On the Role of Arbitrageurs in Rational Markets*

Suleyman Basak
London Business School and CEPR
Institute of Finance and Accounting
Regent’s Park
London NW1 4SA
United Kingdom
Tel: (44) 20 7706 6847
Fax: (44) 20 7724 3317
sbasak@london.edu

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

January 2003

*We are grateful to seminar participants at McGill University and the University of Wisconsin-Madison for their helpful comments. Financial support from FQRSC, IFM2 and SSHRC is gratefully acknowledged. All errors are solely our responsibility.
On the Role of Arbitrageurs in Rational Markets

Abstract

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 subject to position limits. An arbitrageur, indulging in costless, riskless arbitrage is shown to alleviate the effects of position limits 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 numbers: C60, D50, D90, G11, G12.

Keywords: arbitrage, asset pricing, margin requirements, non-competitive markets, risk-sharing.
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 (2000))\(^1\) and the deviations from put-call parity in options markets (Ofek, Richardson and Whitelaw (2002)). Such mispricings clearly lead to arbitrage opportunities that are, in theory, riskless and costless. Not surprisingly, there is ample evidence of market participants engaging in trades designed to reap these arbitrage profits, and 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 position limits.\(^2\) 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 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 allocations. Our setup is comparable to the main workhorse models in asset pricing (Lucas (1978), Cox, Ingersoll and Ross (1985)),

\(^1\)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 (2002)) points to impediments to short-selling comparable to our position limits 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 position limits that could be incorporated into our work without considerably affecting our main message.
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, finite state economy in which Arrow-Debreu securities are traded, and redundant securities are available in at least one state. In this setting, “mispricing” means that redundant securities (here, Arrow-Debreu claims paying-off in the same state) trade at different prices. 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 extends the space of trades that are possible for the investors. The larger the arbitrageur’s trades, the more herelieves the effects of the position limits. The arbitrageur acts a financial intermediary and an innovator who issues a fictitious, additional zero net supply security that is purchased from the investor who desires the lowest exposure to risk and resold, at a higher price, to the other investor. 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 production technology; a risky, zero net supply “derivative” financial security; and a riskless bond. In this setup, “mispricing” refers to any discrepancy in the market prices of risk offered by the risky investment opportunities (technology and derivative), which generates a riskless arbitrage opportunity. For tractability, our investors are assumed to have logarithmic utility preferences. To generate heterogeneity and trade, we assume that the investors have heterogeneous beliefs on the mean return on production. Mispricing is shown to occur when investors are sufficiently heterogeneous in their beliefs.

Our intuition on the arbitrageur’s role from the single-period example extends readily to the continuous time case. Under mispricing, by taking advantage of the resulting arbitrage opportunity, the arbitrageur allows the investors to trade more risk, effectively relieving their position limits. Accordingly, the equilibrium prices and distribution of wealth among investors are affected by the arbitrageur, revealing that the arbitrageur’s presence alleviates the effect of the binding constraints, in addition to making these less likely to bind. The greater the size of his trades, the greater this effect; in particular, the size of the mispricing is decreasing in that of 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 is subject to an exogenously specified position limit. Then, under mispricing he always takes on the largest 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 (1997,1999), who solve for equilibrium in the presence of heterogeneous agents, with and without portfolio constraints, but not with redundant derivative securities. 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 particular form of constraints, that of a segmented market for the risky assets (where different investors hold different risky assets). Our model is somewhat less specialized (and thus easier to relate to observable economic variables or well-understood benchmarks), which prevents us from undertaking an explicit welfare analysis. Gromb and Vayanos’ main findings are consistent with our results, in that the arbitrageur’s presence benefits all investors (albeit not necessarily in a Pareto-optimal way). Loewenstein and Willard (2000) provide an example of 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 papers that are devoted to a constrained arbitrageur’s optimal policy,
given that the arbitrage opportunity is present. Examples include Brennan and Schwartz (1990) and Liu and Longstaff (2001). 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 and Vishny (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 when, in addition to the noise traders, long-term investors behave suboptimally. 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.

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 finite-state, one-period framework. Portfolio decisions are made at time 0 and the uncertainty is resolved and securities pay-off at time 1. We first consider an economy with two heterogeneous investors \( i = 1, 2 \) and no arbitrageur. Consider two Arrow-Debreu securities, \( S \) and \( P \), with time-zero prices also denoted by \( S \) and \( P \), and with identical payoffs of one unit if state \( \omega \) occurs, zero otherwise. The only difference between \( S \) and \( P \) are their net supplies, \( s \) and \( p \) shares respectively, and how trading in \( S \) and \( P \) is constrained. Denoting by \( \alpha^i_j \) investor \( i \)'s investment in security \( j \) (where \( j = S, P \)), measured in number of shares, we assume that both investors’ holdings are constrained as follows:

\[
\alpha^i_S \leq \beta, \quad \alpha^i_P \geq -\gamma, \quad i = 1, 2. \tag{2.1}
\]

For simplicity, we assume that \( S \) and \( P \) are the only two available securities paying-off in state \( \omega \). Then, investor \( i \)'s payoff if state \( \omega \) occurs is given by \( A^i = \alpha^i_S + \alpha^i_P \). The constraints in (2.1) alone do not place any restrictions on the choice of \( A^i \), because there is no lower bound on holdings of \( S \) and no upper bound on \( P \). Nonetheless, only a limited set of payoffs are available to the
investors in equilibrium. For example, if \( P \) is in zero net supply (\( p = 0 \)), no agent can hold (long) more than \( \gamma \) shares thereof, because the other investor cannot provide a counterparty without violating his lower bound on holdings in \( P \). Hence, no investor can receive a state \( \omega \)-payoff greater than \( \gamma + \beta \).

More generally, combining the market clearing conditions (\( \alpha_1^S + \alpha_2^S = s, \alpha_1^P + \alpha_2^P = p \)) and the portfolio constraints (2.1) reveals that, in equilibrium, both agents’ state \( \omega \)-payoffs must satisfy the following constraint:

\[
s - \beta - \gamma \leq A^i \leq p + \beta + \gamma, \quad i = 1, 2.
\]  

We now examine how the set of feasible trades (those that are such that (2.2) holds) is modified in the presence of a third agent, an arbitrageur. The arbitrageur is assumed to only take on riskless, costless arbitrage positions to take advantage of any difference in price 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 holding one share (long) 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 agents. In our subsequent analysis, we provide equilibria in which such mispricings also arise endogenously. For now, we take the presence of mispricing (\( P > S \)) as given and address the issue of the economic role of an arbitrageur who exploits the mispricing. We take the size of the arbitrageur’s position (\( \alpha_3^S > 0 \) shares of \( S \), \( \alpha_3^P = -\alpha_3^S \) shares of \( P \)) as given. Market clearing conditions are then: \( \alpha_1^S + \alpha_2^S + \alpha_3^S = s, \alpha_1^P + \alpha_2^P - \alpha_3^S = p \). Substituting these into the constraints (2.1) and using the definition of \( A^i \) leads to the analogue of (2.2): to be compatible with both market clearing and the portfolio constraints, investors 1 and 2’s state \( \omega \)-payoffs must satisfy

\[
s - \beta - \gamma - \alpha_3^S \leq A^i \leq p + \beta + \gamma + \alpha_3^S, \quad i = 1, 2.
\]  

Equation (2.3) reveals that the presence of the arbitrageur enlarges the set of feasible trades. This is because the arbitrageur provides an additional counterparty to each “normal” investor, over and above what the other investor can provide without violating his portfolio constraints (2.1). The arbitrageur effectively plays the role of a financial intermediary, issuing an additional, zero-net supply security, holdings of which are bounded by the size of his position (\( \alpha_3^S \)). Even though he does not take risk of his own, the arbitrageur thus allows investors 1 and 2 to trade more risk. The larger the arbitrageur’s position, the greater the alleviation of the constraints he enables.

\footnote{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.}

\footnote{In the absence of any constraint, a competitive arbitrageur would optimally choose to hold an unbounded 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.}
It should be pointed that the role of the arbitrageur is specific: adding a third “normal”
investor, 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 constraints
(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 $\omega$-cash-flow
becomes
\[
s - 2\beta - \gamma \leq A^i \leq p + \beta + \gamma - \alpha P^s.
\]
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
constraints (position limits or constraints on portfolio weights; whether they are one- or two-
sided, homogeneous or heterogeneous across agents). Because no assumptions have been made
on the net supplies of the securities, our analysis also is 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 lifetime 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
\[
\frac{dS(t)}{S(t)} = \mu_t dt + \sigma_t dw(t),
\]
where $\mu_S$, $\sigma_S$ (with $\sigma_S > 0$) 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 commonly observe the return process of the technology, but have incomplete (but symmetric) information on its dynamics. They deduce $\sigma_S$ from the technology return’s quadratic variation, but must estimate $\mu_S$ via its conditional expectation, rationally updating their beliefs in a Bayesian fashion with heterogeneous prior beliefs. We denote by $\mu^i_S(t)$ the conditional estimate of $\mu_S$ by investor $i = 1, 2$ at time $t$, given his priors and observation of the technology’s realized return, and by $\bar{\mu}(t) = (\mu^1_S(t) - \mu^2_S(t)) / \sigma_S$ the normalized difference in the two investors’ estimates. The process $\bar{\mu}$ parameterizes investors’ disagreement on their relative “optimism/pessimism” about the technology mean return. The process $\bar{\mu}$ is given exogenously, since all learning about $\mu_S$ comes from observing the (exogenous) realized return on the technology. For simplicity, we shall assume that $\bar{\mu} > 0$, and call agent 1 the “optimistic” investor, and conversely.

Finally, we let

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

denote agent $i$’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 technology return dynamics are given by

$$\frac{dS(t)}{S(t)} = \mu^i_S(t) dt + \sigma_S dw^i(t), \quad i = 1, 2.$$ 

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

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

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

and the other a risky “derivative” with perceived return process

$$\frac{dP(t)}{P(t)} = \mu^i_P(t) dt + \sigma_P dw^i(t), \quad i = 1, 2.$$ 

Similarly to the risky technology $S$, investors observe the derivative return process but not its dynamics coefficients, and hence draw their own inferences about the derivative mean return $\mu_P$. Since the derivative pays no dividends, we take its volatility parameter $\sigma_P$ (assumed to be

$^5$Consider the well-known Gaussian example, in which investor $i$’s prior is normally distributed with mean $\mu^i_S(0)$ and variance $\nu(0)$. Then, investors’ estimates have dynamics $d\mu^i_S(t) = (\nu(t)/\sigma_S) dw^i(t)$, where $\nu(t) = \nu(0)\sigma^2_S/(\nu(0) + \sigma^2_S)$, implying $d\bar{\mu}(t) = -(\nu(t)/\sigma_S)\bar{\mu}(t) dt$, or $\bar{\mu}(t) = \bar{\mu}(0)[\sigma^2_S/(\nu(0) t + \sigma^2_S)]^{\nu}$s. 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 $\bar{\mu} > 0$. 

7
positive for simplicity) to define the derivative contract $P$, as is standard in the literature (e.g., Karatzas and Shreve (1998)). Any security whose price is continuously resettled (e.g., a futures contract) is an example of this derivative. The other security return parameters, the interest rate $r$ and the perceived mean returns $\mu_P^i$, are to be determined endogenously in equilibrium.

3.2. Investors’ Optimization problems

With all the uncertainty generated by a one-dimensional Brownian motion, the technology and derivative returns are perfectly correlated, and so the derivative is redundant. However, position limits on investors’ risky investments generate an economic role for the derivative. Letting $\theta^i \equiv (\theta^i_B, \theta^i_S, \theta^i_P)$, with $\theta^i_B = W^i(t) - \theta^i_S(t) - \theta^i_P(t)$, 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, investments in the technology are bounded from above, and that investments in the derivative are bounded from below at all times:

$$0 \leq \theta^i_S(t) \leq \bar{\beta} W^i(t), \quad \theta^i_P(t) \geq -\gamma W^i(t),$$

where $\bar{\beta} \geq 1$, $\gamma > 0$.

Absent position limits, with one factor of risk no-arbitrage would enforce identical market prices of risk across the risky assets. This motivates our definition of mispricing between the technology and derivative under an investor $i$’s beliefs as the difference between their market prices of risk:

$$\Delta^i_{S,P}(t) \equiv \frac{\mu^i_S(t) - r(t)}{\sigma_S} - \frac{\mu^i_P(t) - r(t)}{\sigma_P}.$$  

We note that the observed return agreement across investors enforces agreement on the mispricing: $\Delta^1_{S,P}(t) = \Delta^2_{S,P}(t) = \Delta_{S,P}(t)$. We say that $S$ is favorable if $\Delta_{S,P}(t) > 0$, and conversely.

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

$$\max_{c^i, \theta^i} E^i \left[ \int_0^T \log (c^i(t)) \, dt \right]$$

s. t. $dW^i(t) = [W^i(t)r(t) - c^i(t)] \, dt + \{ \theta^i_S(t) [\mu^i_S(t) - r(t)] + \theta^i_P(t) [\mu^i_P(t) - r(t)] \} \, dt$

$$+ [\theta^i_S(t)\sigma_S + \theta^i_P(t)\sigma_P] \, dw^i(t)$$

and (3.1)

This differs from the standard frictionless investor’s dynamic optimization problem due to the position limits, potential redundancy and mispricing between the risky assets $S$ and $P$. The

---

6Defining $P$ by its dividend process $\delta_P$ would not entail major changes to our analysis. All of the expressions in the paper would remain valid, but $\sigma_P$ would be endogenously determined, via a present value formula. Our main points and intuition would be unaffected.

7Assuming more constraints for the investors (for example, a lower bound on holdings of the derivative) would only complicate the model without substantially affecting our analysis, with extra equilibrium cases where no arbitrage activity takes place.
solution of the investor $i$’s problem is reported in Proposition 3.1. Those cases where $P$ is favorable are ignored, as they cannot arise in equilibrium.

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

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

when there is no mispricing,

$$\frac{\theta_S^i(t)}{W^i(t)} = \begin{cases} \beta, & \text{if } \frac{\mu_S^i(t) - r(t)}{\sigma_S} > \frac{\mu_P^i(t) - r(t)}{\sigma_P} \\ \bar{\beta}, & \text{if } -\gamma \sigma_P + \bar{\beta} \sigma_S > \frac{\mu_P^i(t) - r(t)}{\sigma_P} \end{cases}, \quad (a)$$

and when there is mispricing with $S$ being more favorable,

$$\frac{\theta_S^i(t)}{W^i(t)} = \begin{cases} \frac{\mu_S^i(t) - r(t)}{\sigma_S} + \frac{\sigma_P}{\sigma_S} \gamma & \text{if } -\gamma \sigma_P + \bar{\beta} \sigma_S \geq \frac{\mu_P^i(t) - r(t)}{\sigma_P} \\ 0 & \text{if } -\gamma \sigma_P > \frac{\mu_P^i(t) - r(t)}{\sigma_P} \end{cases}, \quad (c)$$

$$\frac{\theta_P^i(t)}{W^i(t)} = \begin{cases} \frac{\mu_P^i(t) - r(t)}{\sigma_P} - \frac{\sigma_S}{\sigma_P} \beta & \text{if } (c) \\ -\gamma & \text{if } (d) \\ -\gamma & \text{if } (e) \\ -\gamma & \text{if } (f) \end{cases}, \quad (d)$$

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 market price of risk is low enough, the investor binds on his lower position limit of both risky assets (case (b)), otherwise he is indifferent between the same risky assets provided his position is feasible (case (a)).

Under mispricing, the investor takes on as much of a riskless, costless arbitrage position as allowed by the position limits, hence uniquely determining the allocation between the derivative and the technology. When the risky technology is favorably mispriced and the investor is relatively optimistic about both $S$ and $P$ (case (c)), he desires a high risk exposure, wishing to go long in both risky assets, and so ends up binding on the upper position limit of $S$. Otherwise, his relative pessimism about the derivative drives him to his position limit on $P$; when additionally pessimistic about $S$, he desires a low risk exposure and so does not invest in the technology (case (f)), otherwise he goes long in the technology (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. Denoting his (dollar) investments by $\theta^3 \equiv (\theta^3_B, \theta^3_S, \theta^3_P)$, his problem can be expressed as follows:

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

s.t. $\Psi^3(t) dt = \left\{ \theta^3_S(t) \mu^3_S(t) + \theta^3_P(t) \mu^3_P(t) + \theta^3_B(t) r(t) \right\} dt$

$$+ [\theta^3_S(t) \sigma_S + \theta^3_P(t) \sigma_P] dw^3(t)$$

$$\theta^3_S(t) + \theta^3_P(t) + \theta^3_B(t) = 0$$

(4.1)

$$\theta^3_S(t) \sigma_S + \theta^3_P(t) \sigma_P = 0$$

(4.2)

$$0 \leq \theta^3_S(t) \leq M$$

(4.3)

Conditions (4.1)-(4.2) imply:

$$\theta^3_P(t) = -\frac{\sigma_S}{\sigma_P} \theta^3_S(t), \quad \theta^3_B(t) = \left(\frac{\sigma_S}{\sigma_P} - 1\right) \theta^3_S(t)$$

(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 $\theta^3_S > 0$ the arbitrageur receives a profit that is proportional to $\theta^3_S$. The optimal policy is then to choose the maximal possible value for $\theta^3_S$, namely $M$. From (4.4), the other holdings are given by $\theta^3_P(t) = -M \sigma_S / \sigma_P$ and $\theta^3_B(t) = M (\sigma_S / \sigma_P - 1)$, and the arbitrageur’s instantaneous profit rate by:

$$\Psi^3(t) = \theta^3_S(t) \mu^3_S(t) + \theta^3_P(t) \mu^3_P(t) + \theta^3_B(t) r(t) = M \sigma_S \Delta_{S,P}(t).$$

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

**Definition 4.1 (Competitive Equilibrium).** An equilibrium is a price system \((r, \mu_1, \mu_2)\) and consumption-portfolio processes \((c, \theta)\) such that: (i) agents choose their optimal consumption-portfolio strategies given their beliefs; (ii) security markets clear, i.e.,

\[
\theta^1_P(t) + \theta^2_P(t) + \theta^3_P(t) = 0, \quad \theta^1_B(t) + \theta^2_B(t) + \theta^3_B(t) = 0. \tag{4.5}
\]

We now proceed to the characterization of equilibrium. Quantities of interest will be expressed as a function of two (endogenous) state variables: the aggregate wealth \(W \equiv W^1 + W^2\) (coinciding, in equilibrium, with the total amount invested in the production technology), 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.** If equilibrium exists, the investors’ optimization cases, equilibrium mispricing and interest rate, and distribution of wealth dynamics are as follows.

When \(\hat{\mu}(t) \leq \frac{\gamma \sigma_P}{\lambda(t)} + \sigma_S \min \left\{ \frac{1}{\lambda(t)}, \frac{1}{1-\lambda(t)} \left[ (\beta - 1) + \frac{M}{\lambda(t)W(t)} \right] \right\} \), agents are in \((a,a)\) and

\[
\Delta_{S,P}(t) = 0 \quad \text{and} \quad r(t) = \lambda(t) \mu^1_S(t) + (1-\lambda(t)) \mu^2_S(t) - \sigma_S^2,
\]
\[
d\lambda(t) = \lambda(t)(1-\lambda(t))^2 (\hat{\mu}(t))^2 dt + \lambda(t)(1-\lambda(t))\hat{\mu}(t)dw^1(t). \tag{4.6}
\]

When \(\hat{\mu}(t) > \frac{\gamma \sigma_P + \sigma_S}{\lambda(t)} \) and \(\lambda(t) \geq \frac{1}{\beta} \left( 1 - \frac{M}{W(t)} \right) \), agents are in \((a,b)\) and

\[
\Delta_{S,P}(t) = 0 \quad \text{and} \quad r(t) = \mu^1_S(t) - \sigma_S^2 - \left( \frac{1}{\lambda(t)} - 1 \right) \sigma_S (\sigma_S + \gamma \sigma_P),
\]
\[
d\lambda(t) = \frac{(1-\lambda(t))^2}{\lambda(t)} \left( \sigma_S + \gamma \sigma_P \right) dt + (1-\lambda(t)) \left( \sigma_S + \gamma \sigma_P \right) dw^1(t).
\]

When \(\hat{\mu}(t) > \frac{\gamma \sigma_P}{\lambda(t)} + \frac{\sigma_S}{1-\lambda(t)} \left[ (\beta - 1) + \frac{M}{\lambda(t)W(t)} \right] \) and \(\lambda(t) < \frac{1}{\beta} \left( 1 - \frac{M}{W(t)} \right) \), agents are in \((c,e)\) and

\[
\Delta_{S,P}(t) = \hat{\mu}(t) - \frac{\sigma_S}{1-\lambda(t)} -(\frac{\gamma \sigma_P}{\lambda(t)} - \frac{\sigma_S M}{\lambda(t)(1-\lambda(t))W(t)}) > 0,
\]
\[ r(t) = \mu_S^2(t) - \sigma_S^2 + \Sigma \sigma_S \sigma_P + \sigma_S^2 \left[ (\beta - 1) \frac{\lambda(t)}{1-\lambda(t)} + \frac{M}{(1-\lambda(t))W(t)} \right], \]
\[ d\lambda(t) = \lambda(t) \left\{ \sigma_S \Delta_{S,P}(t) \left[ (\beta - 1) + \frac{M}{W(t)} \right] + [(1 - \lambda(t)) (\bar{\mu}(t) - \Delta_{S,P}(t))] \right\} dt \]
\[ + \lambda(t)(1 - \lambda(t)) (\bar{\mu}(t) - \Delta_{S,P}(t)) dw^1(t). \]

(4.8)

In all cases, the aggregate wealth dynamics follow
\[ dW(t) = \left[ W(t) \left( \mu_S(t) - \frac{1}{T-t} \right) - M \sigma_S \Delta_{S,P}(t) \right] dt + W(t) \sigma_S dw^1(t) \]
and investors 1 and 2’s consumption and the arbitrageur’s profit are given by:
\[ c^1(t) = \frac{\lambda(t)W(t)}{T-t}, \quad c^2(t) = \frac{(1 - \lambda(t))W(t)}{T-t}, \quad \Psi^3(t) = M \sigma_S \Delta_{S,P}(t). \]

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. In (a,a), there is only moderate divergence in beliefs across investors and so no constraints are binding. Heterogeneity is not sufficient for the investors to deviate very much from the no-trade case (where both of them would invest all of their wealth in the technology) and hit the constraints. Accordingly, the economic quantities are as in an unconstrained economy. When heterogeneity in beliefs is high enough, investors find it optimal to trade more, and some constraints are binding (regions (a,b) and (c,e)). The equilibrium with mispricing (c,e) occurs when the more optimistic investor (1) is insufficiently wealthy relative to the pessimistic investor (2) (low \( \lambda \)). Then, absent the mispricing the pessimistic investor would only invest a small amount in the technology and, due to his upper bound on investments in the technology, the optimistic investor would be unable to invest enough for market clearing to obtain (which requires that the whole aggregate wealth be invested in the technology: combining both clearing conditions in the definition of equilibrium ((4.5)) yields: \( \theta_S^1 + \theta_S^2 = W \)). The mispricing, with \( S \) being favorably mispriced, increases investor 2’s demand therein and so allows for market clearing. The intuition for the occurrence of mispricing is elaborated on further in Basak and Croitoru (2000). When, in contrast, agent 1 is sufficiently wealthy, it is possible for him to invest enough in the technology without violating his upper bound, and so it is not necessary to increase agent 2’s investment, and equilibrium (a,b) (without mispricing) obtains.

We henceforth concentrate on region (c,e), where the arbitrageur has an economic impact and so affects the characterization of economic quantities. In region (a,a), the equilibrium is as in an unconstrained economy as analyzed by Detemple and Murthy (1994), 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 available. It is straightforward to check that region (c,e) does occur for plausible parameter values: for example, according to (4.7), if \( \gamma = 0.5, \beta = 1.5, \sigma_S = \sigma_P = 0.1, M/W(t) = 0.1, \lambda(t) = 0.5 \), region (c,e) occurs whenever investor 1’s estimate of the mean instantaneous return of the technology exceeds investor 2’s estimate by
at least 2.4% per year. In our analysis of economic quantities, 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 ($\lambda$). 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. On moving from region (a,a) to region (c,e), the mispricing starts at a value of zero and then increases linearly with heterogeneity ($\bar{\mu}$) while inter-agent 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: $\Delta_{I_{S,P}} < \Delta_{S,P} < \Delta_{II_{S,P}}$. 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 size of the arbitrageur’s position only matters via its ratio to aggregate wealth, because what matters is the size of the counterparty he can offer to the investors, relative to their own wealths. This generates a dependence of the mispricing with respect to aggregate wealth: the higher aggregate wealth is, the less important the arbitrageur is relative to the other agents and so the less valuable the improvement in risk-sharing he generates. This increases the value of the mispricing and the arbitrageur’s profits, which are thus shown to be procyclical.

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.

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 (4.7), 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 $\bar{\beta}$ or $\bar{\gamma}$ (i.e., an alleviation of the constraints.)

All of this suggests that the arbitrageur alleviates the effect of the portfolio constraints. The effect of his presence on the investors’ welfare is ambiguous, however, because his profits reduce the growth of the investors’ aggregate wealth ((4.9)).
4.3. The Economic Role of the Arbitrageur

To clarify the role of the arbitrageur, we examine 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 technology).

In Economy III, agents do not trade and invest all of their wealth in the technology ($\theta_i^S = W^i$).

We introduce $\Phi^i$, investor $i$’s measure of risk exposure, defined as follows:

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

It is straightforward to check that the diffusion of investor $i$’s wealth is given by $\Phi^i W^i$. In the absence of trade (Economy III), all of agent $i$’s wealth is invested in the technology, and so $\Phi_{III}^i = \sigma_S$. 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:

$$
\Phi_1^I(t) = \sigma_S + (1 - \lambda(t))\bar{\mu}(t), \quad \Phi_2^I(t) = \sigma_S - \lambda(t)\bar{\mu}(t).
$$

In the (c,e) equilibrium,

$$
\Phi_1^I(t) = \beta \sigma_S + \frac{\gamma \sigma_P}{\lambda(t)} - \frac{\sigma_S M}{\lambda(t) W(t)}, \quad \Phi_2^I(t) = \sigma_S \frac{1}{\lambda(t)} + \beta \sigma_S - \frac{\sigma_S M}{\lambda(t) W(t)}.
$$

In the presence of constraints, once the amount of heterogeneity ($\bar{\mu}$) is sufficient 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,

$$
\Phi_{III}^I(t) < \Phi_{II}^I(t) < \Phi_1^I(t), \quad \Phi_{III}^I(t) > \Phi_{II}^I(t) > \Phi_2^I(t).
$$

Thus, the presence of the arbitrager 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 $\bar{\mu}$. In our equilibrium (c,e),

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

(4.10)
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. 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 \( \Phi_i \)’s. In the absence of the arbitrageur (benchmark economy II), investors’ risk exposures are:

\[
\Phi_{1I}^I(t) = \beta \sigma_S + \gamma \sigma_P \frac{1 - \lambda(t)}{\lambda(t)}, \quad \Phi_{2I}^I(t) = \sigma_S - \gamma \sigma_P - \left( \beta - 1 \right) \sigma_S \frac{\lambda(t)}{1 - \lambda(t)}.
\]

Even though the optimistic investor 1 would optimally increase his risk exposure (to \( \Phi_1^I \)) by increasing his holding of the derivative (since he already is binding on the technology), 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 \( \Phi_1 \) (by \( \sigma_S M / \lambda W \)) 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 clearing in the bond market (which requires \( \theta_1^1 + \theta_2^1 = W \)) would not obtain. 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 \( \sigma_S M / [(1 - \lambda)W] \)). The introduction of the arbitrageur in the economy is equivalent to that of an extra, zero-net supply security, 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. The difference (the mispricing) generates the arbitrageur’s profit.

Our understanding of the role of the arbitrageur allows us to provide additional insight on some of the characterizations in Proposition 4.1. The distribution of wealth \( \lambda \) dynamics is affected by the presence of the arbitrageur. When no constraints are binding, the volatility of investor 1’s share of aggregate wealth is given by \( \lambda(1 - \lambda)\bar{\mu} \) (from equation (4.6)): the more optimistic investor 1 is (i.e., the higher \( \bar{\mu} \)), the more positively his wealth covaries with the technology return. In region (c,e) of our constrained economy, due to the binding constraints investors cannot trade as much risk as in the unconstrained case and accordingly the volatility of \( \lambda \) is reduced to \( \lambda(1 - \lambda)(\bar{\mu} - \Delta_{S,P}) \) (from (4.8)). In the presence of an arbitrageur, the value of the mispricing is reduced; this increases the covariance of investor 1’s wealth with the technology return, thanks to improved risk-sharing.

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 position limits 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 increases the risk-bearing capacity in an economy and hence improves risk-sharing.\(^8\) 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 \(i = 3\ast\) and, accordingly, denote his wealth by \(W^{3\ast}\) and his estimate of \(\mu_S\) by \(\mu_S^{3\ast}\). (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 \((3\ast)\), but no arbitrageur, if

\[
\mu(t) > \sigma_S \left( \beta - 1 \right) \frac{1 + \frac{W^{3\ast}(t)}{W(t)}}{1 - \lambda(t)} + \gamma \sigma_P \frac{1 + \frac{W^{3\ast}(t)}{W(t)}}{\lambda(t) + \frac{W^{3\ast}(t)}{W(t)}} + \frac{\mu_S(t) - \mu_S^{3\ast}(t)}{\sigma_S} \frac{W^{3\ast}(t)}{W(t)},
\]

\[-\gamma \sigma_P \left( 1 + \frac{W(t)}{W^{3\ast}(t)} \right) \leq \frac{\mu_S(t) - \mu_S^{3\ast}(t)}{\sigma_S} \leq \frac{1 + \frac{W^{3\ast}(t)}{W(t)}}{\gamma \sigma_P \lambda(t)} \quad \text{and} \frac{1 - \lambda(t)}{\lambda(t) + \frac{W^{3\ast}(t)}{W(t)}} \geq \beta - 1,
\]

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

\[\Phi^{1\ast}(t) = \beta \sigma_S + \gamma \sigma_P \frac{1 - \lambda(t)}{\lambda(t) + \frac{W^{3\ast}(t)}{W(t)}} + \frac{\mu_S(t) - \mu_S^{3\ast}(t)}{\sigma_S} \frac{W^{3\ast}(t)}{\lambda(t)W(t) + W^{3\ast}(t)},\]

\[\Phi^{2\ast}(t) = \sigma_S - \gamma \sigma_P - \left( \beta - 1 \right) \frac{\lambda(t) + \frac{W^{3\ast}(t)}{W(t)}}{1 - \lambda(t)}.
\]

Investors’ equilibrium risk exposures reveal that a third risk-taking investor affects risk-sharing differently than the arbitrageur does. While \(\Phi^7 < \Phi^{2\ast} < \Phi^2_{II}\), implying that the presence of the extra investor \(3\ast\) makes investor 2’s risk exposure closer to its value in an unconstrained economy \((\Phi^2)\), the effect on investor 1 is ambiguous. Depending on investor \(3\ast\)’s beliefs, \(\Phi^{1\ast}\) can be either higher or lower than if the third investor were not present. Accordingly, the effect of investor \(3\ast\) on the amount of risk shared between investors 1 and 2 is ambiguous. Our measure of risk-sharing is now equal to

\[\Phi^{1\ast}(t) - \Phi^{2\ast}(t) = \left( \beta - 1 \right) \sigma_S \frac{1 + \frac{W^{3\ast}(t)}{W(t)}}{1 - \lambda(t)} + \gamma \sigma_P \frac{1 + \frac{W^{3\ast}(t)}{W(t)}}{\lambda(t) + \frac{W^{3\ast}(t)}{W(t)}} + \frac{\mu_S(t) - \mu_S^{3\ast}(t)}{\sigma_S} \frac{1}{1 + \lambda(t) \frac{W(t)}{W^{3\ast}(t)}}.
\]

\(^8\)For example, in a standard economy the risk aversion of the representative agent \(A\) satisfies \(1/A = \Sigma_i (1/A^i)\) and so decreases when the number of agents in the economy increases, ceteris paribus. Hence, so does the market price of risk, \(A \sigma_S\), where \(\sigma_S\) is the volatility of aggregate consumption, and the individual consumption volatilities, \(A \sigma_i/A^i\).
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 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: \( \gamma = 0.5 \), \( \beta = 1.5 \), \( \sigma_S = \sigma_P = 0.1 \), \( \lambda(t) = 0.5 \), \( W^{3*}(t)/W(t) = 0.4 \), \( \mu_1(t) = 0.1 \), \( \mu_2(t) = 0.05 \), \( \mu_{3*}(t) = 0.105 \). 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 (4.10)), 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 maximum trade allowed by the position limit. The importance of this non-competitive case is further underscored by the fact that position limits or any other 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.

We now formulate the non-competitive arbitrageur’s optimization problem. Given his market power, the arbitrageur will act as a price leader in the securities markets. He observes other investors’ (of Section 3) demands as functions of prices, and then chooses the size of his own
trades so as to maximize his concurrent profits subject to the condition that the security 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:

\[
\Delta_{S,P}(t) = \bar{\mu}(t) - \frac{\sigma_S (\beta - 1)}{1 - \lambda(t)} - \frac{\gamma \sigma_P}{\lambda(t)} - \frac{\sigma_S \theta^3_S(t)}{\lambda(t)(1 - \lambda(t))} W(t). \tag{5.1}
\]

Hence, the arbitrageur’s influence on security prices manifests itself, via (5.1), as the size of his stock position 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 \( t \):

\[
\max_{\theta^3_S(t), \Delta_{S,P}(t)} \Psi^3(t) = \theta^3_S(t)\sigma_S \Delta_{S,P}(t) \tag{5.2}
\]

s.t. \( \theta^3_S(t) + \theta^3_P(t) + \theta^3_B(t) = 0, \quad \theta^3_S(t)\sigma_S + \theta^3_P(t)\sigma_P = 0, \quad ((4.1)-(4.2)) \)

and \( \Delta_{S,P}(t) \) satisfies (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

**Definition 5.1 (Non-competitive Equilibrium).** Equilibrium in an economy with two risk-averse investors and one non-competitive arbitrageur is a price system \((r, \mu^1_P, \mu^2_P)\) and consumption-portfolio policies \((c^i, \theta^i)\) such that: (i) investors choose their optimal policies given the price system \((r, \mu^1_P, \mu^2_P)\) and their beliefs; (ii) the arbitrageur chooses his optimal trades in (5.2) taking into account that the mispricing responds to clear the security markets; (iii) the price system is such that all the security markets clear.

According to the definition of a non-competitive equilibrium, the non-competitive arbitrageur simultaneously solves for the optimal size of his arbitrage trades and the equilibrium mispricing. Proposition 5.1 reports the possible equilibrium cases and the characterization for each case.
Proposition 5.1. If equilibrium exists, equilibrium with a non-competitive arbitrageur is as follows.

When \( \bar{\mu}(t) \leq \frac{\gamma \sigma_P}{\lambda(t)} + \sigma_S \min \left\{ \beta \frac{1}{\lambda(t)}, \beta \frac{1}{1-\lambda(t)} \right\} \), agents are in (a,a), and when \( \bar{\mu}(t) > \frac{\gamma \sigma_P + \sigma_S}{\lambda(t)} \) and \( \lambda(t) \geq \frac{1}{\beta} \), agents are in (a,b). In both cases, \( \Delta_{S,P}(t) \), \( r(t) \) and \( \lambda(t) \) are as in the competitive arbitrage equilibrium described in Proposition 4.1.

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

\[
\begin{align*}
\theta_S^2(t) & = \frac{\lambda(t)(1 - \lambda(t))W(t)}{2\sigma_S} \left[ \bar{\mu}(t) - \frac{\sigma_S (\beta - 1)}{1 - \lambda(t)} - \frac{\gamma \sigma_P}{\lambda(t)} \right], \\
\psi^3(t) & = \frac{\lambda(t)(1 - \lambda(t))W(t)}{4} \left[ \bar{\mu}(t) - \frac{\sigma_S (\beta - 1)}{1 - \lambda(t)} - \frac{\gamma \sigma_P}{\lambda(t)} \right]^2,
\end{align*}
\]

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

\[
\begin{align*}
\Delta_{S,P}(t) & = \frac{1}{2} \left[ \bar{\mu}(t) - \frac{\sigma_S (\beta - 1)}{1 - \lambda(t)} - \frac{\gamma \sigma_P}{\lambda(t)} \right], \\
r(t) & = \mu_S^2(t) - \sigma_S^2 + \sigma_S \bar{\mu}(t) \frac{\lambda(t)}{2} + \frac{1}{2} \frac{\gamma \sigma_P}{1 - \lambda(t)} \sigma_S^2 (\beta - 1), \\
d\lambda(t) & = \lambda(t) \left\{ \sigma_S \Delta_{S,P}(t) \left[ \frac{\bar{\mu}(t) - \Delta_{S,P}(t)}{W(t)} \right]^2 + [(1 - \lambda(t)) (\bar{\mu}(t) - \Delta_{S,P}(t))] \right\} dt \\
&= \lambda(t)(1 - \lambda(t)) (\bar{\mu}(t) - \Delta_{S,P}(t)) d\lambda^1(t).
\end{align*}
\]

In all cases, the aggregate wealth dynamics follow

\[
dW(t) = \left[ W(t) \left( \mu_S^2(t) - \frac{1}{T - t} \right) - \theta_S^2(t) \sigma_S \Delta_{S,P}(t) \right] dt + W(t) \sigma_S dw^1(t).
\]

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 in equilibrium. 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.\(^9\) As a result, unlike in the competitive case, the conditions for regions (c,e) are as if the arbitrageur were not present.

\(^9\)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.
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, \( \Delta_{S,P}(t) = \frac{1}{2}\Delta_{II,S,P}(t) \). 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 \( \theta_3^S \) so the profit is a quadratic function of the arbitrageur’s position. Therefore, it is maximized at a position equidistant from 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 \( \bar{\mu} \) and the state variables \( W, \lambda \). For sufficiently low investor heterogeneity, the non-competitive arbitrageur’s stock position is lower than in the competitive case \( (\theta_3^S(t) < M) \), and higher for high investor heterogeneity. Hence, for low investor heterogeneity, 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 the transfer of risk amongst investors, is reduced. The opposite holds when there is high heterogeneity 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 exogenous position limit \( M \) is replaced with the endogenous \( \theta_3^S \). Finally, in contrast to the linear competitive case, the non-competitive arbitrageur’s profits are convex in the extent of heterogeneity 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 technology, derivative and bond. Assuming that \( r(t) > 0 \), 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 \( \theta_C^3 \) and \( W^3 \) denote the arbitrageur’s cash holding and capital, respectively,
his holdings must obey:

\[ W^3(t) \geq \eta \max \left\{ -\theta^3_P(t), 0 \right\} + \eta \left( \theta_S(t) + \max \left\{ \theta^3_P(t), 0 \right\} \right), \quad (6.1) \]

\[ \theta^3_C(t) \geq \left( 1 + \epsilon \eta \right) \max \left\{ -\theta^3_P(t), 0 \right\}, \quad (6.2) \]

where \( \eta, \epsilon \in [0, 1] \) and \( \theta^3_S(t) \geq 0 \) (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 \( \epsilon \) 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)).

For simplicity, we assume that there is no difference between initial and maintenance margin. Denoting \( \theta^3 \equiv (\theta^3_B, \theta^3_S, \theta^3_P, \theta^3_C) \), the arbitrageur’s problem is as follows:

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

s.t. \( \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) \mu^3_B(t) \right\} dt 
+ [\theta^3_S(t) \sigma_S + \theta^3_P(t) \sigma_P] dw^3(t), \]
\[ \theta^3_S(t) + \theta^3_P(t) + \theta^3_B(t) + \theta^3_C(t) = W^3(t), \quad (6.3) \]
\[ \theta^3_S(t) \sigma_S + \theta^3_P(t) \sigma_P = 0 \quad (6.4) \]

and (6.1)-(6.2) hold.

Both the amount of arbitrage capital \( (W^3) \) and its time \( t \)-instantaneous return \((\psi^3(t)/W^3(t)) \) dt\) 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 \( \theta^3_P(t) \):

\[ \theta^3_S(t) = -\frac{\sigma_P}{\sigma_S} \theta^3_P(t), \quad \theta^3_B(t) = \left( \frac{\sigma_P}{\sigma_S} - 1 \right) \theta^3_P(t) - \left( 1 + \epsilon \eta \right) \max \left\{ -\theta^3_P(t), 0 \right\}, \quad (6.5) \]

\[ \theta^3_C(t) = \left( 1 + \epsilon \eta \right) \max \left\{ -\theta^3_P(t), 0 \right\}. \quad (6.6) \]

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 \( \theta^3_P(t) \geq 0 \); 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

\[ \text{Our modeling of margin requirements could easily be amended without affecting our main intuition. All that is really needed is the presence of a constraint that limits the size of the arbitrageur’s position in proportion to his capital, and that of costly short sales.} \]
arbitrageur to take on as large a profitable arbitrage position (long in the favorably mispriced stock, and short in the unfavorable derivative) as feasible; hence, (6.1) holds with equality:

\[ \eta \theta^3_S(t) - \eta \theta^3_P(t) = W^3(t), \]

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

\[ \theta^3_P(t) = -\frac{W^3(t)}{\eta + \frac{\sigma_P}{\sigma_S}} < 0, \quad \theta^3_S(t) = \frac{W^3(t)}{\eta + \frac{\sigma_P}{\sigma_S}} > 0, \]

(6.7)

\[ \theta^3_B(t) = W^3(t) \left[ 1 - \frac{\epsilon \eta + \sigma_P}{\eta + \frac{\sigma_P}{\sigma_S}} \right], \quad \theta^3_C(t) = \frac{\left( 1 + \epsilon \eta \right) W^3(t)}{\eta + \frac{\sigma_P}{\sigma_S}}. \]

(6.8)

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

\[ \frac{\Psi^3(t)}{W^3(t)} = r(t) + \left[ \sigma_P \Delta S, P(t) - r(t) \left( 1 + \epsilon \eta \right) \right] \frac{1}{\eta + \frac{\sigma_P}{\sigma_S}}. \]

(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 \( \theta^i_A \) investor \( i \)'s (dollar) investment in the arbitrage firm, his dynamic budget constraint is now:

\[ dW^i(t) = \left[ W^i(t) r(t) - c^i(t) \right] dt + \left\{ \theta^i_S(t) \left[ \mu^i_S(t) - r(t) \right] + \theta^i_P(t) \left[ \mu^i_P(t) - r(t) \right] \right\} dt \]

\[ + \left[ \theta^i_S(t) \sigma_S + \theta^i_P(t) \sigma_P \right] dw^i(t) + \theta^i_A(t) \frac{\Psi^i(t)}{W^3(t)} dt, \]

while the following condition is added to the clearing conditions in the definition of equilibrium (4.5):

\[ \theta^1_A(t) + \theta^2_A(t) = W^3(t). \]

For simplicity, the investors’ holdings of the arbitrage firm are assumed to be unconstrained. 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

\[ \bar{\mu}(t) > \frac{\gamma \sigma_P}{\lambda(t)} + \frac{(\beta - 1) \sigma_S}{1 - \lambda(t)} + \frac{\sigma_S W^3(t)}{\lambda(t)(1 - \lambda(t)) W(t)} \left( \frac{\lambda(t)\left(1 + \epsilon \eta\right)\sigma_S + \sigma_P}{\eta \sigma_S + \eta \sigma_P} \right) \]

and

\[ 0 \leq W^3(t) \leq \lambda(t) W(t) \left( \frac{\eta \sigma_P + \eta \sigma_S}{\sigma_P + (1 + \epsilon \eta) \sigma_S} \right) \left[ 1 - \frac{\lambda(t)}{1 - \lambda(t)} (\beta - 1) \right], \]

where

\[ W^3(t) = \frac{\sigma_P \bar{\mu}(t) - (1 + \epsilon \eta) \mu_S^2(t) - \frac{\sigma_P \sigma_S (\beta - 1)}{1 - \lambda(t)} - \frac{\eta \sigma_P^2}{\lambda(t)} - (1 + \epsilon \eta) \sigma_S \left[ \frac{\sigma_S \left(1 + \frac{\lambda(t) (\beta - 1)}{1 - \lambda(t)}\right) + \sigma_P}{\eta \sigma_S + \eta \sigma_P} \right]}{\frac{\sigma_S}{\eta \sigma_P + \eta \sigma_S} \left[ \frac{(1 + \epsilon \eta) \sigma_S}{1 - \lambda(t)} W(t) \left( 2 \sigma_P + (1 + \epsilon \eta) \sigma_S \right) + \frac{\sigma_P}{\lambda(t)(1 - \lambda(t)) W(t)} \right]} \]

(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:

\[ \Delta_S P(t) = \bar{\mu}(t) - \frac{\sigma_S (\beta - 1)}{1 - \lambda(t)} - \frac{\sigma_S W^3(t)}{\lambda(t)(1 - \lambda(t)) W(t)} \left( \frac{\lambda(t)\left(1 + \epsilon \eta\right)\sigma_S + \sigma_P}{\eta \sigma_S + \eta \sigma_P} \right) > 0, \]

(6.11)

\[ r(t) = \mu_S^2(t) - \sigma_S^2 + 2 \sigma_S \sigma_P + \sigma_S^2 \left[ \frac{\beta - 1}{1 - \lambda(t)} + \frac{W^3(t)}{(1 - \lambda(t)) W(t)} \left( \frac{\sigma_P + (1 + \epsilon \eta) \sigma_S}{\eta \sigma_P + \eta \sigma_S} \right) \right], \]

(6.12)

\[ d\lambda(t) = \left\{ \sigma_S \Delta_S P(t) \left( \beta - \lambda(t) \right) + (\sigma_S + \sigma_{\lambda}(t)) \left[ \sigma_S (1 - 2 \lambda(t)) + \sigma_{\lambda}(t) \right] + \lambda(t) \sigma_S^2 \right\} dt + \lambda(t) \sigma_{\lambda}(t) dw^1(t), \]

where \( \sigma_{\lambda}(t) = (1 - \lambda(t)) (\bar{\mu}(t) - \Delta_S P(t)) - \frac{W^3(t) (1 + \epsilon \eta) \sigma_S^2}{W(t) (\eta \sigma_S + \eta \sigma_P)}, \]

\[ dW(t) = \left\{ W(t) - (1 + \epsilon \eta) \left[ \frac{\gamma (1 - \lambda(t)) W(t) + W^3(t)}{\eta + \eta \sigma_S} \right] \left( \mu_S^2(t) dt + \sigma_S dw^1(t) \right) - \frac{W(t) T - t}{T - t} dt. \]

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 \( \sigma_P \Delta_S P(t) = r(t) \left(1 + \epsilon \eta\right), \) 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 \( W^3(t) \) 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, \( \Delta_{S,P}^{H} = \bar{\mu} - \sigma_S(\beta - 1)/(1 - \lambda) - \gamma \sigma_P / \lambda, \) 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. As a result, the amount of arbitrage profits is convex in investor heterogeneity, as in the non-competitive case. In addition, the amount of arbitrage capital is increasing in the severity of
the margin constraint (6.1) (as measured by \( \pi \sigma_P + \eta \sigma_S \)): 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 \( \theta^3_S = \sigma_P W^3 / \left( \eta \sigma_S + \pi \sigma_P \right) \). Nonetheless, risk-sharing between investors is further impacted by the cash constraint. This is evidenced by the diffusion of the wealth distribution dynamics, \( \sigma \lambda \), 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:

\[
\Phi^1(t) = \sigma_S + \gamma \sigma_P - \frac{1 - \lambda(t)}{\lambda(t)} + \frac{\sigma_S \sigma_P W^3(t)}{\lambda(t) W(t) \left( \pi \sigma_P + \eta \sigma_S \right)}.
\]

\[
\Phi^2(t) = \sigma_S - \frac{(\beta - 1)}{1 - \lambda(t)} \sigma_S \lambda(t) - \frac{\sigma_S \left( \sigma_P + \sigma_S (1 + \epsilon \eta) \right) W^3(t)}{(1 - \lambda(t)) W(t) \left( \pi \sigma_P + \eta \sigma_S \right)}.
\] (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 \( \epsilon \)), 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 \( \theta^1_S + \theta^2_S + \theta^3_S = W^3 - \theta^C_\ell \). 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 constraint on stock investments 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 a percent increase in the amount of arbitrage activity by \( \sigma_S (1 + \epsilon \eta) / \sigma_P \). The effect of the margin constraint (6.1), on the other hand, is as would be expected: it is similar to that of the position limit 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 \( \theta^3_S \).
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 position limits, 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 and acts as a financial intermediary, who can be interpreted as an innovator synthesizing a derivative security specifically designed to relieve investors’ position limits. 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, investors’ position limits) 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.
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 derivative); (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 position limits 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), Cuoco and Cvitanic (1998)) 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 market clearing conditions (4.5), and noting that both investors facing the same price process for the derivative $P$ implies:

$$\frac{\mu_1^P(t) - \mu_2^P(t)}{\sigma_P} dt = dw^2(t) - dw^1(t) = \frac{\mu_1^S(t) - \mu_2^S(t)}{\sigma_S} dt,$$

hence $\frac{\mu_1^P(t) - \mu_2^P(t)}{\sigma_P} = \bar{\mu}(t)$,

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:

$$\sigma_S \frac{\theta_1^S(t)}{W^1(t)} + \sigma_P \frac{\theta_1^P(t)}{W^1(t)} = \sigma_S + \bar{\mu}(t)(1 - \lambda(t)), \quad (A.1)$$

$$\sigma_S \frac{\theta_2^S(t)}{W^2(t)} + \sigma_P \frac{\theta_2^P(t)}{W^2(t)} = \sigma_S - \bar{\mu}(t)\lambda(t), \quad (A.2)$$

$$\theta_3^P(t) = -\frac{\sigma_S}{\sigma_P} \theta_3^S(t), \quad (A.3)$$

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:

$$\theta_1^S(t) + \theta_2^S(t) + \theta_3^S(t) = W^1(t) + W^2(t), \quad \theta_1^P(t) + \theta_2^P(t) + \theta_3^P(t) = 0. \quad (A.4)$$

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, $\theta_1^P(t)$ and $\theta_3^S(t)$), since there are two degrees of freedom.
in the system, and substitute these expressions into the position limits. This results in eight inequalities that $\theta_1^t(t)$ and $\theta_3^t(t)$ 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 $\lambda$ 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 $\Phi^i$, 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 $\Phi^i$ then yields the expressions in the Proposition. Q.E.D.
References


