

**Cardiff Economics
Working Papers**

Laurence Copeland

*Arbitrage Bounds and the Time Series Properties of the
Discount on UK Closed-End Mutual Funds*

E2006/11

Cardiff Business School
Cardiff University
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
www.cardiff.ac.uk/carbs

ISSN 1749-6101
February 2006

Arbitrage Bounds and the Time Series Properties of the Discount on UK Closed-End Mutual Funds

by

Laurence Copeland*

Cardiff Business School

*I am grateful for comments on previous drafts of this paper to Gordon Gemmill, to participants in seminars at Cardiff, Manchester, LaTrobe and Melbourne Universities, the 9th International Conference on Macroeconomic Analysis and International Finance at the University of Crete, and to a referee of this journal. Any remaining errors are the author's responsibility.

Abstract

In a dataset of weekly observations over the period since 1990, the discount on UK closed-end mutual funds is shown to be nonstationary, but reverting to a nonzero long run mean. Although the long run discount could be explained by factors like management expenses etc., its short run fluctuations are harder to reconcile with an arbitrage-free equilibrium. In time series terms, there is evidence of long memory in discounts consistent with a bounded random walk. This conclusion is supported by explicit nonlinearity tests, and by results which suggest the behaviour of the discount is perhaps best represented by one of the class of Smooth-Transition Autoregressive (STAR) models.

Key Words and Phrases: Mutual Funds, ESTAR

1 Introduction

On the whole, anomalies in financial markets tend to melt like snowflakes almost as soon as they are examined closely. The phenomenon of the discount on closed-end mutual funds is different in this respect, as well as a number of others. After decades of published academic work on the UK and US markets, discounts show little sign of shrinking, and in fact at the time of writing this paper, still average more than 11% in the UK. This is true in spite of the fact that no entirely rational explanation for the mispricing has ever been found, while all the obvious explanations have been eliminated. Among the factors which have been adduced as possible explanations have been management expenses and holdings of illiquid or unquoted securities (Ingersoll (1976), Malkiel (1977) for USA, Draper and Paudyal (1991) for UK), tax liabilities on unrealised capital gains² (Malkiel (1977), (1995)) and, as far as country funds are concerned, asymmetric information (Frankel and Schmukler (2000)). Most of these would-be explanations have since been discredited, or at least shown to be no longer relevant.³ However, in an important recent contribution to this debate, Gemmill and Thomas (2002) (henceforth G-T) rationalise the long run level of the discount in terms of the cost of arbitrage, though even they are unable to explain its short run deviations without relying on the idea of noise traders motivated by market sentiment,

²Note that in the UK no tax is levied on the capital gains made by closed-end mutual funds.

³For a survey of the literature, see Dimson and Minio-Kozerski (1999).

thus confirming the views originally put forward by DeLong, Schleifer, Summers and Waldmann (1990) and Lee, Schleifer and Thaler (1991).⁴

Rather than seeking directly to explain the apparent anomaly in pricing behaviour, this paper approaches the problem from a different angle, instead revisiting the problem of modelling the underlying time series processes for evidence consistent or inconsistent with the explanations offered in the literature. Thus, starting from an acceptance that the long run discount must be nonzero in the overwhelming majority of cases, as is obvious from a cursory examination of the data as well as from the published literature, we look for statistical evidence of mean-reversion generated by arbitrage subject to bounds determined by transaction costs. In particular, this paper takes account of the fact that the costs of trading may well vary across different classes of investor. If this is the case, the further the discount wanders from its steady state level, the greater the potential volume of arbitrage trade, as more and more investors are offered capital gains in excess of their costs. In econometric terms, the implication is that mean reversion will be stronger the further the discount strays from its long run level, a mechanism modelled explicitly by Smooth Transition AutoRegressive processes.

The data analysed relate to UK closed-end funds (Investment Trust companies

⁴This paper is exclusively concerned with the discount in the secondary market. No attention is given to the well-known anomaly in the pricing of initial public offerings of investment trust companies (e.g. Levis and Thomas (1995)).

or ITC's), which make up a larger and more diverse market sector than their equivalent in the USA,⁵ the subject of most of the published research. Previous time series analysis of this market established that price and net asset value (NAV) were cointegrated, but that the cointegrating vector was probably not (1, -1) and that the discount tended to follow a distinct characteristic pattern over the life of a fund (Chen, Copeland, O'Hanlon (1994)).⁶ This paper first confirms that, while price and NAV are both $I(1)$ variables, the discount is in most cases nonstationary, at least once we allow for the possibility of noninteger orders of integration. In fact, estimation of ARFIMA models suggests that for most ITC's the discount tends to be nonstationary, but mean-reverting around a level of about 15%. As a proximate explanation of this behaviour, it is shown by simulation that these results are consistent with a random walk within the bounds estimated by G-T by direct methods i.e by computing actual costs of arbitrage. Finally, evidence is produced to demonstrate that ITC discounts adjust to shocks in a nonlinear fashion, and estimates are computed for the smooth-transition autoregressive (STAR) class of models which have been shown to provide an adequate representation of the time series characteristics of a number of other financial variables, including exchange rates and index futures. As far as ITC discounts are concerned, it is found that in over half the cases the symmetric ES-

⁵The value of the total assets of the UK ITC sector is about £40bn. G-T cite figures to show that in the 1990's there were more than 3 times as many closed-end funds in the UK as in the USA

⁶See also Bleaney (2004).

TAR model provides a better description of discount behaviour than the asymmetric LSTAR .

The first section of the paper introduces the dataset and its properties. The following section considers the stationarity properties of the time series for price, NAV and the discount. Estimates of the fractional root d are discussed in Section 3, and an interpretation is offered in the next section. The paper ends in Section 5 with estimates of nonlinear adjustment processes consistent with the results presented in preceding sections.

Throughout the paper, the variables of interest will be the logs of price and net asset value, $p = \ln(P)$ and $v = \ln(V)$, so that the premium is defined as $q = \log P - \log V$.⁷ The text here follows market convention and much of the published literature in dealing with the discount (i.e. the negative of the premium, $-q$), but all data in the tables is given in terms of the premium, so that smaller (more negative) numbers imply a greater discount.

⁷In the UK market, the premium is defined as $(P-V)/V$, the practice followed by G-T among others. This may be less problematic for cross-section work, but it has drawbacks in a time series context. For example, while it is possible to test the premium defined in this way for stationarity, it creates a problem interpreting tests for cointegration between P and V. Also, using the log definition, the change in the discount can be straightforwardly decomposed into the price return and net asset value return.

2 The Dataset

The dataset used in this paper is taken from Datstream, and consists of weekly Wednesday market-closing prices for closed end mutual funds, known in the London market as investment trust companies (ITC's). In most cases, the data start at 05/05/90 and end 12/05/04, a total of 735 observations, but the dataset includes a number of funds which ceased trading before the end of the period, though none provided fewer than 500 observations. Out of the figure of 300 quoted by G-T as the number of closed-end equity funds listed in London at some point in the 1990's, 133 are included in the present study.⁸ The oldest and largest funds in the industry are included, in particular the venerable Foreign and Colonial founded in 1868 and with a current NAV of over £2bn, but also a substantial proportion of smaller and/or newer funds, though none launched later than mid-1989.⁹ As far as investment portfolios are concerned, the ITC's invest overwhelmingly in ordinary shares, though some have small holdings in other assets, notably preference shares, bonds, real estate etc. Geographically, no attempt has been made to distinguish between funds investing in

⁸In fact, the sample consisted of more or less all the funds covered by Datastream and trading at the start of the period in 1990, excepting only a handful which failed to survive for the minimum 500 weeks. This compares with the G-T dataset of 158 companies observed over the years 1992-7.

⁹The median size of fund in the sample was approaching £200m in terms of NAV, with about 10% over £1bn and approximately the same proportion below £100m.

UK and those investing in other parts of the world.¹⁰

The salient facts about the dataset can be seen in Table 1, which documents clearly what is perhaps the most remarkable fact about the UK closed end mutual fund market: the sheer pervasiveness of the discount. Far from being a temporary aberration, the discount is a more or less permanent feature of the ITC market. In a sample of 133 observed over 735 weeks, the typical (i.e. median) ITC traded at a mean discount to net asset value of 13%, ranging from a maximum premium of only a little over 5% to a maximum discount of 30%. Moreover, the discount was highly variable, with a typical standard deviation over the period of 7.2%.¹¹ The asymmetry is striking, insofar as every single ITC share went to a discount at least once during the period, but more than a quarter of companies never once traded at a premium to NAV, and in fact only four out of 133 traded at a premium on average over the 15 years.¹²

¹⁰By default (i.e. unless they specifically commit to investing in a single country or region) British funds usually diversify internationally. In that respect, the industry could be said to regard the UK as simply another national market, so that ITC's restricted to investing in the UK are effectively treated as country funds. Whether or not the same country fund paradox is observed in the UK ITC sector as in the USA (see e.g. Bodurtha, Kim and Lee (1995) or Levy-Yeyati and Ubide (2000)) is a question not pursued here.

¹¹Not surprisingly, Agyei-Ampomah and Davies (2005) report finding excess volatility in a similar dataset.

¹²Even though these are weekly data, there were a number of zero returns for some of the smaller

In summary, the dataset here displays the same basic characteristics which have become familiar from the large literature on USA and UK closed-end funds published in the last twenty years, with little or no evidence that the well-established anomalies are being eliminated with the passage of time.

3 Arbitrage Bounds and the Long Run Discount

This persistence is remarkable, given that the largest ITC's are extremely liquid, and that the situation is well known to researchers in both the academic and practitioner communities. In order to carry the analysis further, the approach taken here involves breaking down the phenomenon into two components: the long run equilibrium level of the discount, and the short run fluctuations around that level.

A number of papers address aspects of the market situation which could justify the existence of a long run discount. One of the more convincing arguments in this regard is that the computation of NAV overstates the present value on which shareholders have a claim, insofar as there is a continual leakage of value into costs, whether in the form of management charges or in more general agency costs. On the other hand, this is to some extent offset by the fact that ITC's typically pay dividends, albeit in most cases at a lower level than the average dividend yield on stocks in general. If we make the assumption that both of these items are a constant proportion of NAV

stocks, probably reflecting thin trade and consequently stale prices.

(which is not completely unrealistic in this context), then it is straightforward to show (see Ross (2002)) that the equilibrium price implies a long run discount given by the ratio:¹³

$$-\bar{q} \cong \frac{\mu}{\mu + y_P} \quad (1)$$

where μ is the management/agency cost as a proportion of NAV, and y_P is the dividend yield paid by the ITC. In practice, most of the components of the long run equilibrium given in (1) are unobservable. For example, although explicit management fees are usually fixed in advance, other cost factors (e.g. nominee account fees) are more variable, both over time and possibly across investors. More importantly, the expected dividend yield is uncertain. Nonetheless, Ross (2002) and G-T offer back-of-the-envelope estimates of this equilibrium discount for the USA and UK respectively, arriving in both cases at a figure in the 10% to 15% range, which is broadly consistent with their datasets and with the results given in Table 1, as well as with the results of formal time series analysis, as will be shown in the next two sections.

If arbitrage were costless, arbitrageurs might have been expected to prevent any persistent deviation from this long run equilibrium. However, as G-T make clear, there are nonnegligible costs to the types of transactions required to exploit mispricing

¹³The approximation is due to the fact that the Ross (2002) presentation has been translated into logs, for consistency with the definition of the premium used in the rest of this paper.

in this market (see also Pontiff (1996)). For example, in the most common scenario, the underpricing could be exploited in a number of ways, most simply by buying the underpriced ITC stock while simultaneously shorting the underlying portfolio. This transaction undertaken at time θ would yield an expected profit of:

$$E(\pi_T) \cong -q_0 + T[(r_L - r_B) - (y_V - y_P)] \quad (2)$$

where T is the time period over which the arbitrage position has to be maintained, r_L and r_B are lending and borrowing rates respectively, and y_V , y_P are dividend yields on the underlying portfolio and the ITC stock respectively.¹⁴ The term in square brackets will almost invariably be negative, since borrowing rates are usually greater than lending rates,¹⁵ and the dividend yield on ITC stock is in most cases lower than on the market as a whole. It follows that, when an ITC is underpriced, the elimination of arbitrage opportunities requires that the expected change in discount be greater in absolute terms than the net cost of carrying the position for T periods.

There are a number of obstacles to implementing this arbitrage strategy, however. In the first place, it may not always be possible – and may be extremely costly – to

¹⁴This is more or less the same equation as in G-T, translated here into continuous time to be consistent with the definition of the premium in terms of logs.

¹⁵This actually understates the case, since the *effective* lending rate is likely to be further reduced in practice, because the proceeds of short sales are not normally made fully available to the seller.

take a short position in the ITC portfolio, though of course it is likely to be easier the more closely the portfolio mirrors one of the market indices, especially if a liquid futures contract is available to provide cheap replication. In general, replication will be cheaper the closer the beta of the ITC portfolio to one.

Perhaps the most obvious and in practice most important problem is that the length of time, T , over which the arbitrage position will need to be held is *ex ante* unknown. In theory, a perfect capital market will only price this uncertainty insofar as it covaries with the market return, a possibility which will be discussed later. In reality, with default risk and de facto limits on lines of credit often imposed on arbitrageurs by their own institutions, it is likely to be a serious obstacle. This is especially so as the arbitrageur entering the market at time 0 cannot be sure that the discount will not actually widen at any point in time t , $0 < t < T$.¹⁶

4 Stationarity

¹⁶Note the comparison with a (rational) bubble, in which the price is above its equilibrium level, as defined by the fundamentals, but short-selling is never profitable until the ultimate collapse, because the price carries on rising at an exponential rate sufficient to compensate speculators for the risk that the bubble may burst during their holding period. Of course, we cannot invoke the same mechanism here, since ITC prices are in most cases below their equilibrium levels, and a rational bubble cannot be negative, because the price itself cannot fall below zero. (Diba and Grossman (1988)).

Given that discounts are apparently nonzero, both in the short and long run, one might at least expect price and NAV to be cointegrated, so that the discount adjusts over time to the unknown (and unobservable) equilibrium value given in (1). In other words, it might reasonably be argued that price and net asset value can hardly diverge without limit. Shocks to the discount ought to be reversed eventually, so that, invoking the Granger-Engle Theorem, the price would adjust to the previous period's discount along the path implied by the error correction mechanism towards its long run level.

On the face of it, this approach looks unpromising in the present context, given the asymmetry between discounts and premia observed in this dataset. In fact, this pessimism is clearly justified, as can be seen from Table 2, which summarises the results of standard Kwiatkowski et al (1992) tests of the stationarity null, alongside Phillips-Perron (1988) tests of the unit root null hypothesis. From the results presented here, the conclusions with regard to price and NAV are unambiguous. Plainly, as anticipated, they are both $I(1)$, since we can decisively reject stationarity of $p = \log(P)$, $v = \log(NAV)$ in levels, whereas we have no reason to reject the null that Δp and Δv are $I(0)$.

In contrast, the results for the premium defined as $q = (p - v)$ are ambiguous. On the one hand, the value of the KPSS statistic overwhelmingly suggests rejection of the $I(0)$ null in almost every case. On the other hand, the Phillips-Perron test

results indicate rejection of $I(1)$ in the majority of cases. In fact, the median P-P statistic of -3.5 falls just below the 1% rejection level for the $I(1)$ null, and for most ITC's we reject a unit root at the 5% level.

On the face of it, these results appear to be consistent with those reported by Cheng, Copeland and O'Hanlon (1994), who looked for evidence of cointegration in a broadly similar population of UK ITC's, albeit over an earlier period, and reached the conclusion that the long run discount was plainly nonzero. However, for two related reasons, the option of looking for cointegration between unit root variables is not one pursued here. First, the standard approach of looking for cointegration between price and NAV (as in Cheng, Copeland and O'Hanlon (1994)) is likely to yield an estimate of the long run relationship involving not only a nonzero intercept but also a slope coefficient significantly different from unity. While the former result is easy to interpret (e.g. in terms of equation (1)), and in fact is entirely consistent with the other results reported in this paper (and in G-T), the latter is far more difficult to understand. Even allowing for a nonzero steady-state discount, it is hard to see why ITC stock prices should fail to respond one-for-one to changes in NAV in the long run.¹⁷ Secondly, as already pointed out, the balance of the evidence suggests

¹⁷Cheng, Copeland and O'Hanlon (1994), working with logs of price and NAV, interpreted the nonhomogeneity as reflecting a life-cycle pattern in the discount, increasing at first in the early years following the flotation of a fund and subsequently falling back as it matured.

the discount is probably not an $I(0)$ process in most cases, so that even if p and v are cointegrated, the cointegrating vector is certainly not $(1, -1)$.

5 Long Memory Tests

If we were to restrict ourselves to the two polar possibilities that discounts are either $I(0)$ or $I(1)$, we would have to conclude that, not only are long run discounts nonzero, they also show no tendency to settle at any particular level, and shocks to discounts tend to persist indefinitely, with no tendency to reversal as time passes. These are not, however, the only two possibilities. One possible interpretation of the results in Table 2 is that discounts are actually long memory processes, in the fractional integration sense of Granger and Joyeux (1980) i.e. that discounts contain a component for which the degree of integration is neither 0 nor 1, so that they can be described as $I(d)$ processes for which $0 < d < 1$. Following this line of approach, the obvious next step would be to derive maximum likelihood estimates of the univariate ARFIMA model:

$$\phi(L)(1 - L)^d q_t = a + \varphi(L)u_t \quad (3)$$

where $u_t \sim D(0, \sigma^2)$ is an error term which is assumed in most cases to follow a Student's t distribution, in order to allow for the fact that the data exhibit far fatter tails than is consistent with Gaussian normality, $\phi(L), \varphi(L)$ are polynomials in the

lag operator, and the fractional difference operator is defined by:

$$(1 - L)^d = \sum_{j=0}^{\infty} b_j L^j \quad (4)$$

where the b_i are given by:

$$b_0 = 1 \quad \text{and for } j \geq 1 \quad b_j = \frac{-d\Gamma(j-d)}{\Gamma(1-d)\Gamma(j+1)} = \frac{j-d-1}{j} b_{j-1} \quad (5)$$

where $\Gamma(x)$ is the gamma function of x .

It should be noted that, in this context, $0 < d < 0.5$ implies stationarity, while $0.5 < d < 1$ indicates nonstationarity, but mean reversion. The choice of model was based on the Akaike Information Criterion, subject to the overriding requirement that the residuals be nonautocorrelated, as evidenced by the Box-Pierce statistic.¹⁸

The results of fitting the ARFIMA model to the ITC's in the dataset can be summarised as follows:¹⁹

1. As anticipated, there was strong evidence that $d > 0$ in all but 5 cases. Even if one took the failure of the numerical estimation process to converge as evidence against long memory, we would still have had to conclude that $d = 0$ could be rejected (at the 5% level) in at least 125 cases out of 134.

¹⁸In a few cases it proved impossible to find an ARFIMA model (at least, of order (4,d,4) or lower) sufficient to capture all of the autocorrelation in the residuals. There were also four cases of nonconvergence. All models were estimated using James Davidson's TSM Version 4.05 software.

¹⁹Detailed results are available from the author.

2. In about 20% of cases, it was impossible to reject a unit root i.e. $d = 1$.
3. The median value of $d = 0.75$ indicated nonstationarity, but also mean reversion.²⁰
4. There were only 8 cases in which the estimated value of $d < 0.5$ was consistent with stationarity.
5. For the typical ITC, the intercept estimate was -0.15, implying a long run discount of 15%. At the same time, in about 25% of cases, the hypothesis that the true value of the intercept is zero could not be rejected.

In general, the interpretation of these results is not straightforward. The implication is that discounts follow a long memory process, which is nonstationary but mean-reverting, so that for the typical ITC, a shock has a half-life of at the very least 2 months, and in most cases far longer.²¹ Ultimately, disequilibrium is eliminated as the discount returns to its long run level, which is most often not zero, however, as would be implied by a naïve market efficiency view.

In fact, the median 15% long run discount emerging from these ARFIMA esti-

²⁰Davidson and Sibbertsen (2005) examined the long memory properties of 5 ITC's out of the current dataset, using a new test involving log-periodogram regressions. Their conclusions were broadly supportive of the "nonstationary but mean reverting" characterisation of the time series processes, as originally stated in the first draft of this paper.

²¹It is impossible to be more specific since, although estimated values of d are clustered in the 0.5 to 0.8 region, there is wide variation in the ARMA components for different ITC's.

mates is quite close to the G-T estimate of 12.0%, as the required compensation for cost loading on ITC portfolios along the lines set out in equation (1). Broadly, the results given here support their view that we actually have two puzzles in need of an explanation. First, why is there a nonzero long run discount? Second, why do disturbances to the discount persist so long? Having rationalised a nonzero discount with reference to expense ratios etc, how can one further explain a readjustment process in the aftermath of a shock as protracted as is implied by the results given here?

The long memory results are only partly consistent with the scenario of an arbitrage cost-determined bound within which the ITC price may fluctuate in response to random shocks caused by day-to-day imbalances between purchases and sales. Instead, they appear on the face of it to suggest that prices are certainly not confined to these arbitrage bounds, and that when random shocks occur, the reversal takes many weeks and in some (unit root) cases is never expected to happen.

However, the key to a reconciliation may lie in noting two points. First, by considering additional factors over and above the arbitrage costs taken into account in equation (2), G-T ultimately derive a sequence of ever wider channels within which they claim that discounts are confined in the long run. The time series evidence reported here can be reconciled with their analysis in the context of a model of a bounded random walk, the long memory implications of which can be demonstrated by a simple simulation exercise.

In a possible time series representation of the G-T scenario, let us take net asset value as exogenously fixed at any moment, and suppose that, given this level of NAV, the maximum and minimum possible price of the ITC share is determined by factors which are only partly observable: arbitrage costs, management fees, open-ending costs etc. Now if, within these limits, the price (and hence the discount) fluctuates randomly under the impact of shocks to net demand, the outcome could be a bounded random walk process:

$$\begin{aligned}
 q_t &= q_{t-1} + \varepsilon_t && \text{if } \underline{q} < q_{t-1} < \bar{q} \\
 q_t &= q_{t-1} && \text{otherwise}
 \end{aligned}
 \tag{6}$$

where \underline{q}, \bar{q} are lower and upper bounds on the discount respectively, and ε_t is a zero-mean IID series. This process is a random walk until it hits either the upper or lower bound.

As already noted, if there are bounds in the present case, they are not directly observable. What then would be the outcome of taking a standard time series approach to modelling this series i.e. of ignoring the bounds and treating it as an unbounded process?

Figure 1 shows the results of an experiment involving fitting a fractional difference process to a series generated by the random walk (6), subject to successively wider bounds. The shocks ε_t are 1000 drawings from the standard normal distribution at

each bound, where the bounds are measured in terms of standard deviations on the horizontal axis. Experiments with non-Gaussian shocks generated results that were qualitatively similar, and, in the relevant region, extremely close numerically to those given here.

As can be seen, the resulting estimates are biased downward along a smooth curve, so that the narrower the bound, the lower the estimated value. So, for example, a random walk between bounds of one standard deviation in either direction would mimic an autoregressive process with a coefficient of 0.61 or a fractional difference process with root $d = 0.49$. Viewed in this light, the median estimate of $d = 0.75$ quoted earlier is consistent with bounds set at approximately two and a quarter standard deviations either side of the mean, or just under $\pm 18\%$ in the median case.²² Interestingly, this is quite close to the 35% range of variation in the discount which was estimated by G-T by the totally different approach of aggregating the different types of cost associated with arbitrage operations in this market.

6 The Adjustment Mechanism

The evidence given in the previous section is obviously not the end of the story. It is merely suggestive of the possibility that the underlying ITC price process may

²²Allowing for asymmetric bounds had minimal effect, generating a central estimate of $d = 0.78$ at the level quoted in G-T (i.e. $+5\%$ to -30% , or $+0.7$ to -4.3 standard errors).

be characterised by some form of nonlinearity generated by arbitrage limits or, conceivably, by other factors as yet unknown. The nature of the nonlinearity is not immediately obvious, but a hint can be found in G-T's characterisation of ITC pricing. In their words, the option of open-ending limits the discount, so that "*like a spring under tension, the further it is pushed, the more strongly it recoils.*" (p. 2575).

While this picture could be consistent with a number of nonlinear adjustment models, the most likely candidate seems to be the class of Smooth Threshold Autoregression (STAR) models, which makes its appearance in the economics literature in two forms, Exponential (ESTAR) and Logistic (LSTAR).

There are a number of reasons to entertain this class of models. First, the ESTAR model has been shown to provide an adequate representation of the adjustment process for several financial variables, notably the real exchange rate (e.g. Michael, Nobay and Peel (1997), Taylor, Peel and Sarno (2001)) and the basis in the index futures market (Taylor (2003)). Secondly, in the present context it seems highly likely that the bounds analysed in detail by G-T are binding on investors at different levels. For example, even in the absence of capital market imperfections, the borrowing and lending rates appearing in equation (2) are unlikely to be the same across agents with heterogeneous credit ratings, access to capital and information. Moreover, unless expectations are completely homogeneous, estimates of the prospective dividend yields are likely to vary across investors, as also are anticipations with respect to

the holding periods required, T . This is particularly true in cases where the payoff from arbitraging an underpriced ITC stock is dependent on open-ending, the timing of which is likely to be especially hard to predict.²³ Thirdly, given that arbitrage involves the uncertainties in equations (1) and (2), among others, the return to arbitrage will involve a risk premium, unless none of these factors covaries with the market return, which seems improbable. For example, if, as seems highly likely, most potential arbitrageurs have a finite maximum holding period, whether as a result of credit market constraints or other factors, they will require compensation for the risk of having to liquidate early. The risk associated with premature liquidation is clearly related to the market as a whole, first, because the higher the market, other things being equal, the higher the ITC share price, and secondly because it is well known that open-ending is more common in periods when the market is buoyant.

For all these reasons, we would expect that, the further the price from net asset value (i.e. the greater the premium or discount to fair value), the higher the proportion of investors who would view the reward to arbitrage as great enough to cover the expense and associated risk. The argument can be seen as a generalization of the well-known work by Yadav, Pope and Paudyal (1994), (1999) on the pricing of index

²³One major reason for expecting costs to vary across arbitrageurs is that there are plainly cost advantages available to the ITC management itself. Even if it is prevented from trading in its own stock or has no cost advantage in doing so, it can open-end far more cheaply than an outside arbitrageur, who must first gain control of the ITC before being able to proceed.

futures. Instead of a single market-wide threshold, each potential arbitrageur has a threshold determined by his/her transaction costs, borrowing and lending rates, assessment of the probable holding-period horizon and likelihood of open-ending. Then if we imagine the community of arbitrageurs arranged in ascending order of threshold size, the greater the deviation of ITC price from net asset value, the larger the number of traders for whom arbitrage is potentially profitable, and hence the greater the market pressure to reinstate equilibrium.

The ESTAR mechanism captures this effect in what amounts to a three-regime setting. In terms of the ITC premium, we postulate the following:

$$q_t = \alpha' \underline{x}_t + \theta' \underline{x}_t \left[1 - e^{-\gamma(q_{t-d} - c)^2} \right] + u_t \quad (7)$$

where \underline{x}_t is a vector of exogenous and/or predetermined variables, usually including a constant, α and θ are parameter vectors, and the critical adjustment function is in square brackets. Stability requires that $\gamma \geq 0$, with a zero value implying linearity.

At one extreme (the “outer regime”), the adjustment function has a maximum value of one, when $(q_{t-d} - c) \rightarrow \pm\infty$ i.e. when the discount d periods back is a long way above or below its long run equilibrium level, c (which may possibly be the equilibrium described in (1)). The delay d could in principle take any value up to the maximum order of lagged dependent variable in the \underline{x}_t vector. In practice, most papers assume $d = 1$, in which case equation (7) reduces to a simple autoregression,

possibly augmented by exogenous variables in \underline{x}_t :

$$q_t = (\alpha' + \theta')\underline{x}_t + u_t \quad (8)$$

In this limiting case, adjustment is at its most rapid, possibly instantaneous, if the RHS above reduces to a random walk i.e. if the elements of α and θ corresponding to q_{t-1} sum to unity.

At the other extreme, as $(q_{t-d} - c) \rightarrow 0$, the adjustment function tends to zero, so that the inner regime (when the discount is in the neighbourhood of c) is characterized by the alternative autoregression:

$$q_t = \alpha'\underline{x}_t + u_t \quad (9)$$

In general, stability in the middle region requires $\gamma \geq 0$, and we anticipate a value for c significantly different from zero, almost invariably negative, with a conjecture that for most ITC's the long run value will not be far from the unconditional mean given in Table 1. As in most of the published work on financial variables, \underline{x}_t is restricted to a constant²⁴ and lagged values of the dependent variable $[1, q_{t-1}, q_{t-2}, \dots]$, so that the equation actually estimated was:

$$q_t = \alpha_0 + \sum_{j=1}^m \alpha_j q_{t-j} + e^{-\gamma(q_{t-d} - \alpha_0)^2} \left[\sum_{j=1}^n \theta_j (q_{t-j} - \alpha_0) \right] + u_t \quad (10)$$

²⁴A polynomial time trend might be justified, given the well-known tendency for discounts to vary with the age of the ITC (Copeland, O'Hanlon and Cheng (1994)) but experiments with a trend were largely unsuccessful. In any case, Paya and Peel (2003) cast doubt on the reliability of estimates of ESTAR processes with trends.

As already mentioned, this model has been extensively applied to a number of economics variables, most notably to real exchange rates. Its applicability in the present context needs to be explored, but it is not the only candidate. Whereas asymmetric adjustment in exchange rates seems a remote possibility, it certainly cannot be ruled out *a priori* in the present context, since there are a number of elements to arbitrage costs which may be different for buyers and sellers.

Consider again the two sides of the arbitrage involved here. On the one hand, when price is above NAV, the arbitrageur seeks to short sell the ITC share while simultaneously buying the underlying portfolio or a proxy, in the form of an index futures contract, an ETF²⁵ or maybe a (fairly priced) index fund. On the other hand, when, as is usually the case, the price is at a discount to NAV, the arbitrageur has to take a short position in the portfolio while buying the stock. Comparing the two situations, it is not at all obvious that transaction costs, broadly defined, will be the same. For example, if the portfolio is not a good match to the market so that the use of proxies is ruled out, the costs involved in creating a short position are unlikely to be the same as the costs of going long. Moreover, in the longer term open-ending a mutual fund to eliminate a discount will cost more than simply issuing more shares to exploit a premium.

The implication is that asymmetric adjustment is a possibility which needs at the

²⁵ETF's are securities intended to track a specific (London) market index.

very least to be entertained. The obvious alternative to ESTAR in this regard is the LSTAR model, given by the following equation:

$$q_t = \alpha_0 + \sum_{j=1}^m \alpha_j q_{t-j} + [1 + e^{-\gamma(q_{t-d} - \alpha_0)}]^{-1} \left[\sum_{j=1}^n \theta_j (q_{t-j} - \alpha_0) \right] + u_t \quad (11)$$

This model differs from ESTAR in allowing for three, rather than two regimes. Instead of simply an inner and an outer regime, we now have an inner regime, and two outer regimes depending on whether the premium is above or below its long run level. Specifically, in the neighbourhood of long run equilibrium, we have:

$$q_t \approx \alpha_0 + \sum_{j=1}^m \alpha_j q_{t-j} + \frac{1}{2} \left[\sum_{j=1}^n \theta_j (q_{t-j} - \alpha_0) \right] + u_t \quad (12)$$

which is the autoregression defining the inner zone.

As far as the two outer regimes are concerned, note that we continue to assume that $\gamma \geq 0$. It follows that, in the upper regime, when $q_{t-1} - \alpha_0 \rightarrow +\infty$, we get:

$$q_t \approx \alpha_0 + \sum_{j=1}^m \alpha_j q_{t-j} + \left[\sum_{j=1}^n \theta_j (q_{t-j} - \alpha_0) \right] + u_t \quad (13)$$

On the other hand, in the lower regime when the premium is a long way below its long run level i.e. when $q_{t-1} - \alpha_0 \rightarrow -\infty$, the model reduces to:

$$q_t \rightarrow \alpha_0 + \sum_{j=1}^m \alpha_j q_{t-j} + u_t \quad (14)$$

As far as model selection is concerned, Escribano and Jorda (1999) formulate a straightforward series of LM-tests for nonlinearity in general, and subsequently to

distinguish between ESTAR and LSTAR processes, which in the present case can be based on the following equation:

$$q_t = \delta_0 + \delta_1 x_t + \lambda_1 x_t q_{t-1} + \lambda_2 x_t q_{t-1}^2 + \lambda_3 x_t q_{t-1}^3 + \lambda_4 x_t q_{t-1}^4 + u_t \quad (15)$$

\underline{x}_t is the vector of pre-determined variables appearing in linear form, which means here the lagged values of q_t in the second term on the RHS of (7) and (8). The test for nonlinearity involves the null hypothesis that all four $\lambda_i = 0 \quad i = 1, 2, 3, 4$. In the same equation, the null hypothesis of no LSTAR (ESTAR) process is accepted if we cannot reject the constraint $\lambda_1 = \lambda_3 = 0$ ($\lambda_2 = \lambda_4 = 0$). The authors of the testing procedure recommend choosing between LSTAR and ESTAR on the basis of how decisive the rejection i.e selecting the model with the lowest p -value for the test statistic on the constraint.

The outcome of applying these tests is given in the first five columns of Table 3. In implementing the tests, the order of autoregression is selected by the AIC criterion subject to a maximum of $m = 6$ and the longest delay entertained is likewise $d = 6$. There are two noteworthy features of the test results. First, for 90% of the ITC's in the sample, the tests reject linearity at the 10% level. Second, in about 60% of cases, ESTAR is preferred to LSTAR. There is no obvious explanation of why some ITC's appear to adjust symmetrically, others asymmetrically. In particular, there was no significant difference between the estimates of the long run discount in the ESTAR

and LSTAR groups.²⁶

Actually fitting ESTAR/LSTAR models involves dealing with a number of complicating factors. In the first place, discounts share some of the same characteristics as other financial variables, in particular a high degree of heteroscedasticity and non-normality in the form of fat-tailed error distributions, so that all estimates are based on t -distributed GARCH error processes. Secondly, the vagaries of numerical methods make it essential to explore the likelihood surface so as to avoid, if possible, settling on a local rather than global optimum. To that end, the estimates given in the table were generated from a sequence of starting values for the transition parameter, γ , so as to ensure as far as possible convergence to a global minimum of the least squares function.²⁷ In six cases, convergence proved impossible to achieve in any case. The remainder produced a wide spread of parameter values, though mostly high, indicating relatively rapid reaction to disequilibrium.²⁸ Unfortunately, there is

²⁶Note that, if we insisted on believing the long run discount to be zero, there would be a clear presumption in favour of asymmetry, in the face of the clear bias towards discount rather than premia. However, once we accept that the long run discount is not necessarily zero, there is no *a priori* reason to prefer an asymmetric to a symmetric formulation of the adjustment process.

²⁷The equilibrium deviation in the nonlinear component was normalised by the standard error for estimation purposes. The coefficients in the table have been adjusted to take account of the normalisation.

²⁸It is impossible to be more precise because, as is well known, the response of STAR models to shocks cannot easily be derived analytically, especially when they are augmented by a GARCH- t

no straightforward way of judging the significance or otherwise of these estimates. In fact, as van Dijk et al (2002) note:

"..the t -statistic [of γ] does not have its customary asymptotic t -distribution under the hypothesis that $\gamma = 0$[since] large changes in γ have only a minor effect on the transition function, high accuracy in estimating γ is not necessary."

Nonetheless, there are a number of noteworthy features of the results in Table 3. First, the intercept estimates are close both to the mean discount observed over the sample period and to the fitted values from the ARFIMA models. Second, the estimates of γ are spread over such a wide range that it can at least be said with confidence that adjustment speeds vary substantially across ITC's. However, closer examination revealed no obvious pattern which might explain the different transition speeds. As regards the adequacy of the fitted model, it is doubtful whether much reliance can be placed on the R^2 statistics in this case. Perhaps more relevant is the fact that the residuals of the fitted models appear nonautocorrelated for the most part, with no sign of any remaining heteroscedasticity.

error process. Instead, simulation is required in order to generate what is known as the Generalised Impulse Response Function.

7 Conclusions

This paper has examined the time series properties of the discount rate on UK ITC's, confirming the results of Cheng, Copeland and O'Hanlon (1994) and G-T that, if there is cointegration between price and net asset value, it is not a relationship implying a zero long run discount. In fact, the evidence presented here indicates a long memory discount process implying in most cases nonstationarity with mean-reversion, a result which could be the outcome of fitting a linear model to any of a number of possible nonlinear processes, for example a bounded random walk. A process of this kind would be consistent both with the anecdotal evidence and also with the detailed estimates of G-T. To make the mechanism explicit, nonlinearity tests were applied to the discount process across a sample of 133 investment trust, resulting in rejection of the null hypothesis of linearity in 90% of the dataset. Two types of smooth transition autoregressive models were estimated, the symmetric ESTAR model dominating the asymmetric LSTAR in 60% of cases.

These results open up a potentially rich research agenda. The most obvious question is whether a similar mechanism is at work in other countries, especially the USA, where there is a sizable closed-end mutual fund sector. Other issues discussed in the published literature which could be investigated in the framework set out here relate to the role played by interest rates and, more importantly, by market sentiment

in affecting ITC pricing, and vice versa.

References

Agyei-Ampomah, S and Davies, J R (2005) Excess Volatility and UK Investment Trusts, *Journal of Business Finance and Accounting*, Vol. 32, No. 5-6, pp. 1033-1062

Bleaney, M (2004) Past Returns and Investment Trust Discounts, *Journal of Business Finance and Accounting*, Vol. 31, No. 9-10, pp. 1505-1523

Bodurtha, J N Jr, Kim, D-S and Lee, C M C (1995) Closed-End Country Funds and US Market Sentiment, *Review of Financial Studies*, 8, 879-918

Cheng A, Copeland L and O'Hanlon J (1994) Investment Trust Discounts and Abnormal Returns: UK Evidence, *Journal of Business Finance and Accounting*, 21: 813-31

Davidson J E and Sibbertsen P (2005) Tests of Bias in Log-Periodgram Regression, Working Paper

DeLong B, Schleifer A, Summers, L and Waldmann (1990) Noise trader Risk in Financial Markets, *Journal of Political Economy*, 98:703-38

Diba, B T and Grossman, H I (1988) Rational Inflationary Bubbles, *Journal of Monetary Economics*, 21, 35-46

Dimson E and Marsh P (1998) The Closed-End Fund Discount and Performance Persistence, London Business School Working Paper

Dimson E and Marsh P (1999) Closed-End Funds: a Survey, *Financial Market Institutions and Instruments*, 9, 1-41

Draper P (1989) *The Investment Trust Industry in the UK: an Empirical Analysis*, Gower Press, London

Draper P and Paudyal K (1991) The Investment Trust Discount Revisited, *Journal of Business Finance and Accounting*, 18: 791-806

Escribano, A and Jorda, O (1999) Improved Testing and Specification of Smooth Transition Models, in *Nonlinear Time Series Analysis of Economic and Financial Data*, Kluwer Academic Publishers, USA

Frankel, J A and Schmukler (2000) Country Funds and Asymmetric Information, *International Journal of Finance and Economics*, 5, 177-95

Gemmill G and Thomas D C (2002) Noise Trading, Costly Arbitrage, and Asset Prices: Evidence from Closed-End Funds, *Journal of Finance*, 57(6), 2571-94

Granger, C W J (1980) Long Memory relationships and the Aggregation of Dynamic Models, *Journal of Econometrics*, 14: 227-38

Granger and Joyeux (1980) An Introduction to Long-Memory Time Series and Fractional Differencing, *Journal of Time Series Analysis*, 1, 15-29

Ingersoll (1976) A Theoretical and Empirical Investigation of the Dual Purpose Funds, *Journal of Financial Economics*, 3: 81-123

Kwiatkowski D, Phillips P C B , Schmidt P and Shin Y (1992) Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root, *Journal of Econometrics*, 54: 159-78

Lee C, Schleifer A and Thaler R (1991) Investor Sentiment and the Closed-End Fund Puzzle, *Journal of Finance*, 46: 75-109

Levis M and Thomas D (1995) Investment Trust IPO's: Issuing Behaviour and Price Performance, *Journal of Banking and Finance*, 19: 1437-58

Levy-Yeyati, E and Ubide, A (2000) Crises, Contagion, and the Closed-End Country Fund Puzzle, *IMF Staff Papers*, 47(1), 54-89

Malkiel (1977) The Valuation of Closed-End Fund Investment Company Shares, *Journal of Finance*, 32: 847-59

Malkiel (1995) The Structure of Closed-End Fund Discounts Revisited, *Journal of Portfolio Management*, 21: 32-8

Michael, P, Nobay, A R and Peel, D A (1997) Transactions Costs and Nonlinear Adjustment in Real Exchange Rates: an Empirical Investigation, *Journal of Political Economy*, 105, 862-79

Paya, I and Peel, D A (2003) Purchasing Power Parity Adjustment Speeds in High Frequency Data when the Equilibrium Real Exchange Rate is Proxied by a Deterministic Trend, *Manchester School (Supplement)*, 39-53

Phillips P C B and Perron P (1988) Testing for a Unit Root in Time Series Regression, *Biometrika*, 75: 335-46

Pontiff, J (1996) Costly Arbitrage: Evidence from Closed-End Funds, *Quarterly Journal of Economics*, 1135-51

Taylor, M P, Peel, D A and Sarno, L (2001) Nonlinear Mean-Reversion in Real Exchange Rates: Towards a Solution to the Purchasing Power Parity Puzzles, *International Economic Review*, 42(4), 1015-42

Taylor, N (2003) A New Econometric Model of Index Arbitrage, CARBS Working Paper

van Dijk, D, Terasvirta, T and Franses, P H (2002) Smooth Transition Au-

toregressive Models - a Survey of Recent Developments, *Econometric Reviews*,
21(1), 1-47

P K Yadav, P F Pope and K Paudyal (1994) Threshold Autoregressive Modelling in Finance : The Pricing of Equivalent Assets, *Mathematical Finance*,
Vol 4, No 2, pp 205-221

P K Yadav, P F Pope and K Paudyal (1999) Non-linear Dependence in Stock Returns: Does Trading Frequency Matter?, *Journal of Business Finance and Accounting*, 26: 651-79

Figure 1: SIMULATIONS OF d FROM EQUATION (2)
 1000 Simulations at Each Bound

