On modeling panels of time series*

Philip Hans Franses†
Econometric Institute, Erasmus University Rotterdam

June 20, 2002

Econometric Institute Report 2002-23

Abstract

This paper reviews research issues in modeling panels of time series. Examples of this type of data are annually observed macroeconomic indicators for all countries in the world, daily returns on the individual stocks listed in the S&P500, and the sales records of all items in a retail store. A panel of time series usually concerns the case where the cross-section dimension and the time dimension are large. Usually, there is no a priori reason to select a few series or to aggregate the series over the cross-section dimension. In that case, however, the use of for example a vector autoregression or other types of multivariate systems becomes cumbersome. Panel models and associated estimation techniques are more useful. This paper discusses representation, estimation and inference in case the data have trends, seasonality, outliers, or nonlinearity. Various examples illustrate the various models.

*This paper has been prepared for presentation at the Royal Netherlands Academy of Arts and Science Colloquium on "Blending theory and practice and the evaluation of economic policy. A Conference dedicated to the memory of Henri Theil" (August 16-18, 2002, Amsterdam). I thank Merel van Diepen for research assistance, and Dick van Dijk and Richard Paap for helpful comments, and Teun Kloek for useful information on aspects of Henri Theil's work. All errors and omissions are my own responsibility. Parts of this paper have been presented at seminars in Singapore (2002), Tokyo (2002), Lisbon (2000) and Perth, Australia (2001). I thank the participants of these seminars for their useful suggestions.

†Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR, Rotterdam, The Netherlands, email: `franses@few.eur.nl`

1 Introduction and motivation

The work of Henri Theil covers many topics in economics, econometrics and forecasting. In various stages of his career, Theil devoted attention to systems of equations. Examples are the well-known demand system, the multinomial logit model, and of course the simultaneous equation model, for which Theil developed useful estimation routines. In those days, the empirical analysis usually concerned, say, two to ten variables, which were annually measured for, say, less than twenty years. An additional feature of the era in which Theil wrote his influential papers, was that economic theory had quite a strong grip on the practical work of applied econometricians.

At present, the world might look a little different for an applied econometrician. First, there is an abundance of available data on a range of economic variables, some of which can be observed at any aggregation level one wants. Naturally, this is perhaps not true for most macroeconomic variables, where the observation frequency went from years to only months, but for many variables in for example finance and marketing, data can nowadays be available at the individual transaction level, see also Granger (1998). Not only did the sampling frequency over time increase, it is also possible to obtain information at the individual level or in highly disaggregated sectors. For example, one can nowadays download US city-specific unemployment figures and industry-specific data for males in certain age categories, see **www.economagic.com**. And, in marketing one can keep track of daily household-specific purchase behavior.

The second change that can be noticed since the days of Theil, is that economic theory is now often seen as only partially useful to specify an econometric model. Indeed, the data play an important role, the choice of data and their transformations too, and also the statistical properties of the model. Next, if the data can be collected as such a highly disaggregated level, one may wonder how informative economic theories are, as these usually lead to predictions at a highly aggregate level.

My personal view on econometrics, which is laid out in more detail in Franses (2002a), is that econometric analysis nowadays is mainly driven by research questions which require an econometric model. These questions may be motivated by theory, but can also be based on practical matters. The main premise is that economic theory is not precise enough to establish the model, so the econometrician has substantial freedom. The second notion is that nowadays there are often no precise reasons why one type of data is better than another. For example, if one wants to say something about international economic growth, which variable does one use? If the focus is on international convergence, should one think of income levels or on health issues or schooling?

The research questions I have in mind, some of which will also be touched upon in the rest of this paper, are like the following. Are cyclical fluctuations in unemployment getting smaller?, Do African countries follow a world business cycle and do they grow as fast as Asian countries do?, Do emerging markets converge to mature markets (in terms of volatility)?, Do temperatures within a year increase with a similar speed?, Do quarterly industrial production series display the same seasonal fluctuations?, and Does display promotion lead to more sales in a retail store? These questions typically concern data that can be measured as panels of time series, with cross-section dimension N and time series dimension T. There are various unemployment series, there are more than twenty African countries, there are more than ten emerging markets, there are twelve monthly temperatures within a year, there are hundreds of industrial production series, and there are thousands of items sold in regular-sized retail stores, and these are regularly put on display.

One might now be inclined to analyze the data after cross-sectional aggregation, that is, aggregation over N or over subsamples of N. The problem with such aggregation, however, is that it is uncertain what its consequences are. What does aggregation do if the data are nonlinear, or if they are seasonal, have outliers or structural breaks? And, which features are preserved after aggregation? Also, how does one determine the weights of the individual series upon aggregation? Finally, why should one want to aggregate in the first place? Why not use all data available? Of course, many data have already been aggregated somehow, but in the end, one can try to minimize the overall error by looking at disaggregated data.

In sum, the type of data that can be studied nowadays are time series data for

25 to 250 periods where the cross-section dimension concerns 10 to 100 individual cases (people, firms, countries). This is in contrast with typical panel data where N is large like 5000 and T is small like 3 to 5, see Baltagi (1995) and Hsiao (1986). Hence, one nowadays can have data with large N and large T.

It is well known that the use of a Vector Autoregression [VAR] might not be sensible now. Indeed, VAR models are typically used for cases where T ranges from, say, 50 to 250, but where N is 4 or 5. Note that VAR models for large N and small T have been considered in Holtz-Eakin, Newey and Rosen (1988). The main reason is that a VAR includes lags of all other variables in each equation, thereby generating a wealth of parameters. This wealth comes at the cost of estimation precision. Additionally, the model parameters are not easy to interpret, and one usually has to resort to derivative measures like Impulse Response Functions [IRFs]. VAR models for large N also make statistical tests to have low power and incorrect empirical size, see for example Gonzalo and Pitarakis (1999).

In sum, one needs different models to answer the questions mentioned of the type above, and these can be summarized as models for panels of time series. An additional advantage of these models is that for some important econometric techniques, the commonly observed time series samples are too small for having much confidence in the statistical outcomes. An example concerns tests for unit roots. Analyzing panels allows one then to consider N series at the same time, and this increases statistical power, see also Papell (1997) and O'Connell (1998).

This paper aims to focus attention to the analysis of currently available data in economics, finance and marketing by models for panels of time series. In some sense these models are more restrictive, but their merits lie in interpretation and ease of use. In Section 2, I discuss the generic representation of models for panels of economic time series data. Next, I discuss parameter estimation. In Section 3, I focus on the extension of the basic models in case one has (economic) data with trends, seasonality, outliers and nonlinearity. Section 4 concludes with an outline of further research topics.

It must be stressed that sometimes this paper contains just ideas and loose notes. The topic is rather new, and not much research has been dedicated to it, although

I intend to refer to all the relevant ones. This paper should be seen as an invitation to a scientific discussion.

2 Representation and Estimation

This section deals with a generic representation of a model for a panel of time series. This representation will be modified in the next section in order to incorporate specific features of (economic) data. This section also discusses various ideas behind parameter estimation. It is not the intention to outline the technical details of these estimation routines, as some of these can be found in the relevant literature.

2.1 Representation

Consider observations $y_{i,t}$, where i runs from 1 to N and t runs from 1 to T. A model for a panel of time series looks like

$$y_{i,t} = \mu_{i,t} + u_{i,t}, \tag{1}$$

where $\mu_{i,t}$ can capture explanatory variables and also deterministic terms, such as an intercept, seasonal cycles and a trend, represented as, for example,

$$\mu_{i,t} = \mu_i + \alpha_{1,i}\cos(\pi t) + \alpha_{2,i}\cos(\frac{\pi t}{2}) + \alpha_{3,i}\cos(\frac{\pi(t-1)}{2}) + \delta_i t, \tag{2}$$

in case of quarterly data, where $u_{i,t}$ is an error term. This notation is slightly different from the one typically used, that is, with seasonal dummies, but for analyzing unit root properties it is more convenient, see Franses and Kunst (1999). The deterministic part can be extended to include quadratic trends, and also by sine and cosine functions for higher frequency data, that is, weekly data or the like.

For the error process $u_{i,t}$ it can be assumed that

$$\Phi_{p_i,i}(L)u_{i,t} = \Theta_{q_i,i}(L)\varepsilon_{i,t}, \tag{3}$$

where $\varepsilon_{i,t} \sim \mathrm{N}(0,\Omega)$, with Ω a $N \times N$ matrix. Naturally, the size and form of this covariance matrix will determine how the eventual parameters can get estimated. It might be that Ω is block-diagonal, due to specific spatial structures. In other cases it is perhaps best to assume that $\Omega = \sigma_i^2 I_N$, where I_N is an $N \times N$ identity matrix. The lag polynomials are defined by

$$\Phi_{p_i,i} = 1 - \phi_{1,i}L - \phi_{2,i}L^2 - \ldots - \phi_{p_i,i}L^{p_i}, \tag{4}$$

and

$$\Theta_{q_i,i} = 1 + \theta_{1,i}L + \theta_{2,i}L^2 + \ldots - \theta_{q_i,i}L^{q_i}. \tag{5}$$

The moving average part is seldom seen in practice, and I also do not foresee its intensive use in the near future. The lag polynomials indicate that each variable can be described by an ARMA model with individual-specific lag orders (p_i, q_i). If there are no moving average error terms, the model will be called a panel autoregression.

The pure time series models above can be extended by including explanatory variables, in which case one has a panel AR-X model. The $\mu_{i,t}$ term can then for example be

$$\mu_{i,t} = \beta_i X_{i,t} = \beta_{0,i} + \beta_{1,i}x_{1,i,t} + \ldots + \beta_{K,i}x_{K,i,t}, \tag{6}$$

indicating that there are K explanatory variables. The explanatory variables can also concern the same variables across the N equations, that is, one can consider

$$\mu_{i,t} = \beta_{0,i} + \beta_i x_t, \tag{7}$$

where x_t appears in all equations. For example, when looking at economic growth in N African countries, one can include as x_t the growth in the US economy, thereby assuming that this variable could have a country-specific influence on all countries. A particularly useful motivation to include such a variable is that it puts the contemporaneous cross-sectional correlation between the countries into the conditional mean part of the model, instead of in a potentially very large covariance matrix.

One may even further reduce the model by assuming that there are effects over time influencing all the series, without explicitly assigning any prior interpretation to these effects. An example of such a model is considered in Hjellvik and Tjostheim (1999), who consider the estimation of parameters for

$$y_{i,t} = \phi y_{i,t-1} + \eta_t + \varepsilon_{i,t}, \tag{8}$$

where the error terms are assumed to be independent and identically distributed.

The main difference between models for panels of time series and VAR models is the absence of other lagged endogenous variables on the right hand side, and also further cross-equation parameter restrictions. The presence of these variables naturally follows from extending a univariate AR model to a multivariate model. However, a drawback of a VAR model becomes the enormous amount of parameters and the problems with interpreting the parameters. In fact, the interpretation of the VAR model is cumbersome in the first place, and hence one usually has to resort to derivative measures such a variance decompositions and impulse-response functions. By the very nature of the VAR model, these derivative measures tend to have large standard errors in many practical applications, and hence might not very informative. Additionally, the analysis of VAR models for more than five or six variables already becomes quite impractical. Hence, the models for panels of time series are, admittedly, more restrictive in terms of parameters, but there might be a gain in terms of model interpretation. With examples in Section 3, I hope to be able to make some constructive arguments in that direction.

2.2 Estimation

There is not a single strategy for estimating the parameters of a model for a panel of time series. Depending on the size of T and of N, the quality of the data, the number of explanatory variables, the number of lagged endogenous variables, and, of course, the particular application at hand, the model representation will vary, and so will the estimation method. This section deals with only a few methods for a few specific models. For ease of notation, the basic model is assumed to be

$$y_{i,t} = \mu_{i,t} + \beta_i x_{i,t} + \rho_{i,1} y_{i,t-1} + \varepsilon_{i,t}, \tag{9}$$

that is, a first order dynamic model with a single explanatory variable.

In case one has large T and large N, there are in principle no objections to considering N separate time series models, although one might want to assume that the error terms are contemporaneously correlated. When N is large, one should however somehow restrict the covariance matrix of $\varepsilon_{i,t}$, as it will not be easy to invert unrestricted $N \times N$ matrices. Particular applications may suggest specific structures of the covariance matrix Ω, but I suspect that there will not be too many of these cases. The literature sofar seems to assume that Ω is a diagonal matrix with idiosyncratic error variances. If there would be a strong coherence across the N equations, one may then choose to incorporate variables which are common across all equations, perhaps with equation-specific effects, as in (7).

In many cases, an equation-by-equation estimation routine is perhaps unsatisfactory. This can happen when the data are not so informative, as they have little variance, or when the size of T is rather short, say like 20. Also, one might be explicitly interested in somehow summarizing the potentially huge number of parameters, to be obtained from N regressions, into a smaller amount, also to facilitate interpretation. For example, one can estimate N parameters β_i, but in the end be interested in the average value of these parameters and its associated standard error.

There are various ways to impose restrictions on the parameters in (9). A first example, which is usually coined as a fixed-effects model, is

$$y_{i,t} = \mu_i + \beta x_{i,t} + \rho_1 y_{i,t-1} + \varepsilon_{i,t}, \tag{10}$$

where the β and ρ_1 parameters are restricted to be equal across equations. The parameter μ_i takes equation-specific values, as it is unlikely that all $y_{i,t}$ variables have the same mean, even after correction for $x_{i,t}$ and $y_{i,t-1}$. Clearly, the number of parameters gets drastically reduced. Nickell (1981) shows that when T is small, there is a non-negligible bias in the parameter estimators for (10). But, Kunst and Franses (1999) show that when T takes values like 10 or more, this bias rapidly vanishes towards zero. This fixed-effects model is particularly useful in case one aims to test hypotheses on the ρ_1 parameter, for example, whether ρ_1 is equal to 1. These corresponding so-called unit root tests have notoriously low power for single equation models, but when N equations are analyzed at the same time, this power rapidly increases, see Phillips and Moon (1999) among others.

It is quite conceivable that imposing the restrictions that $\beta_i = \beta$ and $\rho_{i,1} = \rho_1$ cannot be validated by statistical tests. Indeed, when N is large, it is unlikely that all β_i parameters are equal in a statistical sense. Hence, poolability tests would then entail that the probability of rejecting the null hypothesis that the parameters can be collected into single parameters gets close to 1, no matter how good or bad the test is. On the other hand, if one has to resort to pooling, simply because there are not sufficient observations in certain dimensions, one does not want to test for poolability in the first place. This is because the possible rejection of the null hypothesis cannot be succeeded by the estimation of an unrestricted model, as it was already decided that this model was not plausible. Indeed, pooling can also be seen as a favorable approach, as one, so to speak, lets some equations help to get the parameters estimated in other equations. Hoogstrate et al. (2000) consider the consequences of pooling on forecasting. Naturally, one may decide to only partially pool the parameters, like in Pesaran, Shin and Smith (1999). This might require knowledge as to which equations are involved in pooling, although one might make such a decision based on first-round estimation results. Finally, pooling can also be used to examine if equations share common properties, like the same trend or the same deterministic seasonal fluctuations.

A second quite popular panel model is the random-coefficients model. There are many variants possible, and an example concerning the model in (9) is

$$y_{i,t} = \mu_i + \beta_i x_{i,t} + \rho_1 y_{i,t-1} + \varepsilon_{i,t}, \tag{11}$$

when $\mu_{i,t}$ is set at μ_i, and with

$$\beta_i \sim \mathrm{N}(\beta, \sigma^2). \tag{12}$$

This model says that the average effect of $x_{i,t}$ is β, but that there is equation-specific variation that can be captured by a normal distribution with variance σ^2. Again, the number of parameters get drastically reduced. A test for poolability can also be derived for this model, as it amounts to σ^2 being equal to 0.

It should be mentioned here that random-coefficients models for panels of pure time series might be difficult to interpret. Indeed, the basic notion behind the model is that the equation-specific parameters are collected in a single distribution. When this concerns the parameters of an explanatory variable $x_{i,t}$, the collected parameters have a natural interpretation. However, when it concerns the parameters of lagged endogenous variables, matters change. For example, consider

$$y_{i,t} = \mu_{i,t} + \rho_{i,1} y_{i,t-1} + \varepsilon_{i,t}. \tag{13}$$

If one has N time series, it might be that some of these show explosive behavior, others might display a trend, others might look like a random walk, while again others would cycle around a constant mean. These data can all be captured by a first order autoregression. For the first, one would find a $\rho_{1,i}$ parameter to exceed 1, the second set has $\rho_{i,1} = 1$, the third additionally has that $\mu_{i,t} = 0$, while the last set of series have $\rho_{i,1}$ smaller than unity. It may even happen that some time series show seasonal fluctuations, which would correspond with $\rho_{i,1} = -1$, or close to it. Hence, the time series might look all very different, and it may then seem odd in the first place, to try to collect the parameters into a single distribution. Of course, when the equation-specific $\rho_{i,1}$ parameters would all look very similar, one may be tempted to do so. But then still, if the model is not an autoregression of order 1, but of order 2, that is,

$$y_{i,t} = \mu_{i,t} + \rho_{1,i} y_{i,t-1} + \rho_{2,i} y_{i,t-2} + \varepsilon_{i,t}, \tag{14}$$

even small differences between the ρ parameters would correspond with very different time series patterns. In short, the main reason for these difficulties is that for time series models, the value of the parameters determines how the data, that could be generated from these models would look like, see Franses (1998), among others. Hence, straightforwardly estimating a random-coefficients model for a panel autoregression of order P might not be a good strategy. Indeed, the interpretation of the parameters is cumbersome, if not impossible.

The random-coefficients model with (12) can also be extended by including explanatory variables, like for example,

$$\beta_i \sim \mathrm{N}(\beta_0 + \beta_1 z_i, \sigma^2), \tag{15}$$

where z_i is an observed variable. Fok, van Nierop and Franses (2002) consider this model for two years of weekly sales on 23 items in the same product category. The effects of promotions and distribution in $x_{i,t}$ are made a function of the size of an item and its location on a shelf. Estimation of the model parameters is not straightforward, as it needs modern simulation-based techniques, like those surveyed in Paap (2002). In empirical marketing research, one often sees the approach where in a first round the β_i parameters get estimated for the N equations, and next, a regression of these estimated parameters on z_i is considered. This approach introduces randomness in the β_i only in the second stage, and hence overestimates confidence in this second stage regression.

An approach which goes a little beyond the random-coefficients models as given above is proposed in Franses, Paap and van Dijk (2002). These authors consider the model

$$y_{i,t} - \mu_{i,s} = \phi_i(y_{i,t-1} - \mu_{i,s}) + \beta_i(x_t - \delta) + \varepsilon_{i,t}, \tag{16}$$

where $y_{i,t}$ is a country-specific growth rate in GNP per capita with mean $\mu_{i,s}$, where x_t is the corresponding growth rate in the US, with mean δ, and where the $\varepsilon_{i,t}$ are independent and identically distributed. The notation $\mu_{i,s}$ corresponds with the idea that Franses, Paap and van Dijk (2002) assume that the N countries, in terms of their GNP per-capita growth rates, can be classified into S classes, within which the growth rates obey

$$\mu_{i,s} \sim \mathrm{N}(\mu_s, \sigma_s^2), \tag{17}$$

where $\mu_{i,s}$ corresponds with observation i in class s, with $s = 1, 2, ..., S$. Each country has a probability p_s, with $p_1 + p_2 + ... + p_S = 1$, to get assigned to a class s. This idea builds on the latent class methodology, recently summarized in Wedel and Kamakura (1999). Using simulation-based estimation techniques, the probabilities can be computed, and one can also decide on the empirical value of S. The application in Franses, Paap and van Dijk (2002) deals with a comparison of growth patterns in Africa and Asia, and, by example, it leads to statements that Kenia has a probability of 0.85 to get assigned to the fast-growing class of countries, with mean growth of 10 per cent, and a probability of 0.15 to be associated with the low-growth class with 1 per cent. Also, it can be seen how many countries of which continent get assigned to classes of fast- or slow-growing countries.

There are perhaps many other ways to estimate parameters, and new methods need to be developed. It is clear though that for useful, sensible and practical models for panels of time series, one needs to resort to simulation-based estimation techniques. Standard panel estimation methods seem to require many assumptions, that are implausible from either an economic point of view or from statistical considerations.

3 Incorporating time series features

In this section, I will review a few extensions of a basic model for a panel of time series, which address the particular properties of time series data.

3.1 Trends

Economic time series data can display trending patterns. Such a trend can be described by a model where the autoregressive parameter is equal to unity, or where the sum of autoregressive parameters is equal to unity, and the model contains a non-zero intercept. It can also be described by a model which includes the term t. In time series econometrics, it can be important to diagnose if the data have a unit root or not. Tests for unit roots in panel data are developed in Levin, Lin and Chu (2002) (which is a revised version of a working paper that circulated since 1992) and in Harris and Tzavalis (1999), Moon and Phillips (2000), among others, while these are applied in for example Papell (1997) and O'Connell (1998). These tests usually consider the model

$$y_{i,t} = \mu_i + \delta_i t + \rho_{i,1} y_{i,t-1} + \varepsilon_{i,t}. \tag{18}$$

and examine if $\rho_{i,1} = 1$ for all i, against the alternative that some or all are not equal to unity. Interestingly, and in contrast to single time series, it is found that the asymptotic theory for these tests is again standard normal, provided that N and T are both large. This also suggests that the potentially occurrence of spurious regression results might be less relevant for panels of time series. If so, one can then also analyze panel error correction models, like

$$y_{i,t} - y_{i,t-1} = \alpha_i(x_{i,t} - x_{i,t-1}) + \beta_i(y_{i,t-1} - \gamma_i x_{i,t-1}) + \varepsilon_{i,t}, \tag{19}$$

with the possibly plausible restriction that $\gamma_i = \gamma$ for all i, see also Groen and Kleibergen (2002).

Other time series do not display unit root type trend behavior, and hence their trends might be best captured by deterministic trend functions. In some cases, one might then be interested in investigating if two or more trending series, where the trend can be described by a linear deterministic trend, have the same trend. Vogelsang and Franses (2001) consider N trend-stationary time series denoted by $y_{1,t}$ to $y_{N,t}$ with $t = 1, 2, \ldots, T$, and assume that the representation

$$\begin{aligned} y_{1,t} &= \mu_1 + \beta_1 t + u_{1,t} \\ y_{2,t} &= \mu_2 + \beta_2 t + u_{2,t} \\ &\ldots \\ y_{N,t} &= \mu_N + \beta_N t + u_{N,t}. \end{aligned} \tag{20}$$

Define the three $N \times 1$ vectors u_t, μ and β by $(u_{1,t}, u_{2,t}, \ldots, u_{N,t})'$, $(\mu_1, \mu_2, \ldots, \mu_N)'$ and $(\beta_1, \beta_2, \ldots, \beta_N)'$, respectively. It is assumed that a functional limit theorem applies to u_t, that is,

$$T^{-\frac{1}{2}} \sum_{t=1}^{[rT]} u_t \Rightarrow \Lambda W_N(r), \tag{21}$$

where $\Rightarrow$ denotes weak convergence in distribution, where $W_N(r)$ is an $N \times 1$ standard independent Wiener process, and $[rT]$ is the integer part of rT. Denoting Ω as the long-run variance of u_t, that is,

$$\begin{aligned} \Omega &= \Lambda\Lambda' \\ &= \sum_{j=-\infty}^{+\infty} \Gamma_j, \end{aligned} \tag{22}$$

where $\Gamma_j = \text{Cov}[u_t u'_{t-j}]$. The hypotheses of interest in Vogelsang and Franses (2001) are

$$\begin{aligned} H_0 &: R\beta = r \\ H_1 &: R\beta \neq r, \end{aligned} \tag{23}$$

where R is a $q \times n$ matrix, and r is a $n \times 1$ vector.

Vogelsang and Franses (2001) derive the asymptotic theory, and show that this is not equivalent to the standard theory. They also provide the relevant critical values. The method is applied to a sample of about 300 annual data on monthly temperatures in The Netherlands. One of the conclusions is that there is evidence of warming in the winter, especially in January, and cooling in the summer, especially September, and that there are no significant trends in the other seasons.

3.2 Seasonality

Before one can say something about deterministic seasonality, one first has to decide on the number of seasonal unit roots. The basic test regression to test for seasonal unit roots is

$$\begin{aligned}\Phi_{p_i(,i}(L)\Delta_4 y_{i,t} &= \alpha_i + \beta_i t + \sum_{s=1}^{3} \gamma_{i,s} D_{s,t} \\ &\quad + \rho_{i,1} y^{(1)}_{i,t-1} + \rho_{i,2} y^{(2)}_{i,t-1} + \rho_{i,3}\Delta_2 y_{i,t-1} + \rho_{i,4}\Delta_2 y_{i,t-2} + \varepsilon_t, \quad (24)\end{aligned}$$

where Δ_k is the $k-$th order differencing filter, where $y^{(1)}_{i,t} = (1 + L + L^2 + L^3)y_{i,t}$ and $y^{(2)}_{i,t} = (1 - L + L^2 - L^3)y_{i,t}$, and where $D_{s,t}$ are seasonal dummies for the first, second and third quarter. The model assumes that the series $y_{i,t}$ can be described by a $(p_i + 4)-$th order autoregression. A version of this test regression, which enables setting the parameters of deterministic variables equal to zero when they correspond with certain seasonal unit roots, is given by replacing the seasonal dummies by the relevant cosine functions, see Franses and Kunst (1999). An appropriate test for a seasonal unit root at the bi-annual frequency is now given by a joint $F-$test for $\rho_{i,2}$ and $\alpha_{1,i}$. An appropriate test for the two seasonal unit roots at the annual frequency is then given by a joint $F-$test for $\rho_{3,i}$, $\rho_{4,i}$, $\alpha_{2,i}$ and $\alpha_{3,i}$. Franses and Kunst (1999) use these F-tests in a pure fixed-effects model (with all common parameters) and in a model where the autoregressive parameters are pooled over the equations. It is seen that the power of this panel test procedure is rather large. Additionally, Franses and Kunst (1999) examine if, once one has take care off seasonal unit roots, two or more series have the same deterministic fluctuations.

Along the lines of this last study, Franses (2002b) analyzes two years of weekly sales data of brands in the FMCG category. There it is of interest to see if prices and promotion activities display seasonal variation, and if sales do so too. Next, it is studied if price elasticity is constant throughout the year. As it concerns weekly data, one might be tempted to consider fifty-one dummies. However, one can also use the representation

$$y_{i,t} = \mu_i + \sum_{k=1}^{26}[\alpha_{i,k}\cos(\frac{2\pi kt}{52}) + \beta_{i,k}\sin(\frac{2\pi kt}{52})] + \varepsilon_{i,t}, \qquad (25)$$

and decide to focus only on interpretable cycles, like within four weeks or within twenty-six weeks. The interest also lies in tests for for example $\alpha_{i,k} = \alpha_k$ and $\beta_{i,k} = \beta_k$.

3.3 Outliers

Many economic time series display single or sequences of aberrant observations. Examples are innovation and additive outliers, permanent structural breaks and level shifts. As single outliers occur only once for each time series, one might examine if there are common outliers across series. For example, consider the innovation outlier model for a panel of time series, that is,

$$y_{i,t} = \mu_{i,t} + \rho_{i,1}y_{i,t-1} + \omega_i I_t[t = \tau] + \varepsilon_{i,t}, \qquad (26)$$

then one can gain precision for example by assuming that $\omega_i \sim \mathrm{N}(\omega, \sigma_\omega^2)$. This model can be used to test for the presence of common outliers.

It is well known that time series data with many occasional level shifts can perhaps best be described by fractionally integrated models. The key parameter in these models is the fractional integration parameter. It might now well be that a panel of time series can be described by

$$(1 - L)^{d_i} y_{i,t} = \mu_{0,i} + u_{i,t}, \qquad (27)$$

where d_i is the parameter of interest, and that parameter estimation improves by assuming that $d_i = d$ for all i.

3.4 Nonlinearity

Finally, there are many economic variables that display nonlinear features, which can usually be characterized by the presence of two or more regimes, across which parameters can differ. For macroeconomic variables, one can think of recessions and expansions. For asset returns, one can think about high and low volatile periods. There are various models that one can use, and it is yet unclear which models for single variables lend themselves best for extension to the panel case. It is easy to write down nonlinear models for panels, but whether these models make sense is unclear. For example, a nonlinear random-coefficient panel time series model could look like

$$y_{i,t} = \alpha_{i,t} y_{i,t-1} + u_{i,t}, \tag{28}$$

with

$$\alpha_{i,t} = \rho_i \alpha_{i,t-1} + v_{i,t}, \tag{29}$$

where

$$\rho_i \sim \mathrm{N}(\rho, \sigma^2), \tag{30}$$

and other examples can easily be given.

In case one has access to a panel of nonlinear time series, one might want to examine if the series share their nonlinearity feature. Typically, common nonlinearity gets defined as that a linear combination of two otherwise nonlinear time series is again a linear time variable. An alternative definition may be that series have the same parameters or variables in the nonlinear components of the model. For example, consider a panel Smooth Transition AR [STAR] model

$$y_{i,t} = \mu_{0,i} + \phi_{1,i} y_{i,t-1} + \phi_{2,i} F(z_{t-1}, \theta) y_{i,t-1} + \varepsilon_{i,t}, \tag{31}$$

where

$$F(z_{t-1}, \theta) = \frac{1}{1 + \exp[-\gamma_i (z_{t-1} - \gamma_0)]}, \tag{32}$$

with z_t is some leading indicator variable, which is common to all series. In words, this model entails that once this leading variable exceeds a common threshold, the model parameters change, but with a speed (determined by γ_i) which differs across, say, countries.

If one would want to allow for a world business cycle and a idiosyncratic business cycle, one can consider

$$y_{i,t} = \mu_{0,i} + \mu_{1,i}F(z_{t-1}, \theta) + \mu_{2,i}F(y_{i,t-1}, \theta_1) + \phi_{1,i}y_{i,t-1} + \varepsilon_{i,t}, \tag{33}$$

where

$$F(y_{i,t-1}, \theta_1) = \frac{1}{1 + \exp[-\gamma_{0,i}(y_{i,t-1} - \gamma_{2,i})]}. \tag{34}$$

This model effectively is a panel artificial neural network model. Franses and Van Dijk (2002) consider the above models to capture a large panel of unemployment data, in order to examine the common business cycle in unemployment and the importance of sector-specific variations.

4 Conclusion

This paper has discussed various aspects of the analysis of panels of time series. New models have been proposed and new estimation routines discussed. Various illustrations were discussed to show the potential usefulness of the models. It is expected that many further developments are possible, both for econometric theory and for practice.

One such development might concern the analysis of time series of cross sections. For example, one can obtain weekly survey data on thousands of individuals, and it might be of interest to examine transition probabilities and to see how these change over time.

References

Baltagi, B. (1995), *Econometric Analysis of Panel Data*, New York: Wiley.

Fok, D., E. van Nierop and P.H. Franses (2002), Sales models for many items using attribute data, Report ERS-2002-??-MKT, Erasmus Research Institute of Management.

Franses, P.H. (1998), *Time Series Models for Business and Economic Forecasting*, Cambridge: Cambridge University Press.

Franses, P.H. (2002a), *A Concise Introduction to Econometrics, An Intuitive Guide*, Cambridge: Cambridge University Press.

Franses, P.H. (2002b), On seasonality in weekly FMCG scanner data, Working paper.

Franses, P.H. and R.M. Kunst (1999), Testing Common deterministic seasonality, with an application to industrial production, Econometric Institute Report 9905, Erasmus University Rotterdam.

Franses, P.H., R. Paap and D.J.C. van Dijk (2002), Do African countries grow slower than Asian countries?, Econometric Institute Report 2002-??, Erasmus University Rotterdam.

Gonzalo, J. and J.Y. Pitarakis (1999), Dimensionality effects in large dimensional systems, in R.F. Engle and H. White (eds.), *Cointegration, Causality and Forecasting: Volume in Honour of C.W.J. Granger*, Oxford: Oxford University Press.

Granger, C.W.J. (1998), Extracting information from mega-panels and high-frequency data, *Statistica Neerlandica*, 52, 258-272.

Groen, J.J.J. and F. Kleibergen (2002), Likelihood-based cointegration analysis in panels of vector error correction models, *Journal of Business and Economic Statistics*, in print.

Harris, R.D.F. and E. Tzavalis (1999), Inference for unit roots in dynamic panels where the time dimension is fixed, *Journal of Econometrics*, 91, 201-226.

Hjellvik, V. and D. Tjostheim (1999), Modelling panels of intercorrelated autoregressive time series, *Biometrika*, 86, 573-590.

Holtz-Eakin, D., W. Newey and H.S. Rosen (1988), Estimating vector autoregressions with panel data, *Econometrica*, 56, 1371-1395.

Hsiao, C. (1986), *Analysis of Panel Data*, Cambridge: Cambridge University Press.

Kunst, R.M. and P.H. Franses (1999), Testing for converging deterministic seasonal variation in European industrial production, Econometric Institute 9917, Erasmus University Rotterdam.

Levin, A., C.-F. Lin and C.-S. J. Chu (2002), Unit root tests in panel data: asymptotic and finite-sample properties, *Journal of Econometrics*, 108, 1-24.

Moon, H.R. and P.C.B. Phillips (2000), Estimation of autoregressive roots near unity using panel data, Econometric Theory, 16, 927-997.

Nickell, S. (1981), Biases in dynamic models with fixed effects, *Econometrica*, 49, 1399-1416.

O'Connell, P. (1998), The overvaluation of PPP, *Journal of International Economics*, 44, 1-19.

Paap, R. (2002), What are the advantages of MCMC based inference in latent variable models? *Statistica Neerlandica*, 56, 2-22.

Papell, D.H. (1997),Searching for stationarity: PPP under the current float, *Journal of International Economics*, 43, 313-332.

Pesaran, M.H., Y.C. Shin and R.P. Smith (1999), Pooled mean group estimation of dynamic heterogeneous panels, *Journal of American Statistical Association*, 94, 621-634.

Phillips, P.C.B. and H.R. Moon (1999), Linear regression limit theory for nonstationary panel data, it Econometrica, 67, 1057-1111.

Van Dijk, D.J.C. and P.H. Franses (2002), Common cycles in a panel of nonlinear time series, Econometric Institute Report 2002-??, Erasmus University Rotterdam.

Vogelsang, T.J. and P.H. Franses (2001), Testing for common deterministic trend slopes, Econometric Institute Report 2001-16, Erasmus University Rotterdam.

Wedel, M. and W.A. Kamakura (1999), *Market Segmentation: Conceptual and Methodological Foundations*, Boston: Kluwer.