# DISCUSSION PAPER SERIES

No. 2672

## ARBITRAGE WITH INELASTIC LIQUIDITY DEMAND AND FINANCIAL CONSTRAINTS

Antonio S Mello and Mukarram Attari

***FINANCIAL ECONOMICS***

**Centre for Economic Policy Research**

**www.cepr.org**

**An online version of this Paper can be found at** www.cepr.org/pubs/dps/DP2672.asp

ISSN 0265-8003

# ARBITRAGE WITH INELASTIC LIQUIDITY DEMAND AND FINANCIAL CONSTRAINTS

**Antonio S Mello,** University of Wisconsin-Madison and CEPR
**Mukarram Attari,** University of Wisconsin-Madison

Discussion Paper No. 2672
January 2001

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 2672

January 2001

# ABSTRACT

## Arbitrage with Inelastic Liquidity Demand and Financial Constraints*

This Paper derives arbitrage trading strategies taking into account the fact that the actions of arbitrageurs impact prices. This avoids the difficulty of having to rely on exogenous position limits to prevent infinite arbitrage profits. When arbitrageurs are financially constrained their trading strategies can be expressed as feedback functions of their capital, which in turn depends on the optimal amount traded. An important component of the trading by financially constrained arbitrageurs is done to guarantee future financial flexibility. It is this hedging component that explains why price deviations persist in spite of arbitrage. Financial constraints are also responsible for periods of excessively volatile prices and for the time variation in the correlation of price deviation across markets. The fact that the actions of regulated firms can influence the dynamics of prices on which minimum capital requirements are based raises important implications for the regulation of securities firms.

JEL Classification: G10, G20
Keywords: arbitrage, large traders; market liquidity; financial supervision

Antonio S Mello
Research Fellow in Financial Economics
School of Business
University of Wisconsin-Madison
975 University Avenue
Madison
WI 53706-1323
USA
Tel: (1 608) 263 3423
Fax: (1 608) 265 4195
Email: amello@bus.wisc.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=115766

Mukarram Attari
Finance Department
School of Business
University of Wisconsin-Madison
975 University Avenue
Madison
WI 53706-1323
USA
Tel: (1 608) 262 6508
Fax: (1 608)265 4195
Email: mattari@bus.wisc.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=153069

*Antonio S Mello acknowledges the financial support from the Aschenbrener Fellowship, Puelicher Center for Banking Education at the School of Business, University of Wisconsin, Madison.

Submitted 19 September 2000

# NON-TECHNICAL SUMMARY


The Paper uses standard multi-period optimization techniques to derive optimal trading strategies for rational, profit-maximizing investors (which we call arbitrageurs), who trade in a market with discretionary liquidity traders and noise traders. One distinct feature of the model is that arbitrageurs understand that the quantities traded have an impact on the path of prices, and therefore scale their trades to take into account how marginal trading opportunities affect the value of their positions. Departing from the standard assumption that the impact of trades on price can be considered before the optimal trading strategy is derived poses the difficult problem of endogenizing the behaviour of arbitrageurs. The challenge is justified since we want to provide an accurate representation of large arbitrageurs in less than perfectly liquid markets. Given the size of the positions currently being managed by large securities houses and large hedge funds, we believe that the model can better describe a more common situation than at first appears, and that it applies to developed markets and not just to emerging and shallow markets.

We find that important anomalies can be explained in the context of our model. For example, why prices are excessively volatile and why there are patterns of volatility characterized by stable periods followed by turbulence. We also show that financially weak arbitrageurs can contribute to problems of contagion across different markets, and why in periods of instability returns tend to be significantly more correlated across markets. In this respect our Paper is an addition to the recent literature on financial crisis.

Our work discusses rules of capital requirements used to assess the risk of traders' positions. Measures used to determine the risk of firms' capital are not exogenous and are themselves influenced by the actions of the regulated firms. When this happens, VaR models incorrectly measure the extent of forced reductions in capital with important externalities and welfare distortions. This is of relevance, since how good capital adequacy rules measure risk of financial distress depends above all on liquidity. In markets that lack depth, as is often the case during periods of instability, traders face extraordinary difficulties in rearranging positions. Attempts to trade in shallow markets to ensure compliance of regulations can impose large losses on traders' capital. Capital adequacy requirements may then have unintended effects at times when they are most needed, because in their design it is ignored that supervised traders impact prices when they rearrange their positions.

In a letter to investors announcing the recasting of his Quantum Fund, George Soros wrote: "Quantum Fund is far too big and its activities too closely watched to be able to operate successfully in the market".[1] A month earlier another hedge fund arbitrageur, Julian Robertson, echoed this same feeling when deciding to close Tiger Management which had became too big, undermining its ability to trade in and out of positions. Both these statements occurred in a period of considerable market volatility and although pricing anomalies were frequent according to both arbitrageurs, opportunities to exploit discrepancies were dwindling. Given the sheer size of some hedge fund portfolios it should not be surprising that their bets are often noticed by the market.[2]

Financial arbitrage is carried out with very large positions in assets taken by relatively few investors for short periods of time. Though sophisticated arbitrageurs understand that their behavior might modify the path of prices, the academic finance literature still assumes that the optimal strategy for arbitrageurs can be derived ignoring the impact of arbitrage on prices. The few exceptions are the empirical studies of Mackinlay and Ramaswamy (1988) and Merrick (1989), which find evidence that path dependence in the basis of stock index futures contracts can be associated with the trading activities of arbitrageurs. Given that it is common practice for arbitrageurs when trading to consider the effects of marginal

[1]See "High-flyers come back down to earth", The Financial Times, April 29-30, 2000.

[2]Another interesting case is that of Pimco's actions reportedly associated with "Bloody Thursday", Feb 3, 2000, when the 30-year Treasury bond finished off a week with a steep rise of more than 2%, while mortgage bonds plummeted. According to the Wall Street Journal, Pimco bought large amounts of the 30-year Treasury bond and dumped mortgage bonds and triggered panicked trading. In Mr. William Gross, Pimco's manager, own words, "I think we were the spark to the market tumult. Others found out about the trades too quickly, and we couldn't do enough without driving the long bond to ridiculous levels." The Wall Street Journal, March 2000.

trading opportunities on the total value of their trades, it appears that competition between arbitrageurs can be viewed as less than atomistic. Deviations from the perfectly competitive case are not easily incorporated in a model of arbitrage trading because this raises the challenge of endogenizing the behavior of arbitrageuers given the impact of trades on market prices, a point first made by Brennan and Schwartz (1990).

In this paper we propose an analysis of arbitrage trading that attempts to capture some of these features in a simple and tractable way. Arbitrageurs are large and trade with noise traders, as well as with discretionary liquidity traders. This last group of traders is characterized as having inelastic demand functions for the asset. Asset prices deviate from the fundamental value and arbitrageurs aim to gain from buying cheap and selling expensive. But, because arbitrageurs know that their actions affect market prices, it is necessary to depart from the assumption that the impact of arbitrage on price is known before the optimal arbitrage strategy is derived. When arbitrageurs are unconstrained they supply a level of liquidity to other traders that maximizes their expected total trading profits.

A realistic model of arbitrage trading must take into account the fact that arbitrage traders have to back their trading activities with own capital. Examples include the capital adequacy directives establishing minimum capital requirements for the trading books of banks and securities firms and the system of margins in futures contracts and equity positions. Financial constraints can make a significant difference to the optimal level of arbitrage trading, as well as to the time series properties of market prices. We show that in the presence of financial constraints based on market prices for securities, arbitrage trading strategies can be expressed as feedback functions of the arbitrageur's current capital, which in turn is related to the optimal amount the arbitrageur chooses to trade. As a result, the arbitrageur evaluates his trading options as if the costs of the leverage constraints were endogenous.[3]

In addition to paying careful consideration as to how their marginal trades impact the total trade, financially constrained arbitrageurs try to avoid violations of the capital constraints that can result from adverse shocks to prices. Reducing the likelihood of violating the capital requirements effectively means biasing arbitrage positions towards zero. This is because a flat position in the asset is effectively immune to adverse shocks that affect the value of the portfolio. Although the bias towards the zero position has opportunity costs associated with forgone profits it has the benefit of reducing the likelihood of default, and therefore increasing financial flexibility. Interestingly, this hedging motive for trading makes financially constrained arbitrageurs at times trade larger quantities than unconstrained arbitrageurs faced with similar arbitrage opportunities.

In comparing the behavior of prices, we find that when arbitrageurs are not financially constrained, the expected future asset price is always equal to its expected fair value. However, in the presence of financially constrained arbitrageurs the market price is inefficient and tends to deviate from its fundamental value for a sequence of time periods. The reason is that in the desire to gain financial flexibility, a long arbitrageur sells more of the security when liquidity traders want to buy it than he buys of the security when liquidity traders want to sell it. Therefore, the mean value of the trade of a constrained arbitrageur with a long position is negative. This has the effect of driving the expected price below the ex-

[3]Several other articles have also addressed the effects of financial constraints on the trading strategies of rational traders - see, for example, Grossman and Vila (1992) and Zariphopoulou (1994). However, in all these models, trading strategies do not modify the price process and leverage constraints are evaluated in the presence of exogenously imposed leverage costs.

pected fundamental value. The opposite occurs when the arbitrageur is initially short in the asset. Persistent price deviations occur and are then the result of arbitrageurs being financially constrained, since no price deviations are expected if arbitrageurs are absent from the market or if they are unconstrained.[4]

Although it is not our objective to explicitly model financial crisis, in our setting financial constraints are shown to be responsible for anomalies such as excessively volatile prices and for calm periods turning into turbulent periods and vice versa. Contrary to the case where arbitrageurs are unconstrained where volatility is constant, in a market with financially constrained arbitrageurs price volatility is dependent on the position held by the arbitrageurs. On average, arbitrageurs reduce price volatility and improve market liquidity. However, there are periods when prices become especially volatile, because arbitrageurs, after an adverse shock, become constrained and are forced to liquidate positions to comply with capital requirements. Instead of contributing to reverse the effect of the liquidity shock, in these situations arbitrageurs cause the destabilization of market prices.[5] Higher price volatility also occurs when net liquidity in the market drops to low levels as a result of high volumes of noise trading and low elasticity of discretionary liquidity demand.[6] Another interesting feature of our model is that patterns of changing volatility occur endogenously, whereby stable periods are succeeded by more turbulent ones, and vice versa. The possibility of such scenarios

[4]Long run price ineficiencies have also been explained by transaction costs by Tuckman and Vila (1992), and short horizons by Dow and Gordon (1994).

[5]This paper is not the first to point out this anomaly. Gennotte and Leland (1990) show that rational traders facing portfolio insurance constraints put downard pressure on prices and amplify volatility in declining markets.

[6]Yuan (1999) attemps to explain financial crisis arising from informational differences between rational speculators facing borrowing constraints and liquidity traders with upward sloping demand functions. When the asset price is high liquidity traders absorb more of the random asset supply. When the asset price is low liquidity traders are increasingly resistant to absorb any supply because they do not know whether arbitrage trading is informational driven or because of borrowing constraints.

suggests that different forms of institutional constraints imposed on arbitrage trading, such as margin calls or tighter capital requirements, need to be carefully considered. When financially fragile arbitrageurs find it impossible to raise additional capital from reluctant investors these constraints can leave markets in a funk. In addition poorly designed rules such as the standard Value at Risk measure can contribute to the perverse regulation of trading risk and cause inefficiencies in the financial system.[7]

The actions of financially constrained arbitrageurs are also associated with correlations of prices across different markets. In a multi-asset environment we show that correlations between markets increase significantly in periods of instability, a fact that has been documented in recent empirical studies. Shocks that are unique to one market can affect the financial condition of arbitrageurs with positions in that market, forcing them to liquidate their holdings in other assets, and contributing to the propagation volatility across various markets. This idea of contagion is related to that found in Kodres and Prisker (1998). There informed investors reduce asset positions in various assets during market declines to meet portfolio balancing needs. Undiversified investors in unrelated markets are not clear about the motives of informed investors and sell as well. The result is that a negative shock in an asset spreads to other assets even when nothing has fundamentally changed. In our model these effects are more pronounced, because the actions by arbitrageurs impact both the level and volatility of market prices.

Some of the points made in our paper are consistent with the arguments in Shleifer

[7]In an interesting quote George Soros remarks that "Markets have become extremely unstable and historical measures of VaR no longer apply" (The Financial Times, 29-30 April 2000). We assert that VaR rules that are set ignoring that trading by financial firms impact on prices do not measure well the trading risk of these firms' portfolios, and can lead to serious distortions.

and Vishny (1997). Their paper considers arbitrage in a delegated portfolio management setting by modelling the arbitrageur's strategy when he obtains capital from investors who can observe returns on the portfolio but not the investment opportunities available. Their results are similar to some of those presented in this paper, namely that the fully invested arbitrageur may liquidate positions after a negative shock that results in immediate losses due to increased mispricing, mispricing which improves the future returns to the position. This will happen when the arbitrageur is unable to raise the additional capital needed to back the positions. The authors argue, without an explicit model, that risk averse arbitrageurs may liquidate positions even when they are not fully invested.

The remainder of the paper is organized as follows. Section I describes the economy and the optimal trading strategy of a monopolist unconstrained arbitrageur. The section also contains our main results on financially constrained arbitrage strategies and presents the moments of the distribution of price deviations Section II extends the results to a market with multiple arbitrageurs competing with each other. Section III considers the case of arbitrage trading in multiple risky securities and derives implications for price behavior across markets. Section IV discusses policy issues related to the regulation of arbitrage activity. Section V concludes.

## I. The Model

This section presents a dynamic strategy for an arbitrage trader who trades in a market for a security with a population of liquidity traders. The fundamental value of the security,

$I_t$, follows a random walk

$$I_t = I_{t-1} + \sigma_I \varepsilon_t \tag{1}$$

where, $\sigma_I$ is the per period volatility and $\varepsilon_t$ is the innovation in the fundamental value of the security at time $t$. This fair price is never observed directly but is revealed by the market price of the security, $F_t$. We assume that there are two components to liquidity trades in the market for the security. The first is a pure noise component, $A_t = \sigma_F \varepsilon_t$. The other component represents the actions of discretionary traders, who respond to the deviation of the market price from its fair price. In particular, a liquidity trader contemplating making a discretionary purchase of the asset will do so if $F_t < I_{t-1}$, since this means that the asset looks "cheap".[8] We therefore assume that the total demand by liquidity traders is given by

$$D_t^L = -\beta \left( F_t - I_{t-1} \right) + A_t \tag{2}$$

where $D_t^L$ is the flow of liquidity trades in the interval $(t-1, t)$, and $\beta > 0$. Notice that the first component of the demand is the price sensitive demand and the second is a flow of random liquidity orders from those who trade and are not price sensitive. For simplicity we assume that same shock affects both the asset value and the liquidity demand and that it has the following distributional properties: $E[\varepsilon_t] = 0$ and $E[\varepsilon_t^2] = 1$.[9]

The price of the security is set by market clearing. In the absence of arbitrage traders this price is determined by setting $D_t^L$ to zero, giving $F_t = I_{t-1} + \frac{\sigma_F}{\beta} \varepsilon_t$.

[8] The assumption here is that discretionary liquidity traders submit demand schedules, that is, they specify the quantity that they will trade as a function of price.

[9] This assumption can be relaxed very easily - two correlated shocks can be used, where the shocks have a correlation $\rho$. This makes the notation complicated and does not add to the intuition of the problem.

## A. Trading Strategy of an Unconstrained Monopolist arbitrageur

The price of the security in the absence of arbitrageurs is different from its value. The magnitude of the difference is $\left|\left(\frac{\sigma_F}{\beta} - \sigma_I\right)\varepsilon_t\right|$. The difference results from the fact that information related to the value of the security is incorporated into prices via the demand of liquidity traders. But the sensitivity of the liquidity trader demand to the changes in the value of the security is not one.[10] Arbitrageurs, will attempt, in light of their knowledge of the process generating liquidity trades to profit from these price deviations.

Let, $\theta_{t-1}$, represent the number of units of the security that the arbitrageur holds after trading at date $t-1$ (prior to trading at date $t$); the quantity that he trades at date $t$ is $\theta_t - \theta_{t-1}$. Given that the demand for the security from liquidity traders is not perfectly elastic, trading by the arbitrageur affects the market price of the security. If, as before, the liquidity trader demand for the security is given by $D_t^L$, the market price is set by the market clearing condition that $D_t^L + \theta_t - \theta_{t-1} = 0$. This gives the following expression for the price of the security:

$$F_t = I_{t-1} + \frac{A_t}{\beta} + \frac{\theta_t - \theta_{t-1}}{\beta} \tag{3}$$

The expression above describes how the demand flow from arbitrageurs affects the market price of the asset. If the arbitrageur is buying the security ($\theta_t - \theta_{t-1} > 0$) the price, $F_t$, will rise relative to its fundamental value. The reverse will be true if the arbitrageur is selling the security ($\theta_t - \theta_{t-1} < 0$).

[10]If $\beta = \frac{\sigma_F}{\sigma_I}$, the asset price adjusts to reflect the value of the asset. If $\beta > \frac{\sigma_F}{\sigma_I}$, the asset price under adjusts to changes in value, while $\beta < \frac{\sigma_F}{\sigma_I}$, causes asset prices to over adjust to changes in value. The $\beta < \frac{\sigma_F}{\sigma_I}$ case is consistent with observed short-run negative autocorrelation in asset prices.

The arbitrageur's profit from trading the security is given as:

$$TP(\theta_t) = \theta_{t-1}(I_t - I_{t-1}) + (\theta_t - \theta_{t-1})(I_t - F_t) \tag{4}$$

where, the first term represents the gains on the existing portfolio position and the second term represents the profits from trading the security at a price that is different from its value. The first part is unaffected by the trading of the arbitrageur, while the second part is sensitive to the choice of quantity traded. Using (3), the expression for trading profit can be re-written as

$$TP(\theta_t) = \theta_{t-1}\sigma_I\varepsilon_t - (\theta_t - \theta_{t-1})\left(\frac{\sigma_F}{\beta} - \sigma_I\right)\varepsilon_t - \frac{(\theta_t - \theta_{t-1})^2}{\beta} \tag{5}$$

This is the arbitrageur's profit from trading the security in a single period indexed $t$. The arbitrageur's objective function which is the discounted, at discount rate $\delta$ per period, expected trading profit over an infinite horizon starting period $t$, is given as

$$\begin{aligned} V_U &= \underset{\theta_\tau}{Max} \sum_{\tau=t}^{\tau=\infty} \delta^{\tau-t} E_t\left[TP(\theta_\tau)\right] \\ &= \theta_{t-1}\sigma_I\varepsilon_t - \underset{\theta_\tau}{Max} \sum_{\tau=t}^{\tau=\infty} \delta^{\tau-t} E\left[(\theta_\tau - \theta_{\tau-1})\left(\frac{\sigma_F}{\beta} - \sigma_I\right)\varepsilon_\tau + \frac{(\theta_\tau - \theta_{\tau-1})^2}{\beta}\right] \end{aligned} \tag{6}$$

The solution to this problem is $\theta_\tau^{U*} = \theta_{\tau-1} - \frac{\beta}{2}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)\varepsilon_\tau$. If $\frac{\sigma_F}{\beta} > \sigma_I$, the security price over adjusts in the absence of arbitrageurs and the arbitrageur will buy the security, that is $\theta_\tau^{U*} > \theta_{\tau-1}$, when liquidity traders are net sellers, $\varepsilon_\tau < 0$; and the arbitrageur will sell the security, that is $\theta_\tau^{U*} < \theta_{\tau-1}$, when liquidity traders are net buyers, $\varepsilon_\tau > 0$. In this situation both discretionary liquidity traders and the arbitrageur profit from trading against the noise traders.

On the other hand, when $\frac{\sigma_F}{\beta} < \sigma_I$ the liquidity shocks cause prices to under adjust for changes in the value of the asset. The discretionary liquidity traders are supplying too much liquidity to the market, that is, $\beta$ is too large. In this situation discretionary liquidity traders lose money and the arbitrageur profits from trading against them. In the long-run this is unsustainable since the discretionary liquidity traders would be forced out of the market. Therefore, $\frac{\sigma_F}{\beta} < \sigma_I$ is unlikely to be observed.

Note that in this simple setting the quantity traded is not a function of the current position or of any other variable. Substituting the optimally traded quantity into the expression for the price of the security, gives $F_t = I_{t-1} + \frac{1}{2}\left(\frac{\sigma_F}{\beta} + \sigma_I\right)\varepsilon_t$, and the into the objective function of the monopolist arbitrageur gives the maximized profit as

$$V_U = \theta_{t-1}\sigma_I\varepsilon_t + \frac{1}{(1-\delta)}\frac{\beta}{4}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)^2 \tag{7}$$

Arbitrage traders in financial markets are often required to back their trading positions with own capital. Examples include banks subject by regulators minimum capital requirements and margin requirements in futures contracts. Let us next analyze the problem of an arbitrageur who faces such capital requirements.

## B. Trading Strategy of a Financially Constrained Monopolist arbitrageur

The most common constraints faced by traders in financial markets are: (a) a bankruptcy constraint; (b) a symmetric margin requirement and (c) an asymmetric margin requirement. In this sub-section we consider the case of an arbitrageur who faces a bankruptcy constraint; discussion of the other types of constraints is delayed until a later section.[11]

In order to understand the implications of the constraints we start by modelling the evolution of the arbitrageur's capital balances. The capital balance, $C_t$ at time $t$, reflects the initial capital of the arbitrageur and the market value of the sum of the gains and losses from trading the asset. Changes to the capital balances occur due to gains and losses resulting from changes in the price of the security,

$$C_t = C_{t-1} + \theta_{t-1} \left( F_t - F_{t-1} \right) \tag{8}$$

An important measure of the capital balances is the amount of capital remaining after all positions have been liquidated,[12]

$$C_t(\theta_t = 0) = C_{t-1} + \theta_{t-1} \left( F_t(\theta_t = 0) - F_{t-1} \right) \tag{9}$$

[11] A very simple form of the bankruptcy constraint is considered here - the arbitrageur is forced to leave the market immediately liquidating his position if one measure of his capital becomes zero. More complicated forms of the constraint would include different liquidation strategies and different liquidation triggers. Almgren and Chriss (1998) show that a liquidation strategy of the type considered here is the strategy that has the minimum variance of liquidation proceeds.

[12] For very liquid assets, for example on-the-run US Treasury securities, liquidation of the arbitrageurs position will have a very small impact on the capital balance and $C_t\left(\theta_t = 0\right) \simeq C_t$. For relatively less liquid assets, for example stocks of small firms, liquidation of the arbitrageurs position may have a substantial negative impact on the capital balance giving $C_t\left(\theta_t = 0\right) << C_t$.

The evolution of the above quantity can be given as[13]

$$C_t(\theta_t = 0) = C_{t-1}(\theta_{t-1} = 0) + \frac{2}{\beta}\theta_{t-1}\left(\theta^U_{t-1} + \frac{\sigma_F}{2}\varepsilon_t - \theta_{t-1}\right) \tag{10}$$

where, $\theta^U_{t-1} = \theta_{t-2} - \frac{\beta}{2}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)\varepsilon_{t-1}$. The LHS of the above expression, $C_t(\theta_t = 0)$, is the capital available to the arbitrageur if he were to liquidate all positions and $F_t(\theta_t = 0)$ is the market price of the asset if the arbitrageur liquidates his position.

Equation (9) allows us to solve for the maximum position that can be held by the arbitrageur as a function of capital available, $C_{t-1}$, given that the arbitrageur is willing to bear a small risk of forced liquidation. Figure 1 plots this relationship. The horizontal axis is the capital available, $C_{t-1}$, while the vertical axis is the value of $\theta_{t-1}$ that ensures that $\Pr(C_t(\theta_t = 0) > 0) = k$. The plot solves expression (9), setting the LHS=0 and assuming adverse shocks, that is, $\varepsilon_t = -z_k$, when $\theta_{t-1} > 0$ and $\varepsilon_t = +z_k$, when $\theta_{t-1} < 0$. As can be seen the maximum position is a concave function of the capital. The figure also plots the position limits under a system where the position held by the arbitrageur has to be backed by a margin amount. This is represented by the straight lines in the figure. As can be seen, for low levels of capital, the margin requirements give tighter bounds on the position, while for larger amounts of capital the margin requirements are looser than the bankruptcy constraint.

<Insert Figure 1 here>

Ultimately, it is the amount of capital remaining on liquidation, $C_t(\theta_t = 0)$, that must

[13]This is obtained as follows: from equation (8) and (9) we have $C_t = C_t(\theta_t = 0) + \frac{\theta_t \theta_{t-1}}{\beta}$. A similar expression can be obtained at $t-1$. Substituting into (8) and simplifying gives (10).

satisfy some exogenously imposed reservation value. For the analysis here we assume that the arbitrageur can continue trading the security as long as this measure of capital is positive and that he is forced to liquidate all positions and leave the market if the above measure of capital is non-positive. This allows us to write the constrained monopolist arbitrageur's objective function as

$$V_C\left(\theta_{t-1}, C_t\left(\theta_t=0\right)\right)=\begin{cases} \underset{\theta_t}{Max}\ \ \theta_{t-1}\sigma_I\varepsilon_t-\frac{(\theta_t-\theta_{t-1})^2}{\beta} & C_t\left(\theta_t=0\right)>0 \\ \qquad -(\theta_t-\theta_{t-1})(\frac{\sigma_F}{\beta}-\sigma_I)\varepsilon_t & \\ \qquad +\delta E\left[V_C\left(\theta_t, C_{t+1}\left(\theta_{t+1}=0\right)\right)\right] & \\ \theta_{t-1}\sigma_I\varepsilon_t+\theta_{t-1}\left(\frac{\sigma_F}{\beta}-\sigma_I\right)\varepsilon_t-\frac{\theta_{t-1}^2}{\beta} & C_t\left(\theta_t=0\right)\leq 0 \end{cases} \tag{11}$$

In the above expression, the case where the arbitrageur initially has a positive capital balance is the problem that we would like to solve (the other case is trivial). The current level of capital available to the arbitrageur depends on the entire history of positions held by the arbitrageur, that is, it is path-dependent. As a result, a large number of state variables are required to make the problem Markovian, making it difficult to solve the problem analytically. For this reason we propose a simplification to the problem that retains all of its interesting aspects while allowing us to obtain an analytical solution.

The simplification is based on the following observation: The optimized value of the objective function described above is known at the two boundaries described by a financially unconstrained, $C_{t+1}\left(\theta_{t+1}=0\right)=\infty$, and a zero capital, $C_{t+1}\left(\theta_{t+1}=0\right)=0$, arbitrageur. In the first case the optimized value function is given by $V_U$ (as in equation (7)). In the second case, the arbitrageur is forced to liquidate his position and not trade ever again, giving

$$V_L(\theta_t) = \theta_t \sigma_I \varepsilon_{t+1} + \theta_t \left( \frac{\sigma_F}{\beta} - \sigma_I \right) \varepsilon_{t+1} - \frac{\theta_t^2}{\beta} \tag{12}$$

The value of the optimal trading strategy of a financially constrained arbitrageur with positive capital reserves $0 < C_{t+1}(\theta_{t+1} = 0) < \infty$, must lie between $V_L$ and $V_U$. This value of arbitrage trading is a concave function of the amount of capital balances available, because it must remain finite over the entire range of capital balances. We use this fact to approximate the $V_C(\theta_t, C_{t+1}(\theta_{t+1} = 0))$ term in the above optimization problem with a weighted sum of the value function at the two boundaries, $V_U$ and $V_L$. The problem is then written as:

$$\begin{aligned} V_C(\theta_{t-1}, C_t(\theta_t = 0) > 0) = \underset{\theta_t}{Max} \quad & \theta_{t-1}\sigma_I \varepsilon_t - (\theta_t - \theta_{t-1})(\tfrac{\sigma_F}{\beta} - \sigma_I)\varepsilon_t - \tfrac{(\theta_t - \theta_{t-1})^2}{\beta} \\ & + \delta E \left[ \begin{array}{c} (1 - \pi(C_{t+1}(\theta_{t+1} = 0, \varepsilon_{t+1}))) V_U(\theta_t) \\ + \pi(C_{t+1}(\theta_{t+1} = 0, \varepsilon_{t+1})) V_L(\theta_t) \end{array} \right] \end{aligned} \tag{13}$$

where the weights, $\pi(C_{t+1}(\theta_{t+1} = 0, \varepsilon_{t+1}))$ and $1 - \pi(C_{t+1}(\theta_{t+1} = 0, \varepsilon_{t+1}))$, given to each component of the value function next period, $V_U$ and $V_L$, are themselves function of the arbitrageur's capital balances next period. A consistent choice of $\pi's$ would be one which equals one if the trader has zero capital balance on liquidation and zero if the trader has infinite capital available. The $\pi's$ should be monotonically decreasing functions of the capital balance available to the trader, since a higher level of capital should imply a lower likelihood of forced liquidation. Also, the $\pi's$ should be such that the approximation used above should be a concave function of available capital.

The arbitrageur's capital balance next period is a function of his current capital balance, the existing position in the asset and the realization of the liquidity shock. A higher level of existing capital balance, $C_t(\theta_t = 0)$ translates into a higher future capital balance and, therefore, should translate into lower $\pi's$. If the arbitrageur has an initial long position, a negative liquidity shock decreases his capital and should be reflected in a higher $\pi$. Thus, we must have that $\pi(C_{t+1}(\theta_{t+1} = 0, \varepsilon_{t+1} < 0)) > \pi(C_{t+1}(\theta_{t+1} = 0, \varepsilon_{t+1} > 0))$ if $\theta_t > 0$ and $\pi(C_{t+1}(\theta_{t+1} = 0, \varepsilon_{t+1} > 0)) < \pi(C_{t+1}(\theta_{t+1} = 0, \varepsilon_{t+1} < 0))$ if $\theta_t < 0$.

A functional form that satisfies the above requirements for the $\pi's$ is

$$\pi(C_{t+1}(\theta_{t+1} = 0, \varepsilon_{t+1})) = \frac{a}{a + \bar{\theta}_t(C_{t+1}(\varepsilon_{t+1}))} \tag{14}$$

where, $a$ is a constant and $\bar{\theta}_t(C_{t+1}(\varepsilon_{t+1}))$ is the maximum position that the trader can hold in the risky asset without being forced to liquidate at $t+1$; that is, $\bar{\theta}_t(C_{t+1}(\varepsilon_{t+1}))$ solves

$$\begin{aligned} 0 &= C_{t+1}(\theta_{t+1} = 0) \\ &= C_t(\theta_t = 0) + \tfrac{2}{\beta}\overline{\theta}_t\left(\theta_t^U + \tfrac{\sigma_F}{2}\varepsilon_{t+1}\right) - 2\tfrac{\overline{\theta}_t^2}{\beta} \\ &= \overline{\theta}_t^2 - \overline{\theta}_t\left(\theta_t^U + \tfrac{\sigma_F}{2}\varepsilon_{t+1}\right) - \tfrac{\beta}{2}C_t(\theta_t = 0) \end{aligned} \tag{15}$$

where, $\theta_t^U = \theta_{t-1} - \frac{\beta}{2}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)\varepsilon_t$. For a given level of $\theta_t$ and $\varepsilon_{t+1}$, $C_{t+1}(\theta_{t+1} = 0)$ increases as $C_t(\theta_t = 0)$ increases. From (15) it can be seen that $\overline{\theta}_t$ is increasing in $C_t(\theta_t = 0)$, this will cause the $\pi's$ to be decreasing in $C_{t+1}(\theta_{t+1} = 0)$ since $\overline{\theta}_t$ is in the denominator of $\pi$. Since (15) is quadratic in $\overline{\theta}_t$, we have that $\overline{\theta}_t$ is proportional to $(C_t(\theta_t = 0))^{0.5}$, which implies that $\pi$ is proportional to $(C_t(\theta_t = 0))^{-0.5}$. Given a level of $\theta_t$, $C_{t+1}(\theta_{t+1} = 0)$ can be expressed as $C_t(\theta_t = 0)+$*constant*, we get that $\pi$ is proportional to $(C_{t+1}(\theta_{t+1} = 0))^{-0.5}$. The value functions, $V_U$ and $V_L$, are not functions of the available capital giving that $V_C$ is a concave function of capital.

Notice also that $\bar{\theta}_t(\varepsilon_{t+1} > 0) > \bar{\theta}_t(\varepsilon_{t+1} < 0)$ when $\theta_t^U > 0$ and the weight given to the likelihood of liquidation next period is greater for a negative liquidity shock, $\pi(C_{t+1}(\varepsilon_{t+1} > 0)) < \pi(C_{t+1}(\varepsilon_{t+1} < 0))$. Also, when $C_t(\theta_t = 0) \to \infty$, $\overline{\theta}_t \to \infty$ giving $\pi \to 0$ which is consistent with the intuition that a constrained arbitrageur with a very large amount of capital behaves similar to an unconstrained arbitrageur. When $C_t(\theta_t = 0) = 0$, $\overline{\theta}_t = 0$ giving $\pi = 1$ which means that the zero capital arbitrageur's profits from future trading equal the cost associated with liquidating his position in the asset.

The maximization program (13), with $\pi(C_{t+1}(\varepsilon_{t+1}))$ given by (14) has a first order condition:

$$\begin{aligned} 0 = \; & -(\tfrac{\sigma_F}{\beta} - \sigma_I)\varepsilon_t - \tfrac{2}{\beta}(\theta_t - \theta_{t-1}) \\ & +\delta(\tfrac{\sigma_F}{\beta} - \sigma_I)E\left[\pi(C_{t+1}(\varepsilon_{t+1}))\varepsilon_{t+1}\right] - \tfrac{2\delta}{\beta}\theta_t E\left[\pi(C_{t+1}(\varepsilon_{t+1}))\right] \end{aligned} \tag{16}$$

Solving for $\theta_t$ gives:

$$\theta_t^{C*} = \frac{\theta_{t-1} - \frac{\beta}{2}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)\left\{\varepsilon_t - \delta E\left[\pi(C_{t+1}(\varepsilon_{t+1}))\varepsilon_{t+1}\right]\right\}}{1 + \delta E\left[\pi(C_{t+1}(\varepsilon_{t+1}))\right]} \tag{17}$$

To interpret the above expression let us first rewrite it in terms of the quantity traded by the arbitrageur. This gives

$$\begin{aligned} \theta_t^{C*} - \theta_{t-1} = \; & -\tfrac{\beta}{2D(\varepsilon_t)}\left(\tfrac{\sigma_F}{\beta} - \sigma_I\right)\varepsilon_t - \theta_{t-1}\tfrac{(D(\varepsilon_t)-1)}{D(\varepsilon_t)} \\ & +\tfrac{\delta\beta}{2D(\varepsilon_t)}\left(\tfrac{\sigma_F}{\beta} - \sigma_I\right)E\left[\pi(C_{t+1}(\varepsilon_{t+1}))\varepsilon_{t+1}\right] \end{aligned} \tag{18}$$

where $D(\varepsilon_t) = 1 + \delta E\left[\pi(C_{t+1}(\varepsilon_{t+1}))\right] > 1$. The first term is similar to the case where the arbitrageur is unconstrained, except that we now have a $D(\varepsilon_t)$ in the denominator. Since $D(\varepsilon_t) > 1$ this damps the quantity traded by the constrained arbitrageur. The second term contributes a component that is always of opposite sign to $\theta_{t-1}$. The third term is also of opposite sign to $\theta_{t-1}$ because $\pi(C_{t+1}(\varepsilon_{t+1} > 0)) < \pi(C_{t+1}(\varepsilon_{t+1} < 0))$ when $\theta_{t-1} > 0$ and $\pi(C_{t+1}(\varepsilon_{t+1} > 0)) > \pi(C_{t+1}(\varepsilon_{t+1} < 0))$ when $\theta_{t-1} < 0$. These two terms can be interpreted as a drift towards a flat position.

Consider the case where the arbitrageur has an initial flat position, that is, $\theta_{t-1} = 0$. In this case the second term is zero. Also, $\pi(C_{t+1}(\varepsilon_{t+1} > 0)) < \pi(C_{t+1}(\varepsilon_{t+1} < 0))$ when $\varepsilon_t < 0$ and $\pi(C_{t+1}(\varepsilon_{t+1} > 0)) > \pi(C_{t+1}(\varepsilon_{t+1} < 0))$ when $\varepsilon_t > 0$. The intuition for this is as follows: If the current shock is negative the arbitrageur will have a long position after trading in the current period, hence a positive shock next period will increase his capital balance and lead to a lower likelihood of future liquidation. This third term also tends to bias the arbitrageur's position towards zero. Thus, when $\theta_{t-1} = 0$, the constrained arbitrageur trades a quantity that can be expressed as a linear function of the amount that the unconstrained arbitrageur trades and the quantity traded is symmetric.

If the arbitrageur's initial position is long, $\theta_{t-1} > 0$, the second and third terms are negative since $\pi(C_{t+1}(\varepsilon_{t+1} > 0)) < \pi(C_{t+1}(\varepsilon_{t+1} < 0))$. In addition to this the denominator is greater than in the case where the arbitrageur has an initial flat position. If $\varepsilon_t < 0$, this results in the arbitrageur buying a smaller quantity of the asset than the constrained arbitrageur with a initial flat position. On the other hand, if $\varepsilon_t > 0$ this results in the arbitrageur selling a larger quantity than the constrained arbitrageur with an initial flat position. The opposite holds true when $\theta_{t-1} < 0$.

The movement towards a flat position seen above has a cost associated with it by way of forgone profits, but, also has the benefit of increased flexibility resulting from a lower likelihood of forced liquidation.

## C. Price Patterns - Expected Deviation, Variance of Deviation and AutoCovariance of Deviation

In this sub-section we examine the impact on the price of the security by the trading of the arbitrageur. We consider three different properties of price: (a) the expected price deviation next period, $E[I_t - F_t]$; (b) the expected squared price deviation and the variance of the price deviation, $E[(I_t - F_t)^2]$ and $Var(I_t - F_t)$; and (c) the expected serial price deviation and the auto-covariance in price deviations, $E[(I_{t+1} - F_{t+1})(I_t - F_t)]$ and $Cov(I_{t+1} - F_{t+1}, I_t - F_t)$.[14] We do this for the three cases considered above: the absence of arbitrageurs, the case of the unconstrained monopolist arbitrageur and the financially constrained monopolist arbitrageur. The security price is given as $F_t = I_{t-1} + \frac{A_t}{\beta} + \frac{\theta_t - \theta_{t-1}}{\beta}$, where we set the final term to zero when there are no arbitrageurs trading the security. This gives $I_t - F_t = I_t - I_{t-1} - \frac{A_t}{\beta} - \frac{\theta_t - \theta_{t-1}}{\beta} = -\left(\frac{\sigma_F}{\beta} - \sigma_I\right)\varepsilon_t - \frac{\theta_t - \theta_{t-1}}{\beta}$.

### 1. Expected Price Deviation

In this sub-section we compare $E[I_t - F_t]$ for the three cases. In the absence of arbitrageurs $\theta_t = \theta_{t-1}$, giving $E[I_t - F_t] = -\left(\frac{\sigma_F}{\beta} - \sigma_I\right)E[\varepsilon_t] = 0$. When the arbitrageurs are unconstrained substituting for $\theta_t$ gives $E[I_t - F_t] = -\frac{1}{2}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)E[\varepsilon_t] = 0$. Thus, in both

[14] These are the basic moments of the asset price. We could also examine the mean, variance and autocovariance of $F_t$, but these are similar to the expressions reported below.

these cases the expected price deviations are zero.

When a financially constrained arbitrageur trades the security, the above is no longer true. This can be seen by substituting the quantity traded from (18) into the above expression. This gives

$$\begin{aligned} E\left[I_t - F_t\right] &= -\tfrac{1}{2}\left(\tfrac{\sigma_F}{\beta} - \sigma_I\right) E\left[\tfrac{\varepsilon_t}{D(\varepsilon_t)}\right] + \theta_{t-1} E\left[\tfrac{D(\varepsilon_t)-1}{D(\varepsilon_t)}\right] \\ &\quad -\tfrac{\beta\delta}{2}\left(\tfrac{\sigma_F}{\beta} - \sigma_I\right) E\left[\tfrac{\pi(C_{t+1}(\varepsilon_{t+1}))\varepsilon_{t+1}}{D(\varepsilon_t)}\right] \\ &< 0 \quad \theta_{t-1} < 0 \\ &= 0 \quad \theta_{t-1} = 0 \\ &> 0 \quad \theta_{t-1} > 0 \end{aligned} \tag{19}$$

The intuition for the above result is rather simple: a financially constrained arbitrageur always trades towards a flat position to get himself away from the constraint. Thus, if he has a long position he sells more of the security when liquidity traders want to buy it, than be buys of the security when liquidity traders want to sell. Thus, if he has an initial long position the mean quantity that he trades is negative. This has the effect of driving expected prices below expected value and represents a cost of the constraint. Figure 2 plots the expected price deviation. Figure 2 assumes a fixed amount of capital and plots the expected price deviation as a function of the arbitrageur's initial position. As can be seen from the figures the bias is increasing in the initial position of the arbitrageur.

<Insert Figure 2 here>

## 2. Expected Squared Price Deviation and Variance of Price Deviation

In this sub-section we examine two related properties of the price deviations - the expected squared deviation and the variance of the deviation.[15] When arbitrageurs are absent from the market or when there are unconstrained arbitrageurs the expected squared deviation and the variance of the deviations are the same. This is due to the fact that the expected price deviation is zero. When a constrained arbitrageur trades the security, these two quantities are different from each other because the expected price deviation is no longer zero.

The expected squared deviation can be written as $E\left[(I_t - F_t)^2\right]$, while the variance of deviations can be written as $Var(I_t - F_t) = E\left[(I_t - F_t - E\left[I_t - F_t\right])^2\right]$. When arbitrageurs are absent from the market $E\left[(I_t - F_t)^2\right] = Var(I_t - F_t) = \left(\frac{\sigma_F}{\beta} - \sigma_I\right)^2$. Trading by unconstrained arbitrageurs reduces the variability of prices, giving $E\left[(I_t - F_t)^2\right] = Var(I_t - F_t) = \frac{1}{4}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)^2$. This is 25% of the variance in the absence of the arbitrageur.

The expected squared deviation when a constrained arbitrageur trades the security is given as

$$E\left[(I_t - F_t)^2\right] = E\left[\left(\begin{array}{c} -\left(\frac{\sigma_F}{\beta} - \sigma_I\right)\varepsilon_t + \frac{1}{2D}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)\varepsilon_t + \theta_{t-1}\frac{(D-1)}{\beta D} \\ -\frac{\delta}{2D}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)E\left[\pi(C_{t+1}(\varepsilon_{t+1}))\varepsilon_{t+1}\right] \end{array}\right)^2\right] \tag{20}$$

As can be seen from the above equation the expected squared deviation, like the expected deviation, has an explicit dependence on the initial position of the arbitrageur.

Trading by a financially constrained arbitrageur results in a reduction in the variance of

[15] In the absence of information on the position held by the arbitrageur, the forecast of the deviation will be zero. This makes the expected squared deviation, which is the unconditional variance of price deviation, the more relevant measure of uncertainty in this setting.

the price deviation, which is given as

$$
\begin{aligned}
Var(I_t - F_t) &= \left(\tfrac{\sigma_F}{\beta} - \sigma_I\right)^2 + 2\left(\tfrac{\sigma_F}{\beta} - \sigma_I\right) E\left[\tfrac{\theta_t^{C*}\varepsilon_t}{\beta}\right] + \frac{E\left[\left(\theta_t^{C*}\right)^2\right] - E\left[\theta_t^{C*}\right]^2}{\beta} \\
&= \left(\tfrac{\sigma_F}{\beta} - \sigma_I\right)^2 + \tfrac{1}{\beta} E\left[\theta_t^{C*}\varepsilon_t\right]\left[\left(\tfrac{\sigma_F}{\beta} - \sigma_I\right) + \tfrac{1}{4\beta} E\left[\theta_t^{C*}\varepsilon_t\right]\right] \\
&\tfrac{1}{4}\left(\tfrac{\sigma_F}{\beta} - \sigma_I\right)^2 < Var(I_t - F_t) < \left(\tfrac{\sigma_F}{\beta} - \sigma_I\right)^2
\end{aligned} \tag{21}
$$

This results from the fact that the second term is always negative since $\beta\left(\frac{\sigma_F}{\beta} - \sigma_I\right) = -E\left[\theta_t^{U*}\left(\varepsilon_t\right)\varepsilon_t\right] > -E\left[\theta_t^{C*}\left(\varepsilon_t\right)\varepsilon_t\right] > 0$.[16]

Figure 3 plots the expected squared deviation (upper line) and the variance of the price deviation (lower line) as a function of the initial position of the constrained arbitrageur. From the plot it can be seen that the expected squared price deviation is increasing as the arbitrageur's initial position approaches its maximum sustainable level. Consider $\theta_{t-1} > 0$: As $\theta_{t-1}$ increases the arbitrageur sells an increasing amount of the asset to allow him to move back towards a flat position. This causes an increasing deviation in between the asset's price and value. When $\theta_{t-1} < 0$: the arbitrageur buys an increasing amount of the asset to help him move back to a flat position, causing price to deviate in the opposite direction relative to value.

<Insert Figure 3 here>

### 3. AutoCovariance of Price Deviations

In this sub-section we again examine two related properties of the price deviations - the expected serial co-deviation $E\left[\left(I_{t+1} - F_{t+1}\right)\left(I_t - F_t\right)\right]$ and the autocovariance of the price

[16] If $E\left[\theta_t^{C*}\left(\varepsilon_t\right)\varepsilon_t\right] = -\beta\left(\frac{\sigma_F}{\beta} - \sigma_I\right)$, the second term equal $-\frac{3}{4}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)^2$, while if $E\left[\theta_t^{C*}\left(\varepsilon_t\right)\varepsilon_t\right] = 0$, the second term is 0.

deviations, $Cov\left(I_{t+1}-F_{t+1}, I_t - F_t\right)$. In the first two cases, absence of arbitrageur and trading by an unconstrained arbitrageur, both of the above quantities are zero. In both these cases the deviations depend only on the current period shock, and the shocks are not auto-correlated giving zero expected serial co-deviation and zero autocovariance in price deviations.

The presence of a financially constrained arbitrageur induces serial correlation in the price deviations. The expected serial co-deviation and autocovariance of price deviations depend on the initial position of the arbitrageur. This is due to the fact that the quantity traded by the constrained arbitrageur depends on his initial position in the asset - the constrained arbitrageur continuously trades down to a flat position. The expected serial co-deviation is given as

$$\begin{aligned} E_t\left[\left(I_{t+1}-F_{t+1}\right)\left(I_t-F_t\right)\right] \quad & E_t\left[\left(\left(\tfrac{\sigma_F}{\beta}-\sigma_I\right)\varepsilon_{t+1}+\tfrac{\theta_{t+1}^{C*}-\theta_t^{C*}}{\beta}\right)\left(\left(\tfrac{\sigma_F}{\beta}-\sigma_I\right)\varepsilon_t+\tfrac{\theta_t^{C*}-\theta_{t-1}}{\beta}\right)\right] \\ & = E_t\left[\left(\tfrac{\theta_{t+1}^{C*}-\theta_t^{C*}}{\beta}\right)\left(\left(\tfrac{\sigma_F}{\beta}-\sigma_I\right)\varepsilon_t+\tfrac{\theta_t^{C*}-\theta_{t-1}}{\beta}\right)\right] \\ & > 0 \end{aligned} \tag{22}$$

The intuition for this result is that the arbitrageur is always trading down towards a flat position. This will cause successive price deviations to be in the same direction. Mathematically, the autocovariance is given by

$$
\begin{aligned}
Cov(I_{t+1}-F_{t+1}, I_t-F_t)\ & E\left[\left(\frac{\theta_{t+1}^{C*}-E\left[\theta_{t+1}^{C*}\right]}{\beta}+\left(\frac{\sigma_F}{\beta}-\sigma_I\right)\varepsilon_{t+1}\right)\left(\frac{\theta_t^{C*}-E\left[\theta_t^{C*}\right]}{\beta}+\left(\frac{\sigma_F}{\beta}-\sigma_I\right)\varepsilon_t\right)\right] \\
&= E\left[\left(\frac{\theta_{t+1}^{C*}-E\left[\theta_{t+1}^{C*}\right]}{\beta}\right)\left(\frac{\theta_t^{C*}-E\left[\theta_t^{C*}\right]}{\beta}+\left(\frac{\sigma_F}{\beta}-\sigma_I\right)\varepsilon_t\right)\right] \\
&> 0
\end{aligned} \tag{23}
$$

Figure 4 plots the expected serial co-deviation as a function of the constrained arbitrageurs initial position in the asset. As can be seen the expected serial co-deviation is a quadratic function of the initial position of the constrained liquidity trader when the initial position is not too large. For traders that have initial positions close to the maximum sustainable the expected serial co-deviation rises sharply because of the high likelihood of forced liquidation. This is consistent with the explanation provided above.

<Insert Figure 4 here>

This section considered the case of a monopolist arbitrageur trading a single risky security. In the next section we examine the impact of trading by multiple financially constrained arbitrageurs competing with one another. The section following that considers the case where a single arbitrageur trades multiple risky securities.

## II. Oligopolistic Competition Among arbitrageurs

Here we consider the case where there are $n$ arbitrageurs supplying liquidity. The market clearing condition is modified to

$$D_t^L + \theta_t - \theta_{t-1} + (n-1)\,\Delta_t = 0$$

where, the arbitrageur under consideration trades $\theta_t - \theta_{t-1}$ and $\Delta_t$ is the quantity traded by each of the remaining $(n-1)$ arbitrageur. The market price of the asset is given as

$$F_t = I_{t-1} + \frac{A_t}{\beta} + \frac{\theta_t - \theta_{t-1}}{\beta} + \frac{(n-1)\,\Delta_t}{\beta}$$

In this setting, there is competition among arbitrageurs. Since it is not atomistic competition, each arbitrageur will consider the impact of his own trading as well as that of competing arbitrageurs in determining the quantity to trade. A financially unconstrained arbitrageur optimally trades to the following position:

$$\theta_t^{UO*} = \theta_{t-1} - \frac{\beta}{(n+1)}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)\varepsilon_t \tag{24}$$

which solves the maximization problem after substituting for $F_t$ into the expression for the total profits, $TP(\theta_t)$, when arbitrageurs are symmetric[17].

If arbitrageurs are financially constrained, we need to examine the evolution of their

[17] Assuming that the arbitrageurs are symmetric is not unrealistic. Assymetry is generated by two sources: (i) the initail position of arbitrageurs and (ii) differential information or beliefs. In markets where differential information is not important, the impact of differences in the initail position will die out over time as all arbitrageurs positions will converge.

capital balances. The capital balance of an arbitrageur evolves as $C_{t+1} = C_t + \theta_t (F_{t+1} - F_t)$. Substituting for $F_{t+1}$ gives $C_{t+1} = C_t + \theta_t \left( I_t + \frac{A_{t+1}}{\beta} + \frac{\theta_{t+1} - \theta_t}{\beta} + \frac{(n-1)\Delta_{t+1}}{\beta} - F_t \right)$. If each trader accounts for the positions of the remaining traders when calculating the liquidation value of his capital, we get $C_{t+1}(\theta_{t+1} = 0) = C_t(\theta_t = 0) + \frac{(n+1)}{\beta}\theta_t \left[ \theta_t^U + \frac{\sigma_F}{(n+1)}\varepsilon_{t+1} \right] - \frac{(n+1)\theta_t^2}{\beta}$, where $\theta_t^U = \theta_{t-1} - \frac{\beta}{(n+1)} \left( \frac{\sigma_F}{\beta} - \sigma_I \right) \varepsilon_t$. Let $\overline{\theta}_t$ be the limit position that the arbitrageur can hold without being forced to liquidate next period, that is $\overline{\theta}_t$ solves $C_{t+1}(\theta_{t+1} = 0) = 0$, and use as before weight functions $\pi(\varepsilon_{t+1}) = \frac{a}{a + \overline{\theta}_t(\varepsilon_{t+1})}$. This allows us to write the problem of a financially constrained arbitrageur as

$$\begin{aligned} \underset{\theta_t}{Max}\ & \theta_{t-1}\sigma_I\varepsilon_t - (\theta_t - \theta_{t-1}) \left( \frac{\sigma_F}{\beta} - \sigma_I \right) \varepsilon_t - \frac{(\theta_t - \theta_{t-1})^2}{\beta} - (\theta_t - \theta_{t-1}) \frac{(n-1)\Delta_t}{\beta} \\ & + \delta E \left[ \pi(\varepsilon_{t+1}) V_L(\varepsilon_{t+1}) + (1 - \pi(\varepsilon_{t+1})) V_U(\varepsilon_{t+1}) \right] \end{aligned} \tag{25}$$

where, $V_L(\varepsilon_{t+1}) = \theta_t\sigma_I\varepsilon_{t+1} + \theta_t \left( \frac{\sigma_F}{\beta} - \sigma_I \right) \varepsilon_{t+1} - \frac{n\theta_t^2}{\beta}$, which is obtained by setting $\Delta_{t+1} = -\theta_t$ because of simultaneous forced liquidation at $t+1$, and $V_U(\varepsilon_{t+1}) = \theta_t\sigma_I\varepsilon_{t+1} + \frac{1}{1-\delta} \frac{\beta}{(n+1)^2} \left( \frac{\sigma_F}{\beta} - \sigma_I \right)^2$. Solving for the optimal position gives

$$\theta_t^{CO*} = \frac{\theta_{t-1} - \frac{\beta}{(n+1)} \left( \frac{\sigma_F}{\beta} - \sigma_I \right) \{ \varepsilon_t - \delta E[\pi(\varepsilon_{t+1}) \varepsilon_{t+1}] \}}{\left( 1 + \frac{2\delta n}{(n+1)} E[\pi(\varepsilon_{t+1})] \right)} \tag{26}$$

It is not surprising that the financially constrained oligopolistic arbitrageur holds a smaller position in the asset than if he were unconstrained. Also, the difference between the oligopolist case and the monopolist case is the standard one. Individually, each competitive arbitrageur holds less than the monopolist arbitrageur, but on aggregate the quantity traded by oligopolistic arbitrageurs is larger than that of the monopolist, and therefore the impact on market prices is higher under the oligopoly, and the profit per unit and the total arbitrage profits are also lower.

From the above it can be seen that the oligopolist arbitrageur's optimal quantity traded is determined in a similar way to the monopolist arbitrageur's. The results on price patterns are also very similar in this case.

The above results were obtained assuming that the arbitrageurs held the same initial positions and had the same amount of capital available to them. A situation that may sometimes be observed is one where the arbitrageurs have different amounts of capital available to them or/and have different initial positions. The discussion here is an attempt to understand how the above results would change in this setting.

We can consider the heterogenity of capital availability and initial positions together since the quantity that is important is the initial position relative to the maximum sustainable. Let $n$ arbitrageurs trade an asset. As seen above, the quantity traded by constrained arbitrageurs is made up of two parts - the purpose of one component of the trade is to help move the arbitrageur's position down to zero, the second component is the supply of liquidity that generates trading profits.

First consider the case where the total initial position held by the competing arbitrageurs is close to zero. In this setting the arbitrageurs start out with initial positions on both sides of the market - some have long positions in the asset while others have short positions. Each arbitrageur will be aggressive in determining the magnitude of the first component of the trade since the cost of moving towards a flat position will be relatively low - competing arbitrageurs will be on different sides. This means that situations with arbitrageurs on both sides of the market with total initial position close to zero will tend to disappear rather quickly. The second component of the trade, the supply of liquidity will also be determined more aggressively than in the symmetric case, though much less aggressively than the first component of the trade. The aggressiveness in the determination of the second component of the trade is lower since the arbitrageurs would not rationally expect the initial conditions to persist.

Now, let us consider a second case where the arbitrageurs have positions of various sizes but the total position of the arbitrageurs is significantly different from zero, say a net long position. In this setting the total of the of the first component of the trade across arbitrageurs will be negative. Arbitrageurs who have an initial short position will be able to aggressively liquidate their positions, while arbitrageurs with initial long positions will be forced to be much less aggressive in their movement towards a flat position. On the second component of the trade, the arbitrageurs with a greater level of flexibility, those with smaller initial positions relative to the amount of capital available will be able to supply liquidity more aggressively allowing them to earn greater profits.

We next consider the case of an arbitrageur who can trade multiple assets.

## III. Arbitraging with Multiple Assets

In this section we consider the problem of a monopolist arbitrageur who can trade multiple risky assets. For simplicity we assume that the only link between the markets is through the trading of the arbitrageur and the shocks to the values of the assets and to the liquidity demand in the markets. If the arbitrageur is unconstrained his optimization problem is similar to that of two arbitrageurs each of whom trades a single asset. This allows us to write the value function for the two-asset unconstrained problem as

$$V_U\left(\varepsilon_{1,t+1},\varepsilon_{2,t+1}\right)=\sum_{i=1}^{2}\left[\theta_{i,t}\sigma_{iI}\varepsilon_{i,t+1}+\frac{\beta_i}{4\left(1-\delta\right)}\left(\frac{\sigma_{iF}}{\beta_i}-\sigma_{iI}\right)^2\right] \tag{27}$$

Consider next how the financially constrained monopolist arbitrageur splits his trades when there two assets. Position limits are computed from the capital available which depends on the portfolio holdings:

$$C_{t+1}\left(\theta_{1,t+1}=\theta_{2,t+1}=0\right)=C_t\left(\theta_{1,t}=\theta_{2,t}=0\right)+\sum_{i=1}^{2}\frac{2}{\beta_i}\theta_{i,t}\left[\theta_{i,t}^U+\frac{\sigma_{iF}}{2}\varepsilon_{i,t+1}-\theta_{i,t}\right] \tag{28}$$

where, $\theta_{i,t}^U=\theta_{i,t-1}-\frac{\beta_i}{2}\left(\frac{\sigma_{iF}}{\beta_i}-\sigma_{iI}\right)\varepsilon_{i,t}$. Define the weight function $\pi\left(\varepsilon_{1,t+1},\varepsilon_{2,t+1}\right)=\frac{a}{a+\overline{\theta}_{1t}\left(\varepsilon_{1,t+1},\varepsilon_{2,t+1}\right)}$ where $\overline{\theta}_{1t}\left(\varepsilon_{1,t+1},\varepsilon_{2,t+1}\right)$ is obtained by setting to zero the amount of capital left after liquidating the portfolio next period, $C_{t+1}\left(\theta_{1,t+1}=\theta_{2,t+1}=0\right)=0$. We assume that if the arbitrageur is not forced into liquidation he will trade to maintain next period a ratio of holdings, $\frac{\theta_{2,t}}{\theta_{1,t}}=\frac{\theta_{2,t}^U}{\theta_{1,t}^U}=\alpha_t$[18]. $\overline{\theta}_{1t}\left(\varepsilon_{1,t+1},\varepsilon_{2,t+1}\right)$ then solves

$$0=C_t\left(\theta_{1,t}=\theta_{2,t}=0\right)+\frac{2}{\beta_1}\overline{\theta}_{1t}\left[\left(\theta_{1t}^U-\overline{\theta}_{1t}\right)\left(1+\alpha_t^2\frac{\beta_1}{\beta_2}\right)+\frac{\sigma_{1F}}{2}\varepsilon_{1,t+1}+\alpha_t\frac{\beta_1}{\beta_2}\frac{\sigma_{2F}}{2}\varepsilon_{2,t+1}\right] \tag{29}$$

The objective function can then be written as

[18]The expression for the evolution of the capital has two variables, $\theta_{1,t}$ and $\theta_{2,t}$. This assumption allows us to fix the value of one variable relative to the other. In the single asset case $\overline{\theta}_t$ is the maximum position that can be sustained without triggering bankruptcy next period. In the two asset case we would have to solve for $\overline{\theta}_{1,t}$ and $\overline{\theta}_{2,t}$, the maximum in each of the two assets, where $\overline{\theta}_{1,t}$ and $\overline{\theta}_{2,t}$ would be related. As a result of the assumption we consider only a particular combination of $\overline{\theta}_{1,t}$ and $\overline{\theta}_{2,t}$. The combination that we consider has the same ratio of holdings as held by the unconstrained arbitrageur.

$$\underset{\theta_{1,t},\theta_{2,t}}{Max} \sum_{i=1}^{2} \left[\theta_{i,t-1}\sigma_{i,I}\varepsilon_{i,t} - (\theta_{i,t} - \theta_{i,t-1})(\tfrac{\sigma_{i,F}}{\beta_i} - \sigma_{i,I})\varepsilon_{i,t} - \tfrac{(\theta_{i,t}-\theta_{i,t-1})^2}{\beta_i}\right] + \delta E \left[ \begin{array}{c} \pi\left(\varepsilon_{1,t+1}, \varepsilon_{2,t+1}\right) V_L\left(\varepsilon_{1,t+1}, \varepsilon_{2,t+1}\right) \\ + \left(1 - \pi\left(\varepsilon_{1,t+1}, \varepsilon_{2,t+1}\right)\right) V_U\left(\varepsilon_{1,t+1}, \varepsilon_{2,t+1}\right) \end{array} \right] \tag{30}$$

and the value of the trading strategy when the arbitrageur is forced to liquidate because of zero capital is also similar to the single asset example:

$$V_L\left(\varepsilon_{1,t+1}, \varepsilon_{2,t+1}\right) = \sum_{i=1}^{2} \left[\theta_{i,t}\sigma_{iI,t}\varepsilon_{i,t+1} + \theta_{i,t}\left(\frac{\sigma_{i,F}}{\beta_i} - \sigma_{i,I}\right)\varepsilon_{i,t+1} - \frac{\theta_{i,t}^2}{\beta_i}\right]$$

Consider correlated liquidity shocks across markets, $E\left[\varepsilon_{1,t}\varepsilon_{2,t}\right] = \rho$. The solution to the constrained optimization problem for asset 1 is

$$\theta_{1,t} = \frac{\theta_{1,t-1} - \frac{\beta_1}{2}\left(\frac{\sigma_{1,F}}{\beta_1} - \sigma_{1,I}\right)\left(\varepsilon_{1,t} - \delta E\left[\pi\left(\varepsilon_{1,t+1}, \varepsilon_{2,t+1}\right)\varepsilon_{1,t+1}\right]\right)}{\left(1 + \delta E\left[\pi\left(\varepsilon_{1,t+1}, \varepsilon_{2,t+1}\right)\right]\right)}$$

and similarly for asset 2,

$$\theta_{2,t} = \frac{\theta_{2,t-1} - \frac{\beta_2}{2}\left(\frac{\sigma_{2,F}}{\beta_2} - \sigma_{2,I}\right)\left(\varepsilon_{2,t} - \delta E\left[\pi\left(\varepsilon_{1,t+1}, \varepsilon_{2,t+1}\right)\varepsilon_{2,t+1}\right]\right)}{\left(1 + \delta E\left[\pi\left(\varepsilon_{1,t+1}, \varepsilon_{2,t+1}\right)\right]\right)}$$

The above solutions do not contain an explicit dependence of the quantities traded of one asset on the direction of the shock to the other asset. The dependence results from the fact that the expressions for the $\pi' s$ contain the current shocks and holdings of both assets. The arbitrageur's capital balance is affected by the direction of the shock and the current holdings of both assets which affects $\overline{\theta}_t$ and this affects the weights used to approximate future profits. Thus, the quantity traded of each asset does depend on the liquidity shock to the other asset and the current holding of the other asset.

## A. Price Patterns - Expected Cross Deviation

In this sub-section we examine the expected cross deviation, $E\left[\left(I_{1t}-F_{1t}\right)\left(I_{2t}-F_{2t}\right)\right]$, in the two asset prices. In the absence of the arbitrageur the two asset prices are linked only by the fact that the liquidity and value shocks are correlated. Using the expressions for the price of the assets from section 2 we can write

$$\begin{aligned} E\left[\left(I_{1t}-F_{1t}\right)\left(I_{2t}-F_{2t}\right)\right] &= \left(\tfrac{\sigma_{1F}}{\beta_1}-\sigma_{1I}\right)\left(\tfrac{\sigma_{2F}}{\beta_2}-\sigma_{2I}\right)E\left[\varepsilon_{1,t}\varepsilon_{2,t}\right] \\ &= \left(\tfrac{\sigma_{1F}}{\beta_1}-\sigma_{1I}\right)\left(\tfrac{\sigma_{2F}}{\beta_2}-\sigma_{2I}\right)\rho \end{aligned} \tag{31}$$

The presence of an unconstrained arbitrageur does not change the results significantly. There is a reduction the magnitude, but this results purely from the fact that the arbitrageur absorbs part of the liquidity shock. We can now write

$$E\left[\left(I_{1t}-F_{1t}\right)\left(I_{2t}-F_{2t}\right)\right]=\frac{1}{4}\left(\frac{\sigma_{1F}}{\beta_1}-\sigma_{1I}\right)\left(\frac{\sigma_{2F}}{\beta_2}-\sigma_{2I}\right)\rho \tag{32}$$

This reduction in the magnitude of the expected cross deviation is similar to the reduction in the magnitude of the variance of the deviation in the single asset unconstrained case. If we standardize the above expressions by the standard deviation of the price deviation, we obtain the correlation of the price deviation, which in both cases is equal to $\rho$. Thus, the absence of an arbitrageur or trading by an unconstrained arbitrageur does not affect the correlation between the asset prices.

This is no longer true if there is a constrained arbitrageur trading the assets. In this case, the expected cross deviation depends on the arbitrageurs holdings of the two assets. The expression for the expected cross deviation is given as

$$\begin{aligned} E\left[\left(I_{1t}-F_{1t}\right)\left(I_{2t}-F_{2t}\right)\right] &= E\left[\left(\left(\tfrac{\sigma_{1F}}{\beta_1}-\sigma_{1I}\right)\varepsilon_{1t}+\tfrac{\left(\theta_{1t}-\theta_{1t-1}\right)}{\beta_1}\right)\left(\left(\tfrac{\sigma_{2F}}{\beta_2}-\sigma_{2I}\right)\varepsilon_{2t}+\tfrac{\left(\theta_{2t}-\theta_{2t-1}\right)}{\beta_2}\right)\right] \\ &= \left(\tfrac{\sigma_{1F}}{\beta_1}-\sigma_{1I}\right)\left(\tfrac{\sigma_{2F}}{\beta_2}-\sigma_{2I}\right)\rho+E\left[\left(\tfrac{\sigma_{1F}}{\beta_1}-\sigma_{1I}\right)\tfrac{\left(\theta_{2t}-\theta_{2t-1}\right)}{\beta_2}\varepsilon_{1t}\right] \\ &\quad +E\left[\left(\tfrac{\sigma_{2F}}{\beta_2}-\sigma_{2I}\right)\tfrac{\left(\theta_{1t}-\theta_{1t-1}\right)}{\beta_1}\varepsilon_{2t}\right]+E\left[\tfrac{\left(\theta_{2t}-\theta_{2t-1}\right)}{\beta_2}\tfrac{\left(\theta_{1t}-\theta_{1t-1}\right)}{\beta_1}\right] \end{aligned} \quad (33)$$

To interpret this expression let us start by rewriting the expressions of the quantity that the arbitrageur trades as

$$\begin{aligned} \theta_{1,t}-\theta_{1,t-1} = &-\tfrac{\beta_1}{2D}\left(\tfrac{\sigma_{1,F}}{\beta_1}-\sigma_{1,I}\right)\varepsilon_{1,t}-\theta_{1,t-1}\tfrac{(D-1)}{D} \\ &+\tfrac{\delta\beta_1}{2D}\left(\tfrac{\sigma_{1,F}}{\beta_1}-\sigma_{1,I}\right)E\left[\pi\left(\varepsilon_{1,t+1},\varepsilon_{2,t+1}\right)\varepsilon_{1,t+1}\right] \end{aligned}$$

and

$$\begin{aligned} \theta_{2,t}-\theta_{2,t-1} = &-\tfrac{\beta_2}{2D}\left(\tfrac{\sigma_{2,F}}{\beta_2}-\sigma_{2,I}\right)\varepsilon_{2,t}-\theta_{2,t-1}\tfrac{(D-1)}{D} \\ &+\tfrac{\delta\beta_2}{2D}\left(\tfrac{\sigma_{2,F}}{\beta_2}-\sigma_{2,I}\right)E\left[\pi\left(\varepsilon_{1,t+1},\varepsilon_{2,t+1}\right)\varepsilon_{2,t+1}\right] \end{aligned}$$

where, $D=\left[1+\delta E\left[\pi\left(\varepsilon_{1,t+1},\varepsilon_{2,t+1}\right)\right]\right]>1$.

For the remainder of the discussion in this section assume that $\sigma_{1F}=\sigma_{2F}=\sigma_F$, $\sigma_{1I}=\sigma_{2I}=\sigma_I$, $\rho=0$ and $\beta_1=\beta_2=\beta$. Consider first the case when $\theta_{1,t-1}=\theta_{2,t-1}=0$. In this case, the arbitrageur trades the two assets symmetrically and the expected cross deviation is zero[19]. The expected cross deviation is also zero if either $\theta_{1,t-1}=0$ or $\theta_{2,t-1}=0$. In

[19]Let $z_1,z_2\in(0,\infty)$. The set of liquidity shocks to the two assets can be $(z_1,z_2)$, $(-z_1,z_2)$, $(z_1,-z_2)$ and $(-z_1,-z_2)$. If the correlation between the shocks is zero, the four outcomes are equally probable. For $\theta_{1,t-1}=\theta_{2,t-1}=0$, we have $\theta_{1,t}(-z_1,z_2)=\theta_{1,t}(-z_1,-z_2)=-\theta_{1,t}(z_1,z_2)=-\theta_{1,t}(z_1,-z_2)$ and

these cases there is zero price deviation for one of the assets, while there is a price deviation for the other asset, giving an expected cross deviation of zero.

Next, consider the case where $\theta_{1,t-1}$ and $\theta_{2,t-1}$ are of the same sign, that is, the arbitrageur holds initial long positions in both assets or he holds initial short positions in both assets. Assume that he has long positions in both the assets. In this case, part of the arbitrageurs trade is directed towards trading down to a flat position in both assets. For any magnitude of the liquidity shock to the two assets the arbitrageur always sells more of the asset if the shock is positive than he buys of the asset if the shock is negative. This results in an expected cross deviation across the two assets that is positive. The same results when the arbitrageur holds initial short positions in the two assets. The expected cross deviation is highest when the positions in the two assets are equal.

A similar line of reasoning gives the result that the expected cross deviation is negative when the arbitrageur is initially long one asset and short the other. In this case the arbitrageur in an attempt to trade down to a flat position is biasing his trade in the asset that he is long towards selling and towards buying in the asset that he is short. This gives an expected cross deviation in the assets that is negative.

<Insert Figure 5 here>

Figure 5 plots the expected cross deviation of price as a function of the initial position in the two assets. The horizontal axis is a measure of the relative position in the two assets, $\tan^{-1}\left(\frac{\theta_{2,t-1}}{\theta_{1,t-1}}\right)$ and goes from 0 to $2\pi$. The case where $\theta_{1,t-1} > 0$ and $\theta_{2,t-1} > 0$ is from 0 to

$\theta_{2,t}(z_1, z_2) = \theta_{2,t}(-z_1, z_2) = -\theta_{2,t}(-z_1, -z_2) = -\theta_{2,t}(z_1, -z_2)$. This implies that the expected cross-deviations computed over these four points is zero. Integrating over all possible combinations of $(z_1, z_2)$ gives that the expected cross deviation is zero.

$\frac{\pi}{2}$, $\theta_{1,t-1} < 0$ and $\theta_{2,t-1} > 0$ is from $\frac{\pi}{2}$ to $\pi$, $\theta_{1,t-1} < 0$ and $\theta_{2,t-1} < 0$ is from $\pi$ to $\frac{3\pi}{2}$ and $\theta_{1,t-1} > 0$ and $\theta_{2,t-1} < 0$ is from $\frac{3\pi}{2}$ to $2\pi$. The three lines in the graph correspond to initial positions, $\sqrt[2]{\theta_{1,t-1}^2 + \theta_{2,t-1}^2}$, that are 30%, 60% and 90% of maximum sustainable.

As can be seen from figure 5 the expected cross deviation is positive when the initial positions are of the same sign, between 0 and $\frac{\pi}{2}$ and $\pi$ and $\frac{3\pi}{2}$, and negative when the initial positions are of opposite signs, between $\frac{\pi}{2}$ and $\pi$ and $\frac{3\pi}{2}$ and $2\pi$. The line with the smallest variation corresponds to the case where the total initial position is 30% of maximum sustainable, while the line with the greatest variation corresponds to the case where the initial position is 90% of maximum sustainable.

## IV. Implications for the Regulation of Arbitrage Activities

In the last decade banks have greatly increased their holdings of traded assets. The increase in the relative importance of trading risk in bank portfolios poses new challenges to regulators. In 1996 the Basle Committee addressed the issue by requiring banks to hold capital equivalent to a percentage of their holdings in different asset categories, where the percentages are set to reflect the price volatilities of standard assets in the relevant categories. The benchmark used to define risk levels, value at risk (VaR) is a cut-off level that the loss over a given time horizon exceeds with some probability. The Bank for International Settlements (BIS), for example, sets this probability to 1% over a ten day period for purposes of determining the adequacy of bank capital. Banks themselves have developed value at risk measures for internal purposes, such as the J.P. Morgan's daily VaR at the 5% probability (95% confidence level).

As many have pointed out[20] whether the VaR of a firm's portfolio of positions is a good measure of the risk of financial distress over a short period depends on the liquidity of the portfolio, as well as the risk of adverse extreme cash outflows or of severe disruptions to market liquidity. Market liquidity is indeed a relevant precondition for VaR measures of risk to work, since in markets that lack the necessary depth, traders may face difficulties in rearranging their portfolio of positions. In more extreme circumstances, attempts to trade in a temporarily shallow market to ensure compliance with regulations may impose large losses on traders' capital, which is exactly what regulations are supposed to help avoid. Capital adequacy requirements which appear to work under normal circumstances then fail at times when they are most needed.

VaR measures currently in use seem to ignore that trading by large financial institutions can have an decisive impact on market prices in less than perfectly liquid markets. The common assumption in existing VaR models, that, price volatilities are exogenous is problematic for two reasons. First, VaR incorrectly measures the extent of potential forced reductions of the bank's capital at some confidence level. Second, it creates distortions. Below we show that this underestimates the required capital for some banks at the expense of overestimating the amount of capital required for other banks. The problem arises because the available regulatory benchmarks devote little attention to the fact that the actions of regulated firms affect the distribution of assets returns which serve as the basis in determining the minimum capital requirements. This is especially critical in the case of large banks, those for which good oversight is most important to safeguard the financial system.

[20] See, for example, Duffie and Pan (1997), page 9.

VaR models rely on the specification of random changes in the market prices of the assets and on a model for computing the sensitivity of the bank's capital to prices of the underlying securities held. To highlight the importance of building VaR models that correctly reflect trading risk of financially constrained institutions, we will separate the impact of financial constraints from the impact of the assumption that arbitrageurs do not effect prices when they execute their transactions.

Consider the case of the unconstrained arbitrageur presented in Section II.A. Gains and losses from trading are given by $\theta_{t-1}(F_t - F_{t-1})$, where $F_t$ is normally distributed with mean $I_{t-1}$ and variance $\frac{1}{4}(\frac{\sigma_F}{\beta} - \sigma_I)^2$. These gains and losses are also normally distributed, $N\left(\theta_{t-1}\left(I_{t-1} - F_{t-1}\right), \frac{\theta_{t-1}^2}{4}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)^2\right)$. The VaR is then simply given as

$$VaR = \theta_{t-1}\left(\left(I_{t-1} - F_{t-1}\right) + \frac{z_n}{2}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)\right) \tag{34}$$

where $z_n$ is the number of standard deviations from the mean that gives a loss in capital that is exceeded with probability $n$ over a period of time.

In the case of the constrained arbitrageur in Section II.B, gains and losses are expressed by the second term in expression (8), with $\theta_t^C$ in (18) entering the equation for $F_t$. Substituting gives an expression for the gains/losses:

$$\theta_{t-1}\begin{pmatrix} I_{t-1} - F_{t-1} - \theta_{t-1}\frac{(D(\varepsilon_t)-1)}{D(\varepsilon_t)} \\ +\frac{\delta}{2D(\varepsilon_t)}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)E\left[\pi(C_{t+1}\left(\varepsilon_{t+1}\right))\varepsilon_{t+1}\right] + \left[\frac{\sigma_F}{\beta} - \frac{1}{2D(\varepsilon_t)}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)\right]\varepsilon_t \end{pmatrix} \tag{35}$$

and the corresponding VaR measure that obtains is:

$$VaR = \theta_{t-1} \begin{pmatrix} I_{t-1} - F_{t-1} - \theta_{t-1} \frac{(D(z_n)-1)}{D(z_n)} \\ + \frac{\delta}{2D(z_n)} \left( \frac{\sigma_F}{\beta} - \sigma_I \right) E\left[ \pi (C_{t+1} (\varepsilon_{t+1})) \varepsilon_{t+1} \right] + \left[ \frac{\sigma_F}{\beta} - \frac{1}{2D(z_n)} \left( \frac{\sigma_F}{\beta} - \sigma_I \right) \right] z_n \end{pmatrix} \tag{36}$$

Consider now that the regulators in setting the minimum capital requirements overlook the fact that arbitrageurs' trades impact asset prices. These prices are then assumed to be given by $F_t = I_t + \frac{\sigma_F}{\beta} \varepsilon_t$ and in turn the gains and losses are assumed to be: $\theta_{t-1}(F_t - F_{t-1}) = \theta_{t-1} \left( (I_{t-1} - F_{t-1}) + \frac{\sigma_F}{\beta} \varepsilon_t \right)$. The amount of capital required that satisfies the VaR at the $n$ per cent level of confidence is then set equal to

$$C_t \geq M \theta_{t-1} \left( (I_{t-1} - F_{t-1}) + z_n \frac{\sigma_F}{\beta} \right)$$

where $M$ is an arbitrary coverage ratio. Note that the amount of capital is effectively like a margin requirement with a variable component and a fixed component. The first component depends on the price deviations assumed under no influence from arbitrageurs, and the fixed component depends both on an level of price volatility, assuming that arbitrage trading does not contribute to this volatility, and on the liquidity traders' demands. The only reason why this component is fixed lies on the incorrect assumption that arbitrageurs' trading strategies do not impact prices, and therefore it can be exogenously determined.[21]

[21] It is possible to see that the capital required to stay in business depends only on the previous two shocks, after replacing both $F_{t-1}$ and $I_{t-1}$ by their corresponding values:

$$CR_t \geq M \theta_{t-1} \left( \left( \sigma_I \varepsilon_{t-2} - \frac{\sigma_F}{\beta} \varepsilon_{t-1} \right) + z_n \frac{\sigma_F}{\beta} \right)$$

Financially constrained arbitrageurs take this VaR measure in their capital requirements when deciding on their optimal trading strategies, knowing that their trades impact on prices. The optimization problem in Section II.B remains the same, but where (9) is replaced by an equivalent capital adequacy constraint, $C_{t-1} + \theta_{t-1}(F_t - F_{t-1}) \geq M\theta_{t-1}[(I_{t-1} - F_{t-1}) + z_n \frac{\sigma_F}{\beta}]$ Expression (11) is then replaced by

$$V_C\left(\theta_{t-1}, C_t\left(\theta_t = 0\right)\right) = \begin{cases} \underset{\theta_t}{Max} \;\; \theta_{t-1}\sigma_I\varepsilon_t - \frac{(\theta_t - \theta_{t-1})^2}{\beta} & C_t > VaR_t \\ \quad -(\theta_t - \theta_{t-1})(\frac{\sigma_F}{\beta} - \sigma_I)\varepsilon_t & \\ \quad +\delta E\left[V_C\left(\theta_t, VaR_t\right)\right] & \\ \theta_{t-1}\sigma_I\varepsilon_t + \theta_{t-1}\left(\frac{\sigma_F}{\beta} - \sigma_I\right)\varepsilon_t - \frac{\theta_{t-1}^2}{\beta} & C_t \leq VaR_t \end{cases} \tag{37}$$

Solving for the maximum position that the trader can hold in the risky asset without violating the capital adequacy constraint, gives now $\bar{\theta}_t(C_{t+1}\left(\varepsilon_{t+1}\right)) = C_t/M(I_{t-1} - F_{t-1} + \frac{\sigma_F}{\beta}\varepsilon_{t+1})$, which can be seen to differ from (15). The optimal holding in the asset remains as the solution to (13), but where the weight $\pi$ is now different.

When regulators ignore the impact of arbitrageurs actions on the level and the volatility of prices the maximum sustainable position is a linear function of the capital required to stay in business. If followed the correct regulatory capital requirement would produce results consistent with those shown in Figure 1. The concavity of the correct function means that it lies above the linear function implied by the standard VaR capital adequacy rule that ignores the effects of arbitrage trading for small levels of capital, and it lies below the linear function for large levels of capital. What this means is that standard VaR rules force small institutions, those with lower levels of capital, to take less aggressive positions than they could if the correct measure was used. Similarly, standard VaR rules allow large institutions, those with higher levels of capital, to take more aggressive positions than they could if the correct capital requirement measure was used. This creates inefficiencies in the financial system because the social costs created by the unexpected default of a large institution, while it still seems solvent under the standard VaR measure, are much greater than those associated with the default of a small arbitrageur. The effects on competition are also important, since small arbitrageurs can justifiably complain than theirs is not a level playing field position.

The analysis performed here helps to see that a system of capital requirements that ignores that arbitrageurs affect prices has clear drawbacks in its treatment of trading risk. If after a negative shock, financially fragile securities firms are forced to liquidate positions to comply with regulatory requirements, markets can become more volatile and this can contribute to further losses in the value of the arbitrageurs' portfolios. This is also the case when a cash margin call precipitates sales which further depress prices. The intention of creating a stable environment that might have motivated the idea of capital adequacy standards is then put into question. This is of even greater importance since events in one market that affect the wealth of financially constrained arbitrageurs may be transmitted to other markets where nothing has fundamentally changed. In these situations what may have been considered a low risk arbitrage opportunity, due to the low correlations of the asset returns, can turn into a very risky investment. The correlations instead of being exogenously determined as assumed are dependent on the state of the arbitrageurs' financial health and events in one market that negatively impact the arbitrageurs capital may drive the correlations towards unity.

## V. Conclusion

We have developed a dynamic arbitrage strategy which takes into account the impact of arbitrage on prices, thus avoiding the need of risk aversion or position limits to prevent infinite arbitrage positions. The stochastic process for the asset price deviations has been endogenized in a way that appears to be consistent with empirical evidence.

The model consists of stochastic demand by liquidity traders that induces noise and trading by arbitrageurs that modifies the mispricing. The results demonstrate the importance of different competitive structures in the arbitrage market. They also suggest that financially constrained arbitrageurs may contribute to excess volatility in prices, and that they are responsible for the variation in the correlation of asset prices across different markets. In some cases arbitrageurs themselves contribute to the propagation of turbulence.

The findings also raise issues that are important from a regulatory view point. Specifically, it shows that measures that are used to determine the risk of arbitrage capital are not exogenous and are themselves influenced by the actions of arbitrageurs.

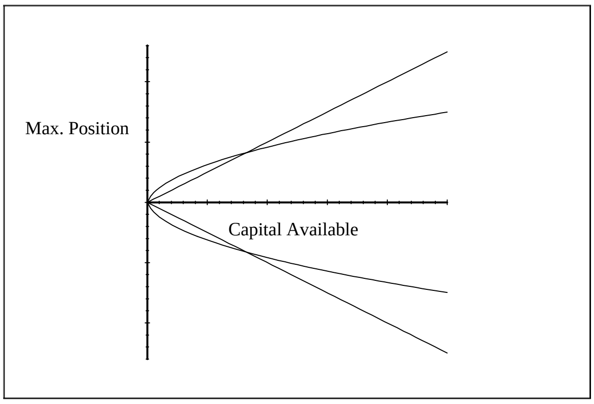



Figure 1: **Maximum sustainable position** as a function of the capital balance ($C_{t-1}$). The curved lines are the values of $\theta_{t-1}$ that solve (9) with the LHS set to zero for a given level of probability. The straight lines correspond to the maximum sustainable position for a given margin requirement.

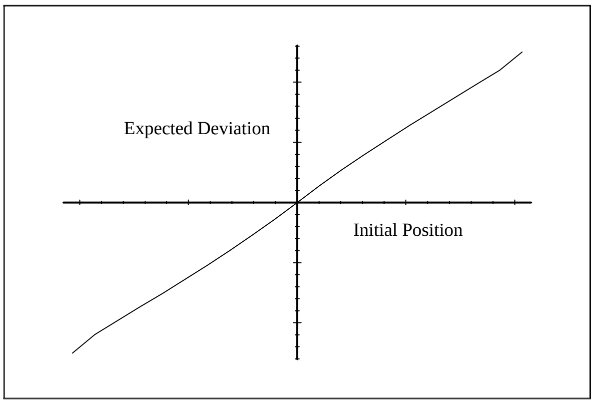



Figure 2: **Expected Price Deviation** as a function of the initial position.

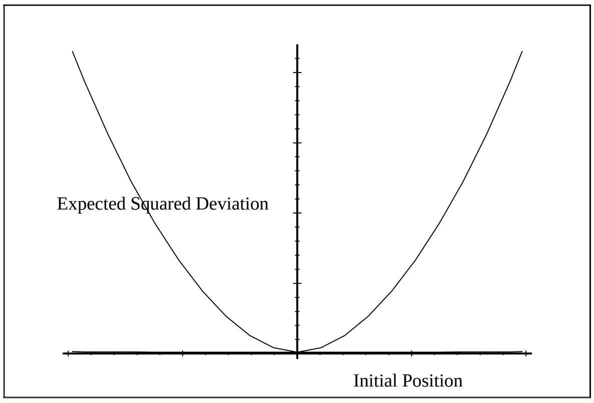



Figure 3: **Expected Squared Price Deviation and Variance of Price Deviation** as a function of the initial position. The variance of price deviation is parallel to the horizontal axis and is indistinguishable from it in the figure..

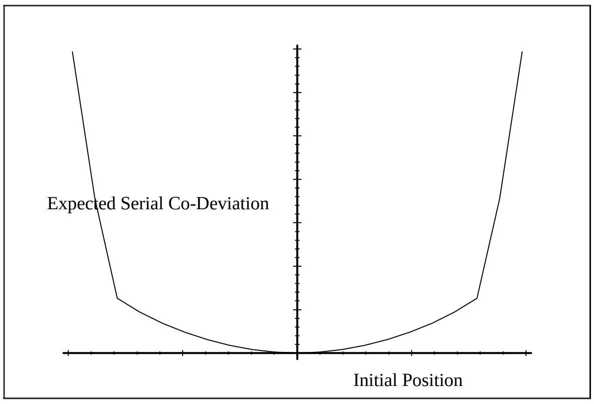



Figure 4: **Expected Serial Co-Deviation** as a function of the initial position.

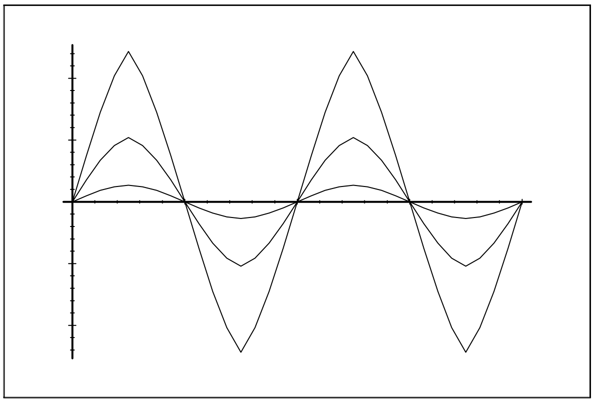


Figure 5: **Expected Cross Deviation** as a function of the arbitrageur's position in the two assets. The figure uses $\rho = 0$. The horizontal axis is $\tan^{-1}\left(\frac{\theta_{2,t-1}}{\theta_{1,t-1}}\right)$ over the range $(0, 2\pi)$. The three lines in the figure are for three different levels of $\sqrt[2]{\theta_{1,t-1}^2 + \theta_{2,t-1}^2}$, corresponding to 30%, 60% and 90% of maximum sustainable position.

## References

[1]Almgren, R., and N. Chriss, 1998, Optimal liquidation, mimeo.

[2]Brennan , M., and E. Schwartz, 1990, Arbitrage in Stock Index Futures, *Journal of Business,* 63, 7-31.

[3]Dow, J. and G. Gordon, 1994, Noise trading, delegated portfolio management, and economic welfare, NBER Working Paper 4858.

[4]Duffie, D. and J. Pan, 1997, An overview of Value at Risk, *Journal of Derivatives*, 4,7-49.

[5]Gennotte, G., and H. Leland, 1990, Market liquidity, hedging and crashes, *America Economic Review,* 80, 999-1021.

[6]Grossman, S., and J.-L. Vila, 1992, Optimal dynamic trading with leverage constraints, *Journal of Financial and Quantitative Analysis*, 27(2), 151-168.

[7]Kodres, L. E., and M. Pritsker, 1998, A rational expectations model of financial contagion, working paper, *Board of Governors of the Federal Reserve Bank.*

[8]MacKinlay, A. C., and K. Ramaswamy, 1987, Program trading and the behavior of stock index futures contracts, *Review of Financial Studies*, 1(1), 137-158.

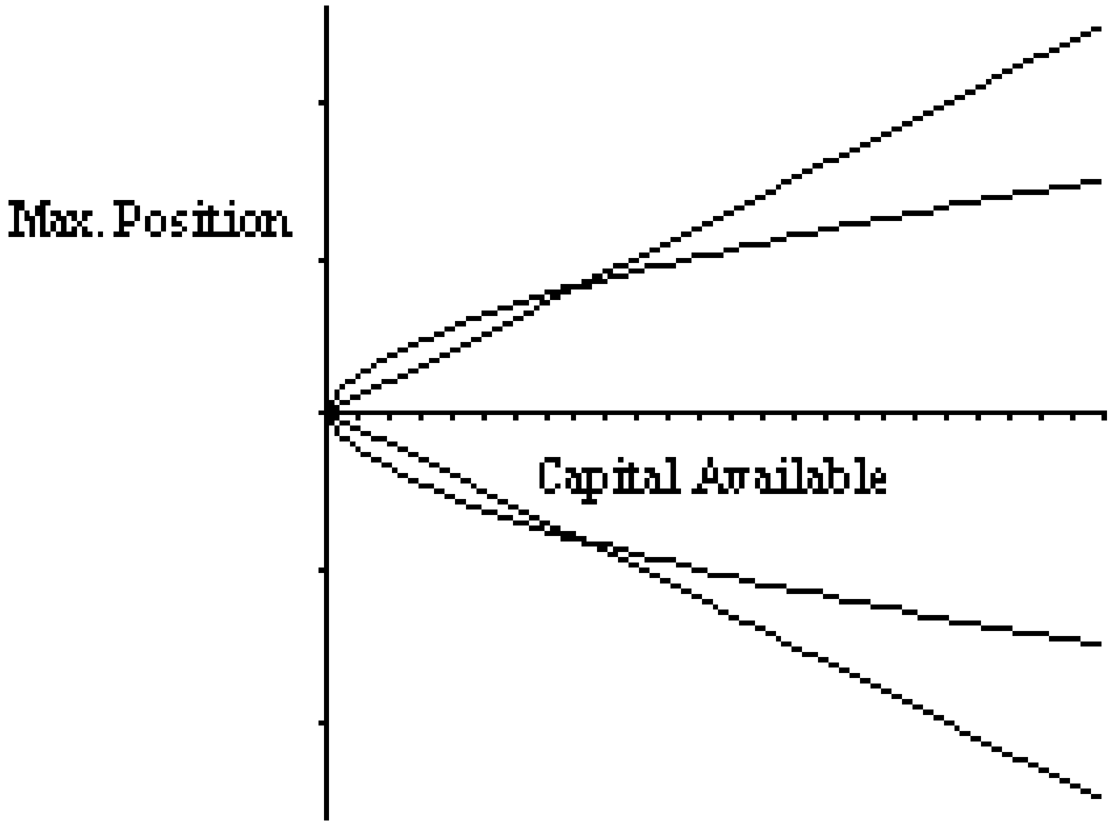
[9]Merrick, J., 1989, Early unwindings and rollovers of stock index futures arbitrage program analysis and implications for predicting expiration day effects, *Journal of Futures Markets*, 9, 101-111.

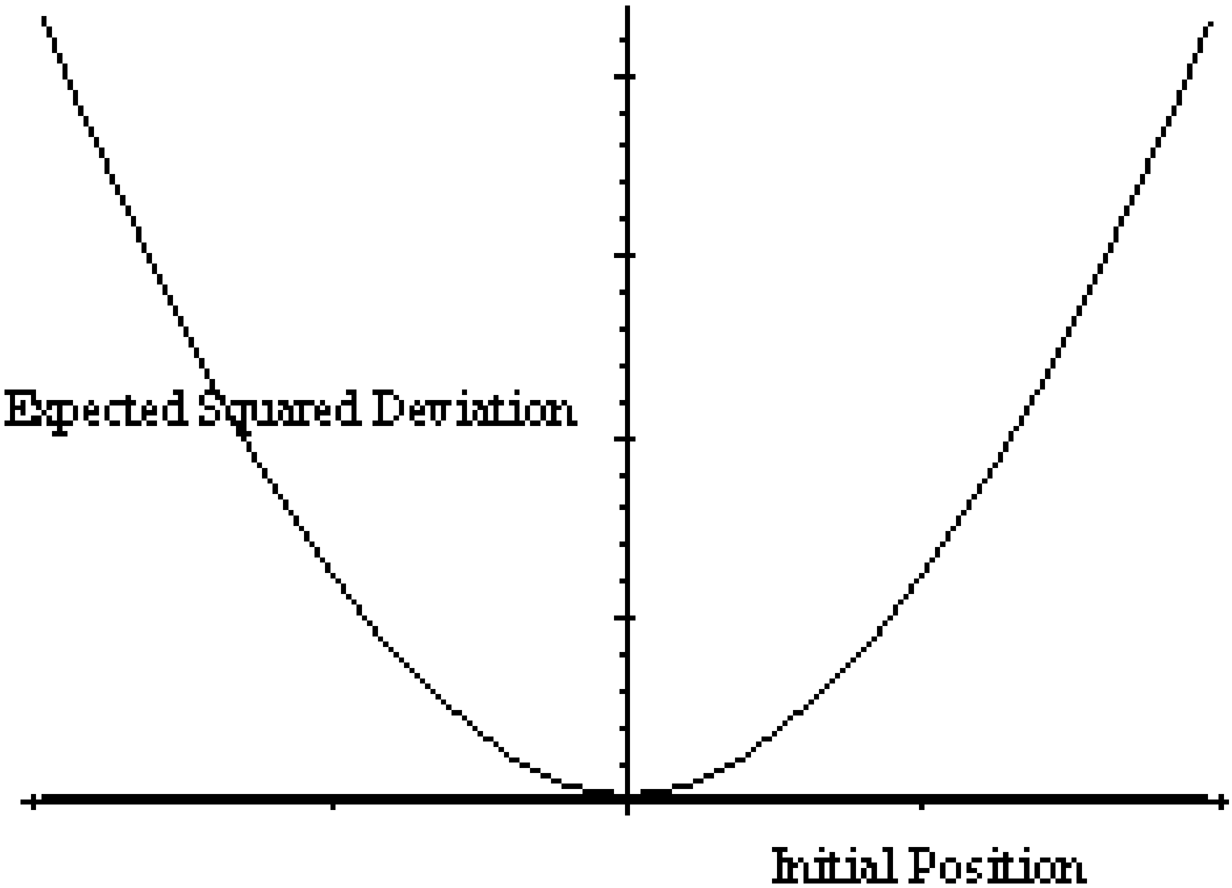
[10]Schleifer, A., and R.W. Vishny, 1997, The limits of arbitrage, *Journal of Finance*, 52(1), 35-55.

[11]Tuckman, B., and J.-L. Vila, 1992, Arbitrage with holding costs: a utility based approach, *Journal of Finance*, 47(4), 1283-1302.

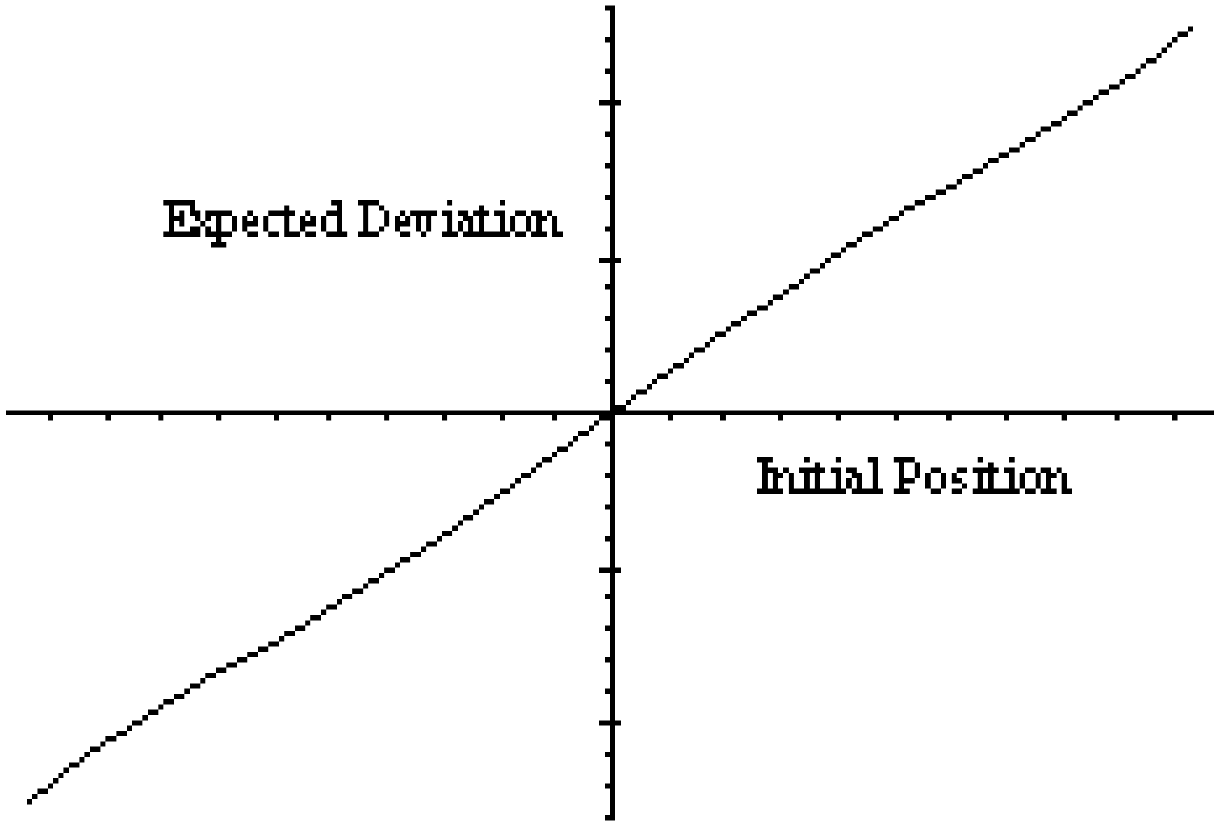
[12]Yuan, K., 1999, Asymmetric price movements and borrowing constraints: a REE model of crisis, contagion and confusion, mimeo.

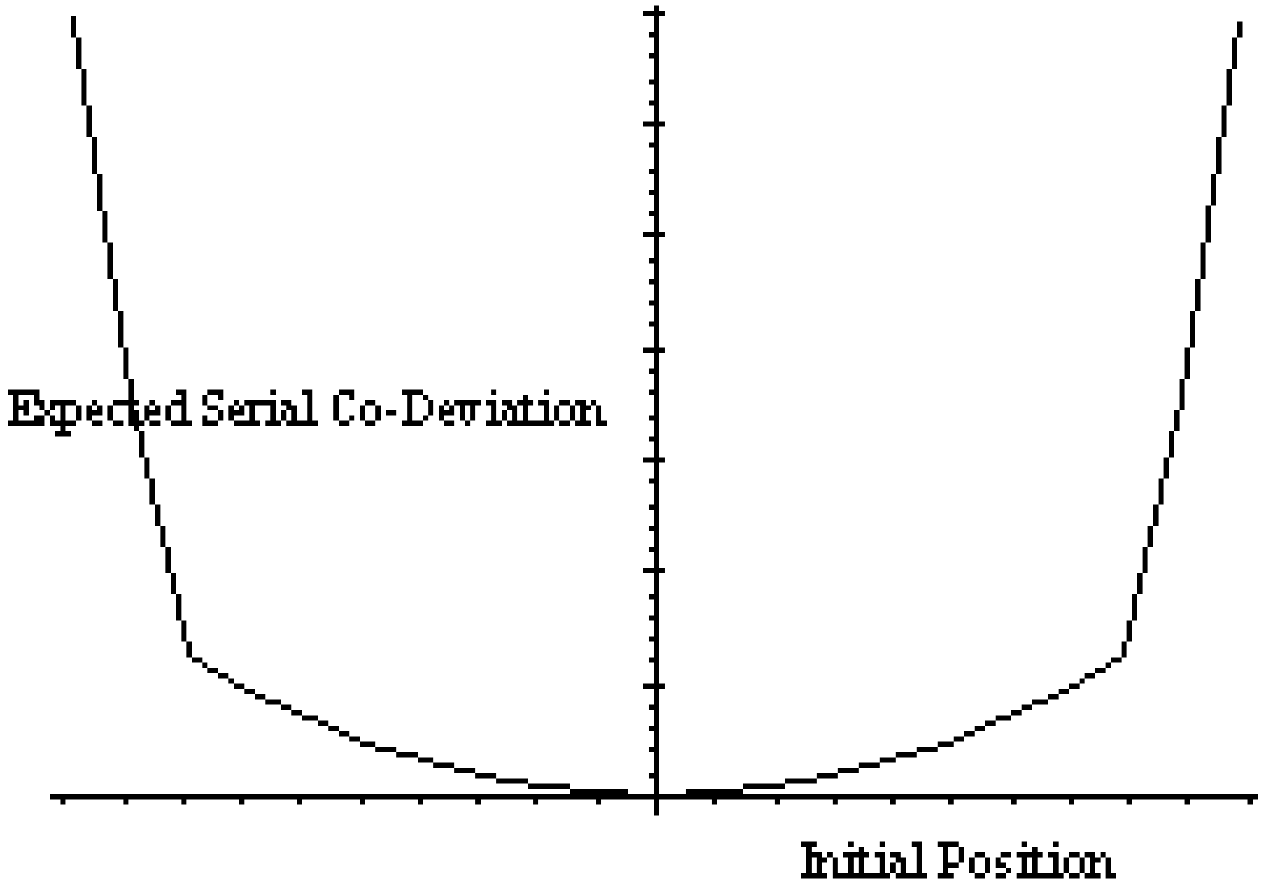
[13]Zariphopoulou, T., 1994, Consumption-investment models with constraints, *SIAM Journal of Control and Optimization*, 32(1), 59-85.


Max. Position
Capital Available



Expected Squared Deviation
Initial Position


Expected Deviation
Initial Position


Expected Serial Co-Deviation
Initial Position