WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

ARBITRAGE WITH HOLDING COSTS:
A UTILITY-BASED APPROACH

by

Bruce Tuckman and Jean-Luc Vila

Latest Revision: December 1991
Working Paper No. 3364-91-EFA

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139
ARRITRAGE WITH HOLDING COSTS: A UTILITY-BASED APPROACH

by

Bruce Tuckman and Jean-Luc Vila

Latest Revision: December 1991
Working Paper No. 3364-91-EFA
Arbitrage With Holding Costs: A Utility-Based Approach

by

Bruce Tuckman
Stern School of Business, New York University

and

Jean-Luc Vila
Sloan School of Management, M.I.T.

This version: December, 1991

We would like to thank participants in the N.Y.U. and Cornell Finance Seminar Series, participants in the 1990 AFA and AFFI meetings, and an anonymous referee for helpful comments and suggestions.
Abstract

Unit time costs, or holding costs, are incurred in many arbitrage contexts. Well known examples include losing the use of short sale proceeds and lending funds at below market rates in reverse repurchase agreements. This paper analyzes the investment problem of a risk averse arbitrageur who faces holding costs. The proposed model allows prices to deviate from their "fundamental" values without allowing for riskless arbitrage opportunities. After characterizing an arbitrageur's optimal strategy, the model is examined in the context of the Treasury bond market. The analysis reveals that holding costs are an important friction in this market and that they can be expected to be a significant determinant of arbitrageur behaviour.
Arbitrage With Holding Costs: A Utility-Based Approach

Financial economists have made great use of the notion of arbitrage in frictionless markets. Researchers define arbitrage as a set of transactions which costs nothing and yet risklessly provides positive cash flows. They then assume that arbitrage opportunities never exist and derive precise implications about the relative prices of traded securities.

The presence of market frictions complicates the story. In the simplest of models, trading costs make it impossible to risklessly exploit small price deviations from arbitrage-free relationships. Therefore, when market prices do not admit riskless arbitrage opportunities, arbitrageurs do nothing. When market prices do admit such opportunities, arbitrageurs bet all they have on the sure proposition that prices will be back in line at the maturity date of the underlying securities.

While this simple model can explain the empirical regularity of price deviations from "fundamental value," more careful portrayals of market frictions make for richer models of arbitrage activity. Brennan and Schwartz (1990), for example, show that trading costs and position limits in the stock index futures market can make it valuable to close a position before maturity. As a result, it may be worthwhile to open a position even when the costs of opening it plus the costs of closing it at maturity exceed the price deviation from fundamental value. This paper contributes to the theory of arbitrage pricing with market frictions in two ways. First, it focuses on a somewhat ignored aspect of true-to-life arbitrage activity, namely unit time costs, or holding costs. Second, it builds a model which rules out riskless arbitrage opportunities without removing all incentives to exploit market mispricings. As a result, risk averse arbitrageurs do not sit idly by or bet all the have. They take a finite, risky position if the mispricing is large enough and adjust its size as the mispricing changes. This behavior is consistent with the casual empirical observation that professional arbitrageurs

---

1 The phrase "fundamental value" is to mean the arbitrage-free price when there are no market frictions. In no sense do we mean that the price, in the presence of frictions, should equal its fundamental value.
routinely find and take advantage of deviations from fundamental values.

Section I describes how holding costs can transform riskless arbitrage opportunities into potentially profitable but risky investment projects. It then argues that holding costs are important in many different arbitrage contexts.

Section II presents a model of arbitrage activity. The scene is set with a holding cost structure and an exogenous mispricing process that, in combination, do not admit riskless arbitrage opportunities. The paper then derives the partial differential equation governing a risk averse arbitrageur's optimal, dynamic strategy. For the case of negative exponential utility it is shown that 1) arbitrageurs will hold a position if and only if the potential gains are large enough, 2) the position is finite, and 3) arbitrageurs will take a position even with an instantaneously negative expected return since any future losses will be accompanied by greater arbitrage opportunities.

Section III examines the model's assumptions in the context of the U.S. Treasury bond market. estimates the parameters of the model, and numerically solves the partial differential equation of section II using those parameters. The data support the contention that holding costs are quite important, relative to trading costs, in preventing riskless arbitrage. The data also support the assumption of an autoregressive process for the deviation of prices from fundamental value. Finally, the numerical solutions reveal an "option effect" and a "risk effect" in the determination of position size and allow for an assessment of the relative importance of the two effects.

Section IV concludes and discusses avenues for future research. In particular, this work can be viewed as a first step towards building an equilibrium model of arbitrage activity in the presence of holding costs. Such a model holds the promise of being able to provide tighter price bounds around fundamental value than those supplied by riskless arbitrage arguments.
I. Holding Costs and Risky Arbitrage

Consider two portfolios, A and B, which provide identical cash flows over time and, at the time of the last cash flow, must have identical market prices. Nevertheless, for some reason, the market prices of the portfolios, \( P^A \) and \( P^B \), differ by an amount \( x \). Assuming that \( P^A > P^B \), investors in frictionless markets will short portfolio A and purchase portfolio B, realizing \( x \) today and incurring no future cash inflows or outflows.

Although this paper could have assumed that institutions were such as to allow the arbitrage just described, the model to be presented in section II builds on a more realistic description of short sale agreements. In order to short portfolio A, an investor must borrow the securities from somewhere. Furthermore, the lender of the securities will require collateral in order to ensure the eventual return of his securities. Therefore, the arbitrage must be arranged as follows: lend \( P^A \) dollars to the holder of the securities (or, equivalently, post an interest-bearing security worth \( P^A \)), take the securities in A as collateral, sell these securities for \( P^A \), borrow \( P^B \), and purchase portfolio B. In this more realistic setting, the investor neither pays nor receives any money when establishing his position, but realizes the future value of \( x \) upon closing the position.

In frictionless markets, investors would wish to arbitrage *ad infinitum*, leading to the conclusion that, in equilibrium, \( x \) must equal 0. Introducing holding costs changes arbitrage behavior in two important ways. First, the no-riskless arbitrage condition only requires that \( x \) be no larger than the present value of the accumulated holding costs, from the opening of the position to the time when the two portfolio prices must be equal. Second, for smaller \( x \), investors face a risky investment opportunity: if \( x \) falls to 0 quickly enough, then the costs of effecting the arbitrage will be small and the transaction will have proved profitable. If, on the other hand, \( x \) does not fall quickly enough, the accumulated costs will wipe out the realized gains.

Holding costs appear in many arbitrage contexts. First, the short sale of any spot security or
commodity will usually incur unit time costs since investors must often sacrifice the use of at least part of the short sale proceeds. In the case of stocks, it is true that institutions with large client bases can take short positions almost costlessly. Nevertheless, smaller firms engaged in arbitrage activity do incur holding costs when short-selling stocks. Second, trading in futures markets may generate holding costs since at least 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, deviations from desired investment strategies caused by having to meet collateral requirements can be thought of as unit time costs.

The model of this paper focuses on holding costs and ignores trading costs for a number of reasons. First, despite their ubiquitousness, holding costs have not received much scholarly attention. Second, trading costs would obscure the qualitative nature of this paper's results. Trading costs transform strategies which continuously adjust holdings into strategies which discretely adjust holdings, where the frequency of adjustment decreases in the trading costs. Since this effect has been studied elsewhere, there is little danger that focusing on holding costs will prove misleading. Third, while the trading costs of professional traders and arbitrageurs can be extremely small, the above examples and the discussion to follow reveal that their holding costs need not be trivial.

One of the easiest contexts in which to establish both the importance of holding costs relative to trading costs and the absolute magnitude of holding costs is the Treasury bond market. The major

---

2 For the case of stocks, see, for example, Cox and Rubinstein (1985), pp. 98-103.

3 We would like to thank Billy Rao of Susquehanna Investment Group for this comment.

4 See, for example, Duffie (1989), pp. 58-68.

5 See, for example, Anderson (1987), p. 198.

6 We would like to thank Peter Carr for this suggestion.

7 See, for example, Dumas and Luciano (1990), Hodges and Neuberger (1989), and Fleming, Grossman, Vila, and Zanphopoulou (1989).
trading costs in this market, namely bid-ask spreads and brokerage fees, are typically small, totalling about 3/32 of one percent and 1/128 of one percent, respectively. The major holding costs in the government bond market are reverse spreads. Recall that bonds are sold short through reverse repurchase agreements in which the short seller lends money and takes the security he wants to short as collateral. The reverse spread is defined as the difference between the lending rate on general collateral and the lending rate on specific collateral. In other words, the reverse spread is the rate available when any Treasury bond will serve as collateral minus the rate available when only a specific Treasury bond will serve as collateral. This spread is usually positive because the difficulties in finding the owner of a particular bond and the likely attempts of other would-be shorters to find the same bond translates into an opportunity loss on the money lent through reverses. Stigum (1983) reports average spreads of .25% to .65%. Now, to compare the magnitudes of trading costs and holding costs, consider shorting $100 face value of a par bond and maintaining the position until maturity. The trading costs come to about 10 cents and, with a spread of .5%, the holding costs exceed the trading costs for maturities longer than about 2 1/2 months. The price data presented in section III show that, for all but the shortest maturities, holding costs do more than trading costs to inhibit riskless arbitrage activity.

Two recent empirical papers serve as excellent motivations for the present analysis. Cornell and Shapiro (1989) document both the persistent mispricing of a particular Treasury bond and the attempts of arbitrageurs, facing holding costs, to profit from that mispricing. Amihud and Mendelson (1990), studying the effects of liquidity by comparing the prices of matched-maturity Treasury bills and notes, find that notes are cheap relative to bills and proceed to examine whether these price

---


9 See Stigum (1990), pp. 596-597.

10 p. 414.
differences can be exploited. Considering only trading costs, as captured by the bid-ask spread and brokerage costs, riskless arbitrage opportunities seem rampant. Adding holding costs, however, most of these money machines break down.\footnote{The authors actually consider the bid-ask spread and then add both brokerage costs and holding costs. Nevertheless, because brokerage costs are so small in this market, the result stated in the text holds as well.}

Another motivation for the analysis of holding costs stems from the relatively recent literature about the source of market mispricings. Some authors (e.g. De Long et al. (1990) and Lee et al. (1991)) have hypothesized that investors’ short term horizons allow for persistent deviations from fundamental values. This paper reveals that holding costs, in effect, generate myopia. Even though arbitrageurs have an investment horizon equal to the maturity of the underlying securities, holding costs discourage the maintenance of long-term arbitrage positions. Therefore, holding costs may be viewed as a particular way to endogenize investor impatience.

II. The Model

This section models the behavior of risk averse arbitrageurs who incur holding costs and observe prices that deviate from fundamental value. To allow for the possibility of arbitrage opportunities, begin by assuming that there exist two portfolios, A and B, which are characterized by identical cash flows through their common maturity date, T.\footnote{While this assumption seems innocuous enough, it does rule out concurrent investments in other risky assets. This simplification allows the model to abstract from the portfolio effects which would arise from holding other risky assets in addition to an arbitrage position.} Let their date-t prices be $P^A_t$ and $P^B_t$, respectively.

Let $x_t = P^A_t - P^B_t$ denote the difference between the prices of these portfolios on date $t$. While it would be ideal to endogenize the stochastic process governing the evolution of $x_t$, this paper follows
others\textsuperscript{13} in taking the process as exogenous in order to focus on the investment problem facing arbitrageurs. The assumed evolution of $x_i$ should exhibit the following properties:

1) $x_T = 0$, i.e. the price deviation disappears at maturity.

2) $x_i$ is never so large as to admit riskless arbitrage opportunities.

3) $x_i$ tends towards zero.

4) The larger the absolute value of $x_i$, the faster $x_i$ tends toward zero.

The last two properties attempt to infuse the process with an equilibrium flavor. Property 3) captures the idea that deviations are transient. Property 4) reflects the notion that larger deviations will be exploited more readily, and, as a result, tend to vanish at a faster rate than smaller deviations.

In order to find a process which satisfies (1) and (2), begin by determining the permissible range of $x_i$. Along the lines of the motivations discussed in section I, let $c > 0$ be the holding cost per unit time per portfolio unit.\textsuperscript{14} Also, let $r$ denote the arbitrageur's cost of capital. Then, the present value cost of maintaining a unit short position from date $t$ until maturity is

$$
\int_t^T c \ e^{-r(s-t)} \, ds .
$$

Letting $\tau = T-t$, the time to maturity, and denoting this cost function by $s(\tau)$,

$$
s(\tau) = \frac{c}{r} (1 - e^{-r\tau})
$$

(1)

If the absolute value of $x$ exceeds $s(\tau)$ at any time, then a riskless profit could be made through the arbitrage described at the start of the previous section. Therefore, to guarantee that $|x_i| < s(\tau)$, generate the process $x_i$ in the following way:

\textsuperscript{13} See, for example, Brennan and Schwartz (1990). Recently, Holden (1990) has endogenized a mispricing process by assuming that a clientele effect splits the market and that each segment is periodically hit by liquidity shocks.

\textsuperscript{14} In the context of the U.S. Treasury market, the holding cost can change as the arbitrageurs roll over expiring repurchase agreements. It seems best, however, to postpone modelling $c$ until it is endogenized to reflect the collective activity of arbitrageurs. Furthermore, in order to model the riskless arbitrage boundary in a simple way, it has further been assumed that the holding cost is per portfolio unit. In the U.S. Treasury market, however, the cost is usually per dollar of portfolio value.
where \( z_t \) is an underlying state variable and \( \phi \) is a function from \((-\infty, +\infty)\) to \((-1,1)\) which preserves the sign of \( z \) and increases in \( z \). Notice that this setup also ensures that \( x_T = 0 \) since \( s(0) = 0 \).

The process for \( x_t \) can be made to satisfy properties 3) and 4) listed above if \( z_t \) is an autoregressive process. To this end, assume that the state variable evolves according to an Ornstein-Uhlenbeck process

\[
dz_t = -\rho z_t \, dt + \sigma dB_t
\]

where \( dB_t \) is the increment of a standard Brownian motion.

Now consider an investor who notices that \( x_t \) is not equal to zero. If \( P_t^A \) exceeds \( P_t^B \), for example, he will short some quantity, \( I_t \), of \( A \), lend \( I_t P_t^A \), borrow \( I_t P_t^B \), and buy \( I_t \) units of \( B \). If, for convenience, \( I_t \) is defined as the negative of the position size when \( P_t^B \) exceeds \( P_t^A \), then the evolution of the investor's wealth, \( W \), can be written as

\[
dW_t = rW_t \, dt + I_t [rP_t^A - rP_t^B - dP_t^A + dP_t^B] - c|I_t| \, dt
\]

The first term represents the interest received from previously accumulated wealth. The second term gives the interest gain plus the capital gain or loss from the arbitrage position. The last term reflects the holding cost incurred for shorting \( I_t \) units of portfolio \( A \) or \( B \).\(^{15}\)

Assuming that the arbitrageur will maximize the expected utility of his terminal wealth completes the model's specification. More formally, he maximizes

\[ E [U(W_T)] \]

\(^{15}\) While equation (4) assumes that the arbitrageur's holding cost equals that which sets the riskless arbitrage bounds, the problem could be set up so that the arbitrageur had a higher cost. An equilibrium model, then, might include several classes of arbitrageurs, each with a different holding cost.
for some von Neumann-Morgenstern utility function, $U$.\(^{16}\)

The first step in solving for the optimal investment policy \(\{I_t\}\) is to rewrite the wealth equation (4) in terms of the underlying state variable, $z$. Using two transformed variables, $w_t = W_t e^{rt}$ and $i_t = I_t s(\tau) e^{rt}$, (4) can be rewritten as

$$\text{dw}_t = i_t \left[ \mu(z,\tau) dt + \sigma db_t \right]$$

where

$$\mu(z,\tau) = \rho z - \frac{1}{2} \sigma^2 \phi'/\phi' + c(\phi' e)/\phi'$$

and $\epsilon$ is the sign of $i_t$, i.e. 1 if $i_t > 0$, 0 if $i_t = 0$, and -1 if $i_t < 0$.

Let $V(w,z,\tau) = \max E_t \left\{ U(W_T) \right\}$, where $E_t$ denotes the expectation at $\tau$. By the principle of optimality in dynamic programming,

$$V(w,z,\tau) = \max_i E_t \left[ V(w + \text{dw},z + \text{dz},\tau + \text{d}\tau) \right]$$

By Ito's Lemma,

$$V(w + \text{dw},z + \text{dz},\tau + \text{d}\tau) = V(w,z,\tau) + V_w \text{dw} + V_z \text{dz} + V_\tau \text{d}\tau$$

$$+ \frac{1}{2} V_{ww} \text{dw}^2 + \frac{1}{2} V_{zz} \text{dz}^2 + V_{wz} \text{dw} \text{dz}$$

where subscripts denote partial derivatives. Substituting (3), (5) and (8) into (7) and taking expectations gives

$$0 = \max_i \left\{ (\mu V_w - \sigma^2 V_{wz}) i + \frac{1}{2} V_{ww} \sigma^2 i^2 - \rho z V_z - V_\tau + \frac{1}{2} V_{zz} \sigma^2 \right\}$$

Solving (9) numerically is by no means a trivial exercise. Imposing the boundary conditions $V(w, + \infty, \tau) = V(w, - \infty, \tau) = U(+ \infty)$ does not characterize the function $V(\cdot)$ because the growth conditions for large $|z|$ have not been specified. Furthermore, for arbitrary functions $U$ and $\phi$, it is

---

\(^{16}\) One might argue that because an arbitrage position based on deviations from fundamental value carries no market risk, discounted expected return evaluates the opportunity correctly. Two replies seem appropriate: 1) Even if investors are risk neutral with respect to the risk of this type of arbitrage, leverage constraints can induce risk aversion. Grossman and Vila (1991) show that the possibility that an investor's future opportunity set will be limited by borrowing restrictions causes risk neutral investors to display some risk aversion. Since leverage constraints probably apply to most potential arbitrageurs, this consideration could justify the assumption of risk aversion. 2) The most important arbitrageurs are probably judged on the performance of their arbitrage portfolios, as opposed to the performance of some diversified portfolio of longer-term holdings. In that case, they will be risk averse with respect to the risks of arbitrage so long as there are only a few arbitrage opportunities during most time periods.
not at all clear that there exists a bounded solution for \( V \). Appendix 1 finds the growth conditions for the case of a CARA utility function and \( \phi(z) = z/(1+z^2)^{1/2} \) and proves that a bounded solution exists. This section continues, therefore, by specializing the model to this case.

Letting \( U = -e^{-aw} \) for some positive constant \( a \) implies that

\[
V(w,z,\tau) = -e^{-aw}F(z,\tau)
\]  

(10)

for some function \( F \).\(^{17}\) Using (10) to calculate the partial derivatives of \( V \) in terms of the partial derivatives of \( F \) and substituting into (9) yields the following partial differential equation:

\[
F_{\tau} = \max_i \{ (\mu + \sigma^2 F_z)ai - \frac{1}{2}\sigma^2(ai)^2 \} - \rho z F_z + \frac{1}{2} \sigma^2(F_{zz}F_z^2)
\]  

(11)

The appropriate boundary conditions for equation (11) can be derived as follows. When \( \tau = 0 \), \( V(w,z,\tau) \) must equal \( U(w) = -e^{-aw} \), so \( F(z,0) = 0 \). Furthermore, as mentioned above, \( V(w,+,\infty) = V(w,-,\infty) = U(\infty) \) since, as the mispricing approaches the riskless arbitrage bound, the value function should approach its maximum. Since this maximum is zero, \( F(\infty,\tau) = F(-\infty,\tau) = 0 \).

Equation (11) will be solved numerically in the following section. The analysis has proceeded far enough, however, to derive some properties of an arbitrageur's optimal investment policy in the presence of holding costs. These properties are presented in proposition 1.

**PROPOSITION 1:** An arbitrageur following the optimal dynamic strategy implied by equation (11) and its boundary conditions

i) will take a position if and only if the mispricing, \( x \), is large enough,

ii) will take a finite position, and

iii) may take a position even if the instantaneous expected return from the position is negative.

\(^{17}\) Because the utility function exhibits constant absolute risk aversion, it is separable in the value of wealth and the value of arbitrage opportunities (see Hodges and Neuberger (1989) for a similar simplification).
Proof: See appendix 1.

The first and second part of proposition 1 tell a story consistent with casual empiricism about arbitrage activity. While price deviations from fundamental values encourage arbitrage activity, risk aversion and holding costs prevent investors from taking infinite positions. Consequently, the arbitrage activity of these investors might not be great enough to force prices back into line. This raises the possibility that consistent arbitrage activity can be sustained in equilibrium.

The third part of proposition 1 reveals an interesting fact about the optimal investment policy. The intuition parallels that of Merton's (1973) intertemporal CAPM with opportunity set changes. CARA investors like to hold assets which are negatively correlated with favorable changes in their opportunity set. Here, arbitrageurs do lose if the absolute value of $x$ rises, but then they have a better arbitrage opportunity available to them. Consequently, they are willing to take an arbitrage position even if it is expected to lose over the next instant.

III. An Application to the Treasury Bond Market

There are many redundant securities in the Treasury bond market, i.e. there are many bonds whose cash flows can be replicated by buying and holding a portfolio of other bonds. One class of such redundancies is formed by three Treasury bonds of the same maturity. If their coupons are $c_1$, $c_2$, and $c_3$, with $c_1 > c_2 > c_3$, then a portfolio composed of $(c_2-c_3)/(c_1-c_3)$ units of the first bond and $(c_1-c_2)/(c_1-c_3)$ units of the third bond exactly replicates the cash flows of one unit of the second bond.\footnote{Note that the accrued interest on the portfolio also matches the accrued interest on the second bond.}

All Treasury triplets which traded from January 1960 to December 1990 were identified from
the CRSP data files. Those containing callable bonds or flower bonds were discarded. If more than three bonds of the same maturity traded on a particular date, the three most recently issued were selected as the triplet for that date. A detailed list of the 42 triplets forming the data set is given in Appendix 2.

Month-end bid prices were obtained from CRSP for all triplets.\(^\text{19}\) Define \(P_2\) to be the price of the second bond in the triplet and \(P_R\) to be the price of its replicating portfolio. Define the mispricing, \(x\), as in the text, i.e. \(P_2-P_R\). For the opportunity cost of funds, \(r\), this paper uses a 3-year Treasury rate as reported by Moody's, but the results would not substantially change for other choices of \(r\).

Without data available for holding costs, it seems reasonable to choose a number from the range .25\% to .65\%, cited above. According to the theory described in the previous section, holding costs should be large enough so that no price deviations exceed the riskless arbitrage bound, \(s(\tau)\). As it turns out, there are some violations in the data set even at \(c=.65\%). Table I reports the number and percentage of observations which violates the no arbitrage conditions for different levels of the holding cost.

**INSERT TABLE I**

Table I reveals that a model of holding costs alone compares most favorably with a model of trading costs alone. The brokerage costs estimate of 1/128\%, for both the long and short legs of the arbitrage, plus a typical bid-ask spread of 3/32\% cited above, would result in 419, or 35.39\%, price deviations which exceed the level of trading costs. So, holding costs seem to be quite important, relative to trading costs, in foiling riskless arbitrage activity. Of course, a model which incorporates both holding costs and trading costs does best. With \(c=.006\) and these trading costs, for example, only 4 observations, or .34\%, violate the riskless arbitrage bounds.

\(^{19}\) The analysis was also done using the average of bid and ask prices. The results were qualitatively identical and quantitatively similar.
Returning to the model of this paper, it seemed reasonable to settle on $c = .006$. This level is in the range of prior beliefs and generates relatively few violations of the riskless arbitrage bounds. Using this level of holding costs and ignoring trading costs, figure 1 plots the individual mispricings as a function of the riskless arbitrage bound, $s(\tau)$. The two rays are the lines $x=s$ and $x=-s$.

The plot provides some additional evidence that holding costs are important inhibitors of arbitrage activity. In the model presented earlier, mispricings can take on any value between $s(\tau)$ and $-s(\tau)$. Consequently, one would expect the variability of the mispricings to increase in $s$. Looking at figure 1, this seems to be the case. More precisely, the standard deviation of $x$ is about .09 for $s<1$, .31 for $1<s<2$, .29 for $2<s<3$, and .43 for $s>3$.

The next step in applying the model to the data at hand is to estimate its parameters $\rho$ and $\sigma$. Recalling that $\phi(z) = z/(1+z^2)^{1/2}$, one can use (2) to transform $x$ and $s$ values into values of the underlying state variable, $z$. Then, changes in $z$ can be regressed on $z$ in order to estimate the parameters of the $z$ process given by (3). The results of the regression are reported in table II.

**INSERT TABLE II**

The estimate of $\rho$ is significantly negative, confirming the autoregressive nature of the deviation process.

The partial differential equation in (11) can now be solved numerically. Converting the parameters estimated from monthly data to an annual basis gives values of about 5.42 and .88 for $\rho$ and $\sigma$, respectively. The holding cost, $c$, was set to .6%, as discussed above. The opportunity cost of funds, $r$, was taken to be 9%, a value close to the sample average. Finally, the original maturity of

---

20 Figure 1 also shows that the price deviations are more often negative than not. In other words, the price of the replicating portfolio usually exceeds the price of the middle bond. This is consistent with a tax timing option or a tax clientele effect. See Constantinides and Ingersoll (1984), Litzenberger and Rolfo (1984), and Schaefer (1982).

21 The observations for which $|x| > s$ were dropped from the sample since $z$ is undefined in these cases.

22 Regressing changes in $x$ on $x$ also gives a significantly negative coefficient. Therefore, the functional form of $\phi$ is not responsible for the autoregressive nature of the $z$ process.
the triplet was set at 5 years.\textsuperscript{23}

The relevant range for the underlying state variable, \(z\), turns out to be well within \((-2,2)\). In other words, values of \(z\) which are larger in absolute value correspond to mispricings which are larger than those which appear in the data set.

Figure 2 graphs the optimal position size as a function of \(z\) for two different times to maturity. As predicted by Proposition 1, part i), for small mispricings it is not optimal to take a position. If the mispricing is large enough, however, position size increases in the absolute value of \(z\).

Figure 2 reveals that position size is not monotonic in time to maturity. There are two effects at work here. First, since positions of potentially longer maturities are likely to incur larger holding costs, investors would take on smaller positions. This might be called a risk effect. Second, since potentially longer horizons are likely to provide better arbitrage opportunities in the future, investors would be willing to take on larger positions. This might be called an option effect, and is related to the hedging properties of an arbitrage position raised by claim 3. When \(z\) is large, realization of current arbitrage profits is relatively more important than the promise of future mispricings, so the risk effect should dominate. When \(z\) is small, however, the potential of future arbitrage profits is relatively more important and the option effect should dominate. Consequently, as shown in the figure, when \(|z|\) is large the position size with 3.04 years to maturity is greater than the position size with 5 years to maturity. When \(|z|\) is small the position size with 5 years to maturity is the larger of the two.

Although not shown in the figure, the hedging effect described in part iii) of Proposition 1 can be examined through the numerical solutions. With 3.04 years to maturity for example, an arbitrageur will take a position even when \(|z|<.04\). The instantaneous expected return from the

\textsuperscript{23}From equation (11), one only need solve for \(a_i\), where \(a\) is the coefficient of absolute risk aversion. Therefore, without loss of generality, \(a\) was set equal to 1.
position is positive, however, only when $|z| > .2$. In terms of the dollar mispricing, $x$, the arbitrageur will take a position even for a mispricing of less than 6 cents for every $100 face value bought or shorted. The expected return over the next instant is positive, however, only when the mispricing exceeds 31 cents on the $100 position.

A similar comparative static might have arisen with respect to $\sigma$. Larger values of $\sigma$ make the arbitrage position riskier but raise the likelihood of future profit opportunities. One might reason, as before, that the risk effect would dominate for large mispricings while the option effect would dominate for small mispricings. As it turns out, however, the parameters determined by the data generate position sizes which increase in $\sigma$: the option effect dominates for all relevant levels of mispricing.

The certainty equivalent of arbitrage activity can be defined as the amount one would pay to become an arbitrageur in a particular triplet.\(^{24}\) Figure 3 graphs this certainty equivalent as a function of $z$ for three different times to maturity. As expected, the certainty equivalent increases in the absolute value of $z$. More interestingly, the certainty equivalent increases in time to maturity. It might have been the case that the risk effect would cause the certainty equivalent to fall with time to maturity since longer maturities run the chance of incurring larger holding costs. Once again, however, for the parameters in the Treasury market, the option effect reigns.

IV. Conclusion and Suggested Extensions

This paper models the effect of holding costs on the activities of risk averse arbitrageurs. The exogenous mispricing process was chosen to allow prices to deviate from fundamental value without allowing for riskless arbitrage opportunities. Solving for the optimal investment policy of arbitrageurs

\(^{24}\)Note that the certainty equivalent is measured in date-T dollars and is therefore comparable across maturities.
in such markets reveals behavior consistent with the casual empirical observation that professional arbitrageurs consistently take limited arbitrage positions. Furthermore, a risk effect and an option effect were identified as driving the comparative statics of optimal position size and of the value of being an arbitrageur. Data from the Treasury bond market supported the importance of holding costs in arbitrage activity, provided parameter estimates of the model, and led to some insights about the relative importance of the risk and option effects.

Two paths of future research emerge from this analysis. First, there is a need to collect better data about the size of holding costs in the various markets mentioned in section I. Second, the analysis of this paper concentrated on the investment problem facing arbitrageurs. The next step would be to model the factors generating price deviations from fundamental value and thus endogenize the deviation process. This avenue of research would clearly benefit from some preliminary progress in the first.
Appendix 1

Proof of PROPOSITION 1

Recall that the Bellman equation associated with the CARA arbitrageur's optimization problem is

$$F_r = \max_i \{ (\mu + \sigma^2 F_z)ai - \frac{1}{2} \sigma^2 (ai)^2 \} - \rho z F_z + \frac{1}{2} \sigma^2 (F_{zz} - F_z^2)$$  \hspace{1cm} (11)

with \( \mu \) given by

$$\mu(z,\tau) = \rho z - \frac{1}{2} \sigma^2 \phi''/\phi' + c(\phi' - \epsilon)/s\phi'$$  \hspace{1cm} (6)

and with \( \epsilon \) being the sign of \( i \).

The boundary conditions are

$$F(z,0) = 0; \hspace{1cm} (A1)$$

$$F(+\infty,\tau) = F(-\infty,\tau) = +\infty. \hspace{1cm} (A2)$$

Unfortunately, condition (A2) is not precise enough to characterize the value function \( F(\cdot,\cdot) \). Indeed it may be possible, a priori, that the arbitrageur reaches an unbounded value of \( F \) for every \( z \). It is therefore necessary to show that \( F \) can be bounded by the value function of a control problem which admits an explicit solution. This will be done later by the choice of a particular function \( \phi \). For the moment, assume that the function \( F \) and its derivatives are well defined and that \( \phi \) is a smooth increasing mapping from \((-\infty,\infty)\) into \((-1,1)\) such that \( \phi(-z) = -\phi(z) \).

The symmetry of the problem yields that

$$F(-z,\tau) = F(z,\tau) \hspace{1cm} (A3)$$

so that

$$F_z(0,\tau) = 0 \hspace{1cm} \text{for every } \tau. \hspace{1cm} (A4)$$

We can now proceed with the proof of Proposition 1.

Proof of i): From equation (9) the optimal investment \( i(z,\tau) \) is given by
\[ i(z, \tau) = i^L(z, \tau) \text{ if } i^L(z, \tau) > 0 \]  
(A5a)

\[ i(z, \tau) = i^S(z, \tau) \text{ if } i^S(z, \tau) < 0 \]  
(A5b)

\[ i(z, \tau) = 0 \text{ otherwise} \]  
(A5c)

with

\[ i^L(z, \tau) = \frac{1}{a} \frac{\rho}{\sigma^2} z - \frac{1}{2} \frac{\phi''}{\phi'} - \frac{c}{\sigma^2} \frac{1}{s(\tau)} \frac{1-\phi}{\phi'} + F_z \text{ when } i > 0; \]  
(A5d)

\[ i^S(z, \tau) = \frac{1}{a} \frac{\rho}{\sigma^2} z - \frac{1}{2} \frac{\phi''}{\phi'} + \frac{c}{\sigma^2} \frac{1}{s(\tau)} \frac{1+\phi}{\phi'} + F_z \text{ when } i < 0. \]  
(A5c)

From (A4), it follows that

\[ i^L(0, \tau) < 0; i^S(0, \tau) > 0. \]  
(A6)

Therefore \( i(0, \tau) = 0 \) and, by continuity, the optimal investment decision for small mispricings is to take no position. This proves i).

**Proof of ii):** ii) follows from (A5) above.

**Proof of iii):** From the symmetry of \( F(z, \tau) \) with respect to \( z \) it follows that \( F_z(z, \tau) \) has the same sign as \( z \), i.e. arbitrageurs are better off with large \( |z| \). As a result,

\[ i = \mu/\sigma^2 + F_z \]

can be positive even if \( \mu \) is negative and vice versa.

**Boundedness of \( F(z, \tau) \) and growth conditions:** As mentioned above, growth conditions for large \( z \) are critical for the numerical analysis. If \( \phi(z) = z/(1+z^2)^{1/2} \) then

\[ \mu(z, \tau) = \rho z + \frac{3}{2} \frac{\sigma^2}{1+z^2} - \frac{r}{1-e^{-\tau r}} \frac{1+z^2}{z+\epsilon \sqrt{1+z^2}}. \]  
(A7)

To provide the bounds and growth conditions of the value function, consider a class of control problems which admits an explicit solution. More precisely, for every function \( M(\tau) \), consider the
value function

\[ V^M(t)(w,z,t) = \max_i E[U(W_t)] \]

subject to the dynamics

\[ dw = \ i_i (M(t)z_i dt - \sigma db_i) \quad \text{and} \quad dz = -\rho z dt + \sigma db_i \quad t \leq T. \]

The function \( V^M(t) \) can be explicitly derived and is of the form

\[ V^M(t)(w,z,t) = - \exp \{-aW - F^M(t)(z,t)\} \]

where \( F^M(t) \) is quadratic with respect to \( z \):

\[ F^M(t)(z,t) = \frac{1}{2} K_2(t)z^2 + K_1(t)z + K_0(t). \]

By symmetry \( K_1(t) = 0 \). The functions \( K_2(t) \) and \( K_0(t) \) solve the ordinary differential equations

\[
\frac{dK_0}{dt} = \frac{1}{2} \sigma^2 K_2 \quad \text{and} \quad \frac{dK_2}{dt} = \frac{1}{\sigma^2} \left( M(t)^2 + 2(M(t) - \rho)K_2(t) \right) \quad \text{(A8)}
\]

and are easy to calculate.

Returning to the original problem, observe that there exists a positive constant \( \bar{M} \), such that

\[ |\mu(z,t)| \leq \bar{M}|z|, \quad \text{for every } z. \quad \text{(A9)} \]

Recalling that \( i_i \) and \( z_i \) must have the same sign, it follows that for every policy \( i_i \)

\[ dw = i_i [\mu(z_i,t)dt - \sigma db_i] \leq \bar{M} i z dt - \sigma i db_i. \quad \text{(A10)} \]

Hence

\[ V(w,z,t) = V^\bar{M}(w,z,t) < U(\infty). \]

It follows that the control problem faced by the arbitrageur is bounded and admits a smooth value function. (See Vila and Zariphopoulou (1989) on proving the regularity of the value function in a similar context.)

Growth conditions can be derived by considering the dominant term in \( \mu(z,t) \) for large \( z \).

From (A7) it follows that for large \( z \), \( \mu \) is approximately linear in \( z \) i.e.

\[ \mu(z,t) \approx (\rho - \frac{r}{1 - e^{-\rho t}}) z \equiv R(t)z. \quad \text{(A11)} \]
Let \( \tau^* \) be such that \( R(\tau^*) = 0 \). Then under certain parameter restrictions, easily satisfied by the data, for every \( \tau \leq \tau^* \) and every \( z \) the expected gain \( \mu(z, \tau) i \) is nonpositive for every arbitrage position \( i \). It follows that the function \( F(z, \tau) \equiv 0 \) for \( \tau \leq \tau^* \). For \( \tau \geq \tau^* \), growths conditions are obtained by considering the function \( V^{M(\tau)}(w, z, \tau) \) with

\[
M(\tau) = \max \{ R(\tau); 0 \}.
\]

Therefore the numerical analysis used the growth condition

\[
F(z, \tau) - \frac{1}{2} K_2(\tau) z^2 \quad \text{for large} \ z.
\]
## Appendix 2
### Treasury bond triplets

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Coupons</th>
<th>Start</th>
<th>End</th>
<th>Original Maturity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/15/62</td>
<td>3.25,3.625,4.0</td>
<td>11/60</td>
<td>1/62</td>
<td>1.21</td>
</tr>
<tr>
<td>11/15/71</td>
<td>3.875,5.375,7.75</td>
<td>5/70</td>
<td>10/71</td>
<td>1.46</td>
</tr>
<tr>
<td>2/15/72</td>
<td>4.0,4.75,7.5</td>
<td>8/70</td>
<td>1/72</td>
<td>1.46</td>
</tr>
<tr>
<td>11/15/79</td>
<td>6.25,6.625,7.0</td>
<td>11/76</td>
<td>10/79</td>
<td>2.96</td>
</tr>
<tr>
<td>8/15/81</td>
<td>7.0,7.625,8.375</td>
<td>8/78</td>
<td>7/81</td>
<td>2.96</td>
</tr>
<tr>
<td>5/15/82</td>
<td>7.0,8.0,9.25</td>
<td>11/78</td>
<td>4/82</td>
<td>3.46</td>
</tr>
<tr>
<td>5/15/84</td>
<td>9.25,13.25,15.75</td>
<td>5/81</td>
<td>4/84</td>
<td>2.96</td>
</tr>
<tr>
<td>8/15/84</td>
<td>6.375,7.25,13.25</td>
<td>2/81</td>
<td>7/84</td>
<td>3.46</td>
</tr>
<tr>
<td>8/15/85</td>
<td>8.25,9.625,13.125</td>
<td>8/82</td>
<td>7/85</td>
<td>2.96</td>
</tr>
<tr>
<td>5/15/86</td>
<td>7.875,9.375,13.75</td>
<td>5/83</td>
<td>4/86</td>
<td>2.96</td>
</tr>
<tr>
<td>11/15/86</td>
<td>11.0,13.875,16.125</td>
<td>11/83</td>
<td>10/86</td>
<td>2.96</td>
</tr>
<tr>
<td>2/15/87</td>
<td>9.0,10.875,12.75</td>
<td>2/84</td>
<td>1/87</td>
<td>2.96</td>
</tr>
<tr>
<td>5/15/87</td>
<td>12.0,12.5,14.0</td>
<td>5/84</td>
<td>4/87</td>
<td>2.96</td>
</tr>
<tr>
<td>11/15/87</td>
<td>7.625,11.0,12.625</td>
<td>11/84</td>
<td>10/87</td>
<td>2.96</td>
</tr>
<tr>
<td>5/15/88</td>
<td>8.25,9.875,10.0</td>
<td>5/85</td>
<td>4/88</td>
<td>2.96</td>
</tr>
<tr>
<td>11/15/88</td>
<td>8.625,8.75,11.75</td>
<td>11/85</td>
<td>10/88</td>
<td>2.96</td>
</tr>
<tr>
<td>5/15/89</td>
<td>6.875,9.25,11.75</td>
<td>5/86</td>
<td>4/89</td>
<td>2.96</td>
</tr>
<tr>
<td>11/15/89</td>
<td>6.375,10.75,12.75</td>
<td>11/86</td>
<td>10/89</td>
<td>2.96</td>
</tr>
<tr>
<td>5/15/90</td>
<td>8.25,7.875,11.375</td>
<td>5/87</td>
<td>4/90</td>
<td>2.96</td>
</tr>
<tr>
<td>8/15/90</td>
<td>7.875,9.875,10.75</td>
<td>8/87</td>
<td>7/90</td>
<td>2.96</td>
</tr>
<tr>
<td>11/15/90</td>
<td>8.0,9.625,12.96</td>
<td>11/87</td>
<td>10/90</td>
<td>2.96</td>
</tr>
<tr>
<td>8/15/91</td>
<td>7.5,8.75,14.875</td>
<td>8/88</td>
<td>12/90</td>
<td>2.96</td>
</tr>
<tr>
<td>11/15/91</td>
<td>6.5,8.5,14.25</td>
<td>11/88</td>
<td>12/90</td>
<td>2.96</td>
</tr>
<tr>
<td>2/15/92</td>
<td>6.625,9.125,14.625</td>
<td>2/89</td>
<td>12/90</td>
<td>2.96</td>
</tr>
<tr>
<td>5/15/92</td>
<td>6.625,9.0,13.75</td>
<td>5/89</td>
<td>12/90</td>
<td>2.96</td>
</tr>
<tr>
<td>8/15/92</td>
<td>7.25,7.875,8.25</td>
<td>8/89</td>
<td>12/90</td>
<td>2.96</td>
</tr>
<tr>
<td>11/15/92</td>
<td>7.75,8.375,10.5</td>
<td>11/89</td>
<td>12/90</td>
<td>2.96</td>
</tr>
<tr>
<td>2/15/93</td>
<td>6.75,7.875,10.875</td>
<td>2/83</td>
<td>11/87</td>
<td>9.97</td>
</tr>
<tr>
<td>2/15/93</td>
<td>7.875,8.25,10.875</td>
<td>12/87</td>
<td>1/90</td>
<td>5.13</td>
</tr>
<tr>
<td>2/15/93</td>
<td>8.25,8.375,10.875</td>
<td>2/90</td>
<td>12/90</td>
<td>2.96</td>
</tr>
<tr>
<td>3/15/93</td>
<td>7.625,8.625,10.125</td>
<td>5/90</td>
<td>12/90</td>
<td>2.96</td>
</tr>
<tr>
<td>8/15/93</td>
<td>8.625,8.75,11.875</td>
<td>8/88</td>
<td>7/90</td>
<td>5.13</td>
</tr>
<tr>
<td>8/15/93</td>
<td>8.0,8.75,11.875</td>
<td>8/90</td>
<td>12/90</td>
<td>2.96</td>
</tr>
<tr>
<td>11/15/93</td>
<td>8.625,9.0,11.75</td>
<td>9/88</td>
<td>12/90</td>
<td>5.13</td>
</tr>
<tr>
<td>8/15/94</td>
<td>8.625,8.75,12.625</td>
<td>6/89</td>
<td>12/90</td>
<td>5.13</td>
</tr>
<tr>
<td>11/15/94</td>
<td>8.25,10.125,11.625</td>
<td>9/89</td>
<td>12/90</td>
<td>5.13</td>
</tr>
<tr>
<td>2/15/95</td>
<td>7.75,10.5,11.25</td>
<td>12/89</td>
<td>12/90</td>
<td>5.13</td>
</tr>
<tr>
<td>5/15/95</td>
<td>10.375,11.25,12.625</td>
<td>5/85</td>
<td>2/90</td>
<td>9.96</td>
</tr>
<tr>
<td>5/15/95</td>
<td>8.5,10.375,11.25</td>
<td>3/90</td>
<td>12/90</td>
<td>5.13</td>
</tr>
<tr>
<td>11/15/95</td>
<td>8.5,9.5,11.5</td>
<td>9/90</td>
<td>12/90</td>
<td>5.13</td>
</tr>
</tbody>
</table>
References


Andersen, T., (1987), Currency and Interest-Rate Hedging, New York Institute of Finance.


Table 1
Frequency of riskless arbitrage bound violations
as a function of the level of holding costs

Table I reports the number and percentage of observations which violates the no arbitrage condition $|x_t| < s(\tau)$ for different levels of the holding cost. There are 1184 observations in all.

<table>
<thead>
<tr>
<th>Holding Cost</th>
<th># of Observations</th>
<th>% of Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>.0025</td>
<td>65</td>
<td>5.49</td>
</tr>
<tr>
<td>.003</td>
<td>49</td>
<td>4.14</td>
</tr>
<tr>
<td>.004</td>
<td>29</td>
<td>2.45</td>
</tr>
<tr>
<td>.005</td>
<td>23</td>
<td>1.94</td>
</tr>
<tr>
<td>.006</td>
<td>14</td>
<td>1.18</td>
</tr>
<tr>
<td>.0065</td>
<td>14</td>
<td>1.18</td>
</tr>
</tbody>
</table>
Table II
OLS regression results from estimating the process
\( dz = -\rho z dt + \sigma db \) with monthly observations.

The regression model is:
\[
Z_{t+1} - Z_t = \rho Z_t + \epsilon_t
\]
with 1125 observations.

The estimated value of \( \rho \) is:
\[
\hat{\rho} = -.452
\]
with a standard error of .0257.

The R squared is:
\[
R^2 = .216
\]

The estimated value of \( \sigma \) is taken to be the standard error of the residuals:
\[
\hat{\sigma} = .253
\]
Figure 1

Riskless Arbitrage Bound per $100 Face Value

The Mispricing of Treasury Triples

as determined by holding costs alone
Underlying State Variable, z

Position Size

- 5 Years to Maturity
- 3.04 Years to Maturity

Position Size in Face Value (Thousands)

$ p = 5.42; \theta = 8.8; c = 0.06; t = 5; r = 0.09 $