The Public Defense

of the Doctoral Thesis in Economics

by

Laszlo Balazsi

on

The Econometrics of Linear Models for Multi-dimensional Panels

will be held on

Tuesday, March 28, 2017 at 9:00 am

in

N13 Room 118,

Central European University,

Nador utca 13, Budapest, Hungary

Thesis Committee

Julius Horvath (chair)
Andras Sugar (external member)
Adam Szentpeteri (external member)
Miklos Koren (internal member)
Gabor Kezdi (internal member)

Supervisor

Laszlo Matyas

Examiners

Patrick Sevestre, Professor at Aix-Marseille University, Faculty of Economics and Management
(external examiner)
Gabor Kezdi, Professor at Central European University, Department of Economics
(internal examiner)

The doctoral thesis is available for inspection
at the CEU Economics Department

Abstract

Recent advances in information technology have been constantly bringing down the barriers of collecting and managing data sets with sizes and representativeness unimaginable before. These data sets are typically arranged in the forms of panels, comprising tens of thousands, perhaps millions of entities, observed over a long time span. The new ways of data management, the comprehensive registry of transactions and other activities, and the attempts at the international harmonization of the data lead to the massive presence and direct accessibility of multi-dimensional panels.

The econometrics of standard, two-dimensional panel data is well-developed: it has been the subject of practically limitless research in the past fifty-sixty years. As much as efforts devoted to two-dimensional panels are admirable, multi-dimensional panels challenge analysts in several new ways. First, two-way models and toolsets are usually insufficient to fully describe and address problems in this three-dimensional context, where the unobserved heterogeneity can take on several new and interesting forms. Second, various new or existent, but increasingly present, data-related issues emerge, like feasibility of the estimators due to the sheer size of the data, incompleteness of observations, variable index deficiencies, or the large number of economically feasible model specifications.

Despite the massive presence of multi-dimensional data sets, the econometrics of three-dimensional panels remains grossly underdeveloped. Luckily, an increasing number of econometricians understand its importance, and aid empiricists with menus of modelling techniques and estimators capable of extracting the excess information embedded in the data. This thesis contributes to the literature by collecting several appealing model formulations, fixed effects, random effects and varying coefficients models, and proposing suitable estimation techniques. The comprehensiveness of the results lies in the diversity of issues discussed (both theoretical and data-related), and the fact that most techniques are feasible practically and so have a strong potential for empirical use.

Chapter 1: The Estimation of Multi-dimensional Fixed Effects Panel Data Models

Sections 1.2–1.6 are joint works with Laszlo Matyas and Tom Wansbeek, Sections 1.7 and 1.8 are solely my own.

The first chapter of the thesis formulates the excess heterogeneity in the data with fixed, observable parameters. Several such three-dimensional fixed effects models are collected from the literature, all of which correspond to empirically relevant cases. The models are estimated with Least Squares Dummy Variable (LSDV) estimator. In order to prevent the joint estimation of possibly (hundreds of) thousands

of parameters, the estimators are also expressed separately for each model parameter. It is also shown that the so-called *Within estimator*, which first wipes out the fixed effect parameters with a linear transformation, then performs a Least Squares on the transformed model, is numerically equivalent to the LSDV. The Within estimator reaches estimates at no costs, as long as the data at hand is complete. Typically, however, the data contains "holes". It is discussed how the Within estimator alleviates the dimensionality issue (the high cost of the estimation) completely, for structured incompleteness (like the no self-flow phenomenon), and partially, when it comes to handling incompleteness in general. This chapter also contributes to the literature by considering dynamic autoregressive specifications with fixed effects, first, by showing how the presence of various lags of the dependent variable violates the consistency of the Within estimator, then, by proposing Arellano-Bond-type instrumental variable estimators to correct for the arising inconsistency. Somewhat surprisingly, not all three-way model specifications carry this asymptotic bias. Eventual heteroscedasticity and the cross-correlation of the disturbance terms are also accounted for by proposing appropriate Feasible Generalized Least Squares (FGLS) estimators. The chapter ends with a generalization to four- and higher-dimensional fixed effect models, and intuitively argues that the results of the study can easily be generalized to any fixed effects specifications in any dimensions.

Chapter 2: Modelling Multi-dimensional Panel Data: A Random Effects Approach

Sections 2.2–2.4 are joint works with Badi H. Baltagi, Laszlo Matyas and Daria Pus, Sections 2.5 and 2.6.2–2.6.3 are joint works with Mark N. Harris, Felix Chan and Maurice Bun, Sections 2.6.1 and 2.7 are solely my own.

The second chapter of the thesis proposes several random effects model specifications. The chapter first assumes that the strict exogeneity assumption holds for the regressors, and derives optimal (F)GLS estimators for all models accordingly, discussing the estimation processes in depth. This is utterly important, as with the proposed methods the performed spectral decompositions and variance components estimations, needed for feasibility reasons and to complete the estimation process, can be easily generalized to any random effects model specification in any dimension. As the data can now grow in not only two, but three dimensions at the same time, it is crucial to collect the exact properties under which the FGLS estimator is consistent. Some of the consistency properties also carry a *convergence property*, which means that the FGLS estimator of a model converges to that model's specific Within estimator. For some models, consistency even *implies* convergence. While this phenomenon by itself does not violate the feasibility of the

estimators or their properties, the parameters of some fixed regressors – just like in case of fixed effects models – become unidentified, rendering the estimation of such parameters impossible. Apart from this identification problem, inconsistency in many of the several semi-asymptotic cases persists. To correct for this, so-called *mixed models* are proposed, combining both fixed and random components. One of the main reasons why random effects lag behind in popularity, is that the strict exogeneity assumption is hard to fulfill. The chapter also considers the case of endogenous regressors, and proposes Hausman-Taylor IV estimators to reach a full set of parameter estimates. The main results of the chapter are also extended to higher dimensions and to incomplete data, to argue for their wide applicability and easy generalizability. Finally, some basic insights on testing for random effects model specifications, for exogeneity, and for instrument validity are considered.

Chapter 3: The Estimation of Varying Coefficients Multi-dimensional Panel Data Models

The third chapter of the thesis considers several new varying coefficients models, and derives appropriate Least Squares estimators for them. The varying slope coefficients are assumed to be fixed, rather than random, and the slope parameters are assumed to comprise a universal part, common for all entities and time periods, as well as a varying component, which can be individual and/or time specific. In order to disentangle these two effects in these under-identified models, some parameter restrictions are to be assumed. As it turns out, the Least Squares estimation of the restricted model is simple theoretically, but cumbersome in practice due to the many complex functional forms and large matrices to work with. Further, as alternative parameter restrictions mean the full repetition of the calculation, alternative solutions are proposed. Luckily, the so-called *Least Squares of incomplete rank*, on the other hand, is easy to implement even in practice, and derives the part of the estimator which is model-specific before arriving at the restriction. In this way, the flexible exchange of various parameter restrictions is guaranteed. Some insights on the identification issues, and on the interpretation of models with variables with index deficits are considered, as well as some preliminary results on varying coefficient autoregressive models. Mixed coefficients models, having both fixed and random coefficients, are also briefly visited, and some of their estimation issues considered.

Chapter 4: Empirical Applications for Multi-dimensional Panels

Section 4.3 is joint work with Janos Kollo and Istvan Boza, Sections 4.1–4.2 are solely my own.

The fourth chapter of the thesis merges two distinct empirical studies employed

on three-way data: an international trade application, "*Regularities of Panel Estimators: A Trade Application*", and a study on wage returns, "*Contemporaneous and Lagged Wage Returns to Foreign-Firm Experience – Evidence from Linked Employer-Employee Data*". The former contributes to the literature by (i) comparing several fixed and random effects estimators, reflecting the typical estimation issues and some further regularities detailed in Chapters 1 and 2; (ii) by considering a new data set and taking into account data related issues, such as incompleteness, improving the results of several earlier papers which measured the effect of trade membership on real trade activity. The second study falls in line with several international studies capturing the (contemporaneous and lagged) wage returns of foreign experience on workers and on their colleagues. Foreign capital in emerging economies is subject to many criticisms, such as displacing local businesses, expatriating profits, or reducing tax liabilities. It is not clear, however, to what extent the domestic market gains from FDI. Apart from the fact that foreign wages are spent in the host country, and that domestic firms can imitate foreign-owned enterprises, workers of foreign-owned firms are usually more productive and are paid higher (contemporaneous effects). This wage premium can then be preserved when the worker re-enters the domestic market (lagged effect). Further, the presence of the accumulated knowledge of ex-foreign workers can also raise the productivity of their colleagues with no foreign experience (spillover effect). These advantages of FDI may in fact outweigh its losses. To elaborate on these ideas, several, mostly fixed effects models are formulated and regressed on a matched employer-employee data set covering half of the Hungarian working-age population.

Curriculum Vitae

Laszlo Balazsi

CONTACT

Address:	1118. Bakony utca 8., Budapest, Hungary
Phone:	+36 30/339 08 09
E-mail:	balazsi_laszlo@phd.ceu.edu
	lac.balazsi@gmail.com

EDUCATION AND DEGREES

2013-2017	Central European University, Hungary Phd in Economics, Department of Economics (to be defended in March, 2017) Supervisor: Laszlo Matyas Thesis Title: "The Econometrics of Linear Models for Multi-dimensional Panel Data"
2011-2013	Central European University, Hungary Master of Arts in Economics, Department of Economics Supervisor: Laszlo Matyas Thesis Title: "The Estimation of Multi-dimensional Fixed Effects Panel Data Models"
2008-2011	Budapest University of Technology and Economics, Hungary BSc in Mathematics, Institute of Mathematics Supervisor: Krisztina Kiss Thesis Title: "Cross-diffusion Modelling in Macroeconomics"
2004-2008	Eotvos Jozsef Gimnazium, Hungary (highschool)

WORKING PAPERS AND PUBLICATIONS

2017	"Random Coefficient Models", Laszlo Balazsi, Jaya Krishnakumar and Monika Avila Marquez, in Laszlo Matyas (ed.), *The Econometrics of Multi-dimensional Panels Theory and Applications*, Springer Verlag (forthcoming).
2017	"Models with Endogenous Regressors", Laszlo Balazsi, Maurice Bun, Felix Chan and Mark Harris, in Laszlo Matyas (ed.), *The Econometrics of Multi-dimensional Panels Theory and Applications*, Springer Verlag (forthcoming).

2017 "Linear Random Effects Models", Laszlo Balazsi, Badi Baltagi, Laszlo Matyas and Daria Pus, in LaszloMatyas (ed.), *The Econometrics of Multi-dimensional Panels Theory and Applications*, Springer Verlag (forthcoming).

2017 "Linear Fixed Effects Models", Laszlo Balazsi, Laszlo Matyas and Tom Wansbeek, in LaszloMatyas (ed.), *The Econometrics of Multi-dimensional Panels Theory and Applications*, Springer Verlag (forthcoming).

2016 "Contemporaneous and Lagged Wage Returns to Foreign-Firm Experience – Evidence from Linked Employer-Employee Data", Laszlo Balazsi, Istvan Boza, Janos Kollo *(Working Paper)*.

2015 "The Estimation of Multi-dimensional Fixed Effects Panel Data Models", Laszlo Balazsi, Laszlo Matyas and Tom Wansbeek, *Econometric Reviews*, DOI: 10.1080/07474938.2015.1032164 Published online: 07 April 2015, forthcoming in print.

Other versions available:

2015 "The Estimation of Multi-dimensional Fixed Effects Panel Data Models", Laszlo Balazsi, Laszlo Matyas and Tom Wansbeek, *CESifo Working Paper No. 5251, March 2015*.

2014 "The Estimation of Multi-dimensional Fixed Effects Panel Data Models" Revisited, Laszlo Balazsi, Laszlo Matyas and Tom Wansbeek, *CEU Working Papers No. 2014/1*.

2012 "The Estimation of Multi-dimensional Fixed Effects Panel Data Models", Laszlo Balazsi, Laszlo Matyas, *CEU Working Papers No. 2012/2*.

PUBLICATIONS IN HUNGARIAN

2016 "Spatial Economics, Cities, Regions, Trade and Environment; Survey from the COEURE-project III." (A gazdaság térbelisége, városok, régiók, kereskedelem és természeti környezetünk; Összegzés a COEURE-projekt tanulmányaiból III), Erika Domotor, Laszlo Matyas, Laszlo Balazsi, *Közgazdasági Szemle (Hungarian Economic Review)*, **LXIII**, July, 2016 (pp. 673-696).

2016 "Economic Growth, Crisis and Policy Making; Survey from the COEURE-project II." (Gazdasági növekedés, válság és szabályozás; Összegzés a COEURE-projekt tanulmányaiból II), Erika Domotor, Laszlo Matyas, Laszlo Balazsi, *Közgazdasági Szemle (Hungarian Economic Review)*, **LXIII**, May, 2016 (pp. 524-547).

2016 "The Human Aspects of Economics; Survey from the COEURE-project I." (Emberi tényező a gazdaságban; Összegzés a COEURE-projekt tanulmányaiból I.), Erika Domotor, Laszlo Matyas, Laszlo Balazsi, *Közgazdasági Szemle (Hungarian Economic Review)*, **LXIII**, April, 2016 (pp. 407-430).

2014 "Data Revolution and Panel Econometrics" (A közgazdasági adatforradalom és a panel ökonometria), Laszlo Balazsi, Janos Karoly Divenyi, Kezdi Gabor, Laszlo Matyas, *Közgazdasági Szemle (Hungarian Economic Review)*, **LXI**, November, 2014 (pp. 1319-1340).

CONFERENCE PRESENTATIONS

2016 "*Modelling Multi-dimensional Panel Data: A Random Effects Approach*"
31st Annual Congress of the European Economic Association, Geneva, August 22-26, 2016

2015 "*Modelling Multi-dimensional Panel Data: A Random Effects Approach*"
21th Panel Data Conference, Budapest, June 29-30, 2015

2013 "*The Estimation of Multi-dimensional Fixed Effects Panel Data Models*" REVISITED
19th Panel Data Conference, London, July 4-5, 2013

2012 "*The Estimation of Multi-dimensional Fixed Effects Panel Data Models*"
18th Panel Data Conference, Paris, July 5-6, 2012

ACADEMIC VISITS

2017, May	Curtin University, School of Economics and Finance, Visiting Researcher, working with Professor Mark Harris and Associate Professor Felix Chan
2017, January	University of Groningen, Faculty of Economics and Business, Visitor, working with Professor Tom Wansbeek
2016, September	Curtin University, School of Economics and Finance, Visiting Researcher, working with Professor Mark Harris and Associate Professor Felix Chan
2016, February	University of Groningen, Faculty of Economics and Business, Visitor, working with Professor Tom Wansbeek

2016, January University of Geneva, Institute of Economics and Econometrics,
Visitor, working with Professor Jaya Krishnakumar

WORK EXPERIENCE

2013-2016 **Research Experience**

Research Assistant for Janos Kollo, MTA-KRTK (2015, June - October)

Research Assistant, COEURE FP7 Research Project, Central European University (2013-2016)

2014-2016 **Teaching Experience**

ELTE University, Hungary (Fall, 2016)
Mathematical Methods for Economists (Lecturer)

Central European University, Hungary (2014-2015)
Quantitative Methods: Mathematics (Lecturer)
Advanced Microeconomics II. (Teaching Assistant; Lecturer: Prof. Botond Koszegi)
Advanced Microeconomics I. (Teaching Assistant; Lecturer: Prof. Adam Szeidl)
Mathematical Methods for Economics (Teaching Assistant; Lecturer: Prof. Peter Medvegyev)

OTHER EXPERIENCE

2016, August CERGE-EI Teaching Fellows Program
2011-2013 Recruitment Office, Central European University, Hungary
2013-2014 Student Representative, Central European University, Department of Economics

AWARDS AND HONORS

2016 Certificate for completing: Graduate Teaching Fellows Teaching Principles and Practices for Economics Course
CERGE-EI
2015 Academic Achievement Award for First-year Doctoral Students
Central European University
2013 Full CEU Doctoral Fellowship (2013-2017)
Central European University
2011 Full Tuition Scholarship (2011-2013)
Central European University

COMPUTER SKILLS

LaTeX, MatLab, Wolfram Mathematica, STATA, R

LANGUAGES SPOKEN

Hungarian (native)
English (fluent)
Italian (intermediate)

REFERENCES

Professor Laszlo Matyas
Central European University
matyas@ceu.edu

Professor Mark Harris
Curtin University
mark.harris@curtin.au.edu

Professor Tom Wansbeek
University of Groningen
t.j.wansbeek@rug.nl