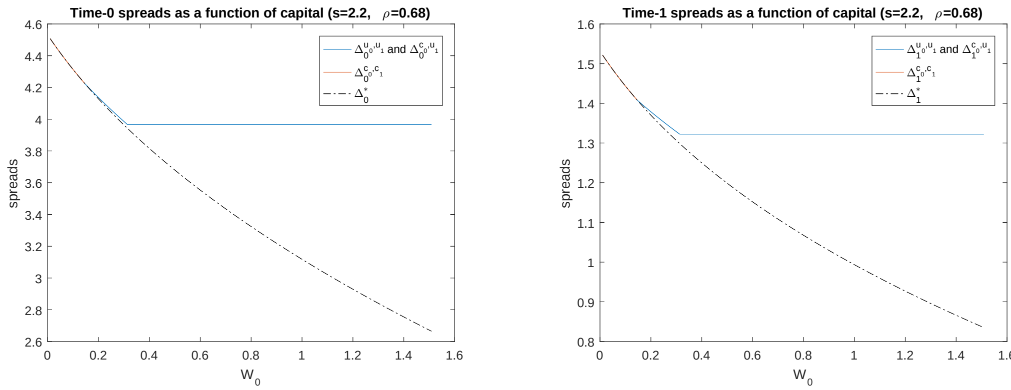
(a) Time 0 (b) Time 1

Figure 6: Spreads as a function of arbitrage capital for $\frac{3}{4} \leq \rho < 1$. The parameters are $a = \bar{e} = \sigma = 1$ and $s = 1.76$.



(a) Time 0 (b) Time 1

Figure 7: Spreads as a function of arbitrage capital for $\frac{7}{10} \leq \rho < \frac{3}{4}$. The parameters are $a = \bar{e} = \sigma = 1$ and $s = 2.1$.



(a) Time 0 (b) Time 1

Figure 8: Spreads as a function of arbitrage capital for $\rho \leq \frac{7}{10}$. The parameters are $a = \bar{e} = \sigma = 1$ and $s = 2.2$.

Appendix

This Appendix contains the main proofs. Additional proofs and the oligopoly case are relegated to the Internet Appendix.

A Benchmarks

A.1 Competitive benchmark

Lemma 2 (Hedgers' Demand and Certainty Equivalent) *At time $t = 0, 1$, in market A hedgers' demand and certainty equivalent are*

$$
\begin{aligned}
Y_t &= \frac{\mathbb{E}(p_{t+1}) - p_t}{a\sigma^2} - s \\
U_t &= w_t + \frac{\sum_{\tau=t}^{T} \left(\mathbb{E}_t(p_{\tau+1}) - p_\tau\right)^2}{2a\sigma^2} - s\left(\sum_{\tau=t}^{T} \left(\mathbb{E}_t(p_{\tau+1}) - p_\tau\right)\right)
\end{aligned}
\tag{7}
$$

Proof. See Internet Appendix. ■

Corollary 4 (Competitive Spreads and Positions in the Constrained Region) *Suppose* *that* $0 \le W_0 < \omega^*$. *Then arbitrageurs' positions in market A and equilibrium spreads are*

$$\begin{aligned} \bar{x}_0 &= \frac{a\sigma^2 s - \bar{e} + \sqrt{d_0^*}}{2a\sigma^2}, \quad \bar{X}_1 = \frac{a\sigma^2 s - \bar{e} + \sqrt{d_1^*}}{2a\sigma^2} \\ \Delta_0^* &= 2\left(a\sigma^2 s + \bar{e}\right) - \sqrt{d_0^*} - \sqrt{d_1^*}; \quad \Delta_1^* = a\sigma^2 s + \bar{e} - \sqrt{d_1^*}, \end{aligned} \tag{8}$$

with $d_0^* = (\bar{e} - a\sigma^2 s)^2 + 2a\sigma^2 W_0$ *and* $d_1^* = (\bar{e} - a\sigma^2 s)^2 + 4a\sigma^2 x_0 \bar{e}$.

Proof. The positions $\bar{x}_0$ and $\bar{X}_1$ are solutions to the system of equations (Gromb and Vayanos, 2002)

$$\bar{x}_0 - \bar{x}_0 \frac{a\sigma^2 (s - \bar{x}_0)}{\bar{e}} = \frac{W_0}{2\bar{e}} \tag{9}$$

$$\bar{X}_1 - \bar{X}_1 \frac{a\sigma^2 s - \bar{X}_1}{\bar{e}} = \bar{x}_0 \tag{10}$$

To obtain equilibrium spreads, subtitute for hedgers' demand Y_t^k and plug these quantities into the market clearing equation

$$Y_t^k + \sum_{i=1}^{n} X_t^{i,k} = 0 \tag{11}$$

■

A.2 Monopoly without Financial Constraints

Price schedules (Lemma 1)

Proof. The result follows from inverting hedgers' demand given in Lemma 2 and imposing market clearing (11). ■

Lemma 1 implies that the spread schedules $\Delta_t(\cdot) \equiv p_t^B(\cdot) - p_t^A(\cdot)$ are given by

$$\Delta_1(X_1) = 2a\sigma^2(s - X_1) = 2a\sigma^2(s - x_0 - x_1) \tag{12}$$

$$\Delta_0(x_0, x_1) = 2a\sigma^2(s - x_0) + \Delta_1(x_0, x_1) \tag{13}$$

B Static Equilibrium ($t = 1$)

At time 1, the arbitrageur solves the following problem:

$$\max_{x_1} \; B_0 + x_1 \Delta_1(X_1) \qquad (14)$$
$$\text{s.t. } f_1^+(X_1)\; \mathbb{1}_{X_1 \geq 0} + f_1^-(X_1)\; \mathbb{1}_{X_1 < 0} \geq 0$$

where B_0 is the position in the risk-free asset, $\Delta_1(X_1)$ is given by (12), and f^+ and f^- are given in the text. The financial constraints define a set $\mathcal{F}_1 = \{X_1 \geq 0 \mid f_1^+(X_1) \geq 0\} \cup \{X_1 < 0 \mid f_1^-(X_1) \geq 0\}$.

Notation 2 (Boundaries of $\mathcal{F}_1$)

- *Let* $\underline{X}_1^+ \equiv \frac{a\sigma^2 s - \bar{e} - \sqrt{d_1^+}}{2a\sigma^2}$ *and* $\bar{X}_1 \equiv \frac{a\sigma^2 s - \bar{e} + \sqrt{d_1^+}}{2a\sigma^2}$ *denote the smallest and largest roots, if they exist, of* f_1^+*, with* $d_1^+ \equiv 2a\sigma^2 W_1 + (a\sigma^2 s - \bar{e})^2$.
- *Let* $\underline{X}_1 \equiv \frac{a\sigma^2 s + \bar{e} - \sqrt{d_1^-}}{2a\sigma^2}$ *denote the smallest root of* f_1^-*, with* $d_1^- \equiv 2a\sigma^2 W_1 + (a\sigma^2 s + \bar{e})^2$.

Notation 3 (Preferred position at $t = 1$) *Let* $X_1^u \equiv x_0 + x_1^u = \frac{s + x_0}{2}$ *denote the arbitrageur's preferred position at time 1, where* $x_1^u \equiv \frac{s - x_0}{2}$ *solves* (14) *without constraints.*

Proposition 12 (Static Equilibrium) *Suppose that the arbitargeur starts with wealth* W_1 *and position* x_0*. There exists a wealth threshold* $\bar{W}_1^+ < 0$ *such that*

1. *If* $W_1 < \bar{W}_1^+$*, or if* $\bar{W}_1^+ \leq W_1 < 0$ *and* $\rho > 1$*, the arbitrageur has not enough capital to hold any position at time 1.*

2. *If* $\bar{W}_1^+ \leq W_1 < 0$ *and* $\rho \leq 1$*, the arbitrageur can hold only long positions in* $\mathcal{F}_1 = \left[\underline{X}_1^+, \bar{X}_1\right]$*. The optimum depends on the initial wealth and the initial position in the asset* x_0*:*

    - *If* $x_0 < -s$*, the arbitrageur's preferred position is a short one, thus his constraint binds downwards* ($X_1^u < 0 < \underline{X}_1^+ < \bar{X}_1$)*. It is optimal to hold* $\underline{X}_1^+$*.*
    - *If* $-s \leq x_0 < -s\rho$*, the constraint binds downwards* ($0 < X_1^u < \underline{X}_1^+ < \bar{X}_1$) *if* $W_1 < \frac{1}{2}a\sigma^2(x_0^2 - s^2) + \bar{e}(x_0 + s)$*. In this case, it is optimal for the arbitrageur to hold* $\underline{X}_1^+$*. Otherwise, the arbitrageur can hold his preferred position* X_1^u*.*

- *If $x_0 \geq -s\rho$, then the arbitrageur's constraint binds upwards ($0 < \underline{X}_1^+ < \bar{X}_1 < X_1^u$) if $W_1 < \frac{1}{2}a\sigma^2(x_0^2 - s^2) + \bar{e}(x_0 + s)$. In this case, it is optimal for the arbitrageur to hold $\bar{X}_1$. Otherwise, the arbitrageur can hold his preferred position X_1^u.*

3. *If $W_1 \geq 0$, then the arbitrageur can choose long and short positions in the segment $\mathcal{F}_1 = \left[\underline{X}_1, \bar{X}_1\right]$, with $\underline{X}_1 < 0$ and $\bar{X}_1 > 0$. The optimum depends on the initial wealth and the initial position in the asset x_0:*

    - *If $x_0 < -s$, then $X_1^u < 0$. If $W_1 \geq \frac{1}{2}a\sigma^2(x_0^2 - s^2) - \bar{e}(x_0 + s)$, the constraint is slack and the arbitrageur holds X_1^u. Otherwise, the arbitrageur's constraint binds downwards, and the arbitrageur chooses $\underline{X}_1$.*
    - *If $x_0 \geq -s$, then*
        - *If $\rho \geq 1$, then if $W_1 \geq \frac{1}{2}a\sigma^2(x_0^2 - s^2) + \bar{e}(x_0 + s) > 0$, the constraint is slack and the arbitrageur holds X_1^u. If $0 \leq W_1 < \frac{1}{2}a\sigma^2(x_0^2 - s^2) + \bar{e}(x_0 + s)$, the constraint binds upwards and the arbitrageur holds $\bar{X}_1$.*
        - *If $\rho < 1$, then*
            * *if $-s \leq x_0 < -s\rho$, the constraint is slack, the arbitrageur holds X_1^u.*
            * *if $x_0 \geq -s\rho$, the arbitrageur holds X_1^u if $W_1 \geq \frac{1}{2}a\sigma^2(x_0^2 - s^2) + \bar{e}(x_0 + s)$, and $\bar{X}_1$ otherwise.*

Proof. See Internet Appendix. ■

Corollary 5 (Subgame spreads) *Equilibrium spreads in the subgame are $\Delta_1^u = a\sigma^2(s - x_0)$, $\Delta_1^{\bar{c}} = a\sigma^2 s + \bar{e} - \sqrt{d_1^+}$, and $\Delta_1^{\underline{c}} = a\sigma^2 s - \bar{e} + \sqrt{d_1^-}$.*

C Equilibrium with Slack Constraint at Time 1

In this section, I conjecture that the arbitrageur holds an unconstrained position at time 1 and verify under which conditions it is optimal to do so. Here is the full version of Proposition 4:

Proposition 13 (Equilibria with slack time-1 constraint) *There exists three thresholds ω_0^u, ω_1^u, and ω^f, with $\omega^f \equiv \omega^c \, \mathbb{1}_{\rho<7} + \hat{\omega} \, \mathbb{1}_{\rho\geq 7}$, that define four regions in terms of initial arbitrage capital:*

1. *In the first region, $W_0 \geq \max(\omega_0^u, \omega_1^u)$, arbitrage capital is abundant, both constraints are slack in equilibrium, and the arbitrageur holds his desired position at time 0 and time 1 $x_0^{u_0,u_1}$ and $X_1^{u_0,u_1}$ given in Proposition 2 ($\mathbf{u_0}, \mathbf{u_1}$)*

2. *In second region, where $\max(\omega^f, \min(\omega_0^u, \omega_1^u)) \leq W_0 < \max(\omega_0^u, \omega_1^u)$, there are two cases:*

   - *If $\rho \geq \frac{7}{10}$, an equilibrium may exist, in which the arbitrageur's constraint binds at time 0 and is slack at time 1 ($\mathbf{c_0}, \mathbf{u_1}$).*
     Namely, if $\tilde{\Omega}_0^{u_1,u_1}(x_0^{c_0,u_1}) \geq \tilde{\Omega}_0^{u_1,\bar{c}_1}(x_0^{d*})$, where $x_0^{d*} = \arg\max_{x_0 \in \mathcal{D}_0^{\bar{c}_1}} \tilde{\Omega}_0^{u_1,\bar{c}_1}(x_0)$, the arbitrageur holds less than his desired position $x_0^{u_0,u_1}$ at time 0 to keep the constraint slack at time 1: $x_0^{c_0,u_1} = \min(\hat{x}_0, \bar{x}_0) < x_0^{u_0,u_1}$, and $X_1^{c_0,u_1} = X_1^u(x_0^{c_0,u_1}) = \frac{s+x_0^{c_0,u_1}}{2}$. Otherwise, there is no equilibrium with a slack constraint at time 1.
   - *If $\rho < \frac{7}{10}$, an equilibrium exists, where the constraint binds at time 0 but not at time 1 ($\mathbf{c_0}, \mathbf{u_1}$), with $x_0^{c_0,u_1} = \bar{x}_0 < x_0^{u_0,u_1}$, and $X_1^{c_0,u_1} = X_1^u(x_0^{c_0,u_1}) = \frac{s+\bar{x}_0}{2}$.*

3. *In the third region, where $\max(0, \omega^f) \leq W_0 < \min(\omega_0^u, \omega_1^u)$, the situation is the same as in the second region with $\rho \geq \frac{7}{10}$. The interval $\left[\omega^f, \min(\omega_0^u, \omega_1^u)\right)$ is non-empty iff $\rho \in \left[0, 3 - \frac{2}{5}\sqrt{30}\right) \cup \left(3 - \frac{2}{5}\sqrt{30}, \infty\right)$.*

4. *In the fourth region, $0 \leq W_0 < \max(0, \omega^f)$, there is little arbitrage capital, and thus there is no equilibrium with a slack constraint at time 1 (**no** $\mathbf{u_1}$).*

Proof. The proof relies on three main steps: i) I first write the arbitrageur's objective function and payoffs from deviating, assuming that hedgers anticipate a slack constraint at time 1. The arbitrageur's maximization involves choosing a position satisfying a set of constraints. ii) I derive the sets of feasible positions and possible deviations. These sets depend on the initial level of arbitrage capital W_0 and the risk benefit ratio ρ. iii) I determine the candidate equilibrium strategy and verify conditions under which it is possible/ optimal for the arbitrageur to follow it.

C.1 Step 1: arbitrageur's problem

Objective functions. Using the notations introduced in the main text, given hedgers' anticipations, the arbitrageur's problem is to choose x_0 (or equivalently an action leading to state $l = \{u_1, \bar{c}_1, \underline{c}_1\}$ at time 1) to maximize expected utility:

$$\Omega_0^{u_1} = \max\ (\Omega_0^{u_1,u_1}, \Omega_0^{u_1,\bar{c}_1}, \Omega_0^{u_1,\underline{c}_1})$$

Of course, x_0 must also satisfy the financial constraint at $t = 0$. The value functions $\Omega_0^{u_1,l}$ associated to each action are defined as follows:

$$\begin{aligned}
\Omega_0^{u_1,u_1} = \max_{x_0} \quad & \tilde{\Omega}_0^{u_1,u_1}(x_0) = W_0 + x_0 \Delta_0^{u_1}(x_0) + \Omega_1^u(x_0) \\
\text{s.t.} \quad & f_0^+(x_0) \geq 0 \ \mathbb{1}_{x_0 \geq 0} + f_0^-(x_0) \ \mathbb{1}_{x_0 < 0} \geq 0 \\
& f_1^+ \left(\frac{s + x_0}{2} \right) \ \mathbb{1}_{x_0 \geq -s} + f_1^- \left(\frac{s + x_0}{2} \right) \ \mathbb{1}_{x_0 < -s} \geq 0 \\
& W_1(x_0) = W_0 + 2a\sigma^2 x_0 (s - x_0) \geq 0
\end{aligned}$$

The objective function $\tilde{\Omega}_0^{u_1,u_1}$ relies on two premises: i) the continuation value $\Omega_1^u(x_0) \equiv \frac{1}{2} a\sigma^2 (s - x_0)^2$ assumes that the arbitrageur chooses his preferred position at time 1, and ii) the spread schedule, based on the price schedule in each market, requires that hedgers correctly anticipate that the arbitrageur's time 1 constraint is slack in equilibrium. Given equation 13, the spread schedule is $\Delta_0^{u_1}(x_0) = \mathbb{E}_0 \left[\Delta_1^u(x_0) \right] + 2a\sigma^2 (s - x_0) = 3a\sigma^2 (s - x_0)$, since $\Delta_1^u(x_0) = 2a\sigma^2 (s - X_1^u) = a\sigma^2 (s - x_0)$.

The first constraint ensures that the arbitrageur has enough capital to hold a position x_0 at time 0. The next constraint ensures that given the position established at time 0, x_0, the arbitrageur can indeed hold his preferred position $X_1^u(x_0)$ at time 1, be it a long or a short position. This requirement ensures that the arbitrageur's strategy is time-consistent. The arbitrageur's ability to satisfy the time-1 constraints requires positive wealth at time 1, which yields the last constraint. Otherwise, Proposition 12 shows that the arbitrageur's constraint is necessarily binding at time 1. Next, I consider the payoff from deviating towards an upward-binding constraint at time 1.

$$\begin{aligned}
\Omega_0^{u_1,\bar{c}_1} = \max_{x_0} \quad & \tilde{\Omega}_0^{u_1,\bar{c}_1}(x_0) = W_0 + x_0 \Delta_0(x_0)^{u_1} + \Omega_1^{\bar{c}}(x_0) \\
\text{s.t.} \quad & f_0^+(x_0) \geq 0 \ \mathbb{1}_{x_0 \geq 0} + f_0^-(x_0) \ \mathbb{1}_{x_0 < 0} \geq 0 \\
& f_1^+ \left(\frac{s + x_0}{2} \right) \ \mathbb{1}_{x_0 \geq -s} + f_1^- \left(\frac{s + x_0}{2} \right) \ \mathbb{1}_{x_0 < -s} < 0 \\
& W_1(x_0) = W_0 + 2a\sigma^2 x_0 (s - x_0) \geq 0
\end{aligned}$$

The objective function includes a different continuation value at time 1. The first constraint ensures that the position is feasible at time 0. The next constraint ensures that, given x_0, the arbitrageur can indeed *not* choose his preferred position at time 1 (time consistency). The last constraint requires that wealth be positive at time 1, as Proposition 12 requires.

Finally, I define the payoff from deviating towards a downward-binding constraint at time 1:

$$\begin{aligned}
\Omega_0^{u_1,\underline{c}_1} = \max_{x_0 \in [-s,-s\rho[\cup]-\infty,-s]} \quad & \tilde{\Omega}_0^{u_1,\underline{c}_1}(x_0) = W_0 + x_0 \Delta_0^{u_1}(x_0) + \Omega_1^{\underline{c}}(x_0) \\
\text{s.t.} \quad & f_0^+(x_0) \geq 0 \; \mathbb{1}_{x_0 \geq 0} + f_0^-(x_0) \; \mathbb{1}_{x_0 < 0} \geq 0 \\
& f_1^+\left(\frac{s+x_0}{2}\right) \; \mathbb{1}_{x_0 \geq -s} + f_1^-\left(\frac{s+x_0}{2}\right) \; \mathbb{1}_{x_0 < -s} < 0 \\
& W_1(x_0) < 0 \quad \text{if } x_0 \in [-s, -s\rho[, \text{ or,} \\
& 0 \leq W_1(x_0) < \frac{1}{2} a\sigma^2 (x_0^2 - s^2) - \bar{e}(x_0 + s) \quad \text{if } x_0 < -s
\end{aligned}$$

The payoff is built as in the previous case. However, the last constraint requires negative wealth and x_0 must be chosen in the interval $[-s, -s\rho[$, as this is necessary for the constraint to bind downwards at time 1, by Proposition 12.

Feasible positions. Suppose first that the arbitrageur chooses x_0 leading u_1. This position must satisfy the following set of constraints. First, the position must satisfy the constraint at time 0, so $x_0 \in \mathcal{F}_0^0$, where $\mathcal{F}_0^0 = \{x_0 < 0 \mid f_0^-(x_0) \geq 0\} \cup \{x_0 \geq 0 \mid f_0^+(x_0) \geq 0\}$. Second, it must be such that at time 1, the arbitrageur can hold his preferred position. I denote $\mathcal{F}_0^1$ the interval determined by the constraints at time 1. It is convenient to write $\mathcal{F}_0^1$ as the union of two intervals, one for long and one for short unconstrained positions at time 1, i.e. $\mathcal{F}_0^1 = \mathcal{F}_0^{1-} \cup \mathcal{F}_0^{1+}$, where $\mathcal{F}_0^{1-} = \{x_0 < -s \mid f_1^-\left(\frac{s+x_0}{2}\right) \geq 0\}$ and $\mathcal{F}_0^{1+} = \{x_0 \geq -s \mid f_1^+\left(\frac{s+x_0}{2}\right) \geq 0\}$. Finally, the positive wealth constraint defines a set $\mathcal{F}_0^{pw} = \{x_0 \mid W_1(x_0) = W_0 + 2a\sigma^2 x_0(s - x_0) \geq 0\}$. The intersection of these sets thus defines a set of feasible positions

$$\mathcal{F}_0^u = \mathcal{F}_0^0 \cap \mathcal{F}_0^{u_1} \cap \mathcal{F}_0^{pw}$$

Therefore we can rewrite $\Omega_0^{u_1,u_1}$ simply as

$$\Omega_0^{u_1,u_1} = \max_{x_0 \in \mathcal{F}_0^u} \quad W_0 + x_0 \Delta_0(x_0) + \Omega_1^{u_1}(x_0)$$

Similarly, we can define the sets $\mathcal{D}_0^{\bar{c}_1}$ and $\mathcal{D}_0^{\underline{c}_1}$ of positions leading to upward -or -downward-binding constraints at time 1. I derive these sets in detail in Section C.2.2 below. Using these notations, we can also rewrite $\Omega_0^{u_1,\bar{c}_1}$ and $\Omega_0^{u_1,\underline{c}_1}$ as follows:

$$\begin{aligned}
\Omega_0^{u_1,\bar{c}_1} &= \max_{x_0 \in \mathcal{D}_0^{\bar{c}_1}} \quad \tilde{\Omega}_0^{u_1,\bar{c}_1}(x_0) = W_0 + x_0\Delta_0(x_0) + \Omega_1^{\bar{c}}(x_0) \\
\Omega_0^{u_1,\underline{c}_1} &= \max_{x_0 \in \mathcal{D}_0^{\underline{c}_1}} \quad \tilde{\Omega}_0^{u_1,\underline{c}_1}(x_0) = W_0 + x_0\Delta_0(x_0) + \Omega_1^{\underline{c}}(x_0)
\end{aligned}$$

C.2 Step 2: Feasible positions and possible deviations at $t = 0$

Definition 2 (Time-0 Boundary Positions)

- *Let $\underline{x}_0$ denote the smallest root of $f_0^-(x_0) = 0$ and $\bar{x}_0$ the largest root of $f_0^+(x_0) = 0$, if they exist.*
- *Let $\hat{x}_0$ and $\hat{x}_0^-$ denote the largest and smallest roots, if they exist, of $f_1^+\left(\frac{s+x_0}{2}\right) = 0$, for all $x_0 \geq -s$.*
- *Let $\underline{\underline{x}}_0$ denote the smallest root of $f_1^-\left(\frac{s+x_0}{2}\right) = 0$, for $x_0 < -s$.*
- *Let $\underline{z}_0$ and $\bar{z}_0$ denote the smallest and largest roots of $W_1 = W_0 + 2a\sigma^2 x_0(s - x_0) = 0$.*

C.2.1 Feasible positions

Note: I use the convention that if $a > b$, then $[a, b] = \emptyset$.

Proposition 14 (Interval $\mathcal{F}_0^u$)

- *If $0 \leq \rho < 1$ or if $1 \leq \rho < 7$, then* $\mathcal{F}_0^u = \begin{cases} \emptyset & \text{if } W_0 < \omega^c \\ \left[\max(\hat{x}_0^-, \underline{x}_0), \min(\hat{x}_0, \bar{x}_0)\right] & \text{if } \omega^c \leq W_0 < \hat{\omega} \\ \left[\underline{x}_0, \min(\hat{x}_0, \bar{x}_0)\right] & \text{if } \hat{\omega} \leq W_0 < \hat{\omega} + \omega^* \\ \left[\max(\hat{x}_0^-, \underline{x}_0), \min(\hat{x}_0, \bar{x}_0)\right] & \text{if } W_0 \geq \hat{\omega} + \omega^* \end{cases}$
- *If $\rho \geq 7$, then* $\mathcal{F}_0^u = \begin{cases} \emptyset & \text{if } W_0 < \omega^c \\ \emptyset & \text{if } \omega^c \leq W_0 < \hat{\omega} \\ \left[\underline{x}_0, \min(\hat{x}_0, \bar{x}_0)\right] & \text{if } \hat{\omega} \leq W_0 < \hat{\omega} + \omega^* \\ \left[\max(\hat{x}_0^-, \underline{x}_0), \min(\hat{x}_0, \bar{x}_0)\right] & \text{if } W_0 \geq \hat{\omega} + \omega^* \end{cases}$

Proof. See Internet Appendix. ■

C.2.2 Possible deviations

Lemma 3 *Suppose hedgers anticipate a slack constraint at time 1. Deviations from the arbitrageur leading to a downward-binding constraint at time 1 are either not feasible or dominated at time 0.*

Lemma 4 *Suppose hedgers anticipate a slack constraint at time 1. Deviations from the arbitrageur leading to an upward-binding constraint at time 1 must belong to the set* $\mathcal{D}_0^{\bar{c}_1}$*, given by*

- *If* $\rho < 1$, $\mathcal{D}_0^{\bar{c}_1} = \begin{cases} [\underline{x}_0, \bar{x}_0] & \textit{if } W_0 < \omega^c \\ [\underline{x}_0, \max(\hat{x}_0, \bar{x}_0)) \cup (\min(\hat{x}_0, \bar{x}_0), \bar{x}_0] & \textit{if } \omega^c \leq W_0 < \hat{\omega} \\ (\min(\hat{x}_0, \bar{x}_0), \bar{x}_0] & \textit{if } \hat{\omega} \leq W_0 < \hat{\omega} + \frac{4\bar{e}^2}{a\sigma^2} \\ \left[\max(\underline{x}_0, -s\rho), \max(\underline{x}_0, -s\rho, \hat{x}_0^-)\right) \cup (\min(\hat{x}_0, \bar{x}_0), \bar{x}_0] & \textit{if } W_0 \geq \hat{\omega} + \frac{4\bar{e}^2}{a\sigma^2} \end{cases}$

- *If* $1 \leq \rho < 7$, $\mathcal{D}_0^{\bar{c}_1} = \begin{cases} [\underline{x}_0, \bar{x}_0] & \textit{if } W_0 < \omega^c \\ [\underline{x}_0, \max(\hat{x}_0, \bar{x}_0)) \cup (\min(\hat{x}_0, \bar{x}_0), \bar{x}_0] & \textit{if } \omega^c \leq W_0 < \hat{\omega} \\ (\min(\hat{x}_0, \bar{x}_0), \bar{x}_0] & \textit{if } \hat{\omega} \leq W_0 < \hat{\omega} + \omega^* \\ \left[-s, \max(-s, \hat{x}_0^-)\right) \cup (\min(\hat{x}_0, \bar{x}_0), \bar{x}_0] & \textit{if } W_0 \geq \hat{\omega} + \omega^* \end{cases}$

- *If* $\rho \geq 7$, $\mathcal{D}_0^{\bar{c}_1} = \begin{cases} [\underline{x}_0, \bar{x}_0] & \textit{if } W_0 < \omega^c \textit{ or } \omega^c \leq W_0 < \hat{\omega} \\ (\min(\hat{x}_0, \bar{x}_0), \bar{x}_0] & \textit{if } \hat{\omega} \leq W_0 < \hat{\omega} + \omega^* \\ \left[-s, \max(-s, \hat{x}_0^-)\right) \cup (\min(\hat{x}_0, \bar{x}_0), \bar{x}_0] & \textit{if } W_0 \geq \hat{\omega} + \omega^* \end{cases}$

The proofs of these Lemmata are in Internet Appendix.

C.3 Step 3: Equilibrium determination

C.3.1 Value functions and candidate equilibrium strategy

Given the results in Proposition 12 and Corollary 5, we can define the objective functions as follows:

$$\begin{aligned}
\tilde{\Omega}_0^{u_1,u_1}(x_0) =& W_0 + x_0\Delta_0^{u_1}(x_0) + \Omega_1^{u}(x_0) = W_0 + 3a\sigma^2 x_0(s - x_0) + \frac{a\sigma^2}{2}(s - x_0)^2 \\
\tilde{\Omega}_0^{u_1,\bar{c}_1}(x_0) =& W_0 + x_0\Delta_0^{u_1}(x_0) + \Omega_1^{\bar{c}}(x_0) = W_0 + x_0\Delta_0^{u_1}(x_0) + \bar{x}_1(x_0)\Delta_1^{\bar{c}}(x_0) \\
\tilde{\Omega}_0^{u_1,\underline{c}_1}(x_0) =& W_0 + x_0\Delta_0^{u_1}(x_0) + \Omega_1^{\underline{c}}(x_0) = W_0 + x_0\Delta_0^{u_1}(x_0) + \underline{x}_1(x_0)\Delta_1^{\underline{c}}(x_0)
\end{aligned}$$

We can now determine under which conditions the arbirtrageur's preferred position satisfies the constraints at time 0 and time 1.

Proposition 15 (Candidate equilibrium strategy with slack time-0 and -1 constraints ($\mathbf{u_0}, \mathbf{u}$ *Let ω_0^u and ω_1^u denote two wealth thresholds, with $\omega_0^u = \frac{4}{5}s\bar{e} - \frac{12}{25}a\sigma^2 s^2$ and $\omega_1^u = \frac{7}{5}s\bar{e} - \frac{9}{10}a\sigma^2 s^2$, and let $\omega^u = \max(\omega_0^u, \omega_1^u)$.*

1. *If $W_0 \geq \omega^u$, then the arbitrageur can hold his preferred positions at time 0 and time 1 $x_0^{u_0,u_1} = \frac{2}{5}s$ and $X_1^{u_0,u_1} = \frac{7}{10}s$*

2. *The arbitrageur's expected utility is denoted $\Omega_0^{u_0,u_1} = \tilde{\Omega}_0^{u_1,u_1}(x_0^{u_0,u_1})$.*

3. *For any x_0 such that $\tilde{\Omega}_1^{\bar{c}}(x_0)$ exists, $\tilde{\Omega}_0^{u_1,\bar{c}_1}(x_0) \leq \tilde{\Omega}_0^{u_1,u_1}(x_0)$.*

Proof. The arbitrageur's objective function $\tilde{\Omega}_0^{u_1,u_1}$ admits a global maximum at $x_0^{u_0,u_1}$ given in the Proposition. Substituting $x_0^{u_0,u_1}$ into X_1^u (Definition 2) gives $X_1^{u_0,u_1}$. Since $x_0^{u_0,u_1} > 0$, the relevant constraints are f_0^+ and f_1^+. Thus, to determine the thresholds ω_0^u, ω_1^u, it suffices to substitute $x_0^{u_0,u_1}$ into $f_0^+(x_0) \geq 0$ and $f_1^+\left(\frac{s+x_0}{2}\right) \geq 0$ and rearrange the terms. The last point follows from the fact that for any x_0 such that $\Omega_1^{\bar{c}}(x_0)$ exists, $\Omega_1^{\bar{c}}(x_0) \leq \Omega_1^u(x_0)$, and from the definition of $\tilde{\Omega}_0^{u_1,\bar{c}_1}$ and $\tilde{\Omega}_0^{u_1,u_1}$. ■

C.3.2 Capital and risk benefit thresholds

Wealth thresholds. Given our analysis so far, we must order the following wealth thresholds: ω_0^u, ω_1^u, ω^c, $\hat{\omega}$, $\hat{\omega} + \omega^*$ and $\hat{\omega} + \frac{4\bar{e}^2}{a\sigma^2}$. It is easy to see that $\hat{\omega} + \omega^*$ is larger than ω_0^u, ω_1^u, ω^c, and $\hat{\omega}$. Similarly, when $\rho < 1$, $\hat{\omega} + \frac{4\bar{e}^2}{a\sigma^2}$ is larger than ω_0^u, ω_1^u, ω^c, and $\hat{\omega}$.

Lemma 5 (Wealth Threshold Ordering in Equilibrium with Slack Time-1 Constraint) *The order of thresholds is given in Table 2.*

Proof. By direct calculation using threshold definition. ■

Releveant ρ thresholds. The relevant thresholds from Table 2 and previous results are $\frac{3}{5}$, $\frac{9}{14}$, $7-2\sqrt{10}$, $\frac{7}{10}$,3- $\frac{2}{5}\sqrt{30}$, 1, $\frac{7}{2}$, $3+\frac{2}{5}\sqrt{30}$, $\frac{28}{5}$, 7, $7+2\sqrt{10}$. If $\rho \geq 7-2\sqrt{10}$, wealth thresholds not necessarily positive but are always in the same order. Thus, for simplicity, I treat all the cases with $\rho < 7 - 2\sqrt{10}$ as one case. Similarly, I ignore the case $\rho > 7 + 2\sqrt{10}$, which determines the positivity of ω^c, but does not affect the order of the thresholds. However, I add 1 and 7, which do not affect the order of tresholds, but affect the set of feasible positions or deviations.

C.3.3 Equilibrium case by case

Equilibrium determination follows from combining the results above. The details are available in the Internet Appendix. ■

Table 2: Wealth Threshold Order for Equilibrium with Slack Time-1 Constraint

Treshold	Greater than	Condition	Value
$\omega_0^u \equiv \frac{4}{5}s\bar{e} - \frac{12}{25}a\sigma^2 s^2$	ω_1^u	$\rho < \frac{7}{10}$	
	ω^c	$\rho < 3 - \frac{2}{5}\sqrt{30}$ or $\rho > 3 + \frac{2}{5}\sqrt{30}$	0.809 — 5.19
	0	$\rho \geq \frac{3}{5}$	
	$\hat{\omega}$	$\rho > \frac{28}{5}$	
$\omega_1^u \equiv \frac{7}{5}s\bar{e} - \frac{9}{10}a\sigma^2 s^2$	ω^c	for all $\rho > 0$	
	0	$\rho > \frac{9}{14}$	0.64
	$\hat{\omega}$	$\rho > \frac{7}{2}$	
$\omega^c \equiv \omega_1^u - \frac{\bar{e}^2}{10a\sigma^2}$	0	$7 - 2\sqrt{10} \leq \rho \leq 7 + 2\sqrt{10}$	0.675—13.32
	$\hat{\omega}$	never, equality for $\rho = 7$	
$\hat{\omega} \equiv 4a\sigma^2 s^2$	0	for any $\rho > 0$	

Table 3: ρ and Wealth Intervals for Equilibrium with Slack Time-1 Constraint

Case	ρ interval	Wealth ordering
1	$0 \leq \rho < \frac{7}{10}$	$0 < (\omega^c, 0)^+ < (\omega_1^u, 0)^+ < (\omega_0^u, 0)^+ < \hat{\omega}$
2	$\frac{7}{10} \leq \rho < 3 - \frac{2}{5}\sqrt{30}$	$0 < \omega^c < \omega_0^u < \omega_1^u < \hat{\omega}$
3	$3 - \frac{2}{5}\sqrt{30} \leq \rho < 1$	$0 < \omega_0^u < \omega^c < \omega_1^u < \hat{\omega}$
4	$1 \leq \rho < \frac{7}{2}$	$0 < \omega_0^u < \omega^c < \omega_1^u < \hat{\omega}$
5	$\frac{7}{2} \leq \rho < 3 + \frac{2}{5}\sqrt{30}$	$\bar{0} < \omega_0^u < \omega^c < \hat{\omega} < \omega_1^u$
6	$3 + \frac{2}{5}\sqrt{30} \leq \rho < \frac{28}{5}$	$0 < \omega^c < \omega_0^u < \hat{\omega} < \omega_1^u$
7	$\frac{28}{5} \leq \rho < 7$	$0 < \omega^c < \hat{\omega} < \omega_0^u < \omega_1^u$
8	$7 \leq \rho$	$0 < \omega^c < \hat{\omega} < \omega_0^u < \omega_1^u$

C.4 Equilibrium spreads

Corollary 6 (Equilibrium spreads in the u_1 equilibria)
In the u_0, u_1 and c_0, u_1 equilibria, spreads are

$$\Delta_0^{u_0,u_1} = \frac{9}{5}a\sigma^2 s, \quad \textit{and} \quad \Delta_1^{u_0,u_1} = \frac{3}{5}a\sigma^2 s \tag{15}$$

$$\Delta_0^{c_0,u_1} = 3a\sigma^2(s - x_0^{c_0,u_1}), \quad \textit{and} \quad \Delta_1^{c_0,u_1} = 2a\sigma^2(s - X_1^u(x_0^{c_0,u_1})) = a\sigma^2(s - x_0^{c_0,u_1}) \tag{16}$$

Proof. Equilibrium spreads follow from substituting the equilibrium quantity (either $x_0^{u_0,u_1}$ or $x_0^{c_0,u_1}$) into the spreads schedule (13)-(12). ■

D Equilibrium with Binding Time-1 Constraint

Here is the full result:

Proposition 16 (Equilibria with binding time-1 constraint)

- *There are no equilibria in which the arbitrageur's constraint binds downwards at time 1.*
- *There are equilibria in which the arbitrageur's constraint binds upwards at time 1, as follows. Let $\omega_0^p \equiv s\bar{e} - \frac{1}{2}a\sigma^2 s^2$ and $\omega_1^p \equiv \frac{3}{2}s\bar{e} - \frac{7}{8}a\sigma^2 s^2$ denote two thresholds.*

  1. *If $0 \leq \rho < \frac{3}{4}$, then $\omega_1^p < \omega_0^p$, and there are three regions in terms of arbitrage capital:*
     (a) *In the first region, with $0 \leq W_0 < \max(0, \omega^f)$, the arbitrageur's constraint binds upwards at time 0 and time 1 in equilibrium ($\mathbf{c_0}, \mathbf{c_1}$ equilibrium). This equilibrium is the same as in the constrained competitive case, for a given level of capital. The arbitrageur holds $x_0^{c_0,c_1} = \bar{x}_0$, and $X_1^{c_0,c_1} = \bar{X}_1(\bar{x}_0)$.*
     (b) *In the second region, with $\max(0, \omega^f) \leq W_0 < \max(0, \omega_1^p)$, there are two cases*
        i. *If $\max(\hat{x}_0, \bar{x}_0) = \hat{x}_0$, there is no equilibrium in which the arbitrageur's constraint binds upwards at time 1 (**no** $\mathbf{c_1}$).*
        ii. *Otherwise, both constraints bind in equilibrium as in (a) iff $\Omega_0^{c_0,c_1} \geq \Omega_0^{\bar{c}_1,u_1}(x_0^{d*})$, where $x_0^{d*} = \arg\max_{x_0 \in \mathcal{D}_0^u} \Omega_0^{\bar{c}_1,u_1}(x_0)$.*
     (c) *In the third region, with $\max(0, \omega_1^p) \leq W_0 < \omega_0^p$ or $\omega_0^p \leq W_0$, there is no equilibrium in which the arbitrageur's constraint binds upwards at time 1 (**no** $\mathbf{c_1}$).*
  2. *If $\rho \geq \frac{3}{4}$, then $\omega_1^p > \omega_0^p$, and there are four regions in terms of arbitrage capital:*
     (a) *In the first region, with $0 \leq W_0 < \omega^f$, the equilibrium is $\mathbf{c_0}, \mathbf{c_1}$, as in case 1a.*
     (b) *In the second region, with $\omega^f \leq W_0 < \omega_0^p$, the equilibrium is the same as in 1b.*
     (c) *In the third region, with $\omega_0^p \leq W_0 < \omega_1^p$, there is an equilibrium in which the abritrageur's constraint binds upwards at time 1 and is slack at time 0 ($\mathbf{u_0}, \mathbf{c_1}$ equilibrium) iff $\Omega_0^{u_0,c_1} \geq \Omega_0^{\bar{c}_1,u_1}(x_0^{d*})$, where $x_0^{d*} = \arg\max_{x_0 \in \mathcal{D}_0^u} \Omega_0^{\bar{c}_1,u_1}(x_0)$.*
     (d) *In the fourth region, with $\omega_1^p \leq W_0$, is no equilibrium in which the arbitrageur's constraint binds upwards at time 1 (**no** $\mathbf{c_1}$), as in 1c.*

The proof is based on the same three steps as in the previous case.

D.1 Arbitrageur's problem

Suppose hedgers anticipate an upward-binding constraint at time 1. Let $\Omega_0^{\bar{c}_1,l}$ denote the arbitrageur's expected utility when hedgers anticipate an upward-binding constraint, and the arbitrageur chooses a trade x_0 subject to the $t = 0$ constraint, leading to state $l \in \{\bar{c}_1, u_1, \underline{c}_1\}$ at time 1, i.e. an upward-binding, slack, or downward-binding constraint at time 1. The maximization problem of the arbitrageur is thus as follows.

$$\Omega_0^{\bar{c}_1} = \max \ (\Omega_0^{\bar{c}_1,\bar{c}_1}, \Omega_0^{\bar{c}_1,\underline{c}_1}, \Omega_0^{\bar{c}_1,\bar{u}_1}) \tag{17}$$

The expected utilities associated with state l are defined as follows:

$$\begin{aligned}
\Omega_0^{\bar{c}_1,\bar{c}_1} = \max_{x_0} \tilde{\Omega}_0^{\bar{c}_1,\bar{c}_1} &= W_0 + x_0 \Delta_0^{\bar{c}_1}(x_0) + \Omega_1^{\bar{c}}(x_0) \\
\text{s.t.} \quad & f_0^+(x_0)\ \mathbb{1}_{x_0 \geq 0} + f_0^-(x_0)\ \mathbb{1}_{x_0 < 0} \geq 0 \\
& f_1^+\left(\frac{s + x_0}{2}\right)\ \mathbb{1}_{x_0 \geq -s} + f_1^-\left(\frac{s + x_0}{2}\right)\ \mathbb{1}_{x_0 < -s} < 0 \\
& W_1 = W_0 + 2a\sigma^2 x_0 (s - x_0) \geq 0
\end{aligned}$$

where $\Omega_1^{\bar{c}}(x_0) = 2a\sigma^2 \bar{x}_1(x_0)(s - \bar{X}_1(x_0))$. The second constraint ensures that in equilibrium, the arbitrageur cannot hold his preferred position because his constraint binds upwards at time 1. The last constraint ensures that equilibrium wealth is positive at time 1, which is required by Proposition 12. Next, I consider the expected utlity from deviations leading to a slack constraint at time 1.

$$\begin{aligned}
\Omega_0^{\bar{c}_1,u_1} = \max_{x_0} \tilde{\Omega}_0^{\bar{c}_1,u_1} &= W_0 + x_0 \Delta_0^{\bar{c}_1}(x_0) + \Omega_1^{u}(x_0) \\
\text{s.t.} \quad & f_0^+(x_0)\ \mathbb{1}_{x_0 \geq 0} + f_0^-(x_0)\ \mathbb{1}_{x_0 < 0} \geq 0 \\
& f_1^+\left(\frac{s + x_0}{2}\right)\ \mathbb{1}_{x_0 \geq -s} + f_1^-\left(\frac{s + x_0}{2}\right)\ \mathbb{1}_{x_0 < -s} \geq 0 \\
& W_1 = W_0 + 2a\sigma^2 x_0 (s - x_0) \geq 0
\end{aligned}$$

The difference with the previous problem is that the constraints at time 1 are slack, leading to continuation value Ω_1^u at time 1. Finally, here is the expected utlity from deviations leading to a downward-binding constraint at time 1.

$$
\begin{aligned}
\Omega_0^{\bar{c}_1,u_1} = \max_{x_0 \in [-s,-s\rho[\cup]-\infty,-s]} \tilde{\Omega}_0^{\bar{c}_1,\underline{c}_1} &= W_0 + x_0 \Delta_0^{\bar{c}_1}(x_0) + \Omega_1^{\underline{c}}(x_0) \\
\text{s.t.} \quad & f_0^+(x_0)\ \mathbb{1}_{x_0 \geq 0} + f_0^-(x_0)\ \mathbb{1}_{x_0<0} \geq 0 \\
& f_1^+\left(\frac{s+x_0}{2}\right)\ \mathbb{1}_{x_0 \geq -s} + f_1^-\left(\frac{s+x_0}{2}\right)\ \mathbb{1}_{x_0<-s} < 0 \\
& W_1(x_0) < 0 \quad \text{if } x_0 \in [-s, -s\rho[, \text{ or,} \\
& 0 \leq W_1(x_0) < \frac{1}{2} a\sigma^2 (x_0^2 - s^2) - \bar{e}(x_0 + s) \quad \text{if } x_0 < -s
\end{aligned}
$$

D.2 Set of feasible positions and possible deviations at $t = 0$

Lemma 6 *There is no equilibrium with a downward-binding constraint at time 1.*

Lemma 7 *Let $\mathcal{F}_0^{\bar{c}_1}$ denote the set of feasible positions with upward-binding constraint at time 1, and $\mathcal{D}_0^{u_1}$ the set of deviations leading to a slack constraint at time 1. We have: $\mathcal{F}_0^{\bar{c}_1} = \mathcal{D}_0^{\bar{c}_1}$, and $\mathcal{D}_0^{u_1} = \mathcal{F}_0^{u}$.*

The proofs of these Lemmata are in Internet Appendix.

D.3 Equilibrium determination

D.3.1 Value functions and candidate equilibrium

Assume that hedgers anticipate an upward-binding constraint at time 1. We have ruled out equilibria with downward-binding constraints. Using Proposition **??** and Corollary 5, the value functions in the remaining cases ($l = \{\bar{c}_1, u_1,\}$) are re

$$
\begin{aligned}
\Omega_0^{\bar{c}_1,\bar{c}_1}(x_0) &= W_0 + x_0 \Delta_0^{\bar{c}_1}(x_0) + \Omega_1^{\bar{c}}(x_0) = W_0 + x_0 \Delta_0^{\bar{c}_1}(x_0) + \bar{x}_1(x_0) \Delta_1^{\bar{c}}(x_0) \\
\Omega_0^{\bar{c}_1,u_1}(x_0) &= W_0 + x_0 \Delta_0^{\bar{c}_1}(x_0) + \Omega_1^{u}(x_0) = W_0 + x_0 \Delta_0^{\bar{c}_1}(x_0) + x_1^u(x_0) \Delta_1^u(x_0)
\end{aligned}
$$

Proposition 17 (Candidate equilibrium strategy - slack time-0 and binding time-1 constrai *Let $\omega_0^p \equiv s\bar{e} - \frac{1}{2} a\sigma^2 s^2$ and $\omega_1^p \equiv \frac{3}{2} s\bar{e} - \frac{7}{8} a\sigma^2 s^2$ denote two wealth thresholds.*

1. *The function $\tilde{\Omega}_0^{\bar{c}_1,\bar{c}_1}$ admits a maximum $x_0^{u_0,c_1}$ iff $W_0 \in [\omega_0^p, \omega_1^p)$, where $x_0^{u_0,c_1} = \frac{s}{2}$, and $X_1^{u_0,c_1} = \bar{X}_1(x_0^{u_0,c_1})$.*

2. *The interval $[\omega_0^p, \omega_1^p)$ is non-empty iff $\rho > \frac{3}{4}$.*

3. At this candidate equilibrium strategy, the arbitrageur's utility is

$$\Omega_0^{u_0,c_1} \equiv \tilde{\Omega}_0^{\bar{c}_1,\bar{c}_1}(x_0^{u_0,c_1}) = \frac{\bar{e}}{a\sigma^2}\left[a\sigma^2 s - \bar{e} + \sqrt{d_1^+(x_0^{u_0,c_1})}\right],$$

$$\text{where } d_1^+(x_0^{u_0,c_1}) = 2a\sigma^2 W_0 + 2a^2\sigma^4 s^2 + \bar{e}^2 - 2a\sigma^2 s\bar{e}$$

4. For any x_0 such that $\tilde{\Omega}_1^{\bar{c}}(x_0)$ exists, $\tilde{\Omega}_0^{\bar{c}_1,\bar{c}_1}(x_0) \leq \tilde{\Omega}_0^{\bar{c}_1,u_1}(x_0)$.

Proof. Let's first rewrite the objective function $\tilde{\Omega}_0^{\bar{c}_1,\bar{c}_1}$ by substituting for $\Delta_0^{\bar{c}_1}$ and $\Omega_1^{\bar{c}}$.

$$\begin{aligned}
W_0 + x_0\Delta_0^{\bar{c}_1}(x_0) + \Omega_1^{\bar{c}}(x_0) &= W_0 + 2a\sigma^2 x_0(s-x_0) + x_0\Delta_1^{\bar{c}}(x_0) + (\bar{X}_1(x_0) - x_0)\Delta_1^{\bar{c}}(x_0)\\
&= W_0 + 2a\sigma^2 x_0(s-x_0) + \bar{X}_1(x_0)\Delta_1^{\bar{c}}(x_0)\\
&= \underbrace{W_0 + 2a\sigma^2 x_0(s-x_0)}_{W_1(x_0)} + \frac{\left(a\sigma^2 s + \bar{e} - \sqrt{d_1^+(x_0)}\right)\left(a\sigma^2 s - \bar{e} + \sqrt{d_1^+(x_0)}\right)}{2a\sigma^2}
\end{aligned}$$

The last line follows from substituting for $\bar{X}_1$ and $\Delta_1^{\bar{c}}$. Then developing the numerator in the last term, substituting for d_1^+ and simplifying, we get:

$$W_0 + x_0\Delta_0^{\bar{c}_1}(x_0) + \Omega_1^{\bar{c}}(x_0) = \frac{\bar{e}}{a\sigma^2}\left[a\sigma^2 s - \bar{e} + \sqrt{d_1^+(x_0)}\right]$$

Therefore maximizing $\tilde{\Omega}_0^{\bar{c}_1,\bar{c}_1}$ is equivalent to maximizing d_1^+ subject to the constraint, which boils down to maximizing $W_1(x_0) = W_0 + 2a\sigma^2 x_0(s-x_0)$, subject to constraints. The solution is $x_0^{u_0,c_1} = \frac{s}{2}$ if $f_0^+(\frac{s}{2}) \geq 0$ and $f_1^+(\frac{3s}{4}) < 0$. The first condition requires $W_0 \geq \omega_0^p$, and the second $W_0 < \omega_1^p$.

These conditions define a non-empty interval iff $\omega_0^p < \omega_1^p$, which is equivalent to $\rho > \frac{3}{4}$.

Substituting $x_0^{u_0,c_1}$ into W_1 yields the equilibrium utility $\Omega_0^{u_0,c_1}$.

Finally, since for any x_0 such that $\Omega_1^{\bar{c}}(x_0)$ exists, $\tilde{\Omega}_1^{\bar{c}}(x_0) \leq \tilde{\Omega}_1^{u}(x_0)$, we also have, by definition of the value functions $\Omega_0^{\bar{c}_1,\bar{c}_1}$ and $\Omega_0^{\bar{c}_1,u_1}$, $\tilde{\Omega}_0^{\bar{c}_1,\bar{c}_1}(x_0) \leq \tilde{\Omega}_0^{\bar{c}_1,u_1}(x_0)$. ■

Corollary 7 *If $\rho > \frac{3}{4}$ and $W_0 < \omega_0^p$, or if $\rho \leq \frac{3}{4}$ and $W_0 < \min(\omega_0^p, \omega_1^p)$, the candidate equilibrium strategy is $x_0^{c_0,c_1} = \bar{x}_0$, and $X_1^{c_0,c_1} = \bar{X}_1(\bar{x}_0)$. At this strategy, the arbitrageur's expected utility is $\Omega_0^{c_0,c_1} \equiv \tilde{\Omega}_0^{\bar{c}_1,\bar{c}_1}(\bar{x}_0)$.*

Proof. Follows immediately from Proposition 17. ■

Note that this strategy is the same as the constrained equilibrium of the competitive case (for a given W_0), but with different wealth thresholds.

D.3.2 Relevant thresholds

Wealth thresholds. Given our analysis of feasible positions and deviations, the relevant wealth thresholds are ω^c, ω_0^p, ω_1^p, $\hat{\omega}$, $\hat{\omega}+\omega^*$, and $\hat{\omega}+\frac{4\bar{e}^2}{a\sigma^2}$.

Lemma 8 (Wealth Tresholds Ordering in Equilibrium with Binding Time-1 Constraint) *The order of the wealth thresholds is given in Table 4.*

Table 4: Wealth Threshold Order for Equilibrium with Binding Time-1 Constraint

Treshold	Greater than	Condition	Numerical value
$\omega_0^p \equiv s\bar{e}-\frac{1}{2}a\sigma^2 s^2$	ω_1^p	$\rho < \frac{3}{4}$	
	ω^c	for any $\rho > 0$ (equality if $\rho = 2$)	
	0	$\rho \geq \frac{1}{2}$	
	$\hat{\omega}$	$\rho > \frac{9}{2}$	
$\omega_1^p \equiv \frac{3}{2}s\bar{e}-\frac{7}{8}a\sigma^2 s^2$	ω^c	for all $\rho > 0$	
	0	$\rho > \frac{7}{12}$	0.583
	$\hat{\omega}$	$\rho > \frac{13}{4}$	
$\omega^c \equiv \omega_1^u - \frac{\bar{e}^2}{10a\sigma^2}$	0	$7-2\sqrt{10} \leq \rho \leq 7+2\sqrt{10}$	0.675 - 13.32
	$\hat{\omega}$	never, equality for $\rho = 7$	
$\hat{\omega} \equiv 4a\sigma^2 s^2$	0	for any $\rho > 0$	

Relevant risk benefit ratio thresholds. The relevant thresholds for ρ are thus, in ascending order, $\frac{1}{2}$, $\frac{7}{12}$, $7-2\sqrt{10}$, $\frac{3}{4}$, 1, $\frac{13}{4}$, $\frac{9}{2}$, 7 and $7+2\sqrt{10}$. The thresholds 1 and 7 correspond to a change in $\mathcal{F}_0^{\bar{c}_1}$. The ordering of wealth thresholds per ρ-interval is given in Table 5. Since the positivity of the wealth thresholds does not affect the equilibrium outcome, I group all the cases where $\rho < \frac{3}{4}$ together. Similarly, I do not distinguish the case with $\rho \geq 7+2\sqrt{10}$, as it does not affect the order.

D.3.3 Equilibrium case by case

Equilibrium determination follows from combining the results above. The details are available in the Internet Appendix.

Table 5: ρ and Wealth Intervals for Equilibrium with Binding Time-1 Constraint

Case	ρ interval	Wealth ordering
1	$\rho < \frac{3}{4}$	$(\omega^c, 0)^+ < (\omega_1^p, 0)^+ < (\omega_0^p, 0)^+ < \hat{\omega}$
2	$\frac{3}{4} \leq \rho < 1$	$0 < \omega^c < \omega_0^p < \omega_1^p < \hat{\omega}$
3	$1 \leq \rho < \frac{13}{4}$	$0 < \omega^c < \omega_0^p < \omega_1^p < \hat{\omega}$
4	$\frac{13}{4} \leq \rho < \frac{9}{2}$	$0 < \omega^c < \omega_0^p < \hat{\omega} < \omega_1^p$
5	$\frac{9}{2} \leq \rho < 7$	$0 < \omega^c < \omega_0^p < \hat{\omega} < \omega_1^p$
6	$7 \leq \rho$	$(\omega^c, 0)^+ < \hat{\omega} < \omega_0^p < \omega_1^p$

D.4 Equilibrium spreads

Corollary 8 (Equilibrium Spreads in the $\bar{c}_1$ equilibria) *In the u_0, c_1 and c_0, c_1 equilibria, spreads are*

$$\begin{aligned}
\Delta_0^{u_0,c_1} &= 2a\sigma^2 s + \bar{e} - \sqrt{d_1^+(x_0^{u_0,c_1})}, \quad \Delta_1^{u_0,c_1} = a\sigma^2 s + \bar{e} - \sqrt{d_1^+(x_0^{u_0,c_1})} \\
\Delta_0^{c_0,c_1} &= 2(a\sigma^2 s + \bar{e}) - \sqrt{d_0^+(x_0^{c_0,c_1})} - \sqrt{d_1^+(x_0^{c_0,c_1})}, \quad \Delta_1^{c_0,c_1} = a\sigma^2 s + \bar{e} - \sqrt{d_1^+(x_0^{c_0,c_1})}
\end{aligned} \tag{18}$$

Proof. Follows from substituting equilibrium positions in the spread schedules (13)-(12). ■

E Coexistence

Proposition 18 (Equilibria with Slack and Binding Time-1 Constraints May Coexist)

- *There is a unique equilibrium when arbitrage capital is either sufficiently low or sufficiently high:*
  - *If $0 \leq W_0 < \max(0, \omega^f)$, the unique equilibrium is $\mathbf{c_0}, \mathbf{c_1}$.*
  - *If $W_0 \geq \max(\omega_0^u, \omega_1^u, \omega_1^p)$, the unique equilibrium is $\mathbf{u_0}, \mathbf{u_1}$.*
- *When capital is intermediate, i.e. if $\max(0, \omega^f) \leq W_0 < \max(\omega_0^u, \omega_1^u, \omega_1^p)$, multiple equilibria may coexist depending on the level of ρ:*
  - *For $0 \leq \rho < \frac{7}{10}$, two equilibria may coexist:*
    - *If $\omega^f \leq W_0 < \max(\omega_0^u, \omega_1^u)$, $\mathbf{c_0}, \mathbf{u_1}$ may coexist with $\mathbf{c_0}, \mathbf{c_1}$.*
    - *If $\max(\omega_0^u, \omega_1^u) \leq W_0 < \omega_1^p$, $\mathbf{u_0}, \mathbf{u_1}$ may coexist with $\mathbf{c_0}, \mathbf{c_1}$.*

    In the special case where $0 \leq \rho < \frac{79}{140}$ and $\omega_1^p \leq W_0 < \omega_0^u$, $\mathbf{c_0}, \mathbf{u_1}$ is the unique equilibrium, with $x_0^{c_0,u_1} = \bar{x}_0$.

– *For* $\frac{7}{10} \leq \rho < \frac{3}{4}$*, two equilibria may coexist:* $\mathbf{u_0}, \mathbf{u_1}$ *with* $\mathbf{c_0}, \mathbf{c_1}$*, or* $\mathbf{c_0}, \mathbf{u_1}$ *with* $\mathbf{c_0}, \mathbf{c_1}$*.*

– *For* $\rho \geq \frac{3}{4}$*,* $\omega_0^u < \omega_1^u < \omega_1^p$*, and*

  * *If* $\omega^f \leq W_0 < \min(\omega_1^u, \omega_0^p)$*,* $\mathbf{c_0}, \mathbf{u_1}$ *may coexist with* $\mathbf{c_0}, \mathbf{c_1}$*.*
  * *If* $\omega_0^p \leq W_0 < \omega_1^u$*,* $\mathbf{c_0}, \mathbf{u_1}$ *may coexist with* $\mathbf{u_0}, \mathbf{c_1}$*.*
  * *If* $\omega_1^u \leq W_0 < \max(\omega_1^u, \omega_0^p)$*,* $\mathbf{u_0}, \mathbf{u_1}$ *may coexist with* $\mathbf{c_0}, \mathbf{c_1}$*.*
  * *If* $\max(\omega_0^p, \omega_1^u) \leq W_0 < \omega_1^p$*,* $\mathbf{u_0}, \mathbf{u_1}$ *may coexist with* $\mathbf{u_0}, \mathbf{c_1}$*.*

Proof. See Internet Appendix. ■

F Welfare

F.1 Proposition 7

Proof. *Price effects.* Given the definition of $d_1^+(x_0)$ and the spread $\Delta_t^{u_0,c_1}$ given in Proposition 17 and Corollary 8, and the definition of $\Delta_t^{u_0,u_1}$ given in Proposition 2 we get: $\Delta_1^{u_0,c_1} < \Delta_1^{u_0,u_1} \Leftrightarrow \frac{2}{5}a\sigma^2 s + \bar{e} < \sqrt{d_1^+\left(x_0^{u_0,c_1}\right)}$. Raising both sides to the square and rearranging terms gives after simplication

$$\Delta_1^{u_0,c_1} < \Delta_1^{u_0,u_1} \quad \Leftrightarrow \quad W_0 > r_1 \equiv \frac{7}{5}s\bar{e} - \frac{71}{50}a\sigma^2 s^2$$

Proceeding in the same fashion for time-0 spreads gives:

$$\Delta_0^{u_0,c_1} < \Delta_0^{u_0,u_1} \quad \Leftrightarrow \quad W_0 > r_0 \equiv \frac{6}{5}s\bar{e} - \frac{37}{25}a\sigma^2 s^2$$

Clearly, $r_0 < r_1$, so $\Delta_1^{u_0,c_1} < \Delta_1^{u_0,u_1} \Rightarrow \Delta_0^{u_0,c_1} < \Delta_0^{u_0,u_1}$. I then determine the position of the thresholds r_0 and r_1 relative to 0, ω_0^u, ω_1^u, ω_0^p, and $\bar{\omega}_0^u$. The results are given in Table 6.

Table 6: Thresholds for price effects in Proposition 7

Threshold	Lower than	Interval	Threshold	Lower than	Interval
r_1	< 0	$\rho < \frac{71}{70}$	r_0	< 0	$\rho < \frac{37}{30}$
	$< \omega_1^u$	always		$< \omega_0^u$	$\rho < \frac{5}{2}$
	$< \omega_1^p$	always		$< \omega_1^u$	always
	$< \omega_0^p$	$\rho < \frac{23}{10}$		$< \omega_0^p$	$\rho < \frac{49}{10}$
	$< \omega_0^u$	$\rho < \frac{47}{30}$		$< \omega_1^p$	always

Further, from Proposition 17, $\omega_0^p < \omega_1^p$ iff $\rho > \frac{3}{4}$, so we only need to consider this region. The potential coexistence region is determined by the position of ρ relative to 1 (Proposition 18). Thus, the relevant thresholds for ρ are $\frac{3}{4}$, 1, $\frac{47}{30}$, $\frac{23}{10}$, $\frac{5}{2}$, and $\frac{49}{10}$ (ignoring the positivity constraints for r_0 and r_1). Therefore, there are six cases:

1. If $\frac{3}{4} \leq \rho < 1$, then $r_0 < r_1 < \omega_0^u < \omega_1^u < \omega_0^p < \omega_1^p$. The potential coexistence region for this interval is $[\omega_0^p, \omega_1^p[$. Thus, if the u_0, u_1 and u_0, c_1 equilibria coexist, then $W_0 \geq \max(r_0, r_1)$, so $\Delta_t^{u_0,c_1} < \Delta_t^{u_0,u_1}$.

2. If $1 \leq \rho < \frac{47}{30}$, then $r_0 < r_1 < \omega_0^u < \omega_0^p < \omega_1^u < \omega_1^p$. The potential coexistence region is now $[\omega_1^u, \omega_1^p[$. If equilibria coexist, then $W_0 \geq \max(r_0, r_1)$, so $\Delta_t^{u_0,c_1} < \Delta_t^{u_0,u_1}$. The remaining cases are the same as case 2:

3. If $\frac{47}{30} \leq \rho < \frac{23}{10}$, then $r_0 < \omega_0^u < r_1 < \omega_0^p < \omega_1^u < \omega_1^p$. 4. If $\frac{23}{10} \leq \rho < \frac{5}{2}$, then $r_0 < \omega_0^u < \omega_0^p < r_1 < \omega_1^u < \omega_1^p$. 5. If $\frac{5}{2} \leq \rho < \frac{49}{10}$, then $\omega_0^u < r_0 < \omega_0^p < r_1 < \omega_1^u < \omega_1^p$. 6. If $\frac{49}{10} \leq \rho$, then $\omega_0^u < \omega_0^p < r_0 < r_1 < \omega_1^u < \omega_1^p$. Thus, if the two equilibria coexist, spreads are always smaller in the u_0, c_1 equilibrium.

Hedgers' welfare. Since $\mathbb{E}_0(p_1) - p_0 = \frac{1}{2}(\Delta_0 - \Delta_1)$ and $\mathbb{E}_0(p_2 - p_1) = \frac{1}{2}\Delta_1$, we can rewrite equation (7) in Lemma 2 as

$$U_0 = \frac{(\Delta_0 - \Delta_1)^2 + \Delta_1^2}{8a\sigma^2} - \frac{s}{2}\Delta_0 \tag{19}$$

The first term represents hedgers' capital gains on their time 0 and time 1 positions. The second term represents the total cost of sharing risk at a discount relative to the fundamental value (the expected value). Equation (19) gives hedgers' welfare in market A. Market B is symmetric. From (19), we get:

$$\begin{aligned} & U_0^{u_0,c_1} > U_0^{u_0,u_1} && (20) \\ \Leftrightarrow\ & (\Delta_0^{u_0,c_1} - \Delta_1^{u_0,c_1})^2 - (\Delta_0^{u_0,u_1} - \Delta_1^{u_0,u_1})^2 + (\Delta_1^{u_0,c_1})^2 - (\Delta_1^{u_0,u_1})^2 > 4a\sigma^2 s(\Delta_0^{u_0,c_1} - \Delta_0^{u_0,u_1}) \end{aligned}$$

Using (18) and Proposition 17, we get $\Delta_0^{u_0,c_1} - \Delta_1^{u_0,c_1} = a\sigma^2 s$, $\Delta_0^{u_0,u_1} - \Delta_1^{u_0,u_1} = \frac{3}{5}a\sigma^2 s$, thus condition (20) becomes

$$\frac{11}{5}a^2\sigma^4 s^2 + 2\bar{e}^2 + 2a\sigma^2 W_0 - 2(a\sigma^2 s + \bar{e})\sqrt{d_1^+(x_0^{u_0,c_1})} > 4a\sigma^2 s\left[\frac{1}{5}a\sigma^2 s + \bar{e} - \sqrt{d_1^+(x_0^{u_0,c_1})}\right]$$

Rearranging the terms, we can rewrite condition (20) as

$$a\sigma^2(W_0 - \omega^h) > (\bar{e} - a\sigma^2 s)\sqrt{d_1^+(x_0^{u_0,c_1})}, \quad \text{with } \omega^h \equiv 2s\bar{e} - \tfrac{7}{10}a\sigma^2 s^2 - \tfrac{\bar{e}^2}{a\sigma^2} \tag{21}$$

I then place ω^h relative to 0, ω_0^u, ω_1^u, ω_0^p, and ω_1^p: see Table 7. Since the interval of interest

Table 7: Thresholds for hedgers' welfare in Proposition 7

Threshold	Greater than	Interval
ω^h	> 0	iff $\rho \in \left]1 - \frac{\sqrt{1.2}}{2}, 1 + \frac{\sqrt{1.2}}{2}\right[$
	$> \omega_0^u$	iff $\rho \in \left]\frac{3}{5} - \frac{\sqrt{14}}{10}, \frac{3}{5} + \frac{\sqrt{14}}{10}\right[$
	$> \omega_1^u$	iff $\rho \in \left]0, \frac{3+\sqrt{29}}{10}\right[$
	$> \omega_0^p$	iff $\rho \in \left]\frac{1}{2} - \frac{\sqrt{5}}{10}, \frac{1}{2} - \frac{\sqrt{5}}{10}\right[$
	$> \omega_1^p$	iff $\rho \in \left]0, \frac{1}{4} + \frac{\sqrt{95}}{20}\right[$

is $\rho \geq \frac{3}{4}$, the only relevant thresholds are $\frac{3+\sqrt{29}}{10}$ (≈ 0.88), $\frac{3}{5} + \frac{\sqrt{14}}{10}$ (≈ 0.97), 1 and $1 + \frac{\sqrt{1.2}}{2}$. I add the threshold 1, as it determines the region of potential coexistence. We have thus five cases:

1. If $\frac{3}{4} \leq \rho < \frac{3+\sqrt{29}}{10}$, then $0 < \omega_0^u < \omega_1^u < \omega^h < \omega_0^p < \omega_1^p$
2. If $\frac{3+\sqrt{29}}{10} \leq \rho < \frac{3}{5} + \frac{\sqrt{14}}{10}$, then $0 < \omega_0^u < \omega^h < \omega_1^u < \omega_0^p < \omega_1^p$
3. If $\frac{3}{5} + \frac{\sqrt{14}}{10} \leq \rho < 1$, then $0 < \omega^h < \omega_0^u < \omega_1^u < \omega_0^p < \omega_1^p$
4. If $1 \leq \rho < 1 + \frac{\sqrt{1.2}}{2}$, then $0 < \omega^h < \omega_0^u < \omega_0^p < \omega_1^u < \omega_1^p$
5. If $1 + \frac{\sqrt{1.2}}{2} \leq \rho$, then $\omega^h < 0 < \omega_0^u < \omega_0^p < \omega_1^u < \omega_1^p$

It is clear that when equilibria potentially coexist under the conditions of Proposition 18, then $W_0 \geq \omega^h$. Therefore for any $\rho \geq \frac{3}{4}$, the left-hand side of condition (21) is positive. Instead, the right-hand side is negative for $\rho < 1$ and positive otherwise. Thus, if $\frac{3}{4} \leq \rho < 1$, condition (21) holds and $U_0^{u_0,c_1} > U_0^{u_0,u_1}$. If $\rho \geq 1$, we can raise both sides of (21) to the square to determine the trade-off. Substituting for $d_1^+\left(x_0^{u_0,c_1}\right)$, we can rewrite (21) as:

$$a^2\sigma^4 W_0^2 - 2a\sigma^2\left[a\sigma^2\omega^h + (\bar{e} - a\sigma^2 s)^2\right] W_0 + (a\sigma^2\omega^h)^2 - (\bar{e} - a\sigma^2 s)^2\left[a^2\sigma^4 s^2 + (a\sigma^2 s - \bar{e})^2\right] > 0$$

Viewing the left-hand side as a polynomial in W_0, we can calculate its discriminant. After a few lines of algebra, we obtiain $4a^2\sigma^4(\bar{e} - a\sigma^2 s)^2\left[2(\bar{e} - a\sigma^2 s)^2 + 2a\sigma^2\omega^h + a^2\sigma^4 s^2\right] > 0$.

After calculating the term in parenthesis, we can write the roots as

$$W_{(1)} = \frac{a\sigma^2\omega^h + (\bar{e} - a\sigma^2 s)^2 - (\bar{e} - a\sigma^2 s)\sqrt{\frac{8}{5}a^2\sigma^4 s^2}}{a\sigma^2} = -\sqrt{\frac{8}{5}}s\bar{e} + \left(\frac{3}{10} + \sqrt{\frac{8}{5}}\right)a\sigma^2 s^2$$

$$W_{(2)} = \sqrt{\frac{8}{5}}s\bar{e} + \left(\frac{3}{10} - \sqrt{\frac{8}{5}}\right)a\sigma^2 s^2$$

In the relevant region for coexistence, $\rho \geq 1$; this implies that $W_{(1)} < W_{(2)}$. It then remains to determine the position of these roots relative to ω_1^u and ω_1^p. We have: for any ρ, $W_{(2)} < \omega_1^p$ and $W_{(2)} < \omega_1^u$. Thus $W_0 \in [\omega_1^u, \omega_1^p[$ implies $W_0 > W_{(2)}$, which implies that $U_0^{u_0,c_1} > U_0^{u_0,u_1}$.

Arbitrageurs' welfare. First, I show that $\Omega_0^{u_0,c_1} > \Omega_0^{u_0,u_1}$ is equivalent to $W_0 \in [\omega^a, \omega^{a'}]$. To see this, note that $\Omega_0^{u_0,c_1} > \Omega_0^{u_0,u_1}$ is equivalent to (equilibrium utilities are given in Propositions 2 and 17)

$$\bar{e}\sqrt{d_1^+\left(x_0^{u_0,c_1}\right)} > a\sigma^2(W_0 - \underline{\omega}^a), \quad \text{with } \underline{\omega}^a \equiv s\bar{e} - \tfrac{9}{10}a\sigma^2 s^2 - \tfrac{\bar{e}^2}{a\sigma^2}$$

Since $W_0 \geq \omega_0^p > \underline{\omega}_0^a$, the right-hand side is positive. So raising both sides to the square preserves the order. After some simple algebra, the condition becomes

$$-a^2\sigma^4 W_0^2 + 2a\sigma^2(\bar{e}^2 + a\sigma^2\underline{\omega}^a)W_0 + \bar{e}^2\left[a^2\sigma^4 s^2 + (a\sigma^2 s - \bar{e})^2\right] - (a\sigma^2\underline{\omega}^a)^2 > 0$$

Viewing the left-hand side as a polynom in W_0, and using the definition of $\underline{\omega}^a$, we compute the discriminant. After some simplification, we obtain $\frac{4}{5}a^4\sigma^8 s^2\bar{e}^2$. Thus, we can write the roots as

$$\omega^a = \left(1 + \frac{1}{\sqrt{5}}\right)s\bar{e} - \frac{9}{10}a\sigma^2 s^2 > \omega^{a'} = \left(1 - \frac{1}{\sqrt{5}}\right)s\bar{e} - \frac{9}{10}a\sigma^2 s^2$$

This establishes the first point. Second, I compare the roots to the equilibrium thresholds. Since $1 + \frac{1}{\sqrt{5}} < \frac{3}{2}$, $\omega_1^p > \omega^a$. Since $1 + \frac{1}{\sqrt{5}} > \frac{7}{5}$, $\omega^a > \omega_1^u$. Further, $\omega^{a'} < \omega_0^p$ and $\omega^a > \omega_0^p$ iff $\rho > \frac{2\sqrt{5}}{5}$. Thus, arbitrageurs are better off in the partly constrained equilibrium iff $W_0 \in [\max(\omega_1^u, \omega_0^p), \omega^a[$. Given that $\omega_0^p > \omega_1^u$ for $\rho < 1$, this interval is not empty iff $\rho > \frac{2\sqrt{5}}{5}$.
■

F.2 Corollary 2

Proof. The first point is obvious. The condition under which the constraint does not bind is $W_0 \geq \max(\omega_0^u, \omega_1^u, \omega_1^p)\mathbb{1}_{\rho<0.7\cap\rho>0.75} + h\mathbb{1}_{0.7\leq\rho\leq 0.75}$, where h is the threshold represented by a dotted line on Figure 7.

The second point follows from Proposition 7 and the fact that, in the absence of constraint, the unique equilibrium is u_0, u_1. Proposition 7 shows that hedgers are better off in the u_0, c_1 than in the u_0, c_1 equilibrium under milder conditions than point 2. Thus, the conditions for Pareto improvement are those for the improvement in the arbitrageur's welfare. We have shown that $\omega^a > \omega_0^p$ iff $\rho > \frac{2\sqrt{5}}{5}$, so we must restrict our attention to this interval. For $W_0 \in [\omega_0^p, \omega_1^u]$, u_0, c_1 is the unique equilibrium in the presence of constraints. Given that $\omega_1^u < \omega^a$, the arbitrageur is better off. For $W_0 \in [\omega_1^u, \omega^a]$, u_0, c_1 coexists with u_0, u_1, thus the arbitrageur is better off only if u_0, c_1 is selected.[43]

The third point follows from comparing hedgers' welfare in the u_0, u_1 vs c_0, c_1 or c_0, u_1 equilibria. From Proposition 4, recall that $x_0^{c_0,u_1} = \bar{x}_0 < x_0^{u_0,u_1}$ and $X_1^{c_0,u_1} = \frac{s+\bar{x}_0}{2}$. This implies that $\Delta_1^{c_0,u_1} = a\sigma^2(s-\bar{x}_0) > \Delta_1^{u_0,u_1}$ and that $\Delta_0^{c_0,u_1} - \Delta_1^{c_0,u_1} = 2a\sigma^2(s-\bar{x}_0)$. Then using (19), $U_0^{c_0,u_1} < U_0^{u_0,u_1}$ iff

$$\frac{5}{8}a\sigma^2(s-\bar{x}_0)^2 - \frac{3}{2}a\sigma^2(s-\bar{x}_0) < \frac{5}{8}a\sigma^2(s-x_0^{u_0,u_1})^2 - \frac{3}{2}a\sigma^2(s-x_0^{u_0,u_1})$$

After a few lines of simple algebra, this condition boils down to

$$\frac{5}{2}(\bar{x}_0^2 - (x_0^{u_0,u_1})^2) < a\sigma^2 s(x_0^{u_0,u_1} - \bar{x}_0)$$

Thus, $x_0^{c_0,u_1} = \bar{x}_0 < x_0^{u_0,u_1}$ implies that this condition is satisfied, so hedgers' welfare decreases when imposing constraints leads to c_0, u_1.

Similarly, we get

$$U_0^{c_0,c_1} = \frac{4a^2\sigma^2(s-\bar{x}_0)^2 + 4a^2\sigma^2(s-\bar{X}_1)^2}{8a\sigma^2} - \frac{s}{2}\left[2a\sigma(s-\bar{x}_0) + 2a\sigma^2(s-\bar{X}_1)\right]$$

Thus the condition $U_0^{c_0,c_1} < U_0^{u_0,u_1}$ can be simplified to

$$\frac{1}{2}a\sigma^2\left[(s-\bar{x}_0)^2 - (s-x_0^{u_0,u_1})^2 + (s-\bar{X}_1)^2 - (s-X_1^{u_0,u_1})^2\right] < a\sigma^2 s\left[x_0^{u_0,u_1} - \bar{x}_0 + X_1^{u_0,u_1} - \bar{X}_1\right],$$

[43]Given that it Pareto-dominates, one may argue that this is the most likely outcome.

which can be further reduced to

$$\frac{1}{2}a\sigma^2\left[\bar{x}_0^2 - (x_0^{u_0,u_1})^2 + \bar{X}_1^2 - (X_1^{u_0,u_1})^2\right] < 0$$

This condition holds true since when the constraint binds, $\bar{x}_0 < x_0^{u_0,u_1}$ and $\bar{X}_1 < X_1^{u_0,u_1}$ (the latter also follows from the analysis in the proof of Proposition 10). Hence hedgers' welfare also decreases when imposing constraints leads to c_0, c_1. ■

F.3 Proposition 8

Proof. *Counterfactual 1.* I compute hedgers' welfare using equation (19), under the assumptions that $x_0^{cf1} = x_0^{u_0,u_1} = \frac{2}{5}s$ and $X_1^{cf1} = X_1^{u_0,c_1} = \frac{a\sigma^2 s - \bar{e} + \sqrt{d_1^+(x_0^{u_0,c_1})}}{2a\sigma^2}$. These quantities imply the following spreads:

$$\Delta_1^{cf1} = \Delta_1^{u_0,c_1}, \quad \Delta_0^{cf1} = 2a\sigma^2(s - x_0^{u_0,u_1}) + \Delta_1^{u_0,c_1} = \frac{6}{5}s + \Delta_1^{u_0,c_1} > \Delta_0^{u_0,c_1}$$

Substituting into (19), we obtain

$$U_0^{cf1} = \frac{\left(\frac{6}{5}s\right)^2 + (\Delta_1^{u_0,c_1})^2}{8a\sigma^2} - \frac{s}{2}\left(\frac{6}{5}s + \Delta_1^{u_0,c_1}\right)$$

Therefore, $U_0^{cf1} < U_0^{u_0,c_1}$ can be simplified into $\frac{27}{25} > \frac{3}{4}$, which holds true.
Counterfactual 2. The quantities are $x_0^{u_0,c_1} = \frac{s}{2}$ and $X_1^{cf2} = X_1^{u_0,u_1} = \frac{7}{10}s$, implying that $\Delta_1^{cf2} = \Delta_1^{u_0,u_1} = \frac{3}{5}a\sigma^2 s$, $\Delta_0^{cf2} - \Delta_1^{cf2} = a\sigma^2(s - x_0^{u_0,c_1}) = a\sigma^2 s$, and $\Delta_0^{cf2} = a\sigma^2 s + \Delta_1^{cf2} = \frac{8}{5}a\sigma^2 s$. Substituting these spreads into equation (19), we get $U_0^{cf2} = -\frac{63}{100}a\sigma^2 s^2$. Comparing this welfare level to $U_0^{u_0,u_1}$, we get $U_0^{cf2} < U_0^{u_0,u_1} < U_0^{u_0,c_1}$. ■

F.4 Proposition 10

See Internet Appendix.

F.5 Corollary 3

Proof. Recall that margins (on long positions) are given by $m_t = \bar{e} - \frac{1}{2}(\Delta_t - \Delta_{t+1}) = \bar{e} - a\sigma^2(s - X_t)$. Besides, the VaR at time t is the product of the position and the margin: $VaR_t = m_t X_t$. Thus, at time 0, $VaR_0^l = \omega_0^l$, $l \in \{*, u, p\}$. Thus, with competitive arbitrageurs, when the constraint is slack, $X_t^* = s$, $m_t^* = \bar{e}$ and $VaR_t^* = \omega^*$. When the constraint binds, the position is reduced and the margins and thus the VaR increase: $X_t^* = \bar{X}_t < s$, $m_t^* = \bar{e} - a\sigma^2(s - \bar{X}_t)$, and $VaR_t^* = \bar{X}_t\bar{e} - a\sigma^2(s - \bar{X}_t) < \omega^*$. With market power, if the risk benefit ratio is low enough ($\rho < 7/10$), a drop in capital switches the equilibrium from u_0, u_1 to c_0, u_1. Since $X_t^{c_0,u_1} < X_t^{u_0,u_1}$, $m_t^{c_0,u_1} < m_t^{u_0,u_1}$, and $VaR_t^{c_0,u_1} < VaR_t^{u_0,u_1}$, so the comparative statics are the same as with competitive arbitrageurs. When ($7/10 \leq \rho < 3/4$), the equilibrium a drop in capital switches the equilibrium from u_0, u_1 to c_0, c_1, and the analysis is similar. When $\rho \geq 3/4$, however, a drop in capital switches the equilibrium from u_0, u_1 to u_0, c_1, and then to c_0, c_1. Between u_0, c_1 and c_0, c_1, positions decrease following a drop in capital, so margins and VaR also decrease. However, between u_0, u_1 and u_0, c_1, the position always increases at time 0, and so margins and VaR also increase. At time 1, the position increases as long as $\rho < 23/10$. Indeed in the proof of Proposition 7 shows that $r_1 < \omega_0^p$ on this interval for the time-1 spread to decrease between u_0, u_1 and u_0, c_1, and a decrease in time-1 spread is equivalent to an increase in time-1 position. ■

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.

Boyarchenko, N., T. Eisenbach, P. Gupta, O. Shachar, and P. V. Tassel (2018): "Bank-Intermediated Arbitrage," FRB of New York Staff Report No. 858.

Bruche, M., and C. F. Kuong (2019): "Dealer Funding and Market Liquidity," INSEAD working paper.

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

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.

Dávila, E., and A. Korinek (2017): "Pecuniary Externalities in Economies with Financial Frictions," *Review of Economic Studies*, 85, 352–395.

DeMarzo, P., and B. Urosevic (2006): "Optimal Trading and Asset Pricing with a Large Shareholder," *Journal of Political Economy.*

Dow, J., J. Han, and F. Sangiorgi (2019): "Why is Capital Slow Moving? Liquidity Hysteresis and the Dynamics of Limited Arbitrage," London Business School working paper.

Eisenbach, T., and G. Phelan (2020): "Cournot Fire Sales in Real and Financial Markets," FRB of New York Staff Report No. 837.

Falato, A., D. Iercosan, and F. Zikes (2019): "Banks as Regulated Traders," Finance and Economics Discussion Series 2019-005. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2019.005.

Gabaix, X., P. Gopikrishnan, V. Plerou, and H. E. Stanley (2006): "Institutional Investors and Stock Market Volatility," *Quarterly Journal of Economics*, 121(2), 461–504.

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.

——— (2018): "The Dynamics of Financially Constrained Arbitrage," *Journal of Finance*, 73, 1713–1750.

Gul, F., H. Sonnenschein, and R. Wilson (1986): "Foundations of Dynamic Monopoly and the Coase Conjecture," *Journal of Economic Theory*, 39.

I., B. D., F. Franzoni, R. Moussawi, and J. Sedunov (2015): "The Granular Nature of Large Institutional Investors," NBER Working Paper No. 22247.

Kahn, C. (1986): "The Durable Goods Monopolist and Consistency with Increasing Costs," *Econometrica*, 54, 275–294.

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.

Kondor, P. (2009): "Risk in Dynamic Arbitrage: Price Effects of Convergence Trading," *Journal of Finance*, 64 (2), 638–658.

Kondor, P., and D. Vayanos (2019): "Liquidity Risk and the Dynamics of Arbitrage Capital," *Journal of Finance*, 74, 1139–1173.

Marinovic, I., and F. Varas (2019): "Dynamic Blockholder Incentives: Liquidity and Reputation," Stanford Working Paper.

McAfee, R. P., and T. Wiseman (2008): "Capacity Choice Counters the Coase Conjecture," *Review of Economic Studies*, 75, 317–332.

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

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

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.

Stockey, N. (1981): "Rational Expectations and Durable Goods Pricing," *Bell Journal of Economics*, 12, 112–128.

Tirole, J. (1988): *The Theory of Industrial Organization*. MIT Press.

van Kervel, V., and A. Menkveld (2019): "High-Frequency Trading around Large Institutional Orders," *Journal of Finance*, 74, 1091–1137.

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

Wallen, J. (2020a): "Intermediary Asset Pricing and Imperfect Competition," Stanford working paper.

——— (2020b): "Markups to Financial Intermediation in Foreign Exchange Markets," Stanford working paper.

Zarinelli, E., M. Treccani, D. Farmer, and F. Lillo (2015): "Beyond the Square Root: Evidence for Logarithmic Dependence of Market Impact on Size and Participation Rate," *Market Microstructure and Liquidity*, 1.