Risk in Dynamic Arbitrage: Price Effects of Convergence Trading*

Job Market Paper

Péter Kondor

London School of Economics

November 10, 2005

Abstract

This paper studies the adverse price effects of convergence trading. We assume two assets with identical cash flows traded in segmented markets. Initially, there is gap between the prices of the assets, because local traders' hedging needs differ. In the absence of arbitrageurs, the gap remains constant until a random period when the difference across local markets disappears. While arbitrageurs' activity reduces the price gap, it also generates potential losses: the price gap widens with positive probability in each period. The size of the gap is determined by the time-varying option value of saving a unit of capital for the next period. In a calibrated example we show that these endogenously created losses alone can explain episodes when arbitrageurs lose most of their capital in a relatively short time.

JEL classification: G10, G20, D5.

Keywords: Convergence trading; Limits to arbitrage; Liquidity crisis; Option value of waiting

1 Introduction

It has been widely observed that prices of fundamentally very similar assets can differ significantly. Perhaps the best known examples are the so-called "Siamese twin stocks" (e.g. Royal Dutch Petroleum /Shell Transport and Trade, Unilever NV/Unilever PC, SmithKline Beckman/Beecham Group) which represent claims on virtually identical cash-flows, yet their price differential fluctuates substantially around the theoretical parity.[1] Financial institutions speculating on the convergence of prices of similar assets (whom we will loosely refer to as "arbitrageurs") can suffer large losses if diverging prices force them to unwind some of their positions. The spectacular collapse of the Long-Term

*Email address: p.kondor@lse.ac.uk. We are grateful for the guidance of Hyun Shin and Dimitri Vayanos and the helpful comments from Margaret Bray, Max Bruche, Zsuzsi Elek, Antoine Faure-Grimaud, Bernardo Guimaraes, Miklós Koren, Arvind Krishnamurthy, John Moore, Jakub Steiner, Gergely Ujhelyi and seminar participants at the LSE and at the 2005 European Winter Meeting of the Econometric Society, Istanbul. All remaining errors are our own. We gratefully acknowledge the EU grant "Archimedes Prize" (HPAW-CT-2002-80054).

[1] See Lamont and Thaler (2003) and Froot and Dabora (1999) for details.

Capital Management hedge fund in 1998 is frequently cited as an example of this phenomenon.[2] In this paper we argue that the possibility of similar episodes is an equilibrium consequence of the competition of arbitrageurs with limited capital. In contrast to previous models,[3] our mechanism is not a result of the amplification of exogenous adverse shocks. Instead, it is based on an efficiency argument. Arbitrageurs' competition generates the possibility of losses, because without these the investment opportunity would be too attractive to exist in equilibrium. In a calibrated example we show that these endogenously created losses alone can explain episodes when arbitrageurs lose most of their capital in a relatively short time.

We present an analytically tractable, stochastic, general equilibrium model of convergence trading. We assume two assets with identical cash flows traded in segmented markets. Initially, there is a gap between the prices of the assets because local traders' hedging needs differ. In each period the difference across local markets disappears with positive probability. Therefore, in the absence of arbitrageurs, the gap remains constant until a random period when it disappears. We label this interval with asymmetric local demand a *window* of arbitrage opportunity. Arbitrageurs can profit from the temporary presence of the window by taking opposite positions in the two markets. Arbitrageurs pay a small unit cost for holding the position and they have limited capital to cover their losses. If their trades did not affect prices, the development of the gap would provide a one-sided bet as prices could only converge. However, by trading, they endogenously determine the size of the gap as long as the window is open. At the same time, arbitrageurs have to decide how to allocate their capital over time given the uncertain characteristics of future arbitrage opportunities, i.e., the development of the price gap.[4] Thus, there is an interdependence between arbitrageurs' optimal strategies and the pattern of future arbitrage opportunities.

Our main result is that arbitrageurs' individually optimal strategies generate losses in the form of widening price gaps. Essentially, opportunities which are "too attractive" have to be eliminated in equilibrium. In particular, in each period expected gains and expected losses per unit of capital

[2] For detailed analysis of the LTCM crisis see e.g. Edwards (1999), Loewenstein (2000), MacKenzie (2003). Although after the collapse of the LTCM many market participants made changes to their risk-management systems to avoid similar events, it is clear that financial markets are still prone to similar liquidity crises. A very recent example is the turbulence in May 2005 connected to the price differential between General Motors stocks and bonds:

> "The big worry is that an LTCM-style disaster is occurring with hedge funds as they unwind GM debt/stock trade (a potential Dollars 100bn trade across the industry) at a loss, causing massive redemptions from convert arb funds, forcing them to unwind other trades, and so on, leading to a collapse of the debt markets and then all financial markets." (Financial Times, US Edition, May 23, 2005)

[3] Apart from the seminal paper of Shleifer and Vishny (1997) and the following literature on limits to arbitrage (e.g. Xiong, 2001, Kyle and Xiong, 2001, Gromb and Vayanos, 2002), there is also a related literature which concentrates on endogenous risk as a result of amplification due to financial constraints (e.g. Danielsson and Shin, 2002, Danielsson et al. 2002, 2004, Morris and Shin, 2004, Bernardo and Welch, 2004).

[4] Our focus on the timing of arbitrage trades connects our work to Abreu and Brunnermeier (2002, 2003). However, our problem is dramatically different. They analyze a model where the development of the gap between a price of an asset and its fundamental value is exogenously given and informational asymmetries cause a coordination problem in strategies over the optimal time to exit the market. In our model, information is symmetric, arbitrageurs are competitive, they want to be on the market when others are not (i.e., there is strategic substitution instead of complementarities), and our focus is the endogenous determination of the price gap. Furthermore, we do not model a bubble, but the endogenous development of a price gap, which cannot increase above a certain level.

invested have to be such that arbitrageurs are indifferent about when to invest that particular unit. Otherwise, no arbitrageur would choose the dominated periods. If in any two consecutive periods when the window is still open the gap does not widen, investing in the first period will dominate saving capital for the second one. This is so because the investment opportunity is not getting more profitable, and there is the additional risk that the window closes by the second period. But if arbitrageurs do not save capital for the second period, the gap will widen.[5] Hence, the competition of arbitrageurs transforms the price process in a fundamental way. Without arbitrageurs the price gap could only converge. While arbitrageurs' activity reduces the price gap, it also generates potential losses: the price gap widens with positive probability in each period as long as arbitrageurs have any capital to invest.

We show that our results are not driven by the size of the holding cost. In particular, we demonstrate with the help of a calibrated example that the losses created by the strategies of arbitrageurs can be quantitatively substantial even in the limit when the holding cost is insignificantly small. These endogenous losses alone are enough to cause arbitrageurs to lose most of their capital in a relatively short time with positive probability. The role of the small holding cost is only to ensure that the equilibrium is unique. If taking positions was free, as long as the aggregate level of capital is not too large, there would exist many equilibria. In one of these equilibria, arbitrageurs fully eliminate the price gap in each period. This equilibrium is ruled out by any positive holding cost. Another one of these equilibria is the limiting case of our unique equilibrium: the price gap is always positive and it gets wider in each period with positive probability.

We also show that the size of the gap is determined by the option value of saving a unit of capital for the next period. It is compensation for the sacrifice of future arbitrage opportunities. This option value determines the premium over the unit return of the risk-free asset, i.e., the expected gain-to-loss ratio corresponding to the arbitrage opportunity. Thus, this time-varying premium term changes the size of the gap as if the risk-premium was changing over time, consistently with the empirical literature on time-varying risk-aversion.[6] The premium is increasing with longer windows, i.e., it is increasing as the aggregate capital of arbitrageurs decreases because of past losses. This is again reminiscent of an observation in the literature that "risk-appetite" is smaller when liquidity is scarce.[7] However, as arbitrageurs are risk-neutral, this premium is not related to preferences. The premium is smaller for betting on shorter windows because it has to make arbitrageurs indifferent across intertemporal arbitrage possibilities and the probability of longer windows is smaller. Thus, arbitrageurs will allocate more capital for early periods, i.e., they bet extensively on short windows which keeps the premium low. At the same time, the different source of this premium term has implications which differ from those of a risk-based explanation. Most importantly, our mechanism implies that the premium is not necessarily monotonic in risk. Although the premium is largest when the aggregate level of capital

[5] Interestingly, our argument is similar to the text-book mechanism of full elimination of price discrepancies by risk-neutral arbitrageurs with unlimited capital. There, price discrepancies providing positive expected profit would attract large investments and would be eliminated. In our case, price gaps which cannot diverge would be similarly attractive possibilities and would be eliminated.

[6] See, e.g., Engle and Rosenberg (2002) and the references therein.

[7] See, e.g., Dungey et al. (2003), Kumar and Persaud (2001), Gai and Vause (2005).

is small, the potential loss from arbitrage trades is typically much larger when the aggregate level of capital is at an intermediate level. Note that this can be a starting point for testing our model.

The analytical tractability of our model stems from the reduced state space in our structure. The state of the window influences the distribution of the price gap in a trivial manner. If the window closes in a given period, then the price gap jumps to zero in that period and our modelled world ends. It is thus sufficient to characterize the path of the price gap conditional on an open window. This result in a one-dimensional system, where one state variable, time (or equivalently the aggregate level of capital at that point in time) determines the equilibrium development of the price gap and the aggregate portfolio of arbitrageurs.

Our model naturally belongs to the literature on general equilibrium analysis of risky arbitrage (e.g., Gromb and Vayanos, 2002, Zigrand, 2004, Xiong, 2001, Kyle and Xiong, 2001, Basak and Croitoru, 2000). However, to the best of our knowledge, our paper is the first to show that competition of arbitrageurs alone can generate losses. In contrast to previous models focusing on potential losses in convergence trading (Shleifer and Vishny, 1997, Xiong, 2001, Gromb and Vayanos, 2003, Liu and Longstaff, 2004), our mechanism is not based on the amplification of exogenous shocks by the capital constraint. Instead, it is based on an efficiency argument. The reason for our unique finding is that we are the first to analyze the price effect of arbitrageurs whose dynamic portfolio choice is influenced by uncertain future arbitrage possibilities over many periods. Shleifer and Vishny (1997) and Gromb and Vayanos (2002) touch upon one of the elements of our mechanism. Allowing uncertainty of future opportunities in one period only, they show that arbitrageurs may be reluctant to take a maximal position as they fear that they will make losses when the arbitrage possibility will be the most attractive. We show that allowing uncertainty over many periods, this consideration is sufficient to transform the price process in a systematic way. In spirit, our paper is close to Liu and Longstaff (2004) who argue that arbitrage with limited capital might lead to substantial losses. However, in their model, this is a result of an exogenously defined price process, while our focus is on the determination of the price process. The closest paper to our model in its stochastic structure is Xiong (2001) as he also considers arbitrage possibilities which are present for an uncertain time span. However, a crucial difference is that in his model, because of their specific preferences, the uncertainty over future arbitrage opportunities does not influence the decisions of arbitrageurs. Furthermore, results rely mostly on numerical results, while our equilibrium can be fully characterized analytically.

This paper proceeds as follows. In section 2, we present the structure of the model. In section 3, we derive the unique equilibrium. In section 4, we discuss the results and in section 5, we analyze the robustness of our equilibrium. Finally, we conclude.

2 A simple model of risky arbitrage

The structure of the model is based on three group of agents: two groups of local traders in two segmented markets and a group of arbitrageurs taking positions on both markets. A temporary asymmetry in the demand curves of local traders creates an arbitrage opportunity. As the asymmetry is bound to disappear sooner or later, so is the arbitrage opportunity. As arbitrageurs can take positions on both markets, they can exploit the price discrepancy. They are in the focus of our analysis. They have limited capital so they have to decide how to allocate it over time given the distribution of future arbitrage opportunities. Their strategies in turn determine the development of the price gap through market clearing. Hence, there is interdependence between the distribution of future arbitrage opportunities and the individual strategies of the arbitrageurs.

First, we describe the available assets in the economy, then we introduce traders and finally we present arbitrageurs.

2.1 Assets

There are two segmented markets represented by two islands, $i = A, B$.[8] On each island a single risky asset is traded. We will call them $A-$asset and $B-$asset respectively. Both assets are in zero net supply. The two assets of the two islands have identical payoff structure: at the end of each period both assets pay the random dividend R_t, where R_t is distributed according to the cumulative distribution function $F(R_t)$ with mean $E(R_t)$ and finite variance σ^2. There is an infinite number of periods, $t = 0, 1, 2....$ A riskless bond with a unit gross return is available on both islands as a storage technology.

2.2 Local traders and the window of arbitrage opportunity

Each island is inhabited by traders: A-traders live on island A and B-traders live on island B. Each trader lives for two periods and a new generation of traders is born in each period.[9] Each trader receives a lifetime endowment and makes an investment decision when young, and liquidates her investment and consumes the proceeds when old. Her lifetime endowment, $e_{t,i}$, consists of a constant wage, e, and a non-tradable stochastic component which is correlated with the pay-off of the risky asset, R_t. In particular, the stochastic component pays $\omega_{t,i}R_t$ units at the end of period t. Hence, the value of her endowment by the end of the period (when R_t is realized) will be

$$e_{t,i} = e + \omega_{t,i}R_t \quad i = A, B.$$

Initially, this stochastic component is different across traders on the two islands as

$$\omega_{t,A} = \omega^A \text{ and } \omega_{t,B} = \omega^B$$

[8] As our focus is not the source of the arbitrage possibility, we take market segmentation as given. Gromb and Vayanos (2002) and Zigrand (2004) use similar assumptions. Nieuwerburgh and Veldkamp (2005) provide a mechanism which results in endogenous market segmentation.

[9] Our OLG formulation is a simplifying assumption to keep our model tractable. It is not necessary for our story to go through. For example, infinitely lived agents of Lucas (1978) would be consistent with the intuition. However, this would make the model much less tractable as in this case asset prices would depend on both future and present consumption, hence they depend on past, present and future asset-holding. This would complicate the analysis without providing virtually any added value.

where $\omega^A \neq \omega^B$. However, in each period the stochastic component disappears with probability $(1-q)$, i.e., $\omega_{t,A} = \omega_{t,B} = 0$, or survives to the next generation born in the next period with probability q. The focus of our analysis will be this time-interval of uncertain length when the asymmetry is present. We will call this interval a "window" of arbitrage opportunity. We will refer to the state when incomes are back to equal as "normal times". The temporary income-difference is the only difference between A-traders and B-traders.

Each young trader born in period t on island i makes the investment decision of buying $\theta_{t,i}$ units of the risk asset when young, before R_t is realized. In period $t+1$, when old, she liquidates her investment and consume her total income: the value of her endowment plus the proceeds of her investment decision. Each of them values her consumption according to a standard, strictly increasing and concave utility function $u(\cdot)$ which is twice continuously differentiable for any achievable consumption level. Hence, an i-trader, $i = A, B$, born in period t in a window solves the problem

$$\max_{\theta_{t,i}} E_{R_t} \left[(1-q) \, u \left(e + \omega^i R_t + \theta_{t,i} \left(R_t + p^n - p^w_{t,i} \right) \right) + q u \left(e + \omega^i R_t + \theta_{t,i} \left(R_t + p^w_{t+1,i} - p^w_{t,i} \right) \right) \right] \quad (1)$$

where $\theta_{t,i}$ is the asset holding of a trader on island i and period t, $p^w_{t,i}$ is the price of the asset on island i at time t in a window, and p^n is the constant price on both islands in normal times. The first order condition of this problem implicitly determines the inverse demand curve $p^w_{t,i} = p^w_i \left(\theta_{t,i}, p^w_{t+1,i} \right)$ for a fixed p^n. For a given p^n, the asset price depends on the supply of the asset and the next period price given that the window is still open in the next period.

We also impose no-ponzi scheme conditions on prices:

$$\lim_{t \to 0} \left| p^w_{t,i} \right| < \infty \text{ for } i = A, B. \quad (2)$$

We assume that normal times do not last forever, and with probability q the economy switches back to a window state. Hence, in normal times, traders solve a very similar problem to (1). As the focus of the analysis is only on the first window of arbitrage opportunity, the existence of subsequent windows serves simply to pin down the price p^n. To economize on the notation in the main text, we delegate the problem of traders in normal times and the determination of p^n to the appendix.

For a moment, let us assume that there are no arbitrageurs who trade between markets, so both markets have to clear separately. We will refer to this case as autarchy. In autarchy, the aggregate supply of the risky asset which has to be held on each island is zero unit. It is also natural to suspect that the price process will remain constant if the income-state does not change, i.e., $\theta_{t,i} = 0$ and $p^w_{t,i} = p^*_i$ for $i = A, B$. In the appendix we show that there is a unique combination of p^*_A, p^*_B and p^n which satisfies the first order conditions of traders under these restrictions. Therefore we know that in the absence of arbitrageurs, the gap between the prices of the two assets is $g^* \equiv p^*_A - p^*_B$ when the window is open, while in normal times it is $0 = p^n - p^n$. Without loss of generality we will assume that $g^* > 0$. Therefore, the autarchy gap process switches from g^* to 0 with conditional probability $(1-q)$ in each period.

To proceed we have to make the following technical assumptions on $u(\cdot)$, ω^A and ω^B.

Assumption 1 *For any future prices,* p^n *and* $p^w_{t+1,i} \in [p^*_B, p^*_A]$*, there exists a positive, finite, minimal* $\bar{\theta} = \bar{\theta}\left(p^w_{t+1,A}, p^w_{t+1,B}\right)$ *such that* $p^w_A\left(\bar{\theta}, p^w_{t+1,A}\right) - p^w_B\left(-\bar{\theta}, p^w_{t+1,B}\right) = 0.$

Assumption 2 *Inverse demand functions* $p^w_A(\cdot)$ *for* $\theta_{t,A} \in \left[\bar{\theta}^{\max}, 0\right]$ *and* $p^w_B(\cdot)$ *for* $\theta_{t,B} \in \left[0, -\bar{\theta}^{\max}\right]$ *and for* $p^w_{t+1,A}$*,* $p^w_{t+1,B} \in [p^*_B, p^*_A]$ *are well defined and continuously differentiable in* $\theta_{t,i}$ *and future prices, where*

$$\bar{\theta}^{\max} = \max_{p^w_{t+1,A}, p^w_{t+1,B}} \bar{\theta}\left(p^w_{t+1,A}, p^w_{t+1,B}\right)$$

for $p^w_{t+1,A}, p^w_{t+1,B} \in [p^*_B, p^*_A]$.

Assumption 3 *In the full domain of* $p^w_i(\cdot)$ *there is no such* $\theta_{t,A} \notin \left[\bar{\theta}^{\max}, 0\right]$ *or* $\theta_{t,B} \notin \left[0, -\bar{\theta}^{\max}\right]$ *that* $p^w_{t,i} = p^w_i\left(\theta_{t,i}, p^w_{t+1,i}\right)$ *if* $p^w_{t,i}, p^w_{t+1,i} \in [p^*_B, p^*_A]$.

Assumption 4 *In the domain defined in Assumption 2, inverse demand functions* $p^w_i(\cdot)$ $i = A, B$ *are decreasing in the supply of the asset,* $\frac{\partial p^w_i(\cdot)}{\partial \theta_{t,i}} < 0$*, and increasing in possible future prices,* $\frac{\partial p^w_i(\cdot)}{\partial p^w_{t+1,i}} > 0$.

Assumption 1 is an innocuous requirement that the price gap can be eliminated with sufficiently large positions on the two markets. We need Assumption 2 to make sure that the demand curves are well defined on the relevant domain. The relevant domain is defined by allowing future prices on both islands to move between the autarchy prices of windows and allowing θ to move between 0 and the smallest position which eliminates the gap for any relevant future prices. Assumption 3 is a technical assumption to ensure that we do not have to worry about $\theta_{t,i}$ because the inverse demand function does not "bend back" into the relevant domain. Assumption 4 ensures that the income effect is not too strong: within the relevant domain traders demand more of the asset if it is cheaper and if its future price is higher.

We present two examples to demonstrate that Assumptions 1-4 are consistent with a wide range of utility functions and that these assumptions tend to be satisfied if traders are not very risk-averse, the volatility of R_t is sufficiently small or if islands are not too different, i.e., $\left|\omega^A - \omega^B\right|$ is sufficiently small.

Example 1 (CARA-symmetric framework) *Let us suppose that the utility of traders with consumption* h_t *when old is* $u(h_t) = -\exp(-\alpha h_t)$*, and* R_t *is normally distributed with zero mean and* σ^2 *variance and* $\omega^A = -\omega^B = -\omega$*. In the appendix, we show that these assumptions create a symmetric structure in the following sense. If arbitrageurs provide the risky asset in opposite aggregate amount for the two islands in each period, i.e.,* $\theta_{t,A} = -\theta_{t,B} = \theta_t$*, then there is a function* $p^w\left(\theta_t, p^w_{t+1,A} - p^n\right)$ *that*

$$p^w_{t,A} - p^n = p^n - p^w_{t,B} = p^w\left(\theta_t, p^w_{t+1,A} - p^n\right)$$

, i.e., opposite positions of size θ_t *push both prices away from* p^n *to the same extent and the absolute slope of the inverse demand functions is the same at that point on both islands. In the appendix we derive the closed form for* $p^w\left(\theta_t, p^w_{t+1,A} - p^n\right)$ *and show that if either* α*,* ω *or* σ^2 *is sufficiently small Assumptions 1-4 are satisfied.*

Example 2 (CRRA framework) *Let us suppose that the utility of traders with wealth h_t when old is $u(h_t) = h_t^{(1-\gamma)}$ and the support of $F(R_t)$ is bounded. In the appendix we show that if $\gamma \leq 1$ and ω^A is sufficiently close to ω^B, then Assumptions 1-4 are satisfied.*

2.3 Arbitrageurs

Because of the difference in endowments of the risky assets, if markets on the two islands clear separately, there will be a price differential of g^* between asset prices, although the assets have identical dividend structure. Arbitrageurs can reduce this gap by taking positions on both markets. Arbitrageurs are the model-equivalent of global hedge funds with the resources and the expertise to discover such price anomalies and to take positions in distant local markets. Arbitrageurs live forever, they are risk neutral and operate in a competitive environment: they are small and they have a unit mass. Arbitrageurs take positions $x_{t,i}$ on island $i = A, B$. We make the following assumption on $x_{t,A}$ and $x_{t,B}$.[10]

Assumption 5 *Arbitrageurs take exactly opposite positions on the two markets, i.e., $x_{t,A} = -x_{t,B} = x_t$.*

Therefore, arbitrageurs engage in "market neutral arbitrage trades". We call the composite asset of one long unit of the B-asset and one short unit of the A-asset the "gap asset". We will show that in equilibrium x_t is non-negative, i.e., arbitrageurs always buy the cheap asset and sell the more expensive one. We will label such strategy as "short selling x_t unit of the gap". In normal times, demand curves in the two markets coincide thus arbitrageurs are not motivated to trade. Consequently, it is sufficient to focus on the dynamic strategies of arbitrageurs during the interval of the open window of arbitrage opportunity.

Note that if arbitrageurs were not financially constrained, the strategy of short selling the gap would be riskless and would lead to unbounded profit as long as the gap is non-zero. However, because of the financial constraints specified below, sometimes arbitrageurs are forced to liquidate before the prices of the assets converge, which can (and in equilibrium will) lead to losses. In effect, their strategy is neutral only to the random payoff of the assets, R_t, but not to the endogenous fluctuations of relative prices caused by the random time of income convergence and arbitrageurs' trades. Consequently, their arbitrage strategy is risky.

[10] Although the assumption seems intuitive, taking exactly opposite positions in the two markets is optimal only if the inverse demand functions are symmetric as in Example 1.

Both Xiong (2001) and Gromb and Vayanos (2002) make assumptions to ensure that arbitrageurs do not take asymmetric positions across local markets. It simplifies the analysis substantially, because it implies that arbitrageurs' trades are neutral to the uncertainty of assets' payoff, R_t. Xiong (2001) makes a direct assumption, while Gromb and Vayanos (2002) assume that traders have CARA utility and opposite endowment shocks vary similarly to the structure in Example 1. We choose the direct way, because – as we argue in footnote 14 – the equilibrium would be almost identical even if we allowed arbitrageurs to take asymmetric positions. However, those readers who prefer the latter solution, can see Example 1 as our formalization of traders and can ignore Assumption 5.

From our assumptions on the demand of traders, it is simple to construct the inverse demand function for the gap asset

$$g\left(\bar{x}_t, p^w_{t+1,A}, p^w_{t+1,B}\right) \equiv p^w_A\left(\bar{x}_t, p^w_{t+1,A}\right) - p^w_B\left(-\bar{x}_t, p^w_{t+1,B}\right)$$

which determines the gap $g_t \equiv p^w_{t,A} - p^w_{t,B}$ as a function of arbitrageurs' aggregate position $\bar{x}_t$ and future prices.

Function $g(\cdot)$ is continuous and decreasing in $\bar{x}_t$. Note that as demand functions are increasing in future prices by Assumption 4, $g(\cdot)$ will also be increasing in the future gap g_{t+1} as long as prices $p^w_{t+1,A}$ and $p^w_{t+1,B}$ move in the opposite direction. As this is the case under Assumption 5, we can simplify the notation by using $g(\bar{x}_t, g_{t+1})$. If there is no arbitrage activity either at present or in the future, the gap is g^*. This is the autarchy price gap.

Arbitrageurs are financially constrained by the following institutional environment. Each arbitrageur starts her activity with the same amount of capital[11] $v_0 = \bar{v}_0$, where $\bar{v}_0$ is the aggregate capital available in the economy. They do not get any extra funds as long as the window is open. In section 5 we will relax this assumption. They need funds for their activity for two reasons. Firstly, there is a small, positive unit cost[12], m, of short selling the gap. This is the carry cost of the position. We assume that m is small in the following sense:

$$(1-q)\, g^* > m. \tag{3}$$

Later we will show that as long as m is positive, it can be arbitrarily small without changing the qualitative results or without having large influence on the quantitative effects. Secondly, arbitrageurs are required to fully collateralize their potential losses. This assumption can also be regarded as the formalization of endogenous margin requirements or VaR constraints. In effect, if $g_{t+1} > g_t$ to take a position of the size $x_t = \frac{v_{t-1}}{g_{t+1}-g_t+m}$, an arbitrageur has to be able to present v_{t-1} cash, i.e., deposit v_{t-1} on a margin account, as the maximal possible loss on each unit is $(g_{t+1} - g_t + m)$.

The problem of each arbitrageur is to find the optimal position, x_t, for each period for the contingency that the price discrepancy does not disappear until period t, subject to constraints of full collateralization. The recursive formalization of the problem is

$$\begin{aligned} V_t(v_t) &= \max_{x_t} (1-q)\left((g_t - m)\, x_t + v_t\right) + qV_{t+1}(v_{t+1}) \\ s.t. \quad v_{t+1} &= v_t - x_t\left(g_{t+1} - g_t + m\right) \\ 0 &\leq v_t - x_t\left(g_{t+1} - g_t + m\right). \end{aligned} \tag{4}$$

[11] There is no significance of v_0 being the same across arbitrageurs. The analysis would be virtually the same if the initial capital was distributed in any other way.

[12] We do not consider the case where m is negative i.e. arbitrageurs holding opposite position in two markets incur a net profit even if prices are unchanged. In contrast, Plantin and Shin (2005) analyze a set up with this property with the illustration of carry trades in foreign exchange markets. They show that speculative dynamics can arise, i.e., in terms of our model, arbitrageurs may bet on the divergence of prices instead of the convergence. The assumption of $m > 0$ rules this possibility out in our model.

The problem shows that if an arbitrageur sells x_t unit of the gap in period t for g_t two events can happen. If the window closes exactly in period $t+1$, she close her position for free and she gets $g_t x_t$ profit minus the carry cost $x_t m$ in addition to her remaining capital v_t. If the window is still open in period $t+1$, her new capital level which can be used for collateralizing trades, v_{t+1}, is adjusted by the loss or gain, $(g_{t+1} - g_t) x_t$, she experienced between the two periods and the carry cost $m x_t$. The value of this capital level is $V_{t+1}(v_{t+1})$.

In equilibrium, each arbitrageur follows a strategy $\{x_t\}_{t=0}^{\infty}$ which solves problem (4) for a given conditional gap path $\{g_t\}_{t=0}^{\infty}$ and the aggregate positions $\{\bar{x}_t\}_{t=0}^{\infty}$ support this conditional gap path, i.e.,

$$g(\bar{x}_t, g_{t+1}) = g_t. \tag{5}$$

In the next section we present the equilibrium.

3 Equilibrium

In this section we present the unique equilibrium of our system. In the first part, we show the intuition behind the proof. Then we state the result, but delegate the details of the proof to the appendix. In the second part of this section, we highlight the role of carry cost, m, by deriving the limit equilibrium as $m \to 0$ and presenting a calibrated example. We show that even in the limit, our main result holds: arbitrageurs can lose most of their capital in a relatively short time due to the endogenously created losses by their individual strategies. This finding is important as it demonstrates that our mechanism is not based on the amplification of an exogenous cost. We also show that when $m = 0$ there is a multiplicity of equilibria. Thus, our assumption of positive carry cost serves as an equilibrium selection mechanism.

We discuss the implications of the equilibrium in section 4.

3.1 Equilibrium in the general case

We show that the equilibrium can be derived in a simple, recursive way. The proof is based on two critical observations. The first one is that if the window of arbitrage opportunity is sufficiently long, arbitrageurs will lose all of their capital. The second one is that the gap path can be determined recursively by the condition that in equilibrium arbitrageurs have to be indifferent to the time of their investment. Let us see these two steps in turn.

The intuition behind the observation that a long enough window wipes out all capital of arbitrageurs is based on two facts. Firstly, in each period t when the window has not closed, arbitrageurs with positive positions x_t, will suffer a loss, i.e., $(g_{t+1} - g_t + m) > 0$. Otherwise, arbitrageurs would make a sure positive profit in period $t+1$, so they would not be constrained in period t. Thus, they would take an infinite position which is inconsistent with a non-negative gap by Assumption 1. Putting it simply, a market where $(g_{t+1} - g_t + m) < 0$ would provide such an excellent investment opportunity that it cannot exist in equilibrium. Even if the loss $(g_{t+1} - g_t + m)$ were very small, in each period when arbitrageur's aggregate position is positive and the window does not close, the level of aggregate capital decreases. The second fact to note is that if the level of the gap is close enough to g^* any arbitrageur will be happy to take a maximal position and lose all their capital if the window does not disappear in the next period. The reason is that g^* is the maximal size of the gap[13], so if g_t is close enough to g^*, the potential gain is close to its maximum and the potential loss, $g_{t+1} - g_t + m$ is close to its minimum. Clearly, there is no point in waiting and risking to miss out the arbitrage possibility, if it cannot get any better. Because of these two facts, there will be a period $T-1$ when the aggregate level of capital decreases below a critical level, so g_t increases to a level sufficiently close to g^*. This makes all arbitrageurs with remaining capital take a maximal position. Consequently, if the window is longer than $T-1$ periods, arbitrageurs lose all their capital and $g_{T+\tau} = g^*$ and $\bar{v}_{T+\tau} = 0$ for all $\tau \geq 0$.

It is also true that the conditional gap path g_t for $t = 0, 1, ...T-1$ has to be such that arbitrageurs are indifferent when to invest. It is so, because if investment in any period before T was dominated, then none of the arbitrageurs would invest at that period so the gap would be g^*. But we just pointed out that a gap of g^* would be sufficient motivation for any arbitrageurs to invest all their money, which is a contradiction. The formal condition for the indifference between period t and $t+1$ is

$$(1-q)\left(\frac{g_t - m}{g_{t+1} - g_t + m} + 1\right) = (1-q) + q(1-q)\left(\frac{g_{t+1} - m}{g_{t+2} - g_{t+1} + m} + 1\right). \tag{6}$$

The left hand side shows the expected gross profit from investing a unit of capital in period t and taking the maximal position of $\frac{1}{g_{t+1}-g_t+m}$, as it pays the gross profit of $\left(\frac{g_t-m}{g_{t+1}-g_t+m} + 1\right)$ with probability $(1-q)$ and 0 otherwise. The right hand side is the expected gross profit of saving the unit until period $t+1$. It is easy to see that for a given final value g_{T-1} we can construct recursively an arbitrarily long indifference path.

There are only two problems left to pin down the equilibrium. The number of necessary steps backwards, T, and the determination of g_{T-1}. We know that g_{T-1} has to be large enough to ensure that arbitrageurs are happy to invest their remaining capital in period $T-1$ instead of waiting until period T. Using equation (6) and the fact that $g_{T+\tau} = g^*$ for all $\tau \geq 0$, this is equivalent to

$$(1-q)\frac{g_{T-1} - m}{g^* - g_{T-1} + m} \geq q(1-q)\left(\frac{g^* - m}{m} + 1\right). \tag{7}$$

The size of T and g_{T-1} are simultaneously determined by the condition that g_{T-1} satisfies (7) and that total losses between $t = 0$ and $t = T-1$ exactly equal the aggregate capital of the economy

$$\sum_{t=0}^{T-1} \bar{x}_t\left(g_{t+1} - g_t + m\right) = \bar{v}_0 \tag{8}$$

where the average positions in each period have to be consistent with the size of the gap.

[13] Our outline of the proof builds on the fact that the aggregate position $\bar{x}_t$ is always non-negative i.e. arbitrageurs never speculate on the widening of the gap. We will show in the formal proof why it is the case. The intuition relies on the point that $\bar{x}_t < 0$ would be consistent only with a bubble path of g_t. But a bubble path is not consistent with the finite level of aggregate capital of arbitrageurs.

Formally, we can state the following proposition.[14]

Theorem 1 *There is a unique equilibrium in our economy in terms of average positions, which consists of a period T, a path of average positions $\{\bar{x}_t\}_{t=0}^{\infty}$ and a conditional gap path $\{g_t\}_{t=0}^{\infty}$. The equilibrium is characterized by the following expressions:*

$$g_t = g^* \text{ and } \bar{x}_t = 0 \text{ for } t \geq T$$

$$g_{T-1} \geq \frac{q\left(g^*\right)^2}{m + qg^*} + m \tag{9}$$

$$\frac{g_t - m}{g_{t+1} - g_t + m} = \sum_{j=1}^{T-1-t} q^j + \frac{q^{T-1-t}\left(g_{T-1} - m\right)}{g^* - g_{T-1} + m} \quad \text{for } \; t < T-1 \tag{10}$$

and

$$g_{T-1} = g\left(\bar{x}_{T-1}, g^*\right) \tag{11}$$

$$g_t = g\left(\bar{x}_t, g_{t+1}\right) \text{ for } t < T-1 \tag{12}$$

$$\sum_{t=0}^{T-1} \bar{x}_t\left(g_{t+1} - g_t + m\right) = \bar{v}_0. \tag{13}$$

Furthermore, $\{g_t\}_{t=0}^{T-1}$ is strictly monotonically increasing in t.

Proof. Details of the proof are in the appendix. ■

[14] A gap path of a very similar structure is an equilibrium in the unrestricted problem, where arbitrageurs can buy and sell different amounts of the two assets. The unrestricted problem is

$$E\left(V\left(v_t\right)\right) =$$
$$= \max_{x_t}\left(1-q\right)\Big(\left(p_{t,A} - p^n + E\left(R_t\right)\right)x_{t,A} + \left(p_{t,B} - p^n + E\left(R_t\right)\right)x_{t,B} - \frac{m}{2}\left(\left|x_{t,A}\right| + \left|x_{t,B}\right|\right) + v_t\Big) + qE\left(V\left(v_{t+1}\right)\right)$$
$$s.t. \quad v_{t+1} = v_t - \left(p_{t+1,A} - p_{t,A}\right)x_{t,A} - \left(p_{t+1,B} - p_{t,B}\right)x_{t,B} - \frac{m}{2}\left(\left|x_{t,A}\right| + \left|x_{t,B}\right|\right) - \left(x_{t,A} + x_{t,B}\right)R_t$$

where $x_{t,A}$ and $x_{t,B}$ are the amounts sold of the two asset at t. We omitted the full collateralization constraint as we focus on the periods when it does not bind. The first order conditions in these periods are

$$\left(1-q\right)\Big(p_{t,i} - p^n + E\left(R_t\right) - \frac{m}{2}\left(sgn\left(x_{t,i}\right)\right)\Big) = qE\left(V\left(v_{t+1}\right)\right)\Big(p_{t+1,A} - p_{t,A} + \frac{m}{2}sgn\left(x_{t,i}\right) + E\left(R_t\right)\Big) \qquad i = A, B.$$

If $p_A^* > p^n > p_B^*$, it is safe to assume in equilibrium $x_{t,A} \geq 0 \geq x_{t,B}$. Thus, if we subtract the first order condition for asset B from the one for asset A, we get

$$\left(1-q\right)\left(g_t - m\right) = qE\left(V\left(v_{t+1}\right)\right)\left(g_{t+1} - g_t + m\right)$$

and the envelope theorem implies

$$E\left(V'\left(v_t\right)\right) = \left(1-q\right) + qE\left(V'\left(v_t\right)\right).$$

This last two equations are identical to the first order conditions describing the equilibrium for the restricted problem for $t < T-1$ (see the proof of Theorem 1 in the appendix) except for the presence of the expectation operator. Hence, until arbitrageurs have sufficient funds to support this path, the gap path in the unrestricted problem is qualitatively the same as in the restricted problem. However, local price paths $p_{t,A}, p_{t,B}$ can differ, because $x_{t,A} \neq -x_{t,B}$. Furthermore, from period $T-1$ the two problems differ as in the unrestricted problem the asymmetric portfolio implies that v_t depends on the realization of R_t. Thus, even if arbitrageurs take maximal positions in $T-1$, if they happen to receive a high R_t, they might end up with positive v_T. Their capital will diminish only gradually, in a stochastic manner.

It is important to see that the equilibrium determines only the average positions, $\bar{x}_t$. Individual positions are arbitrary, as long as the average is $\bar{x}_t$. For example, it is possible that each arbitrageur is passive until a given t, when she takes a maximal position.[15] If in each period a given proportion μ_t chooses to do so while $\mu_t \frac{v_0}{g_{t+1}-g_t+m} = \bar{x}_t$ holds, this will be consistent with the equilibrium.

In Figure 1, we demonstrate the main qualitative features of the equilibrium. The graph shows the development of the gap conditional on the surviving asymmetry of markets and the corresponding conditional average positions. A point g_t represents the value of the gap in period t with probability q^t, which is the probability that the window lasts for at least t periods. With the complementary probability the realized gap will be 0 at t. Therefore, in reality, we would observe only the beginning of the curve. With probability $(1-q)$, we observe only element g_0, with probability $q(1-q)$ we observe g_0 and g_1, and with decreasing probability, longer parts of it. The graph shows that from a certain period T on, arbitrageurs do not take any positions and the gap remains at the autarchy level g^*. Until T, the gap monotonically increases. In line with the increasing gap path, the average position of arbitrageurs, $\bar{x}_t$ monotonically decreases until T. The conditional gap path typically has an S shape: convex for small t and concave for large t. For later reference, we emphasize the main

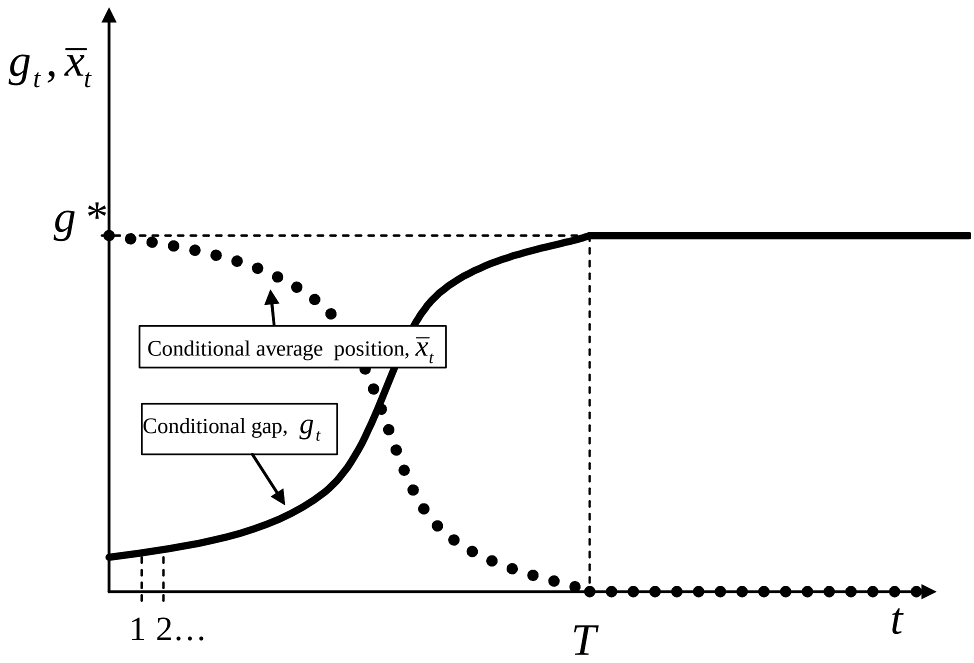

Figure 1: The qualitative features of the equilibrium conditional gap path, g_t, and the average position of arbitrageurs, $\bar{x}_t$. Both variables are plotted conditionally on the window of arbitrage opportunity being still open at period t.

result of Theorem 1 in the following corollary.

[15] Intuitively, this would correspond to a threshold-strategy of entering the market if and only if the gap reaches g_t.

Corollary 1 *For any $m > 0$ and any aggregate level of capital, $\bar{v}_0$, the conditional gap path $\{g_t\}_{t=0}^{T-1}$ is strictly increasing and arbitrageurs lose all their capital in finite time with positive probability.*

In each period before T, the gap widens and arbitrageurs who invest suffer a loss with positive probability. The intuition provided by the outline of the proof is that any period t without a potential loss would be a too attractive possibility for the arbitrageurs. It would encourage them to increase their positions in t to the point when g_t is lower than g_{t+1}, i.e., arbitrageurs create losses. It is interesting to contrast this fact with the intuition of other models of limits to arbitrage, e.g., Shleifer and Vishny (1997), Xiong (2001), Gromb and Vayanos (2002) and Liu and Longstaff (2004). They all emphasize that arbitrageurs might lose money because they might be forced to liquidate early if the gap widens. However, in those models the initial widening of the gap happens for reasons which are exogenous to the arbitrageurs' strategies. In particular, it is a result of noise traders trading against fundamentals (Shleifer and Vishny, 1997, Xiong, 2001, Gromb and Vayanos, 2002) or an exogenously specified price process (Liu and Longstaff, 2004). In our model, if arbitrageurs did not trade, the gap could never widen. But because they trade, it will widen with positive probability in each period. It is purely the individually optimal strategy of the arbitrageurs which is responsible for potential losses.

To support our view that endogenous losses are independent from exogenous shocks, in the next subsection we clear up the role of carry cost, m. We will argue that the main role of m is to select a unique equilibrium. We also quantify our results by a simple calibration exercise and show that the size of m has a very limited influence on the quantitative results. Then, in Section 4, we will have a closer look on the equilibrium and provide further insights into our mechanism.

3.2 Equilibrium in the limit and a calibrated example

It is already apparent from Corollary 1 that the size of m does not influence the main qualitative properties of the equilibrium as long as m is positive. In this subsection we will argue that our mechanism remains quantitatively significant even if $m \to 0$. We are particularly interested in whether losses created endogenously by the strategies of arbitrageurs diminish as $m \to 0$. If the endogenous losses did disappear, this would imply that our mechanism is based on the amplification of an external shock, and in this sense it would be very similar to the mechanism of Shleifer and Vishny (1997). A simple calibration exercise illustrated by Figure 2 shows that this is not the case.

We plot the graphs of the conditional gap path and of the proportion of the aggregate capital left by period t for $m = 10^{-3}$. We also plot the same two graphs for the limiting equilibrium to show that the equilibrium of $m = 10^{-3}$ is close to the limit. We will discuss the characteristics of the limiting equilibrium after the calibration exercise. We use the CARA-symmetric framework defined in Example 1 for the specification of the inverse demand function $g(\bar{x}_t, g_{t+1})$. We choose parameters $q = 0.5$, $\alpha = 0.5$, $\omega = 1$, $\sigma^2 = 1$ which imply $g^* = 2$. With these parameters Assumptions 1-3 are satisfied and a position of $\bar{\theta}^{\max} = 1.62$ would eliminate the gap for any future prices.

Let us suppose that each period corresponds to a week. We choose the aggregate level of capital ($\bar{v}_0 = 1.265$) in a way to ensure that the annualized profit of an average arbitrageur following the optimal strategy is at a reasonable level. In our case it is 10 percent when $m = 10^{-3}$ and 2 percent in the limiting case.[16] It is apparent that the endogenous losses created do not diminish even if m is insignificantly small. In fact, the average arbitrageur lose more than 99 percent of her capital in 13 weeks when $m = 10^{-3}$ and in 15 weeks in the limiting case, if she is unlucky to face a sufficiently long window of arbitrage opportunity. Although the probability of a 15-week-long window is quite small, the graph shows that in both cases arbitrageurs lose significant proportion of their capital even if the window is relatively short.

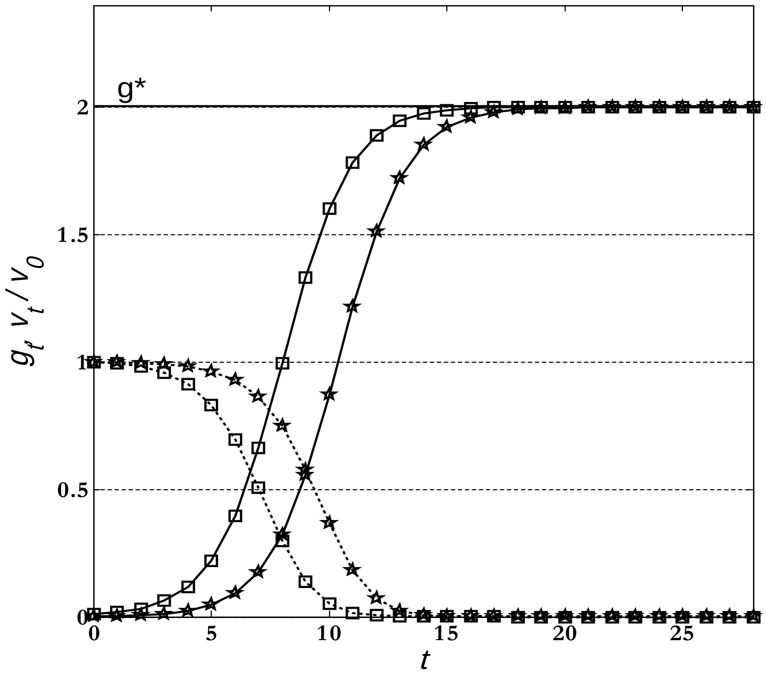

Figure 2: The increasing curves show the conditional gap path, g_t, and the decreasing curves show the proportion of aggregate capital left by period t, $\frac{\bar{v}_t}{\bar{v}_0}$. Graphs with squares are corresponding to the carry cost $m = 10^{-3}$ and graphs with stars are corresponding to the limiting equilibrium. Traders' preferences and shocks are defined by the CARA-symmetric framework of Example 1. Other parameters are $q = 0.5$, $\alpha = 0.5$, $\omega = 1$, $\sigma^2 = 1$, (which imply $g^* = 2$) and $\bar{v}_0 = 1.265$ (which implies the annualized return of 10% if $m = 10^{-3}$ and that of 2% in the limiting case).

Our calibration exercise illustrates that losses created by our mechanism are not proportional to m. To understand better the role of m in the model, let us consider the case of $m = 0$. In the absence of carry costs, there exists an intuitive equilibrium of our model where the gap is always 0. This is possible, since if there is no carry cost, arbitrageurs can commit to eliminate the price discrepancy at

[16] We show in Appendix A.1.2 how to determine g^* and $\bar{\theta}^{\max}$ from the primitives and how to ensure that Assumptions 1-3 are satisfied.

The annualized profit is calculated by valuating the marginal value function at period 0 and by using the fact that $q = 0.5$ implies that the expected length of the window is two weeks. The Matlab 7.0 program of the calibration exercise is available on request from the author.

all periods, because without possible losses their positions are not limited. This might be surprising as this is a very different equilibrium from the general case with $m > 0$. It turns out that as long as the aggregate capital of arbitrageurs, $\bar{v}_0$ is not very large[17], the equilibrium where $g_t = 0$ is not the only one when $m = 0$. In particular, in all of these alternative equilibria the conditional gap paths $\{g_t\}_{t=0}^{\infty}$ are monotonically increasing and they get very flat when t is very large. One of these equilibria is the limiting equilibrium of our model. In the limiting equilibrium the conditional gap path converges to g^*, but always stays below this level, i.e., $T \to \infty$.[18] Consequently, in the limiting equilibrium arbitrageurs never lose all of their capital. However, they will lose almost all relatively fast (as illustrated in Figure 2) as the largest losses are concentrated on the early periods because $(g_{t+1} - g_t)$ will be close to 0 for larger t values. We establish the following formal results.

Theorem 2 *1. If there are no carry costs, $m = 0$, there is an equilibrium of the game where $g_t = 0$ for all t. Additionally, there is a critical level of $\bar{v}_0^{\max} > 0$, that if $\bar{v}_0 < \bar{v}_0^{\max}$ there exist equilibria where conditional gap paths $\{g_t\}_{t=0}^{\infty}$ are strictly monotonically increasing.*

2. There is a critical level of $\bar{v}_0^{\lim}$, $\bar{v}_0^{\max} \geq \bar{v}_0^{\lim} > 0$, that if $\bar{v}_0 < \bar{v}_0^{\lim}$, as $m \to 0$ the equilibrium gap path converges to the strictly monotonically increasing path

$$g_t = \frac{g_0^{\lim} g^*}{q^t g^* + g_0^{\lim} (1 - q^t)}$$

for all $t \geq 0$ where $g_0^{\lim}$ is a function of $\bar{v}_0$.

Proof. The proof is in the appendix. ■

Intuitively, in the $m = 0$ case there is a coordination problem. The $g_t = 0$ equilibrium is achievable, but for this arbitrageurs should coordinate their future actions, i.e., they have to push down the gap in all states. Alternatively, if they are in an equilibrium where the gap increases, then arbitrageurs are not allowed to take unlimited positions because losses are possible. But there are also potential gains and arbitrageurs will be indifferent when to invest. Just as in the benchmark case of $m > 0$, endogenous losses occur in each period with positive probability, because otherwise the arbitrage possibility would

[17] When $\bar{v}_0$ is very large, there is no gap path, $\{g_t\}_{t=0}^{\infty}$ and corresponding average position path, $\{\bar{x}_t\}_{t=0}^{\infty}$ which would satisfy the aggregate budget constraint, $\sum_{t=0}^{\infty} \bar{x}_t (g_{t+1} - g_t) = \bar{v}_0$. It is so, because the left hand side of the budget constraint is bounded from above. If $m = 0$, arbitrageurs cannot lose more than $g^* \bar{\theta}^{\max}$ for any gap path where $\bar{\theta}^{\max}$ is the maximal position which is needed to push down the gap to 0 defined in Assumption 1. In this case, the only equilibrium is the one where $g_t = 0$ for all t. We do not have this problem when $m = 0$. Then, arbitrageurs will lose all of their capital with positive probability, even if $\bar{v}_0$ is very large.

[18] There is a simple way to show how a positive m makes a difference. When $m > 0$, the indifference condition is

$$\frac{g_t - m}{g_{t+1} - g_t + m} = q \left(1 + \frac{g_{t+1} - m}{g_{t+2} - g_{t+1} + m} \right).$$

Observe that the left hand side is bounded from below by $\frac{g_t - m}{g^* - g_t + m}$ and the right hand side is bounded from above by $q \left(\frac{g^* - m}{m} + 1 \right)$. It is easy to see that if g_t is close enough to g^*, the first expression is larger than the second one: the indifference condition will not hold and arbitrageurs invest all their money at t. When $m = 0$, the upper bound of the right hand side is infinity. Hence, there is a path where g_t goes to g^* and still arbitrageurs are always indifferent when to invest.

be too attractive. Thus, the role of a positive m is simply to select the equilibrium corresponding to the given aggregate capital level from this multiplicity.

4 Comparative statics and discussion

In this section we analyze the pattern of potential gains and losses along the equilibrium path, i.e., the expected profit and risk in arbitrage trades. We show that this pattern will be determined by the time-varying premium which arbitrageurs require to invest at a given time point instead of waiting for better opportunities. In the first part of this analysis, we show that the marginal value function $V_t'(v_t)$ is closely connected to this premium. In the second part, with the help of the properties of the marginal value function, we present the results on the role of length of the window and of the level of aggregate capital on this pattern to gain new insights into our mechanism.

4.1 The marginal value function: premium for sacrificed opportunities

In Section 3, we showed that in equilibrium arbitrageurs must be indifferent to the timing of their investments. For expositional purposes, in that section we used a direct approach and derived the formal condition (6) from the expected profit of arbitrageurs investing a dollar at different points of time. We could have derived this in the standard way with the help of the first order condition

$$(1-q)(g_t - m) = (g_{t+1} - g_t + m)\, qV_{t+1}'(v_{t+1}) \tag{14}$$

and the envelope condition

$$V_t'(v_t) = (1-q) + qV_{t+1}'(v_{t+1}), \tag{15}$$

for all $t < T$. Both equations come directly from the problem of arbitrageurs (4). Both equations are intuitive. The envelope condition shows that the value of a saved dollar today is the sum of a unit of return in the event of a closed window in the next period, plus the continuation value, $V_{t+1}'(v_{t+1})$, in the event of an open window in the next period. The first order condition shows that arbitrageurs are indifferent if their potential gain in the next period is equal to the value of their potential loss. Note that the instantaneous expected loss in a dollar position $(g_{t+1} - g_t + m)\, q$ on the right hand side is weighted by the value of a dollar in the next period $V_{t+1}'(v_{t+1})$. It turns out that $V_{t+1}'(v_{t+1})$ can be interpreted in two different ways in connection with the risk and profitability of arbitrage.

Firstly, $V_{t+1}'(v_{t+1})$ can be interpreted along the line of its similarity to the risk-premium. From the first order condition (14) we can see that if $V_{t+1}'(v_{t+1})$ was equal to 1, then the asset would be fairly priced: the expected gain would be the same as the expected instantaneous loss. As we will show, $V_{t+1}'(v_{t+1})$ is always larger than 1, i.e., arbitrageurs require a premium to invest instead of waiting. This premium is not due to risk-aversion as arbitrageurs are risk-neutral. This is a compensation for the sacrifice of future investment possibilities. This is apparent if we derive the closed form of $V_t'(v_t)$

for $0 \leq t \leq T-1$. Solving[19] the first order difference equation in (15), we get that

$$V_t'(v_t) = B\frac{1}{q}^t + 1$$

for an arbitrary B. Using the facts that arbitrageurs take a maximal position in period $T-1$ and that $v_{T+\tau} = 0$ for all $\tau \geq 0$, we can derive the value of $V_{T-1}'(v_{T-1})$ explicitly.[20] This final condition pins down B and we find that

$$V_t'(v_t) = q^{T-1-t}(1-q)\frac{g_{T-1} - m}{g^* - g_{T-1} + m} + 1 - q^{T-t}. \tag{16}$$

This is exactly the expected return on a unit saved until period $T-1$. A saved dollar invested in period $T-1$ gives a return $\frac{g_{T-1}-m}{g^*-g_{T-1}+m}$ with probability $q^{T-1-t}(1-q)$ which is the chance that the window closes exactly at period $T-1$. Additionally this strategy provides a unit return in all states in which the window closes before period T.[21]

The second possible interpretation of the marginal value function is connected to the expected profitability of the market, i.e., the attractiveness of the arbitrage possibility. This is a direct consequence of the fact that the value function is linear in the individual capital level v_t. This can be seen from (16), which shows that the marginal value function is independent from v_t as neither g_{T-1} nor T are influenced by v_t because all arbitrageurs are small. We can thus state that

$$V_t(v_t) = V_t' v_t$$

where V_t' is constant in v_t. This implies that the marginal value function at t shows the expected profit per capital for an arbitrageur who enters the market at t. Therefore, V_0' shows the attractiveness of the market at the beginning of the world.

Before proceeding to the analysis, let us note that there is a close relationship in our model between time and the aggregate capital level. Keep in mind that when referring to time, we refer to the time dependence of the conditional path, i.e., larger t corresponds to a longer window of arbitrage opportunity. As for the connection with the aggregate level of capital, note that in the derivation of the equilibrium T and g_{T-1} were pinned down uniquely by the aggregate initial capital $\bar{v}_0$. With the backward logic that we followed in the derivation of the equilibrium, we can find the price for each period from a period t on if we know the aggregate level of capital at that period, $\bar{v}_t$. Thus, Figure (1) can be interpreted as the plot of g_t corresponding to the discreet capital levels, $\bar{v}_0 > \bar{v}_1 > ... > \bar{v}_{T-1}$

[19] The simplest way to solve (15) is to guess and verify that the solution has the shape of $V_t'(v_t) = B\beta^t + \psi$.

[20] See details in the appendix in the proof of Theorem 1.

[21] The idea that financially constrained agents require a premium for sacrificing future investment possibilities is not new. It has gained popularity in corporate risk management (e.g. Froot et al. 1993, Holmström and Tirole, 2001) and in investment theory (e.g. Dixit and Pindyck, 1994) as well. All these works emphasized that assets should be more valuable if they provide free cash flow in states when investment possibilities are better. This point was also made in Gromb and Vayanos (2002) in relation to the fact that arbitrageurs might not invest fully in the arbitrage expecting better opportunities in the future. Our additional contribution is showing the implication of this idea to the dynamic pattern of this premium and to the change of expected profit and risk over time.

The idea is also closely related to the hedging component of asset demand identified in Merton (1971).

and $\bar{v}_{T+\tau} = 0$ for $\tau \geq 0$. Hence, both the aggregate level of capital and time for given initial capital can be seen as the unique state variable of the system. Therefore, we can write the value function as

$$V_t(v_t) = V(\bar{v}_t, v_t) = V'(\bar{v}_t)\, v_t$$

or

$$V'_t = V'(\bar{v}_t)\,.$$

For deriving the properties of the marginal value function, we establish the following useful lemma which highlights the role of $\bar{v}_t$ as a state variable. We look at the future conditional gap path $\{g_t\}_{t=\tau}^{\infty}$ from an arbitrary time-point τ, and we are interested in what happens with this gap path as the aggregate level of capital at period τ $\bar{v}_\tau$ changes.

Lemma 1 *Each element of the future gap path from period τ on, $\{g_t\}_{t=\tau}^{\infty}$, is non-increasing in the aggregate level of capital at τ, $\bar{v}_\tau$. Furthermore, the period, T_τ, when arbitrageurs ran out of capital is non-decreasing in $\bar{v}_\tau$.*

Proof. The proof is part of the proof of Theorem 1 in the appendix. ■

To see the intuitive content of this lemma, let us choose $\tau = 0$ and let us suppose that we slightly decrease the initial capital level $\bar{v}_0$. It turns out that for a small change T remains unchanged, but g_{T-1} increases. This, because of the indifference condition, implies an increase in all g_t for $t < T - 1$. If we decrease $\bar{v}_0$ further, sooner or later g_{T-1} reaches g^*, its maximum value. But in this case we are in an equilibrium where the new last period T' decreased to $T' = T - 1$.

This lemma implies that past losses correspond to a larger price gap similarly to the model of Xiong (2001), Gromb and Vayanos (2002) or Shleifer and Vishny (1997). However, the mechanism is very different. In Xiong (2001), when arbitrageurs suffer capital losses as a result of exogenous shocks, they choose to cut back their positions because of the wealth effect. In Shleifer and Vishny (1997) and in Gromb and Vayanos (2002) the same happens because of more stringent budget constraint or margin accounts. The gap widens because positions are smaller. In our model, there is no causality from past losses to smaller individual positions. Individual arbitrageurs are indifferent whether to invest more or to invest less after a loss and typically their constraint does not bind. The comovement is simply a result of arbitrageurs' capital allocation decision in the aggregate. In the aggregate they take larger positions in early periods and save less for later periods because otherwise the gap in early periods would be too wide, i.e., investing in early periods would result in extra profit which would be inconsistent with an equilibrium.

In the next subsection, we will analyze the pattern of expected profit and risk. The different interpretations of the value function above will help our intuition.

4.2 Expected profit and risk

First, we show that the marginal value function is increasing in the length of the window and decreasing in the aggregate capital level.

Lemma 2 *The premium and profitability at t, V_t', is decreasing in the aggregate capital level $\bar{v}_t$ and increasing in t.*

Proof. The proof is in the appendix. ■

We saw that the premium is the expected value of saving a dollar until period $T-1$. Hence, the premium has to increase with time, because as period $T-1$ is getting closer, the chance that the window closes before $T-1$ is decreasing and the strategy of waiting till $T-1$ gets more valuable. Similarly, if the aggregate capital decreases, this will not only increase the possibility that the strategy of waiting is profitable by decreasing T, but will also increase the reward by increasing g_{T-1}.

The fact that the premium increases as the gap increases and arbitrageurs suffer losses is consistent with practitioners' impression[22] that the "risk-appetite" of investors decreases after recent losses. However, this is not a result of changing risk aversion, but that of changing future opportunities endogenously determined in equilibrium. This is also consistent with the empirical literature on time-varying risk-aversion (e.g., Engle and Rosenberg, 2001, Campbell, 1996) to a certain extent.[23] For example Engle and Rosenberg (2002) estimate the pattern of risk aversion from the change of the empirical pricing kernel of observed option prices. Consistently with our model, they show that risk-aversion changes together with the credit spread. They connect this result to the habit-formation explanation of changing marginal utilities of Campbell and Cochrane (1999). We argue that it can also be related to our time-varying, liquidity based premium term at least in periods of distressed markets where capital limits are close to binding.

It is also interesting to discuss the same result from the point of view of profitability. It is intuitive that as the aggregate capital in the market increases, so does the competition among arbitrageurs. With more capital, arbitrageurs are able to push down the profit level further. Interestingly, this decline of profit level does not occur solely by reducing the price discrepancy. In fact, the gap can never be fully eliminated, because then no one would trade as losses are always possible. Instead, competition transforms the gap process: it introduces potential losses and changes potential gains. The proportion of gains to losses, i.e., the premium, will determine expected profit. This proportion decreases with larger competition. This is why the profitability of the market decreases with more available capital.

There is another interesting comparison with previous models. Both Shleifer and Vishny (1997) and Xiong (2001) emphasize that arbitrageurs may reduce their positions when arbitrage is the most profitable, i.e., when the gap is the widest. We also have the analogous property that aggregate positions decrease together with the increase of the profitability of the market. However, our observation relies on a different argument. First of all, we distinguish profitability from the size of the gap. We highlight that the profitability of the arbitrage depends on the proportion of potential gains (the size of

[22] See e.g. Dungey et al. (2003), Kumar and Persaud (2001), Gai and Vause (2005). However, these papers see the change of risk-appetite as the cause of liquidity crises. In this paper we argue that changing risk-premium is the symptom.

[23] There is the caveat that this literature refers to the relative-risk aversion of market participants. Our premium term is more closely related to the absolute risk-aversion of our arbitrageurs. It is not clear whether the sensitivity of the premium to the level of capital is large enough to cause a corresponding change in the relative measure as well.

the gap) to potential losses (the potential widening of the gap), not the size of the gap alone. Secondly, just as we argued before, our result is not based on an amplification argument of an exogenous shock. It is a consequence of individually optimal timing decisions. In equilibrium, smaller probability events have to correspond to more profitable investment decisions otherwise arbitrageurs are not willing to save capital for these future opportunities. Therefore, longer windows have to correspond to more profitable investment decisions, simply because they happen with smaller probability.

Now, we turn to the analysis of risk in arbitrage trades. We argued above that there is an analogy between the marginal value function and risk-premium as long as we consider the size of the gap and the profitability of the market. Below we show that this analogy breaks down if we consider risk. The premium based on sacrificed future opportunities is typically non-monotonic in the risk of arbitrage trade. This fact can be the basis of future empirical work to distinguish risk-based mechanisms from our mechanism based on intertemporal capital allocation.

We think of the risk of arbitrage as the size of the potential loss in the next period, $(g_{t+1} - g_t)$, corresponding to the aggregate level of capital $\bar{v}_t$ at that period. More precisely, this is the downside risk on each unit for arbitrageurs taking a position at 0.[24] We show that for sufficiently large $\bar{v}_t$ the downside risk is typically increasing in t, i.e., increases as $\bar{v}_t$ decreases along the equilibrium path. In contrast, for small $\bar{v}_t$ it is deceasing in t, i.e., decreases as $\bar{v}_t$ decreases along the equilibrium path. Consequently, for small $\bar{v}_t$ the premium is increasing (by Lemma 2), while the downside risk is decreasing as arbitrageurs make losses. The non-monotonicity is a result of the typical S-shape pattern of the conditional gap path shown in Figure 1. The reason for its shape is closely related to the property of the marginal value function that it is increasing t. From the first order condition, we can see that the change of the gap $(g_{t+1} - g_t)$ is positively related to the size of the gap in the previous period, g_t, and inversely related to the size of the premium $V'(\bar{v}_{t+1})$. For small t, when $V'(\bar{v}_{t+1})$ is small, the effect of larger g_t dominates and $(g_{t+1} - g_t)$ increases. However, for larger t, the effect of increasing $V'(\bar{v}_{t+1})$ can dominate and $(g_{t+1} - g_t)$ can shrink.

We establish the formal result in the following lemma for the limiting equilibrium defined in Theorem 2.

Lemma 3 *In the limiting equilibrium defined in 2, if q or $\bar{v}_0$ is sufficiently large that*

$$\frac{\sqrt{q}\left(1-\sqrt{q}\right)}{1-q} g^* > g_0^{\lim},$$

then there exists a critical level of $\bar{v}^{\inf} = \bar{v}^{\inf}(\bar{v}_0)$ that if $\bar{v}_t > \bar{v}^{\inf}$ then the downside risk, $g_{t+1} - g_t$ strictly monotonically increases as $\bar{v}_t$ decreases along the equilibrium path, while if $\bar{v}_t \leq \bar{v}^{\inf}$, it strictly monotonically decreases.

Proof. The proof is in the Appendix. ■

[24] We already know from Theorem (1) and Lemma 1 that the upside risk, g_t, is increasing in t, thus decreasing in the aggregate level of capital $\bar{v}_t$.

In Figure 3, we illustrate the intuition behind this result As we argued above the analysis of the size of $(g_{t+1} - g_t)$ at a given $\bar{v}_t$ is equivalent with the analysis of $(g_1 - g_0)$ at a given $\bar{v}_0$. In Figure 3, we plot three curves corresponding to different aggregate capital level, $\bar{v}_0$. We see that downside risk in early periods is largest when the aggregate capital is at an intermediate level. In particular, we can observe the following pattern. When $\bar{v}_0$ is low arbitrageurs cannot take very large positions, potential losses are relatively small. Intuitively, arbitrageurs do not have enough capital to move the gap relative to its autarchy level, g^*, so even if they liquidate, the effect will not be devastating. The large gap combined with small losses is very profitable reflecting the large required premium because of the excellent future opportunities. However, in the intermediate range, large losses occur with relatively large probability. In this range, expected profits are relatively large but they are matched with large possible losses reflecting the intermediate premium. In contrast, in markets with high level of capital, the price gap will be kept in a low level even in long windows. This is the world of small profits and small losses. Large losses are possible, but happen relatively rarely.

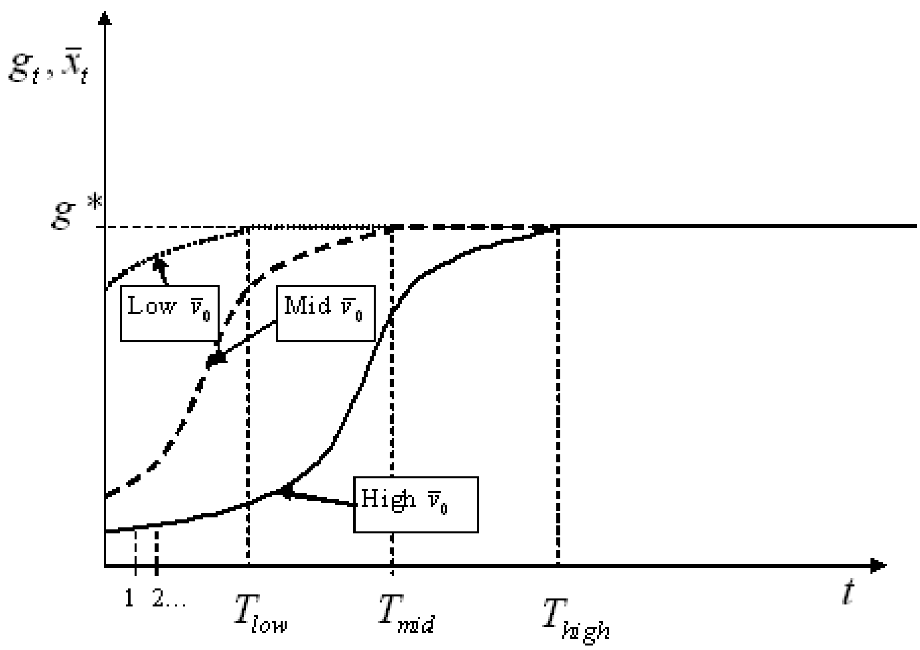

Figure 3: The effect of aggregate speculative capital, $\overline{v}_0$, on the gap path g_t.

5 Robustness

The mechanism of our model is based on the fact that arbitrageurs with limited capital invest only if they are compensated for the sacrifice of future opportunities. As future opportunities endogenously change with time, so does the premium. This effect is expected to be present under quite general assumptions. However, its implications might change with the framework. We consider three of our assumptions particularly important for our specific results. The first one is that the duration of the window of arbitrage opportunity is uncertain and, in particular, that it can be arbitrarily long. An example for the departure from this assumption is Gromb and Vayanos (2002). They assume a window with a fixed length i.e. the gap disappears in an exogenously fixed period. They show – in contrast with our result – that the gap path will typically decrease in that case. The other critical assumption is that arbitrageurs take both prices and the probability of convergence as given. Zigrand (2004) presents a model where there is imperfect competition among arbitrageurs, while Abreu and Brunnermeier (2002,2003) analyze the case where arbitrageurs are strategic and the time of convergence is determined in equilibrium. The third important assumption is that there is no capital inflow into the market during the window of arbitrage opportunity. In this section, we focus on the implications of relaxing this assumption. We argue that even if capital supply is partially flexible, our main intuition remains unchanged.

We present the draft of two scenarios. In the first one, we assume that as the market gets more profitable there are new arbitrageurs with new capital entering the market. The second scenario is based on Kondor (2005), which endogenizes the capital flow of arbitrageurs through a fully specified extension of career concerns.

5.1 First scenario: Reaching for yield

In this scenario, just like in our model, individual arbitrageurs cannot get extra funds during the window. However, there is an external pool of investors who are faced with different costs or outside options. In each period they decide whether to enter the arbitrage market for the given prices and future opportunities. As the arbitrage market gets more profitable, more investors decide to join. This is consistent with the anecdotal evidence that fund managers enter more risky markets when safer opportunities do not provide sufficient profit, i.e., they "reach for yield". Here, we do not model the decision of investors explicitly. Instead, let us assume that there is a threshold level of profitability $\widetilde{V}$ such that if the marginal value function $V'(\bar{v}_t)$ exceeds that level then new arbitrageurs enter instantaneously with the aggregate capital level of Δv.

We first analyze the case where $\widetilde{V}$ is fixed and it is common knowledge, i.e., incumbents expect the new entrants to enter. In this case, as long as the supply of capital is not too flexible in the sense that $V'(\Delta v) \geq \widetilde{V}$, the structure of the equilibrium remains the same. The intuitive content of this constraint is that if only new entrants were on the market then the profitability would be higher than this threshold. Let us suppose that the group of arbitrageurs who enter in period 0 have initial capital of $\bar{v}_0^0$. We will argue that in this case the equilibrium will be the same as it would be in the benchmark model with initial capital $\bar{v}_0 = \bar{v}_0^0 + \Delta v$. The argument goes as follows. Let $\widetilde{t}$ be the first period where $V'\left(\bar{v}_{\widetilde{t}}\right) \geq \widetilde{V}$ in the proposed equilibrium. On the equilibrium gap path all arbitrageurs who are in the market are indifferent to the time of their investment across periods $0...T-1$. It is also true that there is an average portfolio $\{\bar{x}_t\}_{t=0}^{T-1}$ which supports the equilibrium gap path without violating the collateral constraint:

$$\begin{aligned} \bar{v}_0^0 + \Delta v &= \sum_{t=0}^{T-1} \bar{x}_t \left(g_{t+1} - g_t + m \right) \\ g_t &= \left(\bar{x}_t, g_{t+1} \right). \end{aligned}$$

The only question is whether there are individual portfolios $\left\{ x_t^0 \right\}_{t=0}^{T-1}$ and $\{\widetilde{x}_t\}_{t=\widetilde{t}}^{T-1}$ for incumbents and new entrants respectively such that $\bar{x}_t = \bar{x}_t^0$ for $t = 0, ..., \widetilde{t} - 1$ and $\bar{x}_t = \bar{x}_t^0 + \overline{\widetilde{x}}_t$ for $t = \widetilde{t}, ..., T - 1$, where the upper bar denotes the average portfolios of the particular group. It turns out that it is easy to construct such portfolios. For example, if a measure $\mu = \frac{\bar{v}_{\widetilde{t}} - \Delta v}{\bar{v}_0^0}$ of incumbents save their capital until period $\widetilde{t}$ then all other arbitrageurs can choose $\left\{ x_t^0 = \bar{x}_t \right\}_{t=0}^{\widetilde{t}-1}$ and stay inactive after period $\widetilde{t} - 1$ as

$$(1 - \mu) \, \bar{v}_0^0 = \sum_{t=0}^{\widetilde{t}-1} \bar{x}_t \left(g_{t+1} - g_t + m \right).$$

Furthermore, let $\gamma = \frac{\mu \bar{v}_0^0}{\bar{v}_{\widetilde{t}}}$. Then the μ measure of incumbents who saved their capital until $\widetilde{t}$ can choose the portfolio $\left\{ x_t^0 = \gamma \bar{x}_t \right\}_{t=\widetilde{t}}^{T-1}$ as

$$\mu \bar{v}_0^0 = \gamma \bar{v}_{\widetilde{t}} = \sum_{t=\widetilde{t}}^{T-1} \gamma \bar{x}_t \left(g_{t+1} - g_t + m \right),$$

and all new entrants can choose the portfolio $\{\widetilde{x}_t = (1 - \gamma) \, \bar{x}_t\}_{t=\widetilde{t}}^{T-1}$ as

$$\Delta v = (1 - \gamma) \, \bar{v}_{\widetilde{t}} = \sum_{t=\widetilde{t}}^{T-1} (1 - \gamma) \, \bar{x}_t \left(g_{t+1} - g_t + m \right).$$

As incumbents are indifferent along the whole path these portfolios will be optimal for each of them. With the same logic we can create as many groups of arbitrageurs as we like with different thresholds of entry. This changes the equilibrium only in that it increases the initial capital level of our benchmark equilibrium.

We can also consider the possibility when incumbent arbitrageurs do not expect the entries of new arbitrageurs, i.e., it is a zero probability event. For example, it may correspond to a situation when returns on less risky investments decrease as an effect of an unexpected policy shock, i.e., the outside option of potential entrants decrease. In our model, this would simply result in a jump from the original equilibrium gap path to a less profitable one with lower gap corresponding to the higher level of capital at the moment of the surprise. This would explain the comovement of credit spreads and short term interest rates observed (see, e.g., Gerlach, 2005), if we consider the short term interest rate a proxy for the outside option of potential entrants.

5.2 Second scenario: Tournament for funds

In Kondor (2005), we endogenize the initial capital of arbitrageurs and investigate the effect of career concerns on the equilibrium. The intuitive starting point for the extension is that finding an arbitrage possibility is not easy: arbitrageurs need inspiration to be successful. We allow for heterogeneity across arbitrageurs, assuming that some arbitrageurs get inspired and find the arbitrage possibility, while others do not find any. Importantly, we also endogenize the financial constraint of arbitrageurs by introducing investors who entrust their capital to arbitrageurs. However, investors are not able to observe the level of inspiration of arbitrageurs directly nor can they monitor their activity. They observe only realized profit. Hence, non-inspired arbitrageurs might try to gamble with negative expected value investments to pretend that they are inspired. We focus on two arbitrage possibilities arising consecutively: two windows. Inspired arbitrageurs operate on the same market in the two windows, but investors can observe profits from the first window and update their beliefs on the level of inspiration of arbitrageurs. Based on this learning effect, they can reallocate their capital and fire arbitrageurs who do not seem to be inspired. Naturally, the competition of hedge funds for capital distorts their strategies. This in turn changes the characteristics of the premium which influences the short-term price dynamics and the relative profitability of different arbitrage strategies.

The main lesson from the analysis is that arbitrageurs, as an effect of career concerns, will care both about the distribution of profit from their strategies, and about the distribution of other arbitrageurs' strategies. This is in contrast with the benchmark model, where arbitrageurs were interested only in the expected profit of each strategy. We identify two main effects. Firstly, because of the "reward-for-success effect", the premium is smaller in earlier periods compared to the benchmark model. The reason is that in the benchmark equilibrium if an arbitrageur distorted her strategy towards more extensively on shorter windows, she would be among the most successful ones more frequently even if her expected profit was the same. This is so, shorter windows are more probable than longer ones. Hence, in the new equilibrium the premium has to decrease in earlier periods. The second main effect is the "publicity-effect" which implies smaller premium for less popular strategies. Because of the tournament-like set up of our extension, these strategies attract much more capital if successful. Hence, they have to provide less expected profit, i.e., smaller premium.

6 Conclusion

In this paper we present an analytically tractable general equilibrium model of dynamic arbitrage. In our model arbitrage opportunities arise because of a temporary pressure on local demand curves of two very similar assets traded in segmented markets. The temporary demand pressure is present for an uncertain, arbitrarily long time span, but disappears in finite time with probability one. Risk-neutral arbitrageurs can take positions in both local markets, and have to decide how to allocate their limited capital across uncertain future arbitrage opportunities. This allocation – together with the uncertain duration of the local demand pressure – determines the future distribution of the price gap between the two assets. Hence, the individually optimal intertemporal allocation of capital and the distribution of future arbitrage opportunities are determined simultaneously in equilibrium.

We argued that competition among arbitrageurs with limited capital transforms the dynamic properties of arbitrage prices. In particular, even if the fundamental process is riskless in the sense that the price gap could never widen, arbitrageurs' activity introduces potential losses. Interestingly, the intuition behind this result is of the same type as the text-book argument of why unconstrained arbitrageurs eliminate price discrepancies. There, price discrepancies cannot exist, because they are so attractive for arbitrageurs that they would eliminate them. We showed that when arbitrageurs are financially constrained, there cannot be any time period when there are no potential losses, because this would be a too attractive arbitrage opportunity. Arbitrageurs would increase their positions in these periods relative to other periods, which would create potential losses.

Importantly, we highlighted that our mechanism is not based on amplification of an external shock. We presented the limit equilibrium where the cost of trading is negligible, thus, arbitrageurs total losses are created endogenously. In a calibrated example, we showed that even in this case, arbitrageurs lose most of their capital with positive probability in a relatively short time.

We showed that the price gap in our model are determined by the development of a premium term, which arbitrageurs require as a compensation for the sacrifice of future arbitrage opportunities. This premium is the option value of saving a unit for the next period. The idea is that when the price gap widens, arbitrageurs with existing positions lose money exactly when the market is more profitable. Therefore, they require a larger premium to invest in the arbitrage, instead of keeping their capital in liquid instruments when the price gap is wider. Arbitrageurs are more averse to losing money in these times, not because of their preferences, but because they have limited capital and these are the situations when future opportunities are better. The higher premium validates the higher price gap. We derived implications of the time-varying premium term on the pattern of expected profit and on the risk of arbitrage trades. We showed that the premium is increasing after recent capital losses but it can be non-monotonic in the risk of arbitrage trades.

The simplicity of our framework provides the potential of wide applicability to different problems related to limited arbitrage. In Kondor (2005), we demonstrate this potential by building an extension of career concerns on top of our structure. As future work we consider the applicability of our model to a multi-asset set up to analyze contagion across markets and the effects of flight-to-quality and flight-to-liquidity in times of market depression. We believe that our mechanism will shed more light on these issues.

References

[1] Abreu, D., Brunnermeier, M., 2002. Synchronization risk and delayed arbitrage, *Journal of Financial Economics*, 66, 341-360.

[2] Abreu, D., Brunnermeier, M., 2003. Bubbles and crashes, *Econometrica*, 71(1), 173-204.

[3] Basak, S., Croitoru, B., 2000. Equilibrium mispricing in a capital market with portfolio constraints, *Review of Financial Studies* 13, 715-748.

[4] Bernardo, A., Welch, I., 2004. Liquidity and financial market runs, *Quarterly Journal of Economics* 119(1), 135-158.

[5] Campbell, J., 1996. Consumption and the stock market: interpreting international experience, *Swedish Economic Policy Review,* 3, 251-299.

[6] Campbell, J., Cochrane, J., 1999. By force of habit: a consumption based explanation of aggregate stock market behavior, *Journal of Political Economy* 107, 205-251.

[7] Danielsson, J., Shin, H.S., 2002. Endogenous risk, LSE, mimeo.

[8] Danielsson, J., Shin, H.S., Zigrand, J-P., 2004. The impact of risk regulation on price dynamics, *Journal of Banking and Finance*, 28, 1069-108.

[9] Dixit, A. and Pindyck, R., 1994. *Investment under uncertainty*, Princeton: Princeton University Press.

[10] Dungey, M.; Fry, R.; Gonzalez-Hermosillo, B. ; Martin, V., 2003. Characterizing global investors' risk appetite for emerging market debt during financial crises, IMF Working Papers: 03/251.

[11] Edwards, F., 1999. Hedge funds and the collapse of long-term capital management, *Journal of Economic Perspectives* 13, 189-210.

[12] Engle, R. F., Rosenberg, J. V., 2002. Empirical pricing kernels, *Journal of Financial Economics*, 64, 341-372.

[13] Froot, K. A. Dabora, E. M., 1999. How are stock prices affected by the location of trade?, *Journal of Financial Economics*, 53(2), 189–216.

[14] Froot, K., Scharfstein, D., Stein, J., 1993. Risk management: coordinating corporate investment and financing policies, *Journal of Finance*, 48, 1629-1658.

[15] Gerlach, S., 2005. Discussion for Favero, Pagano and von Thadden, Valuation, liquidity and risk in government bond markets, presented at Quatrième Journée de la Fondation Banque de France Paris, 4 February 2005. http://www.banque-france.fr/fr/fondatio/telechar/Gerlach.pdf.

[16] Gai, P., Vause, N., 2005. Measuring investors' risk appetite, Bank of England, mimeo.

[17] Gromb, D., Vayanos, D., 2002. Equilibrium and welfare in markets with financially constrained arbitrageurs, *Journal of Financial Economics*, 66, 361-407.

[18] Holmström, B., Tirole, J., 2001. LAPM: A liquidity-based asset pricing model, *Journal of Finance*, 56(5), 1837-1867.

[19] Kondor, P., 2005. Career concerns and dynamic arbitrage, LSE, mimeo.

[20] Kumar, M. S., Persaud, A., 2001. Pure contagion and investors shifting risk appetite: analytical issues and empirical evidence, IMF Working Papers: 01/134.

[21] Kyle, A. S., Xiong, W., 2001. Contagion as a wealth effect, *Journal of Finance*, 56, 1401-1440.

[22] Lamont, O. A., Thaler, R. H., 2003. Anomalies: the law of one price in financial markets, *Journal of Economic Perspectives*, 17(4), 191–202.

[23] Liu, J., Longstaff, F., 2004. Losing Money on Arbitrage: Optimal Dynamic Portfolio Choice in Markets with Arbitrage Opportunities, *Review of Financial Studies*, 17(3), 611-641.

[24] Loewenstein, R., 2000. *When genius failed: the rise and fall of long-term capital management*, New York, Random House.

[25] Lucas, R. E. Jr., 1978. Asset prices in an exchange economy, *Econometrica*, 46(6), 1426-1445.

[26] MacKenzie, D, 2003. Long-Term Capital Management and the sociology of arbitrage, *Economy and Society* 32(3), 349-380.

[27] Merton, Robert C., 1971. Optimum consumption and portfolio rules in a continuous-time model, *Journal of Economic Theory*, 3, (December):373-413. Reprinted in Merton (1992, Ch. 5).

[28] Merton, Robert C., 1992. *Continuous-time finance*. Rev. ed. Oxford, U.K.: Basil Blackwell.

[29] Morris, S., Shin, H. S., 2004. Liquidity black holes, *Review of Finance*, 8, 1-18.

[30] Nieuwerburgh, S. van, Veldkamp, L., 2005. Information acquisition and portfolio under-diversification, NYU, mimeo.

[31] Plantin, G., Shin, H. S., 2005. "Weight of Money" Pricing and Speculative Dynamics, LSE, mimeo.

[32] Shleifer, A., Vishny, R., 1997. The limits of arbitrage. *Journal of Finance* 52, 35-55.

[33] Xiong, W., 2001. Convergence trading with wealth effects. *Journal of Financial Economics* 62, 247-292.

[34] Zigrand, J-P., 2004. General equilibrium analysis of strategic arbitrage, *Journal of Mathematical Economics*, 40(8), 923-952.

Appendix

A.1 Traders

A.1.1 Traders' problem in normal times and the autarchy prices

Let us assume that in normal times in each period there is a q probability that the income state switches back to a window. In particular, let us assume that with probability $\frac{q}{2}$ it switches to a window described in the main text where $\omega_{t,i} = \omega^i$, and with probability $\frac{q}{2}$ it switches to the mirror image of this window where $\omega_{t,A} = \omega^B$ and $\omega_{t,B} = \omega^A$. Hence, we suppose a transition matrix

$\Pr\left(\left(\omega_{t,A}, \omega_{t,B}\right) \mid \left(\omega_{t-1,A}, \omega_{t-1,B}\right)\right) =$

	ω^-, ω^+	$0,0$	ω^+, ω^-
ω^-, ω^+	q	$1-q$	0
$0,0$	$\frac{q}{2}$	$1-q$	$\frac{q}{2}$
ω^+, ω^-	0	$1-q$	q

This symmetric structure ensures that arbitrageurs do not have any motivation to trade in normal times, which is consistent with the main text. Hence, in normal times traders solve the problem

$$\begin{aligned} & \max E_{R_t}\left[(1-q)\, u\left(e+\theta_{t,i}\left(R_t\right)\right)\right] + \\ & + E_{R_t}\left[\frac{q}{2}\left(u\left(e+\theta_{t,i}\left(R_t + p^w_{t+1,A} - p^n\right)\right) + u\left(e+\theta_{t,i}\left(R_t + p^w_{t+1,B} - p^n\right)\right)\right)\right] \\ i &= A, B. \end{aligned} \tag{17}$$

where we exploited the symmetry in our structure, i.e., that in both islands in one of the possible windows the future price will be $p^w_{t+1,A}$ and in the other window it will be $p^w_{t+1,B}$. In the main text we did not distinguish p^n in for the case of autarchy and when arbitrageurs are present. This is consistent with a set up where arbitrageurs do not return after the end of the initial window, i.e., that $p^w_{t+1,i} = p^*_i$ in problem (17). Instead we could have assumed that in all subsequent windows an other set of arbitrageurs arrive with a given initial capital. This would have resulted in a different p^n for the autarchy case, but otherwise the analysis would have been the same. So under the first scenario, substituting $p^w_{t+1,i} = p^*_i$ and $\theta = 0$ to the first order condition results in the simple expression of

$$p^n = \frac{1}{2}\left(p^*_A + p^*_B\right) + \frac{1}{q} E\left(R_t\right).$$

Similarly, the first order conditions of traders problem in a window with the same substitution implies

$$p^*_A - p^n = \frac{E\left(u'\left(e+\omega^A R_t\right) R_t\right)}{(1-q)\, E\left(u'\left(e+\omega^A R_t\right)\right)} \tag{18}$$

$$p^*_B - p^n = \frac{E\left(u'\left(e+\omega^B R_t\right) R_t\right)}{(1-q)\, E\left(u'\left(e+\omega^B R_t\right)\right)} \tag{19}$$

after straightforward manipulations. It is straightforward to show that these three equations determine unique and well defined p^*_A, p^*_B and p^n values.

A.1.2 Example 1: the CARA-symmetric framework

In the CARA-symmetric framework proposed in Example 1, trader A choosing position θ_t maximizes

$$
\begin{aligned}
&-E_{p_{t+1,A}-p^n,R_t}\left(\exp-\left(\left(e-\omega R_t+\theta_t\left(R_t+p_{t+1,A}-p^w_{t,A}\right)\right)\right)\right)=\\
&=-\exp-\alpha\left(e-\theta_t p^w_{t,A}+x_t p^n\right)E_{p_{t+1}-p^n}\left(\exp-\alpha\left(\theta_t\left(p_{t+1,A}-p^n\right)\right)\right)E_{R_{tz}}\left(\exp-\alpha\left(\left(-\omega+\theta_t\right)R_t\right)\right),
\end{aligned}
$$

which is equivalent with maximizing

$$\alpha\left(e-\theta_t p^w_{t,A}+\theta_t p^n\right)-\ln\left((1-q)+q\exp-\alpha\left(\theta_t\left(p^w_{t+1,A}-p^n\right)\right)\right)-\frac{1}{2}\alpha^2\left(-\omega+\theta_t\right)^2\sigma^2.$$

This, and the analogous problem for trader B gives the inverse demand curves

$$p^w_{t,A}-p^n=p^n-p^w_{t,B}=\frac{g_t}{2}=p^w\left(0,\frac{g_{t+1}}{2}\right)=\frac{\frac{g_{t+1}}{2}}{\left(\frac{(1-q)}{q}\exp\alpha\left(\theta_t\frac{g_{t+1}}{2}\right)+1\right)}-\alpha\left(-\omega+\theta_t\right)\sigma^2, \tag{20}$$

where $g_t=p^w_{t,A}-p^w_{t,B}$. The autarchy prices are given by

$$p^*_A-p^n=p^n-p^*_B=\frac{g^*}{2}=p^w\left(0,p^*_A-p^n\right)=\frac{\alpha\omega\sigma^2}{1-q}>0.$$

Because of $p^w_{t,A}-p^n=p^n-p^w_{t,B}$ for any θ_t, $\bar{\theta}=\bar{\theta}\left(p^w_{t+1,A},p^w_{t+1,B}\right)$ the position which eliminates the gap between current prices for given future prices, is given as the unique and positive solution of

$$\frac{\frac{g_{t+1}}{2}}{\left(\frac{(1-q)}{q}\exp\left(\alpha\left(\bar{\theta}\frac{g_{t+1}}{2}\right)\right)+1\right)}=\alpha\left(-\omega+\bar{\theta}\right)\sigma^2. \tag{21}$$

It is well defined and positive, because the left hand side is a positive and monotonically decreasing function in $\bar{\theta}$, while the right hand side is linear increasing in $\bar{\theta}$ with a negative intercept. Hence, Assumption 1 is satisfied. As the domain of right hand side of (20) contains R^+ for both arguments, Assumption 2 is also satisfied. Both demand functions are clearly downward sloping, so Assumption 3 holds, and we only have to check that they are increasing in the future price to assure that Assumption 4 also holds. We will show that a sufficient condition for this is

$$1>\alpha\bar{\theta}\left(p^*_A,p^*_B\right)\frac{g^*}{2}. \tag{22}$$

First note that $\bar{\theta}\left(p^w_{t+1,A},p^w_{t+1,B}\right)$ is increasing in g_{t+1} if

$$1>\alpha\bar{\theta}\left(p^w_{t+1,A},p^w_{t+1,B}\right)\frac{g_{t+1}}{2} \tag{23}$$

by the implicit function theorem. Hence, by condition (22), $\bar{\theta}\left(p^w_{t+1,A},p^w_{t+1,B}\right)$ is increasing in g_{t+1} at $g_{t+1}=g^*$. It means that condition (22) implies (23) for all $0\leq g_{t+1}\leq g^*$, i.e., $\bar{\theta}\left(p^w_{t+1,A},p^w_{t+1,B}\right)$

is increasing for any $0 \leq g_{t+1} \leq g^*$ and $\bar{\theta}^{\max} = \bar{\theta}\left(p_A^*, p_B^*\right)$. But (23) is a sufficient condition for $\frac{\partial p_i^w\left(\theta_{t,i}, p_{t+1,i}^w\right)}{\partial p_{t+1,i}^w} > 0$ in the symmetric framework as

$$\frac{\partial p_i^w\left(\theta_{t,i}, p_{t+1,i}^w\right)}{\partial p_{t+1,i}^w} = \frac{\partial p^w\left(\theta_t, \frac{g_{t+1}}{2}\right)}{\partial\left(\frac{g_{t+1}}{2}\right)} = \frac{1+\frac{(1-q)}{q}\exp\alpha\left(\theta_t\frac{g_{t+1}}{2}\right)\left(1-\alpha\frac{g_{t+1}}{2}\theta_t\right)}{\left(\frac{(1-q)}{q}\exp\alpha\left(\theta_t\frac{g_{t+1}}{2}\right)+1\right)}.$$

By the implicit function theorem $\bar{\theta}\left(p_A^*, p_B^*\right)$ is increasing in ω. Furthermore, g^* is also increasing in ω and in σ^2. Hence, condition (22) can be equivalently satisfied by a small ω, a small σ^2 or a small α.

A.1.3 Example 2: CRRA framework

First we show the following lemma.

Lemma 4 *If the relative risk-aversion is less than or equal to 1,* $\left(e+\omega^i R_t\right) > 0$ *for all* R_t*, then the inverse demand function* $p_{t,i}^w = p_i^w\left(\theta, p^n, p_{t+1}^w\right)$ *is downward sloping for any fixed* p_{t+1}^w *and increasing in* p_{t+1}^w.

Proof. The first order condition for exercise (1) is

$$E_{R_t}\left[(1-q)\frac{\partial u\left(h_{t+1,i}^n\right)}{\partial h_{t+1,i}^n}\left(R_t+p^n-p_{t,i}^w\right)+q\frac{\partial u\left(h_{t+1,i}^w\right)}{\partial h_{t+1,i}^w}\left(R_t+p_{t+1,i}^w-p_{t,i}^w\right)\right]=0$$

where the variables

$$\begin{aligned} h_{t+1,i}^n &= e+\omega^i R_t+\theta_{t,i}\left(R_t+p^n-p_{t,i}^w\right) \\ h_{t+1,i}^w &= e+\omega^i R_t+\theta_{t,i}\left(R_t+p^n-p_{t,i}^w\right) \end{aligned}$$

are consumption levels in normal times and in windows respectively. Hence, if the conditions of the lemma hold then

$$\begin{aligned} \frac{\partial f.o.c}{\partial p_t} &= E_{R_t}\left[\begin{array}{c} (1-q)\left(-\frac{\partial u\left(h_{t+1,i}^n\right)}{\partial h_{t+1,i}^n}-\frac{\partial^2 u\left(h_{t+1,i}^n\right)}{\partial^2 h_{t+1,i}^n}\theta_{t,i}\left(R_t+p^n-p_{t,i}^w\right)\right)+ \\ +q\left(-\frac{\partial u\left(h_{t+1,i}^w\right)}{\partial h_{t+1,i}^w}-\frac{\partial^2 u\left(h_{t+1,i}^w\right)}{\partial^2 h_{t+1,i}^w}\theta_{t,i}\left(R_t+p_{t+1,i}^w-p_{t,i}^w\right)\right) \end{array}\right]= \\ &= E_{R_t}\left[\begin{array}{c} (1-q)\left(-\frac{\partial u\left(h_{t+1,i}^n\right)}{\partial h_{t+1,i}^n}-\frac{\partial^2 u\left(h_{t+1,i}^n\right)}{\partial^2 h_{t+1,i}^n}h_{t+1,i}^n+\left(e+\omega^i R_t\right)\frac{\partial^2 u\left(h_{t+1,i}^n\right)}{\partial^2 h_{t+1,i}^n}\right)+ \\ +q\left(-\frac{\partial u\left(h_{t+1,i}^w\right)}{\partial h_{t+1,i}^w}-\frac{\partial^2 u\left(h_{t+1,i}^w\right)}{\partial^2 h_{t+1,i}^w}h_{t+1,i}^w+\left(e+\omega^i R_t\right)\frac{\partial^2 u\left(h_{t+1,i}^w\right)}{\partial^2 h_{t+1,i}^w}\right) \end{array}\right]<0 \end{aligned}$$

as $-\frac{\partial u(h)}{\partial h}-\frac{\partial^2 u(h)}{\partial^2 h}h \leq 0$ if $\frac{-\frac{\partial^2 u(h)}{\partial^2 h}h}{\frac{\partial u(h)}{\partial h}} \leq 1$ and $\left(e+\omega^i R_t\right)\frac{\partial^2 u(h)}{\partial^2 h} < 0$. For the effect of future price $p^i_{t+1,i}$,

$$\begin{aligned}
\frac{\partial f.o.c}{\partial p_{t+1,i}} &= E_{R_t}\left[q\left(\frac{\partial^2 u\left(h^i_{t+1,i}\right)}{\partial^2 h^i_{t+1,i}}\left(R_t+p^w_{t+1,i}-p^w_{t,i}\right)\theta_{t,i}+\frac{\partial u\left(h^i_{t+1,i}\right)}{\partial h^i_{t+1,i}}\right)\right] = \\
&= E_{R_t}\left[q\left(\frac{\partial^2 u\left(h^i_{t+1,i}\right)}{\partial^2 h^i_{t+1,i}}h''_{t+1,i}+\frac{\partial u\left(h^i_{t+1,i}\right)}{\partial h^i_{t+1,i}}-\frac{\partial^2 u\left(h^i_{t+1,i}\right)}{\partial^2 h^i_{t+1,i}}e_{t,i}\right)\right] > 0
\end{aligned}$$

with the same logic as before.

By implicit function theorem,

$$\begin{aligned}
\frac{\partial p_t}{\partial \theta_t} &= -\frac{s.o.c}{\frac{\partial f.o.c}{\partial p_t}} < 0 \\
\frac{\partial p_t}{\partial p^w_{t+1}} &= -\frac{\frac{\partial f.o.c}{\partial p_{t+1}}}{\frac{\partial f.o.c}{\partial p_t}} > 0.
\end{aligned}$$

■

Hence, until ω^i is sufficiently small to make $\left(e+\omega^i R_t\right) > 0$ for all realizations of R_t, $\gamma \leq 1$ ensures that the inverse demand function is downward sloping and increasing in future price p^w_{t+1}, so Assumption 3 holds. This is a condition on the absolute size of ω^A and ω^B. Now we show that if ω^B is close enough to ω^A then assumptions 1and 2 hold. The idea is that if we choose $\omega^B = \omega^A$, as we know that the autarchy prices exist, with $\theta^{\max} = 0$ the assumptions are trivially satisfied, but the relevant domain is a single point and $g^* = 0$. If we move ω^B marginally to the direction which makes $g^* > 0$, we have to get a very small relevant domain where the assumptions are satisfied by the continuity of the derivatives of the utility function.

A.2 Equilibrium

Proof of Theorem 1. The proof is given in steps.

1. **The aggregate position, $\bar{x}_t$ has to be in the relevant interval, i.e., $x_t \in \left[0, \bar{\theta}^{\max}\right]$ consequently $g_t \in [0, g^*]$ for all t.** Because of Assumption 3, if $\bar{x}_t < 0$ and the domain of $g(\cdot)$ includes $\bar{x}_t$, $g_t \notin [0, g^*]$. If $x_t < 0$ and $g_t > g^*$, $g_{t+1} > g_t$ must hold, otherwise there would be no possible gain in a $x_t < 0$ trade. But it means that $g_{t+\tau+1} > g_{t+\tau}$ for all $\tau \geq 0$. What is more, arbitrageurs will willing to support this path only if the expected gain is at least as large as the expected loss, i.e., $(1-q)(g_{t+\tau}+m) \leq q(g_{t+\tau+1}-g_{t+\tau}-m)$. This implies a bubble path where $g_{t+\tau+1}$ increases faster as τ grows. Therefore, $\sum_{t=0}^{\infty}\bar{x}_t(g_{t+1}+m)$ will not converge, which is a contradiction to the finite v_0. Similarly, $x_t > \theta^{\max}$ and $g_t < 0$ implies a bubble of the opposite direction, which leads to the same contradiction.

2. **The term $(g_{t+1}-g_t+m)$ is positive for all t.** If $(g_{t+1}-g_t+m) < 0$ in any t then arbitrageurs can take arbitrarily large $x_t > 0$ positions. Also, $(g_{t+1}-g_t) < 0$, so they make positive profit

regardless of whether the window closes in period $t+1$. Hence, it is an unbounded arbitrage possibility and cannot exist in equilibrium.

3. **If g_t is close enough to g^*, arbitrageurs will take a maximal position.** Observe that

$$\begin{aligned} \frac{g_t - m}{g^* - g_t + m} &\leq \frac{g_t - m}{g_{t+1} - g_t + m} \\ q\left(\frac{g^* - m}{m} + 1\right) &\geq q\left(1 + \frac{g_{t+1} - m}{g_{t+2} - g_{t+1} + m}\right). \end{aligned}$$

   Hence, if $g_t < g^*$ is large enough that

$$\frac{g_t - m}{g^* - g_t + m} > q\left(\frac{g^* - m}{m} + 1\right)$$

   then

$$\frac{g_t - m}{g_{t+1} - g_t + m} \geq q\left(1 + \frac{g_{t+1} - m}{g_{t+2} - g_{t+1} + m}\right)$$

   for any g_{t+1} and g_{t+2}. Thus, arbitrageurs will be better of to invest all their capital at period t.

4. **There is a finite T that $\bar{v}_t > 0$ for all $t < T$, but $\bar{v}_{T+\tau} = 0$ for all $\tau \geq 0$.** Observe that

$$\sum_{t=0}^{\infty} \bar{x}_t \left(g_{t+1} - g_t + m\right) \leq \bar{v}_0$$

   is a necessary condition for not violating the collateral constraint. Hence, $\lim_{t\to\infty} \bar{x}_t \left(g_{t+1} - g_t + m\right) = 0$. We know that $\lim_{t\to\infty} \left(g_{t+1} - g_t + m\right) > 0$, because otherwise $\lim_{t\to\infty} \left(g_{t+1} - g_t\right) = -m$, which is impossible as $g_t \geq 0$ for all t. Thus, $\lim_{t\to\infty} \bar{x}_t = 0$. Which implies $\lim_{t\to\infty} g_t = g^*$. Because of the previous point, it means that there is a $T-1$ where all arbitrageurs take a maximal position, i.e., $\bar{v}_T = 0$.

5. **The Lagrangian multiplier of the collateralization constraint $\lambda_t = 0$ for all $t < T$.** If any λ_t were positive for $t < T$, all arbitrageurs would take a maximal position at that period, which would inconsistent with 4.

6. **Characterization.** Because of the last point, the conditional gap path $\{g_t\}_{t=0}^{T-1}$ is characterized by the first order condition

$$(1-q)\left(g_t - m\right) = \left(g_{t+1} - g_t + m\right) q V'_{t+1}\left(v_{t+1}\right)$$

   and the envelope condition

$$V'_t\left(v_t\right) = (1-q) + q V'_{t+1}\left(v_{t+1}\right).$$

   As arbitrageurs take a maximal position in period $T-1$, $x_{T-1} = \frac{v_{T-1}}{g^* - g_{T-1} + m}$. If we plug this into the value function with $\bar{v}_{T+\tau} = 0$ for all $\tau \geq 0$ and differentiate with respect to v_{T-1}, we get

$$V'_{T-1}(v_{T-1}) = (1-q)\left(\frac{g_{T-1} - m}{g^* - g_{T-1} + m} + 1\right).$$

This is a final condition for the recursion described by the first order condition and the envelope theorem. The recursion gives (10). The value of T and g_{T-1} is given by the condition that g_{T-1} is high enough to make arbitrageurs take a maximal position (9) and the budget constraint (13), where the market clearing conditions (11)-(12) are used to determined $\{\bar{x}_t\}_{t=0}^{T-1}$ which support $\{g_t\}_{t=0}^{T-1}$.

7. $\{g_t\}_{t=0}^{T-1}$ **is strictly monotonically increasing.** From point 2, it is apparent that $\bar{v}_t$ is decreasing with t as long as $t \leq T$. Let us construct $T-1$ equilibria of the identical set-ups with the only difference that we change the initial capital level $\bar{v}_0$ to $\bar{v}_0^{\bar{v}_t} = \bar{v}_t$ where $0 < t < T$. From the recursive nature of the determination of the equilibrium, it is clear that the first element of these new equilibria $g_0^{\bar{v}_t}$ will give the original gap path $\{g_t\}_{t=0}^{T-1}$ by $g_t = g_0^{\bar{v}_t}$. So it is enough to see that $g_0^{\bar{v}_t}$ is decreasing in $\bar{v}_0^{\bar{v}_t} = \bar{v}_t$. For later reference, we show the stronger results that in any equilibrium all $\{g_t\}_{t=0}^{T-1}$ are decreasing in $\bar{v}_0$. Let us take a given $\bar{v}_0$ which determines a given price-path, g_i and a given T. More precisely, for g_{T-1}

$$g_{T-1} \geq \frac{q\,(g^*)^2}{(m + qg^*)} + m$$

holds, and this g_{T-1} determines the rest of the path by

$$g_t = m + g_{t+1}c_t \text{ for } t = 0...T-2$$

where

$$c_t(g_{T-1}, T, q) = \left(1 - \frac{(1-q)(g^* - g_{T-1} + m)}{g^* - g_{T-1} + m + q^{T-t-1}(g_{T-1} - m - qg^*)}\right). \tag{24}$$

We will decrease $\bar{v}_0$ to $\bar{v}'_0$, and check how g'_t relates to g_t.It is evident that $g'_t = g_t = g^*$ for all $t \geq T$. Let us first suppose that the decrease is small enough that $T' = T$. If $g_{T'-1}$ would still be equal to g_{T-1} then all g_t, $t < T-1$ would remain equal, but $g_0 < g'_0$ because $\bar{v}'_0$ would be used up sooner, which cannot be an equilibrium. If $g'_{T-1} < g_{T-1}$ then all $g'_t < g_t$ by $\frac{\partial c_t}{\partial g_{T-1}} > 0$ which would require more funds, which is not possible. So the only way to keep $T' = T$ is to increase g'_{T-1} which would increase all g_t by $\frac{\partial c_t}{\partial g_{T-1}} > 0$. which would require less funds.

Let us suppose now that $\bar{v}'_0$ is such that it supports the path $g'_{T-1} = g^*$ and

$$g'_t = m + g'_{t+1}c_t \text{ for } t = 0, ...T-3$$

i.e. $T' = T-1$ and $g'_{T'-1} = \frac{(q)(g^*)^2}{(m+(q)g^*)} + m$ by substituting c_{T-1} and we get the steepest possible $T' = T-1$. Hence, as for a small decrease of $\bar{v}'_0$ T remains the same and all g_t $t < T-1$ increase. Then for a larger increase, when g'_{T-1} reaches g^* the path coincides with the steepest possible

$T' = T - 1$ path. Hence, by the same argument as above, if we decrease $\bar{v}_0'$ further, the path gets flatter until g_{T-2}' reaches g^* as well, when we move to the steepest possible $T' = T - 2$ path etc.

■

Proof of Theorem 2.

1. It is clear that there is an equilibrium with $g_t = 0$ for all t if $m = 0$. In this case $g_{t+1} - g_t = 0$ so arbitrageurs are unconstrained in each period and indifferent when to invest because they can earn 0 expected profit in each period. Possibly, there is also other equilibria where $\{g_t\}_{t=0}^{\infty}$ and $\{\bar{x}_t\}_{t=0}^{\infty}$ satisfy

$$\frac{g_t}{g_{t+1} - g_t} = q\left(1 + \frac{g_{t+1}}{g_{t+2} - g_{t+1}}\right) \tag{25}$$

$$\sum_{t=0}^{\infty} \bar{x}_t \left(g_{t+1} - g_t\right) = \bar{v}_0 \tag{26}$$

$$g\left(\bar{x}_t, g_{t+1}\right) = g_t \tag{27}$$

$$g_t < g^* \tag{28}$$

for all t. If these conditions hold, arbitrageurs are indifferent when to invest (25), their collateral constraint can be satisfied (26) and the market clear (27) and $x_t \in \left[0, \bar{\theta}^{\max}\right]$, $g_t \in [0, g^*]$ for all t. Observe that for any positive constant C starting from any g_0 of the form

$$\frac{q^t + (1-q)\,C}{q^{t+1} + (1-q)\,C} g_t = g_{t+1} \tag{29}$$

equation (25) is satisfied. This implies that

$$g_t = \frac{(1 + (1-q)\,C)}{q^t + (1-q)\,C} g_0$$

which is monotonically increasing and converges to

$$\frac{(1 + (1-q)\,C)}{(1-q)\,C} g_0 \tag{30}$$

and all elements are increasing in g_0.If $g_0 \leq \bar{g}_0(C) = \frac{(1-q)C}{(1+(1-q)C)} g^*$, the series satisfies (28). Note also, that

$$\sum_{t=0}^{\infty} \bar{x}_t \left(g_{t+1} - g_t\right) \leq \sum_{t=0}^{\infty} \bar{\theta}^{\max} \left(g_{t+1} - g_t\right) \leq \bar{\theta}^{\max} g^*$$

so $\sum_{t=0}^{\infty} \bar{x}_t \left(g_{t+1} - g_t\right)$ is limited above in equilibrium. Thus,

$$\bar{v}_0^{\max} = \sup_{g_0 \in [0, \bar{g}_0(C)], C} \sum_{t=0}^{\infty} \bar{x}_t \left(g_{t+1} - g_t\right)$$

subject to (27) and (29) will have a finite solution $\bar{v}_0^{\max}$. As all $\{g_t\}_{t=0}^{\infty}$ and $\{\bar{x}_t\}_{t=0}^{\infty}$ are continuous in g_0 and C and for any C

$$\lim_{g_0 \to 0} \sum_{t=0}^{\infty} \bar{x}_t \left(g_{t+1} - g_t\right) = 0$$

for any $\bar{v}_0 < \bar{v}_0^{\max}$, there must be a solution of (25)-(28).

2. We know that the limit equilibrium will satisfy (25)-(27) because these are the limiting equations of (6), (12), (13) respectively. Fort this, note that $m \to 0$ implies $T \to \infty$, because (9) is never satisfied if $m \to 0$. We also know that as $m \to 0$, $g_t \to g^*$. From (30), this implies

$$C = \frac{g_0}{(1-q)\left(g^* - g_0\right)},$$

consequently

$$g_t = \frac{g_0 g^*}{q^t g^* + g_0 \left(1 - q^t\right)}. \tag{31}$$

The critical value $\bar{v}_0^{\lim}$ is determined by

$$\bar{v}_0^{\lim} = \sup_{g_0 \in [0, g^*]} \sum_{t=0}^{\infty} \bar{x}_t \left(g_{t+1} - g_t\right)$$

subject to (31) and (26). This will have a solution as $0 \leq \sum_{t=0}^{\infty} \bar{x}_t \left(g_{t+1} - g_t\right) < \bar{\theta}^{\max} g^*$. By continuity and as

$$\lim_{g_0 \to g^*} \sum_{t=0}^{\infty} \bar{x}_t \left(g_{t+1} - g_t\right) = 0$$

there will be a $g_0^{\lim}$ that (27) is satisfied if $\bar{v}_0 < \bar{v}_0^{\lim}$.

■

A.3 Comparative statics and discussion

Proof of Lemma 2. This can be seen by differentiating (16) with respect to t and $\bar{v}_t$ to get

$$\begin{aligned}
\frac{\partial V_t'}{\partial t} &= -\left((1-q)\frac{g_{T-1} - m}{g^* - g_{T-1} + m} - q\right) q^{T-1-t} \ln q > 0 \\
\frac{dV'\left(\bar{v}_t\right)}{d\bar{v}_t} &= \frac{\partial V'\left(\bar{v}_t\right)}{\partial g_{T-1}} \frac{\partial g_{T-1}}{\partial \bar{v}_t} + \frac{\partial V'\left(\bar{v}_t\right)}{\partial T} \frac{\partial T}{\partial \bar{v}_t} < 0
\end{aligned}$$

where we used (16) and the results in Lemma (1) that $\frac{\partial g_{T-1}}{\partial \bar{v}_t} < 0$ and $\frac{\partial T}{\partial \bar{v}_t} \geq 0$. ■

Proof of Lemma 3. In the limiting equilibrium

$$g_t = \frac{g_0 g^*}{q^t g^* + g_0^{\lim} \left(1 - q^t\right)}$$

for all t. Simple differentiation shows that

$$\frac{\partial\left(g_{t+1}-g_{t}\right)}{\partial t}=-q^{t}\ln q\left(g^{*}-g_{0}^{\lim}\right)\left(q\frac{1}{\left(q^{t+1}g^{*}+g_{0}^{\lim}\left(1-q^{t+1}\right)\right)^{2}}-\frac{1}{\left(q^{t}g^{*}+g_{0}^{\lim}\left(1-q^{t}\right)\right)^{2}}\right)$$

which is positive if and only if

$$\frac{q^{t}\left(\sqrt{q}-q\right)}{\left(1-q^{t+1}\right)-\sqrt{q}\left(1-q^{t}\right)}g^{*}>g_{0}^{\lim}. \tag{32}$$

This is true at $t=0$ if

$$\frac{\sqrt{q}\left(1-\sqrt{q}\right)}{1-q}g^{*}>g_{0}^{\lim}.$$

If this condition hold, then for $t=0$, the downside risk is increasing in t, or equivalently, for $\bar{v}_0$ it is decreasing in $\bar{v}_0$. As t increases the left hand side of (32) is decreases and as $t\rightarrow\infty$, the left hand side goes to 0. Hence, there must be a unique t when $\frac{\partial(g_{t+1}-g_t)}{\partial t}$ changes its sign. We know that $\bar{v}_t$ is decreasing with t, so the critical $\bar{v}^{\inf}$ also has to exist. ■