

Netherlands Econometrics Study Group 2020 Program

Friday 15 May 2020

Organizer: *Tilburg University*, Tilburg School of Economics and Management

10:00 — 11:00 Keynote Lecture

Chair: Kees Jan van Garderen (University of Amsterdam)

Sophocles Mavroeidis (University of Oxford) - Identification at the Zero Lower Bound ([link](#))

Abstract

I show that the Zero Lower Bound (ZLB) on interest rates can be used to identify the causal effects of monetary policy. Identification depends on the extent to which the ZLB limits the efficacy of monetary policy. I develop a general econometric methodology for the identification and estimation of structural vector autoregressions (SVARs) with an occasionally binding constraint. The method provides a simple way to test the efficacy of unconventional policies, modelled via a 'shadow rate'. I apply this method to U.S. monetary policy using a three-equation SVAR model of inflation, unemployment and the federal funds rate. I reject the null hypothesis that unconventional monetary policy has no effect at the ZLB, but find some evidence that it is not as effective as conventional monetary policy.

11:00 — 11:15 Coffee Break

11:15 — 13:15 Session 1: Econometric Theory

Chair: Otilia Boldea (Tilburg University)

Ingrid Van Keilegom (KU Leuven) - Estimation of the boundary of a variable observed with symmetric error ([link](#))

With: Jean-Pierre Florens (Toulouse School of Economics) and Leopold Simar (UC Louvain)

Abstract

Consider the model $Y = X + \varepsilon$ with $X = \tau + Z$, where τ is an unknown constant (the boundary of X), Z is a random variable defined on \mathbb{R}_+ , ε is a symmetric error, and ε and Z are independent. Based on an iid sample of Y , we aim at identifying and estimating the boundary τ when the law of ε is unknown (apart from symmetry) and in particular its variance is unknown. We propose an estimation procedure based on a minimal distance approach and by making use of Laguerre polynomials. Asymptotic results as well as finite sample simulations are shown. The paper also proposes an extension to stochastic frontier analysis, where the model is conditional to observed variables. The model becomes $Y = \tau(w_1, w_2) + Z + \varepsilon$, where Y is a cost, w_1 are the observed outputs and w_2 represents the observed values of other conditioning variables, so Z is the cost inefficiency. Some simulations illustrate again how the approach works in finite samples, and the proposed procedure is illustrated with data coming from post offices in France.

Denis Kojevnikov (Tilburg University) - Limit theorems for network dependent random variables ([link](#))

With: Vadim Marmer (University of British Columbia) and Kyunghul Song (University of British Columbia)

Abstract

This paper focuses on the bootstrap for network dependent processes under the conditional ψ -weak dependence. Such processes are distinct from other forms of random fields studied in the statistics and econometrics literature so that the existing bootstrap methods cannot be applied directly. We propose a block-based approach and a modification of the dependent wild bootstrap for constructing confidence sets for the mean of a network dependent process. In addition, we establish the consistency of these methods for the smooth function model and provide the bootstrap alternatives to the network heteroskedasticity-autocorrelation consistent (HAC) variance estimator. We find that the modified dependent wild bootstrap and the corresponding variance estimator are consistent under weaker conditions relative to the block-based method, which makes the former approach preferable for practical implementation.

Yi He (University of Amsterdam) - Most powerful test against high dimensional free alternatives ([link](#))

With: Sombut Jaidee (Monash University) and Jiti Gao (Monash University)

Abstract

We propose a powerful quadratic test for the overall significance of many weak exogenous variables in a dense autoregressive model. By shrinking the classical weighting matrix on the sample moments to be identity, the test is asymptotically correct in high dimensions even when the number of coefficients is larger than the sample size. Our theory allows a non-parametric error distribution and the estimation of the autoregressive coefficients. Using random matrix theory, we show that the test has the optimal asymptotic testing power among a large class of competitors against local dense alternatives whose direction is free in the eigenbasis of the sample covariance matrix among regressors. The asymptotic results are adaptive to the predictors' cross-sectional and temporal dependence structure, and do not require a limiting spectral law of their sample covariance matrix. The method extends to general nuisance variables and we provide robustification for irregular scenarios. Monte Carlo studies suggest a good power performance against high dimensional dense alternative for various data generating processes. We apply the test to detect the overall significance of over one hundred exogenous variables in the latest FRED-MD database for predicting the monthly growth in the US industrial production index.

Jan Magnus (VU Amsterdam and Tinbergen Institute) - The perception of climate sensitivity: revealing priors from posteriors

With: Masako Ikefuji (University of Tsukuba and Osaka University)

Abstract

Everybody has priors, Bayesians and non-Bayesians alike. The priors may be vague and difficult to make explicit, but they are there and they may be important. The purpose of this paper is to show that we can make priors explicit from our knowledge of the data and the posterior, and to apply this theory to the perception of climate sensitivity. Imagine a group of people (the 'committee') with a collective prior, perhaps based on knowledge and experience, perhaps on political beliefs, perhaps on short-term profit. The committee meets privately and we have no information about their discussions. But we do have scientific data (official 'objective' statistics and scientific results) and we do have access to their published predictions or policy recommendations, which they present to the public. In other words, we have the data and the posterior,

but not the prior which the committee does not reveal and possibly may not even be able to formulate or quantify.

Given the data and the prior, standard Bayesian theory produces the posterior. But in practice the prior is often not observable, while the data and the posterior are. Can we recover the prior from the data and the posterior? Yes, this is indeed possible and we shall study the details of the recovered prior in some detail.

Then, we apply this theory to the estimation of the equilibrium climate sensitivity (ECS), which is an important diagnostic in climate modeling. The data come from various studies as listed and reported by the Intergovernmental Panel on Climate Change (IPCC). The IPCC then formulates its own conclusions (the posterior). What we are interested in is to recover the prior beliefs of the IPCC.

Our application is more than a mere illustration. We have made a serious attempt to recover the IPCC's priors. But the idea of reversing Bayesian thought and --- rather than obtaining a posterior from data and prior --- recovering the prior from data and posterior, does not seem to have received any attention. The current paper attempts to fill this gap.

13:15 — 14:15 Lunch Break

14:15 — 15:15 Poster Session

Chair: Jochem de Bresser

Ana Moura (Tilburg University) - Long-term care provision and hospital bed-blocking

Abstract

Hospital bed-blocking occurs when patients are medically ready to be discharged from a hospital but require some form of aftercare not readily available, resulting in longer hospital lengths of stay. I assess whether the entry of nursing homes and teams providing home-care in a region alleviates bed-blocking. Using individual data on emergency inpatient admissions at Portuguese hospitals for 2000-2015, I first show that there are specific social factors that put patients at risk of bed-blocking, such as living alone, having no family to care, and having inadequate housing. Then, I use a difference-in-differences design to compare the length of stay of patients at risk of bed-blocking and the length of stay of regular patients, before and after the entry of the first nursing home facility and the first home-care team in a region. I find that the entry of the first home-care team reduces the length of stay of individuals living alone and those with inadequate housing by 4 days. Reductions in length of stay following the entry of the first nursing home occur only for patients with high care needs. Despite these improvements, bed-blocking is not fully eliminated. Additional results suggest that coordination frictions between hospitals and long-term care providers might be driving this result.

Eva Janssens (University of Amsterdam) - Identification in heterogeneous agent models ([link](#))

Abstract

Many representative agent macro models suffer from parameter identification failure. These issues extend to the structural parameters of the standard incomplete markets model of Aiyagari (1994), the workhorse heterogeneous agent model for household and firm behavior. Similar as in representative agent models, the firm-side parameters can display weak or failure of identification. Strong identification of these parameters can, however, be obtained by using a transformation of the data that amplifies the effect of the parameters. The

household preference parameter can, nonetheless, not be identified using aggregate data alone, but it can be identified from the shape of the wealth distribution. This is in contrast with representative agent models, where

the identification of this preference parameter comes from time variation in aggregate data. Using these results, a two-step estimation procedure is proposed, which is analogous to the manner of identification. Its performance is confirmed in a Monte Carlo simulation and applied to US data.

Robert Adamek (Maastricht University) - Desparsified lasso in time series

With: Stephan Smeekes (Maastricht University) and Ines Wilms (Maastricht University)

Abstract

The Desparsified Lasso, introduced by Van de Geer et al. [Ann.Statist. 42 (2014) 1166-1202], is a high-dimensional estimation method which provides uniformly valid inference. In this paper, we extend this method to a time series setting under mixingale assumptions allowing for non-Gaussian, serially correlated and heteroskedastic processes, where the number of regressors can possibly grow faster than the time dimension. We first derive an oracle inequality for the (regular) Lasso, relaxing the commonly made exact sparsity assumption to a weaker alternative, which permits many small but non-zero coefficients. The weak sparsity coupled with the mixingale assumption means this inequality can also be applied to the (inherently misspecified) nodewise regressions performed in the Desparsified Lasso. This allows us to establish the uniform asymptotic normality of the Desparsified Lasso under general conditions. Additionally, we show consistency of a long-run variance estimator, thus providing a complete set of tools for performing inference in high-dimensional linear models. Finally, we perform a simulation exercise to demonstrate the small sample properties of the Desparsified Lasso in common time series settings.

Bram van Os (Erasmus University Rotterdam) - Accelerating peak dating in a dynamic factor Markov-switching model

With: Dick van Dijk (Erasmus University Rotterdam)

Abstract

The dynamic factor Markov-switching (DFMS) model introduced by Chauvet (1998) has proven to be a powerful framework to measure the business cycle. We extend the DFMS framework by allowing for time-varying transition probabilities, with the aim of accelerating the real-time dating of turning points between expansion and recession regimes. Time-variation of the transition probabilities is brought about endogenously using the accelerated score-driven framework and exogenously using the term spread. In a real-time application, using the four components of The Conference Board's Coincident Economic Index for the period 1959-2020, signaling power for recessions is significantly improved.

15:15 — 15:45 Coffee Break + Vote for Posters

15:45 — 17:15 Session 2: Machine Learning and Causal Inference

Chair: Otilia Boldea (Tilburg University)

Ines Wilms (Maastricht University) - Sparse regression for large data sets with outliers ([link](#))

With: Lea Bottmer (Stanford University) and Christophe Croux (EDHEC Business School)

Abstract

The linear regression model remains an important workhorse for data scientists. However, many data sets contain many more predictors than observations. Besides, outliers, or anomalies, frequently occur. This presentation proposes an algorithm for regression analysis that addresses these features typical for big data sets. The resulting regression coefficients are sparse, meaning that many of them are set to zero, hereby selecting the most relevant predictors. A distinct feature of the method is its robustness with respect to outliers in the cells of the data matrix. Simulation results are discussed to show that the proposed robust variable selection method performs very well compared to several benchmarks.

Andrea Nagy (Erasmus University Rotterdam) - The value added of machine learning to causal inference: evidence from revisited studies

With: Anna Baiardi (Erasmus University Rotterdam)

Abstract

A new and rapidly growing econometric literature is making advances in the problem of using machine learning (ML) methods for causal inference questions. The empirical economics literature, however, has not started yet to fully exploit the strengths of these modern causal inference methods, perhaps because the value added of these new tools remains unclear to applied researchers. This paper investigates with concrete empirical examples when and how ML methods can improve economic causal analysis, by revisiting some well-known empirical economics papers from a variety of topics in applied economics, and reanalyzing their main results with causal ML methods. Our focus is on both average and heterogeneous treatment effects. Several unifying aspects with which ML methods add value to traditional econometric methods for causal inference emerge: 1) causal ML methods are suitable and perform well in empirical settings where the number of potential covariates is large; 2) they can recover complex nonlinearities and estimate more flexibly the relationship between the outcome, the treatment and covariates, compared to traditional causal methods; 3) they are useful tools in implementing systematic model selection and offer a solution to the multiple hypothesis testing issue; 4) they can identify heterogeneous treatment effects that are missed by standard causal methods. Monte Carlo simulations inspired by empirically relevant settings support the main findings.

Martin Schumann (Maastricht University) - Difference-in-differences estimation under non-parallel trends

With: Holger Dette (Ruhr-University Bochum)

Abstract

Classic difference-in-differences estimation relies on the validity of the "parallel trends assumption" (PTA), which ensures that the evolution of the variable of interest in the control group can be used to determine its counterfactual development in the treatment group in the absence of treatment. The plausibility of the PTA is usually assessed by a test of the null hypothesis that the difference between the means of both groups is constant over time before the treatment. However, this procedure is problematic as failure to reject the null hypothesis does not imply the absence of differences in time trends between both groups due to low power to detect economically relevant differences. We provide three tests of equivalence leading to a "common range" (CR) condition that replaces the PTA. The tests are designed such that the tendency of standard equivalence tests to be overly conservative is mitigated. We show that the width of the CR naturally reflects differences between treatment and control and provide examples of situations in which the CR can be used to set-identify the average treatment effect on the treated when the PTA does not hold.

17:15 — 17:30 Announcement winning poster and closing statement.