The More Efficient, the More Vulnerable!*

Dong-Hyun Ahn† Soohun Kim‡ Kyoungwon Seo§

April 3, 2019

Abstract

By extending Shleifer and Vishny (1997), we show that, while arbitrageurs enhance the market efficiency during calm periods, they can make the financial market more vulnerable to crashes. These crashes are amplified through their leveraged positions, which are ironically rationalized by the market efficiency that they bring forth. We empirically examine such implications in the U.S. interest rate swap markets. Consistent with the predictions of our model, our efficiency metrics, the mean-reversion speeds of the spreads, are strongly associated with the tail behaviors. We also show that the return characteristics of hedge fund returns are coherent with our model.

JEL classification codes: G01, G23, G28

Keywords: fixed income, hedge fund, leverage, market efficiency, resilience

---

*We are thankful to seminar participants at KAIST, the University of Hawaii, Georgia Tech, Korea University, the Chinese University of Hong Kong and the the Asia Pacific Industrial Organization Conference. Seo gratefully acknowledges the financial support of the Institute of Finance and Banking, and the Institute of Management Research at Seoul National University. Any remaining errors are solely ours. The usual disclaimer applies.

†Department of Economics, Seoul National University, Seoul, South Korea. E-mail: ahnd@snu.ac.kr.

‡Scheller College of Business, Georgia Institute of Technology, Atlanta, GA. E-mail: soohun.kim@scheller.gatech.edu.

§Business School (Finance), Seoul National University, Seoul, South Korea. E-mail: seo8240@snu.ac.kr.

1 Introduction

It is the calm and silent water that drowns a man.

–Ghanaian proverb

The arbitrage transactions, which attempt to monetize the discrepancy of a particular asset's market price from its fundamental value, play a critical role of policing and eliminating such discrepancy. The textbook version of the efficient market hypothesis argues that such arbitrage transactions are strong enough to eliminate any market disequilibrium *instantaneously*. Kyle (1985) defines the market resilience, a facet of liquidity, as the speed with which prices revert to their equilibrium level after a large shock in the transaction flow. The activity of arbitrage transactions *per se* is the key determinant of the market resilience; the more active the arbitrage trades, the more resilient the market becomes. Therefore, textbook efficiency implies the 'infinite' speed of market resilience, which is durable only if arbitrage transactions can be implemented in a perfect market, i.e., no market frictions. However, Shleifer and Vishny (1997) argue that such textbook arbitrage is at odds with reality. They theoretically explore why arbitrageurs may be slow in eliminating the mispricing with agency frictions. The literature has proposed various frictions such as constraints on leverage or short-sale as potential causes for limiting the effectiveness of arbitrageurs.

In this paper, we add a new insight to the 'limits to arbitrage' literature by showing that the enhanced efficiency due to the activity of arbitrageurs during normal time can ironically amplify the market instability during crash period. To do this, we introduce an intertemporal aspect of arbitrage activity. In particular, we extend a version of the Shleifer and Vishny model by adding one more time period. Such an extension is essential to reflect the nature of path-dependency through the wealth effect in endogenous variables such as leverage ratios and market prices. Our model delivers a number of interesting implications for security prices.

Firstly, we show that the past performance of arbitrageurs matters for the security price in a 'crash' state. In a crash state, the arbitrageur has to take a substantially high leverage to take advantage of mispricing. However, this is the very state where the leverage constraint is most binding. How much leverage the arbitrageur can take critically depends upon how much capital loss or gain has been accumulated from past trades. Consequently, the performance of the arbitrageur preceding the crash has a critical impact on the security price. This explains why the response of market price can drastically differ to shocks of seemingly similar sizes.

Secondly, when the evolution of states entails a higher probability of reversion from a mispricing state to a fair pricing state during a normal time, the security price plunges more severely in a crash period. When the security price shows a stronger reversion to a fair price during a tranquil time, arbitrageurs are willing to take a highly leveraged position to exploit any mispricing opportunities. However, once a large shock arrives contrary to the arbitrageurs' expectation, the effect of arbitrageurs on the market price is amplified through their highly leveraged positions.

This striking result necessitates a new approach to market efficiency. The literature on the side of efficient market considers the market efficiency as given (Sharpe (1964); Ross (1976); Lucas (1978)), and the literature on the behavioral finance have offered various causes for the overall market inefficiency (Shleifer and Vishny (1997); Gromb and Vayanos (2002); Brunnermeier and Pedersen (2008)). Although they take the opposite stance of each other in a sense, most of existing studies commonly view the efficiency/inefficiency as a single characteristic of a given market. In contrast, we separate market efficiency in normal times and that in crisis times to examine the relation of those two. In particular, our model shows that a seemingly more efficient market under a normal market condition would be dismantled more severely during times of crises.

Thirdly, we find that the past performance of arbitrageurs matters more for the security price in a 'crash' state when the security price shows a stronger reversion to a fair price during a tranquil time. The intuition directly follows from the first two implications. Given the stronger mean reversion property of prices, arbitrageurs take a more aggressive position. In turn, this leads to a wider variation in the past performance of arbitrageurs. Hence, when a crash arrives, the wider variations in levels of arbitrageur capital lead to more erratic security prices conditional on a crisis.

We empirically investigate the aforementioned implication of our model using two sets of data – (i) the U.S. interest rate swap market and (ii) hedge fund style indexes. We explore the interest rate swap market, one of the most preferred habitat of the arbitragers, to investigate the pricing implication of our model. In contrast, we examine the hedge fund style indexes in order to investigate the distinctive return characteristics of the arbitragers signified by their demands implied by our model.

Let us discuss how popular these so-called 'swap curve trades' are in practice. Figure 1.1 illustrates a replication of a 'heat map' of slope spreads as of November 23, 2018, which was published by one of the top global investment banks.[1] Each cell reports the

[1]Almost all of global IBs report these kinds of heat maps on slope and butterfly spreads on a daily basis. For example, JP Morgan publishes 'JP Morgan Basic ChartPac,' Deutsche Bank publishes RV

Figure 1.1: Heat Map of Slope Spreads

Swap Slopes

3m Z-Scores 11/23/2018

	Spot	1m	3m	6m	1y	2y	3y	5y	10y	15y	20y
2s/3s	-2.2	-2.1	-1.9	-1.3	1.8	2.5	1.9	1.7	0.8	-1.1	2.6
2s/5s	-1.9	-1.7	-1.1	0.2	2.3	2.2	1.9	2.1	0.7	-0.7	1.4
2s/7s	-1.5	-1.1	-0.1	1.3	2.2	2.1	2.0	2.0	0.1	-0.5	1.3
2s/10s	-0.6	-0.2	1.0	1.8	2.2	2.1	2.0	2.0	-0.2	0.2	1.0
2s/30s	0.9	1.3	1.8	2.0	2.2	2.1	2.0	2.0	-0.7	-1.6	-1.9
3s/5s	-1.1	-0.7	0.6	1.6	2.3	2.0	1.8	2.3	0.6	-0.2	0.8
3s/7s	0.0	0.4	1.4	1.9	2.1	2.0	2.0	2.1	-0.1	-0.1	1.0
3s/10s	0.9	1.3	1.8	2.0	2.2	2.0	2.0	2.1	-0.4	0.5	0.7
3s/30s	1.7	1.9	2.0	2.1	2.2	2.1	2.0	2.1	-0.8	-1.6	-2.0
5s/7s	1.2	1.4	1.7	1.9	1.9	2.0	2.1	1.8	-0.7	0.0	0.9
5s/10s	1.6	1.7	1.9	2.0	2.0	2.0	2.0	1.9	-0.7	0.7	0.5
5s/30s	1.9	2.0	2.0	2.1	2.1	2.1	2.1	1.9	-1.1	-1.7	-2.1
7s/10s	1.8	1.9	2.0	2.1	2.1	2.0	1.9	2.1	-0.8	1.1	-0.3
7s/30s	1.9	2.0	2.1	2.1	2.1	2.1	2.0	1.9	-1.1	-1.8	-2.0
10s/20s	2.0	2.0	2.1	2.1	2.1	2.2	2.1	1.8	0.5	-1.8	-1.8
10s/30s	2.0	2.0	2.1	2.1	2.1	2.1	2.0	1.5	-1.2	-2.1	-1.9
20s/30s	1.9	2.0	2.0	2.0	2.0	1.9	1.7	0.1	-1.8	-2.2	-2.0

Forward horizons are on the horizontal axis, while vertical axis indicates the terms

Notes: The table shows a heat map of the normalized slope spreads on November 23, 2018. For example, 2s/3s on the spot is 2.2 standard deviations below from the mean, based on the past 3 months. Darker cells indicate more deviations from the mean.

3-month z-score of a particular slope spread. The 3-month z-score is the standardized normal z-value of the slope spread based on the past three-month estimates of mean and standard deviation. For example, the left top corner cell reports -2.2, which means that the current yield differential between 3 year swap and 2 year swap is 2.2 standard deviation lower than its past 3 month mean. This implies that the spread may be too low, or equivalently, the slope between the two tenors is too flat. As such, some arbitragers may speculate on a curve steepener; paying the 3 year swap and receiving the 2 year swap under the expectation that the spread will revert to its mean.[2] The payoff from the trade is, in general, very small and hence, leverage is widely employed

(Relative Value) PDF package, and Royal Bank of Scotland publishes ‘RBS Daily Chart Pack.’ See also Priaulet (2008) and Tata (2006) for in-depth reviews of a variety of relative value trades.

[2]The length of data employed in estimating the z-score could vary across hedge funds and prop desks. The most popular choice would be 3 months, 6 months and 1 year but some hedge funds (with longer investment horizon) use much longer data.

to magnify potential gains, typically five to fifteen times the asset base's value.

Furthermore, the SDR (Swap data repositories) data reveals the popularity of swap curve trades. The SDRs are new entities created by the Dodd-Frank Wall Street Reform and Consumer Protection Act ("Dodd-Frank Act") in order to provide a central facility for swap data reporting and recordkeeping. Under the Dodd-Frank Act, all swaps, whether cleared or uncleared, are required to be reported to registered SDRs.[3] According the SDR data, slope trades alone account for about ten percent of volumes reported to US SDRs up to August 22 during 2017.[4] In addition, a huge variety of curve trades are reported. In January 2017 alone, 83 different pairs of tenors were traded more than once a day. Thus, arguably the swap curve trades are the most popular trading strategies among fixed-income arbitragers.

We exploit the U.S. interest rate swap market for our empirical study as follows. Using thirteen different tenors from one year to thirty years, we construct various strategies: 78 slope spreads and 286 butterfly spreads of the U.S. interest rate swap from July 23, 1998 to May 11, 2017. We measure the efficiency of each of 364 spreads (78 slope spreads plus 286 butterfly spreads) by the mean reversion speed in AR(1) process of the normalized z-score. Then, we relate the efficiency measure to cross-sectional and time-series implications on tail behaviors from our model.

The empirical study shows an overall compliance with the theoretical predictions. For cross-sectional implications, we regress the various tail properties such as kurtosis, VaR, expected shortfall on the mean-reversion speed. We find that the mean-reversion speed of the spreads is strongly associated with those tail properties. For example, we find that a spread with stronger mean reversion features much higher kurtosis. These results imply that the more seemingly efficient spreads with faster mean-reversion are more susceptible to extreme change in their values, thereby more vulnerable to a crash, which our model predicts.

For time-series implications, we empirically examine the wealth effect in the spread markets through quantile panel regressions. We document the strong effect of the past performance itself as well as its interactions with mean-reversion spread on the tail distribution. Our findings support the predictions of our model: following a path with loss, a market (especially the one with strong resiliency during normal time) becomes extremely vulnerable to large shocks.

Moreover, we provide the empirical evidence of arbitrageurs' activity in various

[3]See the website of the U.S. Commodity Futures Trading Commission for further details: `https://www.cftc.gov/Forms/sdrforms.html`.

[4]`https://www.clarusft.com/curve-trading-in-usd-swaps`.

hedge fund strategies. Because the hedge fund industry keeps developing novel strategies to find mispricing opportunities, there are wide variations in terms of market environment where hedge fund managers implement their strategies. Using the monthly returns from Barclay Hedge Fund Index for sixteen style indexes over the sample period of Jan, 1997 to Aug, 2017, we document a strong relation between frequency of positive returns, interpreted as a tendency of mean reversion to the fair value during normal period, and higher moments, representing the vulnerability to crashes. Especially, our analysis partially explains some return characteristics observed in convertible bond or fixed income arbitrages—picking up nickels in front of a steamroller.

A vast literature has proposed various causes of market inefficiency such as costly information acquisition (Grossman and Stiglitz (1980)), noise traders (De Long et al. (1990)), delegated portfolio management (Shleifer and Vishny (1997)), margin requirements (Brunnermeier and Pedersen (2008)) and coordination failure (Ahn et al. (2017)). We contribute to this literature by proposing an alternative view on the causes of the market efficiency. In particular, we show that the vulnerable market in a crisis can be caused by the market efficiency in normal periods. Hence, by reducing the frictions in the existing literature, we may experience more severe crashes! To examine the interaction of market efficiency/vulnerability across states through arbitrageurs, we revamp Shleifer and Vishny (1997) by extending another time period to explore the path-dependency of the market price and separating out a crash state, to focus on the mispricing in the crash state.

The empirical effect of arbitrageurs' limited activity on the financial market has been an ongoing issue in literature. Mitchell and Pulvino (2001) analyze mergers to characterize the risk and return in risk arbitrage and Mitchell et al. (2002) investigate situations where the market value of a company is less than its subsidiary. Kapadia and Pu (2012) propose limits to arbitrage as an explanation for a low correlation between equity and credit markets and test it. Brunnermeier et al. (2008) attribute the crash risk of currencies to sudden unwinding of carry trades triggered by the lack of funding liquidity. Mancini Griffoli and Ranaldo (2012) investigate how covered interest parity broke down in the aftermath of the global financial crisis by focusing on the funding liquidity. Recently, Jermann (2017) shows that negative swap spread, which implies a risk-free arbitrage opportunity, can be explained by introducing frictions for holding bonds. We contribute to this literature by providing the evidence on the interaction of market efficiency through the channel of arbitrageurs' activity in fixed-income arbitrages and hedge fund strategies. In particular, we find that the assessment of Duarte et al. (2006) on the fixed-income market needs to be reconsidered by extending the sample period to include 2007-2008 crisis period. Kahraman and Tookes (2017) find the positive effect of margin trading on liquidity in general but show that the effect reverses during crises. To the best of our knowledge, existing empirical literature including the aforementioned studies has not investigated a direct relationship between a positive environment for arbitrageurs during normal periods and a tail behavior of arbitrage payoffs.

This paper is organized as follows. In Section 2, we build up a four-period arbitrage transaction model wherein a sequence of equilibrium prices is determined. Section 3 analyze the impact of limits of arbitrage on the efficiency and the vulnerability. The empirical analysis on the major predictions of the model using the U.S. interest rate swap data and hedge fund style indexes is examined in Section 4. Proofs are deferred to Internet Appendix. Section 5 concludes.

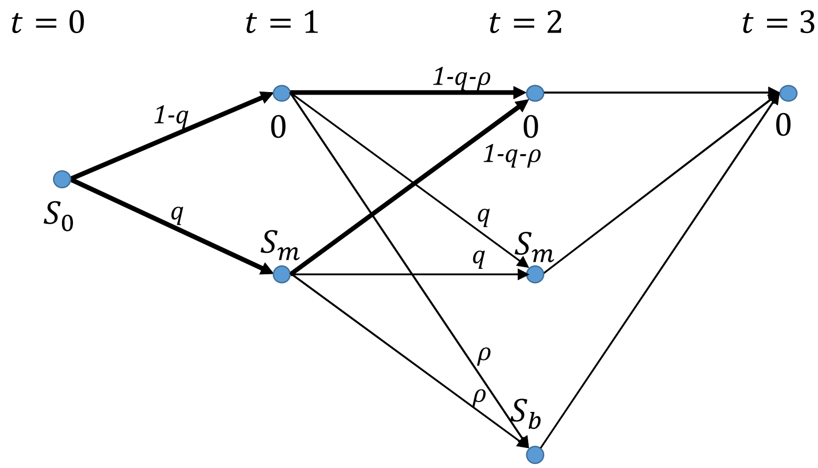
2 The Model

The basic structure of our model follows Shleifer and Vishny (1997) where arbitrageurs exploit underpricing caused by selling pressures from pessimistic noise traders. We extend their model in two aspects: (i) we add one more time period to explore the path-dependency of the market price and (ii) we separate out a crash state, where a pessimistic demand shock is very large, to highlight the effect of market efficiency during normal, or non-crash, states on the mispricing in the crash state. These differences are explained in more detail below.

There are four time periods ($t = 0$, 1, 2 and 3) over which a specific asset of our interest is traded. The aggregate supply of the asset is normalized to 1. The asset pays out the constant amount of V at $t = 3$, and hence there is no long run fundamental risk in this asset. There are three types of market participants: noise traders, arbitrageurs and financiers from which arbitrageurs can borrow short term. The noise traders trade for liquidity reasons not related to the asset's fundamental value thereby making deviation of its market price from the fundamental value. Without loss of generality, we assume that the amount of deviation is random but non-positive; that is, noise traders experience pessimistic shocks as in Shleifer and Vishny.[5] In contrast, the arbitrageurs, the only market participants who know the fundamental value of the

[5]The main result of our analysis holds in a symmetric way when the noise traders may experience optimistic shocks, as long as the funding structure is symmetric between long and short positions.

Figure 2.1: Noise Trader Shocks



Notes: This figure shows the state space of pessimistic noise trader shocks.

asset at $t = 0, 1, 2$, attempt to monetize the mispricing of the asset which is triggered by the noise traders. When the arbitrageurs do not have sufficient capital, they use short term loans provided by the financiers. The financiers do neither observe nor understand the strategy of the arbitrageurs, and they charge higher lending rates as the leverage of arbitrageurs increases, which will be clarified later.

We specify the structure of noise trader shocks. The state space of negative noise trader shocks is illustrated in Figure 2.1. Therein the amount of the shock at $t = 0$ is $S_0 > 0$, which is known to the arbitrageurs at $t = 0$, but the noise trader shocks in the subsequent periods are uncertain. The state space at $t = 1$ is binomial such that the amount of the shock is either zero with probability $1 - q$ or S_m with probability q. Going forward to $t = 2$, the state space is trinomial; the amount of the shock is 0, S_m and S_b with probability $1 - q - \rho$, q and ρ, respectively, regardless of the state at $t = 1$. So we assume path-independence of conditional probability of each state. By this assumption, we make sure that path-dependency of the equilibrium price to be shown later is not due to the exogenously given transition probabilities. Finally, the market price equals V at $t = 3$ as the asset pays out V.

We necessitate the given setup of noise trader shocks in comparison to Shleifer and Vishny (1997) . Figure 2.1 exhibits the state space assumed in their model as thick lines, which is nested in our state space. As such, we add one more time period and one more state at $t = 2$. We consider such extensions for the following two reasons. First, by adding another state at $t = 2$, we can separate the crash state with the extremely 'bad' shock S_b from the intermediate state with the reasonably 'moderate' shock S_m. In addition, the moderate size shock, S_m, may occur at $t = 1$ and $t = 2$ whereas the extremely bad shock, S_b, comes into being only once over the entire periods in our analysis. Hence, the state of S_b can be thought of as a tail event in terms of its magnitude as well as its probability of occurrence. In contrast, the state of S_m or 0 can be regarded as a tranquil or normal time period. Second, adding one more time period enables us to analyze the impact of the past performance on the market price in the crash state. For example, the third crash state with S_b at $t = 2$ can be reached either through the first state with a zero noise trader shock or the second state with S_m at $t = 1$. Note that if the third state at $t = 2$ is reached via the first (second) state at $t = 1$, the arbitrageur earns a profit (makes a loss) at $t = 1$ with $0 < S_0 < S_m$. Then, the intuition of Shleifer and Vishny kicks in. They assume that the available investment amount of the arbitrageur is an increasing function of her past return and they call this 'performance-based arbitrage'. In our model, instead, we assume that the funding cost, charged by the financiers, is proportional to the amount of leverage. Consequently, if an arbitrageur earned a positive profit in the prior period, she affords to utilize more short term loans because the funding cost is cheaper and vice versa. Hence, the intensity of arbitrage activity depends on her past performance as in Shleifer and Vishny, which in turn affects the severity of mispricing in the crash state.

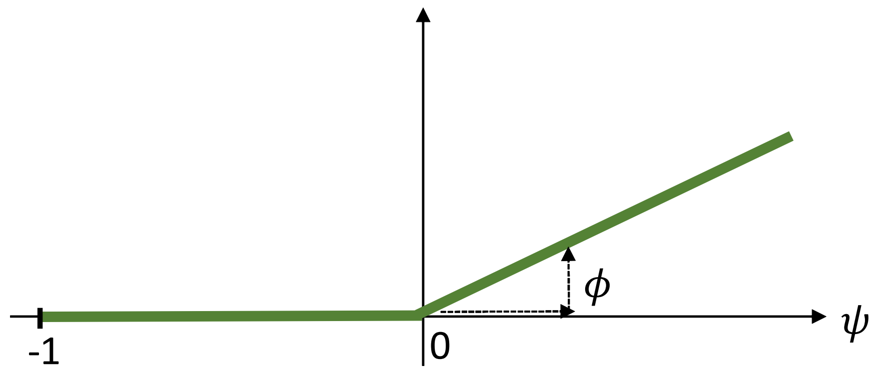
The financiers in our model play the role of providing short term funds to the arbitrageurs. Note that the financiers do not understand the underlying risk of the assets that the arbitrageurs are betting on. Hence, to hedge the delinquency of arbitrageurs, the financiers transact with the arbitrageurs only at the short term basis and also charge higher lending rates for highly leveraged arbitrageurs.[6] To incorporate the realistic features of funding costs that leveraged arbitrageurs face in the financial market, we assume the following structure of short term funding costs.[7]

Assumption 1. *We assume the following funding cost structure. Suppose that an*

[6]In practice, a hedge fund may have an ongoing long-term relationship with its prime brokerage. However, the brokerage firm examines the margin on a daily basis and often adjust the interest rate based on the balance of the margin account. Hence, it is reasonable to examine their relationship with short-term contracts.

[7]See Brunnermeier and Pedersen (2008) for the main sources of leverage for hedge funds. All we need in our model is the feature that the previous negative performance of the arbitrager reduces the arbitrager's leverage capacity in the next period. We can solve a model based on an alternative assumption that the arbitragers are required to deliver more collateral. The main results under this assumption are qualitatively equivalent, but the equilibrium relies upon corner solutions as opposed to interior solutions in our model.

Figure 2.2: Funding Cost Structure



Notes: This figure shows the the funding rate as a function of leverage ratio, $\psi = \frac{L}{W}$. If leverage is employed, $\psi > 0$, a sensitivity of the funding rate to the leverage is determined by $\phi > 0$.

arbitrageur's capital is W and she borrows $L \geq -W$. Then, the funding rate is assumed to be proportional to the leverage ratio, $\psi = \frac{L}{W}$; e.g,

$$c(\psi) = r + \phi\psi 1_{\psi>0}, \tag{2.1}$$

where r is the risk-free rate and $1_{\psi>0}$ is an indicator variable with a value of 1 if leverage is employed, $\psi > 0$, and 0 otherwise. A sensitivity of the funding rate to the leverage is determined by $\phi > 0$. The range of leverage is given by $\psi \in [-1, \infty)$. Without loss of generality, we set $r = 0$.

Figure 2.2 illustrates the funding cost structure given by Assumption 1. When an arbitrageur invests in the risk free asset, or $\psi = \frac{L}{W} < 0$, it provides a normalized rate of zero. When the leverage ratio $\psi = \frac{L}{W}$ is positive, the funding rate increases at the rate of $\phi > 0$. As will be clear later, this simplified structure for the funding cost allows us to avoid corner solutions without losing economic intuitions underlying the model.

Lastly, we introduce the arbitrageurs to our model economy. There is a continuum of arbitrageurs with unit mass. At $t = 0$, the i-th arbitrageur, $i \in [0, 1]$, starts running her fund with the capital of $W_{i,0}$ and in every leverage decision, she maximizes the expected terminal size of fund $W_{i,3}$ at $t = 3$. Let $W_0 = \int_0^1 W_{i,0} di$. Because we restrict our attention to a symmetric equilibrium, we drop the index of arbitrageurs and solve the problem of the aggregate arbitrageurs with W_0 at $t = 0$ who maximize the expected terminal fund size as a price taker.

The strategy of an arbitrageur is a choice of leverage subject to the funding cost given in Assumption 1. Before we characterize the optimal strategy of arbitrageurs, we need to do two preliminary works. First, we need to extend the state space in Figure 2.1 to properly consider the path-dependency. Second, we guarantee the positivity of the arbitrageurs' capital so that they can take meaningful leverage in every state. We handle the first issue as follows. Figure 2.1 shows that there are seven spots defined in the space of (time, noise trader shock). Tracking all possible paths over the seven spots, we find that there are sixteen nodes.[8] We use x as an index for the sixteen nodes. In particular, for the two nodes at $t = 1$, we use $x = 1i$ for $i = 1, 2$, and for the six nodes at $t = 2$, we use $x = 2j|1i$ for $i = 1, 2$, and $j = 1, 2, 3$. The initial node is denoted by $x = 0$ and the terminal nodes are by $x = 3|2j|1i$ for $i = 1, 2$, and $j = 1, 2, 3$. When we do not need to identify nodes at t, we use the time index t to denote the set of all nodes at t.

Next, we provide sufficient conditions for the positive capital of the arbitrageurs. To ensure that the capital of the arbitrageurs is positive but not too large to eliminate the effect of the noise trader shock, we need the following regularity conditions.

Assumption 2. *It holds that* $W_0 < S_0 \leq S_m < S_b < V$, $\phi > \frac{S_m(2S_b - S_m)}{4(V-S_m)(V-S_b)}$ *and* $W_0 \frac{V}{V-S_0} \frac{V}{V-S_m} \left(1 + \frac{S_0^2}{4\phi(V-S_0)V}\right)\left(1 + \frac{S_m^2}{4\phi(V-S_m)V}\right) < S_m$.

The first inequality formalizes the interpretation of the state space described in Figure 2.1. The initial noise trader shock, S_0, is set to be bigger than the initial capital, W_0. The noise trader shock in the mediocre state, S_m, is strictly smaller than that in the crash state, S_b. The next two inequalities restrict that the sensitivity of funding costs to leverage, ϕ, is not too small and the initial capital, W_0, is not too big. The lower bound on ϕ is not very restrictive relative to empirically observed cost structures. For example, in Treasury security markets, it would be very extreme to set S_m and S_b to 5% and 10% of V, respectively but in such a case, ϕ is restricted to be greater than

$$\frac{S_m\left(2S_b - S_m\right)}{4\left(V - S_m\right)\left(V - S_b\right)} = 0.0022.$$

Then, with the leverage $\psi = 1$, the funding cost should be at least 0.22% higher, which is a plausible restriction. The upper bound of the initial capital is necessary so that the capital is not enough to cover the demand shock. The following lemma shows that Assumption 2 provides a sufficient range of parameters for the analysis of the effect of arbitrageurs on the mispricing.

[8] Note that $16 = 1 + 2 \cdot 1 + 3 \cdot 2 \cdot 1 + 1 \cdot 3 \cdot 2 \cdot 1$.

Lemma 2.1. *Under Assumption 2, it holds that* $0 < W_x < S_m$ *for all nodes* x.

The lemma shows that the arbitrageur capital is positive but not too large. In particular, from Assumption 2 and Lemma 2.1, whenever $S_x > 0$, we have $W_0 < S_0$ at $t = 0$ and $W_x < S_x$ at $t = 1, 2$. Thus, the arbitrageurs can absorb partially the noise trader shocks with the positive capital but not completely.

Now, we are ready to describe the optimal leverage decision of arbitrageurs in a specific node x. Let V_x denote the optimized expected terminal fund size, $\mathbb{E}_x[W_3]$, evaluated at node x. We solve for V_x backward. When node x is a final node, it holds that $V_x = W_x$. Fix a non-final node x and assume that we already solved $V_{x'}$ for every node x' following node x. Then, V_x is related to $V_{x'}$ through the leverage decision given by $V_x = \max_{\psi_x} \mathbb{E}_x[V_{x'}]$. It is convenient to introduce $v_x = \frac{V_x}{W_x}$, the expected terminal value of one dollar evaluated at node x. For a final node x, because $V_x = W_x$, it holds that $v_x = 1$. Then, given P_x, $P_{x'}$ and $v_{x'}$, an arbitrageur solves at a non-final node x,

$$V_x = \max_{\psi_x} \mathbb{E}_x[V_{x'}] = \max_{\psi_x} \mathbb{E}_x[W_{x'} v_{x'}].$$

Reflecting the funding cost in Assumption 1, we can write the capital evolution as

$$W_{x'} = W_x \left(\frac{P_{x'}}{P_x} (1 + \psi_x) - \psi_x \left(1 + \phi \psi_x 1_{\psi_x > 0}\right) \right). \tag{2.2}$$

This implies

$$\begin{aligned}
V_x =& \max_{\psi_x} \mathbb{E}_x \left[W_x \left(\frac{P_{x'}}{P_x} (1 + \psi_x) - \psi_x \left(1 + \phi 1_{\psi_x > 0}\right) \right) v_{x'} \right] \\
=& W_x \max_{\psi_x} \mathbb{E}_x \left[\frac{P_{x'} v_{x'}}{P_x} (1 + \psi_x) - \psi_x \left(1 + \phi 1_{\psi_x > 0}\right) v_{x'} \right] \\
=& W_x \max_{\psi_x} \left(\frac{\mathbb{E}_x[P_{x'} v_{x'}]}{P_x} (1 + \psi_x) - \psi_x \left(1 + \phi 1_{\psi_x > 0}\right) \mathbb{E}_x[v_{x'}] \right) \\
=& W_x \mathbb{E}_x[v_{x'}] \max_{\psi_x} \left(\frac{\Gamma_x}{P_x} (1 + \psi_x) - \psi_x \left(1 + \psi_x \phi 1_{\psi_x > 0}\right) \right)
\end{aligned}$$

where

$$\Gamma_x = \frac{\mathbb{E}_x[P_{x'} v_{x'}]}{\mathbb{E}_x[v_{x'}]}. \tag{2.3}$$

Thus,

$$v_x = \mathbb{E}_x[v_{x'}] \max_{\psi_x} \left(\frac{\Gamma_x}{P_x} (1 + \psi_x) - \psi_x \left(1 + \psi_x \phi 1_{\psi_x > 0}\right) \right) \tag{2.4}$$

and

$$\psi_x = \arg\max_{\psi} \left(\frac{\Gamma_x}{P_x} (1 + \psi) - \psi \left(1 + \psi \phi 1_{\psi > 0}\right) \right). \tag{2.5}$$

We interpret the process Γ_x given by (2.3) as the ratio of the expected terminal value of the risky asset to that of cash. If the arbitrageur holds one unit of the risky asset at node x, its dollar value at a following node x' will be $P_{x'}$. Because the expected terminal value of each dollar at node x' is $v_{x'}$, the numerator of Γ_x, $\mathbb{E}_x\left[P_{x'}v_{x'}\right]$, is the expected terminal value of one unit of the risky asset. If, instead, the arbitrageur holds one dollar at node x, its expected terminal value at a following node x' will be $v_{x'}$. Thus, the denominator of Γ_x, $\mathbb{E}_x\left[v_{x'}\right]$, is the expected terminal value of one dollar. Hence, the aforementioned interpretation is justified.

Once we characterize the leverage decision problem as (2.5), we obtain the following lemma whose proof is straightforward and omitted.

Lemma 2.2. *Given Γ_x and P_x, the arbitrageur's optimal leverage ratio is*

$$\psi_x = \begin{cases} \frac{1}{2\phi}\left(\frac{\Gamma_x}{P_x}-1\right) & \text{if } \Gamma_x > P_x \\ \text{any number in } [-1,0] & \text{if } \Gamma_x = P_x \\ -1 & \text{otherwise} \end{cases}.$$

The intuition behind the above lemma is straightforward. Recall that Γ_x is the ratio of the expected terminal value of the risky asset to that of cash. Hence, if $\Gamma_x > P_x$, the expected terminal value of the risky asset is greater than the current value of that. However, due to the quadratic funding cost of $\psi\left(1+\psi\phi 1_{\psi>0}\right)$, she picks $\frac{1}{2\phi}\left(\frac{\Gamma_x}{P_x}-1\right)$ as the optimal leverage, which increases with the ratio $\frac{\Gamma_x}{P_x}$ and decreases with the sensitivity of funding rate to the leverage ϕ. When $\Gamma_x \leq P_x$, the expected terminal value of the risky asset is less than the current value of that, and hence a positive leverage cannot be optimal.

Given the process P_x, we can solve for Γ_x and v_x as well as ψ_x backward from (2.3), (2.4) and Lemma 2.2. Then, the capital process W_x is determined forward by (2.2). We determine the process P_x in the following subsection.

2.1 Equilibrium

We pin down the equilibrium of the economy by imposing market clearing conditions. Furthermore, we will show the existence and uniqueness of the equilibrium.

Consider the market clearing condition at node x. The demand from noise traders is $\frac{V-S_x}{P_x}$ and the demand from arbitrageurs is $\frac{W_x(1+\psi_x)}{P_x}$. Because the aggregate supply of the risky asset is normalized to 1, it should hold that $1=\frac{V-S_x}{P_x}+\frac{W_x(1+\psi_x)}{P_x}$, which is equivalently expressed as

$$P_x = V - S_x + W_x\left(1+\psi_x\right). \tag{2.6}$$

We solve for the optimal leverage and the market clearing price by combining Lemma 2.2 and the condition of (2.6). In describing the equilibrium, it is convenient to use $\widetilde{P}(S, W, \Gamma)$, the unique positive solution to the equation,

$$P = V - S + W(1+\psi),$$

where $\psi = \frac{1}{2\phi}\left(\frac{\Gamma}{P} - 1\right)$. Note that

$$\widetilde{P}(S, W, \Gamma) \equiv \frac{V - S + W\left(1 - \frac{1}{2\phi}\right) + \sqrt{\left(V - S + W\left(1 - \frac{1}{2\phi}\right)\right)^2 + 2\frac{W\Gamma}{\phi}}}{2}. \tag{2.7}$$

Begin with nodes at $t = 2$.

Lemma 2.3. *For each node x at $t = 2$, given the capital W_x at node x, the equilibrium price and leverage are*

$$\psi_x = \begin{cases} -1 & \text{if } S_x = 0 \\ \frac{1}{2\phi}\left(\frac{V}{P_x} - 1\right) & \text{otherwise} \end{cases}$$

$$P_x = \begin{cases} V & \text{if } S_x = 0 \\ \widetilde{P}(S_x, W_x, V) < V & \text{otherwise} \end{cases}$$

where $\widetilde{P}(S, W, \Gamma)$ is defined in (2.7).

At $t = 2$, when there is no shock ($S_x = 0$), the asset price maintains its fair value. If $S_x > 0$, the arbitrageur capital is not enough to cover the shock and thus the equilibrium price is undervalued at $P_x = \widetilde{P}(S_x, W_x, V) < V$.

We move on to nodes at $t = 1$ and $t = 0$. First, the equilibrium at node 11 without a noise trader shock is determined as follows.

Lemma 2.4. $(\psi_{11}, P_{11}) = (-1, V)$.

The above lemma shows that there is no mispricing in the asset and hence no investment from arbitrageurs at node 11.

Next, we proceed to nodes 12 and 0. We say a positive parameter θ is sufficiently small (or small enough) if there is a constant $\overline{\theta} > 0$ such that $\theta < \overline{\theta}$. We assume sufficiently small probabilities for the positive noise trader shocks, i.e. small q and ρ, for the rest of the paper. This restriction is justified in that we are interested in analyzing a modern financial market where the arbitrage opportunities are highly likely to vanish. Finally, the following lemma pins down the equilibrium leverage and prices.

Lemma 2.5. *If q and ρ are sufficiently small,*

$$(\psi_{12}, P_{12}) = \left(\frac{1}{2\phi}\left(\frac{\Gamma_{12}}{P_{12}} - 1\right), \widetilde{P}\left(S_m, W_{12}, \Gamma_{12}\right)\right) \text{ and}$$
$$(\psi_0, P_0) = \left(\frac{1}{2\phi}\left(\frac{\Gamma_0}{P_0} - 1\right), \widetilde{P}\left(S_0, W_0, \Gamma_0\right)\right).$$

Moreover, $P_{12} < V$ and $P_0 < V$.

By the Lemma, we see that whenever the shock exists, the asset is undervalued. Note that Lemmas 2.3, 2.4 and 2.5 describe $(\psi_x, P_x)_x$ as functions of $(W_x)_x$:

- (ψ_x, P_x) at $t = 2$ is computed from W_x at $t = 2$ (Lemma 2.3)
- $(\psi_{11}, P_{11}) = (-1, V)$ is determined (Lemma 2.4)
- Γ_{12} obtains from (W_x, ψ_x, P_x) at $t = 2$ (Equation (2.3))
- Compute (ψ_{12}, P_{12}) from Γ_{12} and W_{12} (Lemma 2.5)
- Γ_0 comes from (W_x, ψ_x, P_x) at $t = 1$ (Equation (2.3))
- Γ_0 gives (ψ_0, P_0) (Lemma 2.5)

Then, (2.2) with given W_0 produces a new capital process. Thus, we have an equilibrium when the initial given capital process equals the new capital process computed as above. The following shows existence and uniqueness of the equilibrium.

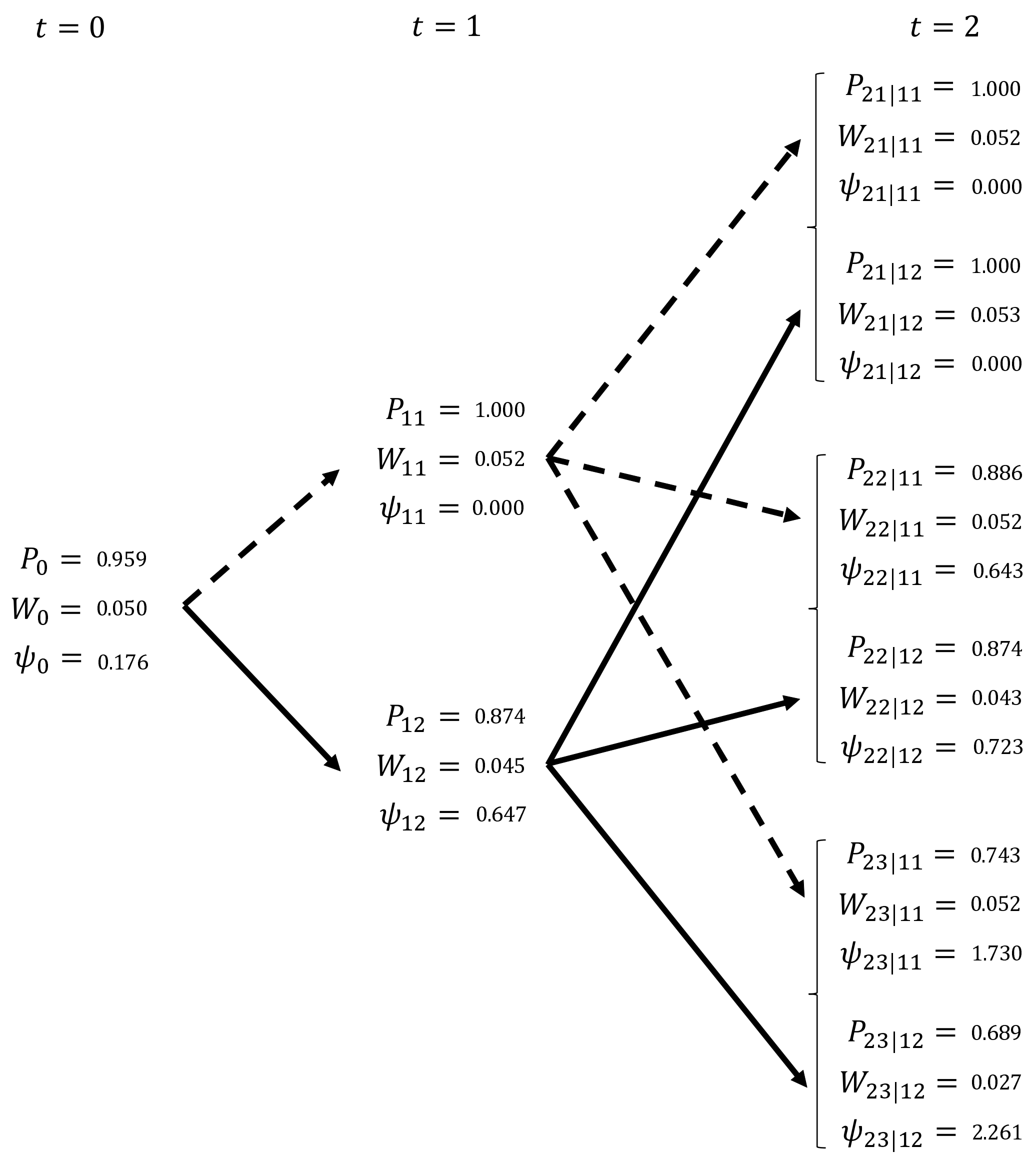
Theorem 2.1. *If q and ρ are sufficiently small, there exists a unique equilibrium.*

Next, we illustrate our model with a numerical example of our economy. We use the following values of parameters:[9]

$$V = 1,\ W_0 = 0.05,\ S_0 = 0.1,\ S_m = 0.2,\ S_b = 0.4,\ \phi = 0.1 \text{ and } q = 5\%,\ \rho = 1\%. \tag{2.8}$$

Note that the moderate noise trader shock $S_m = 0.15$ arrives with probability $q = 5\%$ at $t = 1, 2$ and the extreme noise trader shock $S_b = 0.5$ occurs with probability $\rho = 1\%$ only at $t = 2$. The sensitivity of the funding rate to the leverage is set to be $\phi = 0.1$. For example, the arbitrageur borrows 100% of his capital, its funding rate is 10%. It is easy to check that the combination of the above parameter values satisfy the conditions in Assumption 2, which ensures the positivity of the arbitrageur capital. The numerical

Figure 2.3: Optimal Leverage, Equilibrium Prices, Arbitrageurs' Capital



Notes: This figure illustrates the optimal leverage ratio, ψ, the equilibrium price, P, and the corresponding capital of the arbitrageur, W, across all nodes. Parameters are set as $V = 1,\ W_0 = 0.05,\ S_0 = 0.1,\ S_m = 0.2,\ S_b = 0.4,\ \phi = 0.1,\ q = 0.05$ and $\rho = 0.01$.

solutions to the optimal leverage ratio ψ_x, the equilibrium price P_x and the resulting capital of the arbitrageur W_x are illustrated in Figure 2.3.

At $t = 0$, the (representative) arbitrageur borrows 17.6% of her capital ($=\psi_0$). The resulting equilibrium price of P_0 is 0.959. In the absence of her arbitrage transaction, the price would be 0.9 ($= V - S_0 = 1 - 0.1$). Thus, her investment itself raises the market price by 0.059.

If the noise trader shock disappears at $t = 1$, the market price recovers its fair value, $P_{11} = V = 1$ and her capital increases to 0.052 with 4.8% gain. She is away from the market because the expected future value of the risky asset in the state of S_0 at $t = 1$ is smaller than the market price of V (see Lemma 2.4). In contrast, if the shock S_m realizes at $t = 1$, the market price (P_{12}) drops to 0.874; the realized return of the security is -8.9%. However, she loses more than that due to two driving forces: the leverage itself and the additional funding cost. Altogether her loss is 10% and her capital declines to 0.045. At this state, she employs leverage up to 64.7% to monetize the enlarged undervaluation.

In the state where $S_{21} = 0$ is realized at $t = 2$, the equilibrium price is at its fair value and the arbitrageur leaves the market because of zero expected return. If this state is realized via S_{11}, the first state at $t = 1$, the arbitrageur's capital $W_{21|11}$ is identical to W_{11}. In contrast, if the state comes through S_{12}, the second state at $t = 1$, the arbitrageur's capital $W_{21|12}$ has increased from 0.045 to 0.053.

At $S_{22} = S_m$, the amount of the negative noise shock at this state is identical to $S_{12} = S_m$. The magnitude of the shock is mediocre and could happen before at $t = 1$ so we interpret this state as a tranquil state. If this state is realized via S_{11} at which no investment in the risky asset is made, her capital ($W_{22|11}$) is still W_{11}, 0.052. The market price ($P_{22|11}$) is 0.886 and she employs a leverage ratio of 64.3% ($\psi_{22|11}$). In contrast, if this state is arrived at through S_{12} in which the arbitrageur has already experienced a loss, her capital ($W_{22|12}$) is 0.043; the market price ($P_{22|12}$) is 0.874 and her leverage ratio is 72.3%.

Note that we interpret S_{23} as a 'crash' state given its massive amount of the negative shock ($S_b = 0.4$) coupled with an extremely low probability of occurrence ($\rho = 1\%$). In addition, this amount of the negative noise shock is unprecedented so it could be counted as an exceptionally rare event. At S_{23} via S_{11}, the arbitrageur's capital ($W_{23|11}$) is 0.052 again. The equilibrium price ($P_{23|11}$) is 0.743 and the arbitrageur steps up the leverage ratio to 1.730 ($\psi_{23|11}$). In contrast, at the same state via S_{12}, the arbitrageur

[9] Rather extreme numbers are used (e.g., $\phi = 0.1$ is too large) in the example to emphasize the intuitions of the equilibrium.

becomes penurious; her capital $W_{23|12}$ is merely 0.027, which means 46% of her original capital is wiped out! As such, despite her extensive leverage ratio ($\psi_{23|12} = 2.261$), the equilibrium price $P_{23|12}$ is as low as 0.689.

The results deliver an essential implication on the activity of arbitrageurs. Despite the arbitrageur's aggressive leverage, the price is not boosted much; the equilibrium price is only 0.689 whereas its value without arbitrage transaction is $V - S_b = 0.6$. The arbitrageur's capital is extremely low and consequently she does not afford to make a large amount of investment in the asset. Specifically, her total amount of investment is as small as 0.089 ($= P_{23|12} - (V - S_b) = 0.689 - 0.6$ by (2.6)). This is similar to 0.086 ($= P_{22|11} - (V - S_m) = 0.886 - 0.8$ by (2.6)) at S_{22} via S_{11} despite the fact that the arbitrage opportunity is much more favorable at S_{23} via S_{12} that at S_{22} via S_{11}. Note that the crash itself is a double-edged sword. On one hand, the asset price plunges so that the arbitrageur's capital is precipitated. On the other hand, the setback in the asset price delivers an extraordinary opportunity for arbitrage. However, the funding cost structure limits her leverage capacity; when outside funding is most needed, the funding cost is most binding.

3 Model Prediction

In this section, we examine the impact of q on the features of the equilibrium in a crash state. Note that q is the probability of the moderate noise trader shock S_m. Hence, we interpret $1 - q$ as the strength of reversion to the fair pricing state or, more generally, the efficiency of a financial market during a normal time.

We start our analysis by establishing the importance of past in determining the price in the crash state.

Theorem 3.1. *If q and ρ are sufficiently small, $P_{23|11} > P_{23|12}$.*

In a crash state with extreme noise trader shock S_b, the arbitrageur will take a substantially high leverage to take advantage of mispricing because the mispricing is highly likely to disappear. However, due to the progressive funding rate given in 1, the size of leverage the arbitrageurs can take critically depends on the size of arbitrageurs' capital, which is determined by capital loss or gain from past trades. Hence, the responses of the market price to similar shocks may differ depending on the path of the history. In our model economy, the arbitrageurs experience gains in node 11 but they experience losses in node 12. The latter case introduces a disadvantage in the funding rate compared to the first case and thus we have the above theorem.

Next, we examine the effect of q on the severity of mispricing in the crash state. The following two theorems state that as the strength of reversion to the fair pricing state becomes stronger or the market becomes more efficient in a normal state, the more of the market price diverges away from the fundamentals in a crash state.

Theorem 3.2. *If q and ρ are sufficiently small,* $\frac{dP_{23|11}}{dq} > 0$.

Theorem 3.3. *If q, ρ and W_0 are sufficiently small, then* $\frac{dP_{23|12}}{dq} > 0$.

The small W_0 depicts the nature of arbitrageurs holding a relatively small capital but using extensive leverage to exploit arbitrage opportunities with a greater scale. The intuitions of Theorems 3.2 and 3.3 are described in Figure 3.1. We use the same parameters in (2.8) except q, which takes a value between 0% and 10%.

First, we examine Theorem 3.2. Note that Figure 3.1(a) shows that arbitrageurs take higher leverage as the force of reverting back to the fundamentals increases, or q decreases. Then, P_0 moves along with ψ_0 because investing more in the risky asset induces a higher price, which is shown in Figure 3.1(b). Besides, higher P_0 means a less profitable opportunity which leads to a lower level of capital W_{11} with a smaller q as in Figure 3.1(c).[10] Lastly, from $W_{23|11} = W_{11}$, $P_{23|11}$ also moves in the same direction of W_{11} (Figure 3.1(d)).

We turn to Theorem 3.3. Suppose q gets smaller. With higher leverage of ψ_0 (Figure 3.1(a)) and the increased pessimism of noise traders, $S_m > S_0$, W_{12} decreases as shown in Figure 3.1(e). Although they have smaller capital W_{12}, they take more aggressive leverage due to the belief on the stronger market efficiency from smaller q (Figure 3.1(f)). As a result, the capital in the crash state $W_{23|12}$ following two consecutive losses from S_0 to S_{12} and from S_{12} to S_{23}, is wiped out more severely (Figure 3.1(g)), and the price $P_{23|12}$ diverges more away from the fundamentals (Figure 3.1(h)).

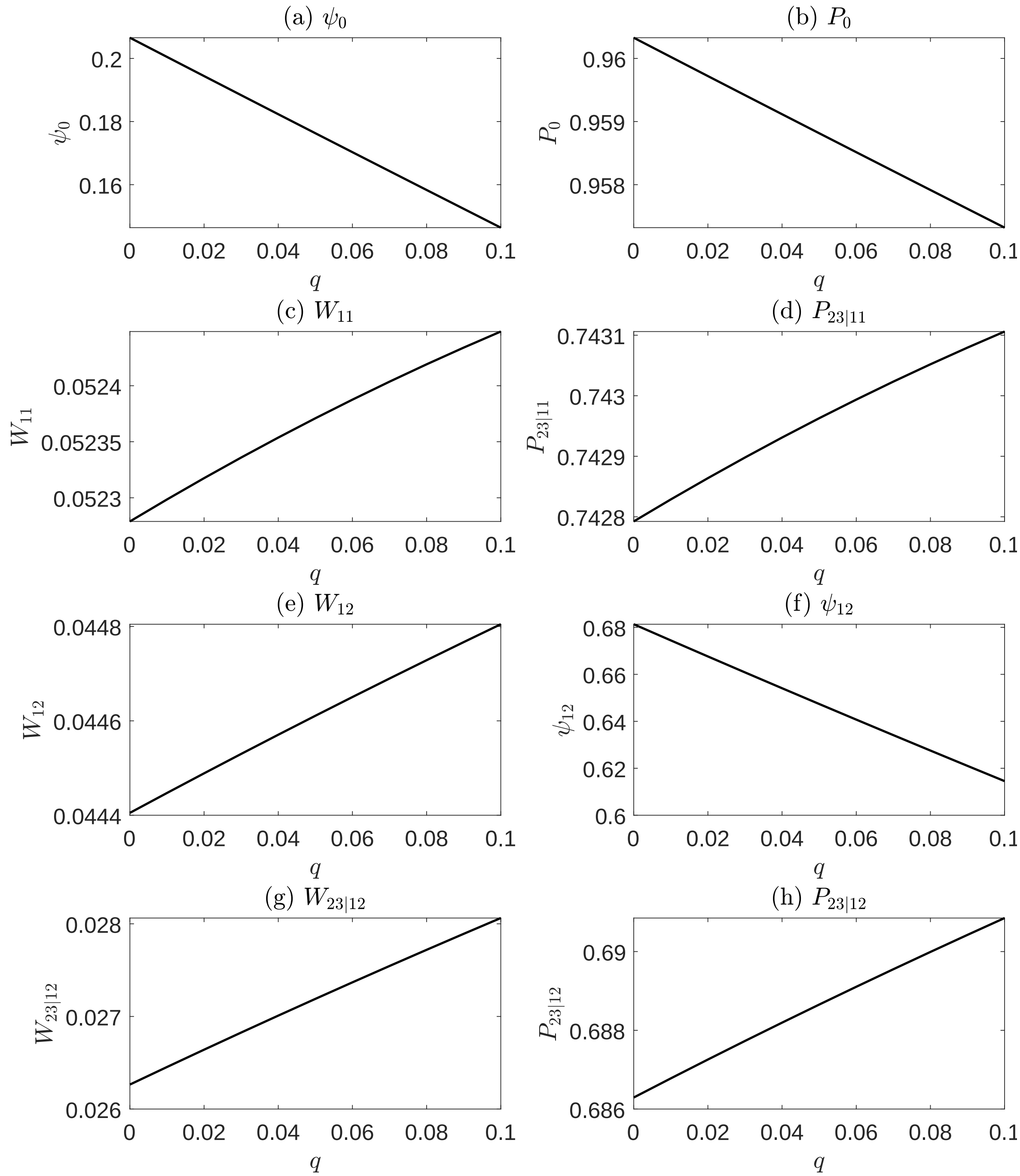
Lastly, we examine the effect of a more efficient market in a normal state on the path-dependency of mispricing in the crash state. In the context of our model, we study the sensitivity of $P_{23|\cdot}$, the price of security in the worst state at $t = 2$, to whether the worst state follows from $S_{11} = 0$ or $S_{12} = S_m$ at $t = 1$. The following theorem shows that the path-dependency of mispricing is stronger in a more efficient market.

Theorem 3.4. *If q, ρ and W_0 are sufficiently small, then* $\frac{d}{dq}\left(P_{23|11} - P_{23|12}\right) < 0$.

Suppose q decreases. Recall that arbitrageurs take higher leverage at $t = 0$ as the force of reverting back to the fundamentals increases, as Figure 3.1(a) shows. Then,

[10] Higher ψ_0 has a positive effect on W_{11} but it is dominated by the effect of P_0. The reason is that $\frac{\partial W_{11}}{\partial \psi_0}$ is very small when q is small. This is proved in Internet Appendix.

Figure 3.1: The effect of q on the equilibrium



Notes: This figure depicts the effect of q on the equilibrium outcomes of ψ_0, P_0, W_{11}, $P_{23|11}$, W_{12}, ψ_{12}, $W_{23|12}$ and $P_{23|12}$. Other parameters are $V = 1$, $W_0 = 0.05$, $S_0 = 0.1$, $S_m = 0.2$, $S_b = 0.4$, $\phi = 0.1$ and $\rho = 1\%$.

$W_{11} - W_{12} > 0$ becomes larger because gains are realized at node 11 and losses worsen at node 12. This is illustrated in Figure 3.2(a). Furthermore, as q decreases, arbitrageurs take higher leverage at node 12 (Figure 3.1(f)) while investing nothing in the risky asset at node 11. Hence, $W_{23|11} - W_{23|12} > 0$ increases further (Figure 3.2(b)). As a result, Figure 3.2(c) shows that the path-dependency in the mispricing $\left(V - P_{23|12}\right) - \left(V - P_{23|11}\right) = P_{23|11} - P_{23|12}$ becomes larger with a smaller q.

4 Empirical Analysis

For the rest of this section, we empirically test Theorems 3.1-3.4 established in the previous section. In particular, we examine the properties of spread strategies in fixed income market (Section 4.1) and various hedge fund style indexes (Section 4.2).

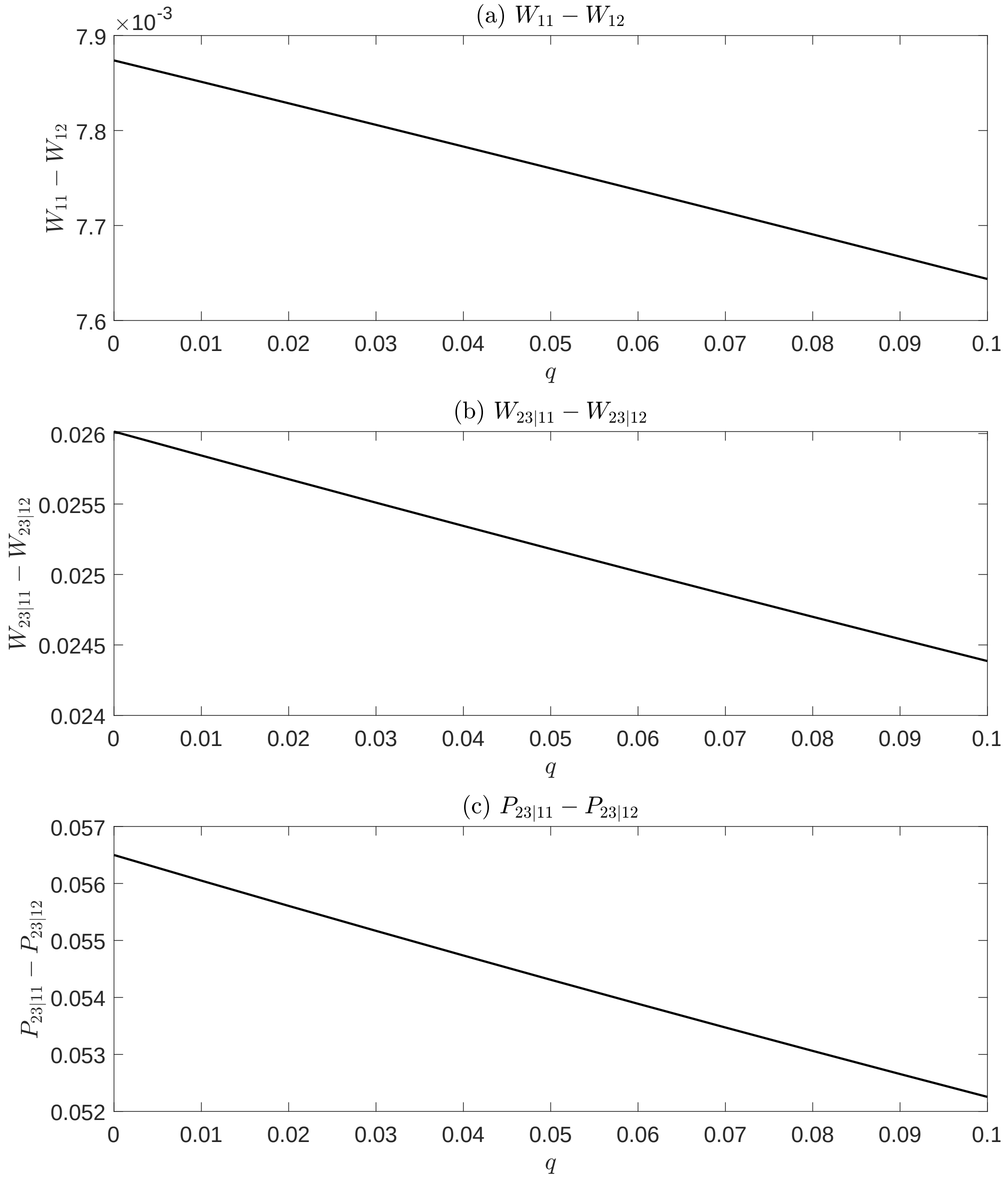
4.1 Fixed Income Arbitrages

We first consider the fixed income market. Fixed-income arbitrage is one of most popular strategies employed by hedge funds. As its name implies, it is an investment strategy that attempts to exploit mispricing which develops among related classes of fixed-income securities. Strictly speaking, it is a statistical arbitrage because mispricing is identified by a statistical analysis rather than by a strict economic reasoning.

In particular, our study on fixed-income arbitrage is closely related to the empirical analysis of Duarte et al. (2006). They document that the fixed-income arbitrage strategies produce significant alphas after controlling for bond and equity market risk factors and many of them produce positively skewed returns. Thus, they conclude that it is not sensible to derogate the fixed-income arbitrageur for 'picking up nickels in front of a steamroller'. However, because they investigate the risk and return characteristics of representative arbitrage trading strategies in the fixed-income sector that they construct, the return of the specific factor may *not* reflect the actual returns of fixed-income arbitrage funds.[11] Besides, their data period ends before the outbreak of the subprime mortgage crisis. Thus, their conclusion might be premature because they did not have a chance to see the genuine steamroller. To see what happened in fixed-income arbitrage funds during the global financial crisis, we investigate a monthly average returns of fixed-income arbitrage index from January 1997 to August 2017, provided by the Barclay Fixed Income Arbitrage Index. Figure 4.1 depicts the time series evolution

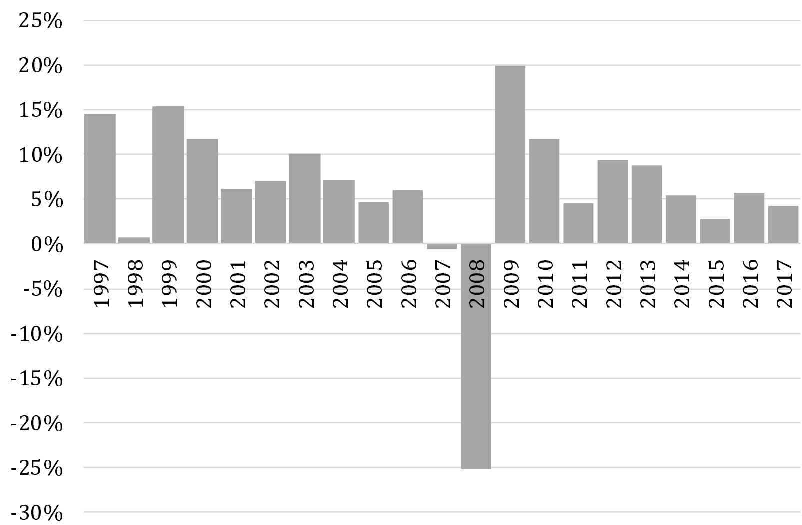
[11] They investigate swap spread arbitrage, yield curve arbitrage (not slope/butterfly spread strategies though), mortgage arbitrage and fixed-income volatility arbitrage.

Figure 3.2: The effect of q on the path-dependency



Notes: This figure illustrates the effect of q on the equilibrium outcomes of $W_{11} - W_{12}$, $W_{23|11} - W_{23|12}$ and $P_{23|11} - P_{23|12}$. Other parameters are $V = 1,\ W_0 = 0.05,\ S_0 = 0.1,\ S_m = 0.2,\ S_b = 0.4,\ \phi = 0.1$ and $\rho = 1\%$.

Figure 4.1: Fixed Income Arbitrage Index Annual Returns



Notes: This figure plots the annual returns of fixed income arbitrage index from 1997 to 2017.

of annual returns of the fixed-income arbitrage funds. Despite the demise of LTCM, the cross-sectional average return in 1998 is still positive, 0.76%. However, the global financial crisis was a much more catastrophic havoc to the industry of fixed-income arbitrage funds. They lost, on average, 0.60% in 2007, for the first time in their history. In the following year, their return nosedived to -25.20%! Before 2007 that Duarte et al.'s data covers, the hedge funds' average monthly return is as high as 66 basis point. Its standard deviation is as small as 87 basis point. Its skewness is -2.08, slightly negative, but not statistically significant. Its excess kurtosis is 9.70.[12] If adding the post-crisis data till August 2017, the above-mentioned figures become quite different. Its mean drops to 49 basis point; its standard deviation jumps up to 140 basis point; the return becomes more negatively skewed, -4.87, and more leptokurtic with excess kurtosis of 42.54 (see column Fixed Income Arbitrage in Table 7). That is, the hidden dark side of the fixed income arbitrage emerged with the outbreak of the global financial crisis!

[12] A number of trading strategies examined in Durate, Longstaff and Yu produce positively skewed returns. In contrast, the actual returns of the fixed income arbitrage funds are negatively skewed as shown in Figure 4.1. This is true even when we use only the data before the global financial crisis. A survivorship bias cannot be an explanation for this discrepancy. Such discrepancy may result from the fact that there are many other strategies not considered by Durarte, Longstaff and Yu and they produce significantly negatively skewed returns. A further study is needed to clarify this.

In implementing fixed income arbitrages, fund managers adopt widely diverse strategies. A representative strategy is to exploit a substantial deviation of a particular spread (such as yield spread, basis between cash and futures, credit spread) from its historical average. To eliminate or minimize its exposure to a fundamental risk, this strategy takes long positions in one asset and short positions in another asset(s). For that reason, the strategy is called 'relative value trading strategy'.[13]

This strategy is generically designed to eliminate its exposure to market risk and credit risk; as such, the expected return on a dollar investment is relatively small so an unusually high degree of leverage is inevitable and often emphasized. In other words, its underlying driver is not to take systematic risk while taking leverage risk: a giant vacuum cleaner sweeping up pennies.

In particular, we employ the U.S. interest rate swap market as a natural candidate for testing the theoretical implications. The interest rate swap market is one of most preferred habitat for hedge funds and proprietary desks of global investment banks. Trading strategies involving interest swaps encompass not only generic yield spreads such as slopes and butterflies of a particular swap curve, but also asset swap margins combined with cash bonds, cross-country basis swaps and basis on futures and their combinations. Given that there are so many different strategies available, most of the trading desks employ a very sophisticated quantitative algorithm called 'trade finder' to detect best opportunities available in a real time basis.

A. Trading Strategies Using the Swap Yield Curve

Herein we investigate the most popularly used yield curve strategies among the fixed-income arbitrage funds: slope and butterfly spreads.

A.1 Slope Spreads

The slope strategies center upon a yield difference between a longer-end and a shorter-end of a yield curve. For example, '2s10s' is a generic industry jargon for a

[13] A Long/short equity strategy is another popular strategy utilized by hedge funds. It also simultaneously takes long and short positions in equity space. However, most of them are long biased (such as 130/30, where long exposure is 130% and short exposure is 30%) and it is composed of a 'long' portfolio by buying equities that are expected to increase in value and a separate 'short' portfolio by shorting equities that are expected to decrease in value. Since it does not eliminate systematic risk completely, it is called a 'fundamental' long/short strategy. There is a pairwise long/short equity strategy, which matches a particular stock that it is long (short) on to another stock with a similar risk profile such as beta. This strategy is not actively adopted by hedge funds though because remaining idiosyncratic risk after controlling the market risk is still sizable and its compensation is relatively small.

spread between the ten-year yield and the two-year yield. There exist many different strategies using a particular slope spread, and the most popular one is the so-called a duration matching strategy. For a short trading horizon, a profit from a fixed leg in the long end swap can be approximated by

$$\Delta V_l \approx -D_l V_l \Delta y_l,$$

where D_l, V_l and y_l are the duration, the value and the yield of the fixed leg, respectively. Thus, a trading position with long on the longer-end and short on the shorter-end with a ratio of $1 : \lambda_s$ is

$$\Delta(V_l - \lambda_s V_s) = \Delta V_l - \lambda_s \Delta V_s \approx -D_l V_l \Delta y_l + \lambda_s D_s V_s \Delta y_s.$$

That is, the trading position's short-term payoff is described as a linear function of the two yields. We want to eliminate the position's exposure to the parallel shift of the yield curve, a directional market risk. Thus we determine λ_s such that

$$-D_l V_l + \lambda_s D_s V_s = 0 \quad \Longrightarrow \quad \lambda_s = \frac{D_l V_l}{D_s V_s} = \frac{D_l}{D_s}.$$

The last equality is based on an assumption that both swaps are par-par swaps. A par-par swap is a generic swap which designates the par notional amount to a fixed leg as well as a floating leg. As such, $V_l = V_s$.[14] Putting this value of λ_s back into the payoff of the position yields

$$\Delta(V_l - \lambda_s V_s) \approx \left[-D_l \Delta y_l + \frac{D_l}{D_s} D_s \Delta y_s\right] V_l = -D_l V_l \Delta(y_l - y_s).$$

Thus, the profit of the aforementioned trading strategy is approximately proportional to the change in the yield spread, $y_l - y_s$. If the yield curve shifts in a parallel manner (i.e., $\Delta(y_l - y_s) = 0$), the arbitrage earns zero. If the curve steepens (i.e., $\Delta(y_l - y_s) > 0$), the strategy loses and vice versa. In the industry, $y_l - y_s$ is called 'pick-up', and the fund managers meticulously monitor its change on a real time basis. Typically, they compute the z-value of the spread based on the past six-month or one-year history and when the z-value is greater or less than a particular threshold level, they consider entering into a position. For example, if the six-month z value is greater than 2.0 (less than -2.0), they believe that the curve is abnormally steep (flat) so the fund managers take a position in a flattener (steepener): i.e., receive (pay) the longer-end and pay (receive)

[14] There is non par-par swap, which is tailored to a specific need of a client; the notional value of each leg is typically not a par value due to a stub period.

the shorter-end. Because the strategy's market exposure (duration risk) is eliminated, its expected profit is quite small and therefore the fund has to use a very high degree of leverage, which could sometimes be as high as ten or twenty times.

There are many variations of the above-mentioned trading strategy. Most of them are based on a belief that the directional move of the curve does not accompany a parallel shift but a certain change in a shape of the curve. For example, when the market interest rate rises in the absence of a monetary policy, the slope upto its belly part tends to steepen while the long-end slope flattens (short-end bear steepening and long-end bear flattening). When the market interest rate falls, the opposite is more likely to occur: short-end bull flattening and long-end bull steepening. Thus, to capture such a statistical relationship between the direction of the yield curve and the corresponding change in the slopes, the quant managers use regression, Principal Component Analysis (PCA) or Independent Component Analysis (ICA) to adjust the ratio of the long-end and the short-end. However, the most popular and representative type of a slope trading strategy is the above-mentioned one.

A.2 Butterfly Spreads

A butterfly strategy involves three tenors as opposed to two tenors in slope strategies. For example, '5s10s20s' refers to a spread between the ten-year yield (middle leg) and the average of the five-year yield (short leg) and the twenty-year yield (long leg). Thus it measures the curvature of the swap yield curve. If the curve is expected to be more concave, the trader 'pays' the butterfly (short on the middle leg and long on the combination of the short and the log legs) and vice versa. Thus the butterfly is associated with the third factor of the PCA, e.g., the curvature factor whereas the slope is related with the second factor.[15]

Below, we investigate the most widely used strategy, a double duration matched butterfly. The value change of the butterfly can be again approximated as

$$-D_m V_m \Delta y_m + \lambda_s D_s V_s \Delta y_s + \lambda_l D_l V_l \Delta y_l.$$

We determine λ_s and λ_l such that (i) the duration of the butterfly is zero and (ii) the duration of the short leg is identical to the duration of the long leg:

$$\begin{aligned} -D_m V_m + \lambda_s D_s V_s + \lambda_l D_l V_l &= 0 \\ \lambda_s D_s V_s - \lambda_l D_l V_l &= 0, \end{aligned}$$

[15] A simple 'pay' or 'receive' of a particular tenor, which is equivalent to shorting or longing a cash bond is a directional bet on the yield and thus is related to the first factor of PCA.

which yields

$$\lambda_s = \frac{D_m V_m}{2 D_s V_s} = \frac{D_m}{2 D_s} \text{ and } \lambda_l = \frac{D_m V_m}{2 D_l V_l} = \frac{D_m}{2 D_l}.$$

Again we assume that the swaps are the par-par swaps in the last equalities in the above two equations. Then the payoff can be described as

$$-D_m V_m \Delta y_m + \lambda_s D_s V_s \Delta y_s + \lambda_l D_l V_\Delta y_l = -D_m V_m \Delta \left[y_m - \frac{1}{2}(y_s + y_l) \right].$$

Thus the payoff of the butterfly is proportional to the difference between the yield of the middle leg and the average yield of the short leg and the long leg. If the trader 'receives' the (middle leg of the) butterfly, she gains if the curve becomes less curvatured and loses if the opposite happens. Similar to the slope spreads, there are many other variations which utilize the statistical association among the three legs such as a regression analysis, a PCA and an ICA.

One thing to mention is the funding cost associated with spread strategies. A swap contract does not accompany any exchange of notional values because the notional value of a fixed leg is identical to that of a floating leg. And its funding cost is already embedded within its contract; if you receive the fixed leg, its funding rate is the six-month LIBOR rate, a coupon rate of the floating leg. If you receive the floating leg, its funding rate is the swap rate. As such, even if the size of swap position (notional value) increases, the funding rate itself does not increase. The same argument can be made about a spread position, a combination of swaps with different tenors. However this logic does not reflect the market practice of marking-to-market (MTM) and collateralization. Under the MTM practice, counterparties are required to post collateral in the amount of the mark-to-market value of the contract.[16] When the mark-to-market value of one party in a swap contract is negative, she needs to pay collateral to her counterparty in the amount of loss. In that sense, it is similar to a futures contract as opposed to a forward contract.[17] Collateral is costly to post, so it induces economic costs to the collateral payer. If she continues to lose in the mark-to-market value of her position, she needs to post additional collateral and this cost rises concomitantly. If she fails to post it, she becomes bankrupt. Most of the collateral posted is in the form of cash or Treasury securities. To pay the collateral, she may use her own cash or Treasuries; otherwise she needs to finance it. Since she loses more, she needs to finance it more and hence the financing cost may rise as well.[18] Therefore, *a priori*, a swap position or

[16]See Johannes and Sundaresan (2007) for this market convention and its impact on the swap rates.
[17]See Johannes and Sundaren (2007) for details.
[18]Or her counterparty, typicall a dealer, applies higher haircuts to Treasuries collateralized.

its combinations (including spreads) entails an implicit funding cost; this cost tends to increases with the size of potential loss in the position. In turn, the size of potential loss is proportional to the size of position and also the riskiness of the position. Overall, this feature is in line with Assumption 1 in our model that the funding rate is proportional to the leverage ratio.

B. Data

We use the U.S. swap yield data from July 23, 1998 to May 11, 2017 from the Bloomberg. We eliminate weekends and holidays. The number of daily observations is 4904. The corresponding number of weekly observations (Wednesday) is 980.[19] The tenors are from one year to ten year along with fifteen year, twenty year and thirty year and thus the total number of tenors is thirteen. Consequently, we obtain $78\left(=\frac{13\times 12}{2}\right)$ slope spreads and $286\left(=\frac{13\times 12\times 11}{3\times 2}\right)$ butterfly spreads, yielding total 364 (=78+286) spreads.

C. Efficiency Measure

We explain how to construct a measure of efficiency for each of 364 spreads. First, for comparison across spreads, we normalize the time-series of each spread to their corresponding z values and denote a normalized spread i at time t by $z_{i,t}$. By doing so, we can equalize the scale of potential collateral. Therefore, we can directly compare their tail risk and tail ranges cross-sectionally. Then, we estimate each spread's mean-reversion speed, δ_i by

$$\Delta z_{i,t+1} = a_i - \delta_i z_{i,t} + \epsilon_{i,t+1}. \tag{4.1}$$

Recall that we are interested in market resilience, or the speed at which prices revert to their equilibrium level after a shock hits transaction flow. Because $z_{i,t}$ is normalized, a_i is close to zero by construction and we can interpret δ_i as the speed at which a spread i goes back to the normalized steady state level of zero. Hence, we interpret δ_i as a measure of market efficiency in terms of resilience.

Also, note that mean-reversion speed δ_i is directly related to the probability of positive arbitrage profits from arbitrage positions, $1 - q$ in our theory. Consider a arbitrageur who enters into a strategy when $z_{i,t} < -\underline{z}$ for a given threshold $\underline{z}$, by betting on that the spread will revert back and exit the position after h holding periods. The

[19] We use other weedays to check the robustness of our results. Qualitatively speaking, the results are the same.

arbitrageur will make profits if $z_{i,t+h} > z_{i,t}$ and experience loss otherwise. Hence, the probability of arbitrage profits can be formulated by $\Pr\left(z_{i,t+h} > z_{i,t} | z_{i,t} < -\underline{z}\right)$, which can be computed by simulation.[20] Figure 4.2 plots how the probability of arbitrage profits moves with the mean-reversion speed, δ_i. We find that for various thresholds $\underline{z}$ and holding periods h, the probability of gain monotonically increases with the mean-reversion speed.[21]

To address the potential concerns on our efficiency measure, we also measure mean-reversion speed δ_i using two alternative methods: (i) we apply HP filter to the time-series of spreads to control the effect of time-varying trends in spreads before estimating δ_i with (4.1) and (ii) we estimate δ_i in (4.1) using the winsorized spread to find the market efficiency only during normal periods. As will be discussed later, our results are robust to these variations.

Before we proceed to the empirical findings, we make it clear that the specification of (4.1) itself cannot generate any spurious relation between the mean reversion speed δ_i and the tail properties of spread $z_{i,t}$. We verify this as follows. Pick a specific δ_i. Then, we simulate daily $z_{i,t}$ with the same size of 4904 daily data using (4.1).[22] Using the simulated 4904 daily data, we compute three tail properties of $z_{i,t}$: (a) kurtosis, (b) VaR 1% left-tail and (c) expected shortfall of 1% left-tail. We repeat this exercise 10000 times. Figure 4.3 plots the average (straight line) along with 10-90% intervals (dashed lines) of the three tail properties across 10000 repetitions. The figure clearly shows that mean reversion speed does not affect mechanically any of the considered statistics.[23] Hence, if we find any empirical relations between mean-reversion speed and the tail behaviors, those are not due to mechanical relations but require some economic reasoning, which we will provide in the subsequent analysis.

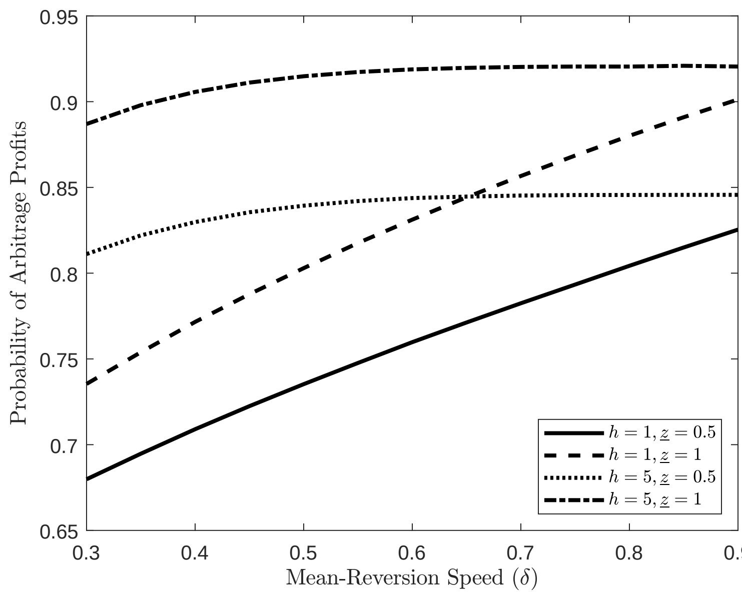
[20] We simulate a time-series $z_{i,t}$ as follows. First, draw $z_{i,1} \sim \mathcal{N}(0,1)$. Then, construct $z_{i,t}$ for $t > 1$ from $\Delta z_{i,t+1} = -\delta_i z_{i,t} + \epsilon_{i,t+1}$, where $\epsilon_{i,t+1} \sim \mathcal{N}\left(0, 1-(1-\delta_i)^2\right)$, so that the long-run variance of $z_{i,t}$ is 1.

[21] Alternatively, we can interpret the relation between mean-reversion speed and market efficiency in terms of conditional expectation. Note that the equation (4.1) can be rewritten as $\Delta z_{i,t+1} = -\delta_i\left(z_{i,t} - \frac{a_i}{\delta_i}\right) + \epsilon_{i,t+1}$. Because of the normalization of z_i, we can treat $\frac{a_i}{\delta_i}$ to be close to zero. Hence, the conditional mean of $E\left[z_{i,t+1}|z_{i,t}\right]$ is approximately $E\left[(1-\delta_i) z_{i,t} + \epsilon_{i,t+1}|z_{i,t}\right] = (1-\delta_i) z_{i,t}$. It follows that the higher mean reversion of $\delta_i \in (0,1)$ pushes the conditional mean of $E\left[z_{i,t+1}|z_{i,t}\right]$ towards zero.

[22] See footnote 20.

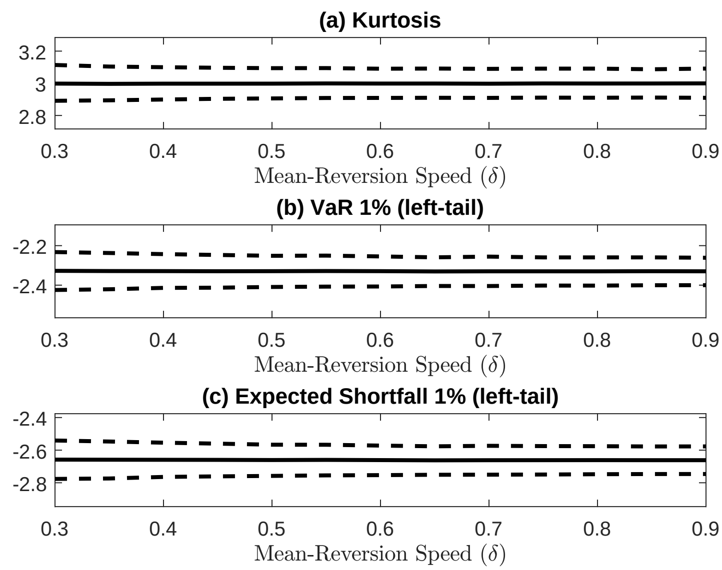
[23] We find similar results when $\epsilon_{i,t+1}$ is drawn from a student-t distribution with degree of freedom of 5.

Figure 4.2: Mean Reversion Speed and Probability of Arbitrage Profits



Notes: This figure illustrates how the probability of arbitrage profits moves with the mean-reversion speed, δ_i. We consider a trading strategy that you enter into a trading strategy when $z_{i,t} < -\underline{z}$ by betting on that the spread will revert back and you exit the position after h holding periods. The probability of arbitrage profits, $\Pr\left(z_{t+h} > z_{i,t} | z_{i,t} < -\underline{z}\right)$, is computed by simulation as follows. First, draw $z_{i,1} \sim \mathcal{N}\left(0,1\right)$. Then, construct $z_{i,t}$ for $t = 2, \cdots, T$ from $\Delta z_{i,t+1} = -\delta_i z_{i,t} + \epsilon_{i,t+1}$, where $\epsilon_{i,t+1} \sim \mathcal{N}\left(0, 1-(1-\delta_i)^2\right)$. We use $T = 10^8$ to compute $\Pr\left(z_{t+h} > z_{i,t} | z_{i,t} < -\underline{z}\right)$.

Figure 4.3: Mean Reversion Speed and Tail Properties



Notes: This figure illustrates that the mean-reversion speed δ_i in the specification of (4.1) does not mechanically affect following tail properties: (a) kurtosis, (b) VaR 1% left-tail and (c) expected shortfall of 1% left-tail. For a given δ_i, we simulate 4904 daily $z_{i,t}$ using (4.1) as follows. First, draw $z_{i,1}$ from a zero-mean and unit-variance distribution. Then, construct $z_{i,t}$ for $t > 1$ from $\Delta z_{i,t+1} = -\delta_i z_{i,t} + \epsilon_{i,t+1}$, where $\epsilon_{i,t+1}$ is drawn from a zero-mean distribution of $1 - (1 - \delta_i)^2$ variance. Using the simulated 4904 daily $z_{i,t}$, we compute the three tail properties of $z_{i,t}$. We repeat this exercise 10000 times. We plot the average (straight line) along with 10-90% intervals (dashed lines) across 10000 repetitions for each of the three tail properties.

D. Hypothesis

We establish the testable hypotheses. The first set of hypotheses focuses on the cross-sectional relation between mean reversion speed and tail properties.

(H1): The spreads with stronger mean reversion are subject to higher tail risk.

(H2): The spreads with stronger mean reversion are subject to higher risk conditional on a tail event.

In the model, (H1) is grounded on Theorems 3.2 and 3.3, showing that $P_{23|11}$ and $P_{23|12}$ decrease with a smaller q. The second hypothesis (H2) is built upon the fact that the discrepancy between $P_{23|11}$ and $P_{23|12}$ increases with a smaller q as stated in Theorem 3.4. Because $P_{23|\cdot}$ is the price in the crash state at $t = 2$, we interpret every realization of $P_{23|\cdot}$ as a tail event. Thus, $P_{23|11} - P_{23|12}$ is viewed as the variation or risk conditional on a tail event mentioned in (H2).

The second set of hypotheses involves time-series implications.

(H3): After experiencing a loss, a spread is subject to higher tail risk.

(H4): After experiencing a loss, the relation in (H1) is stronger.

Note that (H3) is the direct implication of Theorem 3.1, stating that the price in the crash state $\left(S_{23|\cdot}\right)$ is lower following a down market (S_{12}). The fourth hypothesis (H4) comes from Theorem 3.4. Hypothesis (H1) states the effect of the mean reversion on the tail risk regardless of the previous gain/loss. These effects are captured by $\frac{d}{dq}P_{23|11}$ and $\frac{d}{dq}P_{23|12}$, depending on the path. Theorem 3.4 states that the latter (the one after experiencing a loss) is dominant, which leads to (H4).

E. Tests

E.1 Cross-sectional Test

We examine (H1) and (H2) with cross-sectional test. Both hypotheses require the operational definition of 'tail' states (corresponding to S_{23} in our model). We consider $p = 0.5\%$, 1.0% or 1.5% as threshold levels. Even though we theoretically investigate only when the asset is undervalued due to negative noise shocks, a symmetric result holds when the asset is overvalued due to positive noise shocks, as long as the funding cost of shorting increases with the size of short position. Hence, we investigate both tails of a spread distribution. We denote $z_{i|p}$ and $z_{i|1-p}$ the p-percentiles of a specific normalized spread z_i for the left and right tails, respectively.

In testing (H1), we adopt the following three alternative tail risk measures for each spread:

- Kurtosis $\left(= \frac{\widehat{E}(z_{i,t})^4}{\widehat{\sigma}(z_{i,t})^2}\right)$
- p-percentile Value at Risk (VaR): $z_{i|p}$ for the left tail, $z_{i|1-p}$ for the right tail
- Expected Shortfall: $\widehat{E}(z_{i,t}|z_{i,t} < z_{i|p})$ for the left tail, $\widehat{E}(z_{i,t}|z_{i,t} > z_{i|1-p})$ for the right tail.

VaR and expected shortfall are estimated by the historical distribution. Each measure has its own strength and weakness and we do not discuss them in detail. Simply we apply all these measures together.

For (H2), we measure the risk conditional on tail by (i) the tail range and (ii) the tail volatility risk in the following manner:

- Tail Range: $z_{i|p} - \min_t \{z_{i,t}\}$ for the left tail and $\max_t \{z_{i,t}\} - z_{i|(1-p)}$ for right left tail.
- Tail Volatility: standard deviation conditional upon the occurrence of a spread beyond the threshold value, $\widehat{\sigma}\left(z_{i,t}|z_{i,t} < z_{i|p}\right)$ and for the left tail and $\widehat{\sigma}\left(z_{i,t}|z_{i,t} > z_{i|(1-p)}\right)$ for the right tail.

Besides, we highlight that the effects of mean reversion speed on the tail risk are not due to the effects on the overall risk but on tail-specific risk. To this end, we use non-tail volatilities as regressands, which are constructed as follows:

- Non-Tail Volatility: $\widehat{\sigma}\left(z_{i,t}|z_{i,t} > z_{i|p}\right)$ for the left tail and $\widehat{\sigma}\left(z_{i,t}|z_{i,t} < z_{i|(1-p)}\right)$ for the right tail.

Now, we are ready to run the following cross-sectional regressions:

$$y_i = \beta_0 + \beta_1 \delta_i + e_i. \tag{4.2}$$

where y_i is a specific tail characteristic of our interest: kurtosis, VaR, expected shortfall, tail volatility, tail range and non-tail volatility of spread i.[24] The estimated β_1 from

[24]We note that this procedure suffers from a classic errors-in variables (EIV) problem because we use the estimate of mean-reversion speed as an independent variable and the estimates of distributional characteristics (kurtosis, Value-at-Risk, expected short-fall risk and etc.) as dependent variables. Therefore, both regressors and regressands are subject to estimation errors. As is well known, though, the measurement errors in regressands do not cause a problem as long as they are uncorrelated with the regressors and their measurement errors. We apply the methods by Shanken (1992) and Dagenais and Dagenais (1997) to address the errors-in-variable problem in the regressors and find consistent results. The results using EIV correction are available upon request.

(4.2) tests (H1) when we use kurtosis, VaR, expected shortfall for y_i and (H2) when we the tail range and non-tail volatility for y_i.

Table 1 report the estimation results from using daily data (left half) and weekly data (right half) for three different thresholds $p = 0.5\%$, 1.0% or 1.5% for the definition of 'tail' states. Because qualitative results are similar across various cases, we focus on discussing the estimation results from daily data with $p = 1.0\%$.

First, we examine (H1), the spreads with stronger mean reversion are subject to higher tail risk. We find that kurtosis is positively associated with mean-reversion speed with statistical significance and R^2 is as high as 57.0%. Besides, the mean-reversion speeds demonstrate a negative relationship with left VaRs, -0.70 (t-stat=-2.00), and a positive relationship with right VaRs, 0.58 (t-stat=1.97). That is, the spreads with stronger mean reversion is more likely to have fat tail risk on both sides. We find similar findings, albeit stronger when employing expected shortfall as an alternative measure of tail risk. For both daily and weekly data, the results tend to become stronger with smaller p. Although the results for VaR becomes weaker with $p = 1.5\%$, overall we conclude that hypothesis (H1) is strongly supported.

Next, we move to (H2). The range of left and right tails, measured by left VaR minus minimum and maximum minus right VaR, respectively, are strongly related to the mean-reversion speed. The relations are statistically very strong. Furthermore, we find consistent results for both tail volatilities, left tail and right tail. Interestingly, the association of mean-version speed to non-tail volatilities is much smaller in terms of magnitude and often statistically insignificant. The results are robust to different threshold levels and consistent when we use weekly data. So we empirically establish that hypothesis (H2) is also well supported.

Table 2 confirms additional robustness of the results in Table 1. We split the whole 364 spreads into two subsamples of 78 slope spreads and 286 butterfly spreads and repeat the exercise for each subsample. The 'Slope' ('Butterfly') column reports the estimation results for 78 slope (286 butterfly) spreads. Besides, we check the robustness of results to alternative methods for estimating mean-reversion speed. In particular, we control the effect of time-varying trends in spreads by applying the filter by Hodrick and Prescott (1997) before estimating δ_i and report the results using the HP-filtered mean-reversion speed in 'HP filter' column.[25] Furthermore, we also estimate δ_i using the winsorized spread[26] to find the mean-reversion speed only during tranquil periods.

[25]For the smoothing parameter of Hodrick and Prescott's filter, we follow the method by Ravn and Uhlig (2002).

[26]For winsorizing, we use the thresholds of tail events. We drop the left and right p-percentage tails

The mean-reversion speed estimated in this manner is more consistent with our theory model, which relates the efficiency during normal state to the vulnerability in the crash state. We report the estimation results using the (non-tail period) mean-reversion speed in the 'Non-tail δ' column. Across all these variations, we find very consistent results, strengthening the validity of (H1) and (H2).

In summary, we conclude that the two hypotheses, (H1) and (H2), are well in line with the behavior of the U.S. swap curve. Nevertheless, some readers may be concerned about the cross-sectional correlation of residual terms in the specification of (4.2) and have doubts about the true statistical significance of the estimated coefficient β_1. In what follows, we provide quantile regression tests which not only address the concern on the cross-sectionally correlated shock but provide further dynamic implications.

E.2 Quantile regression with panel data

We establish (H1), (H3) and (H4) with the quantile regression. In particular, we investigate those hypotheses by estimating the quantile regression for the left quantile:

$$Q_p\left(z_{i,t}|\delta_i, l_{i,t-1,h}\right) = \gamma_0 + \gamma_1\delta_i + \gamma_2 l_{i,t-1,h} + \gamma_3\delta_i l_{i,t-1,h}, \tag{4.3}$$

where $Q_p\left(z_{i,t}|\delta_i, l_{i,t-1,h}\right)$ is the conditional p-percentile of $z_{i,t}$ given the information of mean-reversion speed of spread i, δ_i and the loss indicator of spread i at time t, $l_{i,t-1,h}$. Note that the LHS in (4.3) is replaced by $Q_{1-p}\left(z_{i,t}|\delta_i, l_{i,t-1,h}\right)$ for the right tail. For the left tail, the loss indicator of $l_{i,t-1,h}$ is computed as follows. Over the periods from $t-1-h$ to $t-1$, we consider a simple holding strategy that if $z_{i,t-1-h} < -\underline{z}$, long the spread i from $t-1-h$ and clear it at $t-1$. We define $l_{i,t-1,h}$ as an indicator function which is 1 if the strategy generates a loss or 0 otherwise.[27] In a similar manner, for the case of right tail, we consider a strategy that if $z_{i,t-1-h} > \underline{z}$, short the spread i from $t-1-h$ and reverse the position at $t-1$, and define the indicator $l_{i,t-1,h}$ accordingly.[28]

We relate the coefficients of γ_1, γ_2 and γ_3 in (4.3) to hypotheses (H1), (H3) and (H4), respectively. Note that when $\gamma_2 = \gamma_3 = 0$, γ_1 in (4.3) captures the relation between the mean-reversion speed δ_i and the tail-quantile. Thus, it directly tests (H1) – the spreads with stronger mean reversion are subject to higher tail risk. Next, when $\gamma_1 = \gamma_3 = 0$, γ_2 in (4.3) measures the effect of past loss on the tail-quantile, and hence addresses (H3) – after experiencing loss, a spread is subject to higher tail risk. Lastly, γ_3 in (4.3) picks up the interaction between the past loss and the mean-version speed. Recalling

in the empirical spread data.

[27] For the left tail, $l_{i,t-1,h} = \mathbf{1}\left(z_{i,t-1} < z_{i,t-1-h} < -\underline{z}\right)$.

[28] For the right tail, $l_{i,t-1,h} = \mathbf{1}\left(z_{i,t-1} > z_{i,t-1-h} > \underline{z}\right)$.

(H4) says that after experiencing loss, the relation between the mean-reversion speed and tail risk becomes stronger, we argue that γ_3 handles (H4).

Table 3 reports the estimation results of (4.3) using the panel data of 364 spreads for 4904 daily observations over the sample period of July 23, 1998 to May 11, 2017. For the columns (1) and (4), we set $\gamma_2 = \gamma_3 = 0$. And for the columns (2) and (5), we set $\gamma_1 = \gamma_3 = 0$. In Panel A, we consider the baseline case $p = 1\%$, $\underline{z} = 1$ and $h = 1$. We find that the coefficient on δ_i is significantly negative (positive) for the left (right) tail, which confirms (H1). Also, we find that the effect of past loss is negative (positive) for the left (right) tail, advocating (H2). Lastly, the coefficient on the interaction of mean-reversion speed is significantly negative (positive) for the left (right) tail, supporting (H3). For Pseudo R^2, we use the method by Koenker and Machado (1999). The t-statistics are computed by bootstrapping standard error. Our bootstrapping design is crucially important to properly address the potential cross-sectional as well as time-series correlation of error terms. To fully address the concern on the cross-sectional correlation, we resample the cross-section of 364 spreads together. In addition, we use time-block sampling with the random time-series block size between 11 and 20 to handle the time-series correlation. The bootstrapping standard error is computed by 200 sets of bootstrap samples.

We check the robustness of results in Panel A of Table 3 in various dimensions. Panels B and C of Table 3 confirm that the results are consistent whether we use five or ten days to compute the loss indicator from the arbitrage strategy. In Table 4, we repeat the exercise for two subsamples of 78 slope spreads (Panel A) and 286 butterfly spreads (Panel B) and find that the results are persistent. We consider alternative tail definitions of $p = 0.5\%$ and $p = 1.5\%$ in Panel A and Panel B, respectively, of Table 5. Lastly, for Table 6, we apply the filter by Hodrick and Prescott (1997)[29] (Panel A) or drop tails in spread data (Panel B) to estimate δ_i, and we use weekly data (Panel C). Several robustness checks support our claim that the hypotheses of (H1), (H3) and (H4) describe the tail properties in U.S. interest swap market accurately.

4.2 Hedge Fund Strategies

Next, we consider various hedge fund strategy indexes. The hedge fund industry provides a plausible environment to examine the implications of our model in that a hedge fund focuses on the mispricing of alternative investment opportunities and uses high

[29]For the smoothing parameter of Hodrick and Prescott's filter, we follow the method by Ravn and Uhlig (2002).

risk methods such as leverage in searching for large gains. Moreover, the hedge fund industry keeps developing novel strategies to find mispricing opportunities. Hence, there exist a numerous trading strategies and the realized returns reveal distinctive characteristics across the strategies.

Table 7 reports the summary statistics of monthly percentage returns from Barclay Hedge Fund Index for one aggregate index and other sixteen style indexes. The monthly returns cover the sample period of Jan, 1997 to Aug, 2017. The reported summary statistics are Mean, Median, Min (minimum), Max (maximum), Std.Dev. (standard deviation), Kurtosis, Skewness and Prob(+). Prob(+) is the frequency of positive monthly returns. For example, for the case of CB Arbitrage, the reported number 0.754 represents that the monthly returns of CB Arbitrage index were positive in 187 months out of the total sample size of 284 months (Jan, 1997 - Aug, 2017).

Now, we establish a testable hypothesis. Our theoretical results show that there is an inherent relation between the favorable market conditions for arbitrageurs during normal periods and the exacerbated mispricing during crash periods. Hence, we empirically examine the association between Prob(+) and higher moments such as skewness and kurtosis.

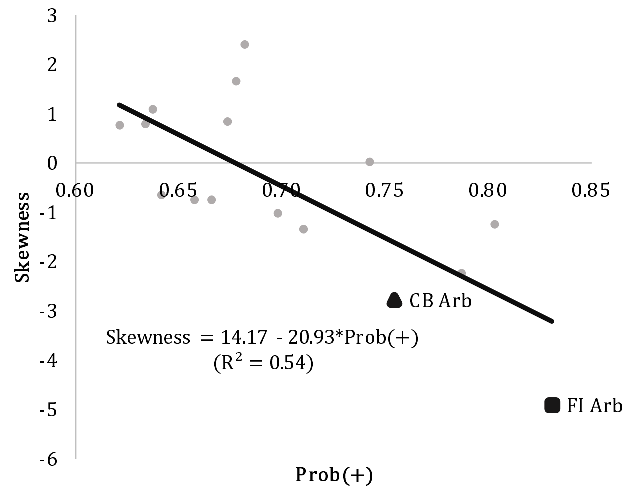
(H5): The hedge fund styles with a higher probability of positive returns are subject to higher levels of skewness and kurtosis.

Note that the above hypothesis is closely related to the commonly used phrases of "picking up nickels in front of steamrollers" or "the stairs up and the elevator down", describing the return characteristics of some trading strategies. Furthermore, Duarte et al. examine a similar question with fixed-income arbitrage. We complement their study by covering post 2007-2008 data as well as considering other various hedge fund strategies.

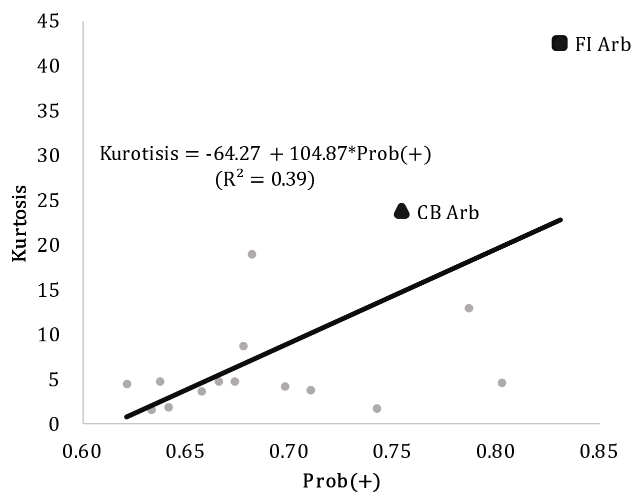
From Table 7, we find that CB Arbitrage and Fixed Income Arbitrage are the top two styles in Kurtosis (23.752 and 42.537, respectively) as well as the bottom two in Skewness (-2.766 and -4.899, respectively). In contrast, they historically experienced positively returns frequently. The Prob(+) of CB Arbitrage is 0.754, fourth in rank, and that of Fixed Income Arbitrage is 0.831, first in rank. These two examples of CB Arbitrage and Fixed Income Arbitrage already manifest the validity of (H5).

We test (H5) more formally by regressing skewness or kurtosis on Prob(+). Figure 4.4 shows the results. We find that a hedge fund style with higher Prob(+) experience more negatively skewed distribution (Figure 4.4-(a)) as well as fatter tailed distribution (Figure 4.4-(b)) over the sample period of Jan, 1997 - Aug, 2017. We conclude that

Figure 4.4: The Relation of Skewness and Kurtosis with Prob(+)



(a) Skewness and Prob(+)



(b) Kurtosis and Prob(+)

Notes: This figure plots the results of regressing skewness and kurtosis on the realized frequency of positive monthly returns over the sample period of 1997:01 to 2017:08 in (a) and (b), respectively. "CB Arb" and "FI Arb" represent convertible bond arbitrage and fixed income arbitrage, respectively.

(H5) depicts the cross-sectional differences in return characteristics over various hedge fund styles.

5 Conclusion

This paper delivers a novel insight on limited arbitrage on top of existing literature by centering upon what kind of market is more likely to attract arbitrage transactions and demonstrating theoretically and empirically that such a market is more susceptible to a crash.

Hedge funds specializing in fixed income arbitrage are extremely similar in their key strategies. They seize a trade opportunity when the gap between the market price of a security and its fair value widens above a pre-specified level. They unwind their positions either when the spread contracts to a certain level (profit realization). Therefore, their entry into and exit from trades are very similar albeit their exact profit realization levels and loss cut levels being slightly different. Simply put, their investment strategies are uni-directional and suffer from the lack of diversity. As a result, regardless of the number of hedge funds, they act as a huddled mass.[30]

Under a tranquil market condition, such synchronized collective actions among hedge funds have the benefits of polishing the market more effectively by eliminating mispricings quickly and sufficiently. However, when the market is embroiled in turmoil, i.e., when it is time that the arbitrageur's demand is most needed, the arbitrage mechanism itself malfunctions and fails to correct dislocations in prices. In our model, an arbitrageur is ensured to survive until the market price converges to its fair value. In addition, we do not introduce a loss-cut practice that is widely employed by hedge funds.[31] As such, in our model, the arbitrageur exits the market only if she earns gains and does not expect any further profit opportunity. If we allow other reasons including the aforementioned ones the arbitrageur leaves the market (so when the gap widens rather than shrinks), the model may amplify the mispricing; for example, $P_{23|12}$ could be even lower than $V - S_b$. In the worst case, the market collapses and fails to be resurrected as evidenced by Japanese floating rate bond market.[32] We reserve this kind of extension for future research.

[30]Thus, our assumption of uniform arbitrageurs is not entirely preposterous.

[31]See Ahn et al. (2017) for a potential equilibrium of disequilibrium in the presence of such a practice.

[32]See Chung et al. (2018) for how the Japanese floater market collapsed in the aftermath of the global financial crisis.

A Proofs

Reference

Ahn, Dong-Hyun, Soohun Kim, and Kyoungwon Seo, 2017, Arbitrage comes hand in hand with the risk of market crash, *Unpublished Manuscript, Seoul National University* .

Brunnermeier, Markus K, Stefan Nagel, and Lasse H Pedersen, 2008, Carry trades and currency crashes, *NBER macroeconomics annual* 23, 313–348.

Brunnermeier, Markus K, and Lasse Heje Pedersen, 2008, Market liquidity and funding liquidity, *The Review of Financial Studies* 22, 2201–2238.

Chung, Ji-Yeong, Dong-Hyun Ahn, In-Seok Baek, and Kyu Ho Kang, 2018, An empirical investigation on funding liquidity and market liquidity, *Review of Finance* 22, 1249.

Dagenais, Marcel G, and Denyse L Dagenais, 1997, Higher moment estimators for linear regression models with errors in the variables, *Journal of Econometrics* 76, 193–221.

De Long, J Bradford, Andrei Shleifer, Lawrence H Summers, and Robert J Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703–738.

Duarte, Jefferson, Francis A Longstaff, and Fan Yu, 2006, Risk and return in fixed-income arbitrage: Nickels in front of a steamroller?, *The Review of Financial Studies* 20, 769–811.

Gromb, Denis, and Dimitri Vayanos, 2002, Equilibrium and welfare in markets with financially constrained arbitrageurs, *Journal of Financial Economics* 66, 361–407.

Grossman, Sanford J, and Joseph E Stiglitz, 1980, On the impossibility of informationally efficient markets, *The American Economic Review* 70, 393–408.

Hodrick, Robert J, and Edward C Prescott, 1997, Postwar us business cycles: an empirical investigation, *Journal of Money, credit, and Banking* 1–16.

Jermann, Urban J., 2017, Negative swap spreads and limited arbitrage, *Unpublished Manuscript, University of Pennsylvania* .

Kahraman, Bige, and Heather E Tookes, 2017, Trader leverage and liquidity, *The Journal of Finance* 72, 1567–1610.

Kapadia, Nikunj, and Xiaoling Pu, 2012, Limited arbitrage between equity and credit markets, *Journal of Financial Economics* 105, 542–564.

Koenker, Roger, and Jose AF Machado, 1999, Goodness of fit and related inference processes for quantile regression, *Journal of the American Statistical Association* 94, 1296–1310.

Kyle, Albert S, 1985, Continuous auctions and insider trading, *Econometrica* 53, 1315–1335.

Lucas, Robert E, 1978, Asset prices in an exchange economy, *Econometrica* 46, 1429–1445.

Mancini Griffoli, Tommaso, and Angelo Ranaldo, 2012, Limits to arbitrage during the crisis: Finding liquidity constraints and covered interest parity, *Unpublished Manuscript, University of St. Gallen* .

Mitchell, Mark, and Todd Pulvino, 2001, Characteristics of risk and return in risk arbitrage, *The Journal of Finance* 56, 2135–2175.

Mitchell, Mark, Todd Pulvino, and Erik Stafford, 2002, Limited arbitrage in equity markets, *The Journal of Finance* 57, 551–584.

Priaulet, Phillipe, 2008, Strategies de courbe à base de swaptions & straddles et stragles de swaptions, *Unpublished Manuscript, Université d'Evry* .

Ravn, Morten O, and Harald Uhlig, 2002, On adjusting the hodrick-prescott filter for the frequency of observations, *Review of Economics and Statistics* 84, 371–376.

Ross, Stephen A, 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.

Shanken, Jay, 1992, On the estimation of beta-pricing models, *The Review of Financial Studies* 5, 1–33.

Sharpe, William F, 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *The Journal of Finance* 19, 425–442.

Shleifer, Andrei, and Robert W Vishny, 1997, The limits of arbitrage, *The Journal of Finance* 52, 35–55.

Tata, Fidelio, 2006, A derivative strategies bag of tricks: Relative value & trade ideas in the u.s. derivatives markets, *Unpublished Manuscript, RBS Greenwich Capital* .

Table 1: Cross-sectional Regression Estimates (t-stats)

Associated Hypothesis	Dependent Variable (y_i)		Daily $p = 0.5\%$ β_1	Daily $p = 0.5\%$ R^2	Daily $p = 1.0\%$ β_1	Daily $p = 1.0\%$ R^2	Daily $p = 1.5\%$ β_1	Daily $p = 1.5\%$ R^2	Weekly $p = 0.5\%$ β_1	Weekly $p = 0.5\%$ R^2	Weekly $p = 1.0\%$ β_1	Weekly $p = 1.0\%$ R^2	Weekly $p = 1.5\%$ β_1	Weekly $p = 1.5\%$ R^2
(H1)	Kurtosis		393.43	0.57	393.43	0.57	393.43	0.57	359.67	0.68	359.67	0.68	359.67	0.68
			(21.82)		(21.82)		(21.82)		(27.97)		(27.97)		(27.97)	
	VaR	Left	-2.18	0.05	-0.70	0.01	-0.10	0.00	-1.60	0.04	-0.46	0.01	-0.09	0.00
			(-4.38)		(-2.00)		(-0.38)		(-3.98)		(-1.59)		(-0.40)	
		Right	1.97	0.10	0.58	0.01	0.10	0.00	0.98	0.04	0.16	0.00	-0.23	0.00
			(6.29)		(1.97)		(0.37)		(3.81)		(0.68)		(-0.97)	
	Expected Shortfall	Left	-6.62	0.26	-4.05	0.15	-2.72	0.10	-4.82	0.23	-2.63	0.11	-1.90	0.08
			(-11.20)		(-8.08)		(-6.32)		(-10.31)		(-6.69)		(-5.54)	
		Right	6.09	0.47	3.58	0.26	2.57	0.17	4.76	0.45	2.80	0.26	1.87	0.14
			(17.83)		(11.32)		(8.53)		(17.36)		(11.16)		(7.77)	
(H2)	Tail Range	Left	34.76	0.69	36.24	0.69	36.83	0.68	12.62	0.46	13.76	0.49	14.13	0.49
			(28.30)		(28.14)		(27.69)		(17.68)		(18.64)		(18.77)	
		Right	40.52	0.68	41.91	0.71	42.38	0.71	19.95	0.53	20.77	0.56	21.15	0.57
			(27.97)		(29.57)		(29.81)		(20.14)		(21.47)		(21.96)	
	Tail Volatility	Left	8.02	0.68	6.39	0.70	5.42	0.66	5.41	0.46	4.07	0.49	3.65	0.50
			(28.09)		(29.13)		(26.60)		(17.63)		(18.85)		(19.08)	
		Right	8.53	0.75	6.62	0.79	5.76	0.81	8.01	0.53	6.26	0.57	5.33	0.59
			(32.63)		(36.99)		(39.64)		(20.32)		(21.82)		(22.85)	
	Non-Tail Volatility	Left	-0.01	0.00	-0.06	0.00	-0.09	0.01	-0.06	0.00	-0.11	0.01	-0.13	0.02
			(-0.15)		(-1.02)		(-1.57)		(-1.10)		(-2.15)		(-2.83)	
		Right	-0.14	0.01	-0.20	0.02	-0.22	0.02	-0.15	0.01	-0.19	0.03	-0.21	0.03
			(-1.71)		(-2.51)		(-2.97)		(-2.30)		(-3.09)		(-3.52)	

Notes: This table reports the results of cross-sectional regression (4.2) for the 364 spreads (78 slope and 286 butterfly spreads). We consider three threshold levels of $p = 0.5\%$, 1.0% and 1.5%. The choice of p is not relevant for Kurtosis. We use daily (left half) as well as weekly (right half) U.S. interest rate swap market data from the Bloomberg over the sample period of July 23, 1998 to May 11, 2017. The t-statistics are reported in parentheses.

Table 2: Robustness of Cross-sectional Regression

Associated	Dependent		Daily, $p = 1.0\%$							
			Slope		Butterfly		HP filter		Non-tail δ	
Hypothesis	Variable (y_i)		β_1	R^2	β_1	R^2	β_1	R^2	β_1	R^2
(H1)	Kurtosis		45.12	0.60	397.14	0.57	459.30	0.54	896.02	0.52
			(10.73)		(19.41)		(20.61)		(19.92)	
	VaR	Left	-8.53	0.13	-0.39	0.00	-0.38	0.01	-4.19	0.05
			(-3.35)		(-1.06)		(-1.78)		(-4.37)	
		Right	7.45	0.07	0.50	0.01	0.08	0.00	2.96	0.04
			(2.41)		(1.55)		(0.41)		(3.65)	
	Expected Shortfall	Left	-22.38	0.36	-3.59	0.14	-3.50	0.21	-13.97	0.23
			(-6.62)		(-6.83)		(-9.87)		(-10.40)	
		Right	9.44	0.08	3.51	0.27	2.84	0.32	10.57	0.29
			(2.53)		(10.22)		(12.93)		(12.12)	
(H2)	Tail Range	Left	175.54	0.77	35.53	0.71	32.82	0.67	95.22	0.60
			(16.10)		(26.14)		(27.06)		(23.31)	
		Right	49.63	0.40	41.87	0.71	37.44	0.71	107.36	0.59
			(7.10)		(26.12)		(29.66)		(22.72)	
	Tail Volatility	Left	27.32	0.84	6.25	0.71	5.75	0.69	16.82	0.62
			(20.48)		(26.57)		(28.74)		(24.08)	
		Right	7.50	0.43	6.62	0.79	6.10	0.82	16.66	0.63
			(7.64)		(32.75)		(40.53)		(25.06)	
	Non-Tail Volatility	Left	-0.08	0.00	-0.09	0.01	-0.21	0.15	0.05	0.00
			(-0.19)		(-1.37)		(-7.95)		(0.32)	
		Right	0.83	0.02	-0.21	0.02	-0.23	0.11	-0.26	0.00
			(1.22)		(-2.48)		(-6.80)		(-1.19)	

Notes: This table reports the results of cross-sectional regression (4.2). We consider $p = 1.0\%$. We use 78 slope (286 butterfly) spreads for the column of 'Slope' ('Butterfly'). We apply the filter by Hodrick and Prescott (1997) before estimating δ_i and report results in 'HP filter' column. We also estimate δ_i using the data within the quantile range of $(p, 1-p)$ and report results in the 'Non-tail δ' column. We use daily U.S. interest rate swap market data from the Bloomberg over the sample period of July 23, 1998 to May 11, 2017. The t-statistics are reported in parentheses.

Table 3: Quantile Regression Estimates (t-stat)

Associated Hypothesis	Independent Variable	Left Tail (1)	Left Tail (2)	Left Tail (3)	Right Tail (4)	Right Tail (5)	Right Tail (6)
Panel A: Baseline							
(H1)	δ	-2.61		-1.55	1.54		1.09
		(-16.14)		(-10.62)	(10.39)		(7.28)
(H3)	l		-0.96	-0.59		0.76	0.47
			(-51.13)	(-32.15)		(43.39)	(24.62)
(H4)	$\delta \times l$			-33.44			21.94
				(-64.19)			(36.46)
	Pseudo R^2	0.0034	0.0652	0.0732	0.0016	0.0582	0.0631
Panel B: loss measure periods = 5							
(H1)	δ	-2.61		-1.58	1.54		1.16
		(-16.14)		(-10.43)	(10.39)		(8.39)
(H3)	l		-1.09	-0.65		0.86	0.57
			(-50.45)	(-28.61)		(48.64)	(24.46)
(H4)	$\delta \times l$			-38.09			21.27
				(-86.21)			(33.41)
	Pseudo R^2	0.0034	0.0817	0.0910	0.0016	0.0774	0.0826
Panel C: loss measure periods = 10							
(H1)	δ	-2.61		-1.52	1.54		1.19
		(-16.14)		(-10.39)	(10.39)		(8.37)
(H3)	l		-1.05	-0.72		0.90	0.60
			(-51.64)	(-30.59)		(37.72)	(25.07)
(H4)	$\delta \times l$			-28.23			21.36
				(-41.17)			(32.96)
	Pseudo R^2	0.0034	0.0769	0.0831	0.0016	0.0869	0.0917

Notes: The table reports the estimation results of quantile regressions:

$$Q_p\left(z_{i,t}|\delta_i, l_{i,t-1,h}\right) = \gamma_0 + \gamma_1\delta_i + \gamma_2 l_{i,t-1,h} + \gamma_3\delta_i l_{i,t-1,h} \text{ for the left tail, and}$$
$$Q_{1-p}\left(z_{i,t}|\delta_i, l_{i,t-1,h}\right) = \gamma_0 + \gamma_1\delta_i + \gamma_2 l_{i,t-1,h} + \gamma_3\delta_i l_{i,t-1,h} \text{ for the right tail.}$$

Here, $z_{i,t}$ is the normalized value of spread i at time t, δ_i is the mean-reversion speed and $l_{i,t-1,h}$ is the loss indicator defined as $\mathbf{1}\left(z_{i,t-1} < z_{i,t-1-h} < -\underline{z}\right)$ for the left tail and $\mathbf{1}\left(z_{i,t-1} > z_{i,t-1-h} > \underline{z}\right)$ for the right tail. For Panel A, we set $p = 1\%$, $\underline{z} = 1$ and $h = 1$. Panel B (C) differs from Panel A in that $h = 5$ (10). We use daily U.S. interest rate swap market data from the Bloomberg over the sample period of July 23, 1998 to May 11, 2017. The t-statistics are computed by a bootstrap method.

Table 4: Quantile Regression using Subsamples of Slope and Butterfly Spreads

Associated Hypothesis	Independent Variable	Left Tail (1)	Left Tail (2)	Left Tail (3)	Right Tail (4)	Right Tail (5)	Right Tail (6)
Panel A: Using only 78 slope spreads							
(H1)	δ	-10.67		-9.37	8.94		7.39
		(-13.65)		(-10.49)	(7.59)		(6.36)
(H3)	l		-0.32	-0.25		0.37	0.22
			(-30.01)	(-14.87)		(20.63)	(11.66)
(H4)	$\delta \times l$			-5.73			42.15
				(-2.76)			(15.44)
	Pseudo R^2	0.0045	0.0277	0.0317	0.0052	0.0331	0.0375
Panel B: Using only 286 butterfly spreads							
(H1)	δ	-1.55		-0.99	0.90		0.78
		(-8.98)		(-6.57)	(5.56)		(5.33)
(H3)	l		-1.04	-0.71		0.81	0.55
			(-43.13)	(-26.37)		(35.26)	(21.94)
(H4)	$\delta \times l$			-26.55			17.01
				(-43.94)			(29.19)
	Pseudo R^2	0.0016	0.0691	0.0749	0.0006	0.0599	0.0634

Notes: The table reports the estimation results of quantile regressions:

$$Q_p\left(z_{i,t}|\delta_i, l_{i,t-1,h}\right) = \gamma_0 + \gamma_1\delta_i + \gamma_2 l_{i,t-1,h} + \gamma_3\delta_i l_{i,t-1,h} \text{ for the left tail, and}$$
$$Q_{1-p}\left(z_{i,t}|\delta_i, l_{i,t-1,h}\right) = \gamma_0 + \gamma_1\delta_i + \gamma_2 l_{i,t-1,h} + \gamma_3\delta_i l_{i,t-1,h} \text{ for the right tail.}$$

Here, $z_{i,t}$ is the normalized value of spread i at time t, δ_i is the mean-reversion speed and $l_{i,t-1,h}$ is the loss indicator defined as $\mathbf{1}\left(z_{i,t-1} < z_{i,t-1-h} < -\underline{z}\right)$ for the left tail and $\mathbf{1}\left(z_{i,t-1} > z_{i,t-1-h} > \underline{z}\right)$ for the right tail. We set $p = 1\%$, $\underline{z} = 1$ and $h = 1$. Panel A uses 78 slope spreads and Panel B uses 286 butterfly spreads. We use daily U.S. interest rate swap market data from the Bloomberg over the sample period of July 23, 1998 to May 11, 2017. The t-statistics are computed by a bootstrap method.

Table 5: Quantile Regression using other Quantile Thresholds

Associated Hypothesis	Independent Variable	Left Quantile (1)	(2)	(3)	Right Quantile (4)	(5)	(6)
		Panel A: $p = 0.5\%$					
(H1)	δ	-5.09		-3.84	4.06		2.89
		(-21.10)		(-16.25)	(20.96)		(14.66)
(H3)	l		-1.14	-0.57		0.76	0.52
			(-33.40)	(-14.00)		(30.24)	(22.95)
(H4)	$\delta \times l$			-42.77			16.69
				(-45.92)			(20.67)
	Pseudo R^2	0.0086	0.0703	0.0879	0.0062	0.0540	0.0624
		Panel B: $p = 1.5\%$					
(H1)	δ	-1.47		-0.63	0.88		0.41
		(-15.73)		(-6.60)	(7.19)		(3.74)
(H3)	l		-0.86	-0.57		0.70	0.47
			(-79.97)	(-42.83)		(48.52)	(28.17)
(H4)	$\delta \times l$			-26.76			19.76
				(-83.58)			(41.86)
	Pseudo R^2	0.0014	0.0595	0.0639	0.0006	0.0583	0.0618

Notes: The table reports the estimation results of quantile regressions:

$$Q_p\left(z_{i,t}|\delta_i, l_{i,t-1,h}\right) = \gamma_0 + \gamma_1\delta_i + \gamma_2 l_{i,t-1,h} + \gamma_3\delta_i l_{i,t-1,h} \text{ for the left tail, and}$$
$$Q_{1-p}\left(z_{i,t}|\delta_i, l_{i,t-1,h}\right) = \gamma_0 + \gamma_1\delta_i + \gamma_2 l_{i,t-1,h} + \gamma_3\delta_i l_{i,t-1,h} \text{ for the right tail.}$$

Here, $z_{i,t}$ is the normalized value of spread i at time t, δ_i is the mean-reversion speed and $l_{i,t-1,h}$ is the loss indicator defined as $\mathbf{1}\left(z_{i,t-1} < z_{i,t-1-h} < -\underline{z}\right)$ for the left tail and $\mathbf{1}\left(z_{i,t-1} > z_{i,t-1-h} > \underline{z}\right)$ for the right tail. We consider $\underline{z} = 1$ and $h = 1$. For Panel A (B), we set $p = 0.5\%$ (1.5%). We use daily U.S. interest rate swap market data from the Bloomberg over the sample period of July 23, 1998 to May 11, 2017. The t-statistics are computed by a bootstrap method.

Table 6: Other Robustness of Quantile Regression

Associated Hypothesis	Independent Variable	Left Quantile (1)	(2)	(3)	Right Quantile (4)	(5)	(6)
		Panel A: Using Hodrick-Prescott filtered spreads					
(H1)	δ	-0.50		-0.40	0.86		0.64
		(-6.17)		(-4.04)	(10.64)		(6.58)
(H3)	l		-1.06	-0.62		0.71	0.55
			(-56.12)	(-36.54)		(42.16)	(28.83)
(H4)	$\delta \times l$			-23.63			7.17
				(-64.45)			(17.47)
	Pseudo R^2	0.0003	0.0654	0.0717	0.0010	0.0609	0.0634
		Panel B: Using δ estimated from non-tail spreads					
(H1)	δ	-13.49		-7.27	15.54		6.86
		(-31.39)		(-15.02)	(43.03)		(17.23)
(H3)	l		-0.96	-0.33		0.76	0.30
			(-54.12)	(-15.55)		(42.15)	(15.69)
(H4)	$\delta \times l$			-100.94			63.43
				(-75.79)			(38.54)
	Pseudo R^2	0.0113	0.0652	0.0830	0.0107	0.0582	0.0710
		Panel C: Using weekly data					
(H1)	δ	-4.20		-2.11	2.56		1.01
		(-16.37)		(-8.44)	(11.11)		(4.17)
(H3)	l		-1.04	-0.26		0.88	0.23
			(-29.85)	(-5.92)		(23.45)	(5.70)
(H4)	$\delta \times l$			-30.80			24.40
				(-38.51)			(23.52)
	Pseudo R^2	0.0081	0.0751	0.0888	0.0041	0.0762	0.0860

Notes: The table reports the estimation results of quantile regressions:

$$Q_p\left(z_{i,t}|\delta_i, l_{i,t-1,h}\right) = \gamma_0 + \gamma_1\delta_i + \gamma_2 l_{i,t-1,h} + \gamma_3\delta_i l_{i,t-1,h} \text{ for the left tail, and}$$
$$Q_{1-p}\left(z_{i,t}|\delta_i, l_{i,t-1,h}\right) = \gamma_0 + \gamma_1\delta_i + \gamma_2 l_{i,t-1,h} + \gamma_3\delta_i l_{i,t-1,h} \text{ for the right tail.}$$

Here, $z_{i,t}$ is the normalized value of spread i at time t, δ_i is the mean-reversion speed and $l_{i,t-1,h}$ is the loss indicator defined as $\mathbf{1}\left(z_{i,t-1} < z_{i,t-1-h} < -\underline{z}\right)$ for the left tail and $\mathbf{1}\left(z_{i,t-1} > z_{i,t-1-h} > \underline{z}\right)$ for the right tail. We set $p = 1\%$, $\underline{z} = 1$ and $h = 1$. For Panel A, we use Hodrick-Prescott filter to detrend each spread. For Panel B, we estimate δ_i using the data within the quantile range of $(p, 1-p)$. We use daily U.S. interest rate swap market data from the Bloomberg over the sample period of July 23, 1998 to May 11, 2017. The t-statistics are computed by a bootstrap method.

Table 7: The Return Characteristics of Barclay Hedge Fund Index

Statistics	Hedge Fund Index	CB Arbitrage	Distressed Securities	Emerging Markets	Equity Long Bias	Equity Long/ Short	Equity Market Neutral	European Equities	Event Driven
Mean	0.71	0.59	0.65	0.77	0.80	0.74	0.43	0.83	0.72
Median	0.77	0.71	0.92	1.09	0.90	0.69	0.41	0.69	0.89
Min	-8.41	-13.76	-8.81	-19.40	-13.46	-6.92	-2.87	-5.95	-9.94
Max	7.73	7.14	5.63	14.54	9.40	10.99	3.07	13.95	6.17
Std.Dev.	2.00	1.73	2.02	3.99	3.22	2.02	0.87	2.30	1.87
Skewness	-0.62	-2.77	-1.32	-0.73	-0.63	0.85	0.03	1.69	-0.99
Kurtosis	3.61	23.75	3.83	3.72	1.91	4.77	1.68	8.74	4.21
Prob(+)	0.70	0.75	0.71	0.66	0.64	0.67	0.74	0.68	0.70
Statistics	Fixed Income Arbitrage	Fund of Funds	Global Macro	Healthcare & Bio-tec	Merger Arbitrage	Multi Strategy	Pacific Rim Equity	Tech	
Mean	0.49	0.43	0.60	1.24	0.60	0.68	0.67	0.94	
Median	0.62	0.52	0.42	1.02	0.67	0.77	0.56	0.83	
Min	-13.55	-6.79	-3.27	-14.27	-4.57	-7.62	-8.51	-11.88	
Max	4.23	6.05	7.53	40.60	3.00	3.88	14.81	19.58	
Std.Dev.	1.40	1.52	1.62	4.84	0.97	1.25	2.73	3.75	
Kurtosis	-4.90	-0.74	0.80	2.41	-1.23	-2.21	1.10	0.79	
Skewness	42.54	4.77	1.62	18.97	4.58	12.89	4.81	4.43	
Prob(+)	0.83	0.67	0.63	0.68	0.80	0.79	0.64	0.62	

This table reports summary statistics of monthly percentage returns from Barclay Hedge Fund Index for various styles. The last row of Prob(+) represents the frequency of positive monthly returns over the sample period of Jan, 1997 to Aug, 2017.

Internet Appendix for

"The More Efficient, the More Vulnerable!"

Proofs

Lemma A.1. *For every node x, it holds that $\Gamma_x \leq V$, $V - S_x \leq P_x \leq V$, $-1 \leq \psi_x \leq \frac{1}{2\phi}\left(\frac{S_b}{V-S_b}\right)$ and $1 \leq v_x \leq \left(\frac{V}{V-S_b} + \frac{1}{4\phi}\left(\frac{S_b}{V-S_b}\right)^2\right)^3$.*

Proof Note that when x is a final node, $P_x = V$ and $v_x = 1$. Thus, it suffices to show the following property: For a non-final node x, if $P_{x'} \leq V$ and $1 \leq v_{x'} \leq k$ for every node x' following node x, then $\Gamma_x \leq V$, $V - S_x \leq P_x \leq V$, $-1 \leq \psi_x \leq \frac{1}{2\phi}\left(\frac{S_b}{V-S_b}\right)$ and $1 \leq v_x \leq k\left(\frac{V}{V-S_b} + \frac{1}{4\phi}\left(\frac{S_b}{V-S_b}\right)^2\right)$. By applying this property recursively to $t = 2, 1, 0$, we will obtain the lemma because $\frac{V}{V-S_b} + \frac{1}{4\phi}\left(\frac{S_b}{V-S_b}\right)^2 > 1$.

Assume that $P_{x'} \leq V$ and $1 \leq v_{x'} \leq k$ for every node x' following node x, and show the followings.

Step 1. $\Gamma_x \leq V$: From the definition of (2.3),

$$\Gamma_x = \frac{\mathbb{E}_x\left[P_{x'}v_{x'}\right]}{\mathbb{E}_x\left[v_{x'}\right]} \leq \frac{\mathbb{E}_x\left[Vv_{x'}\right]}{\mathbb{E}_x\left[v_{x'}\right]} = V,$$

where the inequality is from $P_{x'} \leq V$ and $v_{x'} \geq 1$ for every node x'.

Step 2. $V - S_b \leq P_x$: Because $\psi_x \geq -1$, it follows that $P_x = V - S_x + W_x\left(1 + \psi_x\right) \geq V - S_x \geq V - S_b$.

Step 3. $P_x \leq V$: Assume $P_x > V$ by contradiction. Then, by Step 1, we get $P_x > \Gamma_x$, implying $\psi_x = -1$ by Lemma 2.2. The market clearing condition of (2.6) gives that $P_x = V - S_x \leq V$, contradicting the assumption of $P_x > V$.

Step 4. $-1 \leq \psi_x \leq \frac{1}{2\phi}\left(\frac{S_b}{V-S_b}\right)$: The lower bound of ψ_x is by our model assumption. The upper bound of ψ_x follows from the optimal ψ_x given in Lemma 2.2 along with $\Gamma_x \leq V$ (Step 1) and $V - S_b \leq P_x$ (Step 2).

Step 5. $1 \leq v_{x'} \leq k\left(\frac{V}{V-S_b} + \frac{1}{4\phi}\left(\frac{S_b}{V-S_b}\right)^2\right)$: From (2.4) and the optimal ψ_x given in Lemma 2.2, it holds that when $P_x \leq \Gamma$,

$$v_x = \mathbb{E}_x\left[v_{x'}\right]\left(\frac{\Gamma_x}{P_x}\left(1-1\right) + \left(1 - \phi 1_{-1>0}\right)\right) = \mathbb{E}_x\left[v_{x'}\right] \in \left[1, k\right],$$

and that when $P_x > \Gamma$,

$$
\begin{aligned}
v_x &= \mathbb{E}_x\left[v_{x'}\right]\left(\frac{\Gamma_x}{P_x}\left(1+\frac{1}{2\phi}\left(\frac{\Gamma_x}{P_x}-1\right)\right)-\frac{1}{2\phi}\left(\frac{\Gamma_x}{P_x}-1\right)\left(1+\frac{1}{2\phi}\left(\frac{\Gamma_x}{P_x}-1\right)\phi\right)\right)\\
&= \mathbb{E}_x\left[v_{x'}\right]\left(\frac{\Gamma_x}{P_x}+\frac{1}{4\phi}\left(\frac{\Gamma_x}{P_x}-1\right)^2\right).
\end{aligned}
$$

Because

$$
1 < \frac{\Gamma_x}{P_x}+\frac{1}{4\phi}\left(\frac{\Gamma_x}{P_x}-1\right)^2 < \frac{V}{V-S_b}+\frac{1}{4\phi}\left(\frac{S_b}{V-S_b}\right)^2
$$

by Steps 1 and 2, it follows that $1 \le v_x \le k\left(\frac{V}{V-S_b}+\frac{1}{4\phi}\left(\frac{S_b}{V-S_b}\right)^2\right)$. This completes the proof of the lemma. □

Lemma A.2. *It holds that*

$$
0 < \frac{V-S_b}{V}\left(1-\frac{1}{4\phi}\frac{S_m\left(2S_b-S_m\right)}{\left(V-S_m\right)\left(V-S_b\right)}\right) \le \frac{W_{x'}}{W_x} \le \frac{V}{V-S_x}\left[1+\frac{S_x^2}{4\phi\left(V-S_x\right)V}\right]
$$

for every non-final node x, followed by node x'.

Proof The first inequality is implied by Assumption 2.

Turn to the second inequality. From (2.2), we have that

$$
\frac{W_{x'}}{W_x} = \frac{P_{x'}}{P_x}\left(1+\psi_x\right)-\psi_x\left(1+\phi\psi_x 1_{\psi_x>0}\right).
$$

Consider four cases:

Case 1: If $-1 \le \psi_x \le 0$ and $P_{x'} \ge P_x$,

$$
\frac{W_{x'}}{W_x} = \frac{P_{x'}}{P_x}+\left(\frac{P_{x'}}{P_x}-1\right)\psi_x \ge \frac{P_{x'}}{P_x}-\left(\frac{P_{x'}}{P_x}-1\right) = 1 > \frac{V-S_b}{V}\left(1-\frac{1}{4\phi}\frac{S_m\left(2S_b-S_m\right)}{\left(V-S_m\right)\left(V-S_b\right)}\right).
$$

Case 2: If $-1 \le \psi_x \le 0$ and $P_{x'} < P_x$,

$$
\frac{W_{x'}}{W_x} = \frac{P_{x'}}{P_x}+\left(\frac{P_{x'}}{P_x}-1\right)\psi_x \ge \frac{P_{x'}}{P_x} \ge \frac{V-S_b}{V} > \frac{V-S_b}{V}\left(1-\frac{1}{4\phi}\frac{S_m\left(2S_b-S_m\right)}{\left(V-S_m\right)\left(V-S_b\right)}\right),
$$

where the second inequality is implied by $P_{x'} \ge V-S_b$ and $P_x \le V$ from Lemma A.1.

Case 3: If $\psi_x > 0$ and x' is a terminal node, since $P_{x'} = V \ge P_x$,

$$
\frac{W_{x'}}{W_x} = \frac{V}{P_x}+\left(\frac{V}{P_x}-1\right)\psi_x \ge \frac{V}{P_x} = 1 > \frac{V-S_b}{V}\left(1-\frac{1}{4\phi}\frac{S_m\left(2S_b-S_m\right)}{\left(V-S_m\right)\left(V-S_b\right)}\right).
$$

Case 4: If $\psi_x > 0$ and x' is not a terminal node, Lemma 2.2 says that $\psi_x = \frac{1}{2\phi}\left(\frac{\Gamma_x}{P_x} - 1\right)$ and $\Gamma_x > P_x$, which yields

$$\begin{aligned}
\frac{W_{x'}}{W_x} =& \frac{P_{x'}}{P_x}\left(1 + \frac{1}{2\phi}\left(\frac{\Gamma_x}{P_x} - 1\right)\right) - \frac{1}{2\phi}\left(\frac{\Gamma_x}{P_x} - 1\right)\frac{1}{2}\left(\frac{\Gamma_x + P_x}{P_x}\right) \\
\geq& \frac{V - S_b}{P_x}\left(1 + \frac{1}{2\phi}\left(\frac{\Gamma_x}{P_x} - 1\right)\right) - \frac{1}{4\phi}\left(\frac{\Gamma_x}{P_x} - 1\right)\left(\frac{\Gamma_x + P_x}{P_x}\right) \\
=& \frac{V - S_b}{P_x}\left[1 + \frac{1}{4\phi}\left(\frac{\Gamma_x}{P_x} - 1\right)\frac{2\left(V - S_b\right) - \Gamma_x - P_x}{V - S_b}\right] \\
\geq& \frac{V - S_b}{P_x}\left[1 + \frac{1}{4\phi}\left(\frac{V}{P_x} - 1\right)\frac{V - 2S_b - P_x}{V - S_b}\right] \\
\geq& \frac{V - S_b}{P_x}\left[1 + \frac{1}{4\phi}\left(\frac{V}{V - S_m} - 1\right)\frac{V - 2S_b - V + S_m}{V - S_b}\right] \\
=& \frac{V - S_b}{P_x}\left[1 - \frac{1}{4\phi}\left(\frac{V}{V - S_m} - 1\right)\frac{2S_b - S_m}{V - S_b}\right] \\
\geq& \frac{V - S_b}{V}\left[1 - \frac{1}{4\phi}\left(\frac{V}{V - S_m} - 1\right)\frac{2S_b - S_m}{V - S_b}\right] \\
=& \frac{V - S_b}{V}\left[1 - \frac{1}{4\phi}\left(\frac{S_m}{V - S_m}\right)\left(\frac{2S_b - S_m}{V - S_b}\right)\right],
\end{aligned}$$

where the first and second inequalities are from $P_{x'} \geq V - S_b$ and $\Gamma_x \leq V$ (Lemma A.1). Besides, because $\frac{\partial\left(\frac{V}{P_x} - 1\right)\frac{V - 2S_b - P_x}{V - S_b}}{\partial P_x} = \frac{1}{P_x^2}\frac{1}{V - S_b}\left(-V^2 + 2S_bV + P_x^2\right) > 0$, $P_x \geq V - S_x \geq V - S_m$ (Lemma A.1 and node x is at $t = 0$ or $t = 1$) yields the third inequality. And, the last inequality is from $P_x \leq V$ (Lemma A.1).

Finally, we show that $\frac{W_{x'}}{W_x} \leq \frac{V}{V - S_x}\left[1 + \frac{S_x^2}{4\phi(V - S_x)V}\right]$. Consider three cases:

Case 1: If $-1 \leq \psi_x \leq 0$ and $P_{x'} \geq P_x$,

$$\frac{W_{x'}}{W_x} = \frac{P_{x'}}{P_x} + \left(\frac{P_{x'}}{P_x} - 1\right)\psi_x \leq \frac{P_{x'}}{P_x} \leq \frac{V}{V - S_b} < \frac{V}{V - S_b} + \frac{1}{2}.$$

Case 2: If $-1 \leq \psi_x \leq 0$ and $P_{x'} < P_x$,

$$\frac{W_{x'}}{W_x} = \frac{P_{x'}}{P_x} + \left(\frac{P_{x'}}{P_x} - 1\right)\psi_x \leq \frac{P_{x'}}{P_x} - \left(\frac{P_{x'}}{P_x} - 1\right) = 1 < \frac{V}{V - S_b} + \frac{1}{2}.$$

Case 3: If $\psi_x > 0$, Lemma 2.2 says that $\psi_x = \frac{1}{2\phi}\left(\frac{\Gamma_x}{P_x} - 1\right)$, which yields

$$\begin{aligned}
\frac{W_{x'}}{W_x} =& \frac{P_{x'}}{P_x}\left(1+\psi_x\right) - \psi_x\left(1+\phi\psi_x\right) \\
=& \frac{P_{x'}}{P_x}\left(1+\frac{1}{2\phi}\left(\frac{\Gamma_x}{P_x}-1\right)\right) - \frac{1}{2\phi}\left(\frac{\Gamma_x}{P_x}-1\right)\left(1+\phi\frac{1}{2\phi}\left(\frac{\Gamma_x}{P_x}-1\right)\right) \\
\leq& \frac{V}{P_x}\left(1+\frac{1}{2\phi}\left(\frac{\Gamma_x}{P_x}-1\right)\right) - \frac{1}{4\phi}\left(\frac{\Gamma_x}{P_x}-1\right)\left(\frac{\Gamma_x+P_x}{P_x}\right) \\
=& \frac{V}{P_x}\left[1+\frac{\left(V-P_x\right)^2-\left(V-\Gamma_x\right)^2}{4\phi P_x V}\right] \\
\leq& \frac{V}{P_x}\left[1+\frac{\left(V-P_x\right)^2}{4\phi P_x V}\right] \leq \frac{V}{V-S_x}\left[1+\frac{S_x^2}{4\phi\left(V-S_x\right)V}\right],
\end{aligned}$$

where the first and last inequalities hold because $P_{x'} \leq V$ and $V - S_b \leq P_x \leq V$ (Lemma A.1).

This completes the proof of the lemma. □

Proof of Lemma 2.1 By Lemma A.2,

$$0 < \frac{W_{x'}}{W_x} \leq \frac{V}{V-S_x}\left(1+\frac{S_x^2}{4\phi\left(V-S_x\right)V}\right)$$

for every non-final node x, followed by node x'. It is easy to see $W_x > 0$ for every node. For node x' at $t = 1$ and x'' at $t = 2$,

$$W_{x'} < W_0\frac{V}{V-S_0}\left(1+\frac{S_0^2}{4\phi\left(V-S_0\right)V}\right) < S_m$$

and

$$\begin{aligned}
W_{x''} <& W_{x'}\frac{V}{V-S_m}\left(1+\frac{S_m^2}{4\phi\left(V-S_m\right)V}\right). \\
<& W_0\frac{V}{V-S_0}\left(1+\frac{S_0^2}{4\phi\left(V-S_0\right)V}\right)\frac{V}{V-S_m}\left(1+\frac{S_m^2}{4\phi\left(V-S_m\right)V}\right) < S_m,
\end{aligned}$$

where the last inequalities are from Assumption 2. This completes the proof. □

Lemma A.3. *Assume that $0 < W < S < V$ and $V - S + W < \Gamma \leq V$. Then, $\widetilde{P}\left(S,W,\Gamma\right)$ defined by (2.7) has the following properties:*
(i) $\frac{\partial\widetilde{P}(S,W,\Gamma)}{\partial S} < 0$,

(ii) $V - S < \widetilde{P}(S, W, \Gamma) < \Gamma$,
(iii) $0 < \frac{\partial \widetilde{P}(S,W,\Gamma)}{\partial W} = \frac{1+\psi}{\frac{1}{2\phi}\frac{VW}{P^2}+1} < 1 + \frac{1}{2\phi}\left(\frac{S}{V-S}\right)$,
(iv) $\frac{\partial^2 \widetilde{P}(S,W,\Gamma)}{\partial W^2} < 0$ *and*
(v) $\frac{\partial \widetilde{P}(S,W,\Gamma)}{\partial \Gamma} \to 0$ *as* $W \downarrow 0$.

Proof To see (i), observe

$$\frac{\partial \widetilde{P}(S, W, \Gamma)}{\partial S} < -\frac{1}{2} + \frac{1}{2}\frac{\left|V - S + W\left(1 - \frac{1}{2\phi}\right)\right|}{\sqrt{\left(V - S + W\left(1 - \frac{1}{2\phi}\right)\right)^2 + 2\frac{W\Gamma}{\phi}}} < 0.$$

Show (ii). First, we verify $\widetilde{P}(S, W, \Gamma) < \Gamma$. Recall that $\widetilde{P}(S, W, \Gamma)$ is the unique positive solution of

$$P - V + S - W\left(1 + \frac{1}{2\phi}\left(\frac{\Gamma}{P} - 1\right)\right) = 0.$$

The left-hand side is increasing in P and if $P = \Gamma$, the left-hand side becomes

$$\Gamma - V + S - W\left(1 + \frac{1}{2\phi}\left(\frac{\Gamma}{\Gamma} - 1\right)\right) = \Gamma - V + S - W > 0.$$

Thus, $\widetilde{P}(S, W, \Gamma) < \Gamma$. Second, we examine $V - S < \widetilde{P}(S, W, \Gamma)$. Note that $\widetilde{P}(S, W, \Gamma)$ is the unique positive solution of

$$P^2 - \left(V - S + W\left(1 - \frac{1}{2\phi}\right)\right)P - \frac{W\Gamma}{2\phi} = 0.$$

The left-hand side is increasing in P on $P > 0$ and

$$(V - S)^2 - \left(V - S + W\left(1 - \frac{1}{2\phi}\right)\right)(V - S) - \frac{W\Gamma}{2\phi} = -W(V - S) - \frac{W(\Gamma - V + S)}{2\phi} < 0.$$

Thus, $V - S < \widetilde{P}(S, W, \Gamma)$.

Turn to (iii). A simple algebra can verify $\frac{\partial \widetilde{P}(S,W,\Gamma)}{\partial W} = \frac{1+\psi}{\frac{1}{2\phi}\frac{VW}{P^2}+1}$. Noting that $\widetilde{P}(S, W, \Gamma) < \Gamma$ by (ii), we have $\psi = \frac{1}{2\phi}\left(\frac{\Gamma}{P} - 1\right) > 0$. Hence, we have $\frac{\partial \widetilde{P}(S,W,\Gamma)}{\partial W} > 0$. Also, observe that

$$\frac{1+\psi}{\frac{1}{2\phi}\frac{VW}{P^2} + 1} < 1 + \psi = 1 + \frac{1}{2\phi}\left(\frac{\Gamma}{P} - 1\right) \leq 1 + \frac{1}{2\phi}\left(\frac{S}{V - S}\right),$$

where the last inequality is from the assumption $\Gamma \leq V$ and $V - S < P$ by (ii).

Next, we show (iv). It suffices to show that $\frac{\partial^2 \sqrt{f(W)}}{\partial W^2} < 0$ where $f(W)$ is given by

$$f(W) = \left(V - S + W\left(1 - \frac{1}{2\phi}\right)\right)^2 + 2\frac{W\Gamma}{\phi}.$$

Some algebras show that

$$\frac{\partial^2 \sqrt{f(W)}}{\partial W^2} = -\frac{2f^{-\frac{3}{2}}\Gamma(V-S)}{\phi}\left(1 + \frac{1}{2\phi}\left(\frac{\Gamma}{V-S} - 1\right)\right) < 0.$$

Lastly, note that as $W \downarrow 0$,

$$\frac{\partial \widetilde{P}(S, W, \Gamma)}{\partial \Gamma} = \frac{W}{\phi}\frac{1}{\sqrt{\left(V - S + W\left(1 - \frac{1}{2\phi}\right)\right)^2 + 2\frac{W\Gamma}{\phi}}} \to 0,$$

verifying (v). This completes the proof of the lemma. □

Proof of Lemma 2.3 We use the fact that $\Gamma_x = V$ at $t = 2$ and consider the two cases.

Case 1. $S_x = 0$: Note that $P_x = V - S_x + W_x(1 + \psi_x) \geq V$ because $S_x = 0$ and $\psi_x \geq -1$. Assume $P_x > V$ by contradiction. Then, $\psi_x = -1$ by Lemma 2.2 and $P_x = V - S_x + W_x(1 + \psi_x) = V$, which is a contradiction. Thus, $P_x = V$. The market clearing condition implies $\psi_x = -1$.

Case 2. $S_x > 0$: Note that $S_x > W_x$ by Lemma 2.1.

First, we show that $V > P_x$. Assume $V \leq P_x$. Lemma 2.2 implies $\psi_x \leq 0$. Then, $P_x = V - S_x + W_x(1 + \psi_x) \leq V - S_x + W_x < V$, which is a contradiction. Thus, $P_x \geq V$ cannot be a solution. Hence, it holds that $V > P_x$.

Because $V > P_x$, Lemma 2.2 implies that $\psi_x = \frac{1}{2\phi}\left(\frac{V}{P_x} - 1\right)$, which in conjunction with (2.6) yields

$$P_x = V - S_x + W_x\left(1 + \frac{1}{2\phi}\left(\frac{V}{P_x} - 1\right)\right).$$

Solving it gives $P_x = \widetilde{P}(S_x, W_x, V) < V$. Note that the inequality $V > P_x$ holds by Lemma A.3(ii) and hence $P_x = \widetilde{P}(S_x, W_x, V)$ is an equilibrium price. □

Proof of Lemma 2.4 Using (2.6) and the assumption $\psi_{11} \geq -1$, we have

$$P_{11} = V + W_{11}(1 + \psi_{11}) \geq V \geq \Gamma_{11} = \frac{\mathbb{E}_{11}\left[P_{2i|11} v_{2i|11}\right]}{\mathbb{E}_{11}\left[v_{2i|11}\right]}.$$

Here, the latter inequality holds because $P_{2i|11} \leq V$ for $i = 1, 2, 3$ by Lemma 2.3.

To show $P_{11} = V$, assume $P_{11} > V$. Then, $P_{11} > \Gamma_{11}$ and hence Lemma 2.2 implies $\psi_{11} = -1$ and in turn $P_{11} = V$ by (2.6). This is a contradiction and thus $P_{11} = V$. Then (2.6) gives $\psi_{11} = -1$. $\square$

Lemma A.4. $\Gamma_{12} > P_{12}$ *and* $\Gamma_0 > P_0$ *for sufficiently small* q *and* ρ.

Proof Consider $x = 12$ first. Note that $\Gamma_{12} = \frac{\mathbb{E}_{12}[P_{x'} v_{x'}]}{\mathbb{E}_{12}[v_{x'}]}$ can be arbitrarily close to $P_{21|12}$ by taking small q and ρ. By Lemma 2.3, $P_{21|12} = V$. Noting that $V - S_m + W_0 \left(\frac{V}{V - S_b} + \frac{S_b^2 (V - S_m)}{S_m (2S_b - S_m)(V - S_b)} \right) < V$ by Assumption 2, we have

$$V - S_m + W_0 \left(\frac{V}{V - S_b} + \frac{S_b^2 (V - S_m)}{S_m (2S_b - S_m)(V - S_b)} \right) < \Gamma_{12}$$

for sufficiently small q and ρ. For such q and ρ, we show that $\Gamma_{12} > P_{12}$. Assume $\Gamma_{12} \leq P_{12}$ by contradiction. Lemma 2.2 implies $\psi_{12} \leq 0$ and thus

$$\begin{aligned} P_{12} &= V - S_m + W_{12} (1 + \psi_{12}) \\ &\leq V - S_m + W_{12} \\ &< V - S_m + W_0 \left(\frac{V}{V - S_b} + \frac{S_b^2 (V - S_m)}{S_m (2S_b - S_m)(V - S_b)} \right) < \Gamma_{12}. \end{aligned}$$

where the second inequality follows by Lemma A.2.

Turn to $x = 0$. The proof is similar to the case of $x = 12$. Note that Γ_0 is close to P_{11} for small q and ρ, and $P_{11} = V$ by Lemma 2.4. Because $V - S_m + W_0 < V$ by Assumption 2, $V - S_m + W_0 < \Gamma_0$ for sufficiently small q and ρ. Assume $\Gamma_0 \leq P_0$ by contradiction. Then, Lemma 2.2 implies $\psi_{12} \leq 0$ and

$$P_0 = V - S_0 + W_0 (1 + \psi_0) \leq V - S_m + W_0 < \Gamma_0.$$

This completes the proof. $\square$

Proof of Lemma 2.5 Let $x = 0$ or $x = 12$. By Lemma A.4, $\Gamma_x > P_x$ for sufficiently small q and ρ. Then, $\psi_x = \frac{1}{2\phi} \left(\frac{\Gamma_x}{P_x} - 1 \right)$ by Lemma 2.2. Combine it with (2.6) to obtain

$$P_x = V - S_x + W_x \left(1 + \frac{1}{2\phi} \left(\frac{\Gamma_x}{P_x} - 1 \right) \right)$$

which gives $P_x = \widetilde{P} (S_x, W_x, \Gamma_x)$.

Next, we turn to $P_x < V$. This follows by combining $P_x < \Gamma_x$ (Lemma A.4) and $\Gamma_x \leq V$ (Lemma A.1). This completes the proof of the lemma. $\square$

Lemma A.5. *It holds that* $W_{11} = W_{21|11} = W_{22|11} = W_{23|11}$.

Proof From Lemma 2.4, $\psi_{11} = -1$. Plugging $\psi_{11} = -1$ to (2.2) yields the equalities. $\square$

Before considering small q and ρ, we consider the case $q = \rho = 0$ first. At the non-reachable nodes (i.e., all those except 0, 11, $21|\cdot$, $31|\cdot$) as well as the other nodes, we assume arbitrageurs make rational choices and the market clears. It is easy to see that the equilibrium for $q = \rho = 0$ satisfies $\Gamma_x = V$ for every node x. Then, as proved in Lemmas 2.3, 2.4 and 2.5, we have the following equations:

$$P_{21|11} = P_{21|12} = P_{11} = V, \tag{A.1}$$

$$P_{2i|1j} = \widetilde{P}\left(S_{2i|11}, W_{11}, V\right) \text{ for } i = 2, 3, \tag{A.2}$$

$$P_{2i|1j} = \widetilde{P}\left(S_{2i|12}, W_{2i|12}, V\right) \text{ for } i = 2, 3, \tag{A.3}$$

$$\psi_{2i|1j} = \frac{1}{2\phi}\left(\frac{V}{P_{2i|1j}} - 1\right) \text{ for } i = 2, 3 \text{ and } j = 1, 2, \tag{A.4}$$

$$P_{12} = \widetilde{P}\left(S_m, W_{12}, V\right), \; P_0 = \widetilde{P}\left(S_0, W_0, V\right), \tag{A.5}$$

$$\psi_{12} = \frac{1}{2\phi}\left(\frac{V}{P_{12}} - 1\right), \; \psi_0 = \frac{1}{2\phi}\left(\frac{V}{P_0} - 1\right). \tag{A.6}$$

The capital process W_x is given by (2.2) which completes the system of equations. We show that this system of equation has a unique solution, and then will consider the case that q and ρ are small.

Lemma A.6. *The system of equations, (2.2) and (A.1)-(A.6), has a unique solution. Moreover, for any non-final node x followed by a non-final node x', if $W_x < S_x \leq S_{x'} < V$, it holds that $W_{x'} < W_x$, $0 < 1 - \frac{\partial \widetilde{P}(S_{x'}, W_{x'}, V)}{\partial W_{x'}} \cdot \frac{W_x(1+\psi_x)}{P_x} < 1$ and $\frac{dW_{x'}}{dS_{x'}} < 0$.*

Proof Note that prices and leverage in the system of equations, (A.1)-(A.6), $P_{2i|1j}$ and $\psi_{2i|1j}$ for $i = 2, 3$ and $j = 1, 2$, P_{12}, ψ_{12}, and P_0 and ψ_0, are uniquely determined once we identify W_0, W_{12}, $W_{2j|12}$ for $j = 2, 3$. Note that W_0 is a given constant. Hence, for the uniqueness of solutions for (2.2) and (A.1)-(A.6), it suffices to show that for any non-final node x followed by a non-final node x' such that $W_x < S_x \leq S_{x'}$, when W_x is given, $W_{x'}$ is uniquely determined.

From (2.2), we have

$$f(w) \equiv w - W_x\left(\frac{\widetilde{P}\left(S_{x'}, w, V\right)}{P_x}\left(1 + \psi_x\right) - \psi_x\left(1 + \phi\psi_x\right)\right) = 0. \tag{A.7}$$

Here, W_x is viewed as given and so are $P_x = \widetilde{P}\left(S_x, W_x, V\right)$ and $\psi_x = \frac{1}{2\phi}\left(\frac{V}{P_x} - 1\right)$. Then, (A.7) is an equation for w and its solution w is $W_{x'}$ because $f\left(W_{x'}\right) = 0$. We prove that there is a unique solution to (A.7).

Using $W_x > 0$ from Lemma 2.1, we have

$$f(0) = -W_x \left(\frac{\widetilde{P}(S_{x'}, 0, V)}{P_x} (1 + \psi_x) - \psi_x (1 + \phi \psi_x) \right) < 0$$

because

$$\begin{aligned}
&\frac{\widetilde{P}(S_{x'}, 0, V)}{P_x} \left(1 + \frac{1}{2\phi} \left(\frac{V}{P_x} - 1 \right) \right) - \frac{1}{2\phi} \left(\frac{V}{P_x} - 1 \right) \left(1 + \phi \frac{1}{2\phi} \left(\frac{V}{P_x} - 1 \right) \right) \\
=& \frac{\widetilde{P}(S_{x'}, 0, V)}{P_x} \left(1 + \frac{1}{2\phi} \left(\frac{V}{P_x} - 1 \right) \right) - \frac{1}{2\phi} \left(\frac{V}{P_x} - 1 \right) \frac{1}{2} \left(\frac{V + P_x}{P_x} \right) \\
\geq& \frac{V - S_b}{P_x} \left(1 + \frac{1}{2\phi} \left(\frac{V}{P_x} - 1 \right) \right) - \frac{1}{4\phi} \left(\frac{V}{P_x} - 1 \right) \left(\frac{V + P_x}{P_x} \right) \\
=& \frac{V - S_b}{P_x} \left[1 + \frac{1}{4\phi} \left(\frac{V}{P_x} - 1 \right) \frac{V - 2S_b - P_x}{V - S_b} \right] \\
\geq& \frac{V - S_b}{P_x} \left[1 - \frac{1}{4\phi} \left(\frac{V}{V - S_m} - 1 \right) \frac{2S_b - S_m}{V - S_b} \right] \\
=& \frac{V - S_b}{P_x} \left[1 - \frac{1}{4\phi} \left(\frac{S_m}{V - S_m} \right) \left(\frac{2S_b - S_m}{V - S_b} \right) \right] > 0,
\end{aligned}$$

where the first inequality follows by $\widetilde{P}(S_{x'}, 0, V) \geq V - S_b$ (Lemma A.3(ii)), the second inequality by $\frac{\partial \left(\frac{V}{P_x} - 1 \right) \frac{V - 2S_b - P_x}{V - S_b}}{\partial P_x} = \frac{1}{P_x^2} \frac{1}{V - S_b} \left(-V^2 + 2S_b V + P_x^2 \right) > 0$ and $P_x \geq V - S_m$ (Lemma A.1 and node x is at $t = 0$ or $t = 1$) and the last inequality by Assumption 2, respectively.

On the other hand,

$$\begin{aligned}
f(W_x) =& W_x \left[1 - \frac{\widetilde{P}(S_{x'}, W_x, V)}{P_x} (1 + \psi_x) + \psi_x (1 + \phi \psi_x) \right] \\
\geq& W_x \left[1 - \frac{\widetilde{P}(S_x, W_x, V)}{P_x} (1 + \psi_x) + \psi_x (1 + \phi \psi_x) \right] \\
=& W_x \left[1 - (1 + \psi_x) + \psi_x (1 + \phi \psi_x) \right] > 0
\end{aligned}$$

The first inequality is from Lemma A.3(i) and the assumption $S_{x'} \geq S_x$, and the latter from $\psi_x = \frac{1}{2\phi} \left(\frac{V}{\widetilde{P}(S_x, W_x, V)} - 1 \right) > 0$ implied by Lemma A.3(ii). Because f is continuous, $f(w) = 0$ has a solution on $(0, W_x)$ by the Intermediate Value Theorem. Moreover, f is strictly convex on $(0, \infty)$ by Lemma A.3(iv) and thus the solution should be unique on $(0, \infty)$.

Turn to the property of

$$\frac{d}{dw}f\left(w\right)=1-\frac{\widetilde{P}\left(S_{x'},w,V\right)}{\partial W_{x'}}\cdot\frac{W_{x}\left(1+\psi_{x}\right)}{P_{x}}.$$

It is clear that $\frac{d}{dw}f\left(w\right)<1$ at $w=W_{x'}$ by Lemma A.3(iii). Because $f\left(0\right)<0$ and $f\left(w\right)=0$ has a unique solution, $\frac{d}{dw}f\left(w\right)>0$ at the solution $w=W_{x'}$. Thus

$$0<1-\frac{\widetilde{P}\left(S_{x'},W_{x'},V\right)}{\partial W_{x'}}\cdot\frac{W_{x}\left(1+\psi_{x}\right)}{P_{x}}<1. \tag{A.8}$$

The last property can be seen by differentiating both sides of

$$W_{x'}=W_{x}\left(\frac{\widetilde{P}\left(S_{x'},W_{x'},V\right)}{P_{x}}\left(1+\psi_{x}\right)-\psi_{x}\left(1+\phi\psi_{x}\right)\right).$$

This gives

$$\frac{dW_{x'}}{dS_{x'}}=\frac{W_{x}\left(1+\psi_{x}\right)}{P_{x}}\left(\frac{\partial\widetilde{P}\left(S_{x'},W_{x'},V\right)}{\partial S_{x'}}+\frac{\partial\widetilde{P}\left(S_{x'},W_{x'},V\right)}{\partial W_{x'}}\frac{dW_{x'}}{dS_{x'}}\right)$$

and hence

$$\frac{dW_{x'}}{dS_{x'}}=\frac{\frac{W_{x}\left(1+\psi_{x}\right)}{P_{x}}}{1-\frac{\partial\widetilde{P}\left(S_{x'},W_{x'},V\right)}{\partial W_{x'}}\cdot\frac{W_{x}\left(1+\psi_{x}\right)}{P_{x}}}\cdot\frac{\partial\widetilde{P}\left(S_{x'},W_{x'},V\right)}{\partial S_{x'}}<0,$$

where the inequality is from Lemma A.3(i) and (A.8). □

Lemma A.7. *If $q=\rho=0$, it holds that $W_{11}>W_{0}>W_{12}>W_{22|12}>W_{23|12}$.*

Proof From Lemmas A.6 and A.3(ii), we have $P_{11}=V$, $P_{0}=\widetilde{P}\left(S_{0},W_{0},V\right)<V$ and $\psi_{0}>0$, implying $W_{11}>W_{0}$. The inequality $W_{0}>W_{12}>W_{22|12}$ is implied by Lemma A.6. The last inequality $W_{22|12}>W_{23|12}$ obtains by $S_{22|12}=S_{m}<S_{b}=S_{23|12}$ and $\frac{dW_{x'}}{dS_{x'}}<0$ proved in Lemma A.6. □

Lemma A.8. *If $q=\rho=0$, it holds that $P_{23|11}>P_{23|12}$.*

Proof From Lemmas A.5 and A.7, we have $W_{23|11}>W_{23|12}$, which in conjunction with Lemma A.3(iii) implies that $P_{23|11}=\widetilde{P}\left(S_{b},W_{23|11},V\right)>\widetilde{P}\left(S_{b},W_{23|12},V\right)=P_{23|12}$. □

Lemma A.9. *If $q=\rho=0$, it holds that $P_{x}<V$ for nodes $x=0,\ 12$.*

Proof From Lemmas A.6 and A.3(ii), $P_0 = \widetilde{P}(S_0, W_0, V) < V$. Since $W_{12} < W_0$ by Lemma A.7, we have that $P_{12} = \widetilde{P}(S_m, W_{12}, V) < V$ using (ii) and (iii) of Lemma A.3. $\square$

Recall that in our model, P_x, ψ_x, W_x, Γ_x and v_x are endogenously determined, which are mentioned in the following lemma.

Lemma A.10. *There are $\overline{q}, \overline{\rho}, \overline{M} > 0$ such that if $q < \overline{q}$ and $\rho < \overline{\rho}$, there exists a unique equilibrium and $\frac{d\kappa}{dq} < \overline{M}$ for any endogenous variable κ. Moreover the equilibrium is continuous in (q, ρ) on $[0, \overline{q}) \times [0, \overline{\rho})$.*

Proof Using (2.2) and Lemmas 2.3, 2.4 and 2.5, we write down the system of equations that determines an equilibrium:

$$\mathbf{\Xi}_{(1)} \equiv \begin{bmatrix} -P_0 + V - S_0 + W_0(1+\psi_0) \\ -\psi_0 + \frac{1}{2\phi}\left(\frac{1}{P_0}\frac{(1-q)Vv_{11} + qP_{12}v_{12}}{(1-q)v_{11} + qv_{12}} - 1\right) \\ -W_{11} + W_0\left(\frac{V}{P_0}(1+\psi_0) - \psi_0(1+\phi\psi_0)\right) \end{bmatrix} = \mathbf{0}_3,$$

$$\mathbf{\Xi}_{(2)} \equiv \begin{bmatrix} -P_{22|11} + \widetilde{P}(S_m, W_{11}, V) \\ -P_{23|11} + \widetilde{P}(S_b, W_{11}, V) \\ -v_{11} + (1 - q_{22} - q_{23}) + q_{22}v\left(P_{22|11}\right) + q_{23}v\left(P_{23|11}\right) \end{bmatrix} = \mathbf{0}_3,$$

$$\mathbf{\Xi}_{(3)} \equiv \begin{bmatrix} -W_{12} + W_0\left(\frac{P_{12}}{P_0}(1+\psi_0) - \psi_0(1+\phi\psi_0)\right) \\ -P_{12} + V - S_b + W_{12}(1+\psi_{12}) \\ -\psi_{12} + \frac{1}{2\phi}\left(\frac{1}{P_{12}}\frac{\mathbb{E}_{12}\left[P_{\cdot|12}v\left(P_{\cdot|12}\right)\right]}{\mathbb{E}_{12}\left[v\left(P_{\cdot|12}\right)\right]} - 1\right) \\ -v_{12} + \frac{\mathbb{E}_{12}\left[P_{\cdot|12}v_{\cdot|12}\right]}{P_{12}}(1+\psi_{12}) - \psi_{12}(1+\psi_{12}\phi)\,\mathbb{E}_{12}\left[v_{\cdot|12}\right] \end{bmatrix} = \mathbf{0}_4, \text{ and}$$

$$\mathbf{\Xi}_{(4)} \equiv \begin{bmatrix} -P_{22|12} + \widetilde{P}\left(S_m, W_{22|12}, V\right) \\ -W_{22|12} + W_{12}\left(\frac{P_{22|12}}{P_{12}}(1+\psi_{12}) - \psi_{12}(1+\phi\psi_{12})\right) \\ -P_{23|12} + \widetilde{P}\left(S_b, W_{23|12}, V\right) \\ -W_{23|12} + W_{12}\left(\frac{P_{23|12}}{P_{12}}(1+\psi_{12}) - \psi_{12}(1+\phi\psi_{12})\right) \end{bmatrix} = \mathbf{0}_4.$$

Here, $v(P) = \frac{V}{P} + \frac{1}{4\phi}\left(\frac{V}{P} - 1\right)^2$, $\mathbb{E}\left[P_{\cdot|12}v\left(P_{\cdot|12}\right)\right] = (1 - q_{22} - q_{23})V + q_{22}P_{22|12}v\left(P_{22|12}\right) + q_{23}P_{23|12}v\left(P_{23|12}\right)$, $\mathbb{E}\left[v\left(P_{\cdot|12}\right)\right] = (1 - q_{22} - q_{23}) + q_{22}v\left(P_{22|12}\right) + q_{23}v\left(P_{23|12}\right)$ and $\mathbf{0}_m$ represents the $(m \times 1)$ vector of zeros. Let $\mathbf{\Xi} = \left[\mathbf{\Xi}'_{(1)}\ \mathbf{\Xi}'_{(2)}\ \mathbf{\Xi}'_{(3)}\ \mathbf{\Xi}'_{(4)}\right]'$.

Allocate endogenous variables into four vectors:

$$\boldsymbol{\kappa}_{(1)} = [P_0 \ \psi_0 \ W_{11}]'$$
$$\boldsymbol{\kappa}_{(2)} = \left[P_{22|11} \ P_{23|11} \ v_{11}\right]'$$
$$\boldsymbol{\kappa}_{(3)} = [W_{12} \ P_{12} \ \psi_{12} \ v_{12}]'$$
$$\boldsymbol{\kappa}_{(4)} = \left[P_{22|12} \ W_{22|12} \ P_{23|12} \ W_{23|12}\right]'.$$

Let $\boldsymbol{\kappa} = \left[\boldsymbol{\kappa}'_{(1)} \ \boldsymbol{\kappa}'_{(2)} \ \boldsymbol{\kappa}'_{(3)} \ \boldsymbol{\kappa}'_{(4)}\right]'$.

The following five steps complete the proof

Step 1. $\boldsymbol{\Xi} = \mathbf{0}_{16}$ has a unique solution $\boldsymbol{\kappa}$ when $q = \rho = 0$:

Suppose $q = \rho = 0$. The fourteen variables of prices, leverage and capital in $\boldsymbol{\kappa}$ except v_{11} and v_{12} are uniquely determined by Lemma A.6. Then, the last condition of $\boldsymbol{\Xi}_{(2)} = \mathbf{0}_3$ pins down v_{11} and that of $\boldsymbol{\Xi}_{(3)} = \mathbf{0}_4$ pins down v_{22}.

Step 2. $\frac{\partial \boldsymbol{\Xi}}{\partial \boldsymbol{\kappa}}$ is full rank at $q = \rho = 0$:

Suppose $q = \rho = 0$. Identify the Jacobian of $\boldsymbol{\Xi}$ with respect to $\boldsymbol{\kappa}$ at $q = \rho = 0$. Some algebras show that at $q = \rho = 0$, it holds that

$$\frac{\partial \boldsymbol{\Xi}}{\partial \boldsymbol{\kappa}} = \begin{bmatrix} \frac{\partial \boldsymbol{\Xi}_{(1)}}{\partial \boldsymbol{\kappa}_{(1)}} & \mathbf{0}_{3\times 3} & \mathbf{0}_{3\times 4} & \mathbf{0}_{3\times 4} \\ \frac{\partial \boldsymbol{\Xi}_{(2)}}{\partial \boldsymbol{\kappa}_{(1)}} & \frac{\partial \boldsymbol{\Xi}_{(2)}}{\partial \boldsymbol{\kappa}_{(2)}} & \mathbf{0}_{3\times 4} & \mathbf{0}_{3\times 4} \\ \frac{\partial \boldsymbol{\Xi}_{(3)}}{\partial \boldsymbol{\kappa}_{(1)}} & \mathbf{0}_{4\times 3} & \frac{\partial \boldsymbol{\Xi}_{(3)}}{\partial \boldsymbol{\kappa}_{(3)}} & \mathbf{0}_{4\times 4} \\ \mathbf{0}_{4\times 3} & \mathbf{0}_{4\times 3} & \frac{\partial \boldsymbol{\Xi}_{(4)}}{\partial \boldsymbol{\kappa}_{(3)}} & \frac{\partial \boldsymbol{\Xi}_{(4)}}{\partial \boldsymbol{\kappa}_{(4)}} \end{bmatrix},$$

where

$$
\frac{\partial \boldsymbol{\Xi}_{(1)}}{\partial \boldsymbol{\kappa}_{(1)}}=\begin{bmatrix} -1 & W_0 & 0 \\ -\frac{1}{2\phi}\frac{V}{P_0^2} & -1 & 0 \\ -\frac{W_0V(1+\psi_0)}{P_0^2} & 0 & -1 \end{bmatrix},\ \frac{\partial \boldsymbol{\Xi}_{(2)}}{\partial \boldsymbol{\kappa}_{(1)}}=\begin{bmatrix} 0 & 0 & \frac{\partial \widetilde{P}(S_m,W_{11},V)}{\partial W} \\ 0 & 0 & \frac{\partial \widetilde{P}(S_m,W_{11},V)}{\partial W} \\ 0 & 0 & 0 \end{bmatrix},\ \frac{\partial \boldsymbol{\Xi}_{(2)}}{\partial \boldsymbol{\kappa}_{(2)}}=\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}
$$

$$
\frac{\partial \boldsymbol{\Xi}_{(3)}}{\partial \boldsymbol{\kappa}_{(1)}}=\begin{bmatrix} -\frac{W_0P_{12}(1+\psi_0)}{P_0^2} & \frac{W_0(P_{12}-V)}{P_0} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},\ \frac{\partial \boldsymbol{\Xi}_{(3)}}{\partial \boldsymbol{\kappa}_{(3)}}=\begin{bmatrix} -1 & \frac{W_0(1+\psi_0)}{P_0} & 0 & 0 \\ 1+\psi_{12} & -1 & W_{12} & 0 \\ 0 & -\frac{1}{2\phi}\frac{V}{P_{12}^2} & -1 & 0 \\ 0 & -\frac{V(1+\psi_{12})}{P_{12}^2} & 0 & -1 \end{bmatrix}
$$

$$
\frac{\partial \boldsymbol{\Xi}_{(4)}}{\partial \boldsymbol{\kappa}_{(3)}}=\begin{bmatrix} 0 & 0 & 0 & 0 \\ \frac{W_{22|12}}{W_{12}} & -\frac{W_{12}P_{22|12}(1+\psi_{12})}{P_{12}^2} & \frac{W_{12}\left(P_{12|22}-V\right)}{P_{12}} & 0 \\ 0 & 0 & 0 & 0 \\ \frac{W_{23|12}}{W_{12}} & -\frac{W_{12}P_{23|12}(1+\psi_{12})}{P_{12}^2} & \frac{W_{12}\left(P_{23|12}-V\right)}{P_{12}} & 0 \end{bmatrix}
$$

$$
\frac{\partial \boldsymbol{\Xi}_{(4)}}{\partial \boldsymbol{\kappa}_{(4)}}=\begin{bmatrix} -1 & \frac{\partial \widetilde{P}\left(S_m,W_{22|12},V\right)}{\partial W} & 0 & 0 \\ \frac{W_{12}(1+\psi_{12})}{P_{12}} & -1 & 0 & 0 \\ 0 & 0 & -1 & \frac{\partial \widetilde{P}\left(S_b,W_{23|12},V\right)}{\partial W} \\ 0 & 0 & \frac{W_{12}(1+\psi_{12})}{P_{12}} & -1 \end{bmatrix},
$$

and $\mathbf{0}_{m\times n}$ represents the $(m\times n)$ matrix of zeros.

We verify that $\frac{\partial \boldsymbol{\Xi}}{\partial \boldsymbol{\kappa}}$ is a full rank matrix. Because $\frac{\partial \boldsymbol{\Xi}}{\partial \boldsymbol{\kappa}}$ is a lower triangular block matrix, it suffices to show that the four diagonal blocks of $\frac{\partial \boldsymbol{\Xi}_{(m)}}{\partial \boldsymbol{\kappa}_{(m)}}$ for $m=1,2,3,4$ are full rank matrices. Note that $\det\left(\frac{\partial \boldsymbol{\Xi}_{(1)}}{\partial \boldsymbol{\kappa}_{(1)}}\right)=-1-\frac{1}{2\phi}\frac{VW_0}{P_0^2}<0$ and $\det\left(\frac{\partial \boldsymbol{\Xi}_{(2)}}{\partial \boldsymbol{\kappa}_{(2)}}\right)=-1<0$. Also, we find that

$$
\begin{aligned}
\det\left(\frac{\partial \boldsymbol{\Xi}_{(3)}}{\partial \boldsymbol{\kappa}_{(3)}}\right) &= -1-\frac{1}{2\phi}\frac{V}{P_{12}}\frac{W_{12}}{P_{12}}+\frac{W_0\left(1+\psi_0\right)}{P_0}\left(1+\psi_{12}\right) \\
&= -\frac{1}{2\phi}\frac{V}{P_{12}}\frac{W_{12}}{P_{12}}-1+\frac{W_0\left(1+\psi_0\right)}{P_0}\left(\frac{1}{2\phi}\frac{V}{P_{12}}\frac{W_{12}}{P_{12}}+1\right)\frac{\partial \widetilde{P}\left(S_m,W_{12},V\right)}{\partial W} \\
&= \left(\frac{1}{2\phi}\frac{VW}{P^2}+1\right)\left(-1+\frac{W_0\left(1+\psi_0\right)}{P_0}\frac{\partial \widetilde{P}\left(S_m,W_{12},V\right)}{\partial W}\right)<0,
\end{aligned}
$$

where the second equality is due to the property of $\frac{\partial \widetilde{P}(S_m,W_{12},V)}{\partial W}=\frac{1+\psi_{12}}{\frac{1}{2\phi}\frac{V}{P_{12}}\frac{W_{12}}{P_{12}}+1}$ by Lemma A.3(iii) and the last inequality is due to the inequality of $-1+\frac{\partial \widetilde{P}(S_{x'},W_{x'},V)}{\partial W_{x'}}\cdot\frac{W_x(1+\psi_x)}{P_x}<0$ from Lemma A.6. For the last block of $\frac{\partial \boldsymbol{\Xi}_{(4)}}{\partial \boldsymbol{\kappa}_{(4)}}$, we find that

$$
\det\left(\frac{\partial \boldsymbol{\Xi}_{(4)}}{\partial \boldsymbol{\kappa}_{(4)}}\right)=\left(1-\frac{W_{12}\left(1+\psi_{12}\right)}{P_{12}}\frac{\partial \widetilde{P}\left(S_m,W_{22|12},V\right)}{\partial W}\right)\left(1-\frac{W_{12}\left(1+\psi_{12}\right)}{P_{12}}\frac{\partial \widetilde{P}\left(S_b,W_{23|12},V\right)}{\partial W}\right)>0,
$$

where the last inequality is obtained from Lemma A.6. Hence, $\frac{\partial \boldsymbol{\Xi}}{\partial \boldsymbol{\kappa}}$ is a full rank matrix.

Step 3. For sufficiently small q and ρ, $\boldsymbol{\Xi} = \mathbf{0}_{16}$ has a unique solution $\boldsymbol{\kappa}$: This is true by the implicit function theorem and Steps 1 and 2.

Step 4. For sufficiently small q and ρ, there is a unique equilibrium: The variables in $\boldsymbol{\kappa}$ are uniquely determined by Step 3. It is easy to see that all the other endogenous variables are unique. For example, $P_{11} = V$ by Lemma 2.4.

Step 5. There are $\overline{q}, \overline{\rho}, \overline{M} > 0$ such that if $q < \overline{q}$ and $\rho < \overline{\rho}$, $\left\| \frac{d\kappa}{dq} \right\| < \overline{M}$ for any endogenous variable κ: From Lemmas 2.4 and 2.3, we know that $\frac{d\kappa}{dq} = 0$ for $\kappa = P_{21|\cdot}, P_{11}, \psi_{21|\cdot}$ and ψ_{11}. By Steps 1 and 2, the implicit function theorem implies that $\frac{d\boldsymbol{\kappa}}{dq}$ is bounded on a neighborhood of $q = \rho = 0$. Also, as shown in Step 4, other endogenous variables are continuous and bounded functions of the variables in $\boldsymbol{\kappa}$. □

Proof of Theorem 2.1 This is proved in Lemma A.10. □

Proof of Theorem 3.1 Because the equilibrium is continuous in q and ρ when q and ρ are sufficiently small (Lemma A.10) and $P_{23|11} > P_{23|12}$ at $q = \rho = 0$ (Lemma A.8), it holds that $P_{23|11} > P_{23|12}$ when q and ρ are sufficiently small. □

Proof of Theorem 3.2 Recall $P_{23|11} = \widetilde{P}\left(S_b, W_{23|11}, V\right)$ to get

$$\frac{dP_{23|11}}{dq} = \frac{\partial \widetilde{P}\left(S_b, W_{23|11}, V\right)}{\partial W} \cdot \frac{dW_{23|11}}{dq}. \tag{A.9}$$

Because $W_{23|11} = W_{11} = W_0\left(\frac{V}{P_0}(1+\psi_0) - \psi_0(1+\psi_0\phi)\right)$ from Lemma A.5 and (2.2), it holds that

$$\frac{dW_{23|11}}{dq} = -W_0 \frac{V(1+\psi_0)}{P_0^2}\frac{dP_0}{dq} - W_0\left(\frac{V}{P_0} - 1 - 2\phi\psi_0\right)\frac{d\psi_0}{dq}. \tag{A.10}$$

Because $\psi_0 = \frac{1}{2\phi}\left(\frac{\Gamma_0}{P_0} - 1\right)$ and Γ_0 is close to V, the term $\left(\frac{V}{P_0} - 1 - 2\phi\psi_0\right)$ is close to 0. Then, boundedness of $\frac{d\psi_0}{dq}$ by Lemma A.10 implies $W_0\left(\frac{V}{P_0} - 1 - 2\phi\psi_0\right)\frac{d\psi_0}{dq} \to 0$ as $(q,\rho) \to 0$.

Next, we identify $\frac{dP_0}{dq}$. Lemma 2.5 gives

$$\begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \equiv \begin{bmatrix} -P_0 + V - S_0 + W_0(1+\psi_0) \\ -\psi_0 + \frac{1}{2\phi}\left(\frac{1}{P_0}\frac{(1-q)Vv_{11} + qP_{12}v_{12}}{(1-q)v_{11} + qv_{12}} - 1\right) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

and hence

$$\begin{bmatrix} -1 & W_0 \\ -\frac{1}{2\phi}\frac{V}{P_0^2} & -1 \end{bmatrix} \begin{bmatrix} \frac{dP_0}{dq} \\ \frac{d\psi_0}{dq} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{\partial G_2}{\partial v_{11}}\frac{dv_{11}}{dq} + \frac{\partial G_2}{\partial v_{12}}\frac{dv_{12}}{dq} + \frac{\partial G_2}{\partial P_{12}}\frac{dP_{12}}{dq} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{\partial G_2}{\partial q} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Some algebras show that as $q \downarrow 0$,

$$\frac{\partial G_2}{\partial v_{11}}, \ \frac{\partial G_2}{\partial v_{12}}, \ \frac{\partial G_2}{\partial P_{12}} \to 0 \text{ and } \frac{\partial G_2}{\partial q} \to \frac{1}{2\phi}\frac{v_{12}\left(P_{12}-V\right)}{v_{11}P_0},$$

which, in conjunction with the boundedness of $\frac{dv_{11}}{dq}$, $\frac{dv_{12}}{dq}$ and $\frac{dP_{12}}{dq}$ (Lemma A.10), implies that

$$\begin{bmatrix} \frac{dP_0}{dq} \\ \frac{d\psi_0}{dq} \end{bmatrix} \to - \begin{bmatrix} -1 & W_0 \\ -\frac{1}{2\phi}\frac{V}{P_0^2} & -1 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \frac{1}{2\phi}\frac{v_{12}(P_{12}-V)}{v_{11}P_0} \end{bmatrix} = -\frac{\frac{(V-P_{12})v_{12}}{2\phi P_0 v_{11}}}{1+\frac{VW_0}{2\phi P_0^2}} \begin{bmatrix} W_0 \\ 1 \end{bmatrix}. \tag{A.11}$$

Plug (A.11) to (A.10) and observe that as $q, \rho \downarrow 0$,

$$\frac{dW_{23|11}}{dq} \to \frac{W_0^2 V\left(1+\psi_0\right)}{P_0^2}\frac{\frac{(V-P_{12})v_{12}}{2\phi P_0 v_{11}}}{1+\frac{VW_0}{2\phi P_0^2}} > 0. \tag{A.12}$$

Lastly, because $\frac{\partial \widetilde{P}}{\partial W} > 0$ from Lemma A.3(iii), we conclude that

$$\frac{dP_{23|11}}{dq} > 0$$

from (A.9) and (A.12). □

Lemma A.11. *There are $\overline{q}, \overline{\rho}, \overline{W}_0, \underline{\delta}, \overline{\delta} > 0$ such that if $q < \overline{q}$, $\rho < \overline{\rho}$ and $W_0 < \overline{W}_0$, it holds that $-\left(\underline{\delta}+\overline{\delta}\right) < \frac{d\Gamma_0}{dq} < -\underline{\delta}$.*

Proof Recall that $\Gamma_0 = \frac{(1-q)v_{11}V+qv_{12}P_{12}}{(1-q)v_{11}+qv_{12}}$. Hence,

$$\frac{d\Gamma_0}{dq} = \frac{\partial \Gamma_0}{\partial q} + \frac{\partial \Gamma_0}{\partial v_{11}}\frac{dv_{11}}{dq} + \frac{\partial \Gamma_0}{\partial v_{12}}\frac{dv_{12}}{dq} + \frac{\partial \Gamma_0}{\partial P_{12}}\frac{dP_{12}}{dq}.$$

After some algebras, we find that as $q \downarrow 0$,

$$\frac{\partial \Gamma_0}{\partial q} \to \frac{v_{12}}{v_{11}}\left(P_{12}-V\right), \text{ and } \frac{\partial \Gamma_0}{\partial v_{11}}, \ \frac{\partial \Gamma_0}{\partial v_{12}}, \ \frac{\partial \Gamma_0}{\partial P_{12}} \to 0.$$

Also, from $0 < W_{12} < W_0$ (Lemmas 2.1 and A.7) and the boundedness of ψ_{12} (Lemma A.1), it follows that $W_0 \downarrow 0$, $P_{12} = V - S_m + W_{12}\left(1+\psi_{12}\right) \to V - S_m$. Lastly, because v_{11} and v_{12} are between 1 and $\left(\frac{V}{V-S_b} + \frac{1}{4\phi}\left(\frac{S_b}{V-S_b}\right)^2\right)^3$ by Lemma A.1 and $\frac{dv_{11}}{dq}$, $\frac{dv_{12}}{dq}$ and $\frac{dP_{12}}{dq}$ are bounded by Lemma A.10, we conclude that

$$-\left(\frac{V}{V-S_b} + \frac{1}{4\phi}\left(\frac{S_b}{V-S_b}\right)^2\right)^3 S_m < \frac{d\Gamma_0}{dq} < -\left(\frac{V}{V-S_b} + \frac{1}{4\phi}\left(\frac{S_b}{V-S_b}\right)^2\right)^{-3} S_m.$$

Then, the lemma is proved by setting $\underline{\delta} = \left(\frac{V}{V-S_b} + \frac{1}{4\phi}\left(\frac{S_b}{V-S_b}\right)^2\right)^{-3} S_m$ and $\overline{\delta} = -\underline{\delta} + \left(\frac{V}{V-S_b} + \frac{1}{4\phi}\left(\frac{S_b}{V-S_b}\right)^2\right)^3 S_m$. □

Lemma A.12. *For sufficiently small q, ρ and W_0, it holds that $\frac{d\Gamma_{12}}{dq} < -S_m$.*

Proof Recall that $\Gamma_{12} = \frac{(1-q-q_{23})V + qv_{22|12}P_{22|12} + q_{23}v_{23|12}P_{23|12}}{(1-q-q_{23}) + qv_{22|12} + q_{23}v_{23|12}}$. Hence,

$$\frac{d\Gamma_{12}}{dq} = \frac{\partial \Gamma_{12}}{\partial q} + \frac{\partial \Gamma_{12}}{\partial v_{22|12}}\frac{dv_{22|12}}{dq} + \frac{\partial \Gamma_{12}}{\partial v_{23|12}}\frac{dv_{23|12}}{dq} + \frac{\partial \Gamma_{12}}{\partial P_{22|12}}\frac{dP_{22|12}}{dq} + \frac{\partial \Gamma_{12}}{\partial P_{23|12}}\frac{dP_{23|12}}{dq}.$$

After some algebras, we find that as $q, \rho \downarrow 0$,

$$\frac{\partial \Gamma_{12}}{\partial q} \to v_{22|12}\left(P_{22|12} - V\right), \text{ and } \frac{\partial \Gamma_{12}}{\partial v_{22|12}}\ \frac{\partial \Gamma_{12}}{\partial v_{23|12}}\ \frac{\partial \Gamma_{12}}{\partial P_{22|12}},\ \frac{\partial \Gamma_{12}}{\partial P_{23|12}} \to 0.$$

Also, from $0 < W_{22|12} < W_0$ (Lemmas 2.1 and A.7) and the boundedness of $\psi_{22|12}$ (Lemma A.1), it follows that $W_0 \downarrow 0$, $P_{22|12} = V - S_m + W_{22|12}\left(1 + \psi_{22|12}\right) \to V - S_m$. Lastly, because $v_{22|12} \geq 1$ by Lemma A.1 and $\frac{dv_{22|12}}{dq}, \frac{dv_{23|12}}{dq}, \frac{dP_{22|12}}{dq}, \frac{dP_{23|12}}{dq}$ and $\frac{dP_{12}}{dq}$ are bounded from Lemma A.10, we conclude that $\frac{d\Gamma_{12}}{dq} < -S_m$. □

Lemma A.13. *As $W_0 \downarrow 0$, $\frac{dP_0}{dq} \to 0$ and $\frac{d\psi_0}{dq} - \frac{1}{2\phi(V-S_0)}\frac{d\Gamma_0}{dq} \to 0$.*

Proof Because $P_0 = \widetilde{P}\left(S_0, W_0, \Gamma_0\right)$, it follows that $\frac{dP_0}{dq} = \frac{\partial \widetilde{P}(S_0, W_0, \Gamma_0)}{\partial \Gamma}\frac{d\Gamma_0}{dq}$. By Lemma A.3(v) and the boundedness of $\frac{d\Gamma_0}{dq}$ (Lemma A.10), we have that $\frac{dP_0}{dq} \to 0$ as $W_0 \downarrow 0$.

Next, from $\psi_0 = \frac{1}{2\phi}\left(\frac{\Gamma_0}{P_0} - 1\right)$, we have

$$\frac{d\psi_0}{dq} = -\frac{V}{2\phi P_0^2}\frac{dP_0}{dq} + \frac{1}{2\phi P_0}\frac{d\Gamma_0}{dq}. \tag{A.13}$$

From $P_0 = V - S_0 + W_0\left(1 + \psi_0\right)$ and the boundedness of ψ_0 (Lemma A.1), it follows that as $W_0 \downarrow 0$, $P_0 \to V - S_0$. Plugging in $\frac{dP_0}{dq} \to 0$ and $P_0 \to V - S_0$ to (A.13) yields that $\frac{d\psi_0}{dq} - \frac{1}{2\phi(V-S_0)}\frac{d\Gamma_0}{dq} \to 0$. This completes the proof of the lemma. □

Lemma A.14. *Pick any $k > 0$. Then, there are $\overline{q}, \overline{\rho}, \overline{W}_0, \underline{\lambda}, \overline{\lambda} > 0$ such that if $q < \overline{q}$, $\rho < \overline{\rho}$ and $W_0 < \overline{W}_0$, it holds that $\underline{\lambda} < \frac{1}{W_0}\frac{dW_{12}}{dq} < \underline{\lambda} + \overline{\lambda}$ and $k\frac{dW_{12}}{dq} - \frac{dW_{11}}{dq} > 0$.*

Proof From $W_{12} = W_0\left(\frac{P_{12}}{P_0}\left(1 + \psi_0\right) - \psi_0\left(1 + \psi_0\phi\right)\right)$, it follows that

$$\frac{dW_{12}}{dq} = -W_0\frac{P_{12}}{P_0^2}\left(1 + \psi_0\right)\frac{dP_0}{dq} + \frac{W_0}{P_0}\left(1 + \psi_0\right)\frac{dP_{12}}{dq} + W_0\left(\frac{P_{12}}{P_0} - 1 - 2\phi\psi_0\right)\frac{d\psi_0}{dq}.$$

By plugging $\frac{dP_{12}}{dq} = \frac{\partial \widetilde{P}(S_m, W_{12}, \Gamma_{12})}{\partial W}\frac{dW_{12}}{dq} + \frac{\partial \widetilde{P}(S_m, W_{12}, \Gamma_{12})}{\partial \Gamma_{12}}\frac{d\Gamma_{12}}{dq}$ derived from $P_{12} = \widetilde{P}\left(S_m, W_{12}, \Gamma_{12}\right)$ into the above equation, we have

$$\frac{1}{W_0}\left(1 - \frac{W_0}{P_0}\left(1 + \psi_0\right)\frac{\partial \widetilde{P}\left(S_m, W_{12}, \Gamma_{12}\right)}{\partial W}\right)\frac{dW_{12}}{dq} \tag{A.14}$$
$$= \left(-\frac{P_{12}}{P_0^2}\left(1 + \psi_0\right)\frac{dP_0}{dq} + \frac{1}{P_0}\left(1 + \psi_0\right)\frac{\partial \widetilde{P}\left(S_m, W_{12}, \Gamma_{12}\right)}{\partial \Gamma_{12}}\frac{d\Gamma_{12}}{d} + \left(\frac{P_{12}}{P_0} - 1 - 2\phi\psi_0\right)\frac{d\psi_0}{dq}\right).$$

We show that there exist $\underline{\lambda}, \overline{\lambda} > 0$ such that $\underline{\lambda} < \frac{1}{W_0}\frac{dW_{12}}{dq} < \underline{\lambda} + \overline{\lambda}$ by the following properties.

Property 1. $\frac{W_0}{P_0}(1+\psi_0)\frac{\partial \widetilde{P}(S_m, W_{12}, \Gamma_{12})}{\partial W}$ is negligible with small W_0: Note that $P_0 \geq V - S_0 > 0$ by Lemma A.1, and ψ_0 and $\frac{\partial \widetilde{P}(S_m, W_{12}, \Gamma_{12})}{\partial W}$ are bounded by Lemmas A.1 and A.3(iii), respectively.

Property 2. $-\frac{P_{12}}{P_0^2}(1+\psi_0)\frac{dP_0}{dq} + \frac{1}{P_0}(1+\psi_0)\frac{\partial \widetilde{P}(S_m, W_{12}, \Gamma_{12})}{\partial \Gamma_{12}}\frac{d\Gamma_{12}}{dq}$ is negligible for sufficiently small W_0 : This holds because $\frac{d\psi_0}{dq}$ and $\frac{d\Gamma_{12}}{dq}$ are bounded from Lemma A.10, $\frac{dP_0}{dq} = \frac{d}{dq}(V - S_0 + W_0(1+\psi_0)) = W_0\frac{d\psi_0}{dq}$ and $\frac{\partial \widetilde{P}(S_m, W_{12}, \Gamma_{12})}{\partial \Gamma_{12}} \to 0$ by Lemma A.3(v).

Property 3. $\frac{P_{12}}{P_0} - 1 - 2\phi\psi_0 \to -\frac{S_m}{V-S_0}$: Because $\psi_0 = \frac{1}{2\phi}\left(\frac{\Gamma_0}{P_0} - 1\right)$, $\frac{P_{12}}{P_0} - 1 - 2\phi\psi_0 = \frac{P_{12} - \Gamma_0}{P_0}$. Note that $\Gamma_0 \to V$ as $q \downarrow 0$. Also, from $0 < W_{12} < W_0$ (Lemmas 2.1 and A.7) and the boundedness of ψ_0 and ψ_{12} (Lemma A.1), it follows that as $W_0 \downarrow 0$, $P_0 = V - S_0 + W_0(1+\psi_0) \to V - S_0$ and $P_{12} = V - S_m + W_{12}(1+\psi_{12}) \to V - S_m$. Hence, $\frac{P_{12}}{P_0} - 1 - 2\phi\psi_0 = \frac{P_{12}-\Gamma_0}{P_0} \to \frac{V - S_m - V}{V - S_0} = -\frac{S_m}{V-S_0}$.

Property 4. There exist $\underline{\zeta}, \overline{\zeta} > 0$ such that $-\left(\underline{\zeta} + \overline{\zeta}\right) < \frac{d\psi_0}{dq} < -\underline{\zeta}$: If W_0 is small enough, $\frac{d\psi_0}{dq}$ is close to $\frac{1}{2\phi(V-S_0)}\frac{d\Gamma_0}{dq}$ from Lemma A.13. From Lemma A.11, there exist $\underline{\delta}, \overline{\delta} > 0$ such that $-\left(\underline{\delta} + \overline{\delta}\right) < \frac{d\Gamma_0}{dq} < -\underline{\delta}$. Hence, the claim is proved by setting $\underline{\zeta} = \frac{1}{2\phi(V-S_0)}\underline{\delta}$ and $\overline{\zeta} = \frac{1}{2\phi(V-S_0)}\overline{\delta}$.

Property 5. There exist $\underline{\lambda}, \overline{\lambda} > 0$ such that $\underline{\lambda} < \frac{1}{W_0}\frac{dW_{12}}{dq} < \underline{\lambda} + \overline{\lambda}$: Set $\underline{\lambda} = \frac{S_m}{V-S_0}\underline{\zeta}$ and $\overline{\lambda} = \frac{S_m}{V-S_0}\overline{\zeta}$.

Turn to the latter part of the lemma. First, we show that $\frac{1}{W_0}\frac{dW_{11}}{dq}$ is negligible. From $W_{11} = W_0\left(\frac{V}{P_0}(1+\psi_0) - \psi_0(1+\psi_0\phi)\right)$, we have that

$$\frac{1}{W_0}\frac{dW_{11}}{dq} = -\frac{V}{P_0^2}(1+\psi_0)\frac{dP_0}{dq} + \left(\frac{V}{P_0} - 1 - 2\phi\psi_0\right)\frac{d\psi_0}{dq}. \tag{A.15}$$

Note that $-\frac{V}{P_0^2}(1+\psi_0)\frac{dP_0}{dq}$ is negligible for small W_0 because $\frac{dP_0}{dq} = W_0\frac{d\psi_0}{dq}$ and that $\frac{V}{P_0} - 1 - 2\phi\psi_0$ is negligible for small q. Also, $\frac{dP_0}{dq}$ and $\frac{d\psi_0}{dq}$ are bounded from Lemma A.10. Hence, $\frac{1}{W_0}\frac{dW_{11}}{dq}$ is negligible.

Lastly, we combine the above results to prove the second claim of the lemma. For any k, when q, ρ and W_0 are small, it holds that $k\frac{1}{W_0}\frac{dW_{12}}{dq} > k\underline{\lambda} > \frac{1}{W_0}\frac{dW_{11}}{dq}$, where the first inequality is from $\frac{1}{W_0}\frac{dW_{12}}{dq} > \underline{\lambda}$ and the second inequality holds because $\frac{1}{W_0}\frac{dW_{11}}{dq}$ is negligible. This implies that $k\frac{dW_{12}}{dq} - \frac{dW_{11}}{dq} > 0$, completing the proof of the lemma. $\square$

Proof of Theorems 3.3 and 3.4 First, we prove Theorem 3.3, $\frac{dP_{23|12}}{dq} > 0$ for sufficiently small q, ρ and W_0. Because

$$W_{23|12} = W_{12}\left(\frac{P_{23|12}}{P_{12}}\left(1+\psi_{12}\right) - \psi_{12}\left(1+\phi\psi_{12}\right)\right),$$

it holds that

$$\begin{aligned}\frac{dW_{23|12}}{dq} =& \frac{W_{23|12}}{W_{12}}\frac{dW_{12}}{dq} + \frac{W_{12}}{P_{12}}\left(1+\psi_{12}\right)\frac{dP_{23|12}}{dq} - \frac{W_{12}P_{23|12}}{P_{12}^2}\left(1+\psi_{12}\right)\frac{dP_{12}}{dq} \\ &+ W_{12}\left(\frac{P_{23|12}}{P_{12}} - 1 - 2\phi\psi_{12}\right)\frac{d\psi_{12}}{dq}. \end{aligned} \tag{A.16}$$

Here, the equality $\frac{P_{23|12}}{P_{12}}\left(1+\psi_{12}\right) - \psi_{12}\left(1+\phi\psi_{12}\right) = \frac{W_{23|12}}{W_{12}}$ is used. Plugging (A.16) into $\frac{dP_{23|12}}{dq} = \frac{\partial\widetilde{P}\left(S_b, W_{23|12}, V\right)}{\partial W}\frac{dW_{23|12}}{dq}$ yields

$$\begin{aligned}&\frac{1}{W_{12}}\left(1 - \frac{\partial\widetilde{P}\left(S_b, W_{23|12}, V\right)}{\partial W}\frac{W_{12}}{P_{12}}\left(1+\psi_{12}\right)\right)\frac{dP_{23|12}}{dq} \\ =&\frac{\partial\widetilde{P}\left(S_b, W_{23|12}, V\right)}{\partial W}\left(\frac{W_{23|12}}{W_{12}}\frac{W_0}{W_{12}}\left(\frac{1}{W_0}\frac{dW_{12}}{dq}\right) + A + B\right),\end{aligned} \tag{A.17}$$

where

$$A = -\frac{P_{23|12}}{P_{12}^2}\left(1+\psi_{12}\right)\left(W_0\frac{\partial\widetilde{P}\left(S_m, W_{12}, \Gamma_{12}\right)}{\partial W_{12}}\left(\frac{1}{W_0}\frac{dW_{12}}{dq}\right) + \frac{\partial\widetilde{P}\left(S_m, W_{12}, \Gamma_{12}\right)}{\partial\Gamma}\frac{d\Gamma_{12}}{dq}\right)$$

and

$$B = \left(\frac{P_{23|12}}{P_{12}} - 1 - 2\phi\psi_{12}\right)\frac{d\psi_{12}}{dq_{12}}.$$

Then, the inequality $\frac{dP_{23|12}}{dq_{12}} > 0$ is proved by the following properties.

Property 1. $0 < 1 - \frac{\partial\widetilde{P}\left(S_b, W_{23|12}, V\right)}{\partial W}\frac{W_{12}}{P_{12}}\left(1+\psi_{12}\right) < 1$ for sufficiently small q and ρ: This is implied by Lemma A.6 and the continuity of the economy.

Property 2. $\frac{\partial\widetilde{P}\left(S_b, W_{23|12}, V\right)}{\partial W} > 0$: This is implied by Lemma A.3(iii).

Property 3. There exists $k > 0$ such that $\frac{W_{23|12}}{W_{12}}\frac{W_0}{W_{12}}\left(\frac{1}{W_0}\frac{dW_{12}}{dq}\right) > k$ for sufficiently small q, ρ and W_0: From Lemmas A.7 and A.2 with the continuity of economy, it holds that $\frac{W_{23|12}}{W_{12}}\frac{W_0}{W_{12}} > \frac{W_{23|12}}{W_{12}} > \frac{V-S_b}{V}\left(1 - \frac{1}{2\phi}\left(\frac{S_b}{V-S_b}\right)^2\right)$. Also, by Lemma A.14, there exists $\underline{\lambda} > 0$ such that $\frac{1}{W_0}\frac{dW_{12}}{dq} > \underline{\lambda}$ for sufficiently small q, ρ and W_0. The property is true when $k = \frac{V-S_b}{V}\left(1 - \frac{1}{2\phi}\left(\frac{S_b}{V-S_b}\right)^2\right)\underline{\lambda}$.

Property 4. A is negligible for small enough q, ρ and W_0: The boundedness of $-\frac{P_{23|12}}{P_{12}^2}\left(1+\psi_{12}\right)$ follows from Lemma A.1. Note that $\frac{\partial\widetilde{P}\left(S_m, W_{12}, \Gamma_{12}\right)}{\partial W_{12}}$ and $\frac{1}{W_0}\frac{dW_{12}}{dq}$ are bounded from Lemmas A.3(iii) and A.14. Also, when W_0 is close to 0, so is W_{12} by Lemma A.7. Hence, it holds that $\frac{\partial \widetilde{P}(S_m, W_{12}, \Gamma_{12})}{\partial \Gamma} \to 0$ from Lemma A.3(v) and $\frac{d\Gamma_{12}}{dq}$ is bounded from Lemma A.10.

Property 5. There exists a constant $l > 0$ (that does not depend on q, ρ or W_0) such that $B > l$ for sufficiently small q, ρ and W_0: Because $\psi_{12} = \frac{1}{2\phi}\left(\frac{\Gamma_{12}}{P_{12}} - 1\right)$, it holds that

$$\frac{P_{23|12}}{P_{12}} - 1 - 2\phi\psi_{12} = \frac{P_{23|12} - \Gamma_{12}}{P_{12}} \to \frac{V - S_b - V}{V - S_m} = -\frac{S_b}{V - S_m} \tag{A.18}$$

as $(q, \rho, W_0) \downarrow 0$.

Next, we examine $\frac{d\psi_{12}}{dq}$. From $\psi_{12} = \frac{1}{2\phi}\left(\frac{\Gamma_{12}}{P_{12}} - 1\right)$,

$$\frac{d\psi_{12}}{dq} = -\frac{1}{2\phi}\frac{V}{P_{12}^2}\frac{dP_{12}}{dq} + \frac{1}{2\phi}\frac{\Gamma_{12}}{P_{12}}\frac{d\Gamma_{12}}{dq}.$$

Note that

$$\frac{dP_{12}}{dq} = W_0 \frac{\partial \widetilde{P}\left(S_m, W_{12}, \Gamma_{12}\right)}{\partial W}\left(\frac{1}{W_0}\frac{dW_{12}}{dq}\right) + \frac{\partial \widetilde{P}\left(S_m, W_{12}, \Gamma_{12}\right)}{\partial \Gamma}\frac{d\Gamma_{12}}{dq}.$$

When q, ρ and W_0 are small, $\frac{\partial \widetilde{P}(S_m, W_{12}, \Gamma_{12})}{\partial W}$, $\frac{1}{W_0}\frac{dW_{12}}{dq}$, and $\frac{d\Gamma_{12}}{dq}$ are bounded by Lemmas A.10 and A.14, and $\frac{\partial \widetilde{P}(S_m, W_{12}, \Gamma_{12})}{\partial \Gamma}$ is negligible by Lemma A.3(v) and $0 < W_{12} < W_0$ (Lemma A.7 and continuity of the economy). Thus, $\frac{dP_{12}}{dq}$ is negligible when q, ρ and W_0 are small. Lemma A.12 implies $\frac{d\Gamma_{12}}{dq} < -S_m$ for sufficiently small q, ρ and W_0. Hence, noting that $\Gamma_{12} \to V$ and $P_{12} \to V - S_m$, we have that, for sufficiently small q, ρ and W_0,

$$\frac{d\psi_{12}}{dq} < -\frac{1}{2\phi}\frac{V}{V - S_m}S_m. \tag{A.19}$$

Lastly, combining (A.18) and (A.19), we conclude that $B = \left(\frac{P_{23|12}}{P_{12}} - 1 - 2\phi\psi_{12}\right)\frac{d\psi_{12}}{dq_{12}} > l$ with $l = \frac{1}{2\phi}\frac{S_b V}{(V - S_m)^2}S_m$.

Next, we prove Theorem 3.4, $\frac{dP_{23|12}}{dq} - \frac{dP_{23|11}}{dq} > 0$ for sufficiently small q, ρ and W_0. From $P_{23|11} = \widetilde{P}\left(S_b, W_{23|11}, V\right)$, it follows that

$$\frac{1}{W_{12}}\frac{dP_{23|11}}{dq} = \frac{1}{W_{12}}\frac{\partial \widetilde{P}\left(S_b, W_{23|11}, V\right)}{\partial W}\frac{dW_{23|11}}{dq} = \frac{\partial \widetilde{P}\left(S_b, W_{23|11}, V\right)}{\partial W}\left(\frac{1}{W_{12}}\frac{dW_{11}}{dq}\right), \tag{A.20}$$

where the last equality is due to $W_{23|11} = W_{11}$ from Lemma A.5. We prove the inequality $\frac{dP_{23|12}}{dq} - \frac{dP_{23|11}}{dq} > 0$ by comparing (A.17) and (A.20) through the following properties.

Property 1. $0 < 1 - \frac{\partial \widetilde{P}\left(S_b, W_{23|12}, V\right)}{\partial W}\frac{W_{12}}{P_{12}}(1 + \psi_{12}) < 1$ for sufficiently small q and ρ: This holds by Lemma A.6 and the continuity of economy.

Property 2. $\frac{\partial \widetilde{P}\left(S_b, W_{23|12}, V\right)}{\partial W} > \frac{\partial \widetilde{P}\left(S_b, W_{23|11}, V\right)}{\partial W} > 0$ for sufficiently small q and ρ: From Lemmas A.7 and A.5 along with the continuity of economy, we have $W_{23|12} < W_{11} = W_{23|11}$. Then, use Lemma A.3(iii) and (iv).

Property 3. $\frac{W_{23|12}}{W_{12}} \frac{W_0}{W_{12}} \left(\frac{1}{W_0} \frac{dW_{12}}{dq}\right) > \frac{1}{W_{12}} \frac{dW_{11}}{dq}$ for sufficiently small q, ρ and W_0: From Lemmas A.7, A.2 and A.14 with continuity of economy, we have that $\frac{W_{23|12}}{W_{12}} > \frac{V-S_b}{V}\left(1 - \frac{1}{2\phi}\left(\frac{S_b}{V-S_b}\right)^2\right)$, $\frac{dW_{12}}{dq} > 0$ and $\frac{V-S_b}{V}\left(1 - \frac{1}{2\phi}\left(\frac{S_b}{V-S_b}\right)^2\right) \frac{dW_{12}}{dq} > \frac{dW_{11}}{dq}$ for sufficiently small q, ρ and W_0. Then, the property is proved.

Property 4. $A + B > 0$ for sufficiently small q, ρ and W_0: Recall that we have shown previously that A is negligible and there exists $l > 0$ such that $B > l$ (Properties for Theorem 3.3). This completes the proof of the Theorem. $\square$