DISCUSSION PAPER SERIES

No. 4768

ON THE ROLE OF ARBITRAGEURS IN RATIONAL MARKETS

Suleyman Basak and Benjamin Croitoru

FINANCIAL ECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP4768.asp

ISSN 0265-8003

ON THE ROLE OF ARBITRAGEURS IN RATIONAL MARKETS

Suleyman Basak, London Business School (LBS) and CEPR
Benjamin Croitoru, McGill University

Discussion Paper No. 4768
December 2004

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

CEPR Discussion Paper No. 4768

December 2004

ABSTRACT

On the Role of Arbitrageurs in Rational Markets*

Price discrepancies, although at odds with mainstream finance, are persistent phenomena in financial markets. These apparent mispricings lead to the presence of 'arbitrageurs', who aim to exploit the resulting profit opportunities, but whose role remains controversial. This article investigates the impact of the presence of arbitrageurs in rational financial markets. Arbitrage opportunities between redundant risky assets arise endogenously in an economy populated by rational, heterogeneous investors facing restrictions on leverage and short sales. An arbitrageur, indulging in costless, riskless arbitrage is shown to alleviate the effects of these restrictions and improve the transfer of risk amongst investors. When the arbitrageur lacks market power, he always takes on the largest arbitrage position possible. When the arbitrageur behaves non-competitively, in that he takes into account the price impact of his trades, he optimally limits the size of his positions due to his decreasing marginal profits. In the case when the arbitrageur is subject to margin requirements and is endowed with capital from outside investors, the size of the arbitrageur's trades and the capital needed to implement these trades are endogenously solved for in equilibrium.

JEL Classification: C60, D50, D90, G11 and G12
Keywords: arbitrage, asset pricing, margin requirements, non-competitive markets and risk-sharing

Suleyman Basak
Institute of Finance and Accounting
London Business School
Regents Park
London NW1 4SA
Tel: (44 20) 7706 6847
Fax: (44 20) 7724 3317
Email: sbasak@london.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=148578

Benjamin Croitoru
Faculty of Management
McGill University
1001 Sherbrooke Street West
Montreal
Quebec H3A 1G5
CANADA
Tel: (1 514) 398 3237
Fax: (1 514) 398 3876
Email: benjamin.croitoru@mcgill.ca

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=148594

*We are grateful to an anonymous referee for valuable suggestions and to Viral Acharya, Domenico Cuoco, Michael Gallmeyer, Wei Xiong, seminar participants at McGill University, the University of Wisconsin-Madison, the 2003 GERAD Optimization Days, the Blaise Pascal International Conference on Financial Modeling, the 2004 American Finance Association meetings and the 2004 European Finance Association meetings for their helpful comments. Financial support from ESRC for the first author and from FQRSC, IFM2 and SSHRC for the second author is gratefully acknowledged. All errors are solely our responsibility.

Submitted 18 November 2004

1. Introduction

The presence of apparent inconsistencies in asset prices has long been well-documented. Over the years, various types of securities such as primes and scores, stock index futures, closed end funds have been found to consistently deviate from their no-arbitrage values. Recent examples include the paradoxical behavior of prices in some equity carve-outs (Lamont and Thaler (2003))[1] and the deviations from put-call parity in options markets (Ofek, Richardson and Whitelaw (2004)). Such mispricings clearly lead to opportunities for arbitrage profits. Not surprisingly, there is ample evidence of market participants engaging in trades designed to reap these profits. For example, a *single* type of arbitrage transaction, stock index arbitrage, is estimated by the NYSE (www.nyse.com) to typically make up 7–9% of the program trading volume, itself representing about one third of total NYSE trading (there are many other instruments frequently involved in arbitrages, such as bonds and derivatives.) There is also considerable anecdotal evidence of some investors specializing in such trades. As Shleifer (2000, Chapter 4) puts it, "commonly, arbitrage is conducted by relatively few, highly specialized investors." A key example of such specialized players is that of hedge funds, many of which specialize in arbitrage strategies (Lhabitant (2002)). Their economic role, however, is highly controversial (especially after the infamous collapse of the Long Term Capital Management hedge fund in 1998), underscoring the importance of a rigorous study of the impact of arbitrage.

There is a growing, recent academic literature attempting to study these phenomena. The avenue pursued by much of this literature (see Shleifer (2000) for a survey) has been to postulate the presence of irrational noise traders whose erratic behavior pushes prices out of line and generates mispricings. In such work, an "arbitrageur" is typically synonymous for a rational trader (as opposed to irrational noise traders).

In this paper, we employ a different approach, little explored in the literature so far. We present an equilibrium in which all market participants are rational. Instead of resorting to irrational noise traders to generate arbitrage opportunities, in our model these arise endogenously in equilibrium, following from the presence of heterogeneous rational investors subject to investment restrictions (such as leverage constraints or restrictions on short sales).[2] Irrespective of whether or not noise traders are present in actual financial markets, our viewpoint is that there may exist different categories of rational market participants (e.g., long term pension funds and short term hedge funds) whose interaction is sufficiently rich in implications. Our arbitrageurs are specialized traders who only take on riskless, costless arbitrage positions. Our goal is to study

[1] For example, in the case of 3Com and Palm in March 2000, the market value of 3Com, net of its holdings of Palm, became negative on Palm's first day of trading.

[2] Recent empirical and theoretical work (e.g., Duffie, Garleanu and Pedersen (2002), Ofek, Richardson and Whitelaw (2004)) points to impediments to short-selling comparable to our investment restrictions as a possible cause for the presence of arbitrage opportunities. The restrictions on short-selling in Duffie, Garleanu and Pedersen (2002) can be thought of as stochastic constraints that could be incorporated into our work without considerably affecting our main message. Lamont and Thaler (2003) emphasize the role of short sale constraints, a part of the restrictions we assume, in making mispricings possible.

the equilibrium impact of the presence of these arbitrageurs. More specifically, we consider an economy with two heterogeneous, rational, risk-taking investors and examine how the presence of an arbitrageur affects the way in which risk is traded between the investors, as well as the ensuing equilibrium prices and consumption allocations. The investors could represent real-life retail investors or mutual funds, while the arbitrageur in our model could be an arbitrage desk at an investment bank or a hedge fund specializing in arbitrage strategies. Our setup is comparable to the main workhorse models in asset pricing (Lucas (1978), Cox, Ingersoll and Ross (1985)), which provides the added benefit of having well-understood benchmarks for our results.

Since our main message is not dependent on the particular modeling setup employed, we first convey our insights in the simplest possible framework: a single period economy with a finite set of possible states of nature in which elementary securities (paying-off in a single state) are traded, and redundant securities are available in at least one state. In this setting, "mispricing" means that securities with identical payoffs (here, elementary securities paying-off in the same state) trade at different prices, which generates a riskless arbitrage opportunity. Without making any particular assumptions on the investors, save the fact that their holdings of risky securities are constrained, we show that the presence of an arbitrageur makes it possible for the investors to trade more with each other and partially circumvent the investment restrictions.

The intuition is as follows. Suppose that an investor desires a much higher payoff than the other. He will then purchase securities from the latter investor. If the discrepancy across investors is large enough, the investment restrictions will kick in and prevent investors from attaining their desired payoffs. At this point, a role for the arbitrageur comes into play: he purchases securities from the investor who desires the lower payoff, converts them into a position in the other, redundant security, which is then resold (at a higher price) to the other investor. The larger the arbitrageur's trades, the more he relieves the effects of the constraints. Thus, even though he does not take on any risk, the arbitrageur facilitates risk-sharing amongst investors. The difference between the arbitrageur's "ask" and "bid" prices is the mispricing and generates his arbitrage profit.

To further explore the implications of the arbitrageur's presence, for the remainder of the paper we consider a richer continuous time setup, where investment opportunities include: a risky stock; a riskless bond; and a risky "derivative" financial security that is perfectly correlated with the stock. Real-life examples that are particularly close to the derivative in our model include a stock index futures contract, or a total return swap on a stock index. To make our model more concrete, the latter example is developed in some length; but our model could accommodate any redundant, zero-net supply risky security, and the actual nature of the derivative does not significantly affect our results. For simplicity, the derivative is assumed to have the same exposure to risk (as measured by the volatility) as the stock. Then, "mispricing" refers to any discrepancy in the mean returns offered by the risky investments (stock and derivative). Because these are redundant, such a pricing inconsistency generates a riskless arbitrage opportunity, exploited by short-selling the risky security with the lower mean return and investing the proceeds in the other, generating a profit proportional to the difference in mean returns. For tractability, our investors are assumed to have logarithmic utility preferences. Moreover, the investors face investment restrictions: a leverage constraint that prevents them from taking very large positions in the stock; and a restriction that limits short positions in the derivative. These restrictions may represent margin regulations imposed by an exchange or an intermediary, as well as collateral requirements imposed by a counterparty. To generate heterogeneity and trade, we assume that the investors disagree on the mean stock return. Then, the more optimistic investor desires a higher exposure to risk and so buys risky assets from the more pessimistic investor.

Mispricing is shown to occur when the disagreement across investors is pronounced enough. Then, absent the mispricing the more optimistic agent, who is prevented from investing more in the stock by his leverage constraint, would demand more derivative than what the pessimist can sell (without violating his short-sale limit on the derivative). The mispricing makes the derivative less attractively priced and thus reduces the optimist's demand to a level compatible with the pessimist's constraints, allowing markets to clear.

Our intuition on the arbitrageur's role from the single-period example extends readily to the continuous time case. When the disagreement is large enough for mispricing to occur, the binding investment restrictions prevent the investors from trading as much as they would like. To take advantage of the arbitrage opportunity that is available under mispricing, the arbitrageur buys stock (from the pessimist) and sells derivatives (to the optimist). This allows the pessimist to decrease his exposure to risk and the optimist to increase his. This helps both investors to achieve exposures to risk closer to those they would optimally demand in the absence of investment restrictions. Thus, by taking advantage of the arbitrage opportunity, the arbitrageur allows the investors to trade more risk, effectively relieving their investment restrictions. The greater the size of his trades, the greater this effect; as a result, the size of the mispricing is decreasing in the arbitrageur's position. We further demonstrate that the arbitrageur plays a specific role, different from that of an additional investor: while the arbitrageur *always* improves risk-sharing among investors, a third investor could either improve or impair it, depending on his beliefs.

For the bulk of the analysis, we assume that the arbitrageur behaves competitively and faces a prespecified position limit, which may be linked to a bank's capital requirements or simply represent an internal position limit. Then, under mispricing he always takes on the largest arbitrage position allowed by his position limit. Hence, the size of his position is exogenous. To be able to endogenize the arbitrageur's position size, we provide some extensions of our basic model. First, we consider the case where the arbitrageur is non-competitive, in that he takes into account the impact of his trades on equilibrium prices so as to maximize his profit. In addition to generating much richer arbitrageur behavior, the non-competitive assumption may be more realistic in the case of arbitrageurs that, in real-life, are typically large, few in number, specialized and sophisticated. Due to his decreasing marginal profit, the non-competitive arbitrageur finds it optimal to limit the size of his trades. This makes the occurrence of mispricing more likely, because the non-competitive arbitrageur will never trade enough so as to fully "arbitrage away" the mispricing, making it disappear (and driving his profits to zero). Our main intuition on the role of the arbitrageur, however, is still valid. But the extent to which the arbitrageur alleviates the investors' constraints may be smaller than in the competitive case.

Finally, we consider a case where the (competitive) arbitrageur is endowed with capital and subject to margin requirements that limit the size of his position in proportion to the total amount of his capital. His capital is owned and traded by the investors. This variation appears more realistic than our basic model, and is also richer in implications. Because the investors may now invest in arbitrage capital, in equilibrium the return on arbitrage must be consistent with other investment opportunities; this allows us to explicitly solve for the unique amount of arbitrage capital that is consistent with equilibrium. Our approach is consistent with recent empirical work on the profitability of arbitrage activity (e.g., Mitchell, Pulvino and Stafford (2002)), which suggests that market imperfections severely limit the return on arbitrage, and may drive it to a level that does not dominate other investment opportunities.

The setup of this paper builds on the work of Detemple and Murthy (1994, 1997), who solve for equilibrium in the presence of heterogeneous agents, with and without portfolio constraints, but not with redundant derivative securities. Our framework more closely follows the production economy of Detemple and Murthy (1994), in which the technology pins down the stock price dynamics exogenously, and much of the action is picked up by the bond interest rate. This, along with logarithmic investor preferences, leads to considerable tractability and closed-form solutions in our analysis. Basak and Croitoru (2000) go one step further by endogenizing the arbitrage opportunity and show that in the presence of heterogeneous constrained investors such as ours, the mispricing and arbitrage opportunity will be present under a broad range of circumstances. None of these papers, however, includes a specialized arbitrageur.

To our knowledge, little work exists that examines the impact of specialized arbitrage in an equilibrium where all agents are rational. Related to our work is Gromb and Vayanos (2002). These authors' goals are somewhat different from ours, in that they focus on the issue of a competitive arbitrageur's welfare impact. In order to be able to do this, they examine a different form of constraints, that of a segmented market for the risky assets (where different investors hold different risky assets). Then, without the arbitrageur there is no trade, and the presence of the arbitrageur allows trade and a Pareto improvement to take place (albeit not necessarily in a Pareto optimal way). This is consistent with the arbitrageur's role in our model, which allows more trade to take place between investors. We would not expect to obtain such a clear welfare result, however, since in our model there is trade even in the absence of an arbitrageur and so our benchmark is much more complex. Despite improved risk-sharing, the overall welfare effect of our arbitrageur is ambiguous, because his presence may move prices in a way that can be unfavorable to the investors (Section 4). Our model is also closer to the spirit of standard asset pricing models (and thus easier to relate to observable economic variables or well-understood benchmarks). Loewenstein and Willard (2000) provide an equilibrium where an arbitrageur (interpreted as a hedge fund) plays a similar role when investors are subject to not portfolio constraints, but to an uncertain timing for their consumption. Unlike in our model, arbitrage trades in Loewenstein and Willard are long-lived.

Somewhat related are works, including Brennan and Schwartz (1990) and Liu and Longstaff (2004), which study a constrained arbitrageur's optimal policy, given that the arbitrage opportunity is present. Liu and Longstaff demonstrate that an arbitrageur may optimally underinvest in an arbitrage opportunity, known to vanish at some future time, by taking a smaller position than margin requirement would allow. Finally, our analysis complements an important and growing strand of the finance literature, that is often known as the study of "limits on arbitrage." This approach, inaugurated by De Long, Shleifer, Summers and Waldmann (1990), is thoroughly surveyed in Shleifer (2000). In these papers, arbitrage opportunities are generated by the presence of irrational noise traders, and are allowed to subsist in equilibrium due to various market imperfections (e.g., portfolio constraints, transaction costs, short-termism, model risk). Papers in this area that are particularly related to our work include Xiong (2000), who studies the impact of arbitrageurs (called "convergence traders") on volatility. Kyle and Xiong (2001) extend his model to the study of contagion effects between the markets for two risky assets, one of which only is subject to noise trader risk. Attari and Mello (2002) focus on the effect of trading constraints on the arbitrageur's impact and conclude (as we do) that these may have radical effects. Unlike the bulk of the literature, they assume that arbitrageurs take into account how their own trades affect prices, as we do in part of our analysis.

Our work also complements the growing literature on asset prices in the presence of heterogeneous beliefs and short-sale restrictions, including Harrison and Kreps (1978), Detemple and Murthy (1997) and Scheinkman and Xiong (2003). In these models, an asset price may exceed its fundamental value (its valuation based on investors' own marginal rate of substitution) by a quantity reflecting the investors' anticipated speculative gains. There, investors are willing to pay more than the fundamental value of an asset since ownership of the asset gives them the right to resell the asset to another investor at an even higher price in the future. Some of these models' features are also present in our economy, which contains a short-sale constrained stock and heterogeneous beliefs (but also a derivative security and an arbitrageur, not present in these papers). In the presence of the redundant derivative and investment restrictions, our setting additionally leads to riskless arbitrage opportunities, enabling a study of the role of an arbitrageur, which is our primary focus.

The rest of the paper is organized as follows. Section 2 provides our single period, finite state example on the role of the arbitrageur. Section 3 introduces our continuous time setup, and Section 4 describes the equilibrium with a competitive arbitrageur. Section 5 and 6 are devoted to extensions to the case of a non-competitive arbitrageur, and that of an arbitrageur endowed with capital and subject to margin requirements, respectively. Section 7 concludes, and the Appendix provides all proofs.

2. The Role of the Arbitrageur: an Example

To convey our main intuition on the role of the arbitrageur and show that our main point does not depend on the modeling setup employed, we start with an example in a simple, one-period framework. Portfolio decisions are made at time 0 and securities pay-off at time 1. We first consider an economy with two heterogeneous investors $i = 1, 2$ and no arbitrageur. Consider two zero-net supply securities, S and P, with time-zero prices also denoted by S and P, and with identical payoffs of one unit if state ω occurs, zero otherwise. The only difference between S and P is how trading in S and P is constrained. Denoting by α_j^i investor i's investment in security j (where $j = S, P$), measured in number of shares, we assume that both investors' holdings are subject to the following position limits:

$$\alpha_S^i \leq \beta, \qquad \alpha_P^i \geq -\gamma, \qquad i = 1, 2. \tag{2.1}$$

For simplicity, we assume that S and P are the only two available securities paying-off in state ω. Then, investor i's payoff if state ω occurs is given by $A^i = \alpha_S^i + \alpha_P^i$. The constraints in (2.1) alone do not place any restrictions on the choice of A^i, because one could always use S to take on unlimited short positions and P to take on unlimited long positions. Nonetheless, only a limited set of payoffs are available to the investors in equilibrium: no investor can go long more than γ shares of P, because the other investor cannot provide a counterparty without violating his constraint on short-sales of P. Hence, no investor can receive a state ω-payoff greater than $\gamma + \beta$. A similar argument shows that no investor can take on an unlimited short position, because the other investor cannot provide the necessary counterparty. Thus, in equilibrium, both investors' state ω-payoffs must satisfy the following constraint:

$$-\beta - \gamma \leq A^i \leq \beta + \gamma, \qquad i = 1, 2. \tag{2.2}$$

In short, investors' position limits prevent them from providing a counterparty to each other, and this limits the trades that are possible in equilibrium.

We now examine how the feasible trades (such that (2.2) holds) are affected in the presence of a third agent, an *arbitrageur*. The arbitrageur is assumed to only take on riskless arbitrage positions to take advantage of any price difference between S and P. Whenever $P \neq S$, securities having identical pay-offs trade at different prices, hence a riskless arbitrage opportunity. For example, when $P > S$, an arbitrage position consisting of going long one share of S and short-selling one share of P provides a time 0-profit equal to $P - S > 0$ without any future cash-flows, and hence no risk. Basak and Croitoru (2000) show how such price discrepancies arise in equilibrium, in the presence of constrained heterogeneous investors. In our subsequent analysis, we provide equilibria in which such mispricings also arise endogenously. For now, we take the presence of mispricing $(P > S)$[3] as given and address the issue of the economic role of an arbitrageur who

[3]It is straightforward to check that, with the constraints that we assume for the "normal" investors 1 and 2, the other direction of mispricing $(S > P)$ cannot arise in equilibrium, as agents who then add an unbounded arbitrage position to their portfolio and so there would be no solution to their optimization problems.

exploits the mispricing. We take the size of the arbitrageur's position ($\alpha_S^3 > 0$ shares of S, $\alpha_P^3 = -\alpha_S^3$ shares of P) as given.[4] Then, an investor who desires a large short position can short as many as $\beta + \alpha_S^3$ shares of S (instead of only β if the arbitrageur were not present), because the arbitrageur provides an additional counterparty over and above the other investor. A similar argument applies to large long positions, so that the constraint on investors' state-ω payoffs (the analogue of (2.2)) is now:

$$-\beta - \gamma - \alpha_S^3 \leq A^i \leq \beta + \gamma + \alpha_S^3, \qquad i = 1, 2. \tag{2.3}$$

Equation (2.3) reveals that the presence of the arbitrageur makes it possible for investors to trade more. To see the intuition for this, take the example of investor 1 desiring to buy a large payoff from investor 2 (possibly because he is more optimistic, or less risk-averse). Without an arbitrageur, once he has purchased β shares of S and γ shares of P, trading between the two investors becomes "stuck" no matter how large the heterogeneity between them. The arbitrageur makes it possible for investor 1 to buy an extra α_S^3 shares of P from him. This position is converted by the arbitrageur into a position in S, that is then shorted by investor 2, who is happy to do so because the constraints were preventing him from shorting more. The net result is that the arbitrageur has made it possible for the two investors to trade more risk with each other, even though he does not take on any risk of his own. The arbitrageur plays the role of a financial intermediary. The larger his position, the more he alleviates the constraints.

It should be pointed that the role of the arbitrageur is specific: adding a third "normal" investor (who can take on risk), instead of an arbitrageur, to the economy can either worsen or relieve the constraint in (2.2). In the presence of mispricing, the third investor will always bind on one of his position limits (2.1); which one he binds on depends on his desired cash-flow. If this third investor (indexed by $3*$) binds on his upper constraint on S, the constraint on investor $i \in 1, 2$'s state ω-cash-flow becomes

$$-2\beta - \gamma \leq A^i \leq \beta + \gamma - \alpha_P^{3*}.$$

While the lower bound is lower than in (2.3), showing an alleviation of the lower constraint, the upper bound can be either lower or higher, and the constraint accordingly either more or less stringent. A similar point can be made when agent $3*$ binds on his constraint on P. The effect of a third normal investor is thus ambiguous, highlighting the specific role of an arbitrageur, who *always* improves risk-sharing.

While this section deals with a simplistic economic setup, in Section 4 we show how the exact same points can be made within a much more general, continuous time modeling framework. It should also be noted that the analysis in this section is robust to changes in a large number of assumptions: the preferences, beliefs and endowments of investors 1 and 2, the type of portfolio

[4] In the absence of any constraint, an arbitrageur would optimally choose to hold an infinitely large arbitrage position and so equilibrium would be impossible. To avoid this, later sections either assume that the arbitrageur is constrained or that he is non-price-taking.

constraints (position limits or constraints on portfolio weights; whether they are one- or two-sided, homogeneous or heterogeneous across agents), the net supplies of the securities (meaning that our analysis is also equally valid in production and exchange economies). All that is really needed is the presence of redundant securities, portfolio constraints and heterogeneity within investors.

3. The Economic Setting

This section describes the basic economic setting, which is a variation on the Cox, Ingersoll and Ross (1985) production economy. The economy is populated by two *investors*, who trade so as to maximize the cumulated expected utility of consumption in a standard fashion, and an *arbitrageur*, who is constrained to hold only riskless arbitrage positions. The description of the arbitrageur is relegated to later sections, as assumptions made thereon will be different in each section.

3.1. Investment Opportunities

We consider a finite-horizon production economy. The production opportunities are represented by a single stochastic linear technology whose only input is the consumption good (the numeraire). The instantaneous return on the technology is given by

$$\frac{dS(t)}{S(t)} = \mu_S dt + \sigma dw(t),$$

where μ_S, σ (with $\sigma > 0$) represent the constant mean return and volatility on the technology, S is the amount of good invested and w is a one-dimensional Brownian motion. For convenience, we often refer to the technology as the "stock" S.

The two investors observe the return of the stock, but have incomplete (but symmetric) information on its dynamics. They know σ (from the stock return's quadratic variation), but must estimate μ_S via its conditional expectation, rationally updating their beliefs in a Bayesian fashion with heterogeneous prior beliefs. We denote by $\mu_S^i(t)$ the conditional estimate of μ_S by investor $i = 1, 2$ at time t, given his prior beliefs and observation of the stock's realized return, and by $\bar{\mu}(t) = \left(\mu_S^1(t) - \mu_S^2(t)\right)/\sigma$ the normalized difference in the two investors' estimates. The process $\bar{\mu}$ represents investors' disagreement on their relative "optimism/pessimism" about the stock mean return. The process $\bar{\mu}$ is given exogenously, since all learning about μ_S comes from observing the (exogenous) realized return on the stock. For simplicity, we shall assume that

$\bar{\mu} > 0$, and call investor 1 the "optimistic" investor, and conversely.[5] Finally, we let

$$w^i(t) = \frac{1}{\sigma}\left[\int_0^t \frac{dS(s)}{S(s)} - \int_0^t \mu_S^i(s)ds\right]$$

denote investor i's estimate of the Brownian motion w. By Girsanov's theorem, w^i is a standard Brownian motion under investor i's beliefs and investor i's perceived stock return dynamics are given by

$$\frac{dS(t)}{S(t)} = \mu_S^i(t)dt + \sigma dw^i(t), \qquad i = 1, 2.$$

Though not a central feature of our model, heterogeneity in beliefs is employed to generate trade between the logarithmic investors.

In addition to the stock, investors may invest in two zero net supply, non-dividend paying securities. One is an instantaneous riskless bond with return

$$\frac{dB(t)}{B(t)} = r(t)dt,$$

and the other a risky "derivative" with perceived return process

$$\frac{dP(t)}{P(t)} = \mu_P^i(t)dt + \sigma dw^i(t), \qquad i = 1, 2.$$

Similarly to the stock, investors observe the derivative return process but not its dynamics coefficients, and hence draw their own inferences about the derivative mean return μ_P. Since the derivative pays no dividends, we take its volatility parameter σ (assumed to equal stock volatility for simplicity) to define the derivative contract P, as is standard in the literature (e.g., Karatzas and Shreve (1998)).[6] The interest rate r and derivative mean return μ_P^i, on the other hand, are determined endogenously in equilibrium. Any zero-net supply security is an example of this derivative. Real-life examples that are particularly close in spirit to our idealized derivative include a stock index futures contract, or a total return swap on a stock index, which we take as a leading example in our subsequent discussion. In the case of the derivative being an overnight total return swap on the stock, a long investor receives the total stock return and pays an instantaneously riskless "swap rate" r_P (possibly different from r), with each unit yielding between t and $t + dt$

$$\frac{dS(t)}{S(t)} - r_P(t)dt = \left[\mu_S^i(t) - r_P(t)\right]dt + \sigma dw^i(t), \qquad i = 1, 2.$$

[5]Consider the well-known Gaussian example, in which investor i's prior belief is normally distributed with mean $\mu^i(0)$ and variance $v(0)$. Then, investors' estimates have dynamics $d\mu_S^i(t) = (v(t)/\sigma)dw^i(t)$, where $v(t) = v(0)\sigma^2/(v(0)+\sigma^2)$, implying $d\bar{\mu}(t) = -(v(t)/\sigma)\bar{\mu}(t)dt$, or $\bar{\mu}(t) = \bar{\mu}(0)[\sigma^2/(v(0)t+\sigma^2)]^\sigma$. Assuming $\bar{\mu}(0) > 0$ implies $\bar{\mu}(t) > 0$, $\forall t$. Our subsequent analysis goes through for $\bar{\mu} < 0$, with the only (minor) difference being the presence of another mispricing case arising in equilibrium (Sections 4-6), mirroring the mispricing case arising for $\bar{\mu} > 0$.

[6]The generalization to the case of the derivative having a different volatility than the stock is straightforward and is considered in an earlier version of this manuscript (Basak and Croitoru (2003)). Apart from increased notational complexity, all our main results here are robust to such a generalization. Moreover, defining the derivative by a dividend process would not entail major changes to our analysis. All of the expressions in the paper would remain valid but the derivative volatility σ would be endogenously determined, via a present value formula. Our main points and intuition would be unaffected.

All of our results, derived for the general case of a derivative with mean return μ_P^i, are equally valid in the case of the total return swap, if we take: $r_P(t) = r(t) + \mu_S^i(t) - \mu_P^i(t)$.

3.2. Investors' Optimization Problems

With all the uncertainty generated by a one-dimensional Brownian motion, the stock and derivative returns are perfectly correlated, and so the derivative is redundant. However, restrictions on investors' risky investments generate an economic role for the derivative. Letting $\theta^i \equiv (\theta_B^i, \theta_S^i, \theta_P^i)$, with $\theta_B^i(t) = W^i(t) - \theta_S^i(t) - \theta_P^i(t)$, denote the amounts of investor i's wealth, W^i, invested in B, S and P respectively, we assume that, in addition to short sales being prohibited, stock investments face leverage limits, and that investments in the derivative have limited short-selling restrictions at all times:

$$0 \leq \theta_S^i(t) \leq \beta W^i(t), \qquad \theta_P^i(t) \geq -\gamma W^i(t), \tag{3.1}$$

where $\beta \geq 1$, $\gamma > 0$. For retail investors, the stock investment restriction may be due to margin requirements, while for mutual funds they may represent leverage restrictions. If the derivative security is an over-the-counter contract, such as a total return swap, the derivative investment constraint (3.1) may be due to a collateral requirement imposed by a counterparty. If the derivative is an exchange-traded contract, such as a futures, then the restriction (3.1) may represent margin regulations imposed by the exchange. The one-sided restriction on holdings of the derivative is for expositional simplicity, with the two-sided case being a straightforward extension without affecting our main conclusions.[7]

The risky assets are exposed to the same amount of risk ($\sigma dw^i(t)$), and so it is natural to expect them to be priced accordingly and offer identical mean returns. In fact, if this not the case, there is an arbitrage opportunity available, exploited by short-selling the asset with the lower mean return, and investing the proceeds in the other; this generates a riskless arbitrage profit proportional to the difference in mean returns. This motivates our definition of *mispricing* between the stock and derivative under an investor i's beliefs as the difference between their mean returns:

$$\Delta_{S,P}^i(t) \equiv \mu_S^i(t) - \mu_P^i(t).$$

We note that the observed return agreement across investors enforces agreement on the mispricing: $\Delta_{S,P}^1(t) = \Delta_{S,P}^2(t) \equiv \Delta_{S,P}(t)$. In the presence of investment restrictions, $\Delta_{S,P}$ need not be zero. If $\Delta_{S,P}(t) > 0$, the stock has a higher expected return and we say that it is "cheap" relative to the derivative, and conversely, if $\Delta_{S,P}(t) < 0$, the stock is more "expensive" (and the derivative is cheaper). For example, if the derivative is the overnight total return swap, this mispricing is given by any deviation of the swap rate from the interest rate since, whenever $r_P \neq r$, an arbitrage opportunity is available. If $r_P > r$, the swap is more expensive than the

[7]In particular, suppose the holdings of the derivative was additionally restricted to satisfy a leverage restriction: $\theta_P^i(t) \leq \overline{\gamma} W^i(t)$, $\overline{\gamma} > 0$. Then, there would be additional scenarios in the investors' optimal holdings of Proposition 3.1, leading to additional equilibrium cases, as discussed in Footnote 8 of Section 4.2.

stock, and a strategy consisting of buying the stock, financing the purchase by riskless borrowing and taking a short position in the swap yields an instantaneous arbitrage profit of \$$(r_P - r)$ per dollar in the stock, and vice versa if $r_P < r$.

The investor i is endowed with initial wealth $W^i(0)$. He then chooses a consumption policy c^i and investment strategy θ^i so as to maximize his expected lifetime logarithmic utility subject to the dynamic budget constraint and position limits, that is to solve the problem:

$$\begin{aligned} & \max_{c^i,\theta^i} E^i\left[\int_0^T \log\left(c^i(t)\right) dt\right] \\ \text{s. t.} \quad & dW^i(t) = \left[W^i(t)r(t) - c^i(t)\right] dt + \left\{\theta^i_S(t)\left[\mu^i_S(t) - r(t)\right] + \theta^i_P(t)\left[\mu^i_P(t) - r(t)\right]\right\} dt \\ & \qquad + \left[\theta^i_S(t) + \theta^i_P(t)\right]\sigma dw^i(t) \\ \text{and} \quad & (3.1). \end{aligned}$$

This differs from the standard frictionless investor's dynamic optimization problem due to the investment restrictions, potential redundancy and mispricing between the risky investment opportunities S and P. The solution of the investor i's problem is reported in Proposition 3.1. Those cases where P is cheap are ignored, as they cannot arise in equilibrium.

Proposition 3.1. *Investor i's optimal consumption and investments in the stock S and derivative P are given by:*

$$c^i(t) = \frac{W^i(t)}{T-t};$$

when there is no mispricing, i.e. $\Delta_{S,P}(t) = 0$,

$$\frac{\theta^i_S(t)}{W^i(t)} = \begin{cases} \in [0,\beta], \\ \\ 0, \end{cases} \qquad \frac{\theta^i_P(t)}{W^i(t)} = \begin{cases} \geq -\gamma\sigma, & \text{s.t. } \frac{\theta^i_S(t)}{W^i(t)} + \frac{\theta^i_P(t)}{W^i(t)} = \frac{\mu^i_S(t)-r(t)}{\sigma^2}, \quad (a) \\ & \text{if } \mu^i_S(t) \geq r(t) - \gamma\sigma^2, \\ -\gamma & \text{if } \mu^i_S(t) < r(t) - \gamma\sigma^2, \quad (b) \end{cases}$$

and when there is mispricing with the stock being cheaper than the derivative, i.e. $\Delta_{S,P}(t) > 0$,

$$\frac{\theta^i_S(t)}{W^i(t)} = \begin{cases} \beta & \text{if } \mu^i_S(t) > \mu^i_P(t) \geq r(t) - (\gamma-\beta)\sigma^2, & (c) \\ \beta & \text{if } \mu^i_S(t) > r(t) - (\gamma-\beta)\sigma^2 > \mu^i_P(t), & (d) \\ \frac{\mu^i_S(t)-r(t)}{\sigma^2} + \gamma & \text{if } r(t) - (\gamma-\beta)\sigma^2 \geq \mu^i_S(t) \geq r(t) - \gamma\sigma^2, & (e) \\ 0 & \text{if } r(t) - \gamma\sigma^2 > \mu^i_S(t), & (f) \end{cases}$$

$$\frac{\theta^i_P(t)}{W^i(t)} = \begin{cases} \frac{\mu^i_P(t)-r(t)}{\sigma^2} - \beta & \text{if (c)}, \\ -\gamma & \text{if (d)}, \\ -\gamma & \text{if (e)}, \\ -\gamma & \text{if (f)}. \end{cases}$$

An investor can be in one of six possible scenarios, (a)-(f), depending on the economic environment and whether there is mispricing (cases (c)-(f)) or not (cases (a)-(b)). When there is no mispricing and the perceived mean return is low enough, the investor desires the lowest possible exposure to risk, and so binds on his lower constraint on both risky assets (case (b)). Otherwise, he is indifferent between all feasible investments in the two redundant risky assets that lead to his optimal risk exposure (case (a)) (that is, investments ensuring the volatility of his portfolio to equate to the market price of risk).

Under mispricing (cases (c)-(f)), a riskless, costless, profitable arbitrage opportunity is available, and so the investor indulges in it to the greatest extent possible as allowed by the position limits. Effectively, the investor is adding an arbitrage position, consisting of a long position in the cheap stock and a short position in the expensive derivative, to his portfolio. This implies that he is always binding on at least one of the investment restrictions in (3.1), and thus uniquely determines the allocation between the stock and the derivative. When the investor is relatively optimistic about the mean return of the stock and the derivative is not too expensive (case (c)), he desires a high risk exposure, wishing to go long in both risky assets, and so ends up binding on the leverage constraint on S. Otherwise, the derivative being more expensive ($\Delta_{S,P} > 0$ or, in the case of a total return swap, higher r_P) by a large amount drives him to his short sale limit on P; when additionally pessimistic about the stock, he desires a low risk exposure and so does not invest in the stock (case (f)), otherwise he goes long in the stock (cases (d)-(e)). As will be demonstrated in Section 4.2, mispricing cases (c) and (e) may occur in equilibrium (with or without an arbitrageur). Cases (d) and (f) do not.

4. Equilibrium with a Competitive Arbitrageur

4.1. The Arbitrageur

We assume that, in addition to the two investors $i = 1, 2$, the economy is populated by an *arbitrageur* indexed by $i = 3$. The arbitrageur is assumed to take on only riskless, costless arbitrage positions, and to maximize his expected cumulated profits under a limit on the size of his positions. This natural formulation of the arbitrageur is somewhat simplistic, but it does capture the most important distinguishing characteristics of a specialized arbitrageur, such as an arbitrage desk in an investment bank, or a hedge fund specializing in arbitrage strategies. Denoting the arbitrageur's (dollar) investments by $\theta^3 \equiv (\theta^3_B, \theta^3_S, \theta^3_P)$, his problem can be expressed as follows:

$$\max_{\theta^3} E^3 \left[\int_0^T \Psi^3(t) dt \right]$$

$$\text{s.t.} \qquad \Psi^3(t)dt = \left\{ \theta^3_S(t)\mu^3_S(t) + \theta^3_P(t)\mu^3_P(t) + \theta^3_B(t)r(t) \right\} dt \\ + [\theta^3_S(t) + \theta^3_P(t)]\sigma dw^3(t),$$

$$\theta^3_S(t) + \theta^3_P(t) + \theta^3_B(t) = 0, \tag{4.1}$$

$$\left(\theta^3_S(t) + \theta^3_P(t) \right) \sigma = 0, \tag{4.2}$$

$$0 \leq \theta_S^3(t) \leq M. \tag{4.3}$$

An equivalent interpretation of the no-risk condition (4.2) is that the arbitrageur's position is uncorrelated with the stock return, as a hedge fund's return ideally would. The arbitrageur may be a trader of an investment bank or a hedge fund. For an investment bank, the investment restriction (4.3) imposed on the arbitrage trader may be linked to the capital requirements faced by the bank, with a credit limit controlled by the Federal Reserve. In the case of a hedge fund, the investment restriction imposed on the arbitrageur may simply represent an internal position limit.

Conditions (4.1)-(4.2) imply:

$$\theta_P^3(t) = -\theta_S^3(t), \qquad \theta_B^3(t) = 0, \tag{4.4}$$

so that the arbitrageur's holdings in all securities are pinned down by his stock investment. Therefore, the constraint on holdings of the stock (4.3) also limits the arbitrageur's holdings of the other securities.

In the absence of mispricing ($\Delta_{S,P} = 0$), it is straightforward to verify that the arbitrageur is indifferent between all portfolio holdings that satisfy (4.3)-(4.4), since they all provide him with zero cash-flows (and he can do no better than this). In the presence of mispricing ($\Delta_{S,P} > 0$), however, as long as $\theta_S^3 > 0$ the arbitrageur receives a profit that is proportional to θ_S^3. The optimal policy is then to choose the maximal possible value for θ_S^3, namely M. From (4.4), the other holdings are given by $\theta_P^3(t) = -M$ and $\theta_B^3(t) = 0$, and the arbitrageur's instantaneous profit rate by:

$$\Psi^3(t) = \theta_S^3(t)\mu_S^3(t) + \theta_P^3(t)\mu_P^3(t) = M\Delta_{S,P}(t).$$

For example, in the case of the derivative being a total return swap on the stock, the arbitrageur may be purchasing the stock, financing the purchase by borrowing at the riskless rate and simultaneously entering the swap, whereby he pays the stock return and receives the swap rate r_P. In the absence of mispricing, the two interest rates are equal. Under mispricing, the notional swap rate is higher than the interest rate (by $\Delta_{S,P}$), and the difference generates the arbitrageur's profit.

It should be pointed that the (price-taking) arbitrageur's solution (as well as the value of his profits) is independent of his beliefs and preferences. Due to his constraints, (4.1)-(4.2), the arbitrageur consumes his profits as they are made ($c^3 = \Psi^3$) and hold no net wealth, so that he can be assumed to be risk-neutral, without any time-preference, without loss of generality. Any increasing utility function for the arbitrageur would lead to the same behavior as for the risk-neutral arbitrageur, in this Section as well as Sections 5 and 6.

4.2. Analysis of Equilibrium

We now proceed to the description of the equilibrium (defined, in a standard fashion, as the security prices such that agents' optimal policies clear the security markets, given the prices) in our economy. For convenience, quantities of interest will be expressed as a function of the aggregate wealth $W \equiv W^1 + W^2$, and the proportion of aggregate wealth held by investor 1, $\lambda \equiv W^1/W$. Several cases are possible in equilibrium, depending on investor 1 and 2's binding constraints. We denote each one by the optimization case in which each of investors 1 and 2 is. For example, in equilibrium (a,a) each investor is in case a, in equilibrium (a,b) investor 1 is in case a and investor 2 in case b, etc..

Proposition 4.1 reports the possible equilibrium cases, the conditions for their occurrence, and characterizes economic quantities (prices and consumptions) in each case.

Proposition 4.1. *In equilibrium, the investors' optimization cases, equilibrium mispricing and interest rate, and distribution of wealth dynamics are as follows.*

$$\text{When} \qquad \bar{\mu}(t) \leq \frac{\gamma\sigma}{\lambda(t)} + \min\left\{\frac{\sigma}{\lambda(t)}, \frac{\sigma}{1-\lambda(t)}\left[(\beta-1) + \frac{M}{\lambda(t)W(t)}\right]\right\},$$

investors are in (a,a) and

$$\begin{aligned}
\Delta_{S,P}(t) &= 0 \qquad \text{and} \qquad r(t) = \lambda(t)\mu_S^1(t) + (1-\lambda(t))\mu_S^2(t) - \sigma^2, \\
d\lambda(t) &= \lambda(t)(1-\lambda(t))^2\left(\bar{\mu}(t)\right)^2 dt + \lambda(t)(1-\lambda(t))\bar{\mu}(t)dw^1(t).
\end{aligned} \tag{4.5}$$

$$\text{When} \qquad \bar{\mu}(t) > \frac{(\gamma+1)\sigma}{\lambda(t)} \qquad \text{and} \qquad \lambda(t) \geq \frac{1}{\beta}\left(1 - \frac{M}{W(t)}\right),$$

investors are in (a,b) and

$$\begin{aligned}
\Delta_{S,P}(t) &= 0 \qquad \text{and} \qquad r(t) = \mu_S^1(t) - \left[\frac{\gamma+1}{\lambda(t)} - \gamma\right]\sigma^2, \\
d\lambda(t) &= \sigma(\gamma+1)\left[\frac{(1-\lambda(t))^2}{\lambda(t)}dt + (1-\lambda(t))dw^1(t)\right].
\end{aligned}$$

$$\text{When} \quad \bar{\mu}(t) > \frac{\gamma\sigma}{\lambda(t)} + \frac{\sigma}{1-\lambda(t)}\left(\beta - 1 + \frac{M}{\lambda(t)W(t)}\right) \quad \text{and} \quad \lambda(t) < \frac{1}{\beta}\left(1 - \frac{M}{W(t)}\right) \tag{4.6}$$

investors are in (c,e) and

$$\begin{aligned}
\Delta_{S,P}(t) &= \bar{\mu}(t)\sigma - \frac{(\beta-1)\sigma^2}{1-\lambda(t)} - \frac{\gamma\sigma^2}{\lambda(t)} - \frac{M\sigma^2}{\lambda(t)(1-\lambda(t))W(t)} > 0, \\
r(t) &= \mu_S^2(t) - (1-\gamma)\sigma^2 + \left[(\beta-1)\frac{\lambda(t)}{1-\lambda(t)} + \frac{M}{(1-\lambda(t))W(t)}\right]\sigma^2, \\
d\lambda(t) &= \lambda(t)\left\{\Delta_{S,P}(t)\left[\beta - 1 + \frac{M}{W(t)}\right] + \left[(1-\lambda(t))\left(\bar{\mu}(t) - \Delta_{S,P}(t)/\sigma\right)\right]^2\right\}dt \\
&\quad + \lambda(t)(1-\lambda(t))\left(\bar{\mu}(t) - \Delta_{S,P}(t)/\sigma\right)dw^1(t).
\end{aligned} \tag{4.7}$$

In all cases, the aggregate wealth dynamics follow

$$dW(t) = \left[W(t)\left(\mu_S^1(t) - \frac{1}{T-t}\right) - M\Delta_{S,P}(t)\right]dt + W(t)\sigma dw^1(t)$$

and investors 1 and 2's consumption and the arbitrageur's profit are given by:

$$c^1(t) = \frac{\lambda(t)W(t)}{T-t}, \qquad c^2(t) = \frac{(1-\lambda(t))W(t)}{T-t}, \qquad \Psi^3(t) = M\Delta_{S,P}(t).$$

Three cases are possible in equilibrium: (a,a) and (a,b) where there is no mispricing and the arbitrageur makes no profits, and (c,e) where mispricing occurs. We now describe how, as the divergence in beliefs across investors grows, the equilibrium may move into region (c,e), where the arbitrageur is active and affects the equilibrium.

In our economy, investors trade due to their disagreement on the mean stock return, with the optimist buying stock from the pessimist. If the disagreement is not too large (equilibrium case (a,a)), then the optimist puts more than 100% of his wealth into stock and the pessimist puts some fraction between 0% and 100% into stock. Since the technology return pins down the stock price, to clear the market, the interest rate is set at a value depending on the average degree of optimism across investors (as in the framework of Detemple and Murthy (1994)). As the optimist becomes even more optimistic, he buys more and more stock from the pessimist, until either the optimist hits his leverage constraint or the pessimist hits his no-short sale constraint. Which constraint is hit first depends on the distribution of wealth between the two investors, with the poorer one having a tendency to hit his constraint first. Let us suppose that the pessimist hits his constraint first. Then, as the optimist becomes even more optimistic beyond this point, he buys derivatives from the pessimist, even though the pessimist has more stock he could sell to the optimist, if the optimist were not leverage-constrained. As the optimist becomes even more optimistic, he eventually buys so many derivatives from the pessimist that the pessimist hits his constraint on short derivative positions.

At this point, additional trades between the optimist and the pessimist would be impossible, and risk-sharing would be "stuck", were it not for the presence of the arbitrageur. As the optimist becomes even more optimistic, the arbitrageur buys stock from the pessimist (who still has some left because he did not hit his no-short-sales constraint before the optimist hit his leverage constraint) and sells derivatives to the optimist. He is converting stock positions into derivative positions. As the optimist continues to become more optimistic, he buys more and more derivative, until a point is reached where the arbitrageur hits his position limit M or the pessimist sells him all of his stock. Let us assume that the arbitrageur's position limit is hit first, which is when the economy moves from region (a,a) into region (c,e), and the derivative becomes mispriced relative to the stock. As the optimist becomes more optimistic beyond this point, then the arbitrageur cannot sell him more derivative. Instead, the derivative's mean return is reduced, and the optimist limits the amount of derivative he buys because it becomes more expensive.

Returning to the case of the derivative being a total return swap, a positive mispricing means that the notional riskless rate used to price total return swaps becomes higher than the actual riskless rate in the economy. This is the opposite of what happens when there is a "repo squeeze" (and the notional swap rate is less than the market interest rate.) This higher notional interest rate makes the total return swaps more costly to the optimist, and entices him to reduce his holding therein (to what the arbitrageur is allowed by his position limit to sell him). On the other side of the book, however, the arbitrageur has a plentiful supply of stock for sale by the pessimistic investor. Thus, for markets to clear, the riskless rate needs to drop (below the swap rate); this induces the pessimist to reduce the amount of stock he wants to sell to the amount the arbitrageur can buy (M).

If, on the other hand, the pessimistic investor is much wealthier than the optimist, as divergence in beliefs grows the optimist hits his leverage constraint on the stock before the pessimist hits his constraint on the derivative. If this is the case, market clearing does not require that the optimist's demand of the derivative be reduced, because the pessimist is wealthy enough to provide a counterparty and sell him derivatives without hitting his constraint, and so there is no mispricing in equilibrium (case (a,b)).

In short, the arbitrageur is active under high divergence in beliefs across investors (and provided the pessimist is not much wealthier). The arbitrageur buys stock from the pessimist and sells derivatives to the optimist. Under mispricing (equilibrium case (c,e)), the derivative is more expensive than the stock, and the arbitrageur profits from the difference.[8]

We henceforth focus on the analysis of region (c,e), where the arbitrageur has an economic impact and so affects economic quantities. In region (a,a), the equilibrium is as in an unconstrained economy as analyzed by Detemple and Murthy (1994), because no constraints are binding, and in region (a,b), the equilibrium is similar to a constrained economy without a derivative as in Detemple and Murthy (1997). In both of these cases, there is no arbitrage opportunity available. It is straightforward to check that region (c,e) does occur for plausible parameter values: for example, according to (4.6), if $\gamma = 0.5$, $\beta = 1.5$, $\sigma = 0.1$, $M/W(t) = 0.1$, $\lambda(t) = 0.5$, region (c,e) occurs whenever investor 1's estimate of the mean return of the stock exceeds investor 2's estimate by at least 2.4% per year. In our analysis, we focus on the terms reflecting the effects of the arbitrageur's presence and the size of his position (measured by M). All of our analyses and comparisons are valid only for given aggregate wealth (W) and distribution of wealth across investors (λ). For comparison with our economy, we introduce the following benchmarks:

Economy I: no constraints, no derivative, otherwise as our economy;

Economy II: no arbitrageur, otherwise as our economy.

In Economy I, all equilibrium quantities are as in our region (a,a), while in Economy II, they are as in our economy, for the particular case where $M = 0$.

The value of the mispricing can be interpreted as the amount of heterogeneity (in beliefs) that investors cannot trade on, due to the binding constraints. Whenever no constraints are binding and so risk-sharing is not limited, the mispricing is always equal to zero. However, on moving from region (a,a) to region (c,e), the mispricing starts at a value of zero and then increases

[8] If we also assume that the investors face a leverage constraint on investments in the derivative ($\theta_P^i(t) \le \overline{\gamma} W^i(t)$, $\overline{\gamma} > 0$), an additional case may occur in equilibrium where there is no arbitrage opportunity, and hence no role for the arbitrageur. In that case, the leverage constraint prevents the optimist from purchasing more derivatives than what the pessimist can short, and so the equilibrium does not require the derivative to become less attractive to the optimist via the mispricing. The optimist is prevented by the additional leverage constraint from purchasing derivatives from the arbitrageur, who in turn cannot trade. Our analysis of region (c,e) remains unchanged, however, and our main message on the role of the arbitrageur is thus unaffected by this generalization of our basic setup.

linearly with heterogeneity ($\bar{\mu}$) while inter-agents transfers remain stuck due to the constraints. The mispricing is decreased by the presence of the arbitrageur, and decreases further as the size of his position increases; comparing with our benchmarks, we have: $\Delta^{I}_{S,P} = 0 < \Delta_{S,P} < \Delta^{II}_{S,P}$. Keeping in mind our interpretation of the mispricing as the amount of un-traded heterogeneity, this suggests an improvement in risk-sharing due to the arbitrageur's presence.

The value of the interest rate is increased by the presence of the arbitrageur; we have: $r_I > r > r_{II}$. This is also indirect evidence of an improvement in risk-sharing between investors due to the arbitrageur's presence: improved risk-sharing decreases investors' precautionary savings. For the bond market to clear in spite of this, it is necessary for the interest rate to rise, thus becoming closer to its value in an unconstrained economy. Note that, as in the production economy of Detemple and Murthy (1994), since the technology pins down the stock price dynamics exogenously, it is the interest rate that has to adjust for markets to clear, and so it picks up much of the action. The interest rate can be solved for explicitly because investor's demands are myopic and linear in the stock's Sharpe ratio (that is entirely determined by the interest rate).

In short, the impact of the arbitrageur on all price dynamics parameters is to make them closer to their values in an unconstrained economy; the larger his position, the closer the equilibrium to an unconstrained one. The conditions for region (c,e) to occur (and the constraints to bind) are also affected. From (4.6), it is clear that the larger the arbitrageur's position, the less likely region (c,e) is to occur and the constraints are to bind, perturbing risk-sharing between investors. In all cases, the effect of increasing the size of the arbitrageur's position (M) on the expressions is tantamount to an increase in β or γ (i.e., an alleviation of the constraints.) All of this suggests that the arbitrageur alleviates the effect of the portfolio constraints. The overall effect of his presence on the investors' welfare is ambiguous, however, because his trades move prices in a way that can hurt the investors. The arbitrageur's trades reduce the mean stock excess return (above the riskless rate), which hurts both investors, but increase the mean derivative return, benefiting the optimistic investor (who is long) and hurting the pessimist. The increase in the interest rate benefits the pessimist (long in the bond), but hurts the optimist. Thus, in spite of the unambiguous improvement in risk-sharing due to the arbitrageur's trades, the welfare impact of the arbitrageur's presence is ambiguous.[9]

Finally, interpreting the leverage limit parameter β and the limited short-selling parameter γ as policy variables, either set by an exchange or a regulatory body, we may investigate the impact of leverage and short-selling limits, in particular on the arbitrage activity. From the characterization in Proposition 4.1, we see that increasing these limits (γ, β) shrinks the mispricing region, and hence makes mispricing less likely to occur. Furthermore, increasing these limits increases the interest rate, moving it closer to the perfect risk-sharing case, and reduces the magnitude of the mispricing and hence the arbitrageur's profits: overall, there is less arbitrage activity. This is because increasing these limits alleviates the risky investment restrictions on the investors, and

[9]Note that, however, all agents in our model are better off than in a no-trade economy, as in Gromb and Vayanos (2002), because they could choose to not trade.

hence the two investors are better able to share risk (with or without the arbitrageur). This in turn allows a smaller role for the arbitrageur, reducing the mispricing and the arbitrageur's profits.

4.3. The Arbitrageur and Risk-sharing

To clarify the role of the arbitrageur, we quantify the transfers of risk that take place in region (c,e), as well as in our benchmark economies I and II. We will additionally consider the following benchmark case:

Economy III: no securities (the only investment opportunity is the stock).
In Economy III, investors do not trade and invest all of their wealth in the stock ($\theta^i_S = W^i$).

We introduce Φ^i, investor i's measure of *risk exposure*, defined as the proportion of his wealth held in risky investments:

$$\Phi^i(t) \equiv \frac{\theta^i_S(t) + \theta^i_P(t)}{W^i(t)}.$$

In the absence of trade (Economy III), all of agent i's wealth is invested in the stock, and so $\Phi^i_{III} = 1$ for both investors. When trade is possible, heterogeneity in beliefs leads the more optimistic investor to optimally take on a higher risk exposure than in the no-trade case, and the more pessimistic agent to take on a lower risk exposure. Corollary 4.1 reports the investors' equilibrium risk exposures in the absence of binding constraints (benchmark economy I and region (a,a) of our economy) and in the equilibrium case (c,e) of our economy.

Corollary 4.1. *In the absence of binding constraints, equilibrium risk exposures are as follows:*

$$\Phi^1_I(t) = 1 + (1-\lambda(t))\frac{\bar{\mu}(t)}{\sigma}, \qquad \Phi^2_I(t) = 1 - \lambda(t)\frac{\bar{\mu}(t)}{\sigma}.$$

In the (c,e) equilibrium,

$$\Phi^1(t) = \beta + \gamma\frac{1-\lambda(t)}{\lambda(t)} + \frac{M}{\lambda(t)W(t)}, \qquad \Phi^2(t) = 1 - \gamma - (\beta - 1)\,\frac{\lambda(t)}{1-\lambda(t)} - \frac{M}{(1-\lambda(t))W(t)}.$$

In the presence of constraints, once the disagreement ($\bar{\mu}$) is large enough for the constraints to bind, the amount of risk that is traded becomes "stuck": portfolio holdings do not depend on $\bar{\mu}$ any more. It is straightforward to check that, assuming that the conditions for (c,e) to occur in our economy are fulfilled,

$$\Phi^1_{III}(t) < \Phi^1_{II}(t) < \Phi^1(t) < \Phi^1_I(t), \qquad \Phi^2_{III}(t) > \Phi^2_{II}(t) > \Phi^2(t) > \Phi^2_I(t).$$

Thus, the presence of the arbitrageur makes the allocation of risk closer to what it would be without constraints. A natural measure of the amount of risk that is being shared is the difference $\Phi^1 - \Phi^2$. It is equal to zero in the no-trade case (benchmark economy III) and, in the absence of binding constraints (benchmark economy I and region (a,a) in our economy), grows linearly with heterogeneity: it is then equal to $\bar{\mu}/\sigma$. In our equilibrium (c,e),

$$\Phi^1(t) - \Phi^2(t) = \frac{\beta - 1}{1 - \lambda(t)} + \frac{\gamma}{\lambda(t)} + \frac{M}{W(t)\lambda(t)(1 - \lambda(t))}. \tag{4.8}$$

revealing how the amount of risk that is shared between the investors grows with the size of the arbitrageur's position. The mechanism through which the arbitrageur facilitates risk-sharing between the investors is the one already described in Section 2: the arbitrageur effectively acts as a financial intermediary who provides an additional counterparty to the investors. By converting stock positions into derivative positions, he allows the investors to trade more with each other. The analysis in Section 2 can be repeated with only insubstantial changes, provided one replaces investors' state-ω cash-flows (A^i's) in Section 2 with their risk exposures (Φ^i's). In the absence of the arbitrageur (benchmark economy II), investors' risk exposures are:

$$\Phi^1_{II}(t) = \beta + \gamma\frac{1 - \lambda(t)}{\lambda(t)}, \qquad \Phi^2_{II}(t) = 1 - \gamma - (\beta - 1)\,\frac{\lambda(t)}{1 - \lambda(t)}.$$

Even though the optimistic investor 1 would like to increase his risk exposure (to Φ^1_I) by increasing his holding of the derivative (since he is already leverage-constrained on the stock), this is impossible because investor 2 would then, due to his lower constraint on P, not be able to provide a counterparty. By shorting the derivative, the arbitrageur provides an extra counterparty, which allows for an increase in Φ^1 (by $M/\lambda W$) that is proportional to the arbitrageur's position limit M. Similarly, it is impossible for the pessimistic investor 2 to invest as little in the stock as he optimally would, because then not all of the stock would be held and so markets would not clear. By investing in the stock, the arbitrageur makes it possible for the pessimistic investor to hold less of it and so reduce his risk exposure (by $M/[(1-\lambda)W]$). The introduction of the arbitrageur in the economy is equivalent to that of an extra, zero-net supply security (for example a total return swap on the stock), in which both investors face a position limit proportional to M. This security is being purchased from the pessimistic investor by the innovator, the arbitrageur, and resold to the optimistic investor at a higher price (or a higher notional riskless rate, in the case of the swap). The difference (the mispricing) makes up the arbitrageur's profit.

Finally, we note that much of the above analysis is not dependent on the logarithmic utility assumption. The investors' portfolio holdings in region (c,e) follow directly from the binding constraints and the market clearing conditions. Thus, whenever mispricing occurs, these holdings obtain for arbitrary investor preferences or beliefs, and our analysis of the arbitrageur's contribution to improved risk-sharing is equally valid. Only the occurrence of mispricing is needed for this. Basak and Croitoru (2000, Proposition 5) show how and when mispricing may arise in equilibrium under very general assumptions, demonstrating the robustness of our analysis.

4.4. Comparison with an Economy with Three Investors and no Arbitrageur

It is well-known that increasing, ceteris paribus, the number of agents typically improves risk-sharing in an economy.[10] Thus, a natural question arises as to whether adding a third (risk-taking) investor to our economy would improve risk-sharing in the same way the arbitrageur does. Proposition 4.2 presents an example of equilibrium that can arise in this case. We index the third (logarithmic expected utility-maximizer) investor by $i = 3*$ and, accordingly, denote his wealth by W^{3*} and his estimate of μ_S by μ_S^{3*}. (As before, $W \equiv W^1 + W^2$ and $\lambda \equiv W^1/W$.)

Proposition 4.2. *In the presence of a third optimizing agent (3∗), but no arbitrageur, if*

$$\bar{\mu}(t) > \sigma\,(\beta-1)\,\frac{1+\frac{W^{3*}(t)}{W(t)}}{1-\lambda(t)} + \gamma\sigma\frac{1+\frac{W^{3*}(t)}{W(t)}}{\lambda(t)+\frac{W^{3*}(t)}{W(t)}} + \frac{\mu_S^1(t)-\mu_S^{3*}(t)}{\sigma}\frac{W^{3*}(t)}{W(t)},$$

$$-\gamma\sigma\left(1+\frac{W(t)}{W^{3*}(t)}\right) \le \frac{\mu_S^1(t)-\mu_S^{3*}(t)}{\sigma} \le \gamma\sigma\frac{1+\frac{W^{3*}(t)}{W(t)}}{\lambda(t)} \qquad \textit{and}$$

$$\frac{1-\lambda(t)}{\lambda(t)+\frac{W^{3*}(t)}{W(t)}} \ge \beta - 1,$$

then in equilibrium investors are in optimization cases (c,e,c) and investors 1 and 2's risk exposures are

$$\Phi^{1*}(t) = \beta\sigma + \gamma\sigma\frac{1-\lambda(t)}{\lambda(t)+\frac{W^{3}(t)}{W(t)}} + \frac{\mu_S^1(t)-\mu_S^{3*}(t)}{\sigma}\frac{W^{3*}(t)}{\lambda(t)W(t)+W^{3*}(t)},$$

$$\Phi^{2*}(t) = \sigma(1-\gamma) - (\beta-1)\,\sigma\frac{\lambda(t)+\frac{W^{3*}(t)}{W(t)}}{1-\lambda(t)}.$$

Investors' equilibrium risk exposures reveal that a third risk-taking investor affects risk-sharing differently than the arbitrageur does. While $\Phi_I^2 < \Phi^{2*} < \Phi_{II}^2$, implying that the presence of the extra investor $3*$ makes investor 2's risk exposure closer to its value in an unconstrained economy (Φ_I^2), the effect on investor 1 is ambiguous. Depending on investor $3*$'s beliefs, Φ^{1*} can be either higher or lower than if the third investor were not present. Accordingly, the effect of investor $3*$ on the amount of risk shared between investors 1 and 2 is ambiguous. Our measure of risk-sharing is now equal to

$$\Phi^{1*}(t) - \Phi^{2*}(t) = (\beta-1)\,\sigma\frac{1+\frac{W^{3*}(t)}{W(t)}}{1-\lambda(t)} + \gamma\sigma\frac{1+\frac{W^{3*}(t)}{W(t)}}{\lambda(t)+\frac{W^{3*}(t)}{W(t)}} + \frac{\mu_S^1(t)-\mu_S^{3*}(t)}{\sigma}\frac{1}{1+\lambda(t)\frac{W(t)}{W^{3*}(t)}}.$$

Comparing with the value of this measure in the absence of binding constraints (which remains equal to $\bar{\mu}$, whether or not a third investor is present) reveals that, if investor $3*$ is optimistic

[10] For example, in a standard economy the risk aversion of the representative agent A satisfies $1/A = \Sigma_i\left(1/A^i\right)$ and so decreases when the number of agents in the economy increases, ceteris paribus. Hence, so does the market price of risk, $A\sigma_\delta$, where σ_δ is the volatility of aggregate consumption, and the individual consumption volatilities, $A\sigma_\delta/A^i$.

enough (while still allowing for the conditions in the proposition to hold), risk-sharing can be degraded by his presence. An example of plausible parameters for which this happens is as follows: $\gamma = 0.5$, $\beta = 1.5$, $\sigma_S = \sigma_P = 0.1$, $\lambda(t) = 0.5$, $W^{3*}(t)/W(t) = 0.4$, $\mu_S^1(t) = 0.1$, $\mu_S^2(t) = 0.05$, $\mu_S^{3*}(t) = 0.105$. Then, the equilibrium is as described in Proposition 4.2, and the amount of risk shared between investors 1 and 2 is smaller than if the third investor were not present: $\Phi^1 - \Phi^2$ is equal to 0.1761 versus 0.2 (and 0.5 in an unconstrained economy, whether or not a third investor is present).

Thus, our arbitrageur, whose presence *always* improves risk-sharing across investors (as demonstrated by (4.8)), performs a specific economic function.

5. Equilibrium with a Non-competitive Arbitrageur

In this section, we consider another imperfection for the financial markets, in that the arbitrageur has market power therein. This could naturally arise in markets with limited liquidity or in specialized markets with few participants. There is considerable evidence (see, e.g., Attari and Mello (2002)) suggesting that in such markets, large traders "move" prices.

When the arbitrageur has market power in the securities markets, he will take account of the price impact of his trades on the level of mispricing across the risky assets. This consideration will be shown to induce much richer arbitrageur trades than those in the competitive case, in which under mispricing the arbitrageur simply took on the largest position allowed by his investment restriction. The importance of this non-competitive case is further underscored by the fact that trading restrictions on the arbitrageur are no longer needed to bound arbitrage, and hence attain equilibrium with an active arbitrageur, as will be demonstrated in the sequel.

5.1. The Non-competitive Arbitrageur

For tractability, we consider a *myopic* arbitrageur in the sense that he maximizes his *current* profits (as opposed to lifetime profits). This short-sightedness may be interpreted as capturing in reduced form arbitrageurs as being either short-lived or having short-termism, as often observed in practice. This may for example be due to the separation of "brains and resources" that prevails in real-life professional arbitrage (Shleifer (2000), Chapter 4): if he reports to shareholders who are less sophisticated than himself, the arbitrageur may be enticed to choose short-term profits over a long-term strategy that may return less in the short run.

An arbitrageur with market power faces a problem that is different – and more complex – than the competitive arbitrageur of Section 4. The non-competitive arbitrageur observes the investors' demands as functions of prices (as determined in Section 3), and then chooses the size of his own trades so as to maximize his concurrent profits by taking the exact counterparty of the investors' trades (so that markets clear). Of course, he will only be active when there is mispricing and hence positive profits; otherwise his profits are zero. Accordingly, given the investors' optimal investments (Proposition 3.1), clearing in the security markets implies the following when mispricing (case (c,e)) occurs:

$$\Delta_{S,P}(t) = \bar{\mu}(t)\sigma - \frac{(\beta-1)\,\sigma^2}{1-\lambda(t)} - \frac{\gamma\sigma^2}{\lambda(t)} - \frac{\theta_S^3(t)\sigma^2}{\lambda(t)(1-\lambda(t))W(t)}. \tag{5.1}$$

Hence, the arbitrageur's influence on security prices manifests itself, via (5.1), as the size of his stock position (θ_S^3) affecting mispricing. The larger his position is, the lower the mispricing – in a sense, the prices move against him: the value of his marginal profit (the mispricing) is reduced by his trades.

The non-competitive arbitrageur solves the following optimization problem at time t:

$$\max_{\theta_S^3(t),\Delta_{S,P}(t)} \quad \Psi^3(t) = \theta_S^3(t)\Delta_{S,P}(t) \tag{5.2}$$

$$\begin{array}{lll} \text{s.t.} & \theta_S^3(t)+\theta_P^3(t)+\theta_B^3(t)=0, \qquad \left(\theta_S^3(t)+\theta_P^3(t)\right)\sigma = 0, & \text{((4.1)-(4.2))} \\ \text{and} & \Delta_{S,P}(t) \text{ satisfies (5.1).} & \end{array}$$

The arbitrageur simultaneously solves for the optimal size of his arbitrage position and the equilibrium mispricing, given his understanding of how the equilibrium mispricing responds to his trades ((5.1)). As for the competitive arbitrageur, the costless-riskless arbitrage conditions, (4.1)-(4.2), uniquely determine the non-competitive arbitrageur's holdings in financial securities in terms of his stock investment. Finally, we note that the maximization problem (5.2) is a simple concave quadratic problem.

5.2. Analysis of Equilibrium

Proposition 5.1 reports the possible equilibrium cases and the characterization for each case.

Proposition 5.1. *The equilibrium with a non-competitive arbitrageur is as follows.*

When $\bar{\mu}(t) \le \frac{\gamma\sigma}{\lambda(t)} + \min\left\{\frac{\sigma}{\lambda(t)}, \frac{(\beta-1)\sigma}{1-\lambda(t)}\right\}$*, investors are in (a,a), and when* $\bar{\mu}(t) > \frac{(\gamma+1)\sigma}{\lambda(t)}$ *and* $\lambda(t) \ge \frac{1}{\beta}$*, investors are in (a,b). In both cases,* $\Delta_{S,P}(t)$*,* $r(t)$ *and* $\lambda(t)$ *are as in the competitive arbitrageur equilibrium described in Proposition 4.1.*

When $\bar{\mu}(t) > \frac{\gamma\sigma}{\lambda(t)} + \frac{\sigma(\beta-1)}{1-\lambda(t)}$ *and* $\lambda(t) < \frac{1}{\beta}$*, investors are in (c,e) and the arbitrageur's equilibrium stock investment and profit are given by*

$$\begin{aligned} \theta_S^3(t) &= \frac{\lambda(t)(1-\lambda(t))W(t)}{2\sigma}\left[\bar{\mu}(t) - \frac{\sigma\,(\beta-1)}{1-\lambda(t)} - \frac{\gamma\sigma}{\lambda(t)}\right], \qquad (5.3)\\ \Psi^3(t) &= \frac{\lambda(t)(1-\lambda(t))W(t)}{4}\left[\bar{\mu}(t) - \frac{\sigma\,(\beta-1)}{1-\lambda(t)} - \frac{\gamma\sigma}{\lambda(t)}\right]^2, \end{aligned}$$

and the equilibrium mispricing, interest rate and distribution of wealth dynamics are given by

$$\Delta_{S,P}(t) = \frac{1}{2}\left[\bar{\mu}(t)\sigma - \frac{(\beta-1)\,\sigma^2}{1-\lambda(t)} - \frac{\gamma\sigma^2}{\lambda(t)}\right],$$

$$
\begin{aligned}
r(t) &= \mu_S^2(t) - \sigma^2 + \sigma\bar{\mu}(t)\frac{\lambda(t)}{2} + \frac{1}{2}\gamma\sigma^2 + \frac{1}{2}\frac{\lambda(t)}{1-\lambda(t)}\sigma^2(\beta-1),\\
d\lambda(t) &= \lambda(t)\left\{\Delta_{S,P}(t)\left[(\beta-1) + \frac{\theta_S^3(t)}{W(t)}\right] + \left[(1-\lambda(t))\left(\bar{\mu}(t) - \Delta_{S,P}(t)/\sigma\right)\right]^2\right\}dt\\
&\quad +\lambda(t)(1-\lambda(t))\left(\bar{\mu}(t) - \Delta_{S,P}(t)/\sigma\right)dw^1(t).
\end{aligned}
$$

In all cases, the aggregate wealth dynamics follow

$$
dW(t) = \left[W(t)\left(\mu_S^1(t) - \frac{1}{T-t}\right) - \theta_S^3(t)\Delta_{S,P}(t)\right]dt + W(t)\sigma dw^1(t).
$$

As in the competitive case, three cases are possible in equilibrium, with case (c,e) being the mispricing case. Comparison with the competitive equilibrium of Proposition 4.1 reveals that the region of mispricing is larger in the non-competitive equilibrium. Hence, the presence of the non-competitive arbitrageur makes mispricing more likely to occur. This is intuitive: in the presence of mispricing, the competitive arbitrageur always trades to the full extent that is allowed by the position limit, and so is more likely to fully "arbitrage away" the mispricing, making it disappear. In the non-competitive case, in contrast, he limits his trades so as to make the mispricing (and positive arbitrage profits) subsist under a broader range of conditions.[11] As a result, unlike in the competitive case, the conditions for regions (c,e) are as if the arbitrageur were not present.

As compared with the no-arbitrageur economy (Economy II, Proposition 4.1 with $M = 0$), we see that the equilibrium mispricing is reduced by a half, $\Delta_{S,P}(t) = \frac{1}{2}\Delta_{S,P}^{II}(t)$. That is, instead of providing a perfect counterparty to the two investors (which would eliminate the mispricing), when the arbitrageur is a non-price-taker, he provides only one half of that counterparty. This is also easy to understand: the arbitrageur's profit is proportional to the product of the mispricing and the size of his position. The mispricing is affine in θ_S^3 so the profit is a quadratic function of the arbitrageur's position. Therefore, it is maximized at a position halfway between zero (which would lead to zero profit) and a perfect counterparty to the investors (which would lead to zero mispricing, and hence zero profit).

In contrast to the constant competitive case, the non-competitive arbitrageur's stock position is stochastic, driven by the investors' difference of opinion $\bar{\mu}$, the investors' aggregate wealth W and the distribution of wealth λ. For sufficiently low investor disagreement, the non-competitive arbitrageur's stock position is lower than in the competitive case ($\theta_S^3(t) < M$), and higher for high investor disagreement. Hence, for low investor disagreement, the non-competitive equilibrium mispricing is higher than the competitive one. Consequently, the economic role of the arbitrageur (as discussed in Section 4), in terms of alleviating the effect of portfolio constraints and facilitating trade amongst investors, is reduced. The opposite holds when there is high disagreement amongst investors: the economic role of the non-competitive arbitrageur is more pronounced. Nonetheless, our discussion of the economic role of the arbitrageur in Section 4 remains valid, when the

[11] It is impossible for the arbitrageur to make the mispricing case more likely because, whenever he trades, he improves risk-sharing across investors and reduces the mispricing.

exogenous bound on the arbitrageur's position M is replaced with the endogenous θ_S^3. Finally, in contrast to the linear competitive case, the non-competitive arbitrageur's profits are convex in the amount of disagreement amongst investors.

6. Equilibrium with an Arbitrageur subject to Margin Requirements

We now return to a competitive market, and assume that the arbitrageur is subject to margin requirements. He then needs to be endowed with some capital, which we assume to be held (as an investment) by the investors. In addition to being more realistic, this modification of our model allows us to endogenize the amount of capital that is allocated to arbitrage activity and the size of the arbitrage positions.

6.1. The Arbitrageur under Margin Requirements

The economic setup introduced in Sections 3 and 4 is modified as follows. Riskless, costless storage, henceforth referred to as *cash*, is available to the agents in addition to the stock, derivative and bond. Assuming that $r(t) > 0$,[12] the investors would never find it optimal to invest in cash. However, the arbitrageur may be forced to hold cash due to his margin requirements, that are as follows: letting θ_C^3 and W^3 denote the arbitrageur's cash holding and capital, respectively, his holdings must obey:

$$W^3(t) \geq \eta \left(\theta_S(t) + \left|\theta_P^3(t)\right|\right), \tag{6.1}$$

$$\theta_C^3(t) \geq (1 + \epsilon\, \eta) \max \left\{-\theta_P^3(t), 0\right\}, \tag{6.2}$$

where $\eta, \epsilon \in [0, 1]$ and $\theta_S^3(t) \geq 0$ (hence there is no need to account for short positions in S in (6.1)-(6.2)). While equation (6.1) limits the size of the arbitrageur's position in proportion to his capital, equation (6.2) states that the arbitrageur does not get the use of proceeds from his short position in the derivative; rather, these must be kept in cash as a deposit. In addition, a fraction ϵ of the margin on the short sale must also be met with cash. For clarity, we will refer below to (6.1) as the "margin constraint" and to (6.2) as the "cash constraint". Our modeling of margin requirements is standard (see, e.g., Cuoco and Liu (2000)).[13] For simplicity, we assume that there is no difference between initial and maintenance margin. Denoting $\theta^3 \equiv (\theta_B^3, \theta_S^3, \theta_P^3, \theta_C^3)$, the arbitrageur's problem is as follows:

$$\max_{\theta^3} E^3 \left[\int_0^T \Psi^3(t) dt\right]$$

[12] This is verified to hold in equilibrium, for an appropriate choice of the exogenous parameters, including the pessimistic investor's prior belief distribution on μ_S.

[13] Our modeling of margin requirements could easily be amended without affecting our main intuition. All that is really needed is the presence of a constraint that limits the size of the arbitrageur's position in proportion to his capital, and that of costly short sales.

$$
\begin{aligned}
\text{s.t.} \quad & \Psi^3(t)dt = \left\{\theta_S^3(t)\mu_S^3(t) + \theta_P^3(t)\mu_P^3(t) + \theta_B^3(t)r(t)\right\} dt \\
& \qquad + [\theta_S^3(t) + \theta_P^3(t)]\sigma dw^3(t), \\
& \theta_S^3(t) + \theta_P^3(t) + \theta_B^3(t) + \theta_C^3(t) = W^3(t), && (6.3) \\
& \left(\theta_S^3(t) + \theta_P^3(t)\right)\sigma = 0 && (6.4) \\
\text{and} \quad & \text{(6.1)-(6.2) hold.}
\end{aligned}
$$

Both the amount of arbitrage capital (W^3) and its return ($\Psi^3(t)/W^3(t)$) are to be determined endogenously in equilibrium. Restrictions (6.3), (6.4) and equality in (6.2) (which always holds because cash is dominated by the bond) yield all holdings as a function of $\theta_P^3(t)$:

$$\theta_S^3(t) = -\theta_P^3(t), \quad \theta_B^3(t) = -(1+\epsilon\eta)\max\left\{-\theta_P^3(t), 0\right\}, \tag{6.5}$$

$$\theta_C^3(t) = (1+\epsilon\eta)\max\left\{-\theta_P^3(t), 0\right\}. \tag{6.6}$$

When there is no mispricing ($\Delta_{S,P}(t) = 0$), the arbitrageur is indifferent between all portfolios satisfying (6.1), (6.5)-(6.6) and $\theta_P^3(t) \geq 0$; the last inequality holds because (6.2) penalizes short sales (the arbitrageur loses the interest on his cash deposit). As before, all feasible portfolio holdings yield zero profit. In the presence of mispricing ($\Delta_{S,P}(t) > 0$), it is optimal for the arbitrageur to take on as large a profitable arbitrage position (long in the cheap stock, and short in the more expensive derivative) as feasible; hence, (6.1) holds with equality:

$$\eta\left(\theta_S^3(t) - \theta_P^3(t)\right) = W^3(t),$$

uniquely determining the arbitrageur's holdings as a function of his capital:

$$\theta_P^3(t) = -\frac{W^3(t)}{2\eta} < 0, \qquad \theta_S^3(t) = \frac{W^3(t)}{2\eta} > 0, \tag{6.7}$$

$$\theta_B^3(t) = W^3(t)\left[1 - \frac{1+\epsilon\eta}{2\eta}\right], \qquad \theta_C^3(t) = \frac{(1+\epsilon\eta)\,W^3(t)}{2\eta}. \tag{6.8}$$

Thus, the instantaneous return on the arbitrageur's capital is as follows:

$$\frac{\Psi^3(t)}{W^3(t)} = r(t) + \left[\Delta_{S,P}(t) - r(t)\,(1+\epsilon\eta)\right]\frac{1}{2\eta}. \tag{6.9}$$

6.2. Analysis of Equilibrium

The arbitrageur, in this economic setting, is interpreted as an arbitrage firm, whose stock is held by the investors. In other words, arbitrage is another investment opportunity. Denoting by θ_A^i investor i's (dollar) investment in the arbitrage firm, the amount of the arbitrageur's capital satisfies: $\theta_A^1(t) + \theta_A^2(t) = W^3(t)$, and investor i's dynamic budget constraint is now:

$$
\begin{aligned}
dW^i(t) &= \left[W^i(t)r(t) - c^i(t)\right]dt + \left\{\theta_S^i(t)\left[\mu_S^i(t) - r(t)\right] + \theta_P^i(t)\left[\mu_P^i(t) - r(t)\right]\right\}dt \\
&\quad + \left[\theta_S^i(t) + \theta_P^i(t)\right]\sigma dw^i(t) + \theta_A^i(t)\frac{\Psi^3(t)}{W^3(t)}dt.
\end{aligned}
$$

For simplicity, we assume that the investors face no restriction on their investments in the arbitrage firm. Our assumptions from Sections 3 and 4 are otherwise unaffected. In particular, the arbitrageur is assumed to be competitive, and as before we denote $W \equiv W^1 + W^2$ and $\lambda \equiv W^1/W$.

In an equilibrium with mispricing, the return on arbitrage (6.9) must be consistent with the returns on the other investment opportunities, leading to the equilibrium described by Proposition 6.1.

Proposition 6.1. *Assume that*

$$\begin{aligned} \bar{\mu}(t) &> \frac{\gamma\sigma}{\lambda(t)} + \frac{(\beta-1)\,\sigma}{1-\lambda(t)} + \frac{\sigma W^3(t)}{\lambda(t)(1-\lambda(t))W(t)}\left(\frac{\lambda(t)\,(1+\epsilon\eta)+1}{2\eta}\right) \\ \textit{and} \qquad 0 &\leq W^3(t) \leq \lambda(t)W(t)\left(\frac{2\eta}{1+(1+\epsilon\eta)}\right)\left[1-\frac{\lambda(t)}{1-\lambda(t)}\,(\beta-1)\right], \end{aligned}$$

where

$$W^3(t) = \frac{\bar{\mu}(t)\sigma - (1+\epsilon\eta)\,\mu_S^2(t) - \frac{(\beta-1)\sigma^2}{1-\lambda(t)} - \frac{\gamma\sigma^2}{\lambda(t)} - (1+\epsilon\eta)\,\sigma^2\left(2+\frac{\lambda(t)(\beta-1)}{1-\lambda(t)}\right)}{\frac{\sigma^2}{2\eta}\left[\frac{(1+\epsilon\eta)}{(1-\lambda(t))W(t)}\,(3+\epsilon\eta) + \frac{1}{\lambda(t)(1-\lambda(t))W(t)}\right]}. \tag{6.10}$$

Then, an equilibrium where investors 1 and 2 are in cases (c,e) results, the aggregate amount of capital invested in arbitrage is as in equation (6.10), and the mispricing, interest rate, distribution of wealth dynamics and aggregate wealth dynamics are as follows:

$$\Delta_{S,P}(t) = \bar{\mu}(t)\sigma - \frac{(\beta-1)\,\sigma^2}{1-\lambda(t)} - \frac{\gamma\sigma^2}{\lambda(t)} - \frac{W^3(t)\sigma^2}{\lambda(t)(1-\lambda(t))W(t)}\left(\frac{\lambda(t)\,(1+\epsilon\eta)+1}{2\eta}\right) > 0, \tag{6.11}$$

$$r(t) = \mu_S^2(t) - (1-\gamma)\sigma^2 + \left[\frac{(\beta-1)\,\lambda(t)}{1-\lambda(t)} + \frac{W^3(t)}{(1-\lambda(t))W(t)}\left(\frac{2+\epsilon\eta}{2\eta}\right)\right]\sigma^2, \tag{6.12}$$

$$\begin{aligned} d\lambda(t) &= \left\{\Delta_{S,P}(t)\,(\beta-\lambda(t)) + (\sigma+\sigma_\lambda(t))\,[\sigma(1-2\lambda(t))+\sigma_\lambda(t)] + \lambda(t)\sigma^2\right\}dt + \lambda(t)\sigma_\lambda(t)dw^1(t), \\ \textit{where } \sigma_\lambda(t) &= (1-\lambda(t))\,(\bar{\mu}(t) - \Delta_{S,P}(t)/\sigma) - \frac{W^3(t)\,(1+\epsilon\eta)\,\sigma}{2\eta W(t)}, \\ dW(t) &= \left\{W(t) - (1+\epsilon\eta)\left[\gamma(1-\lambda(t))W(t) + \frac{W^3(t)}{2\eta}\right]\right\}\left(\mu_S^1(t)dt + \sigma dw^1(t)\right) - \frac{W(t)}{T-t}dt. \end{aligned}$$

Because arbitrage is riskless, equilibrium is only possible if the rate of return on arbitrage capital is equal to the riskless interest rate; otherwise, investors would face an unbounded arbitrage opportunity and equilibrium would not obtain. From (6.9), this implies that $\Delta_{S,P}(t) = r(t)\,(1+\epsilon\eta)$, which leads (using (6.11) and (6.12)), to an equation affine in $W^3(t)$, whose solution is given by (6.10). The endogenous amount of arbitrage capital W^3 is provided by equation (6.10)). Interestingly, the main factor that drives the equilibrium amount of arbitrage activity is similar to the non-competitive case of Section 5: the mispricing that would prevail without an arbitrageur, $\Delta_{S,P}^{II} = \bar{\mu}\sigma - (\beta-1)\sigma^2/(1-\lambda) - \gamma\sigma^2/\lambda$, i.e., the amount of heterogeneity that cannot be traded by the investors due to their constraints. This is intuitive: the higher this "un-traded" heterogeneity, the higher the profit opportunities for the arbitrageur/financial intermediary. In addition, the amount of arbitrage capital is increasing in the severity of the margin constraint (6.1) (as measured by η): the more stringent the constraint, the higher the amount of capital needed to achieve a similar result. In contrast, W^3 is decreasing in the severity of the cash constraint (6.2) (i.e., in $(1 + \epsilon\eta)$). The rationale for this will be clarified by our discussion on the respective impact of these two constraints on risk-sharing.

The structure of equilibrium closely resembles that of Section 4, with the exogenous position size M being replaced by the endogenous arbitrageur's stock investment $\theta_S^3 = W^3/2\eta$. Nonetheless, risk-sharing between investors is further impacted by the cash constraint. This is evidenced by the diffusion of the wealth distribution dynamics, σ_λ, which exhibits an additional term proportional to $(1 + \epsilon\eta)$. So do the expressions for the interest rate and the mispricing, suggesting that the more severe the cash constraint, the better risk-sharing between investors.

To investigate risk-sharing among investors, we examine the investors' risk exposures, as provided in Proposition 6.2. We note that the expressions in the proposition are not dependent of our assumption of logarithmic utility for the investors, as they follow from the investors' binding constraints, the arbitrageur's policy ((6.7)-(6.8)) and market clearing. (They would also hold in a pure exchange economy.)

Proposition 6.2. *In the mispriced equilibrium with an arbitrageur subject to the margin requirement (6.1)-(6.2), the investors' risk exposures are as follows:*

$$\begin{aligned}\Phi^1(t) &= \beta + \gamma\frac{1-\lambda(t)}{\lambda(t)} + \frac{W^3(t)}{2\eta\lambda(t)W(t)},\\ \Phi^2(t) &= (1-\gamma) - (\beta - 1)\,\frac{\lambda(t)}{1-\lambda(t)} - \frac{(2+\epsilon\eta)\,W^3(t)}{2\eta(1-\lambda(t))W(t)}.\end{aligned} \tag{6.13}$$

Equation (6.13) reveals that the pessimistic investor's equilibrium risk exposure is decreased by the severity of the cash constraint (measured by ϵ), hence an improvement in risk-sharing due to the cash constraint. This is because the cash deposit is taken out of the aggregate amount of good to be invested in production. Hence, market clearing requires $\theta_S^1 + \theta_S^2 + \theta_S^3 = W^3 - \theta_C^3$. Thus, the higher the arbitrageur's cash holding, the less the pessimistic investor has to invest in the stock for markets to clear. (Investor 2's investment has to adjust to clear the markets, because investor 1's leverage constraint on the stock is binding.) This makes it possible for him to have a lower risk exposure, closer to the value it would have in an unconstrained equilibrium. Investor 1's risk exposure is unaffected by the arbitrageur's cash constraint, leading to a higher value for our measure of risk-sharing $(\Phi^1 - \Phi^2)$. Thus, the cash constraint mitigates the effect of the margin constraint, albeit in an indirect way. From (6.13), the effect of the cash constraint on investor 2's risk exposure is equivalent to an increase in the amount of arbitrage activity. The effect of the margin constraint (6.1), on the other hand, is as would be expected: it is similar to that of the arbitrageur's investment restriction in Section 4, and the intuition therein goes through. In fact, in the absence of any cash requirement $(1+\epsilon\eta=0)$, the expressions from the mispriced equilibrium in Section 4 go through, once M is replaced with the endogenous θ_S^3.

The model of equilibrium arbitrage activity in this section is admittedly simplistic in many respects. In particular, our implication that the return on arbitrage capital equals the riskless interest rate is undoubtedly unrealistic. Nonetheless, we believe that the general approach of this section is consistent with the nature of arbitrage in contemporary financial markets, with specialized agents engaging in arbitrage, such as hedge funds, whose capital is held by outside investors, and could be extended to a more realistic setup. Finally, we demonstrate the different ways in which the constraints imposed on the arbitrageur can affect risk-sharing.

7. Conclusion

In this article, we attempt to shed some light on the potential role of arbitrageurs in rational financial markets. Towards that end, we develop an economic setting with two heterogeneous risk-averse investors subject to leverage and short-sales restrictions, and an arbitrageur engaging in costless, riskless arbitrage trades. In the presence of arbitrage opportunities, the arbitrageur is shown to improve risk-sharing amongst investors. He effectively plays the role of a financial intermediary, or an innovator synthesizing a derivative security specifically designed to relieve investors' restrictions. The improved transfer of risk due to the arbitrageur is shown to be equally valid when the arbitrageur behaves non-competitively, or is subject to margin requirements and needs capital to implement his arbitrage trades. Although we make simplifying assumptions on the primitives of the economy (preferences, investment opportunities, investment restrictions) for tractability, our insights can readily be extended to more general primitives and to a pure exchange economy.

Our framework is amenable to the study of further related issues. For example, the analysis of the arbitrageur under margin requirements could be extended to make arbitrage activity risky by allowing for the possibility of volatility jumps in a discretized version of the model. Then, the risk and return tradeoff of arbitrage capital being consistent with asset prices would pin down the endogenous amount of arbitrage capital. Our model also sheds new light on the issue of who the users of derivatives markets such as futures are, what the motivations of the different categories of market participants are, and what returns they obtain. Typically, the empirical literature on futures markets (e.g., Ederington and Lee (2002), Piazzesi and Swanson (2004)), distinguishes between hedgers and speculators. Our model underscores the need to additionally account for the activity of arbitrageurs, and provides theoretical tools to understand how their positions and their returns are related to underlying economic variables.

Appendix: Proofs

Proof of Proposition 3.1: The investors' optimization problem is non-standard in several respects: (i) the portfolio constraints; (ii) the redundancy in the risky investment opportunities (technology and stock); (iii) the mispricing. The solution technique, developed by Basak and Croitoru (2000), involves using investors' non-satiation (implying that, in the presence of mispricing, the investment restrictions always bind) to convert the original problem with two risky investment opportunities into one with a single, fictitious risky asset with a nonlinear drift. Techniques of optimization in nonlinear markets (Cvitanic and Karatzas (1992)) can then be applied. *Q.E.D.*

Proof of Proposition 4.1: Combining the agents' optimal policies (Proposition 3.1 and Section 4.1) and the conditions for market clearing, and noting that both investors facing the same price process for the derivative P implies:

$$\frac{\mu_P^1(t) - \mu_P^2(t)}{\sigma} dt = dw^2(t) - dw^1(t) = \frac{\mu_S^1(t) - \mu_S^2(t)}{\sigma} dt, \quad \text{hence} \quad \frac{\mu_P^1(t) - \mu_P^2(t)}{\sigma} = \bar{\mu}(t),$$

leads to the equilibrium expressions for r and $\Delta_{S,P}$ in each case. Substituting these equilibrium prices in the conditions for the investors' optimization cases leads to the conditions for the equilibrium case (c,e) and the first condition for (a,b). The conditions for (a,a) and the second condition for (a,b) are those under which it is possible to find portfolio holdings that, among all those between which agents are indifferent, clear markets.

To verify the conditions for the (a,a) case, observe that, if the economy is in (a,a), each of the three agents is indifferent between all the portfolio holdings that satisfy, respectively:

$$\frac{\theta_S^1(t)}{W^1(t)} + \frac{\theta_P^1(t)}{W^1(t)} = \frac{\mu_S^1(t) - r(t)}{\sigma^2} = 1 + \frac{\bar{\mu}(t)(1 - \lambda(t))}{\sigma}, \tag{A.1}$$

$$\frac{\theta_S^2(t)}{W^2(t)} + \frac{\theta_P^2(t)}{W^2(t)} = \frac{\mu_S^2(t) - r(t)}{\sigma^2} = 1 - \frac{\bar{\mu}(t)\lambda(t)}{\sigma}, \tag{A.2}$$

$$\theta_P^3(t) = -\theta_S^3(t), \qquad \theta_B^3(t) = 0, \tag{A.3}$$

where the equilibrium value for the interest rate has been substituted into the investors' portfolio demands. In addition, agents' portfolio holdings must satisfy the clearing conditions, that are equivalent to:

$$\theta_S^1(t) + \theta_S^2(t) + \theta_S^3(t) = W^1(t) + W^2(t), \qquad \theta_P^1(t) + \theta_P^2(t) + \theta_P^3(t) = 0. \tag{A.4}$$

There exist an infinity of solutions to the system formed by equations (A.1)-(A.4). The conditions for case (a,a) provided in the Proposition ensure that there exists a solution that satisfy the agents' position limits ((3.1) and (4.3)). To check this, express the solutions of the system as a function of two free parameters (say, $\theta_P^1(t)$ and $\theta_S^3(t)$), since there are two degrees of freedom in the system, and substitute these expressions into the position limits. This results in eight inequalities that $\theta_P^1(t)$ and $\theta_S^3(t)$ must obey. The conditions for case (a,a) ensure that there is no contradiction among these, so that there exist portfolio holdings that satisfy all the requirements for the unconstrained equilibrium ((a,a)). The second condition for case (a,b) can be obtained in a similar fashion.

It is easy to check that no other cases are possible in equilibrium (the only other cases compatible with market clearing, (b,a) and (e,c), are impossible given that investor 1 is more optimistic). Substituting the prices and optimal policies into the investors' dynamic budget constraints and using Itô's lemma leads to the dynamics of λ and W, while the expressions for the investors' consumptions follow from Proposition 3.1. *Q.E.D.*

Proof of Corollary 4.1: The expressions follow from the definition of Φ^i, the optimal policies in Proposition 3.1, and the equilibrium prices of Proposition 4.1. *Q.E.D.*

Proof of Proposition 4.2: The expressions follow from the optimal policies in Proposition 3.1 and market clearing. Substituting the equilibrium prices into the conditions for the optimality cases leads to the conditions for the (c,e,c) equilibrium. Details are omitted for brevity. *Q.E.D.*

Proof of Proposition 5.1: In the presence of an active arbitrageur, market clearing and agents' optimal policies lead to the expressions for the equilibrium mispricing as a function of the arbitrageur's position ((5.1)). Substitution into the arbitrageur's optimization problem (5.2) leads to a quadratic problem whose solution is provided by (5.3). Substitution into the problem's objective function leads to the optimal value of $\Psi^3(t)$. Given the arbitrageur's position, the equilibrium values of the mispricing, interest rate, distribution of wealth dynamics and aggregate wealth dynamics are deduced in a fashion similar to the competitive case (Proposition 4.1). The conditions for the equilibrium are those under which there exists a solution to the arbitrageur's optimization such that he makes a positive profit, and the investors are optimally in case (c,e). *Q.E.D.*

Proof of Proposition 6.1: The interest rate and mispricing, as a function of W^3, follow from the arbitrageur's holdings in (6.7)-(6.8), the investors' demands (Proposition 3.1) and market clearing, in a fashion similar to Proposition 4.1. To pin down the amount of arbitrage capital W^3, observe that there only exists a solution to the investors' optimization if the (riskless) return on arbitrage (6.9) equals the riskless rate r; otherwise they would face an unbounded arbitrage opportunity. By substituting the interest rate and mispricing into the expression for the return on arbitrage capital (6.9), we obtain (6.10), the only amount of arbitrage capital such that this restriction holds. The conditions in the Proposition ensure that the aggregate amount invested in arbitrage activity is nonnegative, and that the investors, given the trades of the arbitrageur, are in cases (c) and (e). *Q.E.D.*

Proof of Proposition 6.2: The investors' portfolio holdings can be deduced from their binding constraints, the arbitrageur's holdings ((6.7)-(6.8)) and market clearing. Applying the definition of Φ^i then yields the expressions in the Proposition. *Q.E.D.*

References

Attari, M. and A. Mello, 2002, "Financially Constrained Arbitrage in Illiquid Markets," working paper, University of Wisconsin-Madison.

Basak, S. and B. Croitoru, 2000, "Equilibrium Mispricing in a Capital Market with Portfolio Constraints," *Review of Financial Studies*, 13, 715-748.

Basak, S. and B. Croitoru, 2003, "On the Role of Arbitrageurs in Rational Markets," working paper, www.afajof.org/pdf/2004program/UPDF/P371_Asset_Pricing.pdf

Brennan, M. and E. Schwartz, 1990, "Arbitrage in Stock Index Futures," *Journal of Business*, 63, S7-S31.

Cox, J., J. Ingersoll and S. Ross, 1985, "An Intertemporal General Equilibrium Model of Asset Prices," *Econometrica*, 53, 363-384.

Cuoco, D. and H. Liu, 2000, "A Martingale Characterization of Consumption Choices and Hedging Costs with Margin Requirements," *Mathematical Finance*, 10, 355-385.

Cvitanic, J. and I. Karatzas, 1992, "Convex Duality in Constrained Portfolio Optimization," *Annals of Applied Probability*, 2, 767-818.

De Long, B., A. Shleifer, L. Summers and R. Waldmann, 1990, "Noise Trader Risk in Financial Markets," *Journal of Political Economy*, 98, 703-738.

Detemple, J. and S. Murthy, 1994, "Intertemporal Asset Pricing with Heterogeneous Beliefs," *Journal of Economic Theory*, 62, 294-320.

Detemple, J. and S. Murthy, 1997, "Equilibrium Asset Prices and No-Arbitrage with Portfolio Constraints," *Review of Financial Studies*, 10, 1133-1174.

Duffie, D., N. Garleanu and L. Pedersen, 2002, "Securities Lending, Shorting and Pricing," *Journal of Financial Economics*, 66, 307-339.

Ederington, L. and J. H. Lee, 2002, "Who Trades Futures and How: Evidence from the Heating Oil Futures Market," *Journal of Business*, 75, 353-373.

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

Harrison, J. and D. Kreps, 1978, "Speculative Behavior in a Stock Market with Heterogeneous Expectations," *Quarterly Journal of Economics*, 93, 323-336.

Karatzas, I. and S. Shreve, 1998, *Methods of Mathematical Finance*, Springer.

Lamont, O. and R. Thaler, 2003, "Can the Market Add and Subtract? Mispricing in Tech Stock Carve-outs," *Journal of Political Economy*, 111, 227-268.

Lhabitant, S., 2002, *Hedge Funds*, Wiley.

Liu, J. and F. Longstaff, 2004, "Losing Money on Arbitrages: Optimal Dynamic Portfolio Choice in Markets with Arbitrage Opportunities," *Review of Financial Studies*, 17, 611-641.

Loewenstein, and G. Willard, 2000, "Convergence Trades and Liquidity: A Model of Hedge Funds," working paper, Boston University.

Lucas, R., 1978, "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429-1446.

Mitchell, M., T. Pulvino and E. Stafford, 2002, "Limited Arbitrage in Equity Markets," *Journal of Finance*, 57, 551-584.

Ofek, E., M. Richardson and R. Whitelaw, 2004, "Limited Arbitrage and Short Sales Restrictions: Evidence from the Options Markets," *Journal of Financial Economics*, 74, 305-342.

Piazzesi, M. and E. Swanson, 2004, "Futures Prices as Risk-adjusted Forecasts of Monetary Policy," NBER working paper.

Scheinkman, J. and W. Xiong, 2003, "Overconfidence, Short-Sale Constraints and Bubbles," *Journal of Political Economy*, 111, 1183-1219.

Shleifer, A., 2000, *Inefficient Markets*, Oxford University Press.

Xiong, W., 2001, "Convergence Trading with Wealth Effects: An Amplification Mechanism in Financial Markets," *Journal of Financial Economics*, 62, 247-292.

Xiong, W. and A. Kyle, 2001, "Contagion as a Wealth Effect," *Journal of Finance*, 56, 1401-1440.