Forecasting with Spatial Panel Data*

Badi H. Baltagi, Department of Economics and Center for Policy Research

This paper compares various forecasts using panel data with spatial error correlation. The true data generating process is assumed to be a simple error component regression model with spatial remainder disturbances of the autoregressive or moving average type. The best linear unbiased predictor is compared with other forecasts ignoring spatial correlation, or ignoring heterogeneity due to the individual effects, using Monte Carlo experiments. In addition, we check the performance of these forecasts under misspecification of the spatial error process, various spatial weight matrices, and heterogeneous rather than homogeneous panel data models.

JEL Classification: C33

Keywords: forecasting, BLUP, panel data, spatial dependence, heterogeneity

* This paper was presented at a conference in honor of Phoebus Dhrymes in Paphos, Cyprus, June 13, 2007. Also, at the 14th International Conference on Panel Data at the Wang Yanan Institute for Studies in Economics (WISE), Xiamen University, China, July 16-18, 2007, and the 63rd European Meeting of the Econometric Society (ESEM) held at the University of Bocconi in Milan, Italy, August 27-31, 2008.
Level, Change and Social-Structural Determinants of Group-Focused Enmity in Germany

Stefan Thörner*, Stefanie Goosen*, Peter Schmidt** and Eldad Davidov***
* University of Gießen ** University of Marburg *** University of Zurich

In this study we investigate the level and change of group focused enmity (GFE) in Germany between 2002 and 2006. We do it by using a third-order latent growth model (LGM). We study GFE as one generalized higher-order factor, which incorporates prejudices towards various minority groups. GFE is measured as a second order factor, whose elements are theoretically and empirically founded, and include: anti-Semitism, racism, xenophobia, homophobia, exclusion of homeless people and support for rights of the established. We use a representative 4-years panel data of the adult non-immigrant German population collected during the years 2002-2006 to test the development of GFE in East and West Germany while addressing the issue of metric and scalar invariance over time and across groups for the first and second order factors. We find that the level of GFE is stable over time in Germany in the time period between 2002-2006. We try to explain it by a set of social structural variables. Finally, we consider the problem of social desirability response set and how it may affect our conclusions.

Key words: group focused enmity; prejudice; latent growth curve modeling; metric and scalar invariance; social desirability
Longitudinal multilevel modeling. Growth curves and curves of factor models.

Uwe Engel, Universität Bremen, Institut für empirische und angewandte Soziologie (EMPAS)

Starting point of the paper is a hierarchical linear regression model for the analysis of linear and non-linear growth using multi-wave panel data. It is shown how to “translate” this model of linear respectively non-linear growth in equivalent model variants of the structural equation framework (step 1). This step is followed by the additional introduction of latent variables to mediate the manifest indicator variables used and the latent concepts whose means, variances and covariance are used to estimate initial values and expected change over time as well as variances and intercept-slope covariance of involved random effects (step 2). This is done to handle measurement error, in particular to meet the need of separating out true change and unreliability of response. This step is followed by the additional introduction of a corresponding structure of exogenous variables to explain variations in initial values and expected changes over time (step 3).

The paper discusses what is gained by “translating” hierarchical linear regression models into models of the structural equation resp. mixture modeling framework. All modeling is exemplified using data of the German Socio-Economic Panel Study and is carried out by employing the software packages MLwiN, LISREL, and Mplus.
A Panel Data Approach for Program Evaluation — Measuring the Benefits of Political and Economic Integration of Hong Kong with Mainland China*

by

Cheng Hsiao\textsuperscript{a,b,c}, H. Steve Ching\textsuperscript{b,d}, and Shui Ki Wan\textsuperscript{a}

October 7 2009

Abstract

We propose a simple to implement panel data method to evaluate the impacts of social policy. The basic idea is to exploit the dependence among cross-sectional units to construct the counterfactuals for the missing values. We attribute the cross-sectional correlations to the presence of some (unobserved) common factors. However, instead of trying to estimate the impacts of these factors, we propose to use observed data instead. The limited Monte Carlo studies appear to favor our approach over the factor approach when time series dimension is not large. We use our approach to evaluate the impact of political and economic integration of Hong Kong with China. Using the time series data of Australia, Austria, Canada, China, Denmark, Finland, France, Germany, Indonesia, Italy, Japan, Korea, Malaysia, Mexico, Netherlands, New Zealand, Norway, Philippines, Singapore, Switzerland, Taiwan, Thailand, U.K., and U.S. to construct what would have happened to Hong Kong’s real economic growth had there been no interventions, we find that the reversion of sovereignty of Hong Kong from the U.K. to China hardly had any impact. However, the Closer Economic Partnership Agreement (CEPA) between Hong Kong and Mainland China has raised Hong Kong’s real economic growth rate by about 4% a year.

\textsuperscript{a}University of Southern California
\textsuperscript{b}City University of Hong Kong
\textsuperscript{c}WISE, Xiamen University
\textsuperscript{d}Feng Chia University, Taiwan
H. Steve Ching, hsching@cityu.edu.hk
C. Hsiao, chsiao@usc.edu
S. Wan, shuiwan@usc.edu

*We would like to thank T.S. Wang of the Chinese Academy of Social Sciences and Y. Fan and M. Guo of Xiamen University for assistance in obtaining data for China and L.C. Lau of the Pearl River and Yangtze River Delta Collection, City University of Hong Kong for assistance in getting the data for Hong Kong.
The justice evaluation of own earnings in longitudinal perspective

Stefan Liebig, Fakultät für Soziologie, Universität Bielefeld, Jürgen Schupp, Peter Valet. Leonberg

Empirical research on distributive justice seeks to identify justice perceptions in a society which can be considered as patterns that legitimate inequalities. Therefore the analysis of distributive justice evaluations primarily focuses on the perceived justice of (labour) incomes.

The theoretical framework to explain deviating justice evaluations of the personal income consists particularly of socio-psychological models. According to the equity-theory the proportion of contributions and reward determines whether the individual income is considered as just or unjust. Other approaches like the theory of relative deprivation and the status value theory refer to the importance of social comparisons. Individuals evaluate the justice of their own reward by comparing their profits with the profits of a “generalised other” (certain reference persons or reference groups).

So far empirical justice research using large scale survey data was restricted to a cross-sectional perspective. The recent wave of the GSOEP (2009) now provides the third query (after 2005 and 2007) concerning the justice evaluation of the own income. Hence it is now possible for the first time to investigate dynamic aspects of just incomes with a panel approach.

The results of the longitudinal perspective show robust patterns with a significant period effect for 2007. Especially the variations of individual contributions over time have an impact on the subjective justice evaluation. Furthermore there are different effects according to the level of the own income and to the respective phase in the life course.
Measurement of structural and individual change by means of a 3-level IRT model

Herbert Matschinger, Clinic of Psychiatry and Psychotherapy, University of Leipzig

Measuring individual change and inter-individual differences of change on latent constructs is one of the core efforts in many fields of science, for instance in intervention and/or observational studies. The most important task is certainly to evaluate measurement invariance for both the characteristics of the respondents, and the measurement occasions. Only if invariance holds, the impact of time and respondents characteristics as well as the interaction between them can be estimated without bias. Structural invariance for different measurement occasions is frequently evaluated by multiple group confirmatory factor analysis, where each group represents one occasion. Measurement for each occasion implicitly is treated as an independent observation, although this assumption necessarily is violated since each “group” comprises the same individuals. Individual locations of the observation on the latent trait at each occasion are not explicitly part of the model, so characteristics of the respondents and the occasions cannot easily be integrated. Furthermore, in many instances, multi-category manifest variables are treated as ordinal, but the thresholds themselves cannot be subjected to be dependent variables in an explanatory model.

Therefore, we present an approach from the framework of the generalized linear latent and mixed model (Rabe-Hesketh et al., 2004; Skrondal and Rabe-Hesketh, 2004; Rabe-Hesketh et al., 2007) which is employed to estimate and explain both the parameters of the dichotomous and ordinal (partial credit or rating scale) IRT model (De Boeck and Wilson, 2004). Repeated measurement is treated as items in occasions in respondents (3-level). Response categories basically are treated as nominal, thresholds and constraints are defined as predictors in order to define the more restrictive models just mentioned. (Zheng and Rabe-Hesketh, 2007)

Two application, both from psychiatry research, will exemplify this approach by means of the so called EQ-5D - a 5 –item, 3-category questionnaire, developed to measure health related quality of life. The first study comprises 390 patients with anxiety disorder over three points in time, with an intervention between baseline and first follow-up. The second example employs data from an observational longitudinal study on 1208 patients in three different European countries, suffering from schizophrenia, which where observed over five measurement occasions.

It will be shown, that, on the one hand, measurement invariance with respect to item difficulty and category discrimination obviously does not hold, but on the other hand, this variability sheds light on the particular process over time, with and without an intervention.

Reference List


Estimating and Evaluating Dynamic Processes from Panel Data
John J. McArdle, University of Southern California, Los Angeles, CA, USA

Although many behavioral science theories are dynamic in nature, describing processes that unfold over time with multiple interrelated parts, the analytic methods typically used fall short of evaluating such theoretical predictions. In this paper we attempt to clarify this mismatch between theory and methods. We describe current models that focus on an appreciation of dynamic processes, and we place direct emphasis on the change and its underlying mechanisms or antecedents. We specify how these models can be used to estimate and evaluate hypotheses related to time dependency, interrelation with other processes, discontinuities in such processes, as well as individual differences. We also spell out how the proposed “dynamic but structural equation model” (DSEM) approach uses widely available software, and offers a general framework that can accommodate a broad variety of models. We show how SEM embraces the popular cross-lagged, latent growth curve, multi-level, and time-varying covariate models. Moreover, we explain why some of those popular models are not adequate to evaluate dynamic hypotheses that involve developmental sequences and interrelations between variables over time. These demonstrations are based on linear algebra, statistical simulation, and substantive examples.
Level, Change and Social-Structural Determinants of Group-Focused Enmity in Germany
Stefan Thörner*, Stefanie Goosen*, Peter Schmidt** and Eldad Davidov***
* University of Gießen  ** University of Marburg  *** University of Zurich

In this study we investigate the level and change of group focused enmity (GFE) in Germany between 2002 and 2006. We do it by using a third-order latent growth model (LGM). We study GFE as one generalized higher-order factor, which incorporates prejudices towards various minority groups. GFE is measured as a second order factor, whose elements are theoretically and empirically founded, and include: anti-Semitism, racism, xenophobia, homophobia, exclusion of homeless people and support for rights of the established. We use a representative 4-years panel data of the adult non-immigrant German population collected during the years 2002-2006 to test the development of GFE in East and West Germany while addressing the issue of metric and scalar invariance over time and across groups for the first and second order factors. We find that the level of GFE is stable over time in Germany in the time period between 2002-2006. We try to explain it by a set of social structural variables. Finally, we consider the problem of social desirability response set and how it may affect our conclusions.

Key words: group focused enmity; prejudice; latent growth curve modeling; metric and scalar invariance; social desirability
Longitudinal and Panel Data in Labor Market Research

Joachim Möller, Institut für Arbeitsmarkt- und Berufsforschung der Bundesagentur für Arbeit, Nürnberg

The availability of longitudinal and panel data for labor market research has enormously increased in the last one or two decades. This holds true at the micro and macro level as well as for administrative and for survey data.

At the same time the application of panel econometrics has revolutionized the methodology in vast fields of empirical labor economics. One reason is that panel methods help to control for unobserved heterogeneity at the individual level. This is crucial for many applications, for instance in the evaluation of active labor market programs, the determination of earnings, migration analysis and so on. Furthermore, dynamic panel data methods help to tackle the endogeneity problem under some conditions and last but not least panel data approaches might ease the identification of causal effects in non-experimental situations.

For the future several at least two trends are visible. One is the increasing importance of data linkage as, for instance, in the linked employer-employee data sets. The other is the growing number of panel studies including the spatial dimension. In both areas methods and applications are rapidly developing in recent research.
First and second-order stochastic differential equation modeling as an alternative to the LGC and ALT model of panel data

J.H.L. Oud, Behavioural Science Institute, Radboud University Nijmegen, The Netherlands

Abstract. In the structural equation modeling (SEM) literature, two models for the analysis of longitudinal data became very popular in the past: autoregressive (AR) cross-lagged model and latent growth curve (LGC) or latent trajectory (LT) model. Curran and Bollen (2001) and Bollen and Curran (2004), however, argued that, theoretically, there are many instances when both the processes described by the AR model and the processes described by the LT model are plausible. They proposed the autoregressive latent trajectory (ALT) model, which captures features of both. The discrete-time approach in the ALT model has been criticized by Delsing and Oud (2008), who proposed a continuous time version of the ALT model, using stochastic differential equations, called “continuous time autoregressive latent trajectory” (CALT) model. Next, serious problems related to the linear components in the LGC, ALT and CALT models are dealt with. As an alternative for the linear component the first-order derivative in a second-order stochastic differential equation model is proposed. This is applied to Marital Satisfaction data, collected in four consecutive years (2002-2005). It is pointed out that the first-order derivative as explanatory variable has none of the problems associated with the linear component.


Growth Mixture Models in Longitudinal Research

Jost Reinecke / Daniel Seddig, Faculty of Sociology, University of Bielefeld

Latent growth curve models as structural equation models are extensively discussed in various research fields (Duncan, Duncan, Strycker, Li & Alpert, 2006; Curran & Muthén, 1999; Muthén & Muthén, 2000a; Muthén & Muthén, 2002). Recent methodological and statistical extension are focused on the consideration of unobserved heterogeneity in empirical data. Muthén extended the classical structural equation approach by mixture components, i.e. categorical latent classes (Muthén 2001, 2002, 2004).

The paper will discuss applications of growth mixture models with data from one of the first panel studies in Germany which explore deviant and delinquent behavior of adolescents (Reinecke, 2006a, 2006b; Boers et al. (forthcoming)). Observed as well as unobserved heterogeneity will be considered with growth mixture models using the program Mplus (Muthén & Muthén, 2006). Special attention is given to the distribution of the substantive dependent variables as a count variable (Poisson model, zero-inflated Poisson model, negative binomial model and zero-inflated negative binomial model). Different model specifications with respect to substantive questions will also be emphasized.

Keywords: Panel data, Growth Mixture Models, Heterogeneity, Poisson model, Negative Binomial Models
Continuous-Discrete State-Space Modelling of Panel Data with Nonlinear Filter Algorithms

Hermann Singer, FernUniversität in Hagen, Lehrstuhl für angewandte Statistik und Methoden der empirischen Sozialforschung

Continuous time models with sampled data possess several advantages over conventional discrete time series and panel models (cf., e.g. special issue 62:1, 2008, of Statistica Neerlandica). For example, data with unequal time intervals between the waves can be treated efficiently, since the model parameters of the dynamical system model are not affected by the measurement process. The continuous-discrete state space model is a hybrid of continuous time dynamics (stochastic differential equations, SDE) and discrete time noisy measurements.

First I review the maximum likelihood (ML) estimation of linear panel models using Kalman filtering and structural equations models (SEM). Using the Gaussian likelihood function of the SEM equations it is shown that pure time series and correlated panel data (e.g. with random time effects) can also be treated exactly by SEM methods. General nonlinear panel models are estimated by approximate filtering methods such as the extended Kalman filter (EKF), the local linearization filter (LLF), the Gauss-Hermite filter (GHF) and the unscented Kalman filter (UKF). Again, correlated panels are treated by stacking the panel units in a vector Itô equation.

Finally, spatial dynamical models are discussed. The state variables are now random fields given as solutions of stochastic partial differential equations (SPDE), driven by a cylindrical Brownian motion. We present approximate solutions using Itô equations, e.g. the transport equation describing drift, leakage and relaxation effects, and the spatial Ginzburg-Landau model exhibiting a bimodal distribution of the field components and diffusive behavior, as well as bifurcations. Furthermore, fields with noisy and sampled measurements are filtered and ML parameter estimates are computed.

Key Words:
Continuous-discrete state space models; Stochastic differential equations; Itô calculus; Sampling; Kalman filtering; Approximate nonlinear filtering; SEM modeling; Spatial models; Random fields; Stochastic partial differential equations
Estimation of a Panel Selection Model

Martin Spieß, Universität Hamburg, Fakultät für Erziehungswissenschaft, Psychologie und Bewegungswissenschaft, FB Psychologie

Although sample selection bias is a frequent problem of applied research, there has been no generalization of sample selection models with binary dependent variables of interest to data with temporal error correlations. In this paper a generalized estimating equation (GEE) approach is proposed to estimate panel data selection models, considering binary responses (probit) in both equations. Results of a simulation study and of an application to a highly unbalanced data set are presented to demonstrate the usefulness of the model.