Essays on Nonparametric Methods in Econometrics

Takahide Yanagi
Graduate School of Economics, Kyoto University

Abstract:

This dissertation develops three research topics on nonparametric methods in econometrics. First, we present a nonparametric panel data analysis in which dynamic structure of panel data is heterogeneous across individuals. Second, we develop a nonparametric analysis for the sharp regression discontinuity design in which the forcing variable contains measurement error. Third, we study a nonparametric identification analysis in weakly separable models with a binary endogenous explanatory variable.

In the first chapter of the dissertation, we provide a brief review of the literature of nonparametric methods in econometrics. We first explain the reasons why the nonparametric methods are important research consideration in econometrics. There are at least three reasons of this: (i) the nonparametric analyses do not generally force the researchers to assume implausible or restrictive assumptions, (ii) the parametric analyses are sensitive to the problems caused from misspecifications, while the nonparametric analyses are robust to such problems, and (iii) the nonparametric methods allow us to examine the correctness of the parametric specifications. We also describe an outline of the dissertation in the first chapter and the main contribution of the dissertation on the literature of nonparametric econometrics.

In the second chapter (joint with Ryo Okui), we propose the nonparametric analysis of panel data whose dynamic structure is heterogeneous across individuals. The aim in the chapter is to estimate the cross-sectional distributions and/or some distributional features of the heterogeneous mean and autocovariances without assuming any specific model for the dynamics. The asymptotic properties of the proposed estimators are investigated using double asymptotics under which both the cross-sectional sample size and the length of the time series tend to infinity. We prove the functional central limit theorem for the empirical process of the proposed distribution estimator. By using the functional delta method, we also derive the asymptotic distributions of the estimators for various parameters of interest. We show that the distribution estimator exhibits a bias, which is critical when the length of the time series is shorter than the cross-sectional sample size. Conversely, when the parameter of interest can be written as the expectation of a smooth function of the heterogeneous mean and/or autocovariances, the

bias of the order reduces and, furthermore, can be corrected by the jackknife method. The results of Monte Carlo simulations show that our asymptotic results are informative regarding the finite-sample properties of the estimators. They also demonstrate that the proposed jackknife bias correction is successful.

In the third chapter, we develop a nonparametric analysis for the sharp regression discontinuity (RD) design in which the continuous forcing variable may contain measurement error. We show that if the observable forcing variable contains measurement error, the measurement error causes severe identification bias for the average treatment effect given the "true" forcing variable at the discontinuity point. The bias is critical in the sense that even if actually there is significant causal effect, it misleads the researchers into the incorrect conclusion in which there is no causal effect. Furthermore, the measurement error leads the conditional probability of the treatment to be continuous at the threshold. To investigate the average treatment effect using the mismeasured forcing variable, we propose approximating it by the small error variance approximation (SEVA) originally developed by Chesher (1991). Based on the SEVA, the average treatment effect is approximated up to the order of the variance of the measurement error by an identified parameter when the variance is small. We also develop an estimation procedure for the parameter that approximates the average treatment effect based on local polynomial regressions and the kernel density estimation. Monte Carlo simulations reveal the severity of the identification bias due to the measurement error and demonstrate that our approximate analysis is successful.

In the fourth chapter, we present a new identification analysis for the structural function in weakly separable models (Vytlacil and Yildiz, 2007) with a binary endogenous explanatory variable. We focus on identifying the values of the structural function at specified values of the explanatory variables and specific quantiles of the unobservables, in the manner of Chesher (2005) and Jun, Pinkse, and Xu (2011). Our identification analysis is composed of two steps. Firstly, under weak conditions, we provide partial identification results for the values of the structural function based on the modified control variate approach in Chesher (2005). Secondly, by utilizing the results of the first step and the weak separability of the structural function, we show that the values of the structural function are intervally identified by tighter bounds. This second step is based on, using the instrumental variables, identifying the values of the explanatory variables that lead to the values of the structural function being no greater or no less than the value of the structural function we wish to identify. A simple example demonstrates that the identification analysis leads to an informative identification bound for a value of the structural function.