Introduction

The chapters in this volume bring together important recent advances in the areas of (i) econometrics of panel data, (ii) limited dependent variable models, and (iii) limited dependent variable models with panel data. Panel data offers researchers many more possibilities than pure cross-sectional or time series data. Like cross-sectional data, panel data describes each of a number of individuals. Like time series data, it describes changes through time. By blending characteristics of both cross-sectional and time series data, panel data can be used, for example, to (i) expand sample size, (ii) allow the specification of more complicated behavioral hypotheses, (iii) capture cross-sectional variation and dynamic behavior, (iv) lessen the problem of multicollinearity, (v) provide possibilities for reducing omitted variable and estimation biases, (vi) improve accuracy of parameter estimates, (vii) obtain more accurate prediction of individual outcomes. However, the analysis of panel data also raises a number of new issues. For instance, in the case of short dynamic panel data models with large cross-section units, it is known that dealing with the initial values and incidental parameters problem can be complex. In other applications, such as non-linear panel data models with fixed effects, a general solution to the problem may not exist. A notable example is estimation of probit models with fixed effects. Also panel data based on economic surveys are very often qualitative in nature, and have limited variations due to self-selection, and truncation.

This collection focuses on the issues of simplifying complex real-world phenomena into easily generalizable inference from individual outcomes. Since Maddala's contributions in the fields of limited dependent variables and panel data have been particularly influential, it is a fitting tribute to his legacy that we dedicate this volume to him.

Professor G.S. Maddala is one of the leading figures in the econometrics profession and has made highly influential contributions covering almost every area of econometrics. He has been an unerring source of wise counsel to a generation of students, colleagues, and journal editors who have come in contact with him. Moreover, Maddala writes econometrics in plain English. He likes to convey the basic ideas in simple words. His main econometrics textbook, in various editions, has been an important source for the training of students. His Econometric Society monograph on *Limited Dependent and Qualitative Choice Variables* is among the most cited technical books on econometrics. Professor Maddala has also provided invaluable service to practicing econometricians through a large number of timely surveys covering a wide range of topics, including the analysis of sample selectivity bias as it pertains to health care markets, the econometrics of panel data models, limited dependent variable models using panel data, the analysis of expectations using survey data, a perspective on the use of limited dependent and qualitative variable models in accounting research, specification tests in limited dependent variable models, structural change and unit roots, and bootstrapping time series models. These surveys not only summarize the state of the art at the time but have been sources of inspiration. A complete list of these surveys and other publications of Professor Maddala is provided at the end of this volume.

The chapters in this collection can be grouped into two broad categories. The chapters by Amemiya; Arellano, Bover and Labeaga; Geweke and Keane; Lee; and El-Gamal and Grether primarily deal with different aspects of limited dependent variable models and sample selectivity. The second group of papers by Nerlove; Ahn and Schmidt; Kiviet; Davies and Lahiri; Baillie and Baltagi; Hsiao, Pesaran, and Tahmiscioglu; and Pesaran and Zhao consider issues that arise in estimation of dynamic (possibly) heterogeneous panel data models.

The two chapters by Amemiya, and Arellano, Bover and Labeaga consider how to take account of selectivity or censoring issues using panel data. Data censoring can create much difficulty in estimation and inference because of the unobservability of the true state. Surprisingly, panel data can sometimes make an inherently intractable problem easier to solve. The chapter by Amemiya provides a unified treatment of a duration model in which left censoring arises because (i) spells in the middle of continuation at the time of the first observation are either completely observed or partially observed; (ii) spells which start after the time of the first observation are either observed or not observed; or (iii) for a single individual we either observe a single spell or a sequence of spells in different states. Amemiya derives the maximum likelihood estimator when these models are fully specified. He also shows that, in certain situations, a less efficient but more robust method, which does not require the full knowledge of the model specification, may be possible and desirable.

The chapter by Arellano, Bover and Labeaga considers an autoregressive method with random effects for a latent variable which is only partly observed due to a selection mechanism. They show that the intractability of a dynamic model subject to censoring using a single time series can be successfully overcome by noting that the sub-samples of the panel data that only included individuals without censored past observations are exogenously selected. They suggest an easy to implement asymptotic least squares method to estimate features of the distribution of the censored endogeneous variable conditional on its past. They also apply these methods to analyze the dynamics of female labor supply and wages using PSID data.

Geweke and Keane consider the binary choice model where the distribution of the underlying disturbances is assumed to be a mixture of normal densities, which is an important generalization of the standard normal probit model widely used in the literature. The mixture normal specification, by allowing mixing on both the mean and the variance parameters, and by increasing the number of the distributions in the mixtures provides a highly flexible formulation, thus enabling the researcher to explore the possible effects of a wide range of departures from the standard normal probit model. The chapter implements a Bayesian approach showing how Gibbs sampling techniques can be used to carry out the necessary computations. Geweke and Keane contrast their approach to the semiparametric methods developed in the literature for the estimation of the parameters of the probit model, and discuss the pros and cons of their procedure as compared with that of the semiparametric methods. The finite sample performance of the estimation procedure is studied by means of a number of Monte Carlo experiments. A substantive empirical application is also provided where women's labor force participation is investigated using a subset of data from the Panel Study of Income Dynamics.

The chapter by Lung-Fei Lee considers the estimation of limited dependent variable models under rational expectations in the time series context. Serial correlation in disturbances and dynamic structures with lagged dependent variables are considered and incorporated in the estimation. He shows that the simulated maximum likelihood method is feasible for the estimation of such models. A general simulation method with broad applicability is suggested. It is proved that a unique rational expectations solution exists even when the equations characterizing the rational expectations solution are simulated. For a long time series, the potential numerical underflow issue in the simulation of rational expectations solution and likelihood function can be solved with a recursive weighting simulation scheme. Variance reduction in the simulation is possible for models with renewal property. Lung-Fei Lee conducts a number of Monte Carlo experiments to study the finite sample performance of the proposed estimation method.

The chapter by El-Gamal and Grether provides a Monte Carlo study on finite sample performance of their EC (estimation-classification) estimator and algorithm for panel data probit models. The situation under investigation is that each observed individual in a panel may belong to one of a fixed but possibly unknown number of types. The study has found that the EC-estimator can be better than the familiar fixed effects estimator. A diagnostic statistic called the average normalized entropy is also found to be a very useful indicator of possible misclassifications.

The second group of papers deal with panel data models. The chapter by Nerlove re-examines the estimation of dynamic panel data models and studies the sensitivity of the coefficient estimates of both the "state" variable and the other explanatory variables to the econometric method employed. He examines this sensitivity in the context of recent empirical studies of growth rate convergence using panel data from the Penn World Tables. Models with country-specific intercepts and models with country-specific trends are estimated. Even though the primary purpose of the chapter is to assess the performance of alternative estimators, all the results reported support the conventional interpretation of the coefficient of the lagged dependent variable in terms of growth convergence conditional on savings and population growth rates. He shows that the use of the fixed-effects estimator favors the results toward finding a relatively rapid convergence. However, when the maximum likelihood estimation technique is employed, unconditional on the initial observations, very slow convergence is obtained. Biases in the estimates of the coefficient of the "state" variable for all of the usual methods of panel data analysis tend to induce biases in the estimates of the coefficients of other variables as well. Consequently, Nerlove argues that the conclusions of many of the recent studies of the determinants of growth employing dynamic panel data models may largely reflect the econometric methods employed.

In their chapter Ahn and Schmidt consider efficient use of moment conditions in panel data models. In panel data models with strictly exogenous time-varying regressors, the number of moment conditions rapidly increases with the number of time series observations. In terms of asymptotic efficiency, it would always be desirable to use as many moment conditions as possible. However, in finite samples, the biases in generalized method of moments (GMM) estimates tend to increase with the number of moment conditions used. Ahn and Schmidt derive conditions to identify redundant moment conditions. They also propose a modified generalized instrumental variable (MGIV) estimator that is asymptotically equivalent to the GMM estimator when the errors are conditionally homoskedastic. When the errors are conditionally heteroskedastic, the MGIV estimator is less efficient than the full GMM estimator asymptotically. However, their

Monte Carlo results suggest that, in finite samples, the MGIV estimator with heteroskedasticity adjusted asymptotic standard errors performs better than the full GMM estimator.

The chapter by Kiviet provides asymptotic expansions of least squares and instrumental variables estimators for dynamic panel data models. The dynamic panel data model includes lagged dependent variables and a weakly exogenous regressor. Analytic results are obtained by deriving the expectation of higher-order expansions of estimation errors. Those analytical results have strong implications on small sample properties of various estimators of these models.

The chapter by Davies and Lahiri is concerned with the analysis of expectations from the surveys of professional forecasters carried out by American Statistical Association – National Bureau of Economic Research (ASA–NBER). They provide a generalization of the panel data model used by Keane and Runkle (1990), which allows for a more complex correlation structure across the forecast errors over different individuals, target dates, and at different horizons. Within this framework they examine a number of issues raised by Keane and Runkle, notably the problem associated with the use of revised July figures (that could not have been available to the respondents), the proper accounting of aggregate shocks, and the appropriate choice of the forecast horizon. Based on their test results they conclude that on the whole the hypothesis that the ASA-NBER panel have been rational in predicting the inflation rate over the period 1968(4)–1991(4) is rejected. According to their analysis more than 70 percent of the forecasters failed to meet the rationality criteria in the sense of Muth (1961).

The chapter by Baillie and Baltagi considers prediction problems for the regression model with one-way error component disturbances. The expression for the asymptotic mean squared errors of prediction from various predictors are derived. Theoretical and simulation results indicate that it is important to allow for parameter uncertainty in forming prediction confidence intervals.

The chapter by Hsiao, Pesaran, and Tahmiscioglu implements a Bayesian approach to estimate dynamic panel data models when the coefficients are assumed to be randomly distributed across cross-sectional units using Markov Chain Monte Carlo (MCMC) methods. They establish the asymptotic equivalence of the Bayes estimator and the mean group estimator of Pesaran and Smith (1995), and show that the Bayes estimator is asymptotically normal for large N (the number of units) and large T (the number of time periods) so long as $\cdot N/T \to 0$ as both N and $T \to \infty$. The performance of the Bayes estimator for the short-run coefficients is compared against alternative estimators using both simulated and real data. The Monte Carlo results show that the Bayes estimator has better sampling properties than other estimators for both small and moderate T samples. The analysis of the real data yields new results on Tobin's q model.

The problem of estimation for the long-run coefficients in dynamic heterogeneous panels is taken up in the chapter by Pesaran and Zhao. This chapter investigates the finite T bias of the estimators of the mean long-run coefficients of a heterogeneous dynamic panel. Three approaches of bias corrections are applied to derive individual long-run coefficients before taking their averages. The first approach applies to the Kiviet–Phillips bias correction to individual short-run coefficients before deriving their long-run coefficients. This is referred to as "naive" bias corrected (NBC) procedure. The second approach makes the bias correction directly to the individual long-run coefficients. Two variants are considered, DBC_1 and DBC_2. Both are unbiased to order $O(T^{-1})$ but the DBC_1 estimator contains some higher-order bias corrections. The third approach uses the average of the replica of the long-run coefficients derived from bootstrap generated short-run coefficients as the individual long-run coefficients ($BSBC$). Monte Carlo studies are conducted to evaluate the effectiveness of these bias-correction procedures in reducing the small sample bias. It is found that the NBC procedure fails in all cases. The $BSBC$ performs poorly in cases where the true coefficient of the lagged dependent variable is relatively large. The DBC_2 performs reasonably well, although only the DBC_1 outperforms the bootstrap method. When the coefficient of the lagged dependent variable is around 0.8 or above, none of the above-mentioned estimators seems to work.

References

Keane, M.P. and D.E. Runkle (1990), "Testing the Rationality of Price Forecasts: New Evidence from Panel Data," *American Economic Review*, 80: 714–735.

Muth, J.F. (1961), "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29: 315–355.

Pesaran, M.H. and R.P. Smith (1995), "Estimating Long-Run Relationships from Dynamic Heterogeneous Panels," *Journal of Econometrics*, 68: 79–113.

1 A note on left censoring

TAKESHI AMEMIYA

1 Introduction

Left censoring occurs in a duration model when a statistician observes only those spells which either are in the middle of continuation at the time of the first observation or start during the observation period. It is assumed that the statistician has no record of those spells which had ended by the time of the first observation. A special treatment of the problem is necessary because ignoring left censoring will overestimate the mean duration as longer spells tend to be observed more frequently than shorter spells. This is called selectivity bias.

Different cases of left censoring arise depending on the following considerations: (1) Spells in the middle of continuation at the time of the first observation are either completely or partially observed. Suppose such a spell started at s, continued on to 0 (the time of the first observation), and ended at t. The statistician may observe only s (by asking how long the spell had lasted), only t, or both. (2) Spells which start after the time of the first observation are either observed or not observed. (3) For a single individual we either observe a single spell or a sequence of spells in different states.

In each possible case we will consider how the selectivity bias is eliminated. If the model is fully specified, this is accomplished by the method of maximum likelihood estimation, which is fully efficient. However, in certain situations, a less efficient but more robust method, which does not require the full knowledge of the model specification, may be possible and desirable.

Although we treat the case of a homogeneous population, the adjustment for a heterogeneous population is simple as it will be indicated in appropriate places.

The problem of left censoring is dealt with only scantily in the general

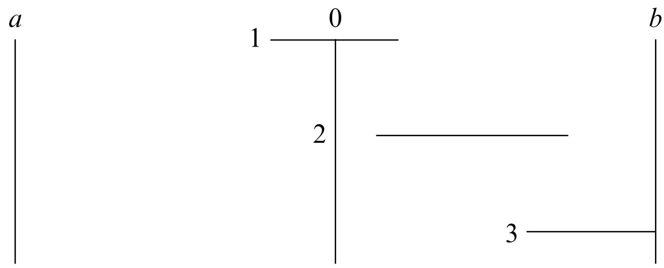
The first version of this chapter appeared as CEPR Publication No. 235 on March 1991. The author is indebted to Aaron Edlin and Tony Lancaster for helpful comments.

statistical literature. For example, standard textbooks on duration analysis such as Kalbfleisch and Prentice (1980) or Cox and Oakes (1984) devote less than a page to the problem. Miller (1981) mentions only a different kind of left censoring from what we discuss here. One can find more discussion in the econometric literature (for example, see Lancaster (1979), Flinn and Heckman (1982), Ridder (1984), and Amemiya (1985)). Here we try to give a more complete, unified treatment of the subject.

2 A single state model

The duration data are generated according to the following scheme: a duration starts in an interval $[a, b]$, which encloses 0, and the starting time X is distributed according to density $h(x)$. Duration T is distributed according to density $f(t)$ and distribution function $F(t)$. We assume that X and T are independent. The statistician observes only those spells which end or are censored after 0. We will consider three types of left censoring and for each type will derive the likelihood function assuming a homogeneous population. The result can be easily modified for the case of a heterogeneous population, as we will indicate below.

Type 1 left censor



Here the spell that was going on at time 0 is completely observed. Three kinds of spells are depicted in the above figure; we will write the likelihood function as a product of three parts corresponding to the three kinds. Each part is to be divided by the probability of observing a spell. Define

$$A_1 = \{x,t | t > -x,\ 0 > x > a\} \text{ and } A_2 = \{x,t | x > 0\}.$$

Then

$$P_1 \equiv P(A_1) = \int_a^0 h(x)[1 - F(-x)]dx \tag{1}$$

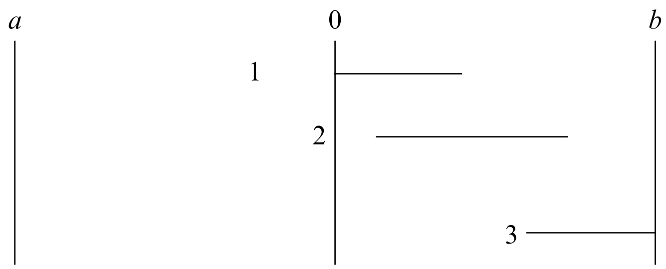
$$P_2 \equiv P(A_2) = \int_0^b h(x)dx. \tag{2}$$

The probability of observing a spell, denoted by P, is $P_1 + P_2$. Finally, the likelihood function can be written as

$$L_1 = \prod_1 h(x_i)f(t_i)\prod_2 h(x_i)f(t_i)\prod_3 h(x_i)[1 - F(b - x_i)]\prod_{all} P^{-1}. \tag{3}$$

Note that the first and second kinds of spells are treated symmetrically. In the next section we will show that dividing the first part by P_1 and the second and third part by P_2 leads to a consistent but less-efficient estimator.

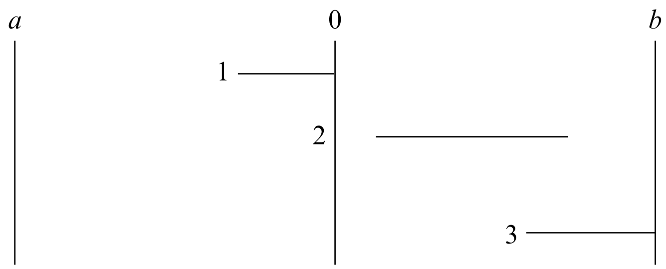
Type 2 left censor



Here the spell that was going on at time 0 is observed only after 0. The likelihood function differs from (3) only in its first part and is given by

$$L_2 = \prod_1 \int_a^0 h(x)f(t_i - x)dx \prod_2 h(x_i)f(t_i)\prod_3 h(x_i)[1 - F(b - x_i)]\prod_{all} P^{-1}. \tag{4}$$

Type 3 left censor



Here the spell that was going on at time 0 is observed only up to 0. Again, the likelihood function differs from (3) and (4) only in its first part.

$$L_3 = \prod_1 h(x_i)[1 - F(t_i)] \prod_2 h(x_i) f(t_i) \prod_3 h(x_i)[1 - F(b - x_i)] \prod_{all} P^{-1}. \quad (5)$$

So far we have assumed a homogeneous population. The necessary adjustment for a heterogeneous population is straightforward. Merely add subscript i to h, f, F, a, and b, and hence also to P. Otherwise, the likelihood function (3), (4), or (5) is unchanged.

3 Why divide by *P*

Now we answer the question posed after equation (3): Why is it less efficient to divide the first part by P_1 and the second and third part by P_2? We will consider Type 1 left censor; the other types can be similarly analyzed. For simplicity we will assume that there is no right censoring. Therefore, there are only two kinds of spells and the spell which reaches b is observed until its end. In this case the correct likelihood function is (3) except the third part. We will first give a heuristic and then a rigorous argument.

Rewrite (3) as

$$\begin{aligned} L_1 &= \prod_1 h(x_i) f(t_i) \prod_2 h(x_i) f(t_i) \prod_{all} P^{-1} \\ &= \prod_1 h(x_i)[f(t_i)] P_1^{-1} \prod_2 h(x_i) f(t_i) P_2^{-1} \prod_1 P_1/P \prod_2 P_2/P \equiv L_{11} L_{12}, \quad (6) \end{aligned}$$

where L_{11} consists of the first two products. From the above it is clear that dividing the two parts separately by P_1 and P_2 means ignoring L_{12}. It means ignoring information that a particular spell is either the first kind or the second kind. The estimator that maximizes L_{11} is a conditional maximum likelihood estimator; therefore, it is consistent but less efficient.

To advance a rigorous argument, we must introduce a parameter vector θ to estimate. Although we will treat θ as a scalar in the subsequent analysis, an extension to the vector case is obvious. Suppose f depends on θ but h does not. Taking the natural logarithm of the first line of (6) and ignoring h because it does not depend on θ, we have

$$\log L_1 = \sum_{i=1}^{n} \log f(t_i) - n \log P \quad (7)$$

where n is the number of observed spells. Differentiating (7) with respect to θ and noting P_2 does not depend on θ, we have

$$\frac{\partial \log L_1}{\partial \theta} = \sum_{i=1}^{n} \frac{1}{f} \frac{\partial f}{\partial \theta} - \frac{n}{P} \frac{\partial P_1}{\partial \theta}. \quad (8)$$