SHORT SELLING, UNWINDING, AND MISPRICING

ALEXANDER KEMPF*

This article highlights the impact of short selling restrictions and early unwinding opportunities on the relation between futures and spot prices. Within a multiperiod equilibrium model, the influence of optimal arbitrage trading on the mispricing is analyzed. Results concerning trade volume and level, mean reversion, and path dependence of the mispricing are provided. The empirical evidence suggests that short selling restrictions and early unwinding opportunities are influential factors for the mispricing behavior. They help explain empirical findings reported in the literature. © 1998 John Wiley & Sons, Inc. Jrl Fut Mark 18:903–923, 1998

INTRODUCTION

Arbitrage trading is an important factor determining the relation between spot and futures prices. In frictionless and competitive markets, arbitrage trading is the only factor. Arbitrageurs take offsetting positions in spot and futures markets so that the law of one price holds. In real markets, however, prices do not always behave as predicted by this simple arbitrage model. Several empirical studies find mispricings different from zero.[1] This suggests that market frictions prevent arbitrageurs from keeping the mispricing zero. However, even in the presence of market frictions, arbitrageurs are certainly expected to have some impact on the mispricing.

*For correspondence, University of Mannheim, Chair of Finance, 68131 Mannheim, Germany.

[1]See for example, Yadav and Pope (1990), Chung (1991), and Buehler and Kempf (1995). A comprehensive literature survey is Sutcliffe (1997), Chapter 5.1.

■ *Alexander Kempf is an assistant professor of finance at the University of Mannheim.*

The Journal of Futures Markets, Vol. 18, No. 8, 903–923 (1998)

CCC 0270-7314/98/080903-21

This article is concerned with the influence of index arbitrage trading on the mispricing process in the presence of market frictions. It specifically highlights the impact of short selling restrictions in the spot market and early unwinding opportunities on the mispricing process. Since short selling restrictions and early unwinding opportunities are empirically found to be important factors of the arbitrage demand, it is expected that they also have a major impact on the mispricing behavior.[2]

A multiperiod equilibrium model of the mispricing between a stock index and index futures is developed to study the impact of optimal arbitrage trading in the presence of market frictions on the mispricing. The structure of the equilibrium model is simple. In absence of arbitrage trading, the mispricing follows a random walk. The arbitrageur has a linear impact on the mispricing. He chooses his optimal dynamic arbitrage strategy. In doing so, he faces trading costs, holding costs, and market impact costs. Their sizes depend on the type of trade for two reasons. Firstly, building up a new arbitrage position leads to higher transaction costs than unwinding an open arbitrage position, as detailed in Brennan and Schwartz (1990). Secondly, building up a short arbitrage position (spot short, futures long) normally leads to higher transaction costs than building up a long arbitrage position (spot long, futures short) due to short selling restrictions in the spot market. Table I summarizes the transaction costs associated with different arbitrage trades.

For all trading strategies, market impact costs have to be taken into account. These costs may result from bid-ask-spreads or from market

[2]For example, Pope and Yadav (1994) focus on short selling restrictions. Sofianos (1993) and Neal (1995) provide evidence that early unwinding is a strategy heavily used.

TABLE I

Transaction Costs

	Building up long positions	Building up short positions	Unwinding long position	Unwinding short position
Market Impact Costs	yes	yes	yes	yes
Trading Costs	yes	yes	no	no
Holding Costs	no	yes	no	yes

Notes: The table shows the transaction costs associated with different arbitrage trading strategies.

power of arbitrageurs. Trading costs cover the commissions in the futures and spot market when building up and closing an arbitrage position. When building up an arbitrage position, trading costs have to be considered. The unwinding decision is independent from the trading costs, since they occur regardless of whether the position is unwound early.[3] Holding costs are important to arbitrageurs, since, in general, traders shorting the spot instrument don't have full use of the proceeds, or have to pay a borrowing fee. Short selling in the futures market is not restricted in this way. Consequently, holding costs only have to be taken into account when building up or unwinding a short arbitrage position.

The optimal arbitrage strategy is complex in the presence of transaction costs. For example, the optimal arbitrage volume differs for long and short arbitrage opportunities of the same absolute size due to short selling costs. The optimal arbitrage volume also depends on arbitrage trading in the past due to early unwinding opportunities and, in addition, on time to maturity. This has implications for the relation between spot and futures prices. For example, the mispricing behavior should depend on whether the mispricing exceeds transaction costs or not. Positive mispricing should behave differently when compared to negative mispricing. The mispricing behavior should depend on time to maturity and on its history. The hypotheses of the model are tested using intraday data for the German stock index DAX and DAX futures. The results provide evidence in support of the model. Overall, the tests suggest that short selling restrictions and early unwinding opportunities are influential factors for the behavior of the mispricing.

This paper is related to the theoretical work of Brennan and Schwartz (1990), Duffie (1990), and Tuckman and Vila (1992), who characterize the optimal dynamic arbitrage strategy with transaction costs and early unwinding opportunities. However, their models do not allow for determining the impact of arbitrage trading on the mispricing, rather they study the optimal strategy for a given mispricing process. The impact of arbitrage trading on the equilibrium mispricing is analyzed in Holden (1990) and Kumar and Seppi (1994) ignoring market frictions. Cooper and Mello (1990) allow for trading costs, and Tuckman and Vila (1993) consider holding costs of an arbitrageur. However, both articles assume that the transaction costs do not differ for long and short arbitrage positions. This assumption is questionable for index arbitrage trading as discussed above.

[3]See Brennan and Schwartz (1990) for a discussion of trading costs associated with index arbitrage trading.

MODEL

This article focuses on the impact of arbitrage trading on the mispricing $X(t) := F(t, T) - S(t)$. $S(t)$ denotes the price of a risky stock portfolio at time t and $F(t,T)$ the simultaneous price of a portfolio consisting of index futures and bonds with maturity T. This portfolio is constructed so that it provides identical cash flows to the index portfolio until T. At maturity, index futures are cash settled at the price $S(T)$. It is assumed that the index portfolio and the futures-bond portfolio are traded as single goods, which will be called index and future. In order to keep the model simple, it is assumed that the mispricing follows a random walk process in the absence of arbitrage and that the arbitrage demand has a linear impact on the mispricing:

$$X(t) = X(t - 1) + Z(t) - \lambda D(t) \tag{1}$$

$Z(t)$ denotes the innovation in the mispricing process at time t and $D(t)$ the arbitrage demand in the spot market. λ is a positive constant which characterizes the depth of the markets. It is assumed that arbitrageurs take opposite positions in the index and index futures so that $D(t)$ is a complete characterization of his arbitrage trading. An economic foundation of model (1) is given in Garbade and Silber (1983).

The mispricing process (1) depends on the optimal arbitrage demand. Next, the optimization problem of the arbitrageur is characterized, and the optimal equilibrium arbitrage demand is determined numerically. Assume that there is only one arbitrageur in this market and that he maximizes his expected total arbitrage profit until maturity T of the futures contract by following his optimal arbitrage strategy $\{D\}$:[4]

$$\max_{\{D\}} E_0 \sum_{t=1}^{T} P(t)(1 + r\Delta t)^{T-t} \tag{2}$$

E_0 denotes the expectation operator given the information at time $t = 0$. The arbitrage strategy consists of trades in the spot and futures markets at the trading dates $t = 1, \ldots, T$. Since the arbitrageur is restricted to pure arbitrage strategies, he invests his arbitrage profits $P(t)$ in a riskless bond yielding a constant interest rate, r, per year. The profit depends on the equilibrium mispricing, $X(t)$, arbitrage demand, $D(t)$, and explicit transaction costs, $C(t)$:

[4]This behavior is consistent with maximizing expected terminal wealth given that the arbitrageur is restricted to pure arbitrage trading, i.e., he invests his initial wealth in a riskless bond, and that he only has arbitrage income.

$$P(t) = X(t)D(t) - C(t) \tag{3}$$

According to (3), the arbitrageur makes, neglecting transaction costs, an arbitrage profit, when the mispricing is positive, i.e., $F(t,T) > S(t)$, and he buys the spot and sells the future, i.e., $D(t) > 0$. The opposite trading strategy is profitable when the mispricing is negative. Since the arbitrageur has market power according to eq. (1), his arbitrage demand does not approach infinity as soon as a nonzero mispricing occurs. The explicit transaction costs cover trading costs, C_1, and short selling costs, C_2.[5] The trading costs are characterized by the following function:

$$C_1(t) = \begin{cases} c_1|D(t)|, & \textit{if } Q(t-1)D(t) \geq 0 \\ 0, & \textit{if } Q(t-1)D(t) < 0 \textit{ and } |D(t)| \leq |Q(t-1)| \\ c_1|Q(t)| & \textit{if } Q(t-1)D(t) < 0 \textit{ and } |D(t)| > |Q(t-1)| \end{cases} \tag{4}$$

The arbitrageur pays round trip transaction costs only when building up new arbitrage positions. There are no (explicit) transaction costs when unwinding. Therefore, the trading costs at time t not only depend on the trade volume $D(t)$ but also on the number of open arbitrage positions, $Q(t-1)$, reflecting arbitrage trades until $t-1$. It is assumed that the arbitrageur cannot hold long arbitrage positions (spot long, future short) and short arbitrage positions (spot short, future long) simultaneously, so the endowment state can be characterized by the single state variable

$$Q(t) := \sum_{s=1}^{t} D(s).^6 \tag{5}$$

Short selling restrictions are only considered as short selling costs, $C_2(t)$. Other institutional constraints, such as tick rules, are neglected. This is done as the model will be tested using data on German markets where short selling costs are the only issue. Short selling costs differ from trading costs in two respects. First, they depend on the period of time over which arbitrage positions are carried. Second, they differ for long and short arbitrage positions. It is assumed that the arbitrageur pays holding costs c_2 for each short arbitrage position (spot short, future long) that he holds. There are no short selling costs when selling futures and buying the index:

[5]Implicit transaction costs (market impact costs) of the arbitrageur are taken into account through eq. (1), which describes the impact of arbitrage trading on the equilibrium mispricing.

[6]This assumption is taken for computational convenience. This assumption can be easily relaxed, but the computational burden increases significantly, since the optimization problem in that case depends on three state variables instead of two.

$$C_2(t) = \begin{cases} c_2|Q(t)| & if\ Q(t) < 0 \\ 0 & if\ Q(t) \geq 0 \end{cases} \tag{6}$$

The optimization problem of the arbitrageur is specified by eqs. (1) to (6). It can be shown that an optimal solution to this problem exists. This solution is determined numerically using dynamic optimization techniques. The state variables are the equilibrium mispricing and the number of open arbitrage positions before the arbitrageur trades. The control variable is the endowment $Q(t){:} = Q(t-1) + D(t)$ after trade at time t. In the numerical procedure, the endowment is restricted to integer values. The market liquidity parameter, λ, is set to 0.1 and the time interval $\Delta t = 1/360$ equals one day. The trading costs, $c_1 = 5$, and the holding costs, $c_2 = 0.15$, are chosen to represent approximately the transaction costs of an index arbitrageur in German markets. The variance estimation of the mispricing process, $\sigma^2 \Delta t = 5.3$, is obtained from data on DAX index and DAX futures. The interest rate per year, r, is assumed to be 0.075. The results presented in the remainder of this section are robust to parameter variations, i.e., the qualitative conclusions do not change.[7]

[7]This is not shown here for economy of space.

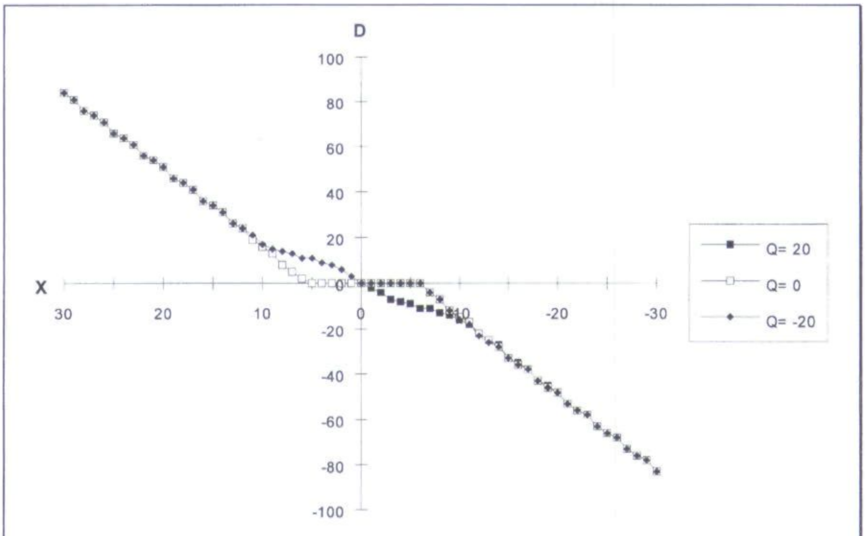

FIGURE 1
Optimal arbitrage trading D plotted against the mispricing X for various open arbitrage positions Q and a given time to maturity $\tau = 3$.

Figure 1 shows the optimal arbitrage demand, $D(t)$, for different combinations of the state variables. From the arbitrage demand, the mispricing process can be easily derived using eq. (1). The time to maturity is assumed to be three days, i.e., $\tau = 3$.

The main result shown in Figure 1 is that the size and location of the no-arbitrage window depends on the number of open arbitrage positions, which reflects the former arbitrage trades. The no-arbitrage window is the largest when the arbitrageur has no open arbitrage positions $[Q(t - 1) = 0]$, since he cannot unwind positions. But even in this case the no-arbitrage window is asymmetric due to holding costs when selling stocks short. Outside the no arbitrage window, the trade volume of the arbitrageur increases with the absolute level of mispricing. If the arbitrageur holds an initial position $[Q(t - 1) \neq 0]$, the no-arbitrage window is smaller and the asymmetry is even more pronounced due to the early unwind opportunity. From those results one can draw several conclusions on the mispricing behavior. First, due to short selling restrictions, the arbitrageur trades less at negative mispricing than at positive mispricing. Consequently, the mean reversion is weaker for negative than for positive mispricing. Second, short selling restrictions lead to a negative average mispricing. Third, the model predicts that the absolute level of the negative mispricing increases with time to maturity, since the holding costs associated with short arbitrage positions increase with time to maturity. Finally, due to the early unwind option, the optimal current arbitrage demand depends on the past arbitrage demand, which was mainly determined by past mispricing. Therefore, the mispricing is path dependent.

Figure 2 shows the optimal arbitrage demand for varying mispricing, given 20 open short arbitrage positions and 30 days to maturity. It can be expected that the influence of the holding costs is stronger due to the longer time to maturity as compared with $\tau = 3$.

Figure 2 shows an asymmetric no arbitrage window, $-25 \leq X(t - 1) \leq -7$. It is optimal for the arbitrageur to unwind short arbitrage positions even when the mispricing is zero or slightly negative. This unwind strategy pays for the arbitrageur, since he reduces the holding costs resulting from his short positions. Therefore, a negative mispricing is not driven back to zero by arbitrageurs, but back to a negative level of mispricing that is determined by short selling costs. A similar trading strategy does not pay for an arbitrageur holding long arbitrage positions. Therefore, the model predicts that the mean reversion of the mispricing is around a negative mean. Since the influence of short selling costs increases with time to maturity, one expects in addition that the absolute value of the negative mean increases with time to maturity.

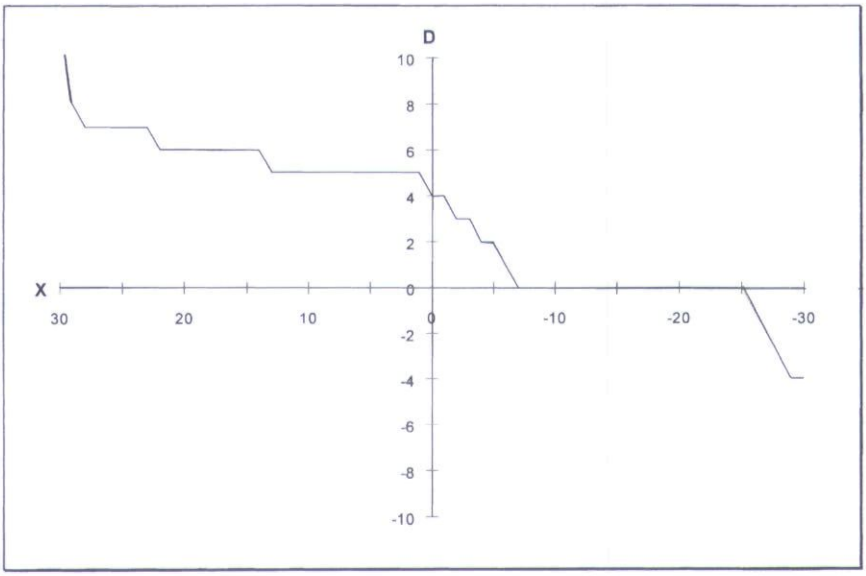

FIGURE 2
Optimal arbitrage trading D plotted against the mispricing X for a given time to maturity $\tau = 30$ and 20 open short arbitrage positions ($Q = -20$).

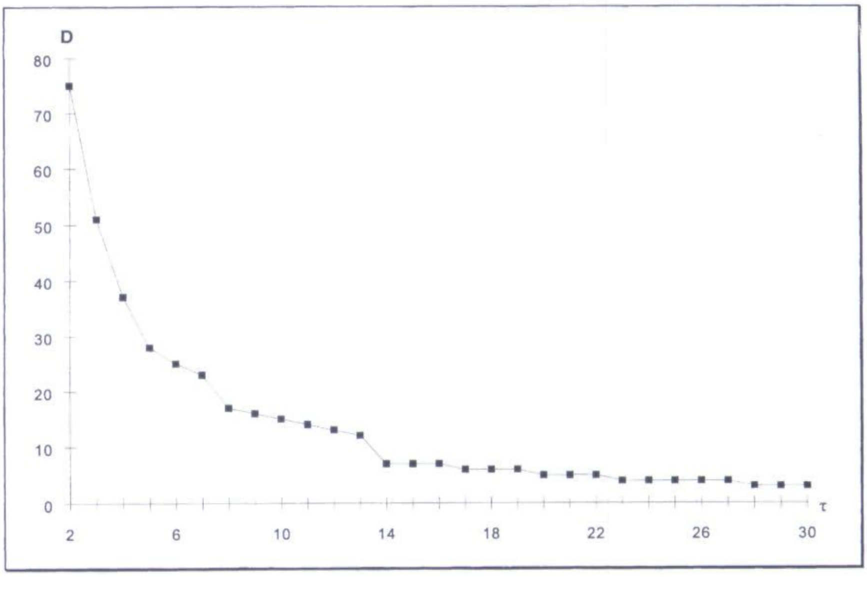

FIGURE 3
Optimal arbitrage trading D plotted against time to maturity τ for a given mispricing $X = 20$ and no open arbitrage positions ($Q = 0$).

In Figure 3, the relation between the arbitrage aggressiveness and time to maturity is analyzed. It shows that, for a given mispricing outside the no arbitrage band, the arbitrage demand is smaller the longer the time to maturity is. This result reflects the fact that the arbitrageur has market power and can increase his total arbitrage profit by splitting his total demand into several orders which are executed consecutively. This strategic behavior of the arbitrageur is similar to the trading strategy of the monopolistic informed investor in Kyle (1985). Therefore, the arbitrage trade volume and consequently the total trade volume decreases with time to maturity. In addition, the model predicts that the mean reversion of the mispricing is smaller, the longer the time to maturity is. Summing up, the theoretical model leads to the following hypotheses which will be tested later in this article:

H1: The average mispricing is negative.

H2: The absolute level of the negative mispricing increases with time to maturity.

H3: The mean reversion is around a negative mean increasing with time to maturity.

H4: The mean reversion is greater for positive than for negative mispricing.

H5: Mean reversion exists only when there is arbitrage trading.

H6: The mean reversion decreases with time to maturity.

H7: Total trading volume decreases with time to maturity.

H8: The mispricing is path dependent.

EMPIRICAL RESULTS

Data Description

The data set consists of 5-minute data from the German stock index DAX and DAX futures nearest to deliver.[8] In addition, futures trading volumes are provided. Interbank interest rates are available on a daily basis for overnight, 1-month, and 3-month money.[9] The research period begins on the first trading day in DAX futures, 23 November 1990, and ends on the last trading day of the DAX future maturing in December 1992, i.e., 17 December 1992. The study is based on the daily time window from 10.45

[8]For a description of the DAX and DAX futures, see, for example, Buehler and Kempf (1995).

[9]The index data are provided by the Frankfurt Stock Exchange, the futures and interest rate data by the German Finance Data Base.

a.m. to 1.30 p.m. in order to avoid possible opening effects in the spot market. The days for which the prices of one market are not available in entirety are eliminated from the data set. There are 492 trading days with 34 observations per day left, i.e., the total number of observations is 16,728.

Each index value $S(t)$ is paired with the valid futures bid, $f^b(t,T)$, and the futures ask quotes, $f^a(t,T)$, of the nearest to deliver contract. From the bid and ask quotes, the future midquote, $f(t,T)$, is calculated. The mispricings $X(t)$, $X^a(t)$, $X^b(t)$ between index and future are constructed using the following formulas:[10]

$$X(t) := f(t, T)e^{-r(t,T)[T-t]} - S(t) + c(t)S(t)\,[T - t] \tag{7a}$$

$$X^a(t) := f^a(t, T)e^{-r(t,T)[T-t]} - S(t) + c(t)S(t)\,[T - t] \tag{7b}$$

$$X^b(t) := f^b(t, T)e^{-r(t,T)[T-t]} - S(t) + c(t)S(t)\,[T - t] \tag{7c}$$

$r(t,T)$ is the interbank interest rate per year for the maturity $[T - t]$ years. It is linearly interpolated from the two interbank rates that match the maturity best. $c(t)S(t)[T - t]$ denotes the payments occurring from holding a stock position for the period of time $[T - t]$. The holding revenue $c(t)S(t)[T - t]$ associated with a long spot position results from delivering the stocks to a security lending company. The total holding costs associated with a short spot position result from borrowing the stocks from a security lending company. They are roughly twice as large as the holding revenues of an arbitrageur. Therefore, 50% of the holding costs, which equals the holding revenues of a long spot position, are considered in the mispricing formula. The other part of the holding costs is taken into account as explicit transaction costs which occur only when shorting the index. The total holding cost rate per year associated with shorting the index, $2c(t)$, decreased over the research period. At the beginning of the research period the borrowing fee per year was 2.5% of the index value. At the end of the research period, the annual borrowing fee has decreased to 0.75%. For each point of time, the borrowing fee per year is constructed by linearly interpolating these borrowing rates. The round trip trading costs are assumed to be 0.3%. The empirical analysis of the mispricing is based on the midquote mispricing, $X(t)$. The ask mispricing, $X^a(t)$, and the bid mispricing, $X^b(t)$, are used to determine arbitrage opportunities.

[10]Dividends are not taken into account in (7a), (7b), and (7c), since the DAX index is a performance index corrected for dividend payments. See also Buehler and Kempf (1995).

Results

The theoretical model provides eight testable hypotheses (summarized earlier in this article), which are tested in this section. The consideration of holding costs in the equilibrium model leads to two hypotheses concerning the level of mispricing. The average mispricing is negative (H1) and its absolute level increases with time to maturity (H2).

Results concerning H1 are reported first. The average mispricing is estimated as -2.57 index points. It is different from zero at the 1% significance level[11] supporting H1. This result is consistent with earlier studies of the mispricing in DAX futures performed by Gruenbichler and Callahan (1994), Bamberg and Roeder (1995), and Buehler and Kempf (1995). In other European markets, similar results are found:[12] Berglund and Kabir (1995) for the Dutch market, Puttonen and Martikainen (1991) and Puttonen (1993) for the Finnish market, Yadav and Pope (1990) and Strickland and Xu (1993) for the British market, and Stulz, Wasserfallen, and Stucki (1990) for the Swiss market all report negative mispricing on average. This indicates that short selling restrictions are—at least in Europe—a main cause of mispricing. For U.S. markets short selling restrictions do not seem to be that important (Neal, 1995), and mixed results concerning the average mispricing are found. An excellent overview is Suttcliffe (1997), Chapter 5.1. A positive average mispricing, as reported for example by Bhatt and Cakici (1990), is inconsistent with the model proposed earlier in this article. To explain a positive mispricing, the random walk assumption has to be given up. For example, suppose informed traders prefer trading in the futures markets due to lower transaction costs than in the spot market. Such behavior is consistent with empirical lead-lag studies reporting a lead of the futures market [see, e.g., Kawaller, Koch, and Koch (1987), Stoll and Whaley (1990), and Chan (1992)]. As a consequence, a positive mispricing occurs when good news reaches the market and a negative mispricing occurs when bad signals are received. Hietala, Jokivoulle, and Koskinen (1995) find empirical support for such a relation between mispricing and private signals.

The second hypothesis (H2) states that the mispricing is more negative the longer the time to maturity. It is tested using the following linear regression model:[13]

[11]Since the mispricing is strongly positively autocorrelated, the standard errors are constructed using the method of Newey and West (1987). The t-statistic is -13.85.

[12]An exception is Yadav and Pope (1994), who find mixed results concerning the sign of the mispricing.

[13]As Durlauf and Phillips (1988) point out, eq. (8) can be estimated only when the mispricing is trend stationary. Therefore, an Augmented Dickey-Fuller test with time trend was performed. The mispricing was found to be trend stationary so that eq. (8) can be estimated. The estimation results are shown following eq. (10).

$$X(t) = \alpha + \beta\tau + \varepsilon(t) \tag{8}$$

where τ denotes the time to maturity in days. The estimations for the parameters α and β are 0.004 and -0.060, respectively. The t-statistics calculated using the method of Newey and West (1987) are 0.01 and -8.65. The constant is not significant and the time parameter is significantly different from zero at the 1% level. The findings support the hypothesis that futures are more undervalued the longer the time to maturity. Again, similar results have been reported earlier. However, most of these studies analyze the relation between absolute mispricing and time to maturity [see the overview in Sutcliffe (1997), Chapter 5.2]. They all find a positive relation. Signed mispricing and time to maturity have been analyzed by Yadav and Pope (1990) and Buehler and Kempf (1995). Whereas Yadav and Pope (1990) find a negative β for early but not for later contracts in the U.K. market, Buehler and Kempf (1995) report $\beta < 0$ for all contracts in the German market.

Overall the results found here are consistent with the interpretation that the holding costs lead to a negative mispricing increasing with time to maturity. However, other explanations for a negative mispricing like the tax timing option have been suggested.[14] In order to distinguish between these possible explanations, the mispricing series is divided into two subsamples, one covering the mispricing of futures maturing before 1992 (8,704 observations) and the other covering the later contracts (8,024 observations). As noted in the data description, the holding costs decreased over the research period, while the tax treatment of profits did not change. If the negative mean and the time dependence of the mispricing are caused by holding costs, it can be expected that the average mispricing is more negative and time dependent in the early contracts compared to the later contracts. The tests support this hypothesis. The average mispricing in the early contracts is -3.83 and in the later contracts -1.20. Both mispricings are significant at the 1% level. The difference between both values is significant at the 1% level, too. The value of the F-test is 58.03. The time dependence of mispricing in the two subsamples is tested allowing for dummy variables, D_j, in the regression eq. (8). The model

$$X(t) = \alpha + \sum_{j=1}^{2} D_j\beta_j\tau + \varepsilon(t) \tag{9}$$

is estimated where dummy variable, D_1, equals one for all contracts maturing before 1992 and zero otherwise. For the contracts maturing in

[14]See Cornell and French (1983) for the tax timing option.

1992, $D_2 = 1$. For all other contracts, $D_2 = 0$. The parameter, β_1, of the early contracts is estimated as -0.091 and the parameter, β_2, of the late contracts as -0.031. Both parameters are significant at the 1% level, and the t-statistics are -12.48 and -4.52, respectively. The difference between both parameters is significant at the 1% level, the value of the F-test is 66.42. The time dependence of the mispricing is stronger in early contracts than in late contracts which was also found in Buehler and Kempf (1995).

Implications of the model concerning the mean reversion of mispricing are tested next. The results presented here complement earlier evidence on mean reversion reported by Yadav and Pope (1992). The first prediction of the model is that the mispricing fluctuates around a negative, time dependent value. This hypothesis (H3) can be tested by performing an Augmented Dickey-Fuller test with time trend. The following model is estimated:

$$\Delta X(t) = \alpha + \beta\tau - \mu X(t-1) + \sum_{i=1}^{p} \phi_i \Delta X(t-i) + \varepsilon(t) \quad (10)$$

In order to correct for autocorrelation in mispricing changes, p lagged mispricing changes are used. Following the procedure of Hall (1994), $p = 11$ lagged variables are chosen. H3 states that the mean reversion parameter, μ, is positive and that the parameter β is negative. The mean reversion parameter estimated is 0.020. The τ_τ-test statistic is 8.96 indicating a significance level of 1%.[15] This result is similar to Yadav and Pope (1992) who report a significant mean reversion estimating (10) for U.S. and U.K. data. The parameter β [not reported by Yadav and Pope (1992)] is estimated as -0.001. It is significantly different from zero at the 1% level (t-value = -3.05). The result supports hypothesis H3.

Due to higher transaction costs when selling stocks short, the theoretical model predicts that the mean reversion is weaker for negative mispricing than for positive mispricing (H4). This hypothesis can be tested by allowing for two dummy variables ($J = 2$) in (10).

$$\Delta X(t) = \alpha + \beta\tau - \sum_{j=1}^{J} D_j \mu_j X(t-1) + \sum_{i=1}^{p} \phi_i \Delta X(t-i) + \varepsilon(t) \quad (11)$$

For a positive mispricing, $D_1 = 1$, and $D_2 = 1$ indicates a negative mispricing. Again, $p = 11$ is chosen. H4 states that $\mu_1 > \mu_2$. The parameters

[15] See Fuller (1976) for the critical values.

estimated are 0.036 and 0.008, respectively. The difference is significant at the 5% level (F-test = 5.72) supporting H4.

The model predicts that mean reversion is arbitrage induced and, consequently, observed only when arbitrageurs trade (H5). In order to test H5 one needs to know when arbitrageurs trade. Since arbitrage trading cannot be identified from the mispricing series, a trading strategy has to be assumed which allows to distinguish whether arbitrageurs trade. It is assumed that arbitrageurs trade whenever the mispricing is outside an asymmetric no-arbitrage bound.[16] If the bid mispricing exceeds the trading costs, i.e. $X^b(t) > 0.003\ S(t)$, the arbitrageur builds up a long arbitrage position. If the ask mispricing is negative and its absolute value larger than trading costs and holding costs, i.e., $-X^a(t) > \{0.003 + c(t)[T - t]\}\ S(t)$, the arbitrageur builds up a short arbitrage position. This strategy is conservative in the sense that it may pay for arbitrageurs to build up arbitrage positions even when the mispricing does not exceed transaction costs as they get a valuable early unwind option (Brennan and Schwartz, 1990). It is found that the arbitrageur can build up long arbitrage positions 36 times, and short arbitrage positions 1,798 times. Since almost all arbitrage opportunities are short arbitrage opportunities, only the unwinding of short positions is considered. Whether an arbitrageur unwinds short positions depends on the current mispricing and the number of open short positions. Since the number of open positions is unobservable, a rough approximation is used. It is assumed that the arbitrageur is always able to unwind short positions when two conditions are met. First, short arbitrage positions could be built up at least once in the past. Second, the current bid mispricing is such that it pays to unwind short positions but not to build up new long positions, i.e., $0.003\ S(t) \geq X^b(t) > -c(t)\ [T - t]S(t)$. This approximation provides an upper bound for the number of unwinding opportunities. Thus, the impact of unwinding trades actually carried out could be larger than the estimated value. Based on this arbitrage trading assumption, hypothesis H5 can be tested by estimating $J = 4$ mean reversion parameters in (11). The definition of the dummy variables D_j follows from the trading strategy described above. If long arbitrage positions are built up, $D_1 = 1$. If short arbitrage positions can be built up, $D_2 = 1$. If short arbitrage positions can be unwound, $D_3 = 1$. If there is no arbitrage trading, $D_4 = 1$. Whenever the dummies are not equal to 1, they are zero. In (11) an augmentation of $p = 11$ is chosen following the procedure of Hall (1994). The

[16]The asymmetric no-arbitrage bound results from short selling costs. Yadav, Pope, and Paudyal (1994) estimate no-arbitrage bounds using a threshold autoregressive model. Their findings support the assumption of an asymmetry.

TABLE II

Mean Reversion and Arbitrage Trading

	$j=1$	$j=2$	$j=3$	$j=4$
number of observations[a]	21	1,173	2,599	7,031
μ_j	0.135	0.011	0.045	0.007
τ_τ-test statistic	3.57*	3.57**	3.76**	1.76

Notes: This table shows the mean reversion parameter μ_j in four different regimes. $j = 1$ ($j = 2$) indicates that arbitrageurs build up long (short) arbitrage positions. $j = 3$ indicates that arbitrageurs unwind short positions. $j = 4$ indicates that arbitrageurs do not trade.

[a]In order to avoid overnight effects when estimating mean reversion coefficients, overnight mispricing changes are excluded. In addition, only lagged mispricing changes of the same trading day are used. This procedure reduces the total number of observations per day by 12 observations. The total number of usable observations reduces to 10,824.

*Significant at the 10-percent level.

**Significant at the 5% level.

results of the estimation are summarized in Table II.

Hypothesis H5 is supported by the data, since μ_1, μ_2, and μ_3 are significantly above zero and μ_4 is insignificant. There is a significant mean reversion when arbitrageurs trade. This is the case not only when they build up arbitrage positions ($\mu_1 > 0, \mu_2 > 0$), but also when they unwind short arbitrage positions ($\mu_3 > 0$).[17] Without arbitrage trading, there is—despite the large number of observations—no significant mean reversion, the mispricing follows a random walk. This result suggests that the mean reversion of the mispricing in Germany is arbitrage induced,[18] and that German arbitrageurs not only build up arbitrage positions but also unwind these positions early. The results here are again similar to those of Yadav and Pope (1992) for U.S. and U.K. markets. They report a strong mean reversion when new arbitrage positions can be built up, but no mean-reversion when arbitrageurs cannot trade. The impact of early un-

[17]Again, the mean reversion is significantly stronger when long arbitrage positions are built up in comparison to building up short positions. The value of the F-test is 10.66. This result supports the earlier interpretation that arbitrageurs trade less aggressively when futures are undervalued due to short selling restrictions in the spot market.

[18]Miller, Muthuswamy, and Whaley (1994) argue that mean reversion might not be arbitrage induced but a statistical illusion due to infrequent trading. They support their model by showing that mean reversion exists even without arbitrage trading. In contrast, our results suggest that mean reversion exists only when arbitrageurs trade. This indicates that mean reversion is arbitrage induced. In addition, Yadav and Pope (1992) show in a simulation study that infrequent trading, as modeled by Miller, Muthuswamy, and Whaley (1994), can lead to negative autocorrelation in the mispricing, but not to mean reversion in the sense of eq. (10). This finding also contradicts the interpretation that the mean reversion reported here is a statistical illusion due to infrequent trading.

winding on mispricing is more pronounced here than in Yadav and Pope (1992), who find only a weak mean reversion in this case. The result that early unwinding has an important influence on the mispricing is consistent with the findings of Sofianos (1993) and Neal (1995) for the U.S. market, who report that arbitrageurs typically unwind their positions before maturity. This should lead to mean-reversion even when building up new arbitrage positions is not profitable. Similarly, Merrick (1989) concludes that early unwinding may explain why arbitrage activity can be heavy even when the mispricing is within the no-arbitrage bounds for building up positions.

The model predicts that arbitrage trading decreases with time to maturity. Therefore, the mean reversion of the mispricing and total trade volume are expected to decrease with time to maturity (H6 and H7). Hypothesis H6 is tested by analyzing the mean reversion parameter in three time to maturity intervals. Since mean reversion results from arbitrage trading, the time dependence should occur only when arbitrageurs trade. No time dependence is expected when there is no arbitrage trading. The hypothesis is tested allowing for six dummies ($J = 6$) in eq. (11). The definition of the dummy variables D_j follows. $D_1 = 1$, if $\tau \leq 30$ and arbitrageurs trade.[19] $D_2 = 1$, if $30 < \tau \leq 60$ and arbitrageurs trade. $D_3 = 1$, if $60 < \tau$ and arbitrageurs trade. $D_4 = 1$, if $\tau \leq 30$ and arbitrageurs do not trade. $D_5 = 1$, if $30 < \tau \leq 60$ and arbitrageurs do not trade. $D_6 = 1$, if $60 < \tau$ and arbitrageurs do not trade. The mean reversion parameters estimated are shown in Table III. Again, $p = 11$ is chosen.

Table III shows that the mean reversion parameter is smaller the longer the time to maturity, when arbitrageurs trade. There is no systematic time dependence of the mean reversion parameter without arbitrage trading. The test statistics, however, show that there is a significant mean reversion at the 1% level only when arbitrageurs trade and time to maturity is less or equal to 30 days. For times to maturity between 30 and 60 days, the significance level is close to 10%. For longer times to maturity there is no significant mean reversion despite arbitrage trading. Without arbitrage trading there is no significant mean reversion in all maturity classes. These results suggest that mean reversion is arbitrage induced and that the influence of arbitrageurs on the mispricing is stronger the shorter the time to maturity. Again, a similar result is reported by Yadav and Pope (1992). They find a significant decrease in the mean reversion with time to maturity for U.S. and U.K. data.

[19]Arbitrageurs are assumed to trade according to the trading strategy detailed above. Arbitrage trading includes building up long or short positions and unwinding short positions.

TABLE III

Mean Reversion, Arbitrage Trading, and Time to Maturity

	$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$
number of observations	1,966	1,086	741	2,060	2,456	2,515
μ_j	0.030	0.017	0.008	0.007	0.015	0.004
τ_τ-test statistic	4.57***	3.10	2.04	0.71	2.77	0.84

Notes: This table shows the mean reversion parameter μ_j in six different regimes. $j = 1$, $j = 2$, and $j = 3$ indicate that arbitrageurs trade and that time to maturity is up to 30 days, between 31 and 60 days, and more than 60 days, respectively. $j = 4$, $j = 5$, and $j = 6$ indicate that arbitrageurs do not trade and that time to maturity is up to 30 days, between 31 and 60 days, and more than 60 days, respectively.
***Significant at the 1% level.

Hypothesis H7 is tested by focusing on the daily trade volume in futures for the three time to maturity intervals defined above. There is earlier evidence for the U.K. market by Board and Sutcliffe (1990) that trade volume decreases with time to maturity supporting H7. For German data, however, H7 is rejected. It is found that the average number of futures traded per day is 6323 for $60 < \tau$, 6794 for $30 < \tau \leq 60$ and 6413 for $\tau \leq 30$. The trade volumes do not differ significantly from each other. A possible explanation for this result is that the volume of non-arbitrage trading changes over time and dominates the impact of changes in arbitrage trading. This dominance can easily happen since arbitrage trading might account only for a small proportion of total trading volume. The proportion of arbitrage trading in German markets is unknown, but Miller, Muthuswamy, and Whaley (1994) and Holden (1995) provide estimates for U.S. markets. They find that only a small proportion of total trading volume is arbitrage induced, namely 3.8% and 2.4%, respectively. A possible explanation for the low total trade volume in the expiration month is the rolling over of non-arbitrage trading. This hypothesis is supported by the finding that the average daily trade volume increases to 6889 when the last two weeks, during which unwinding is probably most important, are neglected in the analysis. Then the daily trade volume in the near period (6889) is larger than the middle period (6794) and in the far period (6323). However, the differences are still insignificant.

The last hypothesis (H8) of the model states that the mispricing is path dependent, since the trading strategy of the arbitrageur depends on the mispricing behavior in the past due to the unwinding opportunity.

Early unwinding is also a possible explanation for the asymmetric distribution of arbitrage opportunities reported above. The hypothesis of path dependence (H8) is tested following MacKinlay and Ramaswamy (1988). For each futures contract separately, it is determined how often the mispricing follows certain paths. The paths are subsequently defined. The mispricing is said to follow the long-long path when a long arbitrage position can be built up next, given that the last arbitrage opportunity was a long arbitrage opportunity and that the mispricing has crossed zero in the meantime. The number of observations for the long-short path is increased by 1 when a short arbitrage position can be built up, given that the last arbitrage opportunity was a long arbitrage opportunity and that the mispricing has crossed zero. The short-long and the short-short paths are defined analogously. If early unwinding has an impact on the mispricing behavior, long-long and short-short paths should occur more often than long-short and short-long paths. Overall, only 3 long-long paths, 3 long-short paths, 3 short-long paths, and 12 short-short paths occur. The very small number of observations indicates, again, a strong persistency of the mispricing. The probability that—at a given mispricing of zero—the next arbitrage opportunity is the same type as the last, is over 70%. This percentage is similar to the number reported by MacKinlay and Ramaswamy (1988). However, the hypothesis that the occurrence of an arbitrage opportunity is independent from the last arbitrage opportunity cannot be rejected here, whereas MacKinlay and Ramaswamy (1988) reject it at the 1% level.

CONCLUSION

This article was concerned with the impact of arbitrage trading on the mispricing process in the presence of market frictions and early unwinding opportunities. A dynamic equilibrium model was developed, which provided eight testable hypotheses:

H1: The average mispricing is negative.

H2: The absolute level of the mispricing increases with time to maturity.

H3: The mean reversion is around a negative mean increasing with time to maturity.

H4: Mean reversion will be greater for positive than for negative mispricing.

H5: Mean reversion exists only when there is arbitrage trading.

H6: The mean reversion decreases with time to maturity.

H7: Futures trading volume decreases with time to maturity.

H8: The mispricing is path dependent.

With the exception of H7, all hypotheses were supported by the empirical findings. The results suggest that short selling restrictions and early unwinding strategies of arbitrage traders have a strong impact on the mispricing behavior. Taking them into account provides a better understanding of the relation between spot and futures prices over time and helps explain several empirical findings reported in the literature.

BIBLIOGRAPHY

Bamberg, G., and Roeder, K. (1995): "The Intraday ex ante Profitability of DAX Futures Arbitrage for Institutional Investors in Germany: The Case of Early and Late Transactions," *Finanzmarkt und Portfolio Management,* 8:50–62.

Berglund, T., and Kabir, R. (1995): "What Explains the Difference Between the Future's Price and Its 'Fair' Value? Evidence from the European Options Exchange," working paper, Tilburg University.

Bhatt, S., and Cakici, N. (1990): "Premiums on Stock Index Futures—Some Evidence," *The Journal of Futures Markets,* 10:367–375.

Board, J. L. G., and Sutcliffe, C. S. M. (1990): "Information, Volatility, Volume, and Maturity: An Investigation of Stock Index Futures," *Review of Futures Markets,* 9:532–549.

Brennan, M. J., and Schwartz, E. S. (1990): "Arbitrage in Stock Index Futures," *Journal of Business,* 63:S7–S31.

Buehler, W., and Kempf, A. (1995): "DAX Index Futures: Mispricing and Arbitrage in German Markets," *The Journal of Futures Markets,* 15:833–859.

Chan, K. (1992): "A Further Analysis of the Lead-Lag Relationship between the Cash Market and Stock Index Futures Market," *Review of Financial Studies,* 5:123–152.

Chung, P. Y. (1991): "A Transaction Data Test of Stock Index Futures Market Efficiency and Index Arbitrage Profitability." *The Journal of Finance,* 46:1791–1809.

Cooper, I., and Mello, A. (1990): "Futures/Cash Arbitrage with Early Unwinding Opportunities and Inelastic Liquidity Demand," paper presented at ESF Network Workshop on Options and Futures, October 1990, Spain.

Cornell, B., and French, K. R. (1983): "Taxes and the Pricing of Stock Index Futures," *The Journal of Finance,* 38:675–694.

Duffie, D. (1990): "The Risk-Neutral Value of the Early Arbitrage Option: A Note," *Advances in Futures and Options Research,* 4:107–110.

Durlauf, S. N., and Phillips, P. C. B. (1988): "Trends versus Random Walks in Time Series Analysis," *Econometrica,* 56:1333–1354.

Fuller, W. A. (1976): *Introduction to Statistical Time Series.* New York, NY: John Wiley & Sons.

Garbade, K. D., and Silber, W. L. (1983): "Price Movements and Price Discovery in Futures and Cash Markets," *The Review of Economics and Statistics,* 65:289–297.

Gruenbichler, A., and Callahan, T. W. (1994): "Stock Index Futures Arbitrage in Germany: The Behavior of the DAX Index Futures Price," *Review of Futures Markets,* 10:180–203.

Hall, A. (1994): "Testing for a Unit Root in Time Series with Pre-Test Data-Based Model Selection," *Journal of Business & Economic Statistics,* 12:461–470.

Hietala, P., Jokivoulle, E., and Koskinen, Y. (1995): "Insider Trading, Short Sales Constraints, and Futures Pricing," working paper, INSEAD.

Holden, C. W. (1990): "Intertemporal Arbitrage Trading: Theory and Empirical Tests," discussion paper, Indiana University.

Holden, C. W. (1995): "Index Arbitrage As Cross-Sectional Market Making," *The Journal of Futures Markets,* 15:423–455.

Kawaller, I. G., Koch, P. D., and Koch, T. W. (1987): "The Temporal Price Relationship between S&P 500 Futures and the S&P 500 Index," *The Journal of Finance,* 42:1309–1329.

Kumar, P., and Seppi, D. J. (1994): "Information and Index Arbitrage," *Journal of Business,* 67:481–547.

Kyle, A. S. (1985): "Continuous Auctions and Insider Trading," *Econometrica,* 53:1315–1335.

MacKinlay, C. A., and Ramaswamy, K. (1988): "Index-Futures Arbitrage and the Behavior of Stock Index Futures Prices," *The Review of Financial Studies,* 1:137–158.

Merrick, J. J. Jr. (1989): "Early Unwindings and Rollovers of Stock Index Futures Arbitrage Programs: Analysis and Implications for Predicting Expiration Day Effects," *The Journal of Futures Markets,* 9:101–111.

Miller, M. H., Muthuswamy, J., and Whaley, R. E. (1994): "Mean-Reversion of Standard & Poor's 500 Index Basis Changes: Arbitrage-induced or Statistical Illusion?" *The Journal of Finance,* 49:479–513.

Neal, R. (1995): "Direct Tests of Index Arbitrage Models," working paper, Federal Reserve Bank of Kansas City.

Newey, W. K., and West, K. D. (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica,* 55:703–708.

Pope, P. F., and Yadav, P. K. (1994): "The Impact of Short Sales Constraints on Stock Index Futures Prices: Evidence from FT-SE 100 Futures," *The Journal of Derivatives,* 1:15–26.

Puttonen, V. (1993): "Stock Index Futures Arbitrage in Finland: Theory and Evidence in a New Market," *European Journal of Operational Research,* 68:304–317.

Puttonen, V., and Martikainen, T. (1991): "Short Sale Restrictions: Implications for Stock Index Arbitrage," *Economics Letters,* 37:159–163.

Sofianos, G. (1993): "Index Arbitrage Profitability," *The Journal of Derivatives,* 1:6–20.

Stoll, H. R., and Whaley, R. E. (1990): "The Dynamics of Stock Index and Stock Index Futures Returns," *Journal of Financial and Quantitative Analysis,* 25:441–468.

Strickland, C., and Xu, X. (1993): "Behaviour of the FTSE 100 Basis," *Review of Futures Markets,* 12:459–502.

Stulz, R. M., Wasserfallen, W., and Stucki, T. (1990): "Stock Index Futures in Switzerland: Pricing and Hedging Performance," *Review of Futures Markets,* 9:577–592.

Sutcliffe, C. M. S. (1997): *Stock Index Futures* (2nd ed). London: International Thomson Business Press.

Tuckman, B., and Vila, J. L. (1992): "Arbitrage With Holding Costs: A Utility-Based Approach," *The Journal of Finance,* 47:1283–1302.

Tuckman, B., and Vila, J. L. (1993): "Holding Costs and Equilibrium Arbitrage," discussion paper, New York University.

Yadav, P. K., and Pope, P. F. (1990): "Stock Index Futures Arbitrage: International Evidence," *The Journal of Futures Markets,* 10:573–603.

Yadav, P. K., and Pope, P. F. (1992): "Mean Reversion in Stock Index Futures Mispricing: Evidence from the US and the UK," working paper, University of Strathclyde.

Yadav, P. K., and Pope, P. F. (1994): "Stock Index Futures Mispricing: Profit Opportunities or Risk Premia?" *Journal of Banking & Finance,* 18:921–953.

Yadav, P. K., Pope, P. F., Paudyal, K. (1994): "Threshold Autoregressive Modeling in Finance: The Price Differences of Equivalent Assets," *Mathematical Finance,* 4:205–221.