WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

HOLDING COSTS AND EQUILIBRIUM ARBITRAGE

Bruce TUCKMAN*
and
Jean-Luc VILA**

WP#3437 Revised December 1993

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139
HOLDING COSTS AND EQUILIBRIUM ARBITRAGE

Bruce TUCKMAN*
and
Jean-Luc VILA**

WP#3437 Revised December 1993

* Stern School of Business, New York University
** Sloan School of Management, Massachusetts Institute of Technology
HOLDING COSTS AND EQUILIBRIUM ARBITRAGE

Bruce TUCKMAN

*Stem School of Business, New York University, New York, NY 10006, USA*

Jean-Luc VILA

*Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

First version: June, 1992

This version: December 1993
Abstract

This paper constructs a dynamic model of the equilibrium determination of relative prices when risk averse arbitrages face holding costs. The major findings are as follows: 1) Arbitrages reduce but do not eliminate mispricings. 2) Because arbitrages optimally take positions when mispricings are within the riskless arbitrage bounds, models based on riskless arbitrage arguments alone may not provide usefully tight bounds on observed prices. 3) Arbitrages are least effective in eliminating the mispricings of long-term assets. 4) Arbitrage activity is most effective when exogenous shocks are transient and conditionally volatile. 5) Arbitrage activity reduces the conditional volatility of the mispricing process but may increase or decrease its mean reversion. Furthermore, in the presence of arbitrages, both conditional volatility and mean reversion are non-linear functions of the level of mispricing.

Acknowledgments: We would like to thank participants in the WFA annual meetings in Whisler, Paul Asquith, Anne Fremault Vila, John Heaton, Francis Longstaff, Jeremy Stein, and Sheridan Titman for helpful comments and suggestions. Also Jean-Luc Vila wishes to acknowledge financial support from the International Financial Services Research Center at the Sloan School of Management.
1. Introduction

In frictionless markets, the law of one price guarantees that securities with identical cash flows sell for the same price. In real markets, however, most agents face frictions of some sort. While the class of agents most active in enforcing the law of one price, namely arbitragers, face particularly low costs, they do face non-trivial costs. The many studies that document deviations from the law of one price\(^1\) testify to the proposition that the existence of arbitragers does not guarantee that market prices equal their fundamental values and, consequently, does not assure an optimal allocation of capital.

Of the many frictions that inhibit arbitrage activity, trading costs have received the most attention. In the simplest of models, the difference between the prices of two securities that have identical cash flows and finite maturities may not exceed the magnitude of the trading costs. Furthermore, arbitrage activity is easy to describe. When market prices do not admit riskless profit opportunities inclusive of trading costs, arbitragers do nothing. When market prices do admit such opportunities, arbitragers stake all they have on the sure proposition that prices will come back into line by the maturity date.

Because arbitrage behavior is not well described by this simplest of models, researchers have sought to enrich that model. Hodges and Neuberger (1989) examine the impact of trading costs when the arbitrage position must be periodically rebalanced. Brennan and Schwartz (1990) impose position limits and allow arbitragers to unwind positions optimally. Fremault (1991, 1993) studies the effects of non-synchronous trading across markets for otherwise identical securities. Holden (1990) explores the impact of oligopolistic behavior on the part of arbitragers.

---

\(^1\) In the case of government bonds, see, for example, Amihud and Mendelson (1991), Cornell and Shapiro (1989), and Tuckman and Vila (1992a). In the case of stock index futures, see, for example, Bacha and Fremault (1993), Chung, Kang, and Rhee (1992), MacKinlay and Ramaswamy (1988), Strickland and Xu (1991), and Pope and Yadav (1992). In the case of closed-end funds, see, for example, Pontiff (1993) and Lee, Shleifer, and Thaler (1991). In the case of scores and primes, see, for example, Jarrow and O'Hara (1989) and Canina and Tuckman (1993).
Another recent approach to enriching models of arbitrage behavior has been the introduction of holding costs or unit-time costs [see Tuckman and Vila (1992a)]. With holding costs, total costs increase with the time over which an arbitrage position is maintained and arbitrage behavior is quite different from that in the simplest of models described above. First, the no-riskless arbitrage condition requires that the difference between the prices of securities with identical cash flows be no greater than the present value of the accumulated holding costs from the time the position is established until the securities mature. Thus, riskless arbitrage bounds widen with maturity. Second, within these bounds, mispricing presents risky, but potentially profitable, investment opportunities. An arbitrager will profit from a long position in the relatively dear security and a short position in the relatively cheap security only if prices come into line soon enough; if prices diverge for too long, the accumulated holding costs will outweigh any realized profit.

Holding costs appear in many arbitrage contexts. First, shorting any spot security or commodity will often result in unit time costs since, for as long as the position is maintained, arbitragers must usually pay a borrowing fee, sacrifice the use of at least some of the short sale proceeds, or earn less than the market rate on posted collateral. Second, futures market positions may generate unit time costs since part of margin deposits may not earn interest. Third, banks making markets in forward contracts often charge a _per annum_ rate over the life of the contract. Fourth, when collateral requirements cause deviations from desired investment strategies, these requirements are essentially generating unit time costs.\(^2\)

Recent empirical work has begun to find that market mispricings may reflect the presence of these various incarnations of holding costs. Several papers find that deviations from fundamental values

\(^2\) See Tuckman and Vila (1992a) for references supporting these institutional details. Also see that paper for a detailed discussion of holding costs in the U.S. Treasury bond market.
increase with a security’s maturity. These findings are consistent with the previous discussion that, under holding costs, riskless arbitrage bounds widen with maturity. Other papers, e.g. Bacha and Fremault (1993) and Pontiff (1993), find that mispricings increase with the level of interest rates. When holding costs arise because of foregone interest, this result follows easily.

Tuckman and Vila (1992a), in focusing on the behavior of arbitragers facing holding costs, take the mispricing process to be exogenous. This paper allows the activity of arbitragers to feed back into the mispricing process and, therefore, allows for an examination of how arbitragers affect market mispricings. The major findings of the paper are as follows.

One, risk averse arbitragers facing holding costs reduce, but do not eliminate mispricings. Thus, mispricings and arbitrage activity are consistent with market equilibrium.

Two, because arbitragers take positions even when no riskless profit opportunities are available, equilibrium prices are kept well within the riskless arbitrage bounds. This implies that models based on riskless arbitrage arguments alone, like many in the option replication literature [see, for example, Boyle and Vorst (1992) and Leland (1985)], may not provide usefully tight bounds on observed prices.

Three, because total position costs increase with the holding period, arbitragers find it most difficult to eliminate mispricings of long-term assets. This finding serves as a microeconomic foundation for the recent literature on the connection between investor impatience and market mispricings. De Long, Shleifer, Summers, and Waldmann (1990) and Lee, Shleifer, and Thaler (1991) have hypothesized that investors’ short-term horizons allow for persistent deviations from fundamental values. Shleifer and Vishny (1990) argue that when mispricings increase in asset maturity, risk averse

---

3 See, for example, Tuckman and Vila (1992a) and MacKinlay and Ramaswamy (1988).
corporate managers will pursue short-term objectives.

Four, given the unconditional volatility of net liquidity shocks, arbitragers are most effective in eliminating mispricings that are transient and conditionally volatile. This finding may be of use in building and testing a theory to predict which markets will be well-arbitraged and which will not.

Five, arbitrage activity lowers the conditional volatility of the mispricing process but may increase or decrease its mean reversion. Furthermore, the conditional volatility and mean reversion of the mispricing process in the presence of arbitragers is highly non-linear in the level of mispricing. While these findings are consistent with recent studies that have found mean reversion in mispricing processes,⁴ and, in particular, with the Pope and Yadov (1992) conclusion that "mean reversion is being caused by trading activity linked directly or indirectly to arbitrage," this paper’s results indicate that regressions which assume a constant coefficient of mean reversion are most likely misspecified.

Section 2 presents the model. Section 3 provides numerical examples to illustrate the model’s implications. Section 4 concludes and suggests avenues for future research.

2. The Model

2.1 Preliminaries

Any model of equilibrium arbitrage must 1) posit the existence of two distinct assets which generate identical cash flows, 2) assume some forces which cause the prices of these assets to differ, and 3) restrict investor and arbitrage behavior so that mispricings do not vanish as soon as they appear.

Markets provide many examples of distinct portfolios which generate identical cash flows. Some common instances are a forward or futures contract vs. a levered position in the spot asset, a bond denominated in one currency vs. a bond denominated in another currency plus a cross-currency forward contract, and a coupon bond vs. a cash-flow matched portfolio of other coupon bonds.

This model assumes that there exist two distinct bond issues with identical coupons and maturities. For ease of exposition, one of the issues will consist of "red" bonds while the other will consist of "green" bonds. Not much effort would be required to recast the model in terms of the other mentioned examples.

The prices of the red and green bonds may tend to differ for a number of reasons. For example, the bonds might be traded by different clienteles with different valuation rules. Clienteles of this sort may have developed for historical reasons or may exist for institutional reasons. Another reason for price differences across these markets might be temporary supply and demand imbalances due to microstructure imperfections. In any case, because this paper aims at explaining the force which constrains these price differences, no serious effort has been made to model the source of price differences. Instead, the model assumes that noise traders occasionally shock the red and green bond markets with buy and sell orders. To ensure that these shocks affect market prices, it is furthermore assumed that the demand for the individual bond issues is not perfectly elastic.
Because close substitutes exist for most financial assets, the more usual assumption is that the demand for individual financial assets is perfectly elastic. But, in the presence of market frictions, considerations other than the existence of substitutes become important. For example, investors in different tax brackets value the same asset differently, thus generating a downward-sloping demand curve. Similarly, to the extent that an asset is purchased after the sale of other assets, investors with low transaction costs will buy at higher prices than investors with high transaction costs, again leading to a downward-sloping demand curve. This model employs the tax motivation because, as shown below, the functional form of the demand curve can be easily derived. But, any downward-sloping demand curve will translate liquidity shocks into price differences across markets.

Finally, price differences must not vanish as they appear. This certainly requires some segmentation of the markets; if investors in one bond market can easily purchase bonds in the other, price differences will result in migrations from the relatively dear market to the relatively cheap market. And these migrations will, in turn, equalize the red and green bond prices. While market segmentation seems a reasonable assumption from the point of view of many investors, arbitragers can usually trade across markets. Nevertheless, arbitrager activity might not be sufficient to force prices back into line. To this end, arbitragers are assumed to be risk-averse and to face holding costs when shorting bonds. As in Tuckman and Vila (1992a), these assumptions imply that optimal arbitrage positions are not necessarily large enough to eliminate price differences across markets.

2.2. Equilibrium pricing without arbitragers

Turning to the model of this paper, begin by assuming that there are two distinct bond issue, one red and one green, trading in two distinct markets. Both bond issues mature n periods from now and entitle holders to $d at the end of each period and to $1+d at maturity. Finally, let r denote the discount rate which is assumed constant through the maturity date.
The appendix derives demand curves for each of the bond issues by assuming that investors face different tax rates on coupon income and that those who buy bonds plan to hold their investments until maturity. The resulting demand curve for each issue can be written in the following functional form

\[ D_n(P_n) = \eta_n - \frac{P_n}{\theta_n} \quad n \geq 1 \]

where

\[ \theta_n = \frac{d}{ar} (1 - \frac{1}{(1+r)^n}) \]

and \( \alpha \) denotes the number of investors in each market. Note that demand elasticity falls as the maturity date approaches. This property reflects the general fact that valuation differences across investors become less and less important as the security matures. In particular, when \( n = 0 \), i.e. at maturity, the demand curve is perfectly elastic: a security that immediately pays $1 and nothing thereafter must sell for $1.

Assume that the quantity of each bond issue available for investor trading is \( Q \). In the absence of noise traders, the market price in each bond market would be the \( P_n \) that solves \( D_n(P_n) = Q \). The presence of noise traders, however, can cause the prices of the red and green bonds to differ.

Noise traders enter the markets to sell or to buy. Let \( L_{R,n} \) and \( L_{G,n} \) be the cumulative amount of noise trading in the red and green bond markets, respectively. By convention, positive quantities denote supply shocks while negative quantities denote demand shocks. If, for instance, \( L_{R,n} > L_{R,n+1} \) there are \( L_{R,n} - L_{R,n+1} \) sellers in the market for red bonds. If, on the other hand, \( L_{R,n} < L_{R,n+1} \), there are \( L_{R,n+1} - L_{R,n} \) buyers. As in other papers,\(^5\) no explicit model of the sources of noise trading will be presented. Finally, note that noise trading shocks affect the quantity of bonds available for investor trading: the

---

supply of red bonds after the shock is \( Q + L_{R,n} \) while the supply of green bonds after the shock is \( Q + L_{G,n} \).

In the absence of arbitragers, the equilibrium prices in the red and green market, \( P_{R,n} \) and \( P_{G,n} \), respectively, are determined by the following equations:

\[
\eta_n - \frac{P_{R,n}}{\theta_n} = Q + L_{R,n}
\]  
(2)

and

\[
\eta_n - \frac{P_{G,n}}{\theta_n} = Q + L_{G,n}
\]  
(3)

Solving for the prices gives,

\[
P_{R,n} = \frac{\eta_n - Q - L_{R,n}}{\theta_n}
\]  
(4)

and

\[
P_{G,n} = \frac{\eta_n - Q - L_{G,n}}{\theta_n}
\]  
(5)

Of particular interest is the difference between the price of red bonds and the price of green bonds. Letting \( \Delta_n = P_{R,n} - P_{G,n} \) and using the pricing equations (4) and (5), the relative mispricing equals

\[
\Delta_n = \theta_n L_n
\]  
(6)

with

\[
L_n = L_{G,n} - L_{R,n}.
\]

2.3. Equilibrium pricing with arbitragers

Arbitragers can now be introduced into the model. Let \( x_n \) be the number of green bonds bought by arbitragers and the number of red bonds sold by arbitragers. The optimal choice of \( x_n \) will be discussed below. For now, consider the effect of arbitrage on the mispricing \( \Delta_n \). If \( L_n > 0 \) and \( \Delta_n > 0 \), i.e. if \( P_{R,n} > P_{G,n} \), arbitragers will want to sell red bonds and buy green bonds, so \( x_n \) will be positive.
Furthermore, \( x_n \) will be added to the supply of red bonds and added to the demand of green bonds.

Adjusting equations (4) and (5) accordingly and subtracting (5) from (4) to obtain \( \Delta_n \) for this case,

\[
\Delta_n = \theta_n (L_n - 2x_n).
\]

Notice that arbitrage activity lowers the relative mispricing.

If \( L_n < 0 \) and \( \Delta_n < 0 \), i.e. \( P_{R,n} < P_{G,n} \), arbitragers will want to buy red bonds and sell green bonds, so \( x_n \) will be negative. Furthermore, \( x_n \) will be added to the demand for red bonds and to the supply of green bonds. In this case also \( \Delta_n = \theta_n (L_n - 2x_n) \) and arbitrage activity reduces the relative mispricing.

Summarizing this discussion, adjusting supply and demand in the bond markets to account for arbitrager activity changes the mispricing \( \Delta_n \) from (6) to

\[
\Delta_n = \theta_n (L_n - 2x_n). \tag{7}
\]

By definition, \( x_n \) is the sum of positions across arbitragers. If, there are \( I \) arbitragers, \( i=1 .. I \), an equilibrium in this model consists of strategies \( x_i \) and a process \( \Delta_n \) such that 1) each arbitrager chooses an optimal strategy given the evolution of \( \Delta_n \), and 2) the resulting \( x_n = \Sigma x_i \), in turn, generates the process \( \Delta_n \) given by (7).

Assume that all arbitragers have negative exponential utility functions with risk tolerance \( \tau_i \). Each arbitrager maximizes his expected utility of wealth as of the date the bonds mature so, letting \( W_n \) denote wealth with \( n \) periods to maturity, each arbitrager maximizes:

\[
- \exp \frac{W_i^n}{\tau_i}. \tag{8}
\]

Assuming that all arbitragers face the same holding cost, the price process in the multi-arbitrager economy is the same as when there is one representative arbitrager with risk tolerance.

---

\(^6\)The extension to several groups of arbitragers facing different holding costs does not present any conceptual difficulties (see section 3.4).
Consider therefore the trading decision of the representative arbitrager and begin with the value of an arbitrage position from one period to the next. Assume for the moment that \( \Delta_n > 0 \), i.e. \( P_{R,n} > P_{G,n} \).

As in Tuckman and Vila (1992), the arbitrager will buy \( x_n \) green bonds financed by borrowing \( x_n P_{G,n} \) dollars, short \( x_n \) red bonds, and lend \( x_n P_{R,n} \) dollars, \( x_n > 0 \). Denote the cost of maintaining a unit short position over one period by \( c \), the position next period is worth

\[
x_n P_{G,n-1} - x_n P_{R,n-1} + x_n P_{R,n}(1+r) - x_n P_{G,n}(1+r) - c|x_n| = x_n [\Delta_n(1+r) - \Delta_{n-1}] - c|x_n|.
\]

(9)

If \( \Delta_n < 0 \), the expression does not change, but \( x_n \) will be negative: the arbitrage position entails buying green bonds and shorting red bonds.

From the above discussion, the evolution of wealth can be described by the following equation

\[
W_{n+1} = W_n(1+r) + x_n [\Delta_n(1+r) - \Delta_{n-1}] - c|x_n|.
\]

(10)

Defining \( w_n = W_n(1+r)^n \), \( \delta_n = \Delta_n(1+r)^n \), and \( c_n = c(1+r)^n \), (10) can be rewritten as

\[
w_{n+1} = w_n + x_n [\delta_n - \delta_{n-1}] - c_n|x_n|.
\]

(11)

Equation (11) and the objective to maximize (8) completes the specification of arbitrager i's investment problem given \( \Delta_n \).

The model is completed by specifying an exogenous stochastic process for the net liquidity shocks, \( L_n \). For simplicity it will be assumed that \( L_n \) evolves as a binomial process: if its value with \( n \) periods to maturity is \( L_n \), then it will take on the value \( L_{n+1} = L_n + u \) with probability \( \pi(n,L_n) \) and a value \( L_{n+1} = L_n - u \) with probability \( 1-\pi(n,L_n) \).

2.4. Model Solution
This section begins by solving arbitrageur’s investment problem. Let:

\[
V(w_n, L_n, n) = \max_{x_n} \mathbb{E}_{n}\left[ -\exp \frac{w_0}{r} \right]
\]

where \( \mathbb{E}_n \) denotes the expectation when there are \( n \) periods to maturity and the maximum is over the strategy \( x_n \). By the principle of optimality in dynamic programming,

\[
V(w_n, L_n, n) = \max \left\{ \pi(n, L_n) V(w_n + x_n [\delta_n(L_n) - \delta_{n-1}(L_n + u)]) - c_{n-1} \right. \left| x_n \right|, L_n + u, n - 1 \right. +
\]

\[
(1 - \pi(n, L_n)) V(w_n + x_n [\delta_n(L_n) - \delta_{n-1}(L_n - u)]) - c_{n-1} \right| x_n \right|, L_n - u, n - 1 \right. \right\}
\]

(12)

Also, because the mispricing must vanish at the maturity date, the initial condition of the problem is \( V(w_0, L_0, 0) = -\exp(-Aw_0) \).

Because of the special form of the utility function, \( V(w_n, L_n, n) \) can be written as \( \exp(-Aw_n) J(L_n, n) \).

Using this fact, (12) becomes

\[
J(L_n, n) = \max \left\{ \pi(n, L_n) e^{-A[\delta_n(h_n(l_n)) - \delta_{n-1}(h_n^{-1}(L_n + u))] - c_{n-1}} \right| x_n \right|, L_n + u, n - 1 \right. +
\]

\[
(1 - \pi(n, L_n)) e^{-A[\delta_n(h_n(l_n)) - \delta_{n-1}(h_n^{-1}(L_n - u))] - c_{n-1}} \right| x_n \right|, L_n - u, n - 1 \right\}
\]

(13)

with initial condition \( J(L_0, 0) = -1 \).

To solve for the optimal strategy as a function of \( \delta_n \), begin as follows. If \( \delta_n > 0 \), the mispricing increases with a move to \( \delta_{n-1}(L_n + u) \) while the mispricing decreases with a move to \( \delta_{n-1}(L_n - u) \). Since the per period arbitrage profits is \( x_n (\delta_n - \delta_{n-1} - c_{n-1}) \), the position will never be profitable if \( \delta_n < \delta_{n-1}(L_n - u) + c_{n-1} \), i.e. if the position is not profitable even when the relative mispricing falls. So, for these value of \( \delta_n, x_n = 0 \). On the other hand, \( \delta_n < \delta_{n-1}(L_n + u) + c_{n-1} \) is inconsistent with equilibrium; if the position is profitable even when the relative mispricing rises, prices furnish a riskless arbitrage inclusive of holding costs and the optimal \( x_n \) would equal \( +\infty \). In the intermediate range,

\[
\delta_{n-1}(L_n + u) + c_{n-1} > \delta_n > \delta_{n-1}(L_n - u) + c_{n-1}
\]

\( x_n \) can be found by solving the optimization problem (13) to obtain

\[
x_n = \left[ \frac{1}{\mathbb{A}[\delta_{n-1}(L_n + u) - \delta_{n-1}(L_n - u)]} \right. \left. \mathbb{Ln} \left( \frac{1 - \pi(n, L_n)}{\pi(n, L_n)} \right) \frac{\delta_n(L_n) - \delta_{n-1}(L_n - u) - c_{n-1}}{c_{n-1} + \delta_{n-1}(L_n + u) - \delta_n(L_n)} J(L_n - u, n - 1) \right] \frac{J(L_n + u, n - 1)}{J(L_n - u, n - 1)} \right\}
\]

(14)
where \((H)^* = \text{Max} \{H, 0\}\).

If \(\delta_n < 0\), similar arguments reveal that \(x_n = 0\) when \(\delta_n > \delta_{n-1}(L_n + u) - c_{n-1}\) while \(x_n = -\infty\) when \(\delta_n < \delta_{n-1}(L_{n-1} - u) - c_{n-1}\). In the intermediate range,

\[
\delta_{n-1}(L_n + u) - c_{n-1} > \delta_n > \delta_{n-1}(L_{n-1} - u) - c_{n-1}
\]

the optimal \(x_n\) is given by

\[
x_n = \left\{ \frac{1}{A[\delta_{n-1}(L_n + u) - \delta_{n-1}(L_n - u)]} \ln \left\{ \frac{1 - \pi(n, L_n)}{\pi(n, L_n)} \frac{\delta_n(L_n) - \delta_{n-1}(L_n - u) + c_{n-1}}{-c_{n-1} + \delta_{n-1}(L_n + u) - \delta_n(L_n)} J(L_n - u, n - 1) \right\} \right\} (15)
\]

where \((H)^* = \text{Min} \{H; 0\}\)

To solve for the equilibrium values of \(x_n\) and \(\delta_n\), rewrite the equilibrium condition (7) in terms of \(\delta_n\):

\[
\delta_n = \theta_n (1 + t)^n (L_n - 2x_n). \quad (16)
\]

Then, solve (16) simultaneously with (14) or (15), as appropriate. For the case \(\delta_n > 0\), figure 1 illustrates the optimal position size and the equilibrium condition as a function of the mispricing \(\delta_n\). The dotted line represents the mispricing which generates infinite arbitrage activity. The simultaneous solution is given by the intersection of the two functions.

Given the solution technique for any \(n\) given the values at \(n - 1\), backward induction will provide the solution for all \(n\). The initial condition of the problem gives \(J(\cdot, 0)\). This allows for the solution of \(x_i\) and \(\delta_i\) along the lines described above. Then, substituting these values into (13) yields \(J(\cdot, 1)\). Proceeding in this fashion produces the entire mispricing process and the accompanying arbitrage strategy.
3. Numerical Examples

This section explores the implications of the model presented in section 2 through several numerical examples. These examples require the specification of particular parameter values and of a particular process governing the evolution of the net liquidity shock, \( L \). The following choices constitute the "base case" scenario.

As mentioned in the previous section, the parameter \( \alpha \) is related to the depth of the market. Since a $10 billion Treasury coupon issue is not uncommon [First Boston (1990), pp. 49-54], set \( \alpha = 10 \) billion. Relying on an often cited estimate of the reverse spread in the special issues market, set the holding cost parameter, \( c \), equal to .5% [Stigum (1983), p. 414].

The other parameters are initially set at reasonable, but more arbitrarily chosen values. Let the aggregate risk tolerance, \( \tau \), be fixed at 1 million, let the bond's maturity at the time of issue be 25 years, and let the liquidity imbalance at that time equal 0. Finally, let the coupon rate and discount rate be 12%.

The net liquidity shock is assumed to follow a discretized version of the Ornstein-Uhlenbeck process. The continuous time version can be written as

\[
dL(t) = -\rho L(t)dt + \sigma dB(t)
\]

where \( \rho \) and \( \sigma \) are nonnegative constants and \( dB(t) \) is the increment of a Brownian motion. To define the discrete version of this process, let \( h \) be the length of the trading interval and set \( u_b \) and \( \pi(n,L_n) \) of the previous section such that

\[
u_b = \sigma \sqrt{h}
\]

and
\[
\frac{1}{2} \left( 1 - \frac{\rho L_n}{\sigma} \sqrt{h} \right) \quad \text{if} \quad \frac{1}{2} \left( 1 - \frac{\rho L_n}{\sigma} \sqrt{h} \right) \in [0,1]
\]

\[
\pi(n, L_n) = 0 \quad \text{if} \quad \frac{1}{2} \left( 1 - \frac{\rho L_n}{\sigma} \sqrt{h} \right) < 0
\]

\[
1 \quad \text{if} \quad \frac{1}{2} \left( 1 - \frac{\rho L_n}{\sigma} \sqrt{h} \right) > 1
\]

[See Nelson and Ramaswamy (1990), pp. 399-400, for the relevant convergence results.]

For the numerical examples, \( \rho \) was set equal to .5 and \( \sigma/\alpha \) to 2%. Note that \( \sigma \) is quoted as a percentage of market size. Finally, the trading interval \( h \) was set at .00125, producing 800 trading intervals each year. This choice provided satisfactory convergence of the discrete process to its continuous time limit.

The first subsection of this section discusses the optimal position size of the arbitrager. The second illustrates how arbitrage activity affects equilibrium mispricing. The third subsection focuses on the dynamics of the mispricing process. The fourth subsection explores the implications of the model with respect to the arbitrage industry.

### 3.1 Arbitrage Positions

Figure 2 graphs the optimal position size of the arbitrager as a function of the equilibrium mispricing when the bond has 5 years remaining until maturity. The domain of this graph, as of all others to follow, captures approximately 95% of the probability distribution of mispricings.

For relatively small mispricings, arbitragers do not take any position: the potential profits are not large enough relative to the potential accumulation of holding costs. For larger levels of mispricing within the riskless arbitrage bounds, optimal position size is finite and increases with the level of mispricing. See Tuckman and Vila (1992a) for a more detailed discussion of optimal position size.
3.2 Equilibrium Mispricing

Figure 3 graphs the equilibrium mispricing with and without arbitragers as a function of the net liquidity shock. Once again, the bond matures in 5 years.

As expected, larger net liquidity shocks increase mispricing. Without arbitragers the mispricing increases linearly, as given by equation (6). For relatively small shocks arbitragers do not optimally trade and mispricing is the same with arbitragers as without arbitragers. For larger shocks, however, arbitragers reduce the level of mispricing that would obtain in their absence. The equilibrium mispricing with arbitragers eventually flattens out at the maximum mispricing consistent with no riskless arbitrage opportunities, i.e. at the present value of the holding costs incurred by maintaining a unit arbitrage position until maturity.

Two lessons emerge from figure 3. First, when the equilibrium mispricing without arbitragers is not zero, the equilibrium mispricing with arbitragers is also not zero. Due to holding costs and risk aversion, arbitragers bring prices closer into line but do not eliminate mispricing altogether.

Second, arbitrage activity reduces mispricing to a level below that which would trigger riskless arbitrage transactions. Therefore, market prices may commonly reflect these smaller mispricings so that the no-riskless arbitrage bound may rarely be binding. Stated another way, mispricing in the presence of arbitragers does not equal the minimum of mispricing without arbitragers and the no-riskless arbitrage bound: mispricing in the presence of arbitragers is less than or equal to that minimum.

Table 1 further explores equilibrium mispricing in the context of the present model. Recall that the steady-state distribution of the Ornstein-Uhlenbeck process is normal with mean 0 and standard deviation $\sigma(2\rho)^{-1/2}$. The three combinations of $\rho$ and $\sigma/\alpha$ considered in table 1 are all such that the
steady-state distribution of the net liquidity shock has a volatility equal to 2% of market size. Furthermore, the values of \( p \) are large enough, and the discretization fine enough, so that this steady-state distribution is achieved for all of the maturities listed in the first column of table 1. Because of this steady-state property, mispricing without arbitragers at a given maturity (i.e. \( \theta_p L_n \)) is the same for all three combinations of \( p \) and \( \sigma \).

In the model of section 2, the sources of mispricing, namely valuation differences across investors, increase with maturity. As a result, table 1 reports that average mispricing without arbitragers increases with maturity as well. More importantly, however, the presence of arbitragers does not eliminate this maturity effect. In other words, average mispricing increases with maturity even in the presence of arbitragers. Because of holding costs, i.e. because arbitrage costs increase with the length of time over which a position is held, arbitragers do not hold mispricing below some fixed level that is independent of maturity.

Another lesson about the relationship between mispricing, arbitrage activity, and maturity can be drawn from table 1. The percentage reduction of mispricing due to arbitrage activity follows an inverted-U shape in maturity. This phenomenon can be explained as follows. Because of holding costs, it is more costly to arbitrage away the mispricings of long-term assets. On the other hand, because mispricing vanishes as a bond matures, potential gains from arbitraging short-term assets are small. These effects, in combination, generate the largest percentage reduction of mispricing at intermediate maturities.

Table 1 also reveals the effects of net liquidity shock dynamics on arbitrage activity and mispricing.

---

7 Because it has been assumed that, at issuance, the net liquidity shock and, therefore, the mispricing equals zero, average mispricing will eventually decline with maturity near the original maturity, i.e. 25 years. But the impact of this initial condition is completely dissipated when the net liquidity shock distribution achieves its steady state. Since this is the case for all of the maturities in table 1, the decline of average mispricing near issuance does not appear in the table.
As $p$ and $\sigma$ increase, without changing the unconditional net liquidity shock distribution, mispricing in the presence of arbitragers decreases and the percentage reduction of mispricing due to arbitragers increases. Increases in $p$ contribute to this effect by returning net liquidity shocks to 0 more rapidly and, consequently, by lowering the expected accumulation of holding costs. Increases in $\sigma$ contribute to this effect by generating higher levels of mispricing.\(^8\) In combination, it may be said that arbitrage is most effective when the underlying mispricing process is transient and conditionally volatile.

3.3 Dynamics: Mean Reversion and Conditional Volatility

This subsection analyzes how arbitrage activity affects the evolution of the mispricing process. Begin by defining the process $Z_n$ to be the liquidity imbalance between the two markets after arbitragers take their positions, i.e.

$$Z_n = L_n - 2x_n.$$  

Recall from equations (6) and (7) that $\Delta_n = \theta_n L_n$ without arbitragers and that $\Delta_n = \theta_n Z_n$ with arbitragers. Therefore, comparing the processes $L_n$ and $Z_n$ is equivalent to comparing the mispricing process with and without arbitragers. This comparison is simpler than comparing the mispricing directly because $L_n$ has very simple conditional moments:

$$E[dL_n|L_n] = -\rho L_n dt$$  and  

$$\text{Var}[dL_n|L_n] = \sigma^2 dt.$$  

Defining $\rho'$ and $\sigma'$ such that

$$\rho' = E \left[ -\frac{dZ_n}{Z_n dt} \mid L_n \right]$$  and

\(^8\) This effect is at least partially offset, however, by the increased riskiness of arbitrage positions at higher levels of $\sigma$.  


the effects of arbitrage on mispricing dynamics can be seen by comparing \( \rho' \) and \( \sigma' \) with \( \rho \) and \( \sigma \), respectively.

Figure 4 graphs the mean reversion coefficients \( \rho \) and \( \rho' \) as a function of the mispricing under the parameters of the base case scenario. Once again, the bond has 5 years remaining before maturity.

For relatively low levels of mispricing, positive or negative, mean reversion is the same with and without arbitragers because arbitragers optimally choose not to take any position. For higher levels of mispricing, arbitrage activity raises the mean reversion of the mispricing process above that which would prevail in the absence of arbitragers. In these regions, arbitragers' selling the dear bond and buying the cheap bond increases the rate at which prices revert to their fundamental values. For relatively high levels of mispricing, however, mean reversion in the presence of arbitragers is less than mean reversion in their absence. In these regions, mispricing approaches the riskless arbitrage bounds. Near these bounds changes in \( L_n \) have little effect on \( Z_n \), i.e. \( Z_n \) becomes nearly constant, as evident from the extreme domain values of figure 3. Therefore, in these regions, mean reversion falls with mispricing and remains below the value that would prevail in the absence of arbitragers.

Figure 5 graphs the annualized conditional volatility of the mispricing process with and without arbitragers. For relatively small levels of mispricing, when arbitragers take no position, the conditional volatility of the two cases are the same. For higher levels of mispricing, however, arbitrage activity reduces conditional volatility. Furthermore, the extent of this reduction increases with the level of mispricing.

Figures 4 and 5 raise a flag of caution with respect to empirical studies of mispricing processes. In
particular, a regression of changes in mispricing on the level of mispricing is likely to be misspecified. The coefficient of mean reversion is not a constant but rather a highly non-linear function of the level of mispricing.

3.4 The Arbitrage Industry

This subsection will refer to the certainty equivalent of arbitrage profits, CE. It is defined such that the value of being an arbitrager with present wealth $W_n$ is equal to the utility of $W_n + CE$. Recalling that $w_n = W_n(1+r)^n$, this condition becomes

$$V(w_n, L_n, n) = \exp \left[ -\frac{w_n + CE(1+r)^n}{\tau} \right].$$

Then, using the separability of $V$ into $J$ and an exponential function of wealth,

$$CE = -\tau \frac{\ln[-J(L_n, n)]}{(1+r)^n}.$$

As noted in section 2, when all arbitragers face the same holding cost, the model with many arbitragers can be viewed as an economy with a representative arbitrager who has a risk tolerance $\tau = \sum \tau_i$. Therefore, increases in $\tau$ can be thought of as increases in the number of arbitragers.

The second column of table 5 reveals that the certainty equivalent of the arbitrage industry is an inverted-U function of $\tau$. This shape can be explained as follows. Increasing the number of arbitragers has two effects on the aggregate certainty equivalent. One, more arbitragers generate larger aggregate positions, larger expected profits, and, consequently, larger certainty equivalents. Two, more arbitragers lower the level of mispricing, potential profits, and certainty equivalents. When there are relatively few arbitragers, mispricing is relatively high and the first effect dominates. When there are relatively many arbitragers, mispricing is relatively low and the second effect dominates.

The third column of table 2 shows that the certainty equivalent per arbitrager decreases with the
number of arbitragers. Because increasing the number of arbitragers reduces mispricing, the value of being one of many arbitragers declines.

Table 3 reports aggregate certainty equivalents as a function of the holding cost. For very small holding costs, mispricings are eliminated so quickly that certainty equivalents are very low. For very large holding costs, arbitrage is so unprofitable that certainty equivalents are again very low. Therefore, certainty equivalents are an inverted-U function of holding costs.

Until now it has been assumed that all arbitragers face the same holding cost. The model, however, can be generalized to allow for different holding costs across arbitragers. In the examples presented below, it was assumed that most arbitragers, with an aggregate risk tolerance of 900,000, face a holding cost of .5%. A small group of arbitragers, with an aggregate risk tolerance of 100,000, face a different holding cost. All other parameters are as in the base case described earlier. Table 4 examines the certainty equivalents of both groups as a function of the holding cost of the smaller group.

When both groups face a cost of .5%, the profit per arbitrager is .53 for each group. If the smaller group has a cost disadvantage, i.e. a holding cost of .7%, its profit per arbitrager falls to .14 while that of the larger group rises very slightly. The smaller group cannot compete effectively because of its higher cost but its inability to do so has little effect on the broader market. On the other hand, if the smaller group has a cost advantage, i.e. a holding cost of .3%, its profit per arbitrager jumps to 2.30 while that of the larger group falls to .31. Not only does the smaller group profit by being able to take advantage of smaller mispricings, but it tightens the riskless arbitrage bounds and reduces mispricings to such a great extent that the vast majority of arbitragers are left with little more than 58% of their previous certainty equivalents.
The analysis of table 3 reveals that the interaction between arbitragers resembles Bertrand competition in that only the lowest cost competitors earn profits. Consequently, there are huge incentives to cost reduction and the arbitrage industry may very well be dominated by a few technologically superior firms.

4. Concluding remarks

The dynamic model of equilibrium arbitrage developed here assumes that noise trading in two segmented markets causes price deviations from fundamental values. Risk averse arbitragers facing holding costs reduce these deviations, but do not eliminate them completely. In contrast to models with an exogenously specified mispricing process, the equilibrium nature of this model allows for conclusions to be drawn about the impact of arbitrage activity on market mispricing. In contrast to static models of arbitrage activity, the dynamic setting of this model allows for conclusions to be drawn about the evolution of the mispricing process.

This paper contributes not only to the arbitrage pricing literature, but also to the growing literature on imperfect financial markets. While the impact of frictions such as trading costs, holding costs, borrowing constraints, or market incompleteness on dynamic investment strategy is relatively well understood, little is known about the impact of these frictions on equilibrium price processes. While some recent papers have addressed this question, they focus on trading costs rather than holding

---


While academic research has focused on riskless arbitrage activity, the process by which market prices are kept close to fundamental values can often be characterized as a risky investment activity. This paper has focused on a risk that arises when an arbitrager facing holding costs trades on price differentials: he looses if the market becomes 'efficient' too slowly.

Delgado and Dumas (1993) focus on a risk that arises when an arbitrager facing trading costs attempts to exploit rate differentials. In their work, the arbitrager loses if the market becomes 'efficient' too quickly, that is if the accumulated rate differential does not compensate for the trading cost. It is no accident that Delgado and Dumas (1993) consider rate arbitrage and trading costs while this paper considers price arbitrage and holding costs. As pointed out in the introduction, the problem of price arbitrage and trading costs has an extremely simple solution. Similarly, the problem of rate arbitrage and holding costs has an extremely simple solution: when the rate differential exceeds the holding cost, take an infinite position. Otherwise, do nothing.

From a theoretical perspective, future research can most improve on the present work by allowing investors and noise traders to behave more rationally than they do here. For example, one might require that traders who buy bonds choose to buy those in the cheaper market while traders who sell bonds sell the very bond they own, cheaper or not. In this scenario, mispricings are reduced by buyers as well as arbitragers and therefore, are likely to be lower and more short-lived than in the present model.

From an empirical perspective, the present model can be used by investigators to frame hypotheses

---


14 In the case of rate arbitrage, trading costs are the costs of moving capital into a different market.

15 See Vayanos and Vila (1993) for a similar situation in the case of trading costs.
about the time series behavior of mispricing processes and about differences in mispricing processes across markets that differ in structural parameters such as size, level of holding costs, etc.
Appendix

This appendix derives the demand curve for a bond issue when investors face different tax rates. As in the text, assume that the bond matures n periods from now and entitles holders to $d$ at the end of each period before maturity and to $1+d$ at maturity. Coupon income is taxed at a rate $\beta$, the after-tax discount rate is constant at $r$, and investors assume that they will hold the bond until maturity.\textsuperscript{16}

Under these assumptions, the value of the bond to an investor with tax rate $\beta$ is

\[
V_n(\beta) = \frac{d(1-\beta)}{r} \left[ 1 - \frac{1}{(1+r)^n} \right] + \frac{1}{(1+r)^n} = V_n(0) - \frac{\beta d}{r} \left[ 1 - \frac{1}{(1+r)^n} \right]. \tag{A1}
\]

For a given bond price, $P_n$, an investor with tax rate $\beta$ will be willing to buy bonds if $P_n \leq V_n(\beta)$. When $n=0$, this condition is $P_n \leq 1$. For $n \geq 1$, solving (A1) for $\beta$ shows that investors with tax rate $\beta$ will be willing to buy bonds so long as

\[
\beta \leq \frac{r}{d} \frac{V_n(0) - P_n}{1 - \frac{1}{(1+r)^n}}. \tag{A2}
\]

To derive a demand curve, assume that i) each investor buys at most 1 bond and ii) the number of investors with tax rates below some $\beta$ is given by $\alpha \beta$.

From these assumptions and (A2), the demand function, $D_n(P_n)$, can be written as

\[
D_n(P_n) = \frac{\alpha r}{d} \frac{V_n(0) - P_n}{1 - \frac{1}{(1+r)^n}}. \tag{3}
\]

Defining $\theta_n = (d/\alpha r)(1-1/[1+r]^n)$ and $n_n = V_n(0)\theta_n$ gives the demand function reported in the text.

\textsuperscript{16} This assumes that investors are myopic in the following sense. When deciding to buy or sell a bond, they compare the market price to their own valuation under a buy and hold strategy. But this strategy is not necessarily optimal, since investors may prefer to delay a sale in the expectation that prices will rise. While one might be tempted to think of the decision to sell in terms of exercising an option, the analogy is misleading: an investor who sells, i.e. exercises, can repurchase the bond later and sell yet again. In fact, in a related context, Tuckman and Vila (1992b) show that the myopic strategy is sometimes optimal. In any case, careful modeling of investor decisions in the context of the present model will be left as a subject for future research.
References


Delgado, F. and B. Dumas, (1993), "How Far Apart Can Two Riskless Interest Rate Be? (One Moves, the Other One Does Not)," mimeo, Fuqua School of Business, Duke University.


Table 1: Expected absolute mispricing with and without arbitragers. At all maturities featured below, the net liquidity shock process (as a percent of market size) is at its steady state distribution with mean 0 and standard deviation \(\sigma/\alpha/(2\rho)^{1/2}\). Also, the three combinations of \(\rho\) and \(\sigma/\alpha\) featured below all produce a volatility equal to 2% of market size. As a result, the expected absolute mispricing without arbitragers at a given maturity, which equals the net liquidity shock times a function of maturity, is the same across the three parameter combinations.

| Maturity | \(\text{E}(\left| \Delta \right|)^{\text{NoArb.}}\) | \(\sigma=2, \rho=.5\) | \(\sigma=4, \rho=2\) | \(\sigma=8, \rho=8\) |
|----------|--------------------------------|-------------------|-------------------|-------------------|
|          | \(\text{E}(\left| \Delta^{\text{arb}} \right|)\) | \% reduc. | \(\text{E}(\left| \Delta^{\text{arb}} \right|)\) | \% reduc. | \(\text{E}(\left| \Delta^{\text{arb}} \right|)\) | \% reduc. |
| 1        | 0.17%                           | 0.16%   5.18%     | 0.13%   25.64%   | 0.04%   75.49%   |
| 5        | 0.69%                           | 0.52%   24.37%     | 0.19%   71.88%   | 0.04%   93.66%   |
| 10       | 1.08%                           | 0.75%   30.98%     | 0.27%   75.10%   | 0.05%   95.65%   |
| 15       | 1.30%                           | 0.96%   26.23%     | 0.53%   59.24%   | 0.06%   95.77%   |
| 20       | 1.43%                           | 1.14%   20.22%     | 0.96%   32.85%   | 0.10%   93.15%   |
Table 2

Table 2: Aggregate arbitrage profits and arbitrage profits per arbitrage as a function of the aggregate risk tolerance, a proxy for the number of arbitragers.

<table>
<thead>
<tr>
<th>Aggregate risk tolerance</th>
<th>Aggregate arbitrage profits</th>
<th>Profits per arbitrage</th>
</tr>
</thead>
<tbody>
<tr>
<td>100,000</td>
<td>181,619</td>
<td>1.82</td>
</tr>
<tr>
<td>500,000</td>
<td>467,468</td>
<td>.93</td>
</tr>
<tr>
<td>1 million</td>
<td>529,866</td>
<td>.53</td>
</tr>
<tr>
<td>5 million</td>
<td>340,260</td>
<td>.07</td>
</tr>
<tr>
<td>10 million</td>
<td>199,957</td>
<td>.02</td>
</tr>
</tbody>
</table>
Table 3

Table 3: Aggregate arbitrage profits as a function of the holding cost c.

<table>
<thead>
<tr>
<th>Holding cost</th>
<th>Aggregate arbitrage profits</th>
</tr>
</thead>
<tbody>
<tr>
<td>.1%</td>
<td>286,547</td>
</tr>
<tr>
<td>.3%</td>
<td>543,354</td>
</tr>
<tr>
<td>.5%</td>
<td>529,966</td>
</tr>
<tr>
<td>.9%</td>
<td>227,544</td>
</tr>
</tbody>
</table>
Table 4: Arbitrage profit per arbitragers when arbitragers have different holding costs. For this table the risk tolerance of the larger group is 900,000 while the risk tolerance of the smaller group is 100,000. The holding cost of the larger group is fixed at .5%.

<table>
<thead>
<tr>
<th>Holding cost of the smaller group</th>
<th>Arbitrage profits per arbitrager in the larger group</th>
<th>Arbitrage profits per arbitrager in the smaller group</th>
</tr>
</thead>
<tbody>
<tr>
<td>.3%</td>
<td>.3076</td>
<td>2.2966</td>
</tr>
<tr>
<td>.5%</td>
<td>.5299</td>
<td>.5299</td>
</tr>
<tr>
<td>.7%</td>
<td>.5744</td>
<td>.1442</td>
</tr>
</tbody>
</table>
Determination of Equilibrium

Figure 1

- Equilibrium Condition
- Optimal Position Size
- Riskless Arbitrage Bound
Equilibrium Mispricing

Figure 3

Mispricing (as a % of Face Value)

Net Liquidity Shock (as a % Market Size)

without arbitrageurs

with arbitrageurs
Conditional Volatility

Equilibrium Mispricing (as a % of Face Value)

-1.1% -0.7% -0.3% 0.3% 0.7% 1.1%

0.0% 0.4% 0.8% 1.2% 1.6% 2.0% 2.4%

with arbitrageurs
without arbitrageurs

Figure 5

Conditional Volatility (as a % of Market Size)