Seven Decades of Econometrics and Beyond

A tribute to the life and work of Marc Nerlove

Editors: Badi H. Baltagi and Laszlo Matyas

Outline

Chapter 1: Klaus F. Zimmermann

Analysis of Business Survey Data: The Mannheim Years

In the 1980s, Marc Nerlove and my doctoral thesis supervisor Heinz König established an innovative joint project on “Business Survey Data Analysis”. It has been the period of a structural shift of applied econometrics towards the study of qualitative micro data. An international team guided by Nerlove made firm data available for microeconometric analysis for the first time. This paper documents the work of the research team, the challenges, spirits, and visions, and provides a review of the academic output achieved. It will reveal the leadership role of Nerlove, his work style, and personality documented by the work in the project and beyond. The author of the chapter has been part of the research team and was his academic guest at the University of Pennsylvania throughout 1987.

König, H., Nerlove, M., and Oudiz, G. (1981). On the formation of price expectations: An analysis of business test data by log-linear probability models, European Economic Review 16, 103-138.

Nerlove, M. (1983). Expectations, Plans, and Realizations in Theory and Practice. Econometrica 51, 1251-1279.

Zimmermann, K. F. (1997). Analysis of Business Surveys, in: M. H. Pesaran and P. Schmidt (Eds.), Handbook of Applied Econometrics, Volume II – Microeconometrics, Blackwell Publishers, Oxford, 407-441.

Chapter 2: John Rust

Predicting World Population with Endogenous Fertility and Uncertain Climate Change: Reflections on Marc Nerlove's Parable of Firewood.

The chapter reflects on Marc Nerlove's 1991 Frederick V. Waugh Memorial Lecture to the American Agricultural Economics Association that analyzed the endogenous dynamics of environmental degradation and population theoretically using an overlapping generations model. Nerlove reached the dismal conclusion that "the possibilities for a stable equilibrium between human population and its environment is quite limited". I review his predictions in light of subsequent developments over the past three decades that have resulted a sharp decline in fertility rates worldwide except in African and middle eastern countries and accelerating environmental damage due to unchecked CO2 emissions. My review of the empirical evidence suggests that Nerlove's lecture was prophetic.
Though his worst case prediction that "under exceptionally adverse environmental

circumstances, rising death rates can bring a halt to further environmental deterioration and/or lead to human extinction." may have seemed alarmist and extreme, I argue that humanity is unlikely to kick its addiction to fossil fuels and transition to green energy technologies quickly enough to avoid a period of massive human suffering and death (though not necessarily extinction). This is especially likely in African and Middle Eastern areas where the decline in fertility is not happening fast enough to offset the effects of global warming on food production in these areas. The chapter concludes by discussing whether technology will be able solve most of our problems including climate change, and lead humanity toward something akin to an utopia, or creates unexpected new challenges that lead humanity toward dystopia or even extinction. Nerlove's lecture did conclude on a more reassuring and optimistic note in this regard: "There is no doubt in my mind that introduction of physical capital formation to offset the environmentally adverse effects of population pressure and of human capital formation to enhance the quality of individual children would result in far more optimistic conclusions."

<u>Chapter 3</u>: Felix Chan, Elizabeth Jackson, Sylvia Soltyk, and László Mátyás

Re-estimating Supply Elasticities of Selected Agricultural Commodities

In his seminal work, Nerlove (1956) highlighted the importance of price expectations in farmers' decisions and demonstrated how these expectations may affect the estimate of supply elasticities. Like most seminal works, it inspired many years of research in both economics and agricultural economics, especially related to price expectations. However, his work was restricted by data and econometric techniques available at the time. Therefore, it seems appropriate to revisit the estimation of supply elasticities by leveraging the recent developments in econometrics, especially in risk modelling, machine learning, and the open data movement. As such, this chapter has three objectives. First, it presents a survey on the supply elasticities of selected agricultural commodities inspired by what is now known as the Nerlovian Model. This part of the chapter can be considered an update of Askari and Cummings (1977), which provided an excellent review of the work inspired by the Nerlovian model until the late 70's. Second, it extends the existing Nerlovian Model by incorporating the effect of price risk in the formulation of price expectation as well as the effect on acreage from the risk of different input prices. Perhaps more importantly, the proposed approach also incorporates the interaction between the different agricultural commodities, which couldn't have been possible in the 1950's due to lack of data availabilities and econometric techniques. Given the number of potential risk factors and different methods of generating risk estimates over different time horizons, the third objective of the chapter is to re-estimates the supply elasticities of selected commodities using the more recent selection techniques to identify the appropriate risk measures that affect price expectations.

Askari, H. and J. T. Cummings (1977, June). *Estimating Agricultural Supply Response with the Nerlove Model: A Survey*. International Economic Review 18 (2), 257.

Nerlove, M. (1956, May). *Estimates of the Elasticities of Supply of Selected Agricultural Commodities.* Journal of Farm Economics 38 (2), 496.

Chapter 4: Paul A. Bjorn, Isabelle Perrigne, and Quang Vuong

Discrete Games: A Historical Perspective

Starting from the seventies, economists have developed interests in the empirical analysis of qualitative variables. Following the statistical literature, Nerlove and Press (1973) pioneered the introduction of log-linear probability models. In parallel, McFadden (1974) initiated discrete choice models based on random utility, whereas Heckman (1978) introduced interactions among agents' discrete decisions extending classical simultaneous equations models. Since then, analysts have relied on game theoretic formulations to account for strategic interactions among agents. This chapter presents a historical perspective with Bjorn and Vuong (1984, 1985) first contributions to the econometrics of discrete games through non-cooperative solution concepts, namely Nash and Stackelberg equilibria. This structural approach to the empirical analysis of agents' binary decisions has led to a rich literature which continues to expand with applications to various domains of economics such as industrial organization, labor, public and development to name a few and beyond, in political science, marketing and management.

Chapter 5: Esfandiar Maasoumi and Yisroel Cahn

Measuring "Income" Inequality and Distribution of Outcomes

We provide a suggestive examination of the state of knowledge on measurement and analysis of "inequality" of outcomes, especially of incomes and earnings. This perspective aims to describe the state of art techniques for identifying the distribution of outcomes as the central object and consider interesting functions of it, such as inequality measures, poverty and mobility indices. We distinguish between "scalar", cardinal functions such as indices, as well as weak uniform rankings, such as by stochastic dominance based on modern rigorous tests. A basic theme permeates the discussion, that of decision theoretic foundations within the potential outcomes paradigm. This permits connectivity and advancement of knowledge that is policy relevant, reveals the essential subjectivity of indices and assessments, and allows important consideration of counterfactual distribution of outcomes. Identification of distributions and their functionals exposes the impact of covariates and different contributing factors to outcomes. A recurrent example is the distribution of earnings of different groups within a population, and decomposition of outcomes by group or other characteristics, and counterfactual states. Modern foundations of inequality measures, dominance rankings, quantile differences/effects, as well as quantile models, instrumental variables and "distribution regressions" are included and analyzed. The hope is to encourage and facilitate adoption of these important new developments at a time of heightened interest in this central socio-political area of policy analysis. The closely related notions of multivariate well-being, mobility and poverty, merit separate treatment and not addressed.

Chapter 6: Dibya Deepta Mishra, Robin C. Sickles, and Yanfei Sun

The Wizard of OZ (Opportunity Zones): Spatial Spillovers in Place Based Programs

The Opportunity Zones (OZ) program, as the largest ongoing place-based development program in the US, aims to stimulate investment and drive economic growth in low-income

areas by lowering capital gains tax rates. This paper investigates the spatial spillover effects of the OZ due to their interconnections with high-income neighboring areas. Using two-way fixed effects, synthetic difference-in-differences, and spatial difference-in-differences, we study the impact of OZ on housing prices and nighttime light emissions in the largest state by area in the continental US, Texas. Our empirical results indicate that census tracts located near more developed regions exhibit a stronger response to the OZ program due to the presence of spillover effects. One of the governing factors of these policies is the number of high-income neighbors. However, they play the role of a double-edged sword. A large number of high-income neighbors will make the tract in question not as attractive for investment, even in the presence of tax breaks. This is because the neighbors will provide higher returns. If a census tract is surrounded by some high-income neighbors and there is scope of future return, it may provide incentives for investing. We provide evidence of this trade-off in our paper, and also show how these effects should be considered carefully when designing place-based policies, especially when providing location-based tax breaks as in the Opportunity Zone program.

Chapter 7: Kajal Lahiri

Churns and Uncertainties in Coping with Health Shocks without Health Insurance: An Analysis Based on SIPP Panels over 2009-2016

Based on SIPP panel data over 2009-2016 and triple differencing method, we unravel two drastically different paths for the uninsured and for people with employer-sponsored health insurance (ESI) to cope with the economic consequences of health shocks. The uninsured individuals support their incomes mostly from other non-labor income sources, public transfers, and borrowing. The overall budget constraint of the uninsured tightens to the extent that they are not able to increase medical consumption enough and maintain necessary consumption in response to shocks. In contrast, people with ESI maintain most of their incomes by keeping their jobs and increasing their medical expenses without sacrificing necessary consumption. Overall, our results show that health shocks inflict significantly worse welfare consequences on the uninsured compared with those with health insurance, highlighting the economic and curative benefits of health insurance

Chapter 8: M. Hashem Pesaran and Ron Smith

Output Convergence: To What and How Fast?

There has been a resurgence of interest in the debate on whether output per-capita has been converging across countries. This paper revisits the output convergence debate by examining, how to model the unobserved steady state of global output (as a measure of the latent global technological factor) to which countries might be converging; how to estimate the effects of lagged country-specific output gaps on growth; and how the heterogeneity in the country-specific speed of convergence to their steady states influences the analysis. Almost all the convergence literature uses panel or cross-section estimators that assume a homogeneous speed of convergence. Not only is this assumption implausible on a priori grounds, many countries have neither the institutions nor the ability to catch up at the same speed, but the homogeneity assumption is also strongly contradicted by the data. The paper shows how

heterogeneity in the speed of convergence causes the usual two-way fixed effects estimators, that neglect this heterogeneity, to result in biased estimates and distorted inferences. The paper also presents new empirical results on output convergence using the Penn World Table data over different time spans, illustrating the sensitivity of the empirical estimates to the estimation methods adopted.

Chapter 9: Andrii Babii, Eric Ghysels and Jonas Striaukas

On Pooling Machine Learning Panel Data Regressions

The work by Marc Nerlove with Pietro Balestra on pooling panel data models is a foundational contribution to the field of econometrics. In this chapter we revisit their work and study structured machine learning regressions for heavy-tailed dependent panel data random effects using LASSO regularization. We take a closer look at the finite sample properties of various estimators put forward in the literature.

Chapter 10: Peter H. Egger and Michaela Kesina

Estimating Dynamic Binary-Choice Models with Higher-Order Network Effects

Estimating dynamic binary-choice models and nonlinear panel probability models with network effects based on standard frequentist methods is problematic. This article builds on ideas proposed by Wooldridge (2005) to rely on control-function estimation to tackle the so-called initial-conditions problem with dynamic binary-choice models in the presence of fixed effects. We address a class of estimating equations, where the concurrent latent outcome variable generating binary outcome of a cross-sectional unit depends not only on lagged own binary outcome but also on lagged outcome of network-linked “neighbors”. In particular, we emphasize the presence of multiple networks regarding the latter and, hence, the presence of so-called higher-order dynamic network effects. We propose variants of control functions towards estimating dynamic probit models with higher-order network effects. Moreover, we demonstrate by way of Monte Carlo simulations the small- and medium-sized sample performance of the proposed approaches.

Chapter 11: Cheng Hsiao, Qiankun Zhou and Yimeng Xie

Horizontal or Vertical Regression to Construct Counterfactuals?

Generating counterfactuals through treating a variable as a function of its own past values or treating a variable as a function of other units, typically being referred as horizontal or vertical regression, respectively, is widely used in the measurement of treatment effects. However, their inferences are often based on different assumptions for the data generating process. We consider unifying the underlying assumptions of the two approaches by a factor approach and considering their respective predictive power in terms of the sample configuration of cross-section dimension N and time dimension T.

Chapter 12: Daniel J. Henderson, Emma Kate Henry, and Alexandra Soberon

Nonparametric Correlated Random-Effects Models

This chapter proposes a method for the estimation of nonparametric panel data models with correlated random effects. Using the Mundlak specification to control for unobserved heterogeneity, this nonparametric estimation procedure can identify both the nonparametric function and a finite-dimensional parameter associated with (potentially) observed time-invariant regressors. We develop the necessary asymptotic theory for our proposed estimator. To assess the validity of our method in practice, we propose a consistent test for whether the model sufficiently controls for the correlation between the unobserved individual effects and the regressors. Monte Carlo simulations support the asymptotic developments. We illustrate the practical utility of our approach via an empirical application

Chapter 13: Maria Elena Bontempi and Jan Ditzen

The Correlated Random Effects GMM-lev Estimation: Monte Carlo Evidence and Empirical Applications

We introduce CRE-GMM, a new estimator that exploits sample splitting, correlated random effects (CRE) and the generalised method of moments on level equations (GMM-lev) in panel data. It efficiently estimates the effects of measurable time-invariant covariates and true persistence in dynamic models, overcoming the limitations of GMM-dif and GMM-sys. CRE-GMM considers explanatory variables possibly affected by double endogeneity (correlation with individual heterogeneity and idiosyncratic shocks), models initial conditions and improves inference. Monte Carlo simulations validate CRE-GMM across panel types and endogeneity scenarios. Empirical applications to R&D, production and wage functions illustrate the advantages of CRE-GMM.

Chapter 14: Badi H. Baltagi, Georges Bresson, and Jean-Michel Etienne

Serial Correlation in the One-way and Two-way Error Components Models and their Estimation using Whittle's Approximate Maximum Likelihood Method

This chapter surveys the estimation of one-way and two-way error components models with serial correlation in the remainder disturbances and proposes a new estimation method based on Whittle's (1953) approximate maximum likelihood method for an ARMA(p; q) in the remainder disturbances.

Chapter 15: Monika Avila Marquez and Jaya Krishnakumar

Dynamic Heterogeneous Linear Models for Three-level Panel Data

National survey data are mostly obtained through a sampling scheme that is stratified at multiple levels such as regions, socio-economic groups, gender, and so on, to ensure adequate representativity of the underlying population. In such a design, the clusters as well as the membership of individuals into a cluster are known. It is more and more the case that such data

are collected for the same units over time, thus resulting in a multilevel panel structure. This chapter studies a three-level dynamic panel data model – the levels being group (cluster), individual (household, firm etc.), and time - with additive cluster fixed effects and a mixed coefficient structure composed of cluster-specific fixed effects and random cluster-individual-time specific effects. We examine the identification and estimation of this dynamic heterogeneous three-level linear panel data model, for known clustering. We propose a Mean Cluster-FGLS estimator and a Mean Cluster-OLS estimator to estimate the mean coefficients. To make the GLS estimation of the cluster-specific parameters feasible, we introduce a ridge estimator of the variance-covariance matrix of the model. We show consistency and asymptotic normality of the Mean Cluster estimators for short panels, under the assumptions that apply to stratified sampling such as the number of clusters (groups) is fixed, all clusters are observed, and the number of individuals per cluster is large (growing to infinity). We also show the consistency of the variance parameter estimators. We discuss similarities and differences between our specification and a dynamic two-level panel data model with random coefficients. Finally, we discuss some extensions in particular unknown clustering and a long time span.

Chapter 16: Badi H. Baltagi and Tom Wansbeek

The Basics of the Mundlak and Chamberlain Projections

One of the best-known results in panel data econometrics, due to Mundlak (1978), is the equality of the random-effects and fixed-effects estimators when the individual effects are correlated with the means over time of the regressors. Chamberlain (1980) showed that the same result holds when the individual effects are correlated with the regressors for all moments in time separately. In this chapter, we review basic elements of the Mundlak and Chamberlain projections and we try to highlight the main ideas spawned by these projections. We emphasize the simplicity that is often obtained when the model is transformed into the within and between model, following Arellano (1993).

Chapter 17: Jeffrey M. Wooldridge

An Algebraic Equivalence between Generalized Fixed Effects and a Generalized Mundlak Regression with Applications to Heterogeneous Trends

The chapter establishes the equivalence between a class of generalized fixed effects estimators that remove multiple-dimensional heterogeneity and a generalized version of a Mundlak regression, which includes fitted values from unit-specific time series regressions. A special case is when the unit-specific time averages are included – the original Mundlak regression. The equivalence has several useful applications, including to testing whether the basic Mundlak (fixed effects) specification is rejected in favor of models with heterogeneous trends. The tests avoid the incidental parameters problem that arises if one tries joint testing of unit-specific dummy variables. The approach has implications when relaxing parallel trends in a difference-in-differences settings with staggered interventions.

Chapter 18: Jean-Marie Dufour and Lynda Khalaf

Simulation-based Finite Sample Tests in Simultaneous Equations

In simultaneous equation (SE) contexts, nuisance parameter and identification problems - resulting from weak instruments - severely complicate exact and asymptotic tests (except for very specific cases). In the past two decades or so, a sizable literature has endeavoured to develop procedures that are robust to these problems. In this chapter, we propose exact tests for possibly nonlinear hypotheses on the coefficients of SE systems. We discuss a number of bounds tests and Monte Carlo simulation based tests that are not restricted to Gaussian models. The latters involve maximizing a randomized p-value function over the relevant nuisance parameter space which is done numerically. Illustrative simulation experiments show that: (i) bootstrapping standard instrumental variable (IV) based criteria may fail to achieve size control, especially (but not exclusively) under near non-identification conditions; (ii) the tests based on IV estimates do not appear to be boundedly pivotal and so no size-correction may be feasible; (iii) by contrast, likelihood ratio type Monte Carlo tests work well in the experiments performed.

Chapter 19: Christian Gourieroux and Nour Meddahi

Dynamic Log-Linear Probability Models with Interaction

Log-linear probability models are parametric models introduced to capture the interactions between a set of binary variables [see, Nerlove and Press (1973), Liang and Zeger (1989)]. Our paper extends these models to a dynamic framework. This extension provides flexible specifications for high dimensional time series of binary variables and for associated processes of random sets, with interactions up to order 2. They allow for conducting a dynamic analysis of both cross-sectional and serial pointwise interactions, for instance.

We compare these models with the dynamic versions of the recursive logit model with latent qualitative variables, known as the Boltzman machine in the neural network literature [see, Achley et al. (1985)] and the Markov determinantal point processes [Gourieroux and Lu (2004)].

Then, we consider the applications to the dynamic analysis of business surveys, and to the improvement of chartist methods in predictive Finance.

Achley, Hinton, Sejnowski (1985): “A Learning Algorithm for Boltzman Machine”, Cognitive Science, 3, 147-169.

Liang, Zeger (1989): “A Class of Logistic Regression Models for Multivariate Binary Time Series”, JASA, 447-451.

Nerlove, Press (1973): “Univariate and Multivariate Log-Linear Logistic Models”, Vol 1306, Rand Corporation

Gourieroux, Lu (2023): “Markov Determinantal Point Process », CREST D.P.

Contributors

Badi H. Baltagi, Syracuse University, bbaltagi@syr.edu

Laszlo Matyas, Central European University, matyas@ceu.edu

Andrii Babii, University of North Carolina at Chapel Hill, babii.andrii@gmail.com.,

Paul A. Bjorn, UH Parma Cleveland,

Maria Elena Bontempi, University of Bologna, mariaelena.bontempi@unibo.it

Georges Bresson, Université Paris Pantheon-Assas, georges.bresson@u-paris2.fr

Yisroel Cahn, New York University, yisroel.cahn@nyu.edu

Felix Chan, Curtin University, Felix.Chan@cbs.curtin.edu.au

Jan Ditzen, University of Bolzano, jan.ditzen@unibz.it

Jean-Marie Dufour, McGill University, jean-marie.dufour@mcgill.ca

Peter H. Egger, ETH Zurich, pegger@ethz.ch

Jean-Michel Etienne, Université Paris-Saclay, jean-michel.etienne@u-psud.fr

Eric Ghysels, Department of Economics and Kenan-Flagler Business School, University of North Carolina, Chapel Hill, eghysels@unc.edu

Christian Gourieroux, University of Toronto and Toulouse School of Business, Christian.Gourieroux@ensae.fr

Daniel J. Henderson, University of Alabama, daniel.henderson@ua.edu

Emma Kate Henry, University of Alabama, ekhenry@crimson.ua.edu

Cheng Hsiao, University of Southern California, Dornsife, chowderlad@gmail.com

Elizabeth Jackson, Curtin University, Elizabeth.Jackson@curtin.edu.au

Lynda Khalaf, Carleton University, Lynda_Khalaf@carleton.ca

Michaela Kesina, University of Groningen, m.kesina@rug.nl

*Jaya Krishnakumar,*University of Geneva, Jaya.Krishnakumar@unige.ch

Kajal Lahiri, SUNY, Albany, klahiri@albany.edu

Esfandiar Maasoumi, Emory University, esfandiar.maasoumi@emory.edu

Nour Meddahi, Toulouse School of Business, nour.meddahi@tse-fr.eu

Monika Avila Marquez, Bristol University, monika.avilamarquez@bristol.ac.uk

*Dibya Deepta Mishra***,** Rice University/Cornerstone Research, Inc., ddm5@rice.edu

Isabelle Perrigne, Rice University*,* iperrigne@gmail.com

M. Hashem Pesaran, University of Southern California Dornsife, pesaran@usc.edu

John Rust, Georgetown University*,* jr1393@georgetown.edu

Robin C. Sickles, Rice University, rsickles@rice.edu

Ron Smith, Birkbeck Business School, r.smith@bbk.ac.uk

Alexandra Soberon, University of Cantabria, alexandra.soberon@unican.es

Sylvia Soltyk, Curtin University, S.Soltyk@cbs.curtin.edu.au

Jonas Striaukas, Department of Finance, Copenhagen Business School, Frederiksberg, Denmark, jonas.striaukas@gmail.com

Yanfei Sun, Toronto Metropolitan University, yanfei.sun@ryerson.ca

Quang Vuong, New York University, qvuong@nyu.edu

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

Jeffrey M. Wooldridge, Michigan State University, wooldri1@msu.edu

Yimeng Xie, Xiamen University, yimengxie@xmu.edu.cn

Klaus F. Zimmermann, Global Labor Organization and Bonn University, klaus.f.zimmermann@gmail.com.

Qiankun Zhou, Louisiana State University, qzhou@lsu.edu