

Preface

Panel data econometrics has evolved rapidly over the last decade. Dynamic panel data estimation, non-linear panel data methods and the phenomenal growth in non-stationary panel data econometrics makes this an exciting area of research in econometrics. The 11th international conference on panel data held at Texas A&M University, College Station, Texas, June 2004, witnessed about 150 participants and 100 papers on panel data.

This volume includes some of the papers presented at that conference and other solicited papers that made it through the refereeing process. *Theoretical econometrics contributions include:* Bai and Kao who suggest a factor model approach to model cross-section dependence in the panel co-integrated regression setting; Lejeune who proposes new estimation methods and some diagnostics tests for a general heteroskedastic error component model with unbalanced panel data; Ullah and Huang who study the finite sample properties of feasible GLS for the random effects model with non-normal errors; Kazemi and Crouchley who suggest a pragmatic approach to the problem of estimating a dynamic panel regression with random effects under various assumptions about the nature of the initial conditions; Krishnakumar who uses a generalized version of the Frisch–Waugh theorem to extend Mundlak’s (1978) results for the error component model. *Empirical applications include:* Sickles and Williams who estimate a dynamic model of crime using panel data from the 1958 Philadelphia Birth Cohort study; Baltagi and Griffin who find that at least 4 structural breaks in a panel data on liquor consumption for 21 Swedish counties over the period 1956–1999; Boumahdi, Chaaban and Thomas who estimate a flexible AIDS demand model for agricultural imports into Lebanon incorporating a three-way error component model that allows for product, country and time effects as separate unobserved determinants of import demand; Biørn, Skjerpen and Wangen who are concerned with the analysis of heterogeneous log-linear relationships (and specifically Cobb–Douglas production functions) at the firm-level and at the corresponding aggregate industry level. They use unbalanced panel data on firms from two Norwegian manufacturing industries over the period 1972–1993; Cermeño and Grier who apply a model that accounts for conditional heteroskedasticity and cross-sectional dependence to a panel of monthly inflation rates of the G7 over the period 1978.2–2003.9; Yasar, Nelson and Rejesus who use plant level panel data for Turkish manufacturing industries to analyze the relative importance of short-run versus long-run

dynamics of the export-productivity relationship; Drine and Rault who focus on developing countries and analyze the long-run relationship between real exchange rate and some macroeconomic variables, via panel unit root and cointegration tests; Harris, Tang and Tseng who quantify the impact of employee turnover on productivity using an Australian business longitudinal survey over the period 1994/5 to 1997/8; Kaltchev who uses proprietary and confidential panel data on 113 public U.S. companies over the period 1997–2003 to analyze the demand for Directors' and Officers' liability insurance; Ortega-Díaz who assesses how income inequality influences economic growth across 32 Mexican States over the period 1960–2002.

Theoretical econometrics contributions

Bai and Kao suggest a factor model approach to model cross-section dependence in the panel co-integrated regression setting. Factor models are used to study world business cycles as well as common macro shocks like international financial crises or oil price shocks. Factor models offer a significant reduction in the number of sources of cross-sectional dependence in panel data and they allow for heterogeneous response to common shocks through heterogeneous factor loadings. Bai and Kao suggest a continuous-updated fully modified estimator for this model and show that it has better finite sample performance than OLS and a two step fully modified estimator.

Lejeune proposes new estimation methods for a general heteroskedastic error component model with unbalanced panel data, namely the Gaussian pseudo maximum likelihood of order 2. In addition, Lejeune suggests some diagnostics tests for heteroskedasticity, misspecification testing using m-tests, Hausman type and Information type tests. Lejeune applies these methods to estimate and test a translog production function using an unbalanced panel of 824 French firms observed over the period 1979–1988.

Ullah and Huang study the finite sample properties of feasible GLS for the random effects model with non-normal errors. They study the effects of skewness and excess kurtosis on the bias and mean squared error of the estimator using asymptotic expansions. This is done for large N and fixed T , under the assumption that the first four moments of the error are finite.

Kazemi and Crouchley suggest a pragmatic approach to the problem of estimating a dynamic panel regression with random effects under various assumptions about the nature of the initial conditions. They find that the

full maximum likelihood improves the consistency results if the relationships between random effects, initial conditions and explanatory variables are correctly specified. They illustrate this by testing a variety of different hypothetical models in empirical contexts. They use information criteria to select the best approximating model.

Krishnakumar uses a generalized version of the Frisch–Waugh theorem to extend Mundlak's (1978) results for the error component model with individual effects that are correlated with the explanatory variables. In particular, this extension is concerned with the presence of time invariant variables and correlated specific effects.

Empirical contributions

The paper by Sickles and Williams estimates a dynamic model of crime using panel data from the 1958 Philadelphia Birth Cohort study. Agents are rational and anticipate the future consequence of their actions. The authors investigate the role of social capital through the influence of social norms on the decision to participate in crime. They find that the initial level of social capital stock is important in determining the pattern of criminal involvement in adulthood.

The paper by Baltagi and Griffin uses panel data on liquor consumption for 21 Swedish counties over the period 1956–1999. It finds that at least 4 structural breaks are necessary to account for the sharp decline in per-capita liquor consumption over this period. The first structural break coincides with the 1980 advertising ban, but subsequent breaks do not appear linked to particular policy initiatives. Baltagi and Griffin interpret these results as taste change accounting for increasing concerns with health issues and changing drinking mores.

The paper by Boumahdi, Chaaban and Thomas estimate a flexible AIDS demand model for agricultural imports into Lebanon incorporating a three-way error component model that allows for product, country and time effects as separate unobserved determinants of import demand. In their application to trade in agricultural commodities the authors are primarily concerned with the estimation of import demand elasticities. Conventionally, such estimates are frequently obtained from time series data that ignore the substitution elasticities across commodities, and thus implicitly ignore the cross-sectional dimension of the data. Exhaustive daily transactions (both imports and exports) data are obtained from the Lebanese customs administration for the years 1997–2002. Restricting their attention to major agricultural commodities (meat, dairy products, cereals, animals and vegetable fats and sugar), they estimate an import share equation

for European products as a function of own-price and competitors prices. Competition is taking place between European countries, Arab and regional countries, North and South America and the rest of the world. The import share equations are estimated by allowing for parameter heterogeneity across the 5 commodity groups, and tests for the validity of the multi-way error components specification are performed using unbalanced panel data. Estimation results show that this specification is generally supported by the data.

The paper by Biørn, Skjerpen and Wangen is concerned with the analysis of heterogeneous log-linear relationships (and specifically Cobb–Douglas production functions) at the firm-level and at the corresponding aggregate industry level. While the presence of aggregation bias in log-linear models is widely recognized, considerable empirical analysis continues to be conducted ignoring the problem. This paper derives a decomposition that highlights the source of biases that arise in aggregate work. It defines some aggregate elasticity measures and illustrates these in an empirical exercise based on firm-level data in two Norwegian manufacturing industries: The pulp and paper industry (2823 observations, 237 firms) and the basic metals industry (2078 observations, 166 firms) observed over the period 1972–1993.

The paper by Cermeño and Grier specify a model that accounts for conditional heteroskedasticity and cross-sectional dependence within a typical panel data framework. The paper applies this model to a panel of monthly inflation rates of the G7 over the period 1978.2–2003.9 and finds significant and quite persistent patterns of volatility and cross-sectional dependence. The authors use the model to test two hypotheses about the inter-relationship between inflation and inflation uncertainty, finding no support for the hypothesis that higher inflation uncertainty produces higher average inflation rates and strong support for the hypothesis that higher inflation is less predictable.

The paper by Yasar, Nelson and Rejesus uses plant level panel data for Turkish manufacturing industries to analyze the relative importance of short-run versus long-run dynamics of the export-productivity relationship. The adopted econometric approach is a panel data error correction model that is estimated by means of system GMM. The data consists of plants with more than 25 employees from two industries, the textile and apparel industry and the motor vehicles and parts industry, observed over the period 1987–1997. They find that “permanent productivity shocks generate larger long-run export level responses, as compared to long-run productivity responses from permanent export shocks”. This result suggests that industrial policy should be geared toward permanent improvements in plant-productivity in order to have sustainable long-run export and economic growth.

The paper by Drine and Rault focuses on developing countries and analyzes the long-run relationship between real exchange rate and some macroeconomic variables, via panel unit root and cointegration tests. The results show that the degrees of development and of openness of the economy strongly influence the real exchange rate. The panels considered are relatively small: Asia ($N = 7$, $T = 21$), Africa ($N = 21$, $T = 16$) and Latin America ($N = 17$, $T = 23$).

The paper by Harris, Tang and Tseng consider a balanced panel of medium sized firms drawn from the Australian business longitudinal survey over the period 1994/5 to 1997/8. The paper sets out to quantify the impact of employee turnover on productivity and finds that the optimal turnover rate is 0.22. This is higher than the sample median of 0.14 which raises the question about whether there are institutional rigidities hindering resource allocation in the labor market.

The paper by Kaltchev uses proprietary and confidential panel data on 113 public U.S. companies over the period 1997–2003 to analyze the demand for Directors' and Officers' liability insurance. Applying system GMM methods to a dynamic panel data model on this insurance data, Kaltchev rejects that this theory is habit driven but still finds some role for persistence. He also confirms the hypothesis that smaller companies demand more insurance. Other empirical findings include the following: Returns are significant in determining the amount of insurance and companies in financial distress demand higher insurance limits. Indicators of financial health such as leverage and volatility are significant, but not corporate governance.

The paper by Ortega-Díaz assesses how income inequality influences economic growth across 32 Mexican States over the period 1960–2002. Using dynamic panel data analysis, with both, urban personal income for grouped data and household income from national surveys, Ortega-Díaz finds that inequality and growth are positively related. This relationship is stable across variable definitions and data sets, but varies across regions and trade periods. A negative influence of inequality on growth is found in a period of restrictive trade policies. In contrast, a positive relationship is found in a period of trade openness.

I hope the readers enjoy this set of 15 papers on panel data and share my view on the wide spread use of panels in all fields of economics as clear from the applications. I would like to thank the anonymous referees that helped in reviewing these manuscripts. Also, Jennifer Broaddus for her editorial assistance and handling of these manuscripts.

Badi H. Baltagi
College Station, Texas and Syracuse, New York