WORKING PAPERS

John R. Cable
Adelina Gschwandtner

On Modelling the Persistence of Profits in the Long Run:
An Analysis of 156 US Companies, 1950-1999

Juli 2004

Working Paper No: 0409

DEPARTMENT OF ECONOMICS

UNIVERSITY OF VIENNA

All our working papers are available at: http://mailbox.univie.ac.at/papers.econ

On Modelling the Persistence of Profits in the Long Run:

An Analysis of 156 US Companies, 1950-1999*

by

John R. Cable[a] and Adelina Geschwandtner[b]

[a]*University of Wales, Aberystwyth*

[b]*University of Vienna*

Abstract

Long run persistence in company profits is analyzed for 156 US companies over a fifty-year period using AR1 and structural time series tests. A statistically significant degree of consistency is found between them in identifying firms persistently above or below the competitive norm. However, the structural time series method detects a higher overall incidence of persistence, with nearly 70% of firms classed as not having converged on zero, compared with 46% under AR1 estimation. The recently proposed structural approach is seen as a useful additional tool in analysing earnings dynamics, in particular where there are complex trends and other dynamic complexities.

Keywords: Profit Persistence; Competition; Structural Time Series

JEL Classifications: L12, C32

---

* The authors would like to thank Jesus Crespo Cuaresma and Dennis C. Mueller for very helpful comments and ideas.
a) E-Mail: joc@aber.ac.uk
b) E-Mail: adelina.gschwandtner@univie.ac.at

1. Introduction

That competition should equalise the returns to all economic activities in the long run is clear, but theory is silent over the time frame within which this should happen in practice. Time series data on the profits of companies engaged upon diverse activities should in principle shed light on this issue and hence on competitiveness and dynamic efficiency. The problem for researchers and practitioners is to determine where, and to what degree, observed differences in returns amongst firms persist 'unduly'.

In a now long empirical literature on the persistence of profits deriving from Mueller (1977, 1986), easily the most widely used approach is by reference to the parameters of first-order autoregressive models of earnings dynamics and, in particular, the company-specific long run projected profit rates (*LRPP*) they imply.[1] On this approach, 'persistence' occurs when the *LRPP* deviates from the competitive norm in a statistically significant way. However, the *AR1* is a restrictive framework and in parallel work Cable and Jackson (2003) propose a more flexible alternative. This utilises structural time series analysis, developed by Harvey (1989, 1997) and others, in order to decompose the overall earnings series into their unobserved long term trend, cyclical, autoregressive and irregular components.[2] Long run persistence is again identified as when, in this case, the estimated long run trend, disentangled from other, potentially confounding components of the series, deviates significantly from a competitive benchmark.

Thus we have two methods, their respective tests for long run persistence offering a point of comparison. This paper applies both to data for a sample of 156 US companies over the period 1950-1999, and carries out the comparison. So far as we are aware, ours is the first such analysis. It benefits from a much longer observation period than has typically been used in previous persistence-of-profits studies, which on average use no more than twenty annual observations, occasionally less. We investigate the incidence of profit deviations from the competitive norm that each model detects, and the extent to which their verdicts agree on an individual firm basis. Where they differ, we explore the reasons. The underlying question, from both policy and research perspectives, is whether

---

[1] Though Mueller (1977) also employed a Markovian framework and polynomial time trend models. See Mueller (1990) for a collection of early studies. Latest additions to the literature include Goddard and Wilson, 1999; Glen, Lee and Singh, 2002, 2003; Maruyama and Odagiri, 2002; Gschwandtner, 2001; and Yurtoglu, 2004. In much of the extant literature, the *AR1* speed of adjustment parameter is also equated with persistence, in this case 'short-run', and interpreted as capturing the intensity of dynamic competition in eroding excess profits. For reasons outlined below we focus only on long run persistence.

[2] No seasonal component is present since the data are annual.

the two approaches are substitutes or complements: whether in any given context it is sufficient to rely on one or the other; whether the choice depends on the particular context or policy / research question at issue; and whether there is synergy in their joint use.

We outline the two test procedures in the next section, commenting briefly on their properties. The empirical results are reported in section 3, and our conclusions follow in section 4. Appendix 1 describes our sample, data sources and the definitions of variables employed, Appendix 2 gives brief technical details of structural time series analysis, and Appendix 3 outlines some procedures used in classifying firms to persistence categories.

1. The AR1 and Structural Time Series Frameworks

As in the previous literature, our focal variable is a measure of ‘excess’, or mean-adjusted profit. Thus $\pi_{i,t} = (\Pi_{it} - \overline{\Pi}_t) / \overline{\Pi}_t$ is the relative deviation of firm i’s profit at time t from the sample mean $\overline{\Pi}_t$. Normalisation by the mean serves two ends. First, it removes the impact of macroeconomic cycles (though, as we shall see, the adjusted series can exhibit residual firm-specific cyclical patterns, and not infrequently do so). Second, following standard practice in the literature and taking the sample mean as a proxy for normal profit, we can interpret $\pi_{i,t}$ as deviations from the competitive norm, with attendant welfare implications.

1.1 AR1

Following Mueller (1986) the dynamics of $\pi_{i,t}$ are modelled as an autoregressive process of first order (*AR1*) given by:[3]

$$\pi_{i,t} = \alpha_i + \lambda_i \pi_{i,t-1} + \varepsilon_{it} \qquad (1)$$

Stability and convergence upon a finite steady state require $\lambda_i \in (-1,1)$, and ε_{it} is a white noise error process with constant variance. The unconditional expectation of $\pi_{i,t}$ in (1) is

[3] Under a widely accepted latent variable interpretation, (1) is regarded as the reduced form of two-equation system where profits are assumed to depend on the threat of entry in the market, and the threat is itself assumed to depend on the profits observed in the last period (Geroski and Jaquemain, 1988; Geroski, 1990). For critiques of this interpretation see Cable and Jackson, 2003, and Cable and Mueller, 2004.

given by $\alpha_i /(1-\lambda_i)$ and is a measure of 'permanent rents', which are not eroded by competitive forces. It is also referred to as the long run projected profit rate (*LRPP*), being the steady-state equilibrium value to which the series is tending asymptotically.

If all firms earn the competitive rate of return, then the $LRPP_i$ should be everywhere equal. The empirical literature of profit persistence therefore usually compares the estimates of the unconditional expectations from (1) (or, occasionally, alternative *AR(p), p>1,* generalizations) and tests the equality of these long run projections of the series across companies. It also tests for deviations of the long-run projections from zero. Since $LRPP = 0$ implies a long run projected return on assets equal to the norm, the percentage of projections significantly different from zero in a given sample is an indicator of the degree of competitiveness within it.

The *AR1* speed of adjustment parameter λ_i, the inverse of which shows how quickly $\pi_{i,t}$ converges to its long run level, is also treated as a measure of persistence in the literature, sometimes designated 'short run persistence'. However, since no direct counterpart is available under the structural time series approach, in this study we focus on the long run projected profit rate $\alpha_i /(1-\lambda_i)$, for which such a counterpart test does exist.[4]

The *AR1* is a tractable model, consuming only two degrees of freedom.[5] It handles neatly for persistence of profits purposes. However it says nothing about other features and properties of the time series, upon which it imposes an inflexible dynamic structure.[6]

2.2 Structural Time Series Analysis

Structural time series analysis (*STS*) decomposes the overall time series under investigation into their unobserved trend, autoregressive, cyclical, irregular and (where

[4] The autoregressive component coefficient in structural time series analysis (see Appendix 1) is not directly comparable with the AR1 regression parameter.

[5] Though the latter advantage may not be all that it seems in that, as we see below, long time series of earnings can exhibit structural breaks and even trend reversals such that estimation over shorter periods - when conserving degrees of freedom would matter - is in any case hazardous, and therefore to be avoided.

[6] Higher order autoregressions naturally afford greater flexibility; for example *AR2* processes (and above) can accommodate cycles. Where reported in the persistence literature, however, the results from such processes are usually either rejected formally or adjudged not to yield much additional explanation (see e.g. Geroski and Jacquemain, 1988; Gschwandtner, 2001; Glen *et. al*. 2002; and Cable, Jackson and Rhys, 2003).

relevant) seasonal components. Implementation is with the aid of the Structural Time Series Analyser, Modeller and Predictor (*Stamp*) due to Koopman, Harvey, Doornik and Shephard (1999). In estimation each one of the components of the series is assumed to follow a random process, the time-varying nature of which is governed by an associated hyperparameter. Maximum likelihood estimates are computed of the variances of each component. After estimation a Kalman filter is run to estimate the state μ_t, $t = 1,2,\ldots,T$, and the final state vector μ_T is reported. Fuller details of the estimation procedure and output are given in Appendix 2.

Using the structural time series approach, Cable and Jackson (2003) develop a comprehensive, 3 x 3 taxonomy of persistence of profits categories based on the level and slope of the long run trend μ_t in the final vector state T. Since 'slope' is constrained asymptotically to zero in the long run profit rate projected under *AR1* estimation, with which we wish to make comparisons, the slope subcategories of this taxonomy are suppressed here. Thus for present purposes we focus simply on whether the level of the long run trend μ_t is significantly greater or smaller than zero. That is, we test the null hypothesis H_0: $\mu_T = 0$, where μ_T denotes the trend in the final vector state T. This is easily done by comparing the estimate of μ_T with its RMSE (Harvey, 2001). The correspondence of this test with that of H_0: $a_i / (1-\lambda_i) = 0$ under *AR(1)* estimation is clear. However, there is a potentially important difference in that, whereas the *STS*-based test relates to where the trend actually stands within (strictly, at the end of) the observation period, the *AR(1)* test is for an eventual, implied steady state, which is out of sample and, in a strict sense, hypothetical. We return to this point when discussing divergences in the results from the two tests in the next section.

The structural approach offers a rich analysis of the time series under investigation in terms of its classical, structural components. It permits the long run trend to adopt a range of alternative, flexible forms; estimates it taking account of cyclical and autoregressive processes; and thereby permits its separation from such otherwise confounding, short run movements. On the debit side, it uses more degrees of freedom and is a more complex procedure than the *AR1*. However, it enjoys the advantage that stationarity of the series under investigation is irrelevant, and implementation with the aid of *Stamp* is straightforward.

3. Empirical Results

AR1 estimates for our 156 companies over the period 1950-99 were obtained using *SAS*, and the corresponding structural time series estimates with the aid of *Stamp*. As previously noted, the purpose was in each case to classify firms according to whether the (*AR1*) long run projected profit rate (*LRPP*), or the level of the trend in excess profits in the final vector state (*STS*), was significantly above or below zero.

Specification and estimation of the *AR1* model was done automatically. Descriptive statistics for $\pi_{i,t}$ and for *LRPP* for the whole sample are given in Tables 1 and 2. The interval (-0.24- 0.00) has the largest number of observations, revealing a stronger tendency for long run projected profit rates below the norm than above. [7]

Though it is possible to estimate the given equations also in *Stamp* automatically, structural time series analysis lends itself better to interactive mode, where the specification of cyclical and autoregressive components, as well as interventions for outliers and structural breaks, can be tailored to individual time series in order to arrive at a 'best' model. This was however impractical with 156 individual series to handle, and so a 'robustness' approach was adopted. Thus, using the batch mode, four models were estimated for each firm, all based on the 'smooth trend' model (see Appendix 2). The first, most general model (1), allowed for up to three cycles of differing period (with initialising periods of 5, 12 and 20 years respectively) plus an autoregressive and an irregular component. [8] Models (2) and (3) suppressed the autoregressive and the cyclical components respectively, and model (4) deleted both. We then looked for consistency of the classifications we sought across the four models. Fully consistent rankings were obtained in 63 instances, and a further 53 were resolved by applying relatively innocuous procedural rules. The remaining 40 cases required individual attention.[9]

The outcomes from the *AR1* and structural time series classifications are summarised in Table 3, in which the matrix elements a_{ij} show the number of firms in *STS* category *i* falling into the *AR1* category *j*, with $i,j = 1,2,3$. Summing the leading diagonal,

[7] Also the number of LRPP that are negative (86) is higher than the number of LRPP that are positive (70).
[8] In a given company's series, there may be more than one cycle superimposed on another, and some of these cycles could have long time periods. For example, there could be fairly short period cycles emanating from eg marketing strategies with existing products (or CEO cycle effects – which are usually short term), longer-period cycles relating to product life cycles and, possibly, very long term industry life cycles (eg long term substitution of steel by plastics, etc).
[9] See Appendix 3 for details.

we see that the two methods classify firms in the same way in 94 of the 156 cases; that is, the classifications match around 60% of the time. Whether consistency of this order is good or bad is not self-evident. One test might be with reference to differences between the group-mean values of the *AR1* long run projected profit rate as between the groups identified by *STS* as significantly above and below the norm, or not significantly different from it. If the group means were not significantly different, the degree of inconsistency would clearly be unacceptable. In fact the mean values in question are 0.630, -0.142 and -0.667 respectively, and as Table 4 records, their differences are all significantly different at better than one per cent (actually, with the probability that the means are equal smaller than one hundredth of one per cent).[10]

Measures of inter-rater agreement, as used in the social sciences (most commonly when dealing with data representing evaluations by two raters on the same individuals) offer a sharper test of congruity between the classifications. When the data are in discrete form, as in the present case, the Kappa test is an appropriate choice. Applied to the data in table 3, and treating *AR1* and *Stamp* as the raters, and the firms as the indivuals, this test yields a simple coefficient of 0.41. By convention this indicates 'moderate agreement', but the coefficient is significantly different from zero meaning that the hypothesis that the two classifications disagree (kappa coefficient=0) could be rejected at 1% level of significance or better.[11]

Notwithstanding the statistically significant consistency between the classification systems, it remains the case that they differ in around 40% of our 156 cases, and analysis of these differences is instructive. Overall, as Table 5 shows, *STS* finds more persistence of profits than does *AR1*, with 55 cases (35.3%) significantly above the norm, and 53 (34.0%) below it, as compared with 39 (25.0%) above, and 34 (21.8%) below in the case of *AR1*. Put the other way round, 108 firms (nearly 70%) are deemed not

[10] In order to assess whether there are significant differences between the AR1 and STS measures of profit persistence we estimated the following equation:

$$LRPP = \alpha + \beta\mu + \varepsilon,$$

where LRPP= long run projected profit rate of the AR1; μ=trend in final vector state of STS

for the whole sample of 156 observations and tested if $\alpha = 0$ and $\beta=1$ jointly. The corresponding F-Test rejects strongly this hypothesis meaning that the two measures differ systematically from one another.

[11] With the data in matrix form, complete agreement ($\kappa = 1$) occurs when all the off-diagonal counts are zero, and agreement by chance only ($\kappa = 0$), which is the null hypothesis, when the distribution is random. By convention a value of 0.4 is seen as indicating moderate agreement, and a value above 0.8 indicates very high agreement. For further details see Fleiss (1981) Ch. 13.

converged on zero by *STS*, compared with 73 firms (46.8%) not converging on zero according to *AR1* estimates. [12]

A possible explanation of this difference could be that the *STS* classification is based on a within-sample test – where the final vector state is – whereas the *AR1* long-run predicted rate is an out-of-sample projection, based on where the series is eventually heading.[13] Depending on the parameter values, the *AR1* process may have more adjustment to make towards steady state beyond the observation period in some cases than others, and in principle at least this might account for some of the *AR1 vs STS* differences in classification. That is to say, if the differences were, on average, significantly greater for the sub-sample where the two classifications disagree, we could claim to have identified at least part of the explanation of the classification differences. To verify this, we test for differences between the latest *AR1* within-sample predicted value (i.e. 'fitted' π_i for 1999) and the long-run, steady state value implied by the estimate of $\alpha / (1-\lambda)$, and look in particular to see if the differences are larger for cases where the classifications disagree than they are where they match**.** In absolute terms, we find that the mean *LRPP*/1999 difference for the disagreed cases is numerically larger than that for the matching categories subgroup by a factor of more than ten (0.103 as opposed to – 0.094), and the difference between these subgroup means is significant at the 5% level. However, taken separately, neither mean is significantly different from zero. Thus while there is some suggestion of an effect here, it is unlikely to provide the major explanation.

Another possibility is that the more flexible *STS* approach may simply deliver better fit, and therefore fewer non-significant cases. Because of differences in estimation method, the respective goodness-of-fit statistics are not directly equivalent. However, relative predictive power is comparable. To make the comparisons we hold back two observations in estimation, and then compute the relevant prediction errors for these observations and apply the Chow test.[14] Neither model predicts well overall.[15] However, *STS* outperforms *AR1* across the full sample at all conventional (cumulative) significance levels (Table 5), and particularly at 1% or higher, the latter no doubt contributing heavily

[12] Note that the significance of the LRPP is influenced by the standard errors of two estimated parameters.
[13] The AR1 long-run projected profit rate can also be interpreted as an in sample concept since it is the long run average of the time series.
[14] For this exercise we used STS model (1) in all cases, without individual level or slope interventions.
[15] In large part, we surmise, due to the sharply increased volatility of many of the individual series after about 1980, as noted earlier, and as illustrated in the *American Home Products Corporation* graphics (Figures 1 and 2). We can infer that STS at least should outperform a random walk, in that the reported coefficients of determination, R^2_D , are never negative, as they would be, given the way they are calculated, if the model was doing worse than a random walk with drift.

to a large difference in the average χ^2 statistics for the two methods (14.01 for *STS* and 44.81 for *AR1*), which is itself significant at better than 1%. Partitioning the sample according the whether or not the *STS* and *AR1* categorisations of persistence agree yields somewhat mixed results (Table 5, columns 2 and 3). *STS* still outperforms *AR1* for both subgroups at the 5 and 1% levels, but is marginally worse at 10% for the matching subgroup. At 5% the *STS* outperformance ratio is greater in the non-matching than the matching set (30/37 as against 45/50), but at the reverse is true at 1% (24/33 versus 31/44). Thus there is some indication that *STS*'s better overall predictive performance originates primarily in the subgroup where the classifications do not match, but this is sensitive to the choice of significance level.

The point emerges more clearly when we compare predictions at the individual firm level (Table 6). The dichotomous variable Stampbest 2 takes a value of 1 when the significance level at which prediction failure occurs in a particular case – or would occur if within conventional acceptance bounds – is lower for *STS* than it is for *AR1*, and zero otherwise. The variable Match is equal to 1 when the *STS* and *AR1* persistence classifications are the same, and zero where they are not, thus partitioning the sample according to whether there is agreement or not. From Table 6 we see that *STS* gives the better prediction in 100 out of 156 cases overall, which is 64% of the time, and in the ratio 1.8 : 1. But in the non-matching subsample the figures are 45 to 17 in terms of cases, which is in the ratio 2.7 : 1, compared with 55 to 39, i.e. only 1.4 : 1, where the classifications agree.

Turning to the pattern of the 'discrepancies' between the two classification procedures (i.e. the off-diagonal cases in Table 3), we note that the large majority (48 cases in all) occur in elements a_{12} and a_{32}, i.e. where *STS* finds cases to be significantly above or below the norm, that are non-significant according to *AR1*. This is consistent with *STS's* greater propensity to detect the presence of persistence in general, as previously noted. However, the traffic is not all one-way: there are also 13 cases which are non-significant under *STS*, which *AR1* classifies as significantly greater or less than zero (elements a_{21} and a_{31}). Either way, such non-matching outcomes from the classification process seem reasonable, inasmuch as they are between adjacent categories, and in that sense are marginal. However, there remains one more extreme case, classified with significantly *positive* LRPP by *AR1*, but significantly *below* the norm according to *STS*.

Inspection of the relevant time series, depicted in Figure 1(a) with the *STS* time trend superimposed, reveals the source of the problem. Evidently, after forty years of gently rising excess profit (albeit with increasing volatility after 1980)[16], the company in question (*American Home Products Corporation*) suffered a sharp reversal of fortunes in the 1990s. This is picked up by *STS's* estimation procedure, in which the parameters evolve over time and more weight is allocated to recent observations when there is rapid evolution. As a result, in this case *STS* fits a non-linear trend, peaking in 1991 and with a relatively steep negative slope thereafter. The *AR1* model, on the other hand, weights all observations equally and, influenced by the majority of early and mid period observations, the long run trend structure is projected through the troubled 1990s, continuing to predict a significantly positive LRPP (Figure 1(b)).

On the evidence of this example, it appears that *STS* is more sensitive to the way that fortunes, and trends, can change, by comparison with the more rigid, long term structure that *AR1* imposes, particularly when these occur late in the series. To investigate this further, and to check for other factors which might be responsible, we carried out individual analyses of six further, randomly selected cases where the *AR1* and *STS* persistence classifications did not match, and a control group of nine where they did (a 10 per cent sampling fraction in each case).

Of the six non-matching cases, the *STS* classification itself was robust over all four models in four cases, and split 2x2 in the others. The *Potlach Corporation* and the *Thomas & Betts Corporation* (Figures 2 and 3) were two of those robustly classified by *STS*, as significantly below the norm and non-significant respectively. Under *AR1* estimation the corresponding classifications were 'non-significant', and significantly positive excess profits. In both cases *STS* fits nonlinear trends, again allowing for sub-periods when profits are rising and falling, especially late in the period as in the *American Home Products* case. As the graphics output reveals, strong cyclical effects are also detected.[17] Evidently there are complex dynamics here. Viewed overall, *STS's* 'non-significant' assessment for *Thomas & Betts* looks more plausible than the *AR1* projection of significant *positive* excess profits (which presumably arises because, as a mean-reverting process, the *AR1* is reflecting the fact that the series is consistently positive up

[16] Also exhibited by a significant number of other companies in our sample post 1980.

[17] Of the four seemingly regular cycles - one and three in the case of *Potlatch*, and one and two in *Thomas & Betts* – three test as stationary and deterministic, the exception being *Thomas & Betts'* cycle two which, interestingly, appears to be emerging with increasing amplitude and frequency. χ^2 values for the other three were 9.04, 14.07 and 18.28 respectively, indicating significance at around or better than 1% in all cases.

to 1993). However, if the *STS* trend is extrapolated, the firm is clearly heading for sub-normal profits and, barring some corrective action, eventual exit. Similarly, *STS's* detection of subnormal profits for *Potlach* at the end of the observation period looks preferable to the *AR1* 'non-significant' verdict. But again, if the sinuous trend fitted by STS is extrapolated, one would predict the firm to oscillate between periods of normal and significantly subnormal profits. In general, these are difficult cases, where long-run predictions are hard to pin down with confidence.

In the other two cases where the *STS* classifications are robust across all models, *Arvin Industries Inc*. (Figure 4) and the *Ferro Corporation* (Figure 5), *STS* fits linear trends and detects significantly positive and negative returns respectively, whereas the *AR1* classification is non-significant in both cases. The fact that the series, on both *STS's* interpretation and visual inspection, appear to be diverging upwards and downwards respectively (having earlier both intersected zero abnormal returns) tends to favour the *STS* interpretation. In the *Arvin Industries* case, it could be that the presence of strong cyclical components, of which *STS* takes account but *AR1* does not, may have inflated the variance of the *AR1* estimates.[18] Very tentatively, something of the sort may also have occurred in the case of *Ferro Corp*, via the 'emerging' cycle three, though in this case *STS* suppresses the second cycle and, as can happen, cycle one appears to have taken on the work of the irregular component (possibly also the AR). Otherwise, and apart from the *AR1'*s general tendency to produce more non-significant cases, no particular explanation for the disagreement suggests itself in this case.

Outlier problems loom large in the two remaining 'disagreed' cases, *Deere & Co* (Figure 6) and *NL Industries* (Figure 7), where the *STS* classification itself had earlier also proved sensitive to specification. In the *Deere & Co* case, *STS* had produced significantly negative profits when cycles were included (models 1 and 2), and non-significant profits otherwise (models 3 and 4). But as Figure 6 shows, the graphical evidence for cycles is not strong, and inclusion of interventions for outliers in observations 37 and 48 confirmed the 'non-significant' *STS* classification even when cycles were included.

In the *NL Industries* case there is no evidence of cycles, but the *STS* classification is significant negative profits when they are suppressed (models 3 and 4), and 'non-

[18] The graphics output for this case (Figure 4) includes two apparently regular cycles of period 5 and just over 7 years respectively, the first of which is actually stationary and deterministic at better than one per cent (χ^2 = 10.115), the second stochastic. Cycle three is perhaps too irregular for consideration, though might be seen as capturing some residual long-swing wavelike motion in the series.

significant' when they are not (models 1 and 2), which is also the *AR1* verdict. Including an outlier intervention for observation 37 (again) under *STS* estimation confirms the significantly negative result, even when the cyclical components are retained. However, if a slope intervention from observation 34 is included as well, as might appear justifiable from inspection, non-significance is again found. By inspection this looks the more plausible result, on this occasion favouring the *AR1* classification over that initially produced by *STS*.

Analysis of the control group of nine randomly selected cases where the *STS* and *AR1* classifications matched proved relatively straightforward, and we report only generic points of comparison, rather than individual details.[19] Little difference emerged in the degree of robustness of the *STS* classifications as between the matching and non-matching groups, four of six being robust over all models in the non-matching group, as previously noted, compared with five of nine in the matching group. The incidence of outliers was also similar.

Where differences did emerge between the groups was in respect of complex trends and, to a lesser degree, the incidence of cycles. In the matching group *STS* fitted simple, linear trends in all cases but one, and in this case the curvature was very slight, disappearing after intervention to correct for an outlier. In the non-matching group, as we have seen, there were complex trends in the initial estimates for two of the six cases, plus two others upon further investigation. A total of six cycles was found in both groups, four in each case testing as deterministic or stationary. However these were distributed over five of nine members of the matching group, compared with four of six in the non-matching group. Though verification in larger samples is necessary, the initial evidence points towards greater dynamic complexity in the cases where the *AR1* and *STS* classifications disagree than where they concur. The inference which then follows, at present tentatively, is that *STS*'s greater ability to detect and deal with the time series properties of earnings where trends change within the observation period, or cyclical factors intrude on the series, may be an important factor in causing the classifications themselves to differ.

2. Conclusions

The widely used *AR1* framework and the recently proposed structural time series approach offer researchers and practitioners alternative ways of testing for the persistence of profits in the long run, an important indicator of the efficacy of market systems, and

[19] Full details are available from the authors.

identifier of potential dominant firm abuse cases. The statistical tests for long-run persistence that they offer form a basis for comparison of the two approaches.

The *AR1* is simple, undemanding in terms of degrees of freedom, and its properties are well known both in the present context and from countless other applications. However it imposes a restrictive dynamic structure. Structural time series analysis is less familiar, more complex, requires longer time series, and is less amenable to batch estimation. However, it offers greater flexibility in terms of the form of the trends it fits and, importantly, enables the long run trend to be disentangled from potentially confounding cyclical, autoregressive and other short run elements. By decomposing the overall series into all structural components, and revealing their nature, it addresses a wider range of phenomena and hence also provides a richer analysis.

Empirically, applying both methods to fifty-year time series for 156 US companies, we found a statistically significant degree of consistency between them in identifying firms persistently above or below the competitive norm. Nevertheless, the two approaches differed in around 40% of cases albeit, with one exception, only as between adjacent classes. Overall, the structural time series analyser *Stamp* detected a higher incidence of persistence, with nearly 70% of firms classed as not having converged on zero, compared with just under half under *AR1* estimation.

STS outperformed *AR1* in comparisons of predictive power, both in terms of prediction failure rates at conventional significance levels and, more particularly, by nearly two to one on an individual basis across the sample as a whole, in terms of the significance level at which failure either actually occurred, or would have done so if within conventional acceptance bounds. Partitioning the sample according to whether *STS* and *AR1* classified firms in the same or different persistence categories revealed that *STS's* relative predictive superiority was markedly higher in the non-matching than in the matching subset, in the ratio 2.7:1 *vs* 1.4:1 (and 1.8:1 overall) while *AR1* marginally outperformed *STS* in the matching subset. Evidence from the random 10% of individual cases investigated in depth was suggestive that this might reflect *STS's* greater ability to deal with series exhibiting complex trends and other dynamic complexities such as cycles, the incidence of which appeared to be higher in the non-matching than the matching set.

We conclude that structural time series analysis adds usefully to the armoury available to practitioners and researchers when tackling persistence of profits issues, particularly where there are changes in trend, or when complicating factors such as cycles intrude. If used as an aid in screening for potential antitrust violations / dominant firm abuses, *STS* would, on the evidence of this study, yield a larger set of possible targets than would an *AR1* model. Moreover, though there would be a considerable overlap, the set of cases where the models disagree seems to throw up particularly complex and sometimes problematic cases, indicating that for this very reason there may be merit in their joint application.

Appendix 1: Sample, Data Sources and Definitions of Variables

The database contains yearly data on profits for 156 surviving companies from the period 1950-1999 and has the advantage that is more than twice as long as the average time period used in the literature (which is around 20 years). The sample corresponds to those among the largest 500 US manufacturing companies (in terms of sales) as of 1950 for which a complete time series on profits spanning the period 1950-1999 existed. The database was compiled using Compustat, Global Vantage (especially for the last years) and Moody's Industrial Manual (for missing data points).

Profit (returns on assets) is defined as net income over total assets, and throughout the study the profit rate of company *i* at time *t* ($\pi_{i,t}$) is defined as the relative deviation from the sample mean at time *t*.

The Compustat (and Global Vantage) variable name corresponding to the proxy for income is "Income before extraordinary items" and it represents the income of a company after all expenses, including special items, income taxes and minority interests, but before provisions for common and/or preferred dividends. Total assets include current assets plus net property, plant and equipment plus other noncurrent assets. Ideally, interest should have been added to income before dividing by total assets, in order to make the profit measure independent of the source of funds used to create total assets. However data for interest were not available especially for the beginning years (1950-1977). A sensitivity analysis has been done for the last period 1980-1999 when interest data were available and the results using interest were not significantly different from the ones without interest.

Appendix 2: Structural Time Series Models[20]

In the most general case available in *Stamp*, the *local linear trend model,* the trend μ_t for the series under investigation y_t comprises stochastic trend and irregular components:

$$y_t = \mu_t + \varepsilon_t, \quad t = 1,2, \ldots, T \qquad (6)$$

The trend is subject to shocks in both level and slope, so that

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t$$
$$\beta_t = \beta_{t-1} + \zeta_t$$

where η_t and ζ_t are normally and independently distributed with mean zero and variance σ^2, and the irregular, level and slope disturbances, ε_t , η_t and ζ_t, are mutually independent.

Setting $\sigma_\eta^2 = \sigma_\varsigma^2 = 0$ yields a *deterministic trend model,* in which both level and slope are non-stochastic, or 'fixed'. When either σ_η^2 or σ_ς^2 is individually zero, the trend becomes, respectively, a random walk with drift (fixed slope, stochastic level) or an integrated random walk, the *smooth trend model* (fixed level, stochastic slope).

Autoregressive and cyclical components are added via a serially correlated stationary component, ψ_t:

$$y_t = \mu_t + \psi_t + \varepsilon_t, \quad t = 1,2, \ldots, T \qquad (7)$$

When ψ_t is an autoregressive process, *Stamp* constrains it to be stationary to avoid it being confounded with the random walk in the trend. Alternatively, ψ_t may be specified as a stochastic cycle:

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} k_t \\ k_t^* \end{bmatrix}, t = 1,2,\ldots,T$$

Here, λ_c is the frequency in radians, in the range $0 < \lambda_c < \pi$; k_t and k_t^* are two mutually uncorrelated white noise disturbances with zero means and common variance σ_k^2; and ρ is a damping factor such that as $\rho \rightarrow 1$ the stochastic cycle reduces to a deterministic, but stationary cycle. Stamp reports significance (χ^2) statistics only where the cycle is deterministic and stationary. Up to three cycles of differing frequencies and the autoregressive component can be incorporated in the same model. However, if λ_c is *0* or π in the cyclical component, the stochastic cycle itself becomes an AR(1).

Stamp output reports convergence performance and also includes the estimated variances of the disturbances and their standard deviations; the estimated autoregressive coefficient; the estimated parameters of the cycles (including period, frequency and amplitude); and the estimated level, slope autoregressive and cycle parameters of the final vector state and their R.m.s.e's. Parameters for 'intervention' (to control for outliers and structural breaks) are also included where these are present, together with their significance levels. Graphics output includes that for components and for residuals.

Diagnostics include the log likelihood statistic and test, the Doornik-Hansen normality test, one heteroscedasticity and three autocorrelation tests (including the DW and Box-Ljung Q-statistics), together with the most appropriate of three alternative coefficients of determination. Of these, R^2_D compares the prediction error variance with the variance of first differences, and is the preferred measure where the series shows trend movements. It can be negative, indicating 'a worse fit than a simple random walk with drift' (Koopman *et al.*, 1999).

---

[20] This appendix draws extensively upon Koopman *et al* (1999) and follows their notation.

Appendix 3: Stamp Classification Procedures

Fully consistent rankings across the structural time series models (1) – (4) (see text) were obtained in 63 instances, and consistency across three (i.e. with just one discrepancy) in a further 52. Often, in the latter group, the outlier was model (4), frequently with low or very low explanatory power. 28 cases were split ‘two-by-two’ between classification categories (the pairings often being between the models with cycles, i.e. (1) and (2), and those without), and 13 cases between all three. Where discrepancies occurred, these never extended to non-adjacent categories (i.e. significantly above’ *vs* ‘significantly below’ the norm); and usually they arose from relatively fine differences, due e.g to fairly marginal effects of the inclusion or exclusion of cycles on r.m.s.e and hence significance levels.

53 of the 93 cases featuring discrepancies were resolved by applying procedural rules, e.g. relaxing significance acceptance levels from 5 to 10 per cent in order to bring one (or very occasionally two) models into line; and discounting model (4) where this was the sole discrepancy, especially if R^2_D was very low or, as in some cases, negative (implying that the model was doing worse than a random walk with drift). Thus nearly three quarters (74.4%) of our total sample was either wholly unambiguous or dealt with procedurally. The remaining 40 cases required individual attention in interactive mode. Amongst these, the appropriate category was often immediately obvious from inspection of the graphics output, and in many cases the discrepancies were clearly due to outliers and (less frequently) structural breaks, intervention for which readily resolved a final category. Ultimately, there were no cases which defied all attempts at classification.

References

Cable, J R. and Jackson, RHG, The Persistence of Profits in the Long Run: A New Approach, *SMBA Research Paper 2003-10*, University of Wales, Aberystwyth, 2003.

Cable, J R, Jackson, RHG. and Rhys, H, The Univariate Time Series Modelling of Earnings: Harmonic Variation, *SMBA Research Paper 2003*-09, University of Wales, Aberystwyth, 2003.

Cable, J R. and Mueller, D C, Testing for the Persistence of Profits' Differences Across Firms, (mimeo), 2004

Fleiss, J. L. , *Statistical Methods for Rates and Proportions.* New York: John Wiley & Sons 1981.

Geroski, P A, and Jaquemain, A, The Persistence of Profits: A European Comparison, *Economic Journal*, 98 (1988) 375-389

Geroski, P A, Modelling persistent profitability, in D C Mueller (ed) *The Dynamics of Company Profits*, Cambridge: Cambridge University Press, 1990

Glen, J., Lee, K. and Singh, A. Persistence of profitability and competition in emerging markets, *Economics Letters*, 72 (2002) 247-253

Glen, J., Lee, K. and Singh, A. Corporate Profitability and the Dynamics of Competition in Emerging Markets: A Time Series Analysis, *Economic Journal*, 113(2003) F456-F484

Goddard, J. A. and Wilson, J. O. S, The persistence of profit: A new empirical interpretation, *International Journal of Industrial Organization*, 17 (1999) 663-687

Gschwandtner, A, *Evolution of the Largest US Corporations 1950-2000*, PhD Thesis, University of Vienna (2001)

Harvey, A.C. *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge: Cambridge University Press, 1989

Harvey, A. Trends, cycles and autoregressions, *Economic Journal*, 107 (1997) 192-201.

Harvey, A. 'The Econometrics of Stability and Convergence', in J.R. Teixeira and F.G. Carneiro (eds) *Growth, Redistribution and Structural Change, Proceedings of the Third International Colloquium*, Universidade de Brasilia e Universidade Catolica de Brasilia, Editora Universa, Taguatinga DF, Brazil, 2001

Koopman, S.J., Harvey, A.C., Doornik, J.A. and Shephard, N. *Stamp: Structural Time Series Analyser, Modeller and Predictor*, London: Timberlake Consultants Ltd., 1999.

Maruyama, N and Odagiri, H, Does the Persistence of Profits Persist? A Study of Company Profits in Japan, 1964-97, *International Journal of Industrial Organisation,* 20 (2002) 1513-1533

Mueller, D C, The persistence of profits above the norm, *Economica*, 44 (1977) 369-380

Mueller, D C, *Profits in the long run*, Cambridge: Cambridge University Press, 1986

Mueller, D C (ed) *The Dynamics of Company Profits*, Cambridge: Cambridge University Press, 1990

Maruyama, N. and Odagiri, H. Does the 'persistence of profits persist?: a study of company profits in Japan, 1964-97, *International Journal of Industrial Organization,* 20 (2002) 1513-1533

Yurtoglu, B Persistence of Firm-level Profitability in Turkey, *Applied Economics*, 36 (2004) 1-11

Table 1: Descriptive statistics for $\pi_{i,t}$ and LRPP, 156 US Companies, 1950-1999

Variable	Mean	Median	Standard Dev.	N
$\pi_{i,t}$	0.0386	0.0117	1.2131	7800
LRPP	0.0395	-0.2922	0.6078	156

Table 2: Distribution of LRPP, 156 US Companies, 1950-1999

Range	# LRPP
> 0.75	0
0.50- 0.74	24
0.25- 0.49	17
0.00- 0.24	29
-0.24- 0.00	43
-0.49- -0.25	22
-0.74- -0.50	10
<-0.75	11
Total	156

Table 3: AR1 and STS Classifications of Long Run Persistence, 156 US Corporations 1950-1999

		AR1 (Long run projected profit rate) *Sig. > 0*	*N.S.*	*Sig. < 0*	**Total**
STS (Level of trend in final vector state)	*Sig. > 0*	33	22	0	55
	N.S.	5	35	8	48
	Sig. < 0	1	26	26	53
Total		39	83	34	156

Table 4: Differences in AR1 Long Run Projected Profit Rates between STS Persistence Categories[(i)]

		Level of Trend in *Sig.> 0*	**Final Vector State** *N.S.*
Level of Trend in	*N.S.*	0.737 (11.88)***	
Final Vector State	*Sig.< 0*	1.138 (12.58)***	0.346 (7.60)***

Note (i): Table entries are between-group differences; figures in parentheses are t statistics; *** denotes significance at better than one per cent.

Table 5: Prediction Failure Rates, AR1 and STS: Full Sample and by Matching versus Non-matching Subsamples(i)

	Significance Level (Cumulative)	Full Sample n = 165	Matching Categories n = 94	Non-matching Categories n = 62
AR1				
	1%	77 (49.4)	44 (46.8)	33 (53.2)
	5%	87 (55.8)	50 (53.2)	37 (59.7)
	10%	96 (61.6)	55 (58.5)	41 (66.2)
STS				
	1%	55 (35.3)	31 (33.0)	24 (38.7)
	5%	75 (48.1)	45 (47.9)	30 (48.4)
	10%	90 (57.7)	57 (60.7)	33 (53.2)

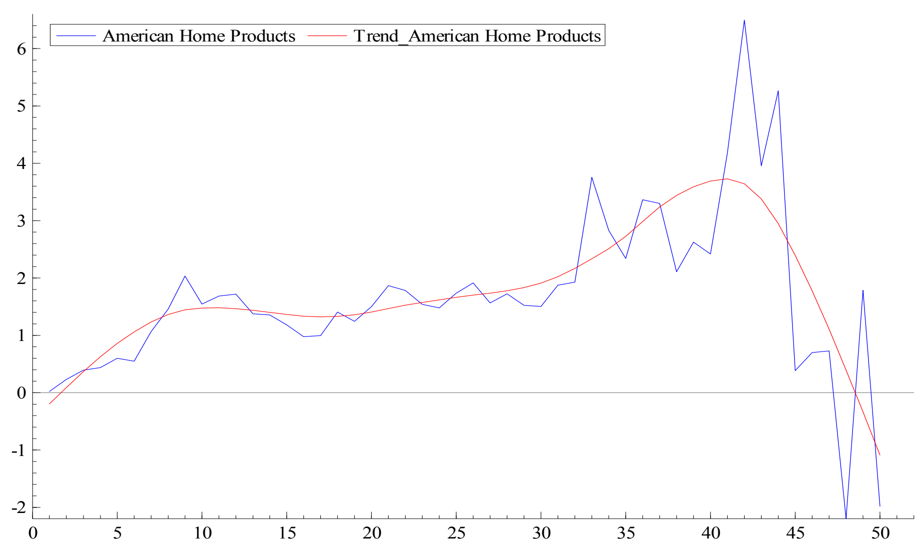
Note (i): Table entries are the number of prediction failures at the cumulative significance levels shown. Figures in parentheses are percentages of the respective samples.

Table 6: Relative Predictive Performance, AR1 and STS, by Matching and Non-matching Subsamples (i)

		Match 0	*Match* 1	**Totals**
Stampbest 2	0	17	39	56
	1	45	55	100
Totals		62	94	156

Note (i): *Stampbest 2* equals 1 if $\chi^2_{\ Stamp} < \chi^2_{\ AR1}$ and zero otherwise; *Match* equals 1 if the *AR1* and *STS* persistence classifications agree, and zero otherwise.

Figure 1(a): American Home Products (STS Estimation)[21]



[21] Only relevant graphs were included.

Figure 1(b): American Home Products (AR1 estimation)

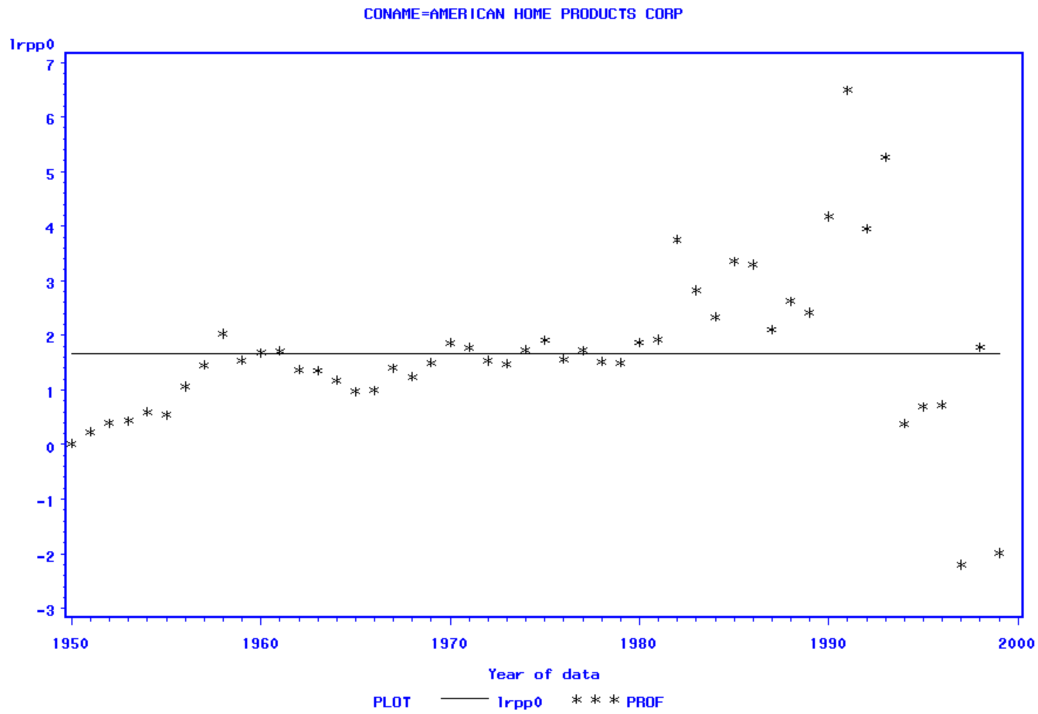


Figure 2: Potlach Corporation

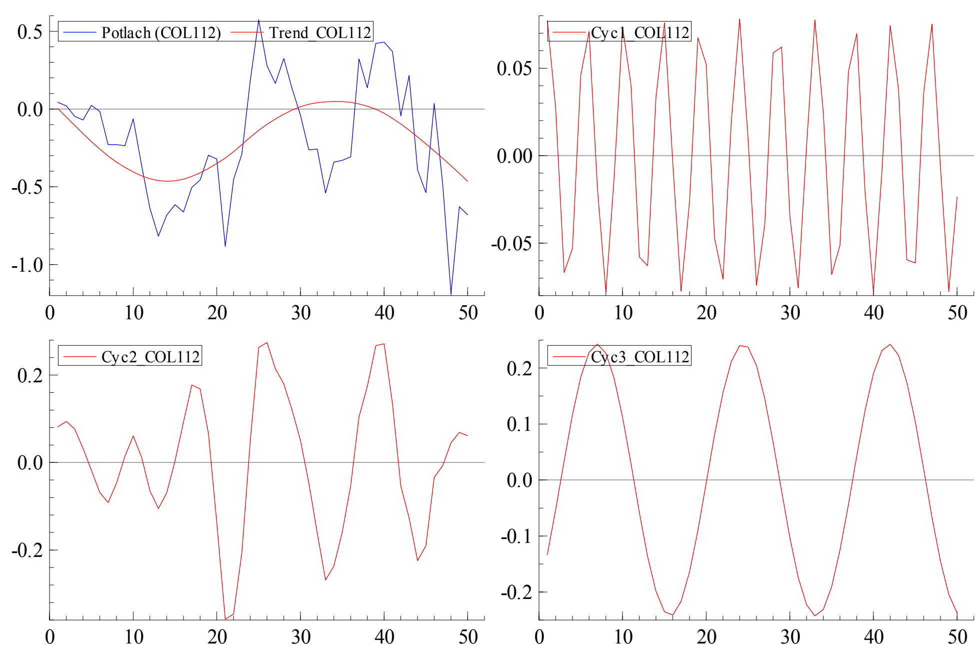


Figure 3: Thomas & Betts Corporation

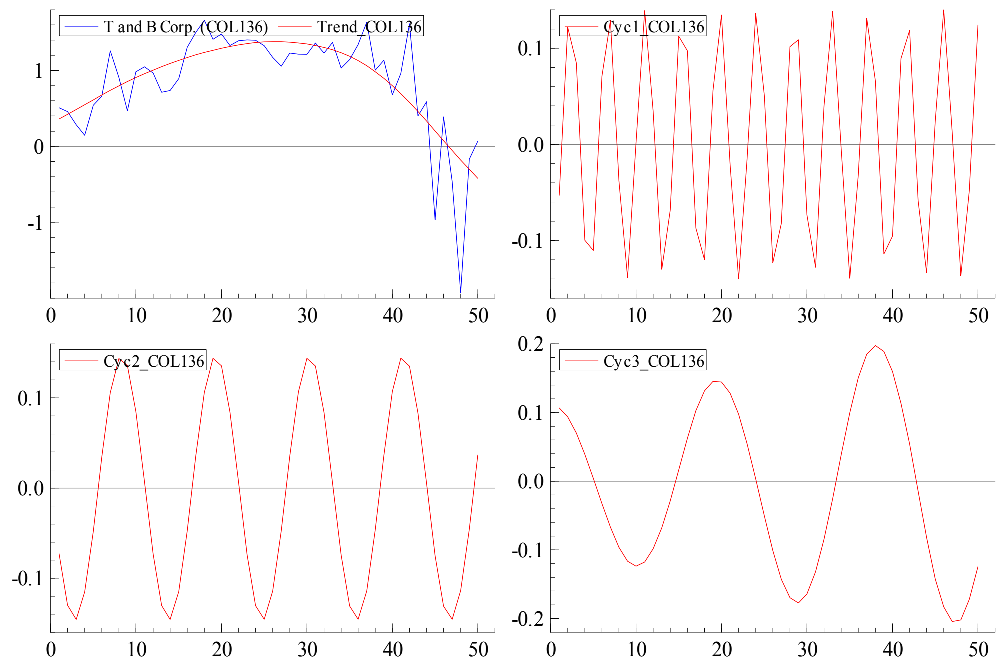


Figure 4: Arvin Industries Inc.

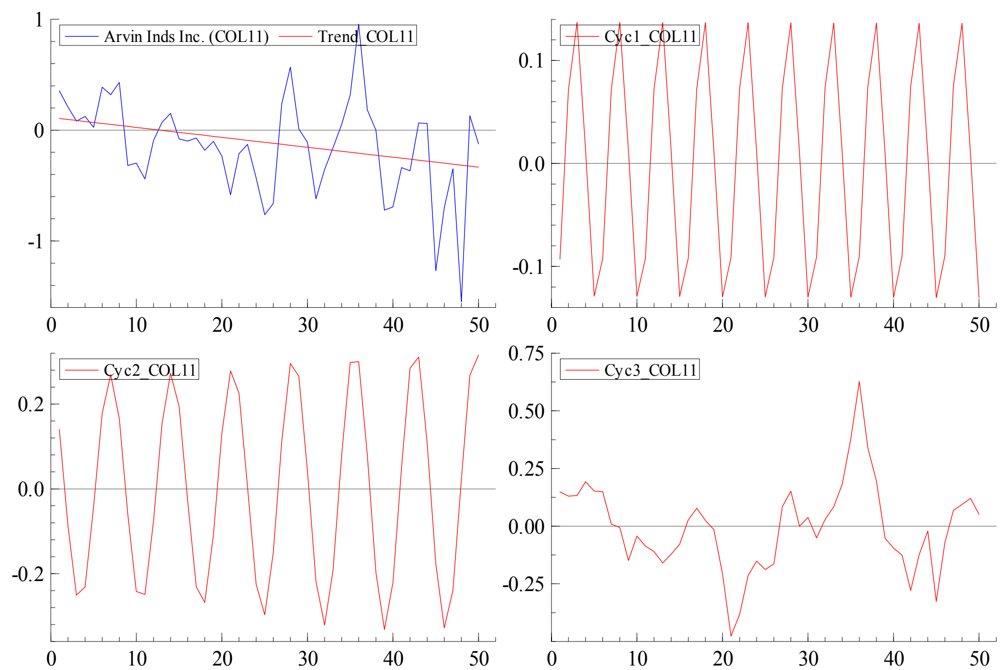


Figure 5: Ferro Corporation

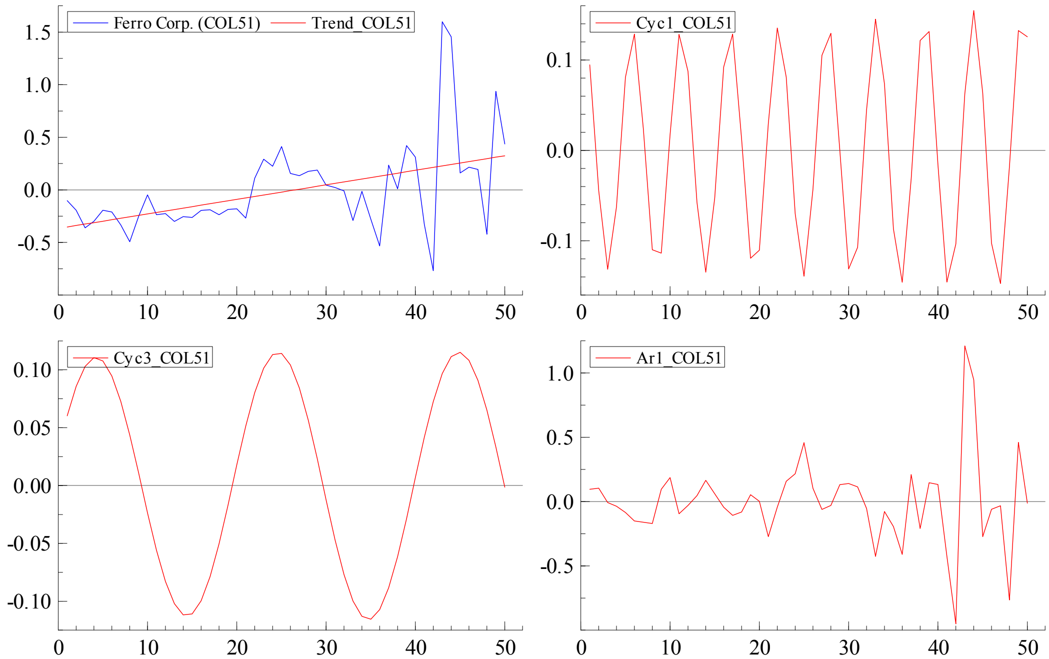


Figure 6: Deere and Co.

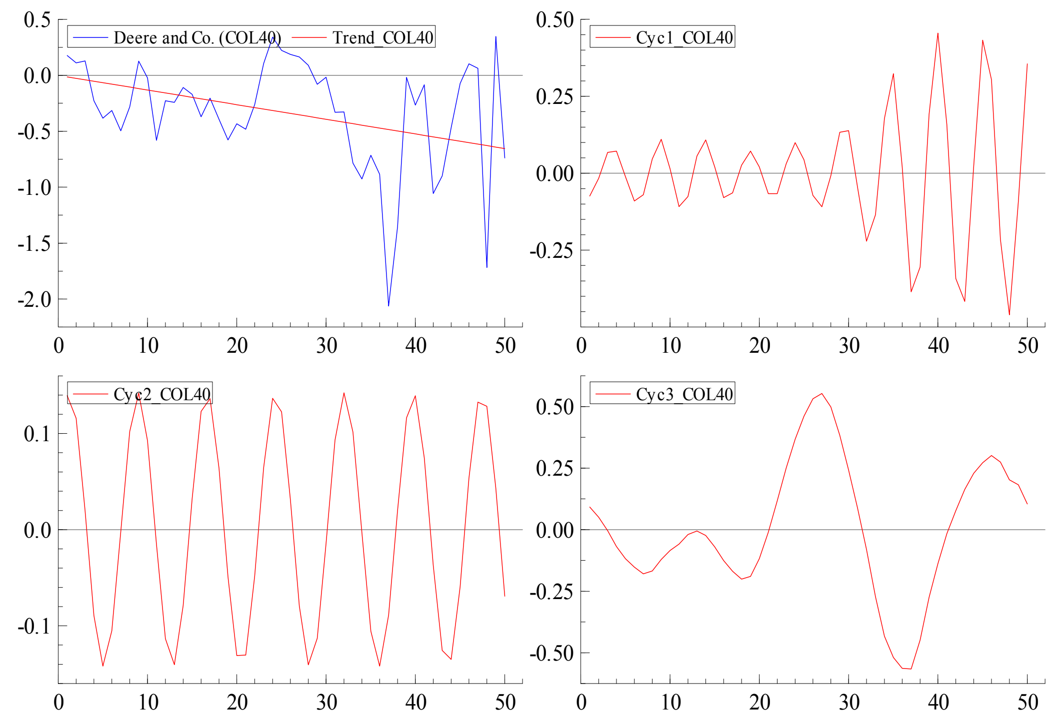


Figure7: NL Industries