Empirical tests of changes in autocorrelation of stock index returns.*

Bernt A. Oedegaard

University of British Columbia

Finance Division
Faculty of Commerce
2053 Main Mall
Vancouver, BC,
Canada V6T 1Z2
Email: bernt@finance.commerce.ubc.ca

April 1994.

Abstract

Recent empirical investigations have found a decrease over time in the estimated first order autocorrelation of stock index returns. It has been suggested that this change in autocorrelation is caused by the introduction of new financial markets, such as options and futures on the index. Froot and Perold (1990).

This paper carries out formal hypothesis tests for changes in index autocorrelation, and evaluates the relationship between the estimates of autoregressive coefficients with those of other second moments. The methods used are tests for structural changes in time series GMM estimation. The small sample characteristics of the test statistics are investigated by Monte Carlo, and the tests are carried out on US stock market data for the period 1976 to 1989.

The results show that if we impose a constant conditional variance, we can reject a null of no change in autoregressive coefficients of the S&P 500 index. If we allow for a changing conditional variance, we no longer reject a null of no change. Hence, correcting for changes in the variance, we can not claim there has been a change in the autoregressive relationship. To explore the causes of these results, we also looked at cross-correlations between different–sized indices, and found that there has been a change in the lead-lag relationship between the portfolio of largest stocks and other size-based portfolios.

*This paper is based on chapter 1 of my Ph.D. dissertation at Carnegie Mellon University. I want to thank the members of my committee, Richard Green (Chairman), Robert Dammon and Stanley Zin. I have also benefited from comments from Burton Hollifield, and seminar participants at Carnegie Mellon University, the European Finance Meetings, the Norwegian School of Economics and Business Administration, Odense University, the Swedish School of Economics, University of Illinois at Chicago, University of Illinois at Urbana-Champaign and Vanderbilt University. Financial support from the Norwegian School of Economics and Business Administration is gratefully acknowledged. All errors are my own.

1 Introduction

A number of authors have shown evidence of changes in the time series properties of S&P 500 index returns at the time of introduction of derivative securities based on the index, such as options and futures. Damodaran and Subrahmanyam (1992) gives an overview of this literature. As examples, Harris (1989) shows evidence of changes in the conditional variance of index returns, and Froot and Perold (1990) show evidence of changes in both the variance and the autocorrelation of index returns. Similarly, there has been a number of investigations of introduction of options on individual stocks and their effects on properties of stock returns.[1] These studies found evidence of changes in the stock volatility.

In this paper we concentrate on changes in autoregressions. We carry out explicit statistical tests for whether changes have occurred in autoregressive relationships in the time series of index returns. Such formal inference are lacking in existing work. By using the Generalised Method of Moments (GMM), the tests allow for dependency and heteroscedasticity in the time series.

Using daily returns on the S&P 500, we show that if we assume a constant conditional variance, we reject the null of a constant autoregressive relationship. However, if we allow for changes in conditional variances, we no longer reject this null. Hence, when we correct for nonconstant conditional variance, we can not reject a null of constant autoregressive relationship in the index throughout the period. As a check of this result we looked at a test statistic that allows us to estimate the time at which an event occurred. We found that there is no evidence of any changes at the time of introduction of new instruments.

As shown by a number of authors, a major source of index autocorrelation is crosscorrelations between the stocks that constitute the index. It is therefore of interest to see if we can find any evidence of structural change in these crosscorrelation, and therefore the observed index changes. By using data on size-sorted portfolios, we find that there is strong evidence for a change in the crosscorrelations between the large sized portfolio and other size-based portfolios.

In the first section we discuss and motivate the problem. Section 3 describes the statistical tests for structural change. Section 4 shows the results from applying the tests to stock market data, and section 5 investigates the small sample properties of the test statistics using Monte Carlo. Section 6 concludes.

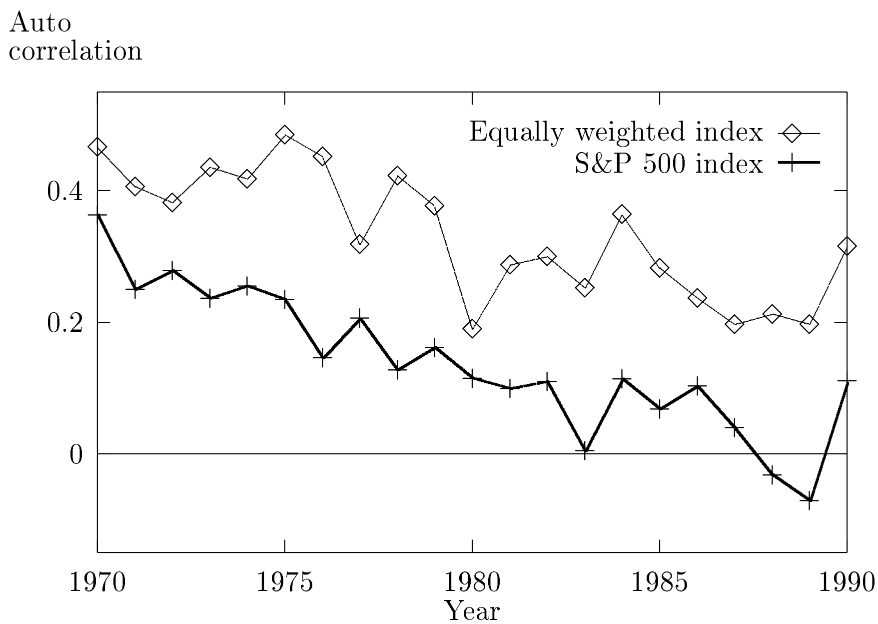
2 Motivation.

Recent empirical investigations have shown evidence of positive first order autocorrelation in returns on broad based stock indices. See for example Lo and MacKinlay (1988). Froot and Perold (1990) show that the measured index autocorrelation has declined over time. Consider figures 1 and 2. They graph yearly estimates of first-order autocorrelation of daily and weekly returns on the S&P

[1] See for example Conrad (1989) and Skinner (1989).

500 stock index and an equally weighted (EW) index of NYSE stocks. For daily returns the sample estimates of autocorrelation has declined.

Figure 1 Yearly estimates of the autocorrelation of daily index returns. 1970 to 1990. Equally weighted index and S&P500 index.

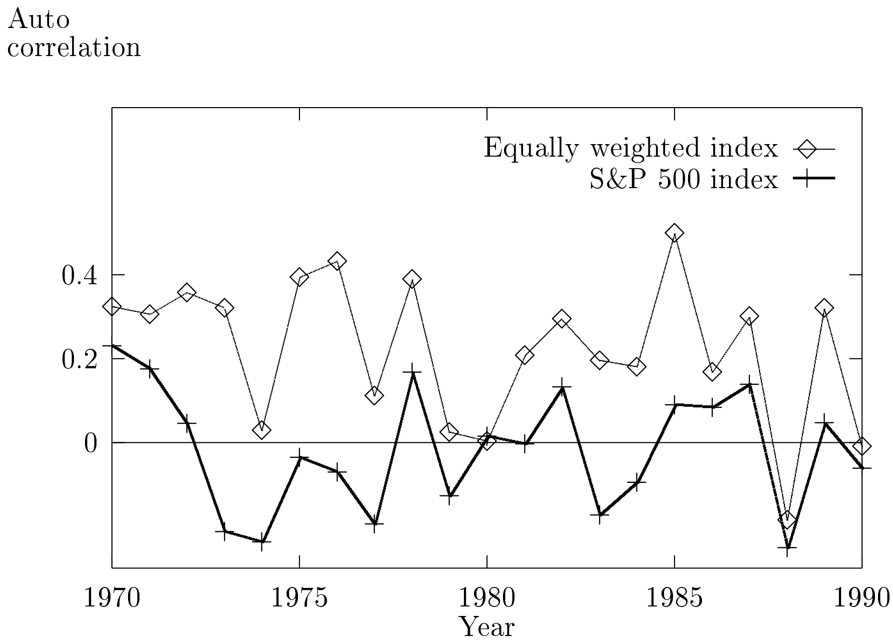


Most of the autocorrelation in index returns is attributed to cross-correlations across stocks. These cross-correlations are shown in e.g. Lo and MacKinlay (1990) and Conrad, Gultekin, and Kaul (1991) to be mainly caused by the stocks of large firms leading the stocks of smaller firms. Suppose some of this behaviour is caused by changes in marketwide 'factors,' the effects of which are reflected faster in prices of larger stocks, which are more heavily traded. If we now introduce a futures contract on a stock index consisting of a large crossection of the stocks on the market, information about marketwide factors will be reflected in prices of this contract. Since the futures price is observable by the whole market, the information about marketwide factors can be disseminated faster into prices of the stocks that constitute the index. Froot and Perold (1990) and Subrahmanyam (1991) are examples of models where introduction of new markets have such informational effects. Market participants can also "hedge" market risk at a lower cost, which may change the arbitrage relationships in the market. From this it follows that the the lead-lag relationships in the market may change, which can be reflected in the autocorrelation of index returns.

The issue of the effects of introduction of new financial instruments is still very much discussed in the popular press, in particular around big swings in stock indices, as seen latest on 15 nov 1991.[2] The question of possible effects of introduction of derivatives is therefore also one of considerable

[2] See for example *"Coming to terms with futures,"* The Economist, 23 nov 1991.

Figure 2 Yearly estimates of the autocorrelation of weekly index returns. 1970 to 1989. Equally weighted index and S&P 500 index.



public interest.

While figures 1 and 2 indicate a decline in sample estimates of autocorrelation, they do not constitute a formal test. The behaviour observed in the graphs may still be generated by a process with constant autoregressive coefficient. One way this could happen is if the variance of the process increased. The intuition for this follows from the formula for autocorrelation:

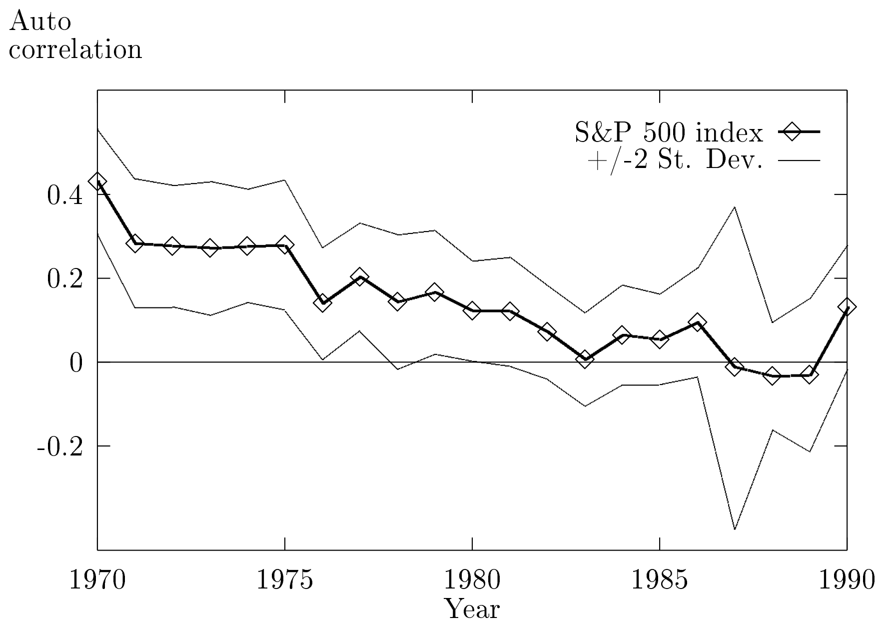
$$\rho(r, r_{-1}) = \frac{\text{cov}(r, r_{-1})}{\sigma(r)\sigma(r_{-1})}$$

The sample autocorrelation can decrease either if $\text{cov}(r, r_{-1})$ decreases or if $\sigma(r)$ increases. In this paper we ask formally whether the observed decline in sample autocorrelations allow us to conclude that the "real" autoregressive coefficient has changed. We formulate this question as a hypothesis test, with a null hypothesis of no parameter change.

To further emphasize the need for formal testing, let us point out the uncertainty in yearly estimates of autocorrelation. This is illustrated in figure 3. The figure shows yearly point estimates of autocorrelation of daily S&P 500 returns, plus and minus 2 standard deviations. As the figure shows, the confidence intervals for the autoregressive coefficient are large.

Possible changes due to the introduction of new markets do not not translate readily into statements about the time series behaviour of returns. Absent a formal model of trading, the parameters of which could be estimated and tested for constancy through time, tests measuring autoregressive

Figure 3 Yearly estimates of the autocorrelation of daily index returns. 1970 to 1990. S&P500 index. The figure also shows +/- 2 standard deviations of the point estimate.



relationships in the context of simple processes seem a reasonable first step to a more complicated test. In this context we also note that the tests carried out in this paper makes no claims about causality. Any observed changes in the time series could be caused by other exogenous events than the introduction of new markets.

3 Test Statistics.

In this section we describe the test statistics. The tests are cast in the framework of Generalised Method of Moments (GMM) estimation Hansen (1982). In two recent papers, Andrews and Fair (1988) and Ghysels and Hall (1990) propose a number of tests for regime changes in structural parameters of models estimated using GMM. In each test, the time of regime change has to be known to the econometrician. The tests involve splitting the data into two subsamples and testing for differences in parameter estimates across the two samples.

We also consider a more recently developed statistic Andrews (1990) that allows the time of parameter change to be unknown.

3.1 Estimating the time series parameters.

First consider estimation of the autoregressive parameters. For illustrative purposes, we look at the case where index returns r_t follow a linear process with constant conditional variance:

$$r_t = \mu + \alpha r_{t-1} + \varepsilon_t, \tag{1}$$

with α and μ constant, and ε_t a noise term satisfying

$$E[\varepsilon_t | I_{t-1}] = 0 \ \ \forall\, t, \tag{2}$$

$$E[\varepsilon_t^2 | I_{t-1}] = \sigma^2 \ < \ \infty \ \ \forall\, t. \tag{3}$$

Here I_{t-1} is the information set at time $t-1$. These are strong assumptions on stock index returns, in particular the assumption of constant conditional variance. It can be viewed as a first try at the problem, and we relax it in later sections of the paper, where we allow the conditional variance to change.

If we made further distributional assumptions on ε_t, the model could be estimated directly by Maximum Likelihood. For example, if the ε_t are *iid* normally distributed, this reduces to a OLS linear autoregression. The tests would then reduce to standard Chow tests. However, we choose to not make any distributional assumptions, and instead estimate the parameters directly from the moment conditions imposed by the model. This allows more flexibility in the possible distributions of ε_t. The most general assumptions where GMM is still consistent are the mixing conditions of Gallant and White (1988). In all the cases we look at we assume these hold. The cost of using GMM is that it is less efficient than ML in cases where the distribution of the error is known. Thus, use of GMM relative to ML will make the tests more conservative in the sense of less ability to reject the null hypothesis. This is caused by the larger uncertainty in the estimates. We can justify the use of GMM by empirical evidence, going back to classical work of Mandelbrot (1963) and Fama (1965), that high frequency stock returns, such as daily returns, deviate from a normal distribution. It is also known that the Chow test is nonrobust to deviations from *iid* disturbances Kramer (1989).

Another advantage of using GMM is that the test of overidentifying restrictions gives a useful specification test of the model formulation.

Let us now show how we would estimate the parameters of this process with GMM. We want to estimate μ, α and σ. The model imposes the orthogonality conditions

$$E[r_t - (\mu + \alpha r_{t-1})] = 0, \tag{4}$$

$$E[\sigma^2 - (r_t - (\mu + \alpha r_{t-1}))^2] = 0. \tag{5}$$

To generate additional moment conditions, we assume that elements Z_{t-1} in the information set

I_{t-1} at time $t-1$ are orthogonal to the error ε_t. This gives us further moment conditions:

$$E[(r_t - (\mu + \alpha r_{t-1}))Z_{t-1}] = 0, \tag{6}$$

$$E[(\sigma^2 - (r_t - (\mu + \alpha r_{t-1}))^2)Z_{t-1}] = 0. \tag{7}$$

We will consider two instruments: (1) Past index returns ($Z_{t-1} = r_{t-1}$), and (2) Past forecast errors ($Z_{t-1} = \varepsilon_{t-1}$). The GMM estimator minimises a quadratic form in the sample counterparts of equations (4) to (7) above, using a weighting matrix in the quadratic form. Hansen (1982) show that the optimal weighting matrix is an estimate of S^{-1}, where S is the variance-covariance matrix of the moment conditions.

3.2 Tests for structural change.

We now consider tests of structural change. The discussion is mainly descriptive, for the explicit definitions of the test statistics we refer to Appendix B and the original papers.

Let the GMM moment conditions be summarised by

$$E[f(x_t, \theta)] = 0,$$

where $f(x_t, \theta)$ is a vector of moment conditions, θ is the parameter vector, and $x_t \in X$ is the data.

We first look at tests where there is a structural change at a known time. We consider three tests due to Andrews and Fair (1988). The data is split into two subsamples

$$X_1 = \{x_t\}_{t=1}^{T_1}$$

$$X_2 = \{x_t\}_{t=T_1+1}^{T}$$

The possible regime change occurs between periods T_1 and $T_1 + 1$. The point T_1 at which the sample is split is assumed known to the econometrician. The matrix X_1 holds data from the period before the regime change, and X_2 the data from the period after the regime change. Similarly, the parameters θ to be estimated are split into $\theta = (\theta_1, \theta_2, \theta_3)$, where θ_3 are parameters that stay unchanged over the whole period, and θ_1 and θ_2 are the parameters of interest before and after the regime change, respectively. The moment conditions are then "stacked" the following way:

$$f(x_t, \theta) = \begin{cases} \begin{bmatrix} f(x_t, \theta_1, \theta_3) \\ 0 \end{bmatrix} & \text{for } x_t \in X_1 \\ \begin{bmatrix} 0 \\ f(x_t, \theta_2, \theta_3) \end{bmatrix} & \text{for } x_t \in X_2 \end{cases}$$

The null hypothesis of no regime change is formulated as

$$\mathcal{H}_0: \quad \theta_1 = \theta_2$$

which is to be tested against an alternative hypothesis:

$$\mathcal{H}_A: \quad \theta_1 \neq \theta_2$$

Andrews and Fair discusses three different tests; a Wald-test W, a Lagrange multiplier-like test LM and a likelihood ratio-like test LR. To calculate the Wald statistic W, the model is estimated without imposing the hypothesis $\theta_1 = \theta_2$. Given this estimate, form a quadratic form where the difference $\hat{\theta}_1 - \hat{\theta}_2$ is used. Under the null, this difference should be zero. Under alternative hypotheses the difference is nonzero. A large value of the test statistic shows the estimated difference $\hat{\theta}_1 - \hat{\theta}_2$ is large, and the null hypothesis is rejected. The 'Lagrange multiplier' statistic LM is found by estimating the model imposing the null hypothesis, and testing the fit of the resulting estimate. The 'Likelihood ratio' statistic LR estimates the model in both restricted and unrestricted form, and compares the difference in fit of the two estimates. The statistics will, under the null hypothesis, all have asymptotic chi–square distributions with degrees of freedom equal to the number of restricted parameters.

We also consider a statistic GH proposed by Ghysels and Hall (1990). Their method is to substiture the estimated parameters from one subperiod into the moment conditions of the other subperiod. If the null hypothesis holds, this will be a good fit. Under alternative hypotheses, the fit will be less good, and will make the test statistic large. Under the null hypothesis, GH has an asymptotic chi-square distribution with degrees of freedom equal to the number of orthogonality conditions in $f(\cdot)$. We consider two versions of the GH statistic. GH_1 uses only part of the data to estimate the covariance matrix, and GH_2 uses all the data. Under the null the covariance matrix is constant, so estimating it using all the data is still consistent, and will result in a better estimate of the covariance matrix.

To illustrate the difference between the Ghysels and Hall statistic and the Andrews and Fair statistics, consider a change of conditional variance in the return process discussed above. The parameter σ only appears in two of the four moment conditions, but the Ghysels and Hall statistic uses all four moment conditions in the test. The Andrews and Fair statistics will only make use of the moment conditions in which σ appears. Since the Ghysels and Hall statistic makes no assumptions about the form of the structural break, we expect it to have low power against a specified alternative, but it is also robust to different forms of structural change.

Finally, in a recent paper, Andrews (1990) discusses tests in which the time of the possible structural change is unknown to the econometrician. It is assumed that there is only one possible regime change, but there is uncertainty as to when it occurred, as well as whether it occurred at all.

Andrews develops Wald, LM and LR-like test statistics for this situation. Essentially, the statistics will take the supremum across time of the statistics for the corresponding tests with a known time of structural change. The statistics will be distributed as suprema of chi-square variables. The distribution of the statistics have no obvious closed form expression, so Andrews uses Monte-Carlo methods to find rejection levels for the statistics. In our work we use his critical values to evaluate the results. In the case where the null is rejected, the method also gives an estimate of the time of structural change.

4 Estimation Results.

In this section we look at results from applying the tests described above to stock index returns. The first subsection considers tests for two broad based stock indices. The second investigate changes in the cross–correlations of stocks.

4.1 Market indices.

We look at two market-wide indices, the equally weighted (EW) index from the CRSP data base and the S&P 500 index. The tests are done using both weekly and daily returns. Table 1 shows some descriptive statistics for the return series. Only the first-order autocorrelation is of a significant magnitude.

For the tests with a fixed time of parameter change, we have to settle on a breakpoint. We are looking for possible effects of new markets. As an indicator of when these effects occurred we have chosen volume numbers on the stock market index futures/options markets. The time at which these volumes became substantial should be a reasonable time to check for changes in the time series. These markets were introduced in 1982, and they had a large volume within a year, so we have chosen to consider the period 1976 to 1989, with a possible breakpoint at 1/1/83.

We will consider two alternative assumptions on the return process. The first is the formulation discussed in the previous section, where the conditional variance is assumed constant. In the second formulation we explicitly parameterise the conditional variance using an ARCH type process, and do the tests for structural change with this process.

To summarise, the results show that if the conditional variance is assumed constant, we reject a null of no change in the autoregressive parameter for daily S&P 500 returns, but in all cases where the conditional variance is allowed to change, we do not reject a null of no change in the autoregressive parameter.

Let us first look at estimating the process described in section 3, where we assumed a constant conditional variance. The returns are assumed to satisfy

$$\mu + \alpha r_{t-1} + \varepsilon_t$$

Table 1 Descriptive statistics for the return series. (OLS) Average returns, standard deviations, and the first through fourth order autocorrelations. Returns in percent. Equally weighted index and S&P 500 index. 1976–1989

Weekly Returns

				Auto Correlation			
Index	Period	$\bar{r} \times 100$	$\sigma(r) \times 100$	ρ_1	ρ_2	ρ_3	ρ_4
EW	1976–89	0.462	2.193	0.228	0.080	0.041	0.059
	1976–82	0.553	2.302	0.205	0.152	0.071	0.094
	1983–89	0.371	2.073	0.249	0.001	0.002	0.011
S&P500	1976–89	0.217	2.121	0.041	0.001	-0.056	-0.005
	1976–82	0.144	2.094	0.046	0.001	-0.053	0.029
	1983–89	0.291	2.145	0.027	0.006	-0.064	-0.043

Daily Returns

				Auto Correlation			
Index	Period	$\bar{r} \times 100$	$\sigma(r) \times 100$	ρ_1	ρ_2	ρ_3	ρ_4
EW	1976–89	0.092	0.790	0.274	0.053	0.022	0.052
	1976–82	0.110	0.764	0.326	0.098	0.079	0.049
	1983–89	0.074	0.815	0.227	0.011	-0.029	0.052
S&P500	1976–89	0.044	0.985	0.067	-0.037	-0.027	-0.046
	1976–82	0.029	0.848	0.129	0.024	0.013	-0.027
	1983–89	0.059	1.106	0.030	-0.073	-0.050	-0.059

with

$$E[\varepsilon_t | I_{t-1}] = 0 \ \ \forall \ t$$
$$E[\varepsilon_t^2 | I_{t-1}] = \sigma^2$$

We consider five tests. In the first three, the null hypothesis is that all three parameters μ, α and σ stay constant throughout the period. This null is tested against the following three alternative hypotheses:

1. The autoregressive parameter α changes, the other parameters μ and σ^2 stay constant.

2. Both the autoregressive parameter α and the mean μ changes, σ^2 stays constant.

3. All parameters μ, α and σ^2 change.

For the first three tests, the null involves a constant conditional variance over the whole period. For the final two tests we allow for changes in some of the parameters while we test for change in the autoregressive coefficient. The fourth test lets the null allow for a one-time change in the conditional variance. We are testing whether we can detect a change in the autoregressive coefficient when we let the conditional variance change once. In the fifth test we allow for a change in the mean parameter μ.

4. $\mathcal{H}_0$: μ and α are constant, σ is allowed to change once.
   $\mathcal{H}_A$: μ is constant, α and σ change.

5. $\mathcal{H}_0$: μ is allowed to change once, α and σ are constant.
   $\mathcal{H}_A$: μ and α change, σ is constant,

Table 2 below lists the results using daily and weekly returns on the market indices and S&P 500. The table lists the values of the LR statistic. The choice of this statistic is justified by Monte Carlo studies of the small sample properties of the various test statistics, which showed the LR statistic to have the best properties. Results for the other test statistics and additional data on the parameter estimates are listed in Appendix A. These tables also report specification tests using the overidentifying restrictions. These show that we do not reject the model specification.

Table 2 Hypothesis test of parameter constancy using the LR statistic. 1976-1989. Change point 1/1/1983. Daily and weekly returns. Numbers in parenthesis are probability values. 1: Test whether α is constant. 2: Joint test of whether α and μ are constant. 3: Joint test of whether α, μ and σ are constant. 4: Test for change in α, allowing for change in the variance σ. 5: Test for change in α, allowing for change in the mean μ.

		LR	
EW	1	2.71	(0.100)
	2	2.72	(0.257)
	3	3.01	(0.390)
	4	2.76	(0.097)
	5	2.24	(0.135)
S&P	1	6.30	(0.012)
	2	9.78	(0.008)
	3	15.12	(0.002)
	4	0.73	(0.392)
	5	8.74	(0.003)

Daily returns

		LR	
EW	1	0.00	(0.949)
	2	0.53	(0.769)
	3	1.09	(0.780)
	4	0.05	(0.829)
	5	0.11	(0.742)
S&P	1	0.67	(0.412)
	2	1.90	(0.387)
	3	1.98	(0.577)
	4	0.71	(0.400)
	5	1.23	(0.267)

Weekly returns

In the case of daily returns we do not reject a null of constant parameters on the EW index. For daily returns on S&P 500, we reject the null of parameter constancy at the 5% level for the first three hypotheses, but not for the fourth. In other words, as long as the conditional variance is assumed constant over the whole period, we reject, but if the conditional variance is allowed to change once, we no longer reject the null of constant autoregressive parameters. For weekly returns we can not reject the null of no change in the autoregressive relationship using any of the test statistics.

Since the previous results showed a sensitivity to the variance process, we next look at doing similar test assuming the variance follows a simple ARCH specification of Engle (1982). A number of authors have found evidence of ARCH effects in stock returns. For a recent summary of applications of ARCH modelling in finance, we refer to Bollerslev, Chou, and Kroner (1992). We formulate the estimation of the ARCH errors in a GMM framework, and test for parameter changes. A clear exposition of estimation of ARCH models by GMM can be found in Rich, Raymond, and Butler (1991). We view the ARCH parameterisation of conditional variance as a parsimonious way of introducing time varying conditional variance. If we can not reject the null of no change in the autoregressive parameter when we allow for a time–varying conditional variance, we may have more confidence that the changes in the sample autocorrelation coefficient are mainly attributable to changes in the conditional variance.

We use the ARCH framework in the simplest possible way. Assume mean returns follow

$$r_t = \mu + \alpha r_{t-1} + \varepsilon_t$$

where

$$E[\varepsilon_t | I_{t-1}] = 0$$
$$E[\varepsilon_t^2 | I_{t-1}] = h_t$$

and the conditional variance h_t follows

$$h_t = \eta_0 + \eta_1 \varepsilon_{t-1}^2 \quad \text{with} \ \ \eta_0 > 0, \ \ \text{and} \ \eta_1 < 1.$$

To estimate the model, we use the moment conditions

$$E[r_t - \mu - \alpha r_{t-1} | I_{t-1}] = 0$$
$$E[\varepsilon_t^2 - \eta_0 - \eta_1 \varepsilon_{t-1}^2 | I_{t-1}] = 0$$

We test a null hypothesis of no parameter change against an alternative

$\mathcal{H}_A$: The AR coefficient α changes, the other parameters stay constant.

Table 3 Hypothesis test of parameter constancy using the LR statistic. 1976-1989. Change point 1/1/1983. Daily and weekly returns. Numbers in parenthesis are probability values. 1: Test whether α is constant.

		LR	
EW	1	0.57	(0.449)
S&P	1	1.98	(0.159)

Daily returns

		LR	
EW	1	0.17	(0.684)
S&P	1	0.12	(0.727)

Weekly returns

The results for daily returns in table 3 show that we are not able to reject a null of no parameter change for either of the indices. The parameter estimates are listed in Appendix A. A specification test using overidentifying restrictions does not reject the model formulation. Hence, we can not reject the null of a constant autocorrelation of daily S&P 500 returns when we parameterise the variance process using the ARCH model.

We justified our choice of the time to test for change in the time series properties from the introduction of options. This assumes that this is the only significant event during the period in question. However, there has been a number of changes to the stock markets during this period. To gauge the importance of the introduction of new instruments versus other possible events, we next consider tests where the time of regime change is unknown. The idea is that we are not quite sure of when the regime change occurred, but we assume there is only one possible time of occurance. We look at the S&P 500 index.

Table 4 Hypothesis test of parameter constancy. S&P 500 index. Unknown time of parameter change. Period 1976 to 1989. See Section 3 for a description of the different tests. See table 12 for the critical probability values. Starred values are significant at the 5% level. Values in parenthesis are estimated dates of parameter change when we reject the null.

	$\sup W(\pi)$	$\sup LM(\pi)$	$\sup LR(\pi)$
1	10.87*	26.35*	25.74*
	(781027)	(781027)	(781020)
2	12.90*	32.13*	32.31*
	(781027)	(781027)	(781027)
3	61.29*	52.79*	74.54*
	(791024)	(791106)	(781020)
4	4.28	6.80	5.82
1	3.28	6.81	5.93
2	4.21	7.26	6.25
1	53.06*	18.15*	23.85*
	(851231)	(781027)	(781020)

Daily Returns

	$\sup W(\pi)$	$\sup LM(\pi)$	$\sup LR(\pi)$
1	6.60	2.36	3.38
2	15.90*	6.23	6.27
	(780216)		
3	25.26*	16.61*	24.56*
	(780302)	(790913)	(780309)
4	3.93	3.51	3.88
1	4.74	5.16	5.47
2	13.73*	7.72	9.07
	(780216)		
1	8.87	4.64	6.29

Weekly returns.

The results are listed in table 4. For daily returns the tests with constant conditional variance reject the null of no parameter change. Note however that the estimated dates are in the 1978–79 period, not around the time of introduction of the options. The tests with an ARCH variance (the last) also strongly reject the null of no parameter change, with the estimated breakpoint in 1978. For weekly returns, we also reject the null of parameter constancy in some cases, with the estimated time of parameter change in the 1978–79 period.

The results cast some doubt on whether we can claim that introduction of options on the index has been the source of observed changes in index time series behaviour. Of other important events in the financial markets we can mention the abandoning of fixed commissions in 1975, and the

Federal Reserve's change of operating procedure in 1979. If there truly was a change in the 1978–79 period, these events could be part of the explanation.

On a cautionary note, we should mention that the statistical properties of these tests are still relatively untested. We give some Monte Carlo evidence on these properties in the next section, and refer to the discussion there.

4.2 Crosscorrelations in size-based indices.

In this section we discuss one obvious way to extend the analysis of the previous sections. One of the well-known stylised facts in this area is that much of the index autocorrelation is related to cross-correlations across stocks. It is therefore natural to test for structural change in cross-correlations as well as the (own) autocorrelations. This may give us more power to detect any changes that have occurred. We therefore look at the cross-correlations of daily returns on 5 size-based indices.[3]

Table 5 lists (OLS) estimates of lagged cross-correlations of daily returns of 5 size-based portfolios for the subperiods 1976–82 and 1983–90. The pattern is well known. Portfolios of smaller stocks are correlated with past returns for other portfolios, and the portfolio of large stocks has low correlation with past returns of other portfolios. If we look across the two periods, the estimated cross-correlations have declined.

Table 5 Descriptive statistics, lagged cross-correlations. Size-based, equally weighted indices of NYSE stocks. Portfolio 1 is the portfolio of smallest firms. Daily returns. 1976–82 and 1983–89.

		R_{t-1}				
	Portfolio	1	2	3	4	5
	small	0.43	0.43	0.42	0.41	0.37
R_t	2	0.31	0.34	0.35	0.36	0.35
	3	0.24	0.27	0.30	0.32	0.32
	4	0.20	0.24	0.27	0.29	0.31
	large	0.11	0.15	0.17	0.20	0.23

1976–82

		R_{t-1}				
	Portfolio	1	2	3	4	5
	small	0.27	0.29	0.32	0.34	0.36
R_t	2	0.21	0.23	0.27	0.31	0.35
	3	0.16	0.19	0.23	0.28	0.33
	4	0.12	0.16	0.20	0.24	0.29
	large	-0.01	0.03	0.05	0.09	0.14

1983–89

We want to do formal testing of whether the declines in estimated cross–autocorrelations reflect a change in the linear predictability of returns. For simplicity, we use similar assumptions to the autoregressive case, and look at estimating the predictability of the return on index i from the return of index j in the previous period:

$$R_{t,i} = \mu + \alpha R_{t-1,j} + \varepsilon_t,$$

allowing for different assumptions about which parameters change and the distribution of the error term. Note that we only look at the two indices i and j in isolation, we do not estimate a joint

[3] The results for 10 size-based portfolios are qualitatively similar.

system.[4]

In this case, we consider four different hypotheses. In the first three tests we assume a constant conditional variance, $E[\varepsilon_t^2|I_{t-1}] = \sigma^2$, and in the fourth we let the variance follow a simple ARCH process, $E[\varepsilon_t^2|I_{t-1}] = h_t$, where $h_t = \eta_0 + \eta_1\varepsilon_{t-1}^2$.

The hypotheses we test are the following:

1. $\mathcal{H}_0$: All parameters μ, α and σ constant.
   $\mathcal{H}_A$: α changes, μ and σ constant.

2. $\mathcal{H}_0$: The conditional variance σ^2 allowed to change once, μ and α constant.
   $\mathcal{H}_A$: α and σ changes, μ constant.

3. $\mathcal{H}_0$: The mean parameter μ changes once, α and σ constant.
   $\mathcal{H}_A$: μ and α changes, σ constant.

4. $\mathcal{H}_0$: All parameters μ, α, η_0 and η_1 are constant.
   $\mathcal{H}_A$: α changes, μ, η_0 and η_1 constant.

The results from these tests are listed in table 6. All the four tests show similar results. The important changes in predictability are related to the large-sized portfolio. The ability of returns of the large sized portfolio to predict future returns of the smaller–sized portfolios has gone down, as seen by the rejections in the last columns. Also, the returns of the large–sized portfolio are less correlated with past returns of the other portfolios, as seen by the rejections in the last row of the tables. Note that these results are robust to assumptions about the conditional variance. Hence,

[4]Theoretically, the tests could be made more efficient by estimating all the parameters simultaneously, using the GMM equivalent of "Seemingly Unrelated" Regressions. Suppose we have p portfolios. Define

$$\mathbf{r}_t = (r_{1,t}, \cdots, r_{p,t})'$$

$$\mathbf{a} = (\mu_1, \cdots, \mu_p)'$$

$$\mathbf{B} = \begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ & & \vdots & \\ 0 & \cdots & 0 & \alpha_p \end{bmatrix}$$

$$\mathbf{e}_t = (\varepsilon_{1,t}, \cdots, \varepsilon_{p,t})'$$

The parameters $\mathbf{a}$ and $\mathbf{B}$ could be estimated simultaneously using the system

$$\mathbf{r}_t = \mathbf{a} + \mathbf{B}\mathbf{r}_{t-1} + \mathbf{e}_t$$

The efficiency would be increased if elements of the error term $\mathbf{e}_t$ are correlated, and this is taken into account in the estimation of the variance-covariance matrix. This framework would also allow for testing for change in only some of the parameters. However, when implementing this approach, we found that the numerical optimisation routine could not find a optimum with any degree of confidence, the surface is extremely flat given the number of parameters we are trying to find. We therefore do not report any results using this approach.

Table 6 Results of hypothesis tests of changes in cross-autocorrelation for index returns. Daily returns. 5 size based indices of NYSE stocks. Index 1 is the index of smallest stocks. Numbers are values of the LR statistic, which have a χ^2 distribution. Numbers in parenthesis are probability values.

		R_{t-1}				
		1	2	3	4	5
R_t	1	3.70	3.18	3.10	**6.10**	**12.98**
		(0.054)	(0.074)	(0.078)	(0.014)	(0.000)
	2	3.13	2.35	2.08	**3.97**	**8.12**
		(0.077)	(0.125)	(0.149)	(0.046)	(0.004)
	3	2.08	1.33	1.36	3.18	**8.22**
		(0.149)	(0.250)	(0.243)	(0.075)	(0.004)
	4	3.26	1.52	0.70	1.53	**5.03**
		(0.071)	(0.217)	(0.403)	(0.217)	(0.025)
	5	**10.81**	**6.95**	**4.76**	**5.50**	**6.61**
		(0.001)	(0.008)	(0.029)	(0.019)	(0.010)

Test 1.

		R_{t-1}				
		1	2	3	4	5
R_t	1	3.54	2.76	2.97	**5.94**	**13.29**
		(0.060)	(0.097)	(0.085)	(0.015)	(0.000)
	2	3.28	2.56	2.31	**4.25**	**8.80**
		(0.070)	(0.109)	(0.128)	(0.039)	(0.003)
	3	2.00	1.57	1.59	3.17	**7.54**
		(0.157)	(0.211)	(0.208)	(0.075)	(0.006)
	4	0.24	0.24	0.35	1.51	**5.23**
		(0.626)	(0.621)	(0.555)	(0.218)	(0.022)
	5	0.95	1.68	1.68	2.52	**3.96**
		(0.329)	(0.195)	(0.195)	(0.112)	(0.047)

Test 2.

		R_{t-1}				
		1	2	3	4	5
R_t	1	1.57	2.06	1.71	3.31	**7.35**
		(0.210)	(0.151)	(0.191)	(0.069)	(0.007)
	2	1.16	0.96	0.61	1.18	2.78
		(0.281)	(0.328)	(0.434)	(0.278)	(0.095)
	3	1.37	0.81	0.51	1.19	3.73
		(0.242)	(0.369)	(0.474)	(0.274)	(0.053)
	4	**4.91**	2.05	0.84	1.40	**4.29**
		(0.027)	(0.152)	(0.358)	(0.237)	(0.038)
	5	**18.89**	**8.77**	**5.58**	**6.27**	**7.24**
		(0.000)	(0.003)	(0.018)	(0.012)	(0.007)

Test 3.

		R_{t-1}				
		1	2	3	4	5
R_t	1	**5.40**	3.78	3.64	**5.89**	**13.45**
		(0.020)	(0.052)	(0.056)	(0.015)	(0.000)
	2	3.27	0.84	0.24	0.53	2.61
		(0.071)	(0.359)	(0.627)	(0.465)	(0.106)
	3	2.60	0.66	0.08	0.16	2.08
		(0.107)	(0.417)	(0.775)	(0.690)	(0.149)
	4	2.37	1.28	0.22	0.43	2.45
		(0.123)	(0.257)	(0.641)	(0.514)	(0.118)
	5	**7.90**	**10.57**	**7.42**	**6.79**	**6.74**
		(0.005)	(0.001)	(0.006)	(0.009)	(0.009)

Test 4.

we conclude that there has been a change in the time series properties of the returns, and that this is related to the large sized portfolio.

5 Small sample properties of the test statistics.

One possible problem with the statistics we have used is the fact that they all rely on asymptotic theory. It is therefore valuable to have some information about their properties in small samples. To this end, we report some Monte Carlo investigations.

5.1 Statistics with a known time of structural change.

We simulate return series following $r_t = \mu + \alpha_1 r_{t-1} + \varepsilon_t$ for $t = 1, \cdots, T_1$ and $r_t = \mu + \alpha_2 r_{t-1} + \varepsilon_t$ for $t = T_1 + 1, \cdots, T$, with $\varepsilon_t \sim N(0, \sigma^2)$. The normality assumption is mainly for convenience, similar simulations have been done with a number of non-normal error assumptions, and the conclusions are not sensitive to the error distribution. We therefore report only the results using a normal error.

These parameters are used to generate series of $T = 200$ return observations, with a possible breakpoint at $T_1 = 100$. We use a small number of observations in order to highlight the differences between the statistics. It can be viewed as a 'worst case' scenario, and the differences between statistics will be smaller for larger sample sizes. The breakpoint was (arbitrarily) picked as the midpoint in the series.

We first do some comparisons across instruments Z_t. We consider two instruments, one is the previous period return ($Z_{t-1} = r_{t-1}$), and the other is the previous period forecast error ($Z_{t-1} = \varepsilon_{t-1}$). The simulation results are listed in tables 7 and 8 below.

Table 7 Percentage of rejections of the null when the null is true. Parameter values $\mu = 0.10$, $\alpha_1 = \alpha_2 = 0.30$ and $\sigma^2 = 0.05$. 2500 simulations.

		Statistic				
		GH_1	GH_2	W	LM	LR
Rejection	10%	25.64	28.60	2.80	5.20	0.96
level	5%	16.44	19.64	1.44	2.08	0.32
	1%	5.44	8.56	0.32	0.36	0.00

Instruments are past returns.

		Statistic				
		GH_1	GH_2	W	LM	LR
Rejection	10%	32.80	32.96	9.20	12.08	11.88
level	5%	23.20	22.08	3.68	6.16	6.08
	1%	10.44	9.44	0.72	0.88	1.20

Instruments are past forecast errors.

Table 8 Percentage of rejections of the null when the null is false. Parameter values $\mu = 0.10$, $\alpha_1 = 0.50$, $\alpha_2 = 0.10$ and $\sigma^2 = 0.05$. 2500 simulations.

		Statistic				
		GH_1	GH_2	W	LM	LR
Rejection	10%	93.80	92.72	20.52	84.60	22.76
level	5%	86.36	86.92	12.48	74.68	18.72
	1%	62.24	67.20	3.92	47.28	11.08

Instruments are past returns.

		Statistic				
		GH_1	GH_2	W	LM	LR
Rejection	10%	88.36	89.52	87.92	92.04	91.72
level	5%	78.80	82.48	76.60	85.56	86.28
	1%	52.64	62.24	42.72	64.76	66.28

Instruments are past forecast errors.

First, consider the results with no parameter change, listed in table 7. The Ghysels and Hall statistics GH_1 and GH_2 over-reject the null, but the Andrews and Fair statistics W, LM and LR are close to the correct rejection percentages. Comparing the two instruments, using past forecast errors as instruments is clearly preferable to using past returns. This point is made even stronger in table 8. This table is generated *with* a parameter change, the autoregressive parameter α changes from 0.1 to 0.5. The statistics using forecast errors have higher power than the ones using past returns. Because of these results, we have in the the paper concentrated on estimation using forecast errors as instruments.

We next do a more detailed investigation of the statistics using forecast errors as instruments. In addition to looking at rejection percentages, we examine the ability of the statistics to estimate power from the sample. For a specified alternative and sample size, the statistics will follow noncentral chi-square distributions with noncentrality parameters that depend on the covariance matrix. If we use the covariance matrix estimated from the data, we can calculate an estimated noncentrality parameter. Using this will give an estimate of the power. In the simulations we compare the results of this calculation with the percentage of rejections in the sample. Appendix B gives further details on the power calculations. Table 9 below shows the results.

Table 9 "Power" calculations for alternative parameter values. The numbers are percentage rejections in the sample under the specified alternative α_2. The "estimated power" calculations are averages of the power calculated from the sample. 2500 simulations used.

					GH Statistics			AF Statistics			
T	α_1	α_2	η	Reject at	GH_1	GH_2	Est.Power	W	LM	LR	Est.Power
200	0.20	0.20	0.000	10%	31.72	32.04	10.00	7.84	9.92	9.64	10.00
				5%	21.72	21.36	5.00	3.68	4.84	4.96	5.00
				1%	9.36	9.20	1.00	0.44	1.20	1.28	1.00
		0.30	1.414	10%	40.72	41.12	32.85	15.68	17.96	17.64	35.91
				5%	30.48	31.32	22.12	7.32	10.40	10.44	24.91
				1%	15.92	16.92	8.37	1.56	2.88	3.08	9.82
		0.40	2.828	10%	59.92	62.00	84.32	34.28	39.92	39.84	81.14
				5%	48.52	51.12	76.38	20.40	28.00	27.60	71.69
				1%	28.56	33.28	56.76	5.56	10.64	10.68	48.78
		0.50	4.243	10%	79.52	81.96	99.28	60.68	67.92	67.52	98.07
				5%	72.04	75.08	98.56	42.92	55.40	54.64	96.14
				1%	52.32	58.36	95.46	12.76	29.72	29.72	88.04

If we first look at the results for the GH statistics, we note that GH_2 gives better results than GH_1. Both the GH statistics seriously over-reject the null. The GH and AF statistics are not directly comparable, since in calculating the AF statistics, the form of the parameter change is known, and the AF statistics are designed to test against this particular form of parameter change. The GH statistics may have better properties in cases with other (unspecified) types of parameter change. The Wald statistic has the lowest power of the three AF statistics. This may be related to the well-known problems in using the Wald statistic, that it is in finite samples not invariant to rotations of the moment conditions.[5] There is little difference between LM and LR.

If we now turn to the power calculations, it is clear that the power estimated from the sample is seriously overstated. For the cases considered here, it seems to of little practical usefulness. We have therefore not reported it in the tests on stock market data.

[5] For some examples in the ML case, see Dagenais and Dufour (1991) and the references therein.

5.2 Statistics with an unknown time of structural change.

We next investigate the tests with an unknown time of parameter change. The moment conditions and the simulated series are the same as above. The time of parameter change is again the midpoint of the series. In table 10 we list the percentages of rejection of the null in the case when there is no actual change in the parameter. Table 11 lists the same percentages when the null is false, the autoregressive parameter α changes.

Table 10 Percentage rejections of the null when the null is true. Parameter values $\mu = 0.10$, $\alpha_1 = \alpha_2 = 0.30$, $\sigma^2 = 0.05$. 2500 simulations.

		W	LM	LR
Rejection	1%	17.00	0.52	1.16
level:	5%	25.56	2.92	5.56
	10%	31.28	6.56	10.68

Table 11 Percentage rejections of the null when the null is false. Parameter values $\mu = 0.10$, $\alpha_1 = 0.50$, $\alpha_2 = 0.10$ and $\sigma^2 = 0.05$. 2500 simulations.

		W	LM	LR
Rejection	1%	48.56	33.96	39.00
level:	5%	65.72	65.16	67.96
	10%	73.40	77.08	79.16

The results show that the Wald statistic seriously over-rejects the null when the null is true. The other statistics have more reasonable behaviour. In particular, the LR statistic performs very well. The problem with the Wald statistic is caused by the small sample. The Wald statistic uses an unrestricted estimate of the parameters, where some of the parameters are estimated in each subperiod. The smallest subperiod contains 15% of the total sample, which in this case is only 30 observations. This may be to be too few observations to obtain a reasonable estimate of the parameters. This is illustrated in figures 4 and 5. The figures shows histograms of estimated time of structural change when the null is rejected, in a case where there actually *is* a structural change. The estimates for the Wald statistics are concentrated in the tails, whereas the LR estimates are concentrated at the true breakpoint of $T_1 = 100$. Unreported results show that for increased sample sizes the behaviour of the Wald statistic improves. However, the LM and LR statistics are clearly preferred to the Wald statistic.

In view of the results of Section 4.1, where we found an estimated breakpoint which was close to one of the endpoints of the period, it may also be of interest to investigate how these tests behaves if the true time of structural change is away from the midpoint of the time period.

Figure 4 Estimates of breakpoint when the null is rejected, using the Wald tests. True breakpoint at $t = 100$. Parameter values $\mu = 0.10$, $\alpha_1 = 0.50$, $\alpha_2 = 0.10$ and $\sigma^2 = 0.05$. 2500 simulations.

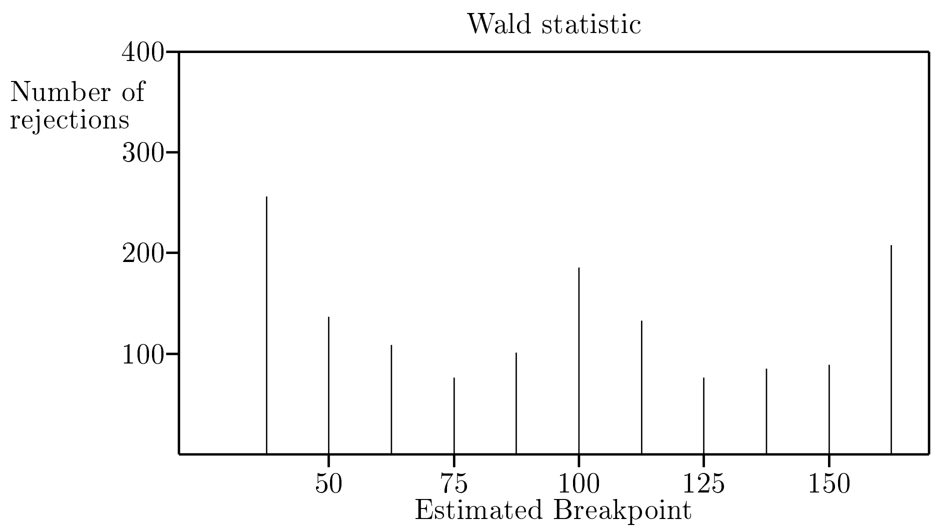
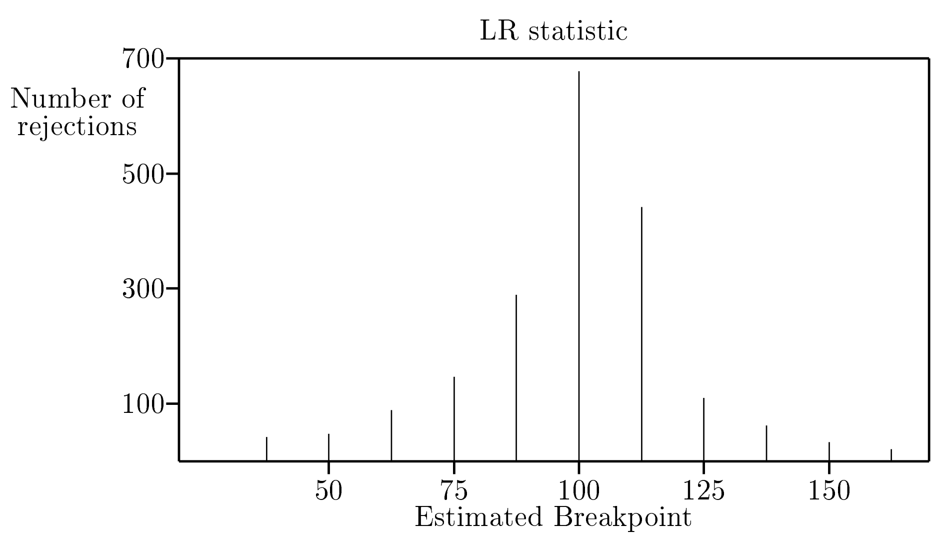


Figure 5 Estimates of breakpoint when the null is rejected, using the Likelihood Ratio test. True breakpoint at $t = 100$. Parameter values $\mu = 0.10$, $\alpha_1 = 0.50$, $\alpha_2 = 0.10$ and $\sigma^2 = 0.05$. 2500 simulations.



6 Conclusion

This paper investigated possible changes in autoregressive relationships of stock indices when derivative securities were introduced. We showed how this could be explicitly tested by applying tests for structural change in a GMM estimation context, due to Andrews and Fair (1988) and Ghysels and Hall (1990) and Andrews (1990).

The small-sample properties of the statistics were investigated using Monte Carlo. The Ghysels and Hall statistics were shown to have bad small sample properties for this problem. For the Andrews and Fair statistics, we note that the "Likelihood Ratio–like" and "Lagrange Multiplier–like" statistics perform better than the Wald statistic. This is in particular true for tests where the time of parameter change is unknown.

We used the tests on US stock market data for the period 1976-1989. We first considered cases with a constant conditional variance. We could then reject a null of parameter constancy of the autoregressive parameter for daily S&P 500 index returns. We did not reject for daily EW returns. By looking at size-based portfolios we could confirm that a major source of the rejections on the value-weighted S&P 500 index is the behaviour of stocks of large-sized firms. We also looked at a parameterisation where we allowed the conditional variance to change by using an ARCH process. With these assumptions we could not reject a null hypothesis of a constant autoregressive parameter on the daily S&P 500 returns. We therefore conclude that a major influence of the observed decline in the sample autocorrelations of S&P 500 returns is changes in conditional variance. This increase in variance will increase the uncertainty in estimates of the autocorrelation, which can result in a decline in the sample autocorrelation.

To further investigate causes of the changes in autoregressive relationships, we looked at cross–correlations across stocks, in view of the well–known "stylised fact" that a source of much of index autocorrelation is lagged cross–correlations across stocks. We found that there has been a change in the return predictability concentrated on the largest-sized portfolio.

APPENDIX.

A Detailed estimation results.

In this appendix we list the results for all the statistics, together with the point estimates of the parameters calculated in the tests with known breakpoint.

A.1 Tests with fixed conditional variance.

Daily Returns.

	GH_1		GH_2	
EW	6.24	(0.182)	7.65	(0.105)
S&P	18.53	(0.001)	30.77	(0.000)

		W		LM		LR	
EW	1	3.03	(0.082)	3.95	(0.047)	2.71	(0.100)
	2	3.02	(0.221)	3.90	(0.142)	2.72	(0.257)
	3	2.87	(0.412)	5.16	(0.160)	3.01	(0.390)
	4	2.64	(0.104)	2.99	(0.084)	2.76	(0.097)
S&P	1	3.09	(0.079)	8.76	(0.003)	6.30	(0.012)
	2	4.71	(0.095)	12.98	(0.002)	9.78	(0.008)
	3	13.83	(0.003)	16.24	(0.001)	15.12	(0.002)
	4	0.61	(0.437)	1.20	(0.274)	0.73	(0.392)

Hypothesis test of parameter constancy. 1976-1989. Change point 1/1/1983. Daily returns. Numbers in parenthesis are probability values. Tests: 1: Test whether α is constant. 2: Joint test of whether α and μ are constant. 3: Joint test of whether α, μ and σ are constant. 4: Test for change in α, allowing for change in the variance σ.

	μ		α		σ^2		χ^2	
OLS	0.0007		0.2739		0.00006			
GMM	0.0007		0.2624		0.00005			
	(0.0001)		(0.0561)		(0.00001)		1.27	(0.26)
OLS 1	0.0007		0.3253		0.00005			
GMM 1	0.0006		0.3838		0.00005			
	(0.0002)		(0.0375)		(0.00000)		3.64	(0.06)
OLS 2		0.0006		0.2282		0.00006		
GMM 2		0.0005		0.2406		0.00005		
		(0.0002)		(0.1027)		(0.00001)	0.87	(0.35)
GMM Restr	0.0005		0.3713		0.00005			
	(0.0001)		(0.0354)		(0.00000)		4.15	(0.13)
GMM Unrestr 1	0.0006		0.3856	0.2250	0.00005			
	(0.0001)		(0.0368)	(0.0930)	(0.00000)		2.80	(0.59)
GMM Unrestr 2	0.0006	0.0006	0.3844	0.2289	0.0000			
	(0.0002)	(0.0002)	(0.0385)	(0.1012)	(0.00000)		2.79	(0.42)
GMM Unrestr 3	0.0006	0.0005	0.3830	0.2275	0.00005	0.00005		
	(0.0002)	(0.0002)	(0.0385)	(0.1012)	(0.00000)	(0.00001)	2.64	(0.27)
GMM Restr 4	0.0005		0.3725		0.00005	0.00005		
	(0.0001)		(0.0356)		(0.00000)	(0.00001)	4.06	(0.26)
GMM Unrestr 4	0.0006		0.3859	0.2182	0.00005	0.00005		
	(0.0001)		(0.0368)	(0.0937)	(0.00000)	(0.00001)	2.68	(0.44)
Equally weighted index								

	μ		α		σ^2		χ^2	
OLS	0.0004		0.0674		0.00010			
GMM	0.0005		0.0576		0.00009			
	(0.0002)		(0.0579)		(0.00001)		1.14	(0.29)
OLS 1	0.0002		0.1286		0.00007			
GMM 1	0.0002		0.1288		0.00007			
	(0.0002)		(0.0262)		(0.00000)		0.07	(0.79)
OLS 2		0.0006		0.0311		0.00012		
GMM 2		0.0007		0.0057		0.00010		
		(0.0003)		(0.0853)		(0.00001)	1.06	(0.30)
GMM Restr	0.0004		0.1074		0.00007			
	(0.0002)		(0.0243)		(0.00000)		8.14	(0.02)
GMM Unrestr 1	0.0005		0.1292	-0.0571	0.00007			
	(0.0002)		(0.0256)	(0.0764)	(0.00000)		4.98	(0.29)
GMM Unrestr 2	0.0003	0.0009	0.1306	-0.0928	0.0001			
	(0.0002)	(0.0003)	(0.0259)	(0.0793)	(0.00000)		3.24	(0.36)
GMM Unrestr 3	0.0002	0.0007	0.1288	0.0163	0.00007	0.00010		
	(0.0002)	(0.0003)	(0.0259)	(0.0938)	(0.00000)	(0.00001)	0.57	(0.75)
GMM Restr 4	0.0004		0.1212		0.00007	0.00011		
	(0.0002)		(0.0248)		(0.00000)	(0.00001)	1.79	(0.62)
GMM Unrestr 4	0.0004		0.1274	0.0522	0.00007	0.00010		
	(0.0002)		(0.0257)	(0.0885)	(0.00000)	(0.00001)	1.42	(0.70)
S&P 500 index								

Parameter estimates from estimation. Daily returns. 1976 to 1989. Breakpoint 1/1/83. GMM is estimated using all the data. GMM 1 is estimated using the first subperiod's data, GMM 2 uses the second subperiod's data. The restricted GMM estimator GMM restr is calculated imposing the null hypothesis. GMM unrestr is calculated without imposing the null. OLS is the corresponding Least Squares estimates. The values below the point estimates are estimated standard deviations. The last two columns lists the chi-square test of overidentifying restrictions.

Weekly returns.

	GH_1		GH_2	
EW	1.35	(0.854)	1.16	(0.885)
S&P	3.08	(0.544)	2.61	(0.626)

		W		LM		LR	
EW	1	0.00	(0.949)	0.00	(0.949)	0.00	(0.949)
	2	0.49	(0.784)	0.52	(0.771)	0.53	(0.769)
	3	1.13	(0.769)	1.14	(0.768)	1.09	(0.780)
	4	0.05	(0.831)	0.04	(0.833)	0.05	(0.829)
S&P	1	0.67	(0.414)	0.68	(0.410)	0.67	(0.412)
	2	1.69	(0.430)	2.03	(0.362)	1.90	(0.387)
	3	1.71	(0.634)	2.13	(0.546)	1.98	(0.577)
	4	0.71	(0.399)	0.71	(0.398)	0.71	(0.400)

Hypothesis test of parameter constancy. 1976-1989. Change point 1/1/1983. Weekly returns. Numbers in parenthesis are probability values. 1: Test whether α is constant. 2: Joint test of whether α and μ are constant. 3: Joint test of whether α, μ and σ are constant. 4: Test for change in α, allowing for change in the variance σ.

	μ		α		σ^2		χ^2	
OLS	0.0034		0.2261		0.00045			
GMM	0.0034		0.2405		0.00039			
	(0.0009)		(0.0658)		(0.00005)		2.44	(0.12)
OLS 1	0.0042		0.2020		0.00049			
GMM 1	0.0036		0.2534		0.00043			
	(0.0014)		(0.0767)		(0.00005)		2.72	(0.10)
OLS 2		0.0026		0.2459		0.00040		
GMM 2		0.0028		0.2568		0.00036		
		(0.0011)		(0.1033)		(0.00008)	0.38	(0.54)
GMM Restr	0.0030		0.2679		0.00041			
	(0.0008)		(0.0597)		(0.00004)		2.03	(0.36)
GMM Unrestr 1	0.0030		0.2658	0.2728	0.00041			
	(0.0009)		(0.0685)	(0.0964)	(0.00004)		2.02	(0.73)
GMM Unrestr 2	0.0037	0.0026	0.2452	0.2879	0.0004			
	(0.0013)	(0.0011)	(0.0742)	(0.0991)	(0.00004)		1.76	(0.62)
GMM Unrestr 3	0.0036	0.0028	0.2496	0.2556	0.00043	0.00036		
	(0.0013)	(0.0011)	(0.0750)	(0.1071)	(0.00005)	(0.00008)	1.47	(0.48)
GMM Restr 4	0.0031		0.2579		0.00044	0.00036		
	(0.0009)		(0.0606)		(0.00005)	(0.00007)	1.63	(0.65)
GMM Unrestr 4	0.0031		0.2649	0.2404	0.00044	0.00035		
	(0.0009)		(0.0686)	(0.1014)	(0.00005)	(0.00008)	1.60	(0.66)
Equally weighted index								

	μ		α		σ^2		χ^2	
OLS	0.0020		0.0395		0.00045			
GMM	0.0022		0.0212		0.00041			
	(0.0008)		(0.0503)		(0.00003)		1.54	(0.21)
OLS 1	0.0012		0.0455		0.00043			
GMM 1	0.0012		0.0975		0.00043			
	(0.0011)		(0.0653)		(0.00004)		2.22	(0.14)
OLS 2		0.0027		0.0246		0.00046		
GMM 2		0.0031		-0.0080		0.00041		
		(0.0011)		(0.0634)		(0.00005)	0.66	(0.42)
GMM Restr	0.0022		0.0354		0.00043			
	(0.0008)		(0.0445)		(0.00003)		2.57	(0.28)
GMM Unrestr 1	0.0022		0.0726	0.0014	0.00043			
	(0.0008)		(0.0632)	(0.0618)	(0.00003)		2.23	(0.69)
GMM Unrestr 2	0.0013	0.0030	0.0835	-0.0053	0.0004			
	(0.0011)	(0.0011)	(0.0664)	(0.0632)	(0.00003)		1.61	(0.66)
GMM Unrestr 3	0.0013	0.0031	0.0852	-0.0073	0.00043	0.00041		
	(0.0011)	(0.0011)	(0.0667)	(0.0638)	(0.00004)	(0.00005)	1.57	(0.46)
GMM Restr 4	0.0022		0.0351		0.00043	0.00043		
	(0.0008)		(0.0446)		(0.00004)	(0.00005)	2.56	(0.47)
GMM Unrestr 4	0.0022		0.0737	-0.0001	0.00044	0.00042		
	(0.0008)		(0.0633)	(0.0623)	(0.00004)	(0.00005)	2.20	(0.53)
S&P 500 index								

Parameter estimates. Weekly returns. 1976 to 1989. Breakpoint 1/1/83. GMM is estimated using all the data. GMM 1 is estimated using the first subperiod's data, GMM 2 uses the second subperiod's data. The restricted GMM estimator GMM restr is calculated imposing the null hypothesis. GMM unrestr is calculated without imposing the null. OLS is the corresponding Least Squares estimates. The values below the point estimates are estimated standard deviations. The last two columns show the results of the chi-square test of overidentifying restrictions (J-test).

A.2 Tests with an ARCH variance.

Daily returns.

	GH_1		GH_2	
EW	10.37	(0.110)	13.72	(0.033)
S&P	21.25	(0.002)	32.09	(0.000)

		W		LM		LR	
EW	1	0.36	(0.549)	0.49	(0.485)	0.57	(0.449)
S&P	1	1.42	(0.234)	1.67	(0.197)	1.98	(0.159)

Hypothesis test of parameter constancy. 1976-1989. Change point 1/1/1983. Daily returns. Numbers in parenthesis are probability values. Tests: 1: Test whether β_1 is constant.

	β_0	β_1	α_0	α_1	χ^2
OLS	0.0007	0.2739	0.00006		
GMM	0.0007	0.2861	0.00004	0.04352	
	(0.0001)	(0.0537)	(0.00000)	(0.02983)	3.50 (0.17)
OLS 1	0.0007	0.3253	0.00005		
GMM 1	0.0007	0.3549	0.00005	-0.05647	
	(0.0002)	(0.0333)	(0.00000)	(0.04243)	7.84 (0.02)
OLS 2	0.0006	0.2282	0.00006		
GMM 2	0.0006	0.2708	0.00005	0.03389	
	(0.0002)	(0.1013)	(0.00001)	(0.02474)	2.56 (0.28)
GMM Restr	0.0006	0.3394	0.00004	0.01224	
	(0.0001)	(0.0298)	(0.00000)	(0.01659)	6.95 (0.43)
GMM Unrestr 1	0.0006	0.3473 0.2928	0.00004	0.01490	
	(0.0001)	(0.0311) (0.0793)	(0.00000)	(0.01847)	6.67 (0.46)
Equally weighted index					

	β_0	β_1	α_0	α_1	χ^2
OLS	0.0004	0.0674	0.00010		
GMM	0.0004	0.0947	0.00008	0.06665	
	(0.0002)	(0.0444)	(0.00000)	(0.04664)	1.43 (0.49)
OLS 1	0.0002	0.1286	0.00007		
GMM 1	0.0002	0.1316	0.00006	0.07887	
	(0.0002)	(0.0260)	(0.00000)	(0.02952)	0.47 (0.79)
OLS 2	0.0006	0.0311	0.00012		
GMM 2	0.0006	0.0601	0.00010	0.06011	
	(0.0002)	(0.0724)	(0.00001)	(0.05044)	1.46 (0.48)
GMM Restr	0.0004	0.1105	0.00007	0.03766	
	(0.0001)	(0.0238)	(0.00000)	(0.02010)	11.51 (0.12)
GMM Unrestr 1	0.0005	0.1211 0.0246	0.00007	0.02539	
	(0.0001)	(0.0256) (0.0600)	(0.00000)	(0.01955)	10.52 (0.16)
S&P 500 index					

Parameter estimates. Daily returns. 1976 to 1989. Breakpoint 1/1/83. GMM is estimated using all the data. GMM 1 is estimated using the first subperiod's data, GMM 2 uses the second subperiod's data. The restricted GMM estimator GMM restr is calculated imposing the null hypothesis. GMM unrestr is calculated without imposing the null. OLS is the corresponding Least Squares estimates. The values below the point estimates are estimated standard deviations. The last two columns show the results of the chi-square test of overidentifying restrictions.

Weekly returns.

	GH_1		GH_2	
EW	3.11	(0.795)	3.72	(0.714)
S&P	6.41	(0.379)	4.82	(0.567)

		W		LM		LR	
EW	1	0.09	(0.768)	0.08	(0.772)	0.17	(0.684)
S&P	1	0.13	(0.723)	0.12	(0.731)	0.12	(0.727)

Hypothesis test of parameter constancy. 1976-1989. Change point 1/1/1983. Weekly returns. Numbers in parenthesis are probability values. Tests: 1: Test whether β_1 is constant.

	β_0	β_1		α_0	α_1	χ^2	
OLS	0.0034	0.2261		0.00045			
GMM	0.0038	0.2119		0.00035	0.10163		
	(0.0009)	(0.0652)		(0.00005)	(0.06275)	2.63	(0.27)
OLS 1	0.0042	0.2020		0.00049			
GMM 1	0.0038	0.2353		0.00041	0.03675		
	(0.0012)	(0.0679)		(0.00007)	(0.09612)	2.84	(0.24)
OLS 2	0.0026	0.2459		0.00040			
GMM 2	0.0029	0.2966		0.00036	0.03529		
	(0.0011)	(0.0998)		(0.00008)	(0.02395)	1.84	(0.40)
GMM Restr	0.0033	0.2558		0.00039	0.04991		
	(0.0008)	(0.0568)		(0.00004)	(0.01846)	2.83	(0.90)
GMM Unrestr 1	0.0033	0.2440	0.2878	0.00039	0.04519		
	(0.0008)	(0.0651)	(0.0942)	(0.00004)	(0.01977)	2.74	(0.91)
Equally weighted index							

	β_0	β_1		α_0	α_1	χ^2	
OLS	0.0020	0.0395		0.00045			
GMM	0.0021	0.0118		0.00038	0.10441		
	(0.0008)	(0.0481)		(0.00003)	(0.05100)	1.61	(0.45)
OLS 1	0.0012	0.0455		0.00043			
GMM 1	0.0006	0.1030		0.00031	0.15399		
	(0.0011)	(0.0702)		(0.00004)	(0.09134)	8.67	(0.01)
OLS 2	0.0027	0.0246		0.00046			
GMM 2	0.0031	0.0265		0.00039	0.05191		
	(0.0011)	(0.0553)		(0.00004)	(0.04031)	1.19	(0.55)
GMM Restr	0.0019	0.0399		0.00036	0.06280		
	(0.0008)	(0.0415)		(0.00003)	(0.03371)	8.39	(0.30)
GMM Unrestr 1	0.0018	0.0570	0.0286	0.00036	0.06312		
	(0.0008)	(0.0641)	(0.0522)	(0.00003)	(0.03346)	8.33	(0.30)
S&P 500 index							

Parameter estimates. Weekly returns. 1976 to 1989. Breakpoint 1/1/83. GMM is estimated using all the data. GMM 1 is estimated using the first subperiod's data, GMM 2 uses the second subperiod's data. The restricted GMM estimator GMM restr is calculated imposing the null hypothesis. GMM unrestr is calculated without imposing the null. OLS is the corresponding Least Squares estimates. The values below the point estimates are estimated standard deviations. The last two columns show the results of the chi-square test of overidentifying restrictions.

B A general description of the GMM estimation and the stability tests.

This appendix hold the formal definitions of the test statistics used in the paper.

B.1 GMM.

To summarise the GMM estimation method, define $f(x_t, \theta)$ as the vector of moment conditions, where $\{x_t\}_{t=1}^T$ is the data, and θ is the vector of parameters. With this notation the moment conditions are expressed as

$$E[f(x_t, \theta_0)] = 0$$

where θ_0 is the true parameter. The optimal GMM estimator $\hat{\theta}_T$ is defined to be

$$\hat{\theta}_T = \arg\min_{\theta} \quad g(X, \theta)' \hat{A}_T g(X, \theta)$$

where

$$g(X, \theta) = \frac{1}{T} \sum_{t=1}^{T} f(x_t, \theta).$$

The weighting matrix $\hat{A}_T$ is a consistent estimator of S^{-1}, where

$$S(\theta) = \sum_{t=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} E[f(x_t, \theta) f(x_{t+s}, \theta)'],$$

the covariance matrix of the moment conditions. The strategy in estimation is to use an initial consistent set of parameters, say θ^*, to generate an initial estimate $\hat{A}_T$. This estimated $\hat{A}_T$ is then used in the further estimation of $\hat{\theta}_T$. In the paper, we use the OLS estimate as an initial consistent estimate.

B.2 The Andrews and Fair statistics.

We first consider the tests formulated by Andrews and Fair (1988). These are tests where the breakpoint is known. The sample is split into two, and the moment conditions are formulated as:

$$f(x_t, \theta) = \begin{cases} \begin{bmatrix} f_1(x_t, \theta_1, \theta_3) \\ 0 \end{bmatrix} & \text{for } x_t \in X_1 \\ \begin{bmatrix} 0 \\ f_2(x_t, \theta_2, \theta_3) \end{bmatrix} & \text{for } x_t \in X_2 \end{cases}$$

The estimation minimises

$$d(X, \theta) = \tfrac{1}{2} g(X, \theta)' \hat{A}_T g(X, \theta)$$

where

$$g(X, \theta) = \frac{1}{T} \sum_{t=1}^{T} f(x_t, \theta).$$

The test statistics proposed by Andrews and Fair (1988) uses the following matrices.

$$S = \sum_{t=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} E[f(x_t, \theta) f(x_{t+s}, \theta)']$$

$$M = E[\frac{\partial}{\partial \theta'} f(x_t, \theta)]$$

$$D = E\left[\frac{\partial^2}{\partial f \partial f'} \tfrac{1}{2} f(x_t, \theta) A f(x_t, \theta)'\right] = A$$

$$\mathcal{I} = M' D M$$

$$\mathcal{J} = M' D S D M$$

$$V = \mathcal{J}^{-1} \mathcal{I} \mathcal{J}^{-1}$$

The estimators of M and D are obtained from their sample counterparts. If there is temporal dependence in the moment conditions, to guarantee a positive definite weighting matrix, the estimator $\hat{S}_T$ of S can be calculated using methods analogous to the estimators of covariance matrices proposed by Andrews (1991), Eichenbaum, Hansen, and Singleton (1988) and Newey and West (1987).

In order to define the tests, describe the hypothesis to be tested in the form

$$h(\theta) = 0$$

In the case of pure parameter change

$$h(\theta) = \theta_1 - \theta_2 = 0$$

Define

$$H = \frac{\partial}{\partial \theta'} h(\theta)$$

On our case $\theta = (\theta_1, \theta_2, \theta_3)$ and $H = [1, -1, 0]$.

Let $\tilde{\theta}_T$ define the restricted GMM estimator of θ, where we impose the null hypothesis $h(\theta) = 0$. Similarly, let $\hat{\theta}_T$ be the unrestricted GMM estimator of θ.

The Wald statistic W_T is defined as

$$W_T = Th(\hat{\theta})'(\hat{H}\hat{V}\hat{H}')^{-1}h(\hat{\theta})$$

The LM statistic is defined as

$$LM_T = T\left(\frac{\partial}{\partial\theta'}d(X,\tilde{\theta})\right)'\tilde{\mathcal{I}}^{-1}\tilde{H}'(\tilde{H}\tilde{V}\tilde{H}')^{-1}\tilde{H}\tilde{\mathcal{I}}^{-1}\left(\frac{\partial}{\partial\theta'}d(X,\tilde{\theta})\right)$$

where

$$\tilde{V} = \tilde{\mathcal{J}}^{-1}\tilde{\mathcal{I}}\tilde{\mathcal{J}}^{-1}$$

and

$$\tilde{\mathcal{I}} = \tilde{M}'\tilde{D}\hat{S}\tilde{D}\tilde{M}$$

In the GMM context, assumption 6a (page 624) of Andrews and Fair is fulfilled, $\mathcal{I} = b\mathcal{J}$ for $b = 1$. The Likelihood Ratio statistic LR is in this case defined as

$$LR_T = 2T[d(X,\tilde{\theta}) - d(X,\hat{\theta})]$$

By theorem 4 of Andrews and Fair, all the statistics W_T, LM_T and LR_T converges in distribution to a chi-square statistic with degrees of freedom equal to the number of restrictions in the hypothesis.

Power calculations. Under the standard case of Pitman Drift, $\theta = \theta_0 + \eta/\sqrt{T}$, the statistics will converge in distribution to a non-central chi-square with noncentrality parameter

$$\delta^2 = \eta' H'(HVH')^{-1}H\eta$$

For a given finite sample size T, and a given alternative θ, the power is approximated from a non-central chi square with noncentrality parameter

$$\delta_T^2 = Th(\theta_T)'(HVH')^{-1}h(\theta_T).$$

Small sample corrections. When calculating the statistics, we do the standard corrections for sample size by replacing the number of observations T with $T - r$, where r is the number of restrictions in the hypotheses to be tested.

B.3 The Ghysels and Hall statistics.

We next consider the test statistics proposed by Ghysels and Hall (1990). Their method is essentially to test the fit the estimated parameters from one period in the moment conditions of the other period. This is very similar to the standard GMM test of overidentifying restrictions. For $i = 1$ to 2, define

$\hat{\theta}_i$ as the GMM estimator of θ, where only the data in X_i is used to do the estimation. Their test statistic GH is calculated as

$$GH_1 = T_2 g_2(\hat{\theta}_1)' \hat{V}_2^{-1} g_2(\hat{\theta}_1)$$

where

$$g_2(\theta) = \frac{1}{T_2} \sum_{t=T_1+1}^{T_1+T_2} f(x_t, \theta)$$

$$\hat{V}_2 = \hat{S}_2 + \frac{T_2}{T_1} \tilde{M}_2 (\hat{M}_1' \hat{S}_1^{-1} \hat{M}_1)^{-1} \tilde{M}_2'$$

$$M_i(\theta) = E\left[\frac{\partial}{\partial \theta} f(x_t, \theta)\right] \quad x_t \in X_i$$

$$\tilde{M}_2(\theta) = E\left[\frac{\partial}{\partial \theta} f(x_t, \theta)\bigg|_{\theta=\theta_1}\right] \quad x_t \in X_2$$

$$S_i = \sum_t \sum_s E[f_i(x_t, \theta) f_i(x_{t+s}, \theta)'] \quad x_t, x_{s+t} \in X_i$$

The parameters estimated using data from the first period is used in the moment conditions for the second period. Under the null hypothesis, GH_1 converges in distribution to a chi-square distribution with p degrees of freedom, where p is the number of orthogonality conditions in $f(\cdot)$.

Since under the null, the variance covariance matrix is stationary, we can replace the estimate $\hat{S}_2$ with an estimate $\hat{S}$ of S calculated from all the data.

$$S = \sum_t \sum_s E[f_i(x_t, \theta) f_i(x_{t+s}, \theta)'] \quad x_t, x_{s+t} \in X$$

This estimated S is then used in the calculations instead of S_2.

$$\hat{V} = \hat{S} + \frac{T_2}{T_1} \tilde{M}_2 (\hat{M}_1' \hat{S}^{-1} \hat{M}_1)^{-1} \tilde{M}_2'$$

which gives the second GH statistic

$$GH_2 = T_2 g_2(\hat{\theta}_1)' \hat{V}^{-1} g_2(\hat{\theta}_1)$$

Power calculations. Similarly to Andrews and Fair, we can do power calculations based on the noncentral chi-square distribution. The noncentrality parameter is

$$\delta_2 = \eta' \tilde{M}_2' V_2^{-1} \tilde{M}_2 \eta,$$

where η is the Pitman drift term.

B.4 Stability test with unknown breakpoint.

These tests are described fully in Andrews (1990). To define the tests, we index the time of possible parameter change by $\pi \in [0, 1]$. For example, $W_T(\pi)$ is the Wald statistic W described in the previous section with a possible breakpoint $T_1 = \lfloor \pi T \rfloor$.

The test statistics proposed by Andrews are

$$\sup_{\pi \in \Pi} W_T(\pi), \quad \sup_{\pi \in \Pi} LM_T(\pi) \text{ and } \sup_{\pi \in \Pi} LR_T(\pi),$$

where Π is a strict subset of $[0, 1]$. All these statistics converges in distribution to $\sup_{\pi \in \Pi} Q_p(\pi)$, where p is the degrees of freedom, and $Q_p(\cdot)$ is the "square of a standardised tied-down Bessel process of order p" Andrews (1990), page 32.

There is not a closed form solution for this probability distribution, so using Monte Carlo approximations, Andrews provides critical values for $\sup Q_p(\pi)$ with a choice of $\Pi = [0.15, 0.85]$. In our work, we use this value of Π, and rely on the critical values provided by Andrews. Table 12 below replicates the relevant critical values.

Table 12 Critical values for the statistics.

Degrees of Freedom	Significance level 1%	2.5%	5%	10%
1	12.3	10.3	8.7	7.2
2	15.3	13.4	11.7	10.2
3	18.3	13.4	11.7	12.3

Source: Andrews (1990)

References

Donald W. K. Andrews. Tests for parameter instability and structural change with unknown change point. Cowles Foundation, Yale University, April 1990.

Donald W K Andrews. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59:817–858, 1991.

Donald W K Andrews and Ray C Fair. Inference in nonlinear econometric models with structural change. *Review of Economic Studies*, 55:615–640, 1988.

Tim Bollerslev, Ray Y Chou, and Kenneth F Kroner. ARCH modelling in finance. A review of the theory and empirical evidence. *Journal of Econometrics*, 52: 5–59, 1992.

Jennifer Conrad. The price effect of option introduction. *Journal of Finance*, 44:487–498, 1989.

Jennifer Conrad, Mustafa N Gultekin, and Gautam Kaul. Assymmetric predictability of conditional variances. *Review of Financial Studies*, 4:597–622, 1991.

Marcel G. Dagenais and Jean-Marie Dufour. Invariance, nonlinear modelling and asymptotic tests. *Econometrica*, 59:1601–17, 1991.

Aswath Damodaran and Marti G Subrahmanyam. The effects of derivative securities on the markets for the underlying assets in the United States. *Financial Markets, Institutions and Instruments*, 1(5):1–22, 1992.

Martin S Eichenbaum, Lars Peter Hansen, and Kenneth J Singleton. A time series analysis of representative agent models of consumption and leisure choice under uncertainty. *Quarterly Journal of Economics*, 102:58–78, 1988.

Robert F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50:987–1007, 1982.

Eugene F. Fama. The behavior of stock market prices. *Journal of Business*, 38:34–105, 1965.

Kenneth A Froot and André F Perold. New trading practices and short-run market efficiency. Working paper, NBER, October 1990.

Ronald A Gallant and Halbert White. *A unified theory of estimation and inference for nonlinear dynamic models*. Basil Blackwell, 1988.

E Ghysels and A Hall. A test for structural stability of Euler conditions parameters estimated via the Generalized Method of Moments estimator. *International Economic Review*, 31:355–364, 1990.

Peter Hackl, editor. *Statistical Analysis and Forecasting of Economic Structural Change*. Springer–Verlag, 1989.

Lars Peter Hansen. Large sample properties of Generalized Method of Moments estimators. *Econometrica*, 50:1029–1054, July 1982.

Lawrence Harris. S&P 500 cash stock price volatilities. *Journal of Finance*, 44:1155–1175, December 1989.

Walter Kramer. The robustness of the Chow test to autocorrelation among disturbances. In Hackl (1989), pages 46–52.

Andrew W Lo and A Craig MacKinlay. When are contrarian profits due to stock market overreaction? *Review of Financial Studies*, 3:175–206, 1990.

Andrew W Lo and A Craig MacKinlay. Stock market returns do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies*, 1:41–60, 1988.

Benoit Mandelbrot. The valuation of certain speculative prices. *Journal of Business*, 36:394–419, 1963.

Whitney K Newey and Kenneth D West. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55: 703–708, 1987.

Robert W Rich, Jennie Raymond, and J S Butler. Generalized instrumental variables estimation of autoregressive heteroskedastic models. *Economics Letters*, 35: 179–185, 1991.

Douglas J Skinner. Options markets and stock return volatility. *Journal of Financial Economics*, 23:61–78, 1989.

Avanidhar Subrahmanyam. A theory of trading in stock index futures. *Review of Financial Studies*, 4:17–52, 1991.