Specification Tests in Econometrics and Their Application
Masamune Iwasawa

Digest

Parametric models are indispensable in most empirical studies of economic events. Estimation results obtained by parametric models may, however, be seriously misleading when the model is misspecified or poorly approximates the true model. This thesis presents two new methods for testing the model specification and their application.

The first chapter proposes a joint specification test that is directly applicable to multinomial choice models with unordered outcome variables. Variables of interest in economic research are often discrete and unordered. Thus, parametric models such as the multinomial logit and probit models proposed by McFadden (1974) and Hausman and Wise (1978), respectively, are widely employed, for example, in structural econometric analysis and as part of econometric methods. These models are referred to as multinomial choice models and consist of multiple response probabilities, each of which may be parameterized differently. This implies that one needs to test multiple null hypotheses to justify the parametric assumptions of these models. The test provided in this chapter allows us to test the specifications of response probabilities jointly for all choice alternatives. The test proposed here is based on moment conditions. We show that the test statistic is asymptotically chi-square distributed, consistent against a fixed alternative, and able to detect a local alternative approaching the null at the rate of $1/\sqrt{nh^{q/2}}$, where n is the sample size, h is a bandwidth for kernel estimation, and q is the number of independent variables.

A simple parametric bootstrap procedure to calculate rejection regions is a practical need because the testing method involves nonparametric estimation and a sufficiently large sample size could be required to establish that the chi-squared distribution is a proper approximation for the distribution of the test statistic. We show that rejection regions can be calculated by a simple parametric bootstrap procedure in the small sample case. A crucial point that makes parametric bootstrap work is that the orthogonality condition holds with bootstrap sampling under both the null and alternative hypotheses. This is different from the specification test for the regression function that requires the wild bootstrapping procedure to calculate the rejection region proved by Härdle and Mammen (1993). It is also noteworthy that the parametric nature of the model leads to substantial savings in the computational cost of bootstrapping. Monte Carlo experiments show that the test has proper size and reasonable power in finite samples.

The second chapter (joint with Kohtaro Hitomi and Yoshihiko Nishiyama) proposes a consistent specification test for instrumental variable models. Empirical economic studies widely employ instrumental variable models to address endogeneity. The test is constructed by using the nearest neighbor approach. We show that the test has prescribed non-trivial power uniformly against a set of alternatives defined on a specific smoothness class that can approach the null hypothesis at the rate $n^{-1/4}$. Our proposed test entails no kernel smoothing and no integration while maintaining consistency against all departures from the null hypothesis. Since the test statistic entails no kernel smoothing and no integration, the test is easily implemented, requires less calculation time, and has no bandwidth choices. The test statistic can be obtained via straightforward calculation that requires only parametric model estimation under the null hypothesis and locating the neighbors nearest to each observation.

We demonstrate the power property of our test by employing Ingster's (1993) minimax approach in which the alternative hypothesis is considered in a set of functions belonging to a smoothness class. This study considers the same smoothness class with Guerre and Lavergne (2002) and shows

that the optimal minimax rate for the instrumental variable model is the same as for nonlinear regression models. The optimal rate remains unchanged under our result even when we allow the model to be linear. Our test can uniformly detect a set of alternatives featuring functions belonging to the small smoothness class, when we confine the set of alternatives to a cone set.

In the third chapter (joint with Seiichi Fukui and Mitsuo Inada), we investigate the migrant heterogeneity in the effect of remittances on child education. The large amount of remittance flows to developing countries and increasing migration have both drawn significant scholarly attention to their role in influencing economic development. In particular, their effect on child education is a crucial research question because the socioeconomic environment in which a large proportion of the population has a low educational level can eventually limit future economic growth. However, measuring the effects of remittances is challenging and demands great caution. Their theoretical positive impacts can be partly or fully offset by the adverse influences of family members' migration that may depend on migrant characteristics. This study explores how migrant heterogeneity, that is, parental, nonparental, and no migration, differently affects the positive effects of remittances on the left-behind child's education.

Our results show that remittances increase educational expenditure for children aged between 3 and 15 years. By contrast, migration decreases this expenditure and the positive effects of remittances are partially cancelled out for nonparental migration and completely eliminated when parental migration occurs. This results may highlight that the countervailing effects of parental migration are caused by a labor shortage in the household as well as insufficient parental input into work/education decision-making process for children. Although our study makes a significant contribution to the better understanding of how migrant heterogeneity in the effects of remittances affects child's education, the application of the specification tests proposed in Chapter 1 and 2 suggests possible limitation of this study in terms of model specification and parameter identification.

References

Guerre, E., & Lavergne, P. (2002). Optimal minimax rates for nonparametric specification testing in regression models. *Econometric Theory*, *18(5)*, 1139–1171.

Härdle, W., & Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *The Annals of Statistics*, *21(4)*, 1926–1947.

Hausman, J. A., & Wise, D. A. (1978). A conditional probit model for qualitative choice: Discrete decisions recognizing interdependence and heterogeneous preferences. *Econometrica*, *46(2)*, 403–426.

Ingster, Y. I. (1993). Asymptotically minimax hypothesis testing for nonparametric alternatives. I, II, III. *Mathematical Methods of Statistics*, *2(2)*, 85–114.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (p. 105-142). New York: Academic Press.