PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.
MAY BE RECALLED with earlier due date if requested.

DATE DUE	DATE DUE	DATE DUE
101708 MAR 2 2 2008	APR 2 2 2009	FEB 1 9 2010
JUL 2 8 2010	012810	

ESTIMATION AND TESTING IN DYNAMIC, NONLINEAR PANEL DATA MODELS

By

Margaret Susan Loudermilk

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Economics

2006

ABSTRACT

ESTIMATION AND TESTING IN DYNAMIC, NONLINEAR PANEL DATA MODELS

By

Margaret Susan Loudermilk

This dissertation consists of three chapters that address issues of estimation and testing in dynamic, nonlinear panel data models. Chapter 1 deals with an example of the peculiar difficulties that can arise in estimation of nonlinear models. Many economic variables occur as fractions and percentages. In these cases, the fractions instead of the level values are the variables of interest. Estimating models with fractional response variables can present challenges due to the presence of corner solution outcomes at 0 and 1 and continuous outcomes in the interval $(0, 1)$. Most standard estimation techniques are inappropriate in this setting because they are designed for variables that are either entirely continuous or take on only a discrete number of values. This chapter demonstrates an easily implemented method for estimating fractional response variables and presents an application of the technique to the determination of firm dividend policy.

Chapter 2 studies the sensitivity and relative performance of average partial effect estimates. Typically, partial effects are the quantities of interest for policy analysis. For linear models, these are often simply the parameter estimates. However, obtaining partial effects is more complicated for nonlinear models because these estimates will depend on all of the model's explanatory variables in a way that is not separable, except in special cases. Therefore, when some important individual specific explanatory variables are unobserved, consistent estimates of the partial effects may not be available. Instead, estimates of the partial effects averaged over the distribution of the unobservables, average partial effects, may be used as the variables of interest for policy analysis. Current estimation techniques

for dynamic, nonlinear panel data models require strong assumptions on economic models. Which assumptions are maintained affects generality, ease of computation, and even which quantities can be estimated, but little evidence exists on the relative performance of different estimation techniques for nonlinear panel data models. Since few economic models conform to such restrictive assumptions, it is important to know how sensitive estimates in these models are to econometric specifications. This chapter includes both simulations and empirical analysis.

Chapter 3 addresses a more general problem of testing the assumption of homoskedasticity in nonlinear models with unobserved effects. As a practical matter, heteroskedasticity is of little concern in linear models since it does not affect consistency or unbiasedness of estimators, and standard errors can easily be corrected to perform inference. However, in many nonlinear models the presence of heteroskedasticity is of greater consequence because it changes the functional form of the estimator. The class of tests known as score tests is ideal for cases in which the alternative hypothesis is complicated or computationally difficult because it only requires estimation under the null for implementation and is invariant to a many alternative hypotheses. Thus, such a test can be formulated for the null hypothesis of homoskedasticity against a general alternative that encompasses many prevalent specifications of the variance as special cases or locally invariant alternatives. In this chapter, a test for heteroskedasticity is proposed for two dynamic latent variable models, namely the panel probit and fractional response models, and applications of the test are presented for each.

DEDICATION

To my family

ACKNOWLEDGEMENTS

The journey to a PhD is a long and often difficult one. Many people have helped me on mine, and I want to take this opportunity to thank them. First, I would like to thank the members of my guidance committee, Leslie Papke, Ana Herrera, and Valentina Bali, for their contributions and support. I especially want to thank my advisor Jeff Wooldridge for educating me not only about econometrics but also about being a researcher, a teacher, and a scientist. Jeff's insight, willingness to talk about his own experiences, and his understanding of the need for balance between career and personal life has made my experience here much more rewarding.

I also want to thank Thomas Jeitschko, Deborah Foster, Mary Hoffman, Nicole Furnari, and Patrice Whitely for their efforts, particularly during my time on the job market. There are too many other people at Michigan State that have provided support and guidance during my years here to name them all. However, they all have my gratitude.

I could never have made it through graduate school without the love and support of my family. My parents were the first scientists in my life and much of the inspiration and motivation for my achievements has come from them. My two brothers have also contributed to my doctorate, each in their own way, and they have my thanks as well.

Without a doubt the person who has contributed the most to my doctorate is my fiance, Alan Bester. Alan has been there for me through all the ups and downs of graduate school and has been unwaivering in his love, support, and belief in me. Following each other through the dissertation process hasn't always been easy, but I have gained trememdously from his insight and advice, whether he knows it or not.

Finally, I would like to acknowledge financial support for this research from the American Association of Univeristy Women Educational Foundation's American Fellowship.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

Estimation of Fractional Dependent Variables in Dynamic Panel Data Models with an Application to Firm Dividend Policy

Many economic variables occur as fractions and percentages. Examples include firm market share, employer 401(k) contribution match rates, and TV Nielsen ratings. In these and similar cases, the fractions, instead of the level values, are the variables of interest. For example, market size varies across industries, but market share remains a meaningful measure of concentration and market power regardless of the absolute market size. What differentiates these variables from an econometric standpoint is that they are not probabilistic outcomes, yet they possess both two-corner solution outcomes and continuous outcomes in the interval $(0, 1)$. Consequently, most standard models are inappropriate for estimation. The approach presented in this paper provides a consistent estimator for fractional dependent variables with panel data in the presence of both lagged dependent variables and unobserved heterogeneity. The technique is then applied to firm dividend policy, which demonstrates the potential effects of ignoring the unique nature of fractional response variables.

In cross-sectional settings, fractional response models have used the standard logit framework under quasi-maximum likelihood estimation (QMLE) to address misspecification issues as proposed by Papke and Wooldridge (1996) or a linear model using the log-odds transformation. However, the log-odds framework is not appropriate for estimating models that have a fractional dependent variable with a substantial number of observations at either zero or one, as is the case in the application undertaken in this paper. In panel data settings with unobserved effects, the argument that provides consistency of the "fixed effects" logit model estimated by conditional maximum likelihood does not carry through when the dependent variable is fractional, evidently making the logit QMLE approach inconsistent for estimation with fixed effects. The consistency of this estimator appears to be entirely a consequence of the logistic distribution's functional form and depends on the binary nature of the dependent variable. (See Wooldridge (2002), p.491 for an illustration of the logit case.)

Estimation of dynamic, nonlinear panel data models is complex even without the addition of a fractional dependent variable. Using logit or probit models in a panel data setting introduces an "incidental parameters" problem since in nonlinear models with unobserved effects it is not possible to separate the unobserved effects from the maximum likelihood estimates (MLE) of the explanatory variables' parameters, and, except in special cases, no known transformations will eliminate the unobserved effects. Chamberlain (1992) and Wooldridge (1997), for example, present transformations for some multiplicative models. For fixed effect models, the number of parameters for the unobserved effects will increase with the number of cross-sectional observations, N, making consistent estimation impossible with a fixed number of time periods, T. MLE is still consistent as the number of time periods approaches infinity, but typically in panel settings the number of cross-sectional observations tends to infinity while the number of time periods is small. In random effects settings, the joint likelihood of the dependent variable $(y_{1t}, ..., y_{Nt})$ can no longer be written solely as the product of the marginal likelihoods of the dependent variables, y_{it}, and will require bivariate integration (Baltagi 2001, Hsiao 1986). Thus, to deal with unobserved heterogeneity in nonlinear models, typically either a distribution for the unobserved effect must be specified or a semiparametric approach will be necessary.

Semiparametric approaches allow consistent estimation of model parameters without assumptions on the distribution of the unobserved effects, but a major limitation of these approaches is their inability to estimate partial effects or average partial effects (APEs).

Even if one is only interested in parameters, currently, semiparametric approaches have some limitations. Honore and Kyriazidou (2000a) consider semiparametric estimation of different types of tobit models with individual specific effects, but the approach requires all of the regressors to be strictly exogenous. This prevents inclusion of lags of the dependent variable. Honore and Kyriazidou (2000b) extend the estimation to a logit framework that allows for lagged dependent variables but with assumptions on strictly exogenous covariates that eliminate the use of time dummy variables. Honore and Lewbel (2002) provide an alternative semiparametric estimator that allows for general predetermined regressors instead of only lagged dependent variables that achieves $\sqrt{N}$ consistency by assuming that at least one of the regressors is independent of both the errors and the unobserved effect. In addition, semiparametric methods are subject to the usual bias-variance trade-off present in nonparametric methods.

If a lagged dependent variable is included as a regressor when unobserved heterogeneity is present, ordinary least squares estimation (OLS) will be inconsistent since the lagged dependent variable will be correlated with the time invariant unobserved effects. Fixed effects estimation will also be inconsistent with fixed T. Monte Carlo evidence from Heckman (1981) shows that the bias created by incidental parameters is quite significant in dynamic models. In these cases, consistent estimation will depend on the treatment of the initial condition, y_{i0}. There are three prevalent parametric methods for dealing with initial conditions in nonlinear models. One approach treats the initial conditions as nonrandom. Another method is to specify a distribution for the initial condition given the unobserved effect so that the joint density of the $\{y_{it}\}$ can be written as $f(y_0, ..., y_T|z, c) = f(y_1, ..., y_T|y_0, z, c)f(y_0|z, c)$. In addition, Heckman (1981) suggests approximating the conditional distribution of the initial condition to avoid having to find it.

This paper follows the methodology proposed by Wooldridge (2005a) to deal with the initial conditions problem in dynamic, nonlinear panel data models and will use a tobit specification with corner solutions at both zero and one to estimate fractional dependent variables. While the approach addresses some problems present in other models, it retains many of the usual drawbacks of parametric models including the need to specify a conditional distribution for the unobserved heterogeneity, referred to herein as the auxiliary distribution. In addition, the model requires the explanatory variables other than the lagged dependent variable to be strictly exogenous; however, it does allow for the use of time dummies.

An application to firm dividend policy is presented to provide a concrete example where the fractional response panel data estimator is appropriate and careful treatment of fractional response variables has potentially important policy implications. The application considers the determination of the share of payouts to firms' shareholders made as share repurchases versus traditional cash dividends. This is an ideal example since 1) share repurchases as a fraction of total payouts is defined on the interval $[0, 1]$; 2) dividend policy theory suggests that both unobserved heterogeneity and state dependence are relevant in determining share repurchases; 3) the variables of interest are not considered to be endogenous in the current literature; and 4) a substantial fraction of observations are observed at both corner solution outcomes. The estimation demonstrates that conclusions drawn when neglecting the dual corner solutions can be misleading.

This chapter proceeds as follows: Section 2 presents the econometric model for dynamic panel data with fractional dependent variables in the presence of unobserved effects. Model specification, calculation of quantities of interest, estimation, and computational issues are all addressed. Section 3 presents an application of the technique to firm dividend policy theory, and section 4 contains conclusions.

1.1 Dynamic Fractional Response Model With Unobserved Effects

1.1.1 Model Specification

Specification of the model begins with a latent variable setup that allows for two corner solution outcomes, zero and one.

$$y_{it}^* = z_{it}\gamma + g(y_{i,t-1})\rho + c_i + u_{it} \tag{1}$$

$$u_{it}|(z_i, y_{i,t-1}, \ldots, y_{i0}, c_i) \sim N(0, \sigma_u^2) \tag{2}$$

$$y_{it} = \begin{cases} 0 & \text{if} \quad y_{it}^* \leq 0 \\ y_{it}^* & \text{if} \quad 0 < y_{it}^* < 1 \\ 1 & \text{if} \quad y_{it}^* \geq 1 \end{cases}$$

This is sometimes referred to as the "Two-Limit" or "Doubly-Censored" Tobit Model. z_{it} is comprised of strictly exogenous regressors, c_i represents the unobserved effect, and u_{it} is a normally distributed error term. For notational simplicity, let $g_{i,t-1} \equiv g(y_{i,t-1})$. The function $g(.)$ allows the effect of $y_{i,t-1}$ to differ depending on its realization in the previous period. For example, the lagged dependent variable might have a different effect if in the previous period there was a corner solution instead of an intermediate value. Such a case is easy to imagine for market share applications in which a firm engages in some degree of competition when it has a positive share but not all of the market, the firm acts as a monopolist at one corner, and has exited the market or is considering entry at the other.

Using the latent variable setup above, the density of y_{it} given z_{it}, $g_{i,t-1}$, and c_i can be derived as follows

$$P(y_{it} = 0|z_{it}, g_{i,t-1}, c_i) \quad = \quad \Phi\left(\frac{-z_{it}\gamma - g_{i,t-1}\rho - c_i}{\sigma_u}\right) \tag{3}$$

$$P(y_{it}=1|z_{it},g_{i,t-1},c_i) = \Phi\left(\frac{z_{it}\gamma + g_{i,t-1}\rho + c_i - 1}{\sigma_u}\right) \tag{4}$$

$$P(y_{it} \leq y|z_{it},g_{i,t-1},c_i) = \Phi\left(\frac{y_{it} - z_{it}\gamma - g_{i,t-1}\rho - c_i}{\sigma_u}\right) \tag{5}$$

$$\frac{\partial P(y_{it} \leq y|z_{it},g_{i,t-1},c_i)}{\partial y} = \frac{1}{\sigma_u}\phi\left(\frac{y_{it} - z_{it}\gamma - g_{i,t-1}\rho - c_i}{\sigma_u}\right) \tag{6}$$

To simplify notation further, define $w_{it} = (z_{it}, g_{i,t-1})$ and $\beta = (\gamma, \rho)$ which, together with (3) - (6), specify the density of y_{it} given $(z_{it}, g_{i,t-1}, c_i)$ as

$$f_t(y_{it}|w_{it}, c_i; \theta) = \Phi\left(\frac{-w_{it}\beta - c_i}{\sigma_u}\right)^{I[y_{it}=0]} \Phi\left(\frac{w_{it}\beta + c_i - 1}{\sigma_u}\right)^{I[y_{it}=1]} \cdots \times \left[\frac{1}{\sigma_u}\phi\left(\frac{y_{it} - w_{it}\beta - c_i}{\sigma_u}\right)\right]^{I[0<y_{it}<1]} \tag{7}$$

where θ represents the vector of parameters.

1.1.2 Estimation

Wooldridge (2005a) proposes estimation of dynamic, nonlinear panel data models with unobserved heterogeneity by modeling the distribution of the unobserved effect conditional on the initial value and any exogenous explanatory variables. There are several advantages to specifying the distribution of c_i conditional on y_{i0}. These include the ability to choose a flexible auxiliary distribution, the ability to specify the auxiliary distribution such that standard software packages can be used for estimation, and the fact that average partial effects are identified and can be estimated with little difficulty. These features are described in greater detail below.

Under the assumptions that the dynamics of the conditional distribution are correctly specified and $z_i = (z_{i1}, ..., z_{iT})$ is strictly exogenous conditional on c_i, the joint density of $(y_{i1}, ..., y_{iT})$ given (y_{i0}, z_i, c_i) is given by

$$f(y_{i1}, ..., y_{iT}|y_{i0}, z_i, c_i) = \prod_{t=1}^{T} f_t(y_{it}|w_{it}, c_i; \theta) \tag{8}$$

The idea behind the methodology is that since the density of $(y_{i1}, ..., y_{iT})$ given (y_{i0}, z_i, c_i) is already available under these assumptions, only the density of c_i given (y_{i0}, z_i) needs to be specified in order to proceed with estimation. In addition, this density is not restricted by the assumptions used above to derive the density of $(y_{i1}, ..., y_{iT})$ given (y_{i0}, z_i, c_i), so it can be chosen based on convenience, flexibility, or any other criteria.

In order to construct the likelihood function, it is necessary to integrate over the distribution of the unobserved effect c_i. This requires specification of the density of c given (y_{i0}, z_i) with parameter vector δ, denoted $h(c|y_{i0}, z_i; \delta)$. Given $h(c|y_{i0}, z_i; \delta)$ is a correctly specified model for the density of c_i given (y_{i0}, z_i), the log likelihood function is

$$l_i(\theta, \delta) = log\left\{\int \left[\prod_{t=1}^{T} f_t(y_{it}|w_{it}, c_i; \theta)\right] h(c|y_{i0}, z_i; \delta)dc\right\} \tag{9}$$

$$L = \sum_{i=1}^{N} log\left\{\int \left[\prod_{t=1}^{T} f_t(y_{it}|w_{it}, c_i; \theta)\right] h(c|y_{i0}, z_i; \delta)dc\right\} \tag{10}$$

In general, the quantities of interest in Tobit models will be $E(y|w)$ and $E(y|w, 0 < y < 1)$, as well as the partial effects of the explanatory variables. Using properties of the normal distribution, the conditional expectations are given by

$$E(y_{it}|w_{it}, c_i) = \Phi\left(\frac{w_{it}\beta + c_i - 1}{\sigma_u}\right) + \left[\Phi\left(\frac{1 - w_{it}\beta - c_i}{\sigma_u}\right) - \Phi\left(\frac{-w_{it}\beta - c_i}{\sigma_u}\right)\right] \cdots \times E(y_{it}|w_{it}, c_i, 0 < y_{it} < 1) \tag{11}$$

where

$$P(0 < y < 1|x) = \left[\Phi\left(\frac{1 - w_{it}\beta - c_i}{\sigma_u}\right) - \Phi\left(\frac{-w_{it}\beta - c_i}{\sigma_u}\right)\right] \tag{12}$$

and

$$E(y_{it}|w_{it}, c_i, 0 < y_{it} < 1) = w_{it}\beta + c_i + \sigma_u \left[\frac{\phi\left(\frac{-w_{it}\beta - c_i}{\sigma_u}\right) - \phi\left(\frac{1 - w_{it}\beta - c_i}{\sigma_u}\right)}{\Phi\left(\frac{1 - w_{it}\beta - c_i}{\sigma_u}\right) - \Phi\left(\frac{-w_{it}\beta - c_i}{\sigma_u}\right)}\right] \tag{13}$$

Combining these terms and defining $\Phi_1 = \Phi\left(\frac{-w_{it}\beta-c_i}{\sigma_u}\right)$, $\Phi_2 = \Phi\left(\frac{1-w_{it}\beta-c_i}{\sigma_u}\right)$, $\phi_1 = \phi\left(\frac{-w_{it}\beta-c_i}{\sigma_u}\right)$, and $\phi_2 = \phi\left(\frac{1-w_{it}\beta-c_i}{\sigma_u}\right)$ produces

$$E(y_{it}|w_{it},c_i) = \Phi\left(\frac{w_{it}\beta + c_i - 1}{\sigma_u}\right) + (w_{it}\beta + c_i)\left[\Phi_2 - \Phi_1\right] + \sigma_u\left[\phi_1 - \phi_2\right] \qquad (14)$$

Notice that the expression in (13) is clearly analogous to that of $E(y_{it}|w_{it}, c_i, y_{it} > 0)$ in the standard Tobit model where the typical Mills Ratio is replaced by a similar quantity for the interval between the two corners.

Then, by taking derivatives of the conditional mean equation with respect to the explanatory variables, the partial effects are given by

$$\frac{\partial E(y_{it}|w_{it},c_i)}{\partial z_{it}} = \frac{\gamma}{\sigma_u}\phi\left(\frac{w_{it}\beta + c_i - 1}{\sigma_u}\right) + \gamma\left[\Phi_2 - \Phi_1\right] + \frac{\gamma}{\sigma_u}(w_{it}\beta + c_i + z_{it})(\phi_1 - \phi_2) \qquad (15)$$

$$\frac{\partial E(y_{it}|w_{it},c_i,0 < y_{it} < 1)}{\partial z_{it}} = \gamma - \frac{\gamma z_{it}}{\sigma_u}\left(\frac{\phi_2 - \phi_1}{\Phi_2 - \Phi_1}\right) + \gamma\left(\frac{\phi_2 - \phi_1}{\Phi_2 - \Phi_1}\right)^2 \qquad (16)$$

$$\begin{aligned} \frac{\partial E(y_{it}|w_{it},c_i)}{\partial y_{i,t-1}} &= \frac{\rho g'_{i,t-1}}{\sigma_u}\phi\left(\frac{w_{it}\beta + c_i - 1}{\sigma_u}\right) + \rho g'_{i,t-1}\left[\Phi_2 - \Phi_1\right]\cdots \\ &+ \frac{\rho g'_{i,t-1}}{\sigma_u}(w_{it}\beta + c_i + g_{i,t-1})(\phi_1 - \phi_2) \end{aligned} \qquad (17)$$

$$\frac{\partial E(y_{it}|w_{it},c_i,0 < y_{it} < 1)}{\partial y_{i,t-1}} = \rho g'_{i,t-1} - \frac{\rho g_{it-1} g'_{i,t-1}}{\sigma_u}\left(\frac{\phi_2 - \phi_1}{\Phi_2 - \Phi_1}\right) + \rho g'_{i,t-1}\left(\frac{\phi_2 - \phi_1}{\Phi_2 - \Phi_1}\right)^2 \qquad (18)$$

However, the partial effects cannot be estimated due to the presence of the unobserved effect. Instead, the partial effect averaged across the distribution of the unobserved effect, the APE, can be estimated in the following way

$$m(z_t, y_{t-1}, c; \theta) \equiv E[y_{it}|z_{it}, y_{i,t-1}, c_i] \qquad (19)$$

$$\mu(z_t, y_{t-1}) = E_{c_i}[m(z_t, y_{t-1}, c_i; \theta)] = E\left\{E[m(z_t, y_{t-1}, c_i; \theta)|y_{i0}, z_i]\right\} \qquad (20)$$

$$= E\left[\int m(z_t, y_{t-1}, c_i; \theta) h(c|y_{i0}, z_i; \delta) dc\right] \qquad (21)$$

This provides a $\sqrt{N}$ consistent estimator

$$\hat{\mu}(z_t, y_{t-1}) = \frac{1}{N}\sum_{i=1}^{N} E[m(z_t, y_{t-1}, c_i; \hat{\theta})|y_{i0}, z_i] \tag{22}$$

from which the APEs can be obtained by taking derivatives with respect to z_t and y_{t-1}.

$$\frac{\partial\hat{\mu}(z_t, y_{t-1})}{\partial z_{it}} = \frac{1}{N}\sum_{i=1}^{N} \frac{\partial E[m(z_t, y_{t-1}, c_i; \hat{\theta})|y_{i0}, z_i]}{\partial z_{it}} \tag{23}$$

$$\frac{\partial\hat{\mu}(z_t, y_{t-1})}{\partial y_{i,t-1}} = \frac{1}{N}\sum_{i=1}^{N} \frac{\partial E[m(z_t, y_{t-1}, c_i; \hat{\theta})|y_{i0}, z_i]}{\partial y_{i,t-1}} \tag{24}$$

1.1.3 Computational Issues

Estimation of the model can be carried out using standard software that allows for two-limit random effects Tobit specifications if the density of c_i is specified in the following way

$$c_i|y_{i0}, z_i = \alpha_0 + \alpha_1 y_{i0} + \alpha_2' z_i + a_i \tag{25}$$

with $a_i \sim N(0, \sigma_a^2)$. Including the entire vector z_i along with y_{i0} allows the unobserved heterogeneity to be correlated with both the initial condition and the exogenous variables. This distribution for the unobserved effect is similar to that employed in the random effects probit model proposed by Chamberlain (1980) but uses the full vector z_i instead of the time averages of the exogenous variables to allow for more general correlation. Substitution for c_i produces

$$P(y_{it} = 0|w_{it}, z_i, g_{i0}, a_i) = \Phi\left(\frac{-z_{it}\gamma - g_{i,t-1}\rho - \alpha_0 - \alpha_1 y_{i0} - \alpha_2 z_i - a_i}{\sigma_u}\right) \tag{26}$$

$$P(y_{it} = 1|w_{it}, z_i, g_{i0}, a_i) = \Phi\left(\frac{z_{it}\gamma + g_{i,t-1}\rho + \alpha_0 + \alpha_1 y_{i0} + \alpha_2 z_i + a_i - 1}{\sigma_u}\right) \tag{27}$$

$$\frac{\partial P(y_{it} \leq y|w_{it}, z_i, g_{i0}, a_i)}{\partial y} = \frac{1}{\sigma_u}\phi\left(\frac{y_{it} - z_{it}\gamma - g_{i,t-1}\rho - \alpha_0 - \alpha_1 y_{i0} - \alpha_2 z_i - a_i}{\sigma_u}\right) \tag{28}$$

The log-likelihood function is then obtained by integrating the density of $(y_{i1}, ..., y_{iT})$ given $(w_{it}, z_i, g_{i0}, a_i)$ against the distribution of a_i

$$L = \sum_{i=1}^{N} log\left\{\int\left[\prod_{t=1}^{T} f_t(y_{it}|w_{it}, z_i, y_{i0}, a_i; \theta)\right]\frac{1}{\sigma_a}\phi\left(\frac{a}{\sigma_a}\right)da\right\} \tag{29}$$

This log-likelihood function has the same form as that for the random effects Tobit model with the explanatory variables w_{it}, z_i, y_{i0}. By iterated expectations and defining $\hat{\Phi}_1 = \Phi\left(\frac{-w_{it}\beta-\alpha_0-\alpha_1 y_{i0}-\alpha_2 z_i}{\sigma_u}\right)$, $\hat{\Phi}_2 = \Phi\left(\frac{1-w\beta-\alpha_0-\alpha_1 y_{i0}-\alpha_2 z_i}{\sigma_u}\right)$, $\hat{\phi}_1 = \phi\left(\frac{-w_{it}\beta-\alpha_0-\alpha_1 y_{i0}-\alpha_2 z_i}{\sigma_u}\right)$, $\hat{\phi}_2 = \phi\left(\frac{1-w_{it}\beta-\alpha_0-\alpha_1 y_{i0}-\alpha_2 z_i}{\sigma_u}\right)$, and $\sigma_v = \sigma_u + \sigma_a$, the conditional mean function, defined as $\hat{\mu}(z_t, y_{t-1})$ in (19), can be written as

$$r(w_{it}, z_i, y_{i0}; \theta) = \Phi\left(\frac{w_{it}\beta + \alpha_0 + \alpha_1 y_{i0} + \alpha_2 z_i - 1}{\sigma_v}\right) + (w_{it}\beta + \alpha_0 + \alpha_1 y_{i0} + \alpha_2 z_i)\cdots \times \left[\hat{\Phi}_2 - \hat{\Phi}_1\right] + \sigma_v\left[\hat{\phi}_1 - \hat{\phi}_2\right] \quad (30)$$

and the APEs are given by

$$\left.\frac{\partial r(w_{it}, z_i, y_{i0}; \theta)}{\partial w_j}\right|_{\theta=\hat{\theta}} = \frac{1}{N}\sum_{i=1}^{N}\beta_j \times \left\{1 + \left[\frac{1-\sigma_v}{\sigma_v}\right]\hat{\phi}_2 - \left[\frac{1-\sigma_v}{\sigma_v}\right](w_{it}\beta + \alpha_0 + \alpha_1 y_{i0} + \alpha_2 z_i)\left(\hat{\phi}_2 - \hat{\phi}_1\right) + \left[\hat{\Phi}_2 - \hat{\Phi}_1\right]\right\} \quad (31)$$

Since the distribution of c_i is now fully specified in terms of observables and a normally distributed error, partial effects on $P(y_{it} = 0|w_{it}, c_i)$ and $P(y_{it} = 1|w_{it}, c_i)$ can also be computed.

While standard software can be used to implement the estimator outlined above under these assumptions on the auxiliary distribution, special programming is required to use more general specifications of the unobserved effect. In addition, since estimation of the fractional response variable employs the truncated normal distribution, better convergence properties should be attainable by exploiting the trimming of the normal distribution's tails in the optimization routine.

1.2 Empirical Application

1.2.1 Dividend Policy Theory

There is a substantial literature, both theoretical and empirical, that investigates the motivations behind firm payouts to shareholders. A small but growing literature has also developed addressing the choice of paying dividends either through cash distributions or through share repurchases. The determination of what fraction of payouts will be made to shareholders in the form of share repurchases is an ideal application of the dynamic fractional response model for several reasons. First, a substantial fraction of dividend paying companies make either 0% or 100% of their payouts in the form of share repurchases in any given year. In the sample used in this application, over 20% of the observations occur at each corner. In addition, there are very few concerns raised in the existing empirical literature related to problems of feedback or endogeneity in the explanatory variables. Also, the concept of state dependence or persistence in dividend policy is central to the theoretical literature. The importance of dynamics is based on the work of Lintner (1956) who showed that firms are reluctant to reduce cash dividend payments since it may be viewed as a negative signal of future performance. Therefore, share repurchase programs may be used to distribute changes in earnings that are expected to be transitory while changes in cash dividends may reflect permanent changes in earnings. Finally, unobserved firm characteristics are recognized as potentially important in explaining firms' dividend policy decisions, but there are few empirical studies that have attempted to use panel data to correct for unobserved heterogeneity. The most common types of analyses performed in the past have been univariate comparisons and least squares estimation on averages of firm level annual data. However, recently some advances have been made in the application of panel data techniques and in addressing other econometric issues.

Jagannathan, Stephans, and Weisbach (2000) propose a number of testable implications of the hypothesis that firms use share repurchases to distribute temporary cash flows. First, they expect that firms with greater uncertainty about future cash flows, as measured by the volatility of operating income, will have a larger percentage of repurchases. Since operating cash flows tend to be more permanent than nonoperating cash flows, they also predict a negative relationship between the ratio of payouts from share repurchases and operating income, and a positive relationship with nonoperating income. In addition, Jagannathan, et al. argue that share repurchases may be used by management when they believe that the stock is undervalued, causing the proportion of repurchases to be negatively related to the market-to-book value of the stock. They provide an analysis of descriptive statistics of firm characteristics by payout method and perform a multinomial logit estimation to predict the choice of payout method. In the univariate comparisons, they find that repurchasing firms have lower operating incomes, higher nonoperating incomes, higher volatility, and poor stock performance as predicted by theory. The multinomial logit model produces the same conclusions with the exception that nonoperating income is not significant in the multivariate analysis. However, it is difficult to interpret how these results generalize to firms' choice between dividends and share repurchases unconditionally since a firm's choice of whether or not to increase payouts and by what method may be related to unobserved characteristics.

Fenn and Liang (2001) investigate the relationship between firms' payout policy and managers' stock incentives. In general, dividend policy theory suggests that insider ownership of stock aligns the incentives of management and shareholders, reducing agency problems and leading to higher payouts of firm cash flows. However, managerial stock incentives may also influence the composition of payouts to shareholders. Since the value of insiders' stock options is negatively related to future dividend payments, stock options create incentives for managers to make payments in the form of share repurchases instead of dividends. Fenn and Liang estimate the ratio of repurchases using a two-limit Tobit model on firm-level averages of annual data for 1993-1997. They use management shares and stock options, net operating cash flow, market-to-book ratio, log of assets, debt-to-assets (leverage), and volatility of operating income as explanatory variables. Of these, only management options, market-to-book ratio, and volatility are significant, and all three exhibit a positive relationship with the percentage of payments made through repurchases. Their finding of a positive and significant coefficient for volatility agrees with that of Jagannathan, et al. (2000), but they find the opposite relationship between market-to-book ratio and share repurchases. In addition, Fenn and Liang find a negative sign on operating income, which is also consistent with the results of Jagannathan, et al., but the coefficient estimate is not significant. Fenn and Liang also provide an alternative explanation for the role of options in share repurchase decisions. They suggest that insider stock options could act as a proxy for unobserved characteristics since firms with substantial growth opportunities may rely more heavily on stock options in providing executive compensation or may be more uncertain about the timing of investment opportunities. However, they believe that they have adequately controlled for growth opportunities through their selection of explanatory variables.

Moh'd, et al. (1995) also study the hypothesis that paying out cash dividends may reduce agency costs by providing outside monitoring of managers. They use an 18 year balanced panel of firms and include both industry effects and dynamics in their analysis. However, since they apply weighted least squares to the panel without instrumenting or using a transformation such as fixed effects or first differencing, their estimates will be inconsistent if unobserved effects are present.

Manos (2002) is perhaps the first paper in this literature to deal with sample selection directly. The study analyzes dividend payouts for a panel of firms from the Bombay Stock

Exchange using a Tobit model. Manos finds evidence of sample selection and applies a Heckman correction. It is uncertain to what degree Manos' results may generalize to firms on the US stock exchanges.

The application presented in this paper looks at some of the most commonly cited determinants of firm dividend policy to assess how estimates of these factors, as well as policy analysis, may be impacted by correcting for dual corner solution outcomes, dynamics, and unobserved heterogeneity. Since OLS estimation has been a prevalent method of analysis of firm dividend policy in the past, the estimation performed in this paper will begin with OLS as a basis of comparison for the results obtained after addressing these issues. Subsequently, the model will be augmented to take into account unobserved effects and dynamics within a linear model, as suggested by theory. Finally, the fractional response panel data model will be implemented to address the econometric problem arising from the presence of two corner solution outcomes.

1.2.2 Data

The data come from Compustat's Industrial Annual and Industrial Quarterly databases for 1992-2002 and includes all companies active during this period without missing data. This data set exhibits a large number of share repurchase outcomes at each corner with approximately 37% at zero and 25% at one. However, the number of observations per firm varies due to a combination of entry, exit, and missing data. Therefore, a sample was created by dropping all firms that were not present at the beginning of the sample and treating attrition as an absorbing state, meaning any observations following a period in which the firm was not observed were dropped. Here $t = 0$ refers to 1992 so firms that did not appear in both 1992 and 1993 were dropped. The underlying assumption when treating attrition in this manner is that at $t = 1$ the data represents a random sample and that any sample selection occurs through exit after this point. Implications of the unbalanced panel for estimation are addressed along with other estimation issues in section 1.2.3. Four observations were also dropped with share repurchase ratios that were either negative or greater than one. These appear to be data entry errors. This leads to a final sample with 11.628 observations on 1,800 firms.

The variables collected include: volatility of earnings, market-to-book ratio, operating income, non-operating income, dividends, and share repurchases. Volatility of earnings is defined as the standard deviation of the ratio of quarterly operating income to assets and is calculated from Compustat Industrial Quarterly File data items 21 and 44. The market-to-book ratio is defined as the ratio of market value of equity multiplied by shares outstanding to the book value of equity from Compustat Industrial Annual File data items 24, 25, and 60. Operating income is the ratio of operating income to assets, Compustat Industrial Annual File data items 13 and 6. Non-operating income is the ratio of non-operating income to assets, Compustat Industrial Annual File data items 61 and 6. The share repurchase and dividend data come from Compustat Industrial Annual data items 115 and 127, respectively. The share repurchase ratio is then computed as repurchases divided by the sum of repurchases and cash dividends. A detailed discussion of share repurchase measurement issues can be found in Jagannathan. et. al. (2000). Volatility, operating income, and nonoperating income are all expressed relative to assets to control for firm size. All Compustat data items are measured in millions of dollars. In addition, firm SIC codes were used to construct industry dummy variables. These definitions as well as summary statistics are provided in the appendix.

1.2.3 Estimation and Results

1.2.3.1 Ordinary Least Squares

Ordinary least squares estimation if a linear model is often a benchmark due to its theoretical and computational simplicity as well as the small number of assumptions required for consistency of the estimates. With a panel of data, OLS can be performed on the observations pooled across the cross-sectional units, i, and the time periods, t. Given the model,

$$y_{it} = y_{i,t-1}\rho + z_{it}\gamma + \epsilon_{it} \tag{32}$$

pooled ordinary least squares (POLS) coefficient estimates will be consistent under a zero conditional mean assumption, $E[\epsilon_{it}|y_{i,t-1}, z_{it}] = 0$, and a rank condition. No other restrictions on the distribution of the error term are required. Use of the OLS model will ignore functional form issues arising from the doubly-censored nature of the dependent variable. Since the share repurchase ratio is a fraction, predicted values, especially those for response probabilities, should always lie in the unit interval in order to be sensible. Predicted values from POLS can occur outside this interval though because a one unit increase in an explanatory variable will always have the same effect on the response probability, regardless of the starting value. In addition, this specification does not take into account potential firm level unobserved effects. Ignoring such unobserved heterogeneity, if present, would lead the POLS estimates to be inconsistent due to omitted variable bias. It is also important to remember that the zero conditional mean assumption will be violated when a lagged dependent variable is present if there are time invariant unobserved effects. The violation arises from the correlation between the unobserved effect, which is a component of ϵ_{it} along with the idiosyncratic error, and the lagged dependent variable.

In the case of this application, y_{it} is the share repurchase ratio and z_{it} is made up of the exogenous covariates (i.e, market-to-book ratio, operating income, non-operating income,

Table 1.1: Pooled Ordinary Least Squares Regression with Industry and Yearly Dummies

Share Repurchase Ratio	Coefficient Estimate
Lagged Repurchase Ratio	**0.675** (0.009)
Market-to-book Ratio	**-0.0001** (0.00003)
Operating Income Ratio	**0.063** (0.018)
Nonoperating Income Ratio	-0.093 (0.085)
Volatility of Earnings	0.011 (0.114)

Number of observations = 9828
Number of Firms = 1800

R-squared = 0.68

Note: Bold type indicates significance at the 1% level.
Quantity in parentheses is standard error.

and volatility). In order to control for industry level differences, the exogenous variables are augmented with 10 industry dummy variables: agriculture, mining & construction, manufacturing, transportation, retail, wholesale, communications, financial, services, and utility. In addition, a set of year dummies are included to control for changes occurring over time that are common to all firms. POLS estimates with heteroskedasticity robust standard errors are computed for the full sample. The estimates are provided in Table 1.1. The estimated coefficients for industry and year dummies are excluded here and in subsequent tables for brevity. The quantities given in parentheses are standard errors. Market-to-book ratio and non-operating income both have negative coefficient estimates, and operating income and volatility both have positive signs. However, only the lagged share repurchase ratio, market-to-book ratio, and operating income are statistically significant. The result for operating income is contrary to that obtained by both Jagannathan, et al. and Fenn and Liang, as well as that predicted by theory. Overall, the POLS estimation only shows empirical support for the ideas that low stock prices contribute to a policy of increased share repurchases and that there is persistence in share repurchase ratios.

1.2.3.2 Linear Dynamic Panel Data Model

Firm dividend theory suggests that both unobserved firm level effects and past dividend policy are determinants of share repurchase ratios. Therefore, the econometric model should be modified to include an unobserved effect such that

$$y_{it} = y_{i,t-1}\rho + z_{it}\gamma + c_i + u_{it} \tag{33}$$

where c_i is a firm specific unobserved effect and u_{it} is an idiosyncratic error term. Since c_i is not observed, estimatation of this model requires a transformation to remove its effect. The treatment of attrition as an absorbing state makes first-differencing (FD) a logical choice, and applying this transformation the model becomes

$$\Delta y_{it} = \Delta y_{i,t-1}\rho + \Delta z_{it}\gamma + \Delta u_{it} \tag{34}$$

where $\Delta x_{it} = x_{it} - x_{i,t-1}$. This removes the unobserved effect, c_i, as well as any time invariant regressors, such as industry dummies, from the estimation equation. Since the FD transformation removes the unobserved effect before estimation, no assumptions need to be made about the form of c_i. However, inclusion of $\Delta y_{i,t-1}$ violates the zero conditional mean assumption required for consistency of the FD estimator, $E[\Delta u_{it}|\Delta y_{i,t-1}, \Delta z_{it}] = 0$. Thus, estimation of this model will require implementation of an instrumental variables procedure. With first-differencing, $(z_i, y_{i,t-2}, y_{i,t-3}, ..., y_{i1})$ are all available as instruments at time t, but, in order to limit the number of overidentifying restrictions, only $(\Delta z_{it}, y_{i,t-2}, y_{i,t-3})$ will be used as instruments in the first stage regression. The first-differenced two-stage least squares (FD 2SLS) estimator corrects the problems inherent in the OLS estimation caused by ignoring the effects of unobserved heterogeneity. The estimator shares the desirable features of the OLS model that no assumptions about the form of the error distribution are needed for consistency beyond those on the conditional mean and that no assumptions about the form of the unobserved effects are required. However, it also shares the negative feature of ignoring functional form problems that can lead to predicted values outside of the unit interval.

In this application, there may also be concern over using a balanced subpanel for estimation since firm entry and exit are likely to be related to financial performance measures. In other words, firms that appear in every period may systematically differ from those which do not. Typically, in the existing literature on firm dividend policy, unbalanced panels are dealt with either by performing estimation on a balanced subpanel from the data or on firm level averages of annual data. Use of the latter precludes dynamics in the econometric model. Even when selection is random or ignorable, estimation on the balanced panel is inefficient since it is in effect throwing away data. Therefore, it is important to determine if sample selection is present. Testing for sample selection was not performed in the previous section because, in addition to the problems with OLS estimation noted above, if unobserved effects are omitted in the econometric model but are correlated with selection, inference on the significance of sample selection may be misleading. In order to produce consistent estimates on an unbalanced panel, selection may be related to z_i or c_i but may not be correlated with the error term. To test for selection in the panel with attrition, variable addition tests can be performed like those outlined by Wooldridge (1995, 2002). The most straightforward method is to include a lead of the selection indicator, $s_{i,t+1}$, as a regressor and test for significance using a t-test. (See, for example, Papke 1994.) The results of this estimation on the full sample are shown in Table 1.2. The estimates show that the lead variable is not significant in explaining share repurchases after conditioning on the other regressors and unobserved effects. It is now reasonable to conclude that sample selection is ignorable in the econometric model. The FD estimator is, therefore, consistent for the unbalanced panel or a balanced subpanel. After dropping the lead of the selection indicator from the regression, the FD 2SLS estimates are given in Table 1.3.

The lagged dependent variable is positive and highly significant with a magnitude similar to that obtained from POLS, indicating that there is substantial persistence in the percentage of payouts made through share repurchases even after accounting for unobserved heterogeneity. The contrast between the other estimates obtained here and under POLS is striking. The market-to-book ratio, which was significant at the 1% confidence level under POLS, is no longer significant after controlling for unobserved effects. The coefficient on non-operating income has changed in sign from negative to positive and is now highly significant. This is the relationship predicted by theory and obtained by Jagannathan, et al., but is opposite that obtained by POLS. Volatility has also become highly significant and has increased substantially in magnitude, making the results under FD 2SLS consistent with previous results, studies, and theory. Consequently, the results of the dynamic, linear model suggest that companies use share repurchases to pay out transitory changes in earnings but do not support a role for stock performance in dividend policy. In addition, the difference between these results and those of the OLS estimation suggests a potentially important role for unobserved firm heterogeneity in determining dividend policy.

1.2.3.3 Fractional Response Panel Data Model

Since the OLS estimates and the estimates from the dynamic, linear model support different but not mutually exclusive theories of the determination of firm payout methods, the natural question is whether one or the other, both, or neither of these theories will be supported once the dual corner solutions are accounted for in the estimation. However, in order to address

Table 1.2: FD 2SLS Estimates Including Lead Selection Indicator

Share Repurchase Ratio	Coefficient Estimate
Lagged Repurchase Ratio	**0.726**
	(0.201)
Market-to-book Value	-0.00002
	(0.00001)
Operating Income Ratio	**0.371**
	(0.094)
Non-operating Income Ratio	*0.385*
	(0.155)
Volatility of Earnings	*0.875*
	(0.397)
Lead Selection Indicator	-0.026
	(0.018)

Number of observations = 4736
Number of Firms = 1129

Note: Bold type indicates significance at the 1% level, italics 5%.
Quantity in parentheses is standard error.

Table 1.3: First-Differenced Two-Stage Least Squares Estimates

Share Repurchase Ratio	Coefficient Estimate
Lagged Repurchase Ratio	**0.606**
	(0.169)
Market-to-book Value	-0.00002
	(0.00001)
Operating Income Ratio	**0.352**
	(0.082)
Non-operating Income Ratio	*0.333*
	(0.138)
Volatility of Earnings	**1.020**
	(0.352)

Number of observations = 5189
Number of Firms = 1129

Note: Bold type indicates significance at the 1% level, italics 5%.
Quantity in parentheses is standard error.

this functional form problem in the context of the model from section 1.1, more restrictions on the error distribution will be required as well as assumptions about the distribution of the unobserved effect. For estimation of the dynamic, fractional response model with unobserved effects, equation (33) is now the latent variable equation in the two-limit Tobit model specified in (1). The auxiliary distribution is specified as described in section 1.1.3 such that

$$c_i = \alpha_0 + \alpha_1 y_{i0} + \alpha_2' z_i + a_i \tag{35}$$

where z_i includes the values of the time varying exogenous regressors (i.e., market-to-book ratio, operating income ratio, non-operating income ratio, and volatility of earnings) in every time period. The inclusion of the exogenous variables from all time periods will limit the number of observations to only those for firms with data available in all time periods, creating a balanced panel for estimation of the fractional response model. Consequently, testing the ignorability of selection will not be possible for the fractional response model under this specification of the unobserved effect. However, tests of selection measures such as the number of periods a firm appears in the sample with alternate specifications failed to show a significant selection effect.

The results of the estimation are displayed in Table 1.4. While the parameter estimates for non-operating income and volatility become insignificant once the nonlinearity is accounted for, a more significant difference between this model and the previous specification is in the size of the coefficient estimates. The parameter estimate for operating income is now almost twice as large, and the coefficient estimate for market-to-book ratio is 100 times larger. However, all of the coefficient estimates maintain the same signs. Looking at the larger implications of the model, it is clear that results that were highly significant under both of the previous specifications continue to be highly significant here (e.g., a positive

Table 1.4: Dynamic, Two-Limit, Random Effects Tobit Estimates

Share Repurchase Ratio	Coefficient Estimate
Lagged Repurchase Ratio	**0.484**
	(0.025)
Market-to-book Value	-0.002
	(0.001)
Operating Income Ratio	**0.687**
	(0.125)
Non-operating Income Ratio	0.192
	(0.245)
Volatility of Earnings	1.107
	(0.689)

Number of observations = 4530
Number of Firms = 453

Note: Bold type indicates significance at the 1% level.
Quantity in parentheses is standard error.

relationship with operating income) while those that were significant in only one now appear as marginally significant. This is the case for both market-to-book value and volatility of earnings with coefficient estimates just below the 10% confidence level. Thus, what appeared to be strong empirical support for different theories in the previous models is now far less conclusive.

Comparing these parameter estimates still provides an incomplete picture of the different implications of these estimators empirically and, more importantly, for policy analysis since the quantities of interest for determining the effects of the explanatory variables are the partial and average partial effects. Under POLS and FD 2SLS, the coefficient estimates are also the partial effects. However, in the fractional response model this is not the case. Instead, the APEs are computed using the parameter estimates from the fractional response model and equation (31). The APE estimates are shown in Table 5. The directions of the effects are the same as those for the parameter estimates, but the size of the effects has increased by 50-100%. Such a large difference in magnitude could be of great importance in policy analysis.

Standard errors for the average partial effects can be obtained by the delta method or bootstrapping. Due to the complicated form of the average partial effects in the case of the doubly-censored Tobit model, bootstrap standard errors are more practical. Bootstrap sampling can also provide asymptotic refinements when the sampling scheme appropriately recreates the dependence structure of the data. This is possible given a correctly specified parametric model under certain regularity conditions. (See Andrews 2001.) Bootstrap standard errors for the APEs, computed with 500 replications, are also included in Table 5. In addition to the difference in magnitudes, the standard errors show that, in contrast to the parameter estimates, all of the APE estimates are highly statistically significant.

1.2.4 Specification Testing

A major drawback of using the Tobit model for estimation of a fractional response variable is that it is only consistent under the assumption of normality of the error distribution. The empirical relevance of departures from normality, however, depends on the extent of their effects on the estimation results. This section attempts to address this concern as well as other issues of misspecification. Another concern raised by this estimation method is that misspecification of the error distribution may also cause inconsistency. The fractional response panel data estimator described above can be implemented under more general distributional assumptions for the unobserved effect, but this specification has two features to recommend it. First, as previously discussed, the adoption of this distribution of the unobserved effect allows estimation to be performed with standard software. In addition, this class of models is prevalent in current empirical work. To evaluate the impact of such

Table 1.5: Average Partial Effect Estimates

	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
Lagged Repurchase Ratio	0.696	0.713	0.728	0.776	0.806	0.836	0.844	0.853	0.789	0.758
	(0.055)	(0.044)	(0.047)	(0.036)	(0.029)	(0.030)	(0.036)	(0.036)	(0.034)	(0.042)
Market-to-book Value	-0.002	-0.002	-0.002	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003
	(0.0002)	(0.0001)	(0.0002)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Operating Income Ratio	0.989	1.012	1.033	1.102	1.145	1.187	1.199	1.211	1.120	1.076
	(0.078)	(0.062)	(0.067)	(0.051)	(0.040)	(0.042)	(0.051)	(0.052)	(0.048)	(0.060)
Non-operating Income Ratio	0.276	0.283	0.289	0.308	0.320	0.332	0.335	0.338	0.313	0.301
	(0.022)	(0.017)	(0.019)	(0.014)	(0.011)	(0.012)	(0.014)	(0.014)	(0.013)	(0.017)
Volatility of Earnings	1.594	1.631	1.666	1.777	1.846	1.914	1.933	1.952	1.805	1.735
	(0.126)	(0.100)	(0.108)	(0.082)	(0.065)	(0.068)	(0.082)	(0.083)	(0.077)	(0.096)

Note: All average partial effect estimates are significant at the 1% level.
Quantity in parentheses is standard error.

specification problems, a functional form test can be performed.

By extending Ramsey's Reset test to index models, Papke and Wooldridge (1996) derive a test that can be performed as a general functional form diagnostic for fractional response models. In this "Reset-type" test, the null model is given by

$$y_{it}^{*} = x_{it}\xi + a_i + u_{it} \tag{36}$$

with $x_{it} = (z_i, y_{i,t-1}, y_{i0})$. The alternative model is then

$$y_{it}^{*} = x_{it}\xi + \eta_1(x_{it}\xi)^2 + \eta_2(x_{it}\xi)^3 + a_i + u_{it} \tag{37}$$

and a Lagrange Multiplier (LM) statistic is computed for the null hypothesis $\eta_1 = 0, \eta_2 = 0$. The idea is that if the model has been correctly specified, nonlinear functions of the explanatory variables will have no additional explanatory power. Table 6 presents the results of the Reset-type specification test described above. Column 1 shows the estimation results and Ramsey's Reset test statistic for the POLS model estimated in section 1.2.3.1. Here the variables included in x are the lagged repurchase ratio, market-to-book value, operating income ratio, nonoperating income ratio, and volatility of earnings. The Reset test statistic clearly indicates rejection of this model specification. This is not surprising given that the model ignores firm specific unobserved effects and nonlinearities. Column 2 shows estimation results and the Reset test statistic for the fractional response model from the previous section using the same set of explanatory variables. The fractional response model takes into account the effects of the nonlinear response probability, lagged dependent variable, and unobserved heterogeneity, but it is subject to the normality assumption discussed above and specific assumptions about the distribution of c_i. The test statistic for this model is substantially smaller than that for the POLS model, but it still clearly indicates misspecification.

What is the cause of the misspecification problem detected though? Unfortunately, the Reset test is not helpful in answering this question. One possiblity is that instead of having

Table 1.6: Reset-type Specification Test

Share Repurchase Ratio	POLS	RE Tobit	RE Tobit
Lagged Repurchase Ratio	**0.675** (0.009)	**0.484** (0.025)	**0.571** (0.076)
Lagged Repurchase Ratio^2			-0.146 (0.081)
Market-to-book Value	**-0.0001** (0.00003)	-0.002 (0.001)	-0.003 (0.002)
Market-to-book Value^2			0.0000004 (0.0000002)
Lagged Repurchases*Market-to-book			0.004 (0.003)
Operating Income Ratio	**0.063** (0.018)	**0.687** (0.125)	**1.080** (0.231)
Operating Income Ratio^2			*-0.994* (0.437)
Lagged Repurchases*Operating Income			0.103 (0.221)
Non-operating Income Ratio	-0.093 (0.085)	0.192 (0.245)	0.547 (0.528)
Non-operating Income Ratio^2			-0.415 (0.555)
Lagged Repurchases*Non-operating Income			-0.535 (0.847)
Volatility of Earnings	0.011 (0.114)	1.107 (0.689)	1.288 (1.248)
Volatility of Earnings^2			-2.708 (9.135)
Lagged Repurchases*Volatility			1.105 (1.469)
Reset statistic	241.77	147.27	4.26
df	2	2	2
5% critical value	5.99	5.99	5.99

Note: Bold type indicates significance at the 1% level, italics 5%.
Quantity in parentheses is standard error.

misspecified the error distribution, there may be nonlinear functions of the explanatory variables which are significant and have been omitted from the model. To test this possibility, column 3 estimates the fractional response model again including in x quadratic functions of the explanatory variables and interactions of the lagged dependent variable with the other explanatory variables. The coefficient estimates for the explanatory variables that were included in both two-limit Tobit estimations are qualitatively the same in terms of sign and significance, and they are quantitatively similar in terms of the size of the coefficient estimates when the quadratic and interaction terms are added. However, now the model passes the Reset test. Due to the similarity of the estimation results across both of the linear models, which are unaffected by assumptions about the form of the errors or unobserved effects, and the fractional response model as well as the results of the Reset-type functional form tests, it seems reasonable to believe that inconsistency due to the lack of robustness of the Tobit model to non-normality of the error distribution is not of great concern in this application.

To test whether it is appropriate to include a firm specific effect, a pooled two-limit Tobit model was estimated and a likelihood ratio test of the significance of the within-panel variance was performed. Results of the pooled two-limit Tobit estimation and likelihood ratio test statistic are provided in Table 1.7. The test clearly rejects the hypothesis of no unobserved firm effects. This is further supported by the differences between the coefficient estimates for the pooled Tobit and those obtained from models that account for unobserved heterogeneity (i.e., RE Tobit and FD 2SLS), as well as by the similarity of the pooled tobit estimates to those of POLS, which ignores these effects.

Comparing results across all four models estimated provides additional support for the appropriateness of the fractional response panel data model in this application. Parameter estimates for the lagged share repurchases are positive and significant across all models and are similar in magnitude. The parameter estimates are somewhat higher for the POLS

Table 1.7: Dynamic, Two-Limit, Pooled Tobit Estimates

Share Repurchase Ratio	Coefficient Estimate
Lagged Repurchase Ratio	**0.962**
	(0.020)
Market-to-book Value	-0.002
	(0.001)
Operating Income Ratio	*0.138*
	(0.060)
Non-operating Income Ratio	-0.081
	(0.203)
Volatility of Earnings	-0.471
	(0.404)
Likelihood Ratio Statistic	279.49

Number of observations = 4530
Number of Firms = 453

Note: Bold type indicates significance at the 1% level, italics at 5%.
Quantity in parentheses is standard error.

and pooled Tobit models though, possibly indicating an upward bias created by neglecting unobserved heterogeneity. Estimates for market-to-book value are consistently negative and small in magnitude across models but are substantially smaller under POLS and FD 2SLS in which the nonlinearity in the dependent variable is not taken into account. The relationship between nonoperating income and share repurchases predicted by theory is supported by the results of the FD 2SLS and fractional response panel data model, but under POLS and pooled Tobit the opposite relationship is obtained. Again, the change in the sign of the parameter estimate is likely caused by neglecting unobserved heterogeneity. Volatility also appears to suffer from a substantial downward bias when unobserved effects are ignored. In summary, estimates are consistent across both the models that are and are not affected by the assumptions of normality of the error terms and of a specific form of the unobserved effect when all factors are taken into account.

1.3 Discussion

This chapter develops a method for estimating fractional dependent variables with panel data in the presence of unobserved effects and lagged dependent variables. The estimator allows for a wide variety of specifications for the density of unobserved heterogeneity. Average partial effects are identified and easy to computable. In addition, a special case of the estimator can be implemented by standard software that includes routines for random effects Tobit models. An application of the technique to firm dividend policy also demonstrates the potential effects of neglecting the doubly-censored nature of fractional responses. Primarily, this application shows that it is important to recognize that average partial effects are the relevant quantity for empirical analysis in dynamic, nonlinear panel data models with unobserved effects. In chapter 2, the robustness of APE estimates to specifications of unobserved effects and intial conditions is examined further using both simulations and an application to household brand choice.

Finally, there are several potential extensions to the results presented here. First, incorporation of heteroskedasticity into both the structural and auxilliary densities could be considered. Also, Wooldridge (2000) presents an approach that could extend the methodology employed in this paper to allow for feedback to future explanatory variables. It might also be interesting to investigate how the proportion of observations occurring at corner solution outcomes affects the differences between the ordinary least squares and dynamic linear and nonlinear model estimates. In addition, programming the general model to allow for any auxiliary distribution and to take advantage of potential computational gains would be desirable.

CHAPTER 2

An Examination of the Sensitivity of Average Partial Effects in Panel Probit Models

In nonlinear panel data models, the usual estimates of interest from regression analysis, partial effects, are not parameter estimates as they typically are in linear models. Instead, partial effects depend on all of the model explanatory variables, including unobserved effects, through a nonlinear function. Consequently, nonlinear panel data models are difficult to estimate because unobserved effects can not be separated from maximum likelihood estimates except in special cases. When lagged dependent variables are present, dynamics add another layer of complexity due to the introduction of initial conditions. Three methods are generally used for estimation of dynamic, nonlinear panel data models that produce consistent parameter estimates: random effects, bias corrected fixed effects, and semiparametrics.[1] However, not all of these methods are able to identify partial effects and average partial effects (APEs), partial effects averaged over the distribution of the unobservables, which are more relevant empirically and for policy analysis.

Wooldridge (2005b) argues that in the literature too much attention has been given to identification of parameters and not enough to partial effects. As discussed by Wooldridge, this criticism is particularly important in the context of latent variable models since only the sign and relative magnitude of the parameters from these models and the total effect of covariates on the response probabilities have quantitative meaning. However, there has been a recent focus in the econometrics literature on estimating limited dependent variable models

[1] For details on bias correction methods, see Hahn and Newey (2004).

using semiparametric methods such as those devised by Honore and Kyriazidou (2000), Honore and Lewbel (2002), and Arellano and Carrasco (2003). Semiparametric methods allow for agnosticism about the form of the unobserved effect and initial conditions in estimation, but current methods do not allow for identification of APEs. In comparison, random effects (RE) models can indentify APEs but require specification of a parametric model for the densities of both the unobserved effect and the initial condition. While it is known that under misspecification of these densities, RE parameter estimates are inconsistent, the behavior of the associated average partial effect estimates is generally unknown. Since partial and average partial effects are typically the true quantities of interest for answering empirical questions. if average partial effects are relatively insensitive to the specification of unobserved effects and initial conditions, the focus on more complicated semiparametric methods may be unnecessary.

This chapter presents an examination of the behavior of average partial effects through simulation analysis and an empirical application as a step in filling this gap in the literature. In addition to addressing an econometric question, understanding the behavior of APEs has important practical implications. This is illustrated in a statement made by Heckman in his 2001 Nobel Lecture: "Different assumptions about the sources of unobserved heterogeneity have a profound effect on the estimation and economic interpretation of empirical evidence, in evaluating programs in place, and in using the data to forecast new policies and assess the effect of transporting existing policies to new environments."[2] The quote addresses the need for care in the choice of assumptions employed in econometric modeling of unobserved heterogeneity in order to obtain valid empirical results. Since the true functional forms of the distributions of the initial condition and unobserved heterogeneity are generally unknown, if RE methods are to be used for policy analysis, it is important to know about the sensitivity

[2]Heckman, 2001, p.686.

of quantities of interest to such assumptions and econometric specifications.

Currently, little evidence on the sensitivity of dynamic, nonlinear panel data estimators to functional form assumptions exists. However, a few studies compare parameter estimates across methods and parametric specifications for the binary response model. Chintagunta, et al (2001) compare traditional logit and probit estimation methods for a discrete choice model of household brand choices with the semiparametric method developed by Honore and Kyriazidou (2000). They compare the parameter estimates and standard errors produced by the various models and also perform Monte Carlo experiments to examine the sensitivity of the estimators to the assumptions made concerning unobserved effects. They conclude that: 1) All methods lead to significant estimates but that the size of the estimates varies greatly across specifications; 2) Conditional logit models produce more robust estimates of the coefficients on the exogenous variables but produce poor estimates of the coefficient on the lagged dependent variable; and 3) Estimates of state dependence greatly depend on the specification of unobserved heterogeneity. However. Chintgunta, et al assume that y_{i0} is exogenous in all specifications where a lagged dependent variable is included, which limits the applicability of their results to a broader class of models. Chay and Hyslop (2000) also compare estimation methods in dynamic, binary response panel data models. They perform both Monte Carlo experiments and estimate a model of female labor force participation to assess the performance of random effects, fixed effects, and the Honore and Kyriazidou methods. In contrast with Chintagunta, et al, Chay and Hyslop include multiple specifications for the initial conditions in their study. They find that when initial conditions are misspecified the degree of state dependence tends to be substantially overstated and the effects of the exogenous covariates are underestimated.

Hyslop (1999) performs a similar comparsion of the linear probability model, a static random effects probit model, and a dynamic random effects probit specification in an empirical application of female labor force participation and by comparing their prediction capabilities. Hyslop finds that the differences between the parameter estimates depend upon the specification of the errors and unobserved heterogeneity but are often comparable. Ham and Lalonde (1996) also conduct a limited comparison of estimates from a duration model application to job training programs when different assumptions are made about the form of unobserved heterogeneity. They do not draw any substantive conclusions about the effect of varying these assumptions though since the focus of their paper is on sample selection issues.

Thus, all previous studies focus on parameter estimates even though the quantities of interest in binary response models are typically the conditional response probabilities and partial effects. Instead, this chapter proposes an examination of the ability of random effects models with flexible specifcations for the conditional distributions of unobserved heterogeneity and intial conditions to produce good estimates of quantities of interest even if the model is incorrectly specified. This will be accomplished through both Monte Carlo experiments and an empirical application. The results of the simulation study are abstract and difficult to put into the proper context alone. Thus, the empirical application helps to clarify the findings of the simulation. The outline of the chapter is as follows: Section 2 describes the model specifications to be considered. Section 3 presents the Monte Carlo study, and section 4 examines APE sensitivity in an empirical application. A discussion of the results and conclusions is presented in section 5.

2.1 Model Specifications

The paper restricts attention to the latent variable model since this is the leading case in which parameters have no meaningful interpretation. This is also the case most commonly considered in previous studies on parameter estimates. Among latent variable models the simplest case is that of the binary response model. The dynamic, binary response panel data model is specified as

$$
\begin{aligned}
y_{it} &= 1(x_{it}\beta + \rho y_{i,t-1} + c_i + u_{it} > 0) && (1)\\
P(y_{it} = 1|x_{it}, y_{i,t-1}, c_i) &= F(x_{it}\beta + \rho y_{i,t-1} + c_i) && (2)\\
P(y_{it} = 0|x_{it}, y_{i,t-1}, c_i) &= 1 - F(x_{it}\beta + \rho y_{i,t-1} + c_i) && (3)\\
P(y_{it} = 1|x_{it}, y_{i,t-1}, ..., y_{i0}) &= P(y_{it} = 1|x_{it}, y_{i,t-1}) && (4)
\end{aligned}
$$

where $F(.)$ is some well-behaved cumulative distribution function, $i = 1, ..., N$ indexes cross-sectional units, $t = 1, ..., T$ indexes time periods, and $1(x_{it}\beta + \rho y_{i,t-1} + c_i + u_{it} > 0)$ is an indicator function equal to 1 when the expression inside the parentheses is greater than zero and is equal to zero otherwise. y_{it} is the dependent variable, x_{it} represents the exogenous variables, and c_i represents unobserved time invariant individual characteristics. u_{it} is an error term that has zero mean, variance σ_u^2, and is normally distributed in the probit case, $F(.) = \Phi(.)$, or has a logistic distribution, $F(.) = \Lambda(.)$, in the logit case. In empirical applications, the choice between the logit and probit models is largely based on convenience since typically there is nothing from economic theory to recommend one form over the other. However, parameter estimates generally differ by a substantial factor between these models. The conditional density of y_{it} is then given by

$$f_t(y_{it}|x_{it}, y_{i,t-1}, c_i) = F\left(x_{it}\beta + \rho y_{i,t-1} + c_i\right)^{y_{it}} \left[1 - F\left(x_{it}\beta + \rho y_{i,t-1} + c_i\right)\right]^{1-y_{it}} \quad (5)$$

and the log-likelihood function for the traditional dynamic random effects model is given by

$$L = \sum_{i=1}^{N} log \int [f(y_{i0}|x_i, c_i) \prod_{t=1}^{T} f(y_{it}|y_{i,t-1}, x_{it}, c_i)] h(c_i|x_i, \delta) dF(c_i) \quad (6)$$

with $h(c_i|x_i, \delta)$ representing the density of c_i given x_i with parameters δ. Average partial effects are then obtained as the partial effect after "integrating out" the unobserved effect. Standard errors for the average partial effects can be obtained by the delta method or bootstrapping.

From the equation for the log-likelihood function, it is clear that specification of $f(y_{i0}|x_i, c_i)$ and $h(c_i|x_i, \delta)$ is required in order to proceed with estimation. Since the data generating processes are unknown and consistency of maximum likelihood depends on correct specification, it is difficult to determine the best way to proceed. Ideally, specifications for unobserved effects and initial conditions that are very simple to estimate would lead to good estimates of quantities of interest even if the true underlying model was quite complicated. To investigate the feasiblity of such an approach for obtaining estimates of APE's, several fairly simple but flexible models for the condtional distributions of c_i and y_{i0} have been chosen. Three different specifications of the intial condition are studied. First, the initial condition is treated as exogenous in the sense that

$$P(y_{i0}|x_i, c_i) = P(y_{i0}). \tag{7}$$

The degree to which this assumption is a concern in practice depends on the application. For example, this assumption may not be distasteful in the context of brand choice, as in Chintagunta, et al, since the initial brand choice may, in fact, be unrelated to most individual specific characteristics. This is in contrast to settings such as estimating a wage equation or labor force participation where this assumption would be difficult to justify at best. Two specifications also approximate the distribution of the initial condition. Approximation of the distribution of the initial condition was proposed by Heckman (1981). The approximations used here are

$$P(y_{i0} = 1|x_i, c_i) \quad = \quad \Phi(x_{i0}\beta) \tag{8}$$

$$P(y_{i0} = 1|x_i, c_i) = \Phi(\eta_0 + \eta_1' \bar{x}_i + \eta_2 c_i) \quad (9)$$

which allow for varying degrees of correlation between the initial condition and individual characteristics. The second specification is more flexible because it allows for a separate intercept and does not impose the same parameter values on the initial period that are estimated for $t \geq 1$. Also, four specifications for the distribution of unobserved heterogeneity are considered. These are given by [3]

$$c_i|x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2) \quad (10)$$

$$c_i|x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2) \quad (11)$$

$$c_i|x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i)) \quad (12)$$

$$c_i|x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i)) \quad (13)$$

The distribution of the unobserved effect in (10) gives rise to what is sometimes called Chamberlain's Random Effects Probit Model or the Mundlak-Chamberlain Device and creates correlation between c and x through the mean[4]. The second specification allows for additional flexibility without attempting to incorporate heteroskedasticity, which complicates estimation, by using a polynomial in the exogenous regressors for the mean function. The last two specifications contain the same mean functions as previosly described but allow for exponential heteroskedasticity as well.

In addition to the traditional RE method, models specifying the distribution of the unobserved effect conditional on the initial condition will be estimated. This approach is proposed by Wooldridge (2005a). Wooldridge points out that, since the density of $(y_{i1}, ..., y_{iT})$ given (y_{i0}, x_i, c_i) is already available without assumptions on the initial condition, only the den-

[3]For notational convenience, when x_i is a vector x_i^2 or x_i^3 indicates that each element of the vector is squared or cubed.

[4]Mundlak (1978), Chamberlain (1980)

sity of c_i given (y_{i0}, x_i), denoted $h(c|y_{i0}, x, \delta)$, needs to be specified in order proceed with estimation. Therefore, given $h(c|y_{i0}, x_i; \delta)$ the log likelihood function for the random effects model is

$$L = \sum_{i=1}^{N} log \left\{ \int \left[\prod_{t=1}^{T} f(y_{it}|y_{i,t-1}, x_{it}, c_i) \right] h(c_i|y_{i0}, x_i; \delta) dF(c_i) \right\} \tag{14}$$

For this method, the distributions of the unoberserved effect are similar to those used in the traditional RE setting but include an additional term taking into account the initial condition

$$c_i|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2) \tag{15}$$

$$c_i|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i{}^2 + \alpha_4 \bar{x}_i{}^3, \sigma_a^2) \tag{16}$$

$$c_i|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i{}^2)) \tag{17}$$

$$c_i|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i{}^2 + \alpha_4 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i{}^2)) \tag{18}$$

By specifying the distribution of the unobserved effect in this way, correlation is allowed between the exogenous variables, unobserved effect, and the initial condition.

2.2 Simulation

Attention in the simulation study is restricted to the probit model with the variance of u_{it}, σ_u^2, normalized to 1 to simplify estimation and comparison of the estimates. A Monte Carlo study was performed for panels of multiple lengths in both the time and cross-section dimensions such that $T = 5, 25$ and $N = 100, 500$.[5] The true data generating process (DGP)

[5]In every case the magnitudes of N and T specified maintain the standard asymptotic properties for panels since T is small relative to N.

in the simulation is given by

$$y_{it} = x_{it}\beta + \rho y_{i,t-1} + c_i + u_{it}, \qquad t = 1, ..., T \tag{19}$$

$$y_{i0} = x_{i0}\beta + c_i + u_{i0} \tag{20}$$

$$u_{it} \sim N(0, 1) \tag{21}$$

$$c_i|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2(\lambda_0 + \lambda_1'\bar{x}_i{}^2)) \tag{22}$$

while estimation is performed using (7)-(9) and (10)-(13) for traditional RE or (15)-(18) for the method proposed by Wooldridge. Thus, the variance of $c_i|x_i$ in the DGP is heteroskedastic but of a different form than in the estimated specifications, and the form of the initial condition in the DGP is a combination of the estimated specifications. The simulations present a comparison of the parameter and average partial effect estimates produced when the correct distribution of the unobserved effect is specified, namely normally distributed unobserved effects, but the moments of the distribution may be incorrectly specified.[6] In addition, the degree of state dependence varies while the coefficients on the exogenous variables remain fixed in order to examine whether the relative size of the parameters influences the senstivity of the estimates. The parameter values used are $\beta_1 = .25$, $\beta_2 = 1.5$, and $\rho = 0, 0.25$, or 0.75. Note that for ρ=0 no lagged dependent variable is included in the data generating process (DGP). Estimating this static model, provides a basis of comparison for determining the effects of misspecification of the unobserved effect alone. All other model parameters (*i.e.*, $\sigma_a^2, \lambda, \psi, \alpha$ and η) will be fixed across designs. The "true" values of the APEs discussed below were obtained by simulating the partial effects for each case and evaluating the average at the mean value of the covariates.

[6] Research is ongoing to check the robustness of the results obtained here by assuming a non-normal distribution for the data generating process of the unobserved effects but performing the estimation using a normal distribution. Thus, the distribution itself, in addition to the mean and variance, is misspecified.

APE results for the static case (i.e. $\rho = 0$) are shown in Table 2.1, which includes 1,000 replications for each of the four panels (A, B, C, and D). Parameter estimates significant at the 10% confidence level or above are obtained in all cases. and the magnitudes of the biases are similar across models and panel sizes, as shown in Table B.2 in the appendix. On average, the parameter estimate for x_1 differs from its true value by .075 and that for x_2 differs by .449. This is approximately 30% of their true values. However, all tests of the hypothesis $\hat{\beta}_2 = 1.5$ are rejected at the 10% level. Similarly, 9 of the 16 cases reject the hypothesis that $\hat{\beta}_1 = 0.25$ at the same level. In Table 2.1, estimates are shown in bold print if they are both significant at greater than or equal to the 5% level and fail to reject the hypothesis that they are equal to their true value. Estimates shown in italics are similarly significant at 10%. In contrast, only 2 of 16 APE estimates for x_1 reject the hypothesis that they are equal to their true value and approximately 1/3 of the APE estimates for x_2 fail to reject this hypothesis. All of the estimated APEs are statistically significant at the 5% level with one exception that is significant at 10%. In addition, the bias of all these estimates is less than the average bias of the parameter estimates as a percentage of their true value. No particular specification of the unobserved effect appears to provide better estimates than any other overall. This could be concerning since statistically significant APE estimates are obtained that are statistically different from their true values in some cases. However, the differences in magnitude are sufficiently small in all cases that the deviation from the true value may not be important from an empirical standpoint, particularly since the size of the bias is not growing either with the sample size or the number of time periods. Also, it is interesting to note that the APE estimates are most robust to misspecification of c_i when the panel is small. As the panel size grows, correct specification of the conditional mean of c_i given x_i appears to become more important as shown in panel D. However, misspecification of the conditional variance does not seem to have a great effect on the estimates even as the panel size changes.

Tables 2.2 and 2.3 present simulation results for the dynamic case using the traditional random effects approach where $\rho = 0.25$ and 0.75, respectively. Again, 1,000 replications were performed in each case. The results for the dynamic models are more complicated than those for the static case and differ slightly according to both the degree of state dependence and the way in which the density of the initial condition is estimated. The parameter estimates associated with Tables 2.2 and 2.3 are given in Tables B.3 and B.4 in the appendix.

For $\rho = 0.25$, all parameter estimates for x_2 are statistically significant and nearly all of those for x_1 are as well. ρ is not well estimated by any model in panel A, which has the smallest sample size, but the estimates are statisitically significant in most other cases. It is interesting to note that the bias of the parameter estimates tends to be smallest when the unobserved effect is specified as in (13), especially for initial conditions specified in columns (i) and (ii). This corresponds to the most flexible specification for c_i and the more simple specifications for y_{i0}, but this effect is smaller in larger panels. Further, no combination of panel length, intial condition specification, and unobserved effect specification leads to a set of parameter estimates that are all significant and fail to reject the hypothesis that they are equal to their true values. The results of the APE estimation are more encouraging though. Insignificant APE estimates only occur when parameter estimates are statistically insignificant. Among the statistically significant APEs all fail to reject the hypothesis that they are equal to their simulated values with one exception, which is for x_2 in panel B under specification (13) for c_i and (i) or (iii) for y_{i0}. In general, the most simple specification for the initial condition, (i), appears to lead to the "worst" APE estimates since no combination of unobserved effect specification and panel size leads to all three APEs being simutaneously well estimated. In addition, no specification is able to accurately estimate all three APEs when $\rho = 0.25$ for the smallest sample size as the average parital effect of ρ is only consistenly well estimated when $T = 25$. However, when $T = 25$ using (10) and (ii) or (iii) all APEs are consistenly estimated for panel C and with (11) or (12) and (iii) all APEs are consistenly estimated in panel D.

For $\rho = 0.75$, which has a higher degree of state dependence, if the parameter estimates are statistically significant then the APE estimates are significant and not statistically different from their simulated values at the 10% confidence level or higher, except for xa_2 in panel B with (12) and (i). However, many fewer of the corresponding parameter estimates are statistically significant and not significantly different from their true values as indicated by the bold and italic print. Overall, for T=5 homoskedastic models seem to produce the best APE estimates. This is seen by observing that for (i)-(iii) and (10)-(12) all three APEs are consistenly estimated in panels A and B by using (10) or (11) to specify the distribution of the unobserved effect. Interestingly, for $\rho = 0.75$ and $T = 25$, all three APEs are only well estimated when (12) and (iii) are specified. This may indicate that the importance of allowing for heteroskedasticity in the variance of the unobserved effect increases as either the degree of state dependence, the number of time periods, or both increase. In contrast, when the number of time periods is small, all specifications for the initial condition and the conditional mean are able to produce a well estimated set of APEs.

Overall, Tables 2.2 and 2.3 share many similarities. Specifically, almost all APEs are statistically significant and close to their true values, except those with corresponding insignificant parameter estimates. In addition, the relative insensitivity of the APE estimates to misspecification is further emphasized by comparison with the findings of earlier studies on the sensitivity of parameter estimates. Chintagunta, et al found that parameter estimates for state dependence varied substantially depending on the specification of unobserved effects. The results in Tables 2.2, B.3,2.3, and B.4 show variation in the parameter estimates across the specifications of $c_i|x_i$, which seems to depend primarily on the specification of the initial condition. More importantly, even when parameter estimates vary the most across the distributions of unobserved heterogeneity, the changes in the magnitudes of the APE estimates are very small. Chay and Hyslop found that when initial conditions are misspecified the effect of state dependence is overstated and the effects of the exogenous regressors are understated. Instead, the simulations presented here show that in larger panels all parameter estimates were underestimated, regardless of the degree of state dependence. In smaller panels, however, when the initial condition was specified as $P(y_{i0} = 1|x_i, c_i) = \Phi(\eta_0 + \eta_1' \bar{x}_i + \eta_2 c_i)$ parameter estimates were all understated when the unobserved effects were estimated as homoskedastic and all overstated when they were heteroskedastic. However, the degree of variation in the APE estimates was much less, and APEs were more sensitive to the specification of the intial condition than that of the unobserved effect.

Table 2.4 shifts to consider the case when the distribution of c_i depends on the initial condition in addition to the exogenous covariates. This simplifies estimation and reduces the number of assumptions required by removing the need to model the distribution of y_{i0} separately. This is of even greater importance in light of the results above, which show that APEs are more sensitive to the specification of y_{i0} than c_i. The results of these simulations, using the specifications of unobserved heterogeneity described in (15)-(18), immediately show two other reasons to prefer this method to that of specifying the distributions of $c_i|x_i$ and $y_{i0}|x_i, c_i$ separately. First, there is never a case where $\hat{\beta}$ is statistically significant and equal to its true value and the associated APE estimate is not as well. Second, using this method, a larger number of the estimated APEs are consistently well estimated than under any single specification for $y_{i0}|x_i, c_i$. Further, when the degree of state dependence is high, for every panel size there is a specification for the unobserved effect that leads to all three APEs being simultaneously consistently estimated. As under the previous method, with $T = 5$ homoskedastic specifications for the unobserved effect seem to produce the best APE estimates and allowing for heteroskedasticity appears to become more important as the sample size increases.

The results obtained in this section are fairly abstract. Taken together, they suggest that if APE estimates are statistically significant, they tend to provide good estimates of their true expected values despite potential misspecification of the distributions of unobserved effects and initial conditions. In cases where dynamics are present and a distribution for the initial condition is specified, there is also a benefit to increased flexibility in the specification of the initial condition and to allowing for heteroskedasticity in the unobserved effect in larger panels. Whereas, in smaller panels, simplicity in estimation is more important. In addition, while the specification of the unobserved effect seems to have only a small effect on APE estimates, the specification of initial conditions has a large effect on APEs. This may be a good argument for using the approach in which the distribution of unobserved effects are specified conditional on the initial condition and exogenous regressors, as proposed by Wooldridge. In the following section, a similar approach is used in an empirical application. The results of the application support the findings described here and make the implications of the simulation study more concrete.

Table 2.1: Static Simulation, ρ=0

DGP $c_i|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2(\lambda_0 + \lambda_1'\bar{x}_i{}^2))$

Estimated Process	APE Estimates	
(A) T=5, N=100	x_1	x_2
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.041** **(0.018)**	**0.247** **(0.020)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2)$	**0.041** **(0.018)**	**0.245** **(0.020)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.041** **(0.018)**	**0.244** **(0.020)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*0.041* *(0.021)*	**0.246** **(0.022)**
Simulated APE	0.037	0.221
(B) T=5, N=500	x_1	x_2
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.041** **(0.008)**	0.247 (0.009)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2)$	**0.041** **(0.008)**	0.244 (0.008)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.041** **(0.011)**	0.244 (0.010)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.037** *(0.011)*	0.229 (0.033)
Simulated APE	0.033	0.199
(C) T=25, N=100	x_1	x_2
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.043** **(0.008)**	0.254 (0.015)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2)$	**0.042** **(0.008)**	0.255 (0.015)

continued on next page

Table 2.1 (continued)

Estimated Process	APE Estimates	
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.042** **(0.008)**	0.253 (0.015)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.037** **(0.011)**	0.229 (0.033)
Simulated APE	0.051	0.304
(D) T=25, N=500	x_1	x_2
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	*0.042* *(0.004)*	0.255 (0.007)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2)$	0.042 (0.003)	0.255 (0.007)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	*0.042* *(0.004)*	0.254 (0.007)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.041 (0.004)	0.249 (0.011)
Simulated APE	0.049	0.292

Quantity in parentheses is standard error.
Estimates in bold are significant at the 5% level & simulated equals estimated APE (italics 10%).

Table 2.2: Dynamic Simulation, $\rho = .25$

DGP $c_i|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2(\lambda_0 + \lambda_1'\bar{x}_i{}^2))$

Estimated Process	APE Estimates		
(i)	$P(y_{i0}\|x_i, c_i) = P(y_{i0})$		
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.041** **(0.019)**	**0.244** **(0.020)**	0.037 (0.037)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2)$	**0.039** **(0.018)**	**0.235** **(0.021)**	0.034 (0.036)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.042 (0.026)	**0.246** **(0.032)**	0.046 (0.070)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.048 (0.032)	**0.251** **(0.041)**	0.057 (0.088)
Simulated APE	0.037	0.219	0.037
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.041** **(0.009)**	0.246 (0.010)	**0.042** **(0.017)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2)$	**0.039** **(0.018)**	*0.233* *(0.020)*	0.035 (0.038)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.047** **(0.025)**	*0.249* *(0.027)*	0.050 (0.068)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*0.047* *(0.027)*	0.246 (0.023)	0.050 (0.095)
Simulated APE	0.033	0.198	0.033
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.045** **(0.008)**	0.273 (0.015)	**0.045** **(0.018)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2)$	**0.040** **(0.008)**	0.0241 (0.016)	**0.039** **(0.017)**

continued on next page

Table 2.2 (continued)

Estimated Process	APE Estimates		
$c_i \mid x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.042** **(0.008)**	0.254 (0.015)	**0.042** **(0.018)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.041** **(0.009)**	0.243 (0.016)	**0.039** **(0.019)**
Simulated APE	0.050	0.302	0.050
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \mid x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.045** **(0.004)**	0.272 (0.006)	**0.045** **(0.008)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2)$	**0.041** **(0.006)**	0.242 (0.009)	**0.040** **(0.010)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.042** **(0.004)**	0.254 (0.007)	**0.042** **(0.008)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.043** **(0.009)**	0.254 (0.010)	**0.043** **(0.017)**
Simulated APE	0.048	0.290	0.048

continued on next page

Table 2.2 (continued)

Estimated Process	APE Estimates		
(ii)	$P(y_{i0}=1\|x_i,c_i)=\Phi(x_{i0}\beta)$		
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2)$	**0.039** **(0.016)**	**0.239** **(0.018)**	0.038 (0.035)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2)$	**0.038** **(0.016)**	**0.228** **(0.019)**	0.033 (0.035)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*0.041* *(0.022)*	**0.239** **(0.032)**	0.049 (0.071)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.042 (0.028)	**0.238** **(0.037)**	0.051 (0.096)
Simulated APE	0.037	0.219	0.037
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2)$	**0.040** **(0.008)**	0.241 (0.009)	**0.040** **(0.017)**
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2)$	**0.038** **(0.016)**	**0.227** **(0.020)**	0.033 (0.036)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.039** **(0.018)**	**0.230** **(0.022)**	0.045 (0.075)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.043 (0.027)	**0.242** **(0.031)**	0.055 (0.088)
Simulated APE	0.037	0.219	0.037
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2)$	**0.045** **(0.008)**	0.272 (0.015)	**0.045** **(0.018)**
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2)$	**0.046** **(0.009)**	0.271 (0.015)	**0.044** **(0.018)**
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.041** **(0.007)**	0.241 (0.015)	**0.040** **(0.016)**

continued on next page

Table 2.2 (continued)

Estimated Process	APE Estimates		
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.043** **(0.011)**	0.252 (0.017)	**0.043** **(0.022)**
Simulated APE	0.050	0.302	0.050
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	**0.045** **(0.004)**	0.272 (0.006)	**0.045** **(0.008)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2)$	**0.046** **(0.007)**	*0.271* *(0.010)*	**0.045** **(0.011)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.042 (0.003)	0.252 (0.007)	**0.042** **(0.008)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.042** **(0.006)**	0.253 (0.008)	**0.044** **(0.015)**
Simulated APE	0.048	0.290	0.048

continued on next page

Table 2.2 (continued)

Estimated Process	APE Estimates		
(iii)	$P(y_{i0}=1\|x_i,c_i)=\Phi(\eta_0+\eta_1'\bar{x}_i+\eta_2 c_i)$		
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2)$	**0.041** **(0.018)**	*0.241* *(0.023)*	0.047 (0.044)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2)$	0.040 (0.025)	**0.236** **(0.026)**	0.046 (0.086)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2exp(\lambda_1'\bar{x}_i))$	**0.041** **(0.029)**	**0.202** **(0.056)**	0.082 (0.080)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2exp(\lambda_1'\bar{x}_i))$	0.055 (0.038)	**0.265** **(0.045)**	0.100 (0.110)
Simulated APE	0.037	0.219	0.037
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2)$	**0.042** **(0.011)**	0.242 (0.016)	**0.051** **(0.031)**
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2)$	**0.040** **(0.021)**	*0.235* *(0.019)*	0.048 (0.085)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2exp(\lambda_1'\bar{x}_i))$	0.043 (0.033)	**0.215** **(0.059)**	0.099 (0.085)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2exp(\lambda_1'\bar{x}_i))$	**0.054** **(0.025)**	0.245 (0.022)	0.079 (0.081)
Simulated APE	0.037	0.219	0.037
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2)$	**0.046** **(0.009)**	*0.272* *(0.016)*	**0.047** **(0.020)**
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2)$	**0.041** **(0.009)**	0.244 (0.016)	**0.045** **(0.022)**
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2exp(\lambda_1'\bar{x}_i))$	**0.038** **(0.013)**	0.220 (0.057)	**0.042** **(0.026)**

continued on next page

Table 2.2 (continued)

Estimated Process	APE Estimates		
$c_i \vert x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.044** **(0.013)**	0.255 (0.019)	0.049 (0.035)
Simulated APE	0.050	0.302	0.050
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \vert x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.046** **(0.005)**	0.273 (0.008)	**0.047** **(0.014)**
$c_i \vert x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2)$	**0.046** **(0.007)**	**0.272** **(0.011)**	**0.049** (0.017)
$c_i \vert x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.039** **(0.010)**	**0.230** **(0.053)**	**0.044** **(0.022)**
$c_i \vert x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.043** **(0.007)**	0.255 (0.012)	0.029 (0.029)
Simulated APE	0.048	0.290	0.048

Quantity in parentheses is standard error.

Estimates in bold are significant at the 5% level & simulated equals estimated APE (italics 10%).

Table 2.3: Dynamic Simulation, $\rho = .75$

DGP $c_i|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2(\lambda_0 + \lambda_1'\bar{x}_i^2))$

Estimated Process	APE Estimates		
(i)	$P(y_{i0}\|x_i, c_i) = P(y_{i0})$		
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.035** **(0.017)**	**0.211** **(0.026)**	**0.100** **(0.028)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2)$	**0.035** **(0.016)**	**0.208** **(0.021)**	**0.099** **(0.034)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*0.041* *(0.023)*	**0.240** **(0.025)**	**0.120** **(0.061)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.046 (0.031)	**0.247** **(0.040)**	0.123 (0.092)
Simulated APE	0.037	0.219	0.109
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.035** **(0.007)**	**0.212** **(0.012)**	**0.106** **(0.011)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2)$	**0.035** **(0.008)**	**0.208** **(0.010)**	**0.102** **(0.017)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*0.039* *(0.020)*	0.241 (0.019)	**0.127** **(0.063)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*0.042* *(0.022)*	**0.219** **(0.021)**	0.107 (0.068)
Simulated APE	0.033	0.196	0.098
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.039** **(0.008)**	0.228 (0.019)	0.113 (0.013)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2)$	0.037 (0.008)	0.227 (0.020)	0.111 (0.013)

continued on next page

Table 2.3 (continued)

Estimated Process	APE Estimates		
$c_i \vert x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.044** **(0.009)**	0.263 (0.015)	**0.131** **(0.018)**
$c_i \vert x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.043** **(0.009)**	0.253 (0.017)	**0.125** **(0.023)**
Simulated APE	0.050	0.299	0.149
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \vert x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	0.038 (0.003)	0.228 (0.009)	0.114 (0.006)
$c_i \vert x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2)$	**0.038** **(0.007)**	0.225 (0.011)	0.113 (0.006)
$c_i \vert x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.044** **(0.004)**	0.263 (0.007)	*0.131* *(0.008)*
$c_i \vert x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.043** **(0.008)**	0.253 (0.009)	**0.126** **(0.014)**
Simulated APE	0.048	0.288	0.144

continued on next page

Table 2.3 (continued)

Estimated Process	APE Estimates		
(ii)	$P(y_{i0} = 1 \| x_i, c_i) = \Phi(x_{i0}\beta)$		
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i \| x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.034**	**0.206**	**0.101**
	(0.015)	**(0.024)**	**(0.027)**
$c_i \| x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2)$	**0.033**	**0.202**	**0.098**
	(0.015)	**(0.020)**	**(0.034)**
$c_i \| x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*0.039*	**0.233**	*0.114*
	(0.020)	**(0.022)**	*(0.060)*
$c_i \| x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*(0.042)*	**(0.236)**	(0.128)
	(0.023)	**(0.037)**	(0.098)
Simulated APE	0.037	0.217	0.109
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i \| x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.034**	**0.207**	**0.105**
	(0.006)	**(0.011)**	**(0.012)**
$c_i \| x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2)$	**0.034**	**0.202**	**0.101**
	(0.007)	**(0.009)**	**(0.017)**
$c_i \| x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.041**	0.235	**0.125**
	(0.016)	(0.018)	**(0.058)**
$c_i \| x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.042**	**0.242**	0.143
	(0.018)	**(0.032)**	(0.102)
Simulated APE	0.033	0.196	0.098
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i \| x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.038**	0.227	0.113
	(0.008)	(0.020)	(0.013)
$c_i \| x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2)$	**0.038**	0.226	0.111
	(0.008)	(0.020)	(0.013)
$c_i \| x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.044**	0.262	**0.130**
	(0.008)	(0.015)	**(0.018)**

continued on next page

Table 2.3 (continued)

Estimated Process	APE Estimates		
$c_i \vert x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.043** **(0.009)**	0.252 (0.016)	**0.126** **(0.025)**
Simulated APE	0.050	0.299	0.149
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \vert x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	0.038 (0.003)	0.227 (0.009)	0.114 (0.006)
$c_i \vert x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2)$	0.037 (0.005)	0.225 (0.010)	0.113 (0.007)
$c_i \vert x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.044** **(0.004)**	0.262 (0.007)	**0.132** **(0.008)**
$c_i \vert x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.042** **(0.006)**	0.252 (0.008)	**0.126** **(0.015)**
Simulated APE	0.048	0.288	0.144

continued on next page

Table 2.3 (continued)

Estimated Process	APE Estimates		
(iii)	$P(y_{i0}=1\|x_i,c_i)=\Phi(\eta_0+\eta_1'\bar{x}_i+\eta_2 c_i)$		
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2)$	**0.035** **(0.017)**	**0.207** **(0.030)**	**0.102** **(0.029)**
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2)$	0.033 (0.022)	**0.209** **(0.027)**	0.094 (0.070)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.038 (0.028)	**0.192** **(0.053)**	*0.125* *(0.072)*
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.052 (0.036)	**0.257** **(0.043)**	0.152 (0.103)
Simulated APE	0.037	0.217	0.109
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2)$	**0.035** **(0.008)**	**0.208** **(0.021)**	**0.108** **(0.017)**
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2)$	*0.034* *(0.019)*	**0.206** **(0.020)**	*0.107* *(0.064)*
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.041 (0.029)	**0.196** **(0.055)**	*0.129* *(0.070)*
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.045 (0.026)	*0.226* *(0.017)*	0.109 (0.074)
Simulated APE	0.033	0.196	0.098
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2)$	**0.039** **(0.008)**	0.228 (0.021)	0.113 (0.014)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2)$	**0.037** **(0.008)**	0.225 (0.022)	0.113 (0.013)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.036** **(0.012)**	**0.210** **(0.055)**	**0.109** **(0.035)**

continued on next page

Table 2.3 (continued)

Estimated Process	APE Estimates		
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.043** **(0.010)**	0.254 (0.018)	**0.134** **(0.038)**
Simulated APE	0.050	0.299	0.1498
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	0.038 (0.005)	0.227 (0.012)	0.114 (0.009)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2)$	0.038 (0.005)	0.225 (0.011)	0.114 (0.006)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.034** **(0.011)**	**0.198** **(0.060)**	**0.104** **(0.035)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.044** **(0.009)**	0.255 (0.012)	**0.133** **(0.031)**
Simulated APE	0.048	0.288	0.144

Quantity in parentheses is standard error.
Estimates in bold are significant at the 5% level & simulated equals estimated APE (italics 10%).

Table 2.4: Dynamic Simulation, $c_i|x_i, y_{i0}$

DGP $c_i|x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2(\lambda_0 + \lambda_1' \bar{x}_i^2))$

Estimated Process	APE Estimates		
		$\rho = .25$	
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	**0.040** **(0.018)**	**0.233** **(0.021)**	0.036 (0.039)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$	*0.049* *(0.027)*	**0.259** **(0.032)**	0.048 (0.069)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	0.050 (0.038)	**0.253** **(0.052)**	0.064 (0.111)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	0.056 (0.041)	**0.262** **(0.042)**	0.065 (0.096)
Simulated APE	0.037	0.219	0.037
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	**0.041** **(0.007)**	0.235 (0.009)	**0.038** **(0.017)**
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$	**0.049** **(0.011)**	0.279 (0.017)	0.046 (0.030)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	**0.041** **(0.007)**	0.235 (0.009)	**0.038** **(0.017)**
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	**0.041** **(0.007)**	0.235 (0.009)	**0.038** **(0.017)**
Simulated APE	0.033	0.198	0.033

continued on next page

Table 2.4 (continued)

Estimated Process	APE Estimates		
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i \| x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	**0.041**	0.243	**0.040**
	(0.008)	(0.015)	**(0.019)**
$c_i \| x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$	**0.049**	0.279	0.046
	(0.011)	(0.017)	(0.030)
$c_i \| x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	**0.043**	**0.244**	0.042
	(0.017)	**(0.038)**	(0.048)
$c_i \| x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	**0.046**	**0.272**	0.047
	(0.014)	**(0.026)**	(0.049)
Simulated APE	0.050	0.302	0.050
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \| x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	**0.039**	**0.240**	**0.040**
	(0.006)	**(0.011)**	**(0.014)**
$c_i \| x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$	**0.049**	*0.276*	0.045
	(0.009)	*(0.015)*	(0.030)
$c_i \| x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	**0.047**	**0.248**	**0.040**
	(0.010)	**(0.033)**	**(0.019)**
$c_i \| x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	**0.043**	**0.280**	*0.047*
	(0.010)	**(0.019)**	*(0.025)*
Simulated APE	0.048	0.290	0.048

continued on next page

Table 2.4 (continued)

Estimated Process	APE Estimates		
		$\rho = .75$	
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	**0.036** **(0.017)**	**0.208** **(0.021)**	**0.100** **(0.035)**
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$	*0.046* *(0.025)*	**0.243** **(0.030)**	*0.117* *(0.061)*
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	**0.042** **(0.031)**	**0.211** **(0.036)**	0.103 (0.085)
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	**0.054** **(0.035)**	**0.248** **(0.041)**	0.115 (0.086)
Simulated APE	0.037	0.217	0.109
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	**0.035** **(0.007)**	**0.210** **(0.009)**	**0.106** **(0.017)**
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$	*0.048* **(0.009)**	0.249 (0.006)	**0.120** **(0.016)**
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	**0.035** **(0.008)**	**0.208** **(0.016)**	**0.105** **(0.019)**
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	**0.035** **(0.008)**	**0.208** **(0.016)**	**0.105** **(0.019)**
Simulated APE	0.033	0.196	0.098

continued on next page

Table 2.4 (continued)

Estimated Process	APE Estimates		
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	*0.036*	0.213	0.106
	(0.008)	(0.016)	(0.017)
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$	**0.046**	*0.265*	**0.129**
	(0.012)	*(0.018)*	**(0.026)**
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x_i}^2))$	**0.037**	0.210	**0.101**
	(0.013)	(0.025)	**(0.039)**
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x_i}^2))$	**0.044**	0.262	**0.126**
	(0.012)	(0.023)	**(0.041)**
Simulated APE	0.050	0.299	0.149
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	0.036	0.214	0.103
	(0.002)	(0.007)	(0.006)
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$	**0.044**	0.267	**0.131**
	(0.003)	(0.010)	**(0.020)**
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x_i}^2))$	**0.038**	0.225	0.109
	(0.009)	(0.036)	**(0.020)**
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x_i}^2))$	**0.045**	**0.273**	**0.124**
	(0.006)	**(0.015)**	**(0.022)**
Simulated APE	0.048	0.288	0.144

Quantity in parentheses is standard error.

Estimates in bold are significant at the 5% level & simulated equals estimated APE (italics 10%).

2.3 Application

The empirical application undertaken is a model of household brand choice. In the context of household brand choice $i = 1, ..., N$ represents the household observed, $t = 1, ..., T$ is the purchase occasion observed, c_i are unobserved time invariant household characteristics, z_{it} are observed explanatory variables, and y_{it} is an indicator variable for whether or not a specific brand was purchased. Only two brands can be considered with the binary model previously specified, but it can easily be extended to encompass all brands with a multinomial specification.

In the dynamic model of brand choice, households are assumed to maximize lifetime expected utility over two goods: the product under study and a composite good. The product studied is comprised of multiple brands. Under standard assumptions, microeconomic theory shows that households will choose the brand, j, that provides both the lowest quality adjusted price and the highest expected utility. Quality is assumed to depend on observed product and individual characteristics, whether the brand was consumed in the previous period (i.e., habit effects), unobserved time invariant characteristics, and random shocks, ϵ_{it}. With only two brands, if utility is quasilinear in the composite good, the utility of the good studied is logarithmic, and a household chooses to consume only brand j in period t, then utility maximization subject to a household budget constraint with prices, q_{jit}, leads to a brand choice decision rule given by

$$y_{1it} = 1\left\{z_{1it}\delta_1 - z_{2it}\delta_2 + \gamma_1 y_{1i,t-1} - \gamma_2 y_{2i,t-1} - ln(q_{1it}) + ln(q_{2it}) + c_i + \epsilon_{1it} - \epsilon_{2it} \geq 0\right\}$$

Redefining the variables such that $x_{it} = z_{1it} - z_{2it}$, $p_{it} = ln(q_{2it}) - ln(q_{1it})$, and $e_{it} = \epsilon_{1it} - \epsilon_{2it}$ the decision rule can be rewritten as

$$y_{it} = 1\left\{\beta x_{it} + \rho y_{i,t-1} + \xi p_{it} + c_i + e_{it} \geq 0\right\} \tag{23}$$

the decision rule can be rewritten as

$$y_{it} = 1\left\{\beta x_{it} + \rho y_{i,t-1} + \xi p_{it} + c_i + e_{it} \geq 0\right\} \tag{23}$$

This equation now mirrors that of (1). Thus, the underlying microeconomic theory leads to a simple econometric model of brand choice by assuming that the probability of purchasing a given brand follows some well-behaved cumulative distribution function as discussed in Section 2.[7]

2.3.1 Data

A.C. Nielsen data on yogurt purchases for a sample of households in Sioux Falls, South Dakota from 1986 to 1988 are used in this investigation.[8] A.C. Nielsen chose Sioux Falls as a representative market because its population demographics resemble those of the United States as a whole and because they were able to monitor purchasing at all major grocery stores in that area. Sample households were issued magnetized cards that were presented at the grocery store when checking out. All of the households' purchases were then scanned and the information provided to the data collection agency. In addition, the agency collected household demographic information and information on weekly marketing from grocery locations.

The full data set consists of 1318 households and 17,679 purchase occasions. Each observation represents a purchasing occasion and contains a household identifier, the date of purchase, the brand of yogurt purchased (Yoplait, Dannon, Nordica, or other), value per ounce of store or manufacturer coupons (if used), price per ounce before applying coupons, indicators for whether the brand had a special display or advertising feature at the time of

[7]See Chintagunta, et al (2001) for a detailed discussion of the theoretical model of dynamic, discrete brand choice.

[8]I would like to thank Ekaterini Kyriazidou, Department of Economics, UCLA for providing the data.

purchase, and data on household characteristics. Nordica has the largest market share for a single brand with 4,739 purchases followed by Yoplait with 3,813 and Dannon at 1,834. However, the majority of yogurt purchases are made of "other" brands, including store brands, which comprise 7,293 purchases together. In order to obtain the largest possible sample, Nordica and Yoplait will be the brands considered for the analysis. Using this data, variables for price per ounce in log differences and variables for special displays and features were created to correspond with the variables p_{it} and x_{it} described at the beginning of section 2.3. The variables for special displays and features are defined as the difference in indicator variables for whether Yoplait or Nordica were featured or on special at the time of the purchase occasion.

Since some of the specifications to be considered include a lagged dependent variable, households without at least 3 purchases were dropped from the data set. This creates an unbalanced panel in which the minimum number of household observations is 3 and the maximum is 305. The household with 305 observations appears to be an outlier since the number of purchase occasions is roughly continuous between 3 and 45 per household with no other observations between 65 and 305. Here $t = 0$ refers to a household's first purchase occasion, and the number of time periods in which a household is observed depends on the number of purchase occasions. In addition, if a household purchases a brand other than Yoplait or Nordica, all subsequent purchases are treated as a new household with the next purchase occasion in which Yoplait or Nordica is purchased defined as $t = 0$. These modifications lead to a final sample with 1,398 households and 5,618 observations for t $\geq$ 1.

2.3.2 Estimation

Standard software, such as Stata, can be used for estimation of some models described in section 2.1, but estimation of these models under most assumptions on the conditional density of the initial condition, $f(y_0|z,c)$, or unobserved heterogeneity, $h(c|\bullet)$, generally requires special programming to allow for specification of the densities, integration over the distribution of the unobserved effect, and calculation of the average partial effects with standard errors. Results for the static brand choice model are shown in Table 2.3.2. All of the estimates obtained are statistically significant at the 5% level. The magnitude of the differences between parameter estimates across unobserved effects specifications is 0.02 to 0.06 and 0.0001 to 0.006 for APE estimates. The APE estimates indicate that a 1% change in the relative price of Nordica or Yoplait leads to a 30% change in the probability of purchasing a particular brand on average. Also, being on a special display or having an advertising feature when the other brand does not increases the probability of purchasing that brand by 8 or 9%, respectively. The estimated effect of a price change on brand choice seems extremely high unless consumers view different brands of yogurt as very close substitutes.

As in the simulations, the results obtained in the dynamic case show more variance across different specifications. However, as with the static case, APE estimates vary by a smaller amount than parameter estimates. The results for $P(y_{i0} = 1|x_i, c_i) = \Phi(x_{i0}\beta)$ and $P(y_{i0} = 1|x_i, c_i) = \Phi(\eta_0 + \eta_1'\bar{x}_i + \eta_2 c_i)$ are shown in Table 2.6. Estimates from the dynamic model indicate a large variance in consumer price sensitivity. After controlling for past purchase behavior, a 1% change in the relative price between brands leads to a change of 8%-18% when $P(y_{i0} = 1|x_i, c_i) = \Phi(x_{i0}\beta)$ and of approximately 30% in the probability of purchase when $P(y_{i0} = 1|x_i, c_i) = \Phi(\eta_0 + \eta_1'\bar{x}_i + \eta_2 c_i)$. Habit effects are shown to be very strong in brand choice decisions. Unlike the effect of price, there is little variation across

Table 2.5: Static Brand Choice

Estimated Process	Coeff. Estimates	APE Estimates
(1) $c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$		
Price	-0.998	-0.305
	(0.138)	(0.009)
Display	0.255	0.078
	(0.098)	(0.032)
Feature	0.300	0.092
	(0.052)	(0.022)
(2) $c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2)$		
Price	-1.060	-0.310
	(0.151)	(0.011)
Display	0.285	0.084
	(0.112)	(0.035)
Feature	0.314	0.092
	(0.057)	(0.025)
(3) $c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$		
Price	-0.997	-0.304
	(0.138)	(0.009)
Display	0.256	0.078
	(0.098)	(0.032)
Feature	0.300	0.092
	(0.052)	(0.022)
(4) $c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$		
Price	-1.056	-0.309
	(0.151)	(0.010)
Display	0.280	0.082
	(0.112)	(0.035)
Feature	0.316	0.092
	(0.057)	(0.025)

Note: All estimates are significant at the 5% level.
Quantity in parentheses is standard error.

specifications in the effect of previous brand choice on current brand choice. Purchasing a particular brand in the previous period leads to an estimated increase in the probability of purchasing the same brand of 36-38%. The effect of a special display on brand choice is essentially unchanged from the static case, varying only between 8 and 10%, and the probability of purchasing a brand when it has an advertising feature is estimated to be 12-16% in the dynamic case. The results from Table 2.6 show that there is substantial variation in the APE estimates across specifications for y_{i0} for some variables. In addition, there is some variation across specifications for unobserved heterogeneity. The variation is small across different specifications of the mean and larger across specifications of the variance.

Table 2.7 shows the estimates obtained by specifing the distribution of the unobserved effects conditional on the initial condition and the exogenous regressors. The APE estimates for lagged brand choice are not statistically significant, and the other estimates are similar to those obtained by the other models. However, the variance of the parameter estimates is larger than in Tables 2.5 or 2.6. The APE estimates indicate that a 1% change in the relative price leads to a 29-37% change in the probability of purchasing a particular brand. Also, being on a special display or having an advertising feature when the other brand does not increases the probability of purchasing that brand by 6-9% or 9-11%, respectively.

Thus, the results of the empirical application mirror those obtained in simulations. Parameter estimates tend to vary more across specifications of the unobserved effects than APE's for both the static and dynamic models. APE's are relatively insensitive to the specification of unobserved effects but tend to vary by a larger amount with the specification of initial conditions.

2.4 Discussion

The results of the simulations and empirical analysis suggest that average partial effects are relatively insensitive to the specification of unobserved heterogeneity, and simple, flexible specifications do a good job of estimating quantities of interest even when the true distribution of the unobserved effects differs from that estimated. However, average partial effects appear to be considerably more sensitive to the specification of initial conditions, and this problem does not seem to be easily solved by choosing a more flexible specification for the distribution of the initial condition. This concern can be mitigated to an extent by using the approach of Wooldridge (2000a) though.

Ongoing research on this topic includes simulation evidence on the performance of APE estimates, specified as previously described, when the unobserved heterogeneity is actually non-normally distributed. For example, the data generating process for the conditional distribution of the unobserved effect could be specified as a mixture of normals where both the mean of the conditional distribution and the mixing probabilities depend on the exogenous regressors. In addition, efforts are being made to improve the computation time of the models described in the paper. Substantial reductions in the computation time are likely to require a change to estimation by simulation methods.

Table 2.6: Dynamic Brand Choice, $c_i|x_i$

	(i)		(iii)	
Estimated Process	Coeff. Est.	APE Est.	Coeff. Est.	APE Est.
(1) $c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$				
Lagged Brand Choice	1.792 (0.047)	0.378 (0.054)	1.830 (0.048)	0.377 (0.095)
Price	-0.842 (0.152)	-0.177 (0.022)	-1.539 (0.157)	-0.317 (0.056)
Display	0.392 (0.108)	0.082 (0.025)	0.447 (0.109)	0.092 (0.032)
Feature	0.695 (0.058)	0.146 (0.026)	0.588 (0.058)	0.121 (0.036)
(2) $c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2)$				
Lagged Brand Choice	1.750 (0.055)	0.362 (0.060)	1.783 (0.056)	0.363 (0.105)
Price	-0.831 (0.176)	-0.172 (0.025)	-1.520 (0.181)	-0.309 (0.064)
Display	0.406 (0.124)	0.084 (0.028)	0.493 (0.125)	0.100 (0.038)
Feature	0.706 (0.066)	0.146 (0.030)	0.563 (0.066)	0.115 (0.040)
(3) $c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$				
Lagged Brand Choice	1.772 (0.047)	0.375 (0.037)	1.791 (0.045)	0.372 (0.077)
Price	-0.392 (0.151)	-0.083 (0.030)	-1.243 (0.162)	-0.108 (0.054)
Display	0.466 (0.108)	0.099 (0.022)	0.476 (0.128)	0.111 (0.036)
Feature	0.771 (0.058)	0.163 (0.020)	0.579 (0.062)	0.149 (0.035)

continued on next page

Table 2.6 (continued)

	(i)		(iii)	
Estimated Process	Coeff. Est.	APE Est.	Coeff. Est.	APE Est.
(4) $c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$				
Lagged Brand Choice	1.731	0.360	1.748	0.369
	(0.055)	(0.041)	(0.051)	(0.081)
Price	-0.379	-0.079	-1.133	-0.102
	(0.173)	(0.034)	(0.144)	(0.057)
Display	0.482	0.100	0.451	0.103
	(0.124)	(0.025)	(0.116)	(0.040)
Feature	0.783	0.163	0.557	0.141
	(0.066)	(0.023)	(0.062)	(0.038)

Note: <u>All</u> estimates are significant at the 5% level.

Quantity in parentheses is standard error.

Column (i) $P(y_{i0} = 1 \mid x_i, c_i) = (\Phi(x_{i0}\beta)$

Column (iii) $P(y_{i0} = 1 \mid x_i, c_i) = \Phi(\eta_0 + \eta_1' \bar{x}_i + \eta_2 c_i)$

Table 2.7: Dynamic Brand Choice, $c_i|x_i, y_{i0}$

Estimated Process	Coeff. Estimates	APE Estimates
(1) $c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$		
Lagged Brand Choice	0.439 (0.760)	0.090 (0.126)
Price	**-1.801** **(0.158)**	**-0.371** **(0.073)**
Display	**0.411** **(0.109)**	**0.085** **(0.019)**
Feature	**0.532** **(0.058)**	**0.110** **(0.009)**
(2) $c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$		
Lagged Brand Choice	0.421 (0.858)	0.085 (0.143)
Price	**-1.775** **(0.183)**	**-0.360** **(0.080)**
Display	**0.442** **(0.126)**	**0.090** **(0.021)**
Feature	**0.548** **(0.066)**	**0.111** **(0.009)**
(3) $c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x_i}^2))$		
Lagged Brand Choice	**1.339** **(0.664)**	0.092 (0.196)
Price	**-4.215** **(0.128)**	**-0.289** **(0.055)**
Display	**0.910** **(0.086)**	**0.062** **(0.031)**
Feature	**1.271** **(0.048)**	**0.087** **(0.026)**

continued on next page

Table 2.7, continued

Estimated Process	Coeff. Estimates	APE Estimates
(4) $c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$		
Lagged Brand Choice	**1.390**	0.102
	(0.691)	(0.194)
Price	**-4.034**	**-0.297**
	(0.138)	**(0.067)**
Display	**0.989**	**0.073**
	(0.096)	**(0.035)**
Feature	**1.239**	**0.091**
	(0.051)	**(0.030)**

Note: Bold type indicates significance at the 5% level.
Quantity in parentheses is standard error.

CHAPTER 3

A Score Test for Heteroskedasticity in Dynamic Latent Variable Models

3.1 Introduction

In linear models heteroskedasticity does not affect consistency or unbiasedness of estimators and can easily be corrected for in standard errors in order to perform inference. However, in many nonlinear models the presence of heteroskedasticity is of greater consquence because it changes the functional form of the estimator. For this reason, the ability to test for heteroskedasticity in nonlinear models is important. The difference between the two cases can be seen by first observing that given a linear panel data model

$$E(y_{it}|x_i, c_i) = x_{it}\beta + c_i \tag{1}$$

with the partial effect of x on β given by

$$\frac{\partial E(y_{it}|x_i. c_i)}{\partial x_i} = \beta \tag{2}$$

does does not depend on any features of the distribution of the unobserved effect, c_i. In addition, it is simple to consistenly estimate β by fixed effects, as discussed in chapter 1. In some cases, such as a random-coefficient model in which there are individual specific slope coefficients

$$E(y_{it}|x_i, \beta_i. c_i) = x_{it}\beta_i + c_i \tag{3}$$

it is not possible to estimate the individual parameters, β_i, in the traditional panel setting with small T. However, under the assumption that $E(\beta_i|\ddot{x_{it}}) = E(\beta_i)$ for $\ddot{x_{it}} = x_{it} - \bar{x_i}$, it is still possible to consistently estimate average partial effects using FE. When a nonlinear model is considered though, such simple solutions are no longer available. The simplest case is that of a panel probit model, given by

$$P(y_{it} = 1|x_{it}, c_i) = \Phi(x_{it}\beta + c_i) \quad (4)$$

Again, the partial effects conditional on c_i will not depend on any features of the distribution of the unobserved effect, but, without knowing a value of the individual specific effect at which to evaluate the partial effect, these quantities can not be obtained. The APE, however, will depend on the entire distribution of c_i given x_i in general. The importance of the assumptions made about the form of this distribution for consistently estimating APEs was shown in chapter 2. For cases in which ignoring potential heteroskedasticity in the unobserved effect may affect estimation of the APEs having a test for heteroskedasticity which is easily implemented could be important.

General test procedures are difficult to construct for dynamic, nonlinear panel data models with unobserved heterogeneity though. As explained by Wooldridge (2005b), it is not possible to identify the effect of covariates on the variance, $\partial Var(y|x,c)/\partial x$, averaged over the distribution of the unobserved effect as it is the average partial effect. Consequently, to learn anything about the conditional variance it must be modeled directly.[1] Thus, tests for heteroskedasticity in limited dependent variable models are generally obtained on a case-by-case basis under specific assumptions about the structure of the unobserved effects. This paper will focus on testing the null hypothesis of homoskedasticity through exclusion restrictions using a score test. This simplifies estimation by avoiding the need to estimate the model as heteroskedastic, which would be necessary using Wald or likelihood ratio statistics. In or-

[1]See Wooldridge (2005b) for further explanation and examples using Poisson regression and duration models.

der to accomplish this, a parametric form for the unobserved heterogeneity will be specified. Throughout the paper, normality of the distribution of the unobserved effect and a given specification for its conditional mean are maintained. Clearly, one can question the validity of these assumptions and their effect on the proceedure. As many authors have noted, the validity of standard test statistics will depend on the auxiliary assumptions made.[2] However, the benefit of this approach is that, by assuming normality and formulating the test through exclusion restrictions, estimation of the restricted model is possible via traditional random effects methods for both the probit and tobit. In addition, the specification of the conditional mean can be made more flexible without any substantive changes to the test proceedure as long as it is a function of the exogenous covariates and the initial condition and is linear in parameters. In theory, the assumption of normality of the unobserved effect could also be tested either before or after testing for heteroskedasticity, but this is beyond the scope of this paper.

The remainder of the chapter proceeds as follows: Section 1 constructs the score test for the dynamic, panel probit model. Section 2 extends the test to the case of a fractional dependent variable, and section 3 presents applications of the test to both the dynamic probit and fractional response models. Section 4 concludes.

[2] See White (1981, 1989), Bera and Jarque (1982), Pagan and Hall (1983), Godfrey (1987), et al.

3.2 Score Test for Heteroskedasticity in the Dynamic Probit Model

The dynamic, panel probit model will be maintained as the model of interest throughout this section. Consider the dynamic, panel probit model of chapter 2 the following form

$$y_{it} = 1(x_{it}\beta + \rho y_{i,t-1} + c_i + u_{it} > 0) \tag{5}$$

$$u_{it}|(x_i, y_{i,t-1}, ..., y_{i0}, c_i) \sim N(0,1) \tag{6}$$

Using the latent variable setup above, the density of y_{it} given x_{it}, $y_{i,t-1}$, and c_i can be derived as follows

$$P(y_{it} = 1|x_{it}, , y_{i,t-1}, c_i) = \Phi(x_{it}\beta + \rho y_{i,t-1} + c_i) \tag{7}$$

$$P(y_{it} = 0|x_{it}, , y_{i,t-1}, c_i) = 1 - \Phi(x_{it}\beta + \rho y_{i,t-1} + c_i) \tag{8}$$

This leads to a conditional density for y_{it} given $(x_{it}, y_{i,t-1}, c_i)$ of

$$f_t(y_{it}|x_{it}, y_{i,t-1}, c_i; \theta) = \Phi\left(x_{it}\beta + \rho y_{i,t-1} + c_i\right)^{y_{it}} \left[1 - \Phi\left(x_{it}\beta + \rho y_{i,t-1} + c_i\right)\right]^{1-y_{it}} \tag{9}$$

with $\theta = (\beta, \rho)$.

Following the method proposed by Wooldridge (2005a) and given $h(c|y_{i0}, x_i; \delta)$ is a correctly specified model for the density of c_i given (y_{i0}, x_i), the log likelihood function is

$$l_i(\theta, \delta) = log\left\{\int \left[\prod_{t=1}^{T} f_t(y_{it}|x_{it}, y_{i,t-1}, c; \theta)\right] h(c|y_{i0}, x_i; \delta)dc\right\} \tag{10}$$

Defining $\Psi = \int \left[\prod_{t=1}^{T} f(y_{it}|x_{it}, y_{i,t-1}, c; \theta)\right] h(c|x_i, y_{i0}; \delta)dc$, the score is given by $s_i(\theta, \delta) = (\frac{\partial l_i(\theta,\delta)}{\partial \theta}, \frac{\partial l_i(\theta,\delta)}{\partial \delta})$ where

$$\frac{\partial l_i(\theta, \delta)}{\partial \theta} = \sum_{i=1}^{N} \frac{1}{\Psi}\left\{\int \left[\prod_{t=1}^{T} \frac{\partial f(y_{it}|x_{it}, y_{i,t-1}, c; \theta)}{\partial \theta}\right] h(c|x_i, y_{i0}; \delta)dc\right\}$$

$$\frac{\partial l_i(\theta, \delta)}{\partial \delta} = \sum_{i=1}^{N} \frac{1}{\Psi}\left\{\int \left[\prod_{t=1}^{T} f(y_{it}|x_{it}, y_{i,t-1}, c; \theta)\right] \frac{\partial h(c|x_i, y_{i0}; \delta)}{\partial \delta}dc\right\}$$

Then a test statistic can be constructed by evaluating the score at the resticted parameter values suggested by the null hypothesis and evaluating

$$S = \left(\sum_{i=1}^{N} \tilde{s}_i'\right)\left(\sum_{i=1}^{N} \tilde{s}_i\tilde{s}_i'\right)^{-1}\left(\sum_{i=1}^{N} \tilde{s}_i\right) \sim \chi_q^2 \tag{11}$$

where q is the number of constraints imposed and $\tilde{s}_i = \tilde{s}_i(\theta, \delta)$ denotes the score evaluated at the restricted estimates. Notice that when the statistic is computed, the $1/\Psi$ terms will cancel so that only the portion of the derivates appearing within the brackets needs to be evaluated in order to obtain it. The outer product form of the asymptotic variance shown in (11) has the advantage of not requiring additional calculations for estimation of the variance matrix and is invariant to reparameterizations. However, the nominal size of the test can be distorted in many cases such that it over rejects the null hypothesis. The expected Hessian form is often a better alternative as an estimate of the asymptotic variance when it is easy to obtain since it is always positive definite if it exists, is invariant to reparameterizations, and has smaller size distortions in many cases. Thus, the estimate of the variance matrix used may affect computational simpicity, finite sample properties, and the power of the test.

To make the test operational, more structure needs to be imposed on the unobserved effect. As explained above, a correctly specified model for the density of c_i given (y_{i0}, x_i) is needed. For this test proceedure, normality of the conditional distribution of the unobserved effect and a specification for the conditional mean are maintained. The conditional mean is only restricted to be a function of the exogenous covariates and the initial condition and to be linear in parameters. Assuming normality of the unobserved effect is neccessary in order to obtain estimates of the restricted model parameters through standard random effects maximum likelihood estimation. A test of the distributional assumption could also be implemented either before or after testing for heteroskedasticity but is not addressed in this paper.

Under the assumption that $c_i|y_{i0}, x_i = \alpha_0 + \alpha_1'\bar{x}_i + \alpha_2 y_{i0} + a_i$, where $\bar{x}_i$ is the time average of the exogenous variables and a_i is an individual specific, zero mean, normally distributed error term,

$$P(y_{it} = 1|x_{it}, y_{i,t-1}, a_i) = \Phi\left(x_{it}\beta + \rho y_{i,t-1} + \alpha_0 + \alpha_1'\bar{x}_i + \alpha_2 y_{i0} + a_i\right) \equiv \Phi_f \quad (12)$$

$$f_t = \Phi_f^{y_{it}}\left[1 - \Phi_f^{1-y_{it}}\right] \quad (13)$$

This specification for the unobserved effect is often referred to as the Mundlak-Chamberlain device and allows for correlation between the exogenous variables and unobserved heterogeneity through the conditional mean. In addition, the conditional variance of the unobserved effect is assumed to be given by a function $g(x_i, y_{i0}; \gamma)$ such that homoskedasticity is obtained, $Var(a_i) = \sigma_a^2$, when exclusion restrictions are imposed on the parameters, γ. This implies that under the null the combined error term will have constant variance, $u + a \sim N(0, \sigma^2)$, as well. These restrictions on the variance of the unobserved effect encompass several commonly used structures for heteroskedasticity. This includes exponential, linear, and quadratic heteroskedasticity as special cases as given by $g(x_i, y_{i0}; \gamma) = exp(\gamma_0 + \gamma_1 x_i + \gamma_2 y_0)^2$ and $g(x_i, y_{i0}; \gamma) = (\gamma_0 + \gamma_1 x_i + \gamma_2 y_0)^k$, $k = 1, 2$, respectively. For all three cases, the exclusion restrictions $\gamma_1, \gamma_2 = 0$ imply $g(\gamma_0) = \sigma_a^2$. Note that now in $h(c|x_i, y_{i0}; \delta)$, $\delta = (\alpha, \gamma)$ where α are the parameters of the conditional mean and γ are the parameters of the conditional variance and that $a_i \sim N(0, g(x_i, y_{i0}; \gamma))$. Under these assumptions the log-likelihood can be written as

$$l_i(\theta, \delta) = log\left\{\int\left[\prod_{t=1}^{T} f_t\right]\left(\frac{1}{\sqrt{g(x_i, y_{i0}; \gamma)}}\right)\phi\left(\frac{a}{\sqrt{g(x_i, y_{i0}; \gamma)}}\right)da\right\} \quad (14)$$

Now the null hypothesis of heteroskedasticity can be tested by applying the exclusion restrictions. Under these restrictions the log likelihood function becomes

$$l_i(\theta, \delta) = log\left\{\int\left[\prod_{t=1}^{T} f_t\right]\left(\frac{1}{\sigma}\right)\phi\left(\frac{a}{\sigma}\right)da\right\} \quad (15)$$

While this is not equivalent to the random effects probit maximum likelihood estimator, consistent parameter estimates can be obtained from the random effects probit with x_{it}, $y_{i,t-1}$, 1, $\bar{x}_i$, and y_{i0} as explanatory variables.

After some algebra, the score for the unrestricted model, $s_i(\theta, \delta)$, with variance function $g(x_i, y_{i0}; \gamma) \equiv g_i$, $z_{it} = (x_{it}, y_{i,t-1})$, $w_i = (1, \bar{x}_i, y_{i0})$, $\theta = (\beta, \rho)$, $\alpha = (\alpha_0, \alpha_1, \alpha_2)$, and $\phi\left(a/\sqrt{g_i}\right) \equiv \phi_a$ can be written as

$$\frac{\partial l_i(\theta, \delta)}{\partial \theta} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\frac{(y_t - \Phi_f)\,\phi_f z_t}{\Phi_f^{1-y_t}\left[1-\Phi_f\right]^{y_t}}\right]\left(\frac{1}{\sqrt{g_i}}\right)\phi_a da\right\} \tag{16}$$

$$\frac{\partial l_i(\theta, \delta)}{\partial \alpha} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\frac{(y_t - \Phi_f)\,\phi_f w_i}{\Phi_f^{1-y_t}\left[1-\Phi_f\right]^{y_t}}\right]\left(\frac{1}{\sqrt{g_i}}\right)\phi_a da\right\} \tag{17}$$

$$\frac{\partial l_i(\theta, \delta)}{\partial \gamma} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\Phi_f^{y_{it}}\left[1-\Phi_f\right]^{1-y_{it}}\right]\left(1+\frac{a^2}{g_i}\right)\left(\frac{\nabla_\gamma \sqrt{g_i}}{g_i}\right)\phi_a da\right\} \tag{18}$$

Details for the test under exponential and quadractic heteroskedasticity are given in the following sections.

3.2.1 Panel Probit with Exponential Heteroskedasticity

Assume that $g(x_i, y_{i0}; \gamma) = exp(\gamma_0 + \gamma_1' \bar{x}_i + \gamma_2 y_{i0})^2$. Under this specification of the conditional variance

$$l_i(\theta, \delta) = log\left\{\int\left[\prod_{t=1}^{T} f_t\right]\left(\frac{1}{exp(\gamma_0 + \gamma_1'\bar{x}_i + \gamma_2 y_{i0})}\right)\phi\left(\frac{a}{exp(\gamma_0 + \gamma_1'\bar{x}_i + \gamma_2 y_{i0})}\right) da\right\} \tag{19}$$

Then heteroskedasticity in the unobserved effect can be tested as $H_0 : \gamma_1 = 0, \gamma_2 = 0$. Under the null, the restricted model estimates can be obtained from the usual random effects probit estimator with variables $x_{it}, y_{i,t-1}, 1, \bar{x}_i, y_{i0}$ as shown in (15). For $\gamma = (\gamma_0, \gamma_1, \gamma_2)$, the score for the unrestricted model can be shown to be

$$\frac{\partial l_i(\theta, \delta)}{\partial \theta} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\frac{(y_t - \Phi_f)\,\phi_f z_t}{\Phi_f^{1-y_t}\left[1-\Phi_f\right]^{y_t}}\right]\left(\frac{1}{\sqrt{g_i}}\right)\phi_a da\right\} \tag{20}$$

$$\frac{\partial l_i(\theta,\delta)}{\partial\alpha} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\frac{(y_t-\Phi_f)\,\phi_f w_i}{\Phi_f^{1-y_t}\,[1-\Phi_f]^{y_t}}\right]\left(\frac{1}{\sqrt{g_i}}\right)\phi_a da\right\} \tag{21}$$

$$\frac{\partial l_i(\theta,\delta)}{\partial\gamma} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\Phi_f^{y_{it}}\,[1-\Phi_f]^{1-y_{it}}\right]\left(1+\frac{a^2}{g_i}\right)\left(-\frac{w_i}{\sqrt{g_i}}\right)\phi_a da\right\} \tag{22}$$

(23)

Imposing the exclusion restrictions $\gamma_1, \gamma_2 = 0$, this reduces to

$$\frac{\partial l_i(\theta,\delta)}{\partial\theta} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\frac{(y_t-\Phi_f)\,\phi_f z_t}{\Phi_f^{1-y_t}\,[1-\Phi_f]^{y_t}}\right]\left(\frac{1}{\sigma_a}\right)\phi\left(\frac{a}{\sigma_a}\right)da\right\} \tag{24}$$

$$\frac{\partial l_i(\theta,\delta)}{\partial\alpha} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\frac{(y_t-\Phi_f)\,\phi_f w_i}{\Phi_f^{1-y_t}\,[1-\Phi_f]^{y_t}}\right]\left(\frac{1}{\sigma_a}\right)\phi\left(\frac{a}{\sigma_a}\right)da\right\} \tag{25}$$

$$\frac{\partial l_i(\theta,\delta)}{\partial\gamma} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\Phi_f^{y_{it}}\,[1-\Phi_f]^{1-y_{it}}\right]\left(1+\frac{a^2}{\sigma_a^2}\right)\left(-\frac{w_i}{\sigma_a}\right)\phi\left(\frac{a}{\sigma_a}\right)da\right\} \tag{26}$$

3.2.2 Panel Probit with Quadratic Heteroskedasticity

Now consider $g(x_i, y_{i0}; \gamma) = (\gamma_0 + \gamma_1'\bar{x}_i + \gamma_2 y_{i0})^2$. The response probability and conditional density of y_{it} will be the same as those given in (12) and (13) above. Then

$$l_i(\theta,\delta) = log\left\{\int\left[\prod_{t=1}^{T} f_t\right]\left(\frac{1}{(\gamma_0+\gamma_1'\bar{x}_i+\gamma_2 y_0)}\right)\phi\left(\frac{a}{(\gamma_0+\gamma_1'\bar{x}_i+\gamma_2 y_0)}\right)da\right\} \tag{27}$$

Again, the restricted parameter estimates can be obtained from random effects probit estimation on $x_{it}, y_{i,t-1}, 1, \bar{x}_i, y_{i0}$, and the null hypothesis of homoskedasticity can be expressed through the exclusion restrictions $\gamma_1 = 0, \gamma_2 = 0$.

$$\frac{\partial l_i(\theta,\delta)}{\partial\theta} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\frac{(y_t-\Phi_f)\,\phi_f z_t}{\Phi_f^{1-y_t}\,[1-\Phi_f]^{y_t}}\right]\left(\frac{1}{\sqrt{g_i}}\right)\phi_a da\right\} \tag{28}$$

$$\frac{\partial l_i(\theta,\delta)}{\partial\alpha} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\frac{(y_t-\Phi_f)\,\phi_f w_i}{\Phi_f^{1-y_t}\,[1-\Phi_f]^{y_t}}\right]\left(\frac{1}{\sqrt{g_i}}\right)\phi_a da\right\} \tag{29}$$

$$\frac{\partial l_i(\theta,\delta)}{\partial\gamma} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\Phi_f^{y_{it}}\,[1-\Phi_f]^{1-y_{it}}\right]\left(1+\frac{a^2}{g_i}\right)\left(-\frac{w_i}{g_i}\right)\phi_a da\right\} \tag{30}$$

Imposing the exclusion restrictions, this reduces to

$$\frac{\partial l_i(\theta,\delta)}{\partial\theta} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\frac{(y_t-\Phi_f)\,\phi_f z_t}{\Phi_f^{1-y_t}\left[1-\Phi_f\right]^{y_t}}\right]\left(\frac{1}{\sigma_a}\right)\phi\left(\frac{a}{\sigma_a}\right)da\right\} \tag{31}$$

$$\frac{\partial l_i(\theta,\delta)}{\partial\alpha} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\frac{(y_t-\Phi_f)\,\phi_f w_i}{\Phi_f^{1-y_t}\left[1-\Phi_f\right]^{y_t}}\right]\left(\frac{1}{\sigma_a}\right)\phi\left(\frac{a}{\sigma_a}\right)da\right\} \tag{32}$$

$$\frac{\partial l_i(\theta,\delta)}{\partial\gamma} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\Phi_f^{y_{it}}\left[1-\Phi_f\right]^{1-y_{it}}\right]\left(1+\frac{a^2}{\sigma_a^2}\right)\left(-\frac{w_i}{\sigma_a^2}\right)\phi\left(\frac{a}{\sigma_a}\right)da\right\} \tag{33}$$

(34)

3.3 Score Test for Heteroskedasticity in the Dynamic Fractional Response Model

In this section, the test statistic is extended to the case of fractional dependent variables. Recall from chapter 1 the dynamic fractional response model given by

$$y_{it}^* = x_{it}\beta + y_{i,t-1}\rho + c_i + u_{it}$$

$$u_{it}|(x_i, y_{i,t-1}, ..., y_{i0}, c_i) \sim N(0,\sigma_u^2)$$

$$y_{it} = \begin{cases} 0 & \text{if} \quad y_{it}^* \le 0 \\ y_{it}^* & \text{if} \quad 0 < y_{it}^* < 1 \\ 1 & \text{if} \quad y_{it}^* \ge 1 \end{cases}$$

With the conditional mean of c_i defined as

$$c_i|y_{i0}, x_i = \alpha_0 + \alpha_1'\bar{x}_i + \alpha_2 y_{i0} + a_i \tag{35}$$

This leads to the response probabilities

$$P(y_{it} = 0|y_{i,t-1}, x_i, y_{i0}, a_i) = \Phi\left(\frac{-x_{it}\gamma - y_{i,t-1}\rho - \alpha_0 - \alpha_1'\bar{x}_i - \alpha_2 y_{i0} - a_i}{\sigma_u}\right) \equiv \Phi_0 \tag{36}$$

$$P(y_{it} = 1|y_{i,t-1}, x_i, y_{i0}, a_i) = \Phi\left(\frac{x_{it}\gamma + y_{i,t-1}\rho + \alpha_0 + \alpha_1'\bar{x}_i + \alpha_2 y_{i0} + a_i - 1}{\sigma_u}\right) \equiv \Phi_1 \tag{37}$$

$$\frac{\partial P(y_{it} \leq y|y_{i,t-1}, x_i, y_{i0}, a_i)}{\partial y} = \frac{1}{\sigma_u}\phi\left(\frac{y_{it} - x_{it}\gamma - y_{i,t-1}\rho - \alpha_0 - \alpha_1'\bar{x}_i - \alpha_2 y_{i0} - a_i}{\sigma_u}\right) \equiv \frac{1}{\sigma_u}\phi_y \quad (38)$$

From which the density of $y_{it}|y_{i,t-1}, x_i, y_{i0}, a_i$ can be written as

$$f_t(y_{it}|y_{i,t-1}, x_i, y_{i0}, a_i; \theta) = \Phi_0^{I[y_{it}=0]}\Phi_1^{I[y_{it}=1]}\frac{1}{\sigma_u}\phi_y^{I[0<y_{it}<1]} \quad (39)$$

By using this density in place of (13), the test outlined above can be implemented. Under the restricted model the log-likelihood function is obtained by integrating the density of $(y_{i1}, ..., y_{iT})$ given $(y_{i,t-1}, x_i, y_{i0}, a_i)$ against the distribution of a_i

$$l_i(\theta, \delta) = log\left\{\int\left[\prod_{t=1}^{T} f_t(y_{it}|y_{i,t-1}, x_i, y_{i0}, a; \theta)\right]\frac{1}{\sigma_a}\phi\left(\frac{a}{\sigma_a}\right)da\right\} \quad (40)$$

Thus, the log-likelihood function has the same general form as in (15) above, and the restricted model parameter estimates can be obtained from the random effects two-limit tobit. The score for the unrestricted model follows similarly to that of the dynamic probit and is given by

$$\frac{\partial l_i}{\partial \theta} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\Gamma\left(-\frac{z_{it}}{\sigma_u}\right)\left(1[y_{it}=0]\frac{\phi_1}{\Phi_1} + 1[y_{it}=1]\frac{\phi_2}{\Phi_2} + 1[0<y_{it}<1]\frac{z_{it}\theta}{\sigma^u}\right)\right]Ada\right\} \quad (41)$$

$$\frac{\partial l_i}{\partial \alpha} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\Gamma\left(-\frac{w_i}{\sigma_u}\right)\left(1[y_{it}=0]\frac{\phi_1}{\Phi_1} + 1[y_{it}=1]\frac{\phi_2}{\Phi_2} + 1[0<y_{it}<1]\frac{w_i\alpha}{\sigma^u}\right)\right]Ada\right\} \quad (42)$$

$$\frac{\partial l_i}{\partial \sigma_u} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\Gamma\left(\frac{1}{\sigma_u^2}\right)\left(-1[y_{it}=0]\frac{\phi_1}{\Phi_1}\xi_1 - 1[y_{it}=1]\frac{\phi_2}{\Phi_2}\xi_2 + 1[0<y_{it}<1]\frac{\xi_3^2}{\sigma^u}\right)\right]Ada\right\} \quad (43)$$

$$\frac{\partial l_i}{\partial \gamma} = \frac{1}{\Psi}\left\{\int\left[\prod_{t=1}^{T}\Gamma\right]\left(1+\frac{a^2}{g_i}\right)\left(\frac{\nabla_\gamma\sqrt{g_i}}{g_i}\right)\phi_a da\right\} \quad (44)$$

where

$$\Gamma = \Phi_0^{I[y_{it}=0]}\Phi_1^{I[y_{it}=1]}\frac{1}{\sigma_u}\phi_y^{I[0<y_{it}<1]}$$

$$\xi_1 = (-x_{it}\gamma - y_{i,t-1}\rho - \alpha_0 - \alpha_1'\bar{x}_i - \alpha_2 y_{i0} - a_i)$$

$$\xi_2 = (x_{it}\gamma + y_{i,t-1}\rho + \alpha_0 + \alpha_1'\bar{x}_i + \alpha_2 y_{i0} + a_i - 1)$$

$$\xi_3 = (y_{it} - x_{it}\gamma - y_{i,t-1}\rho - \alpha_0 - \alpha_1'\bar{x}_i - \alpha_2 y_{i0} - a_i)$$

$$A = \left(\frac{1}{g_i}\right)\phi_a$$

The score test for heteroskedasticity can then be computed for any form of the conditional variance such that $a_i \sim N(0, g(x_i, y_{i0}; \gamma))$ where the variance is constant under the null hypothesis. In this case, the variance is assumed to exhibit exponential heteroskedasticity as specified in section 3.2.1.

3.4 Empirical Applications

3.4.1 Dynamic Probit Model: Persistence of Union Membership

Wooldridge (2005a) estimates a simple model of union membership to illustate the method used above for simultaneously dealing with initial conditions and unobserved heterogeneity in nonlinear, dynamic panel data models. Here, the same application is used to demonstate the test for heteroskedasticity in this setting derived above. The data consists of observations on 545 working men over an 8 year period from 1980 to 1987 with indicator variables for union status, $y_{it} = 0, 1$, and marital status, $x_{it} = 0, 1$. Applying the dynamic probit model to union membership,

$$P(y_{it} = 1 | y_{i,t-1}, ..., y_{i0}, x_i, c_i) = \Phi(\rho y_{i,t-1} + \beta x_{it} + c_i) \quad (45)$$

In addition, a set of year dummies are included to control for changes occurring over time that are common to all workers. Estimation of the restricted model, can be carried out using standard software that allows for random effects probit specifications, and the density of c_i is specified as $c_i = \alpha_0 + \alpha_1' x_i + \alpha_2 y_{i0} + a_i$ and $a_i \sim N(0, \sigma_a^2)$.

The parameter estimates under the restricted model are given in Table 3.1. Only the estimated coefficients for the primary variables of interest are included for brevity. The quantities given in parentheses are standard errors. Given these parameter estimates, the

Table 3.1: Restricted Estimates for the Dynamic Probit Model

Union Membership	Coefficient Estimate
Lagged Union Membership	**0.875**
	(0.094)
Marital Status	0.168
	(0.111)
Test Statistic	10.31
df	7
5% critical value	14.07

Number of observations = 4360
Number of workers = 545

Note: Bold type indicates significance at the 1% level, italics 5%.
Quantity in parentheses is standard error.

scores of the unrestricted model are evaluated at the restricted estimates and the score statistic calculated. The statistic is also shown in the table and is distributed χ^2_2, making the critical value for the test approximately 14 at the 5% level. Thus, the test fails to reject the null hypothesis of homoskedasticity.

3.4.2 Dynamic Fractional Response Model: Determination of Firm Dividend Policy

The dividend policy application from chapter 1 will be used to demonstrate the test as applied to the fractional response model. Recall that in this application y_{it} is the share repurchase ratio, x_{it} is made up of the exogenous covariates (i.e, market-to-book ratio, operating income, non-operating income, and volatility), and c_i is a firm specific unobserved effect. In order to control for industry level differences, the exogenous variables are augmented with 10 industry dummy variables. In addition, a set of year dummies are included to control for changes occurring over time that are common to all firms.[3] As before, estimation of the restricted model, (40), can be carried out using standard software that allows for two-limit random effects tobit specifications if the density of c_i is specified as in (35) and $a_i \sim N(0, \sigma_a^2)$. Here $\bar{x}_i$ includes the time averages of the exogenous regressors (i.e., market-to-book ratio, operating income ratio, non-operating income ratio, and volatility of earnings).

The parameter estimates under the restricted model are given in Table 3.2. Only the estimated coefficients for the primary variables of interest are included for brevity. The quantities given in parentheses are standard errors. Given these parameter estimates, the scores of the unrestricted model are evaluated at the restricted estimates and the score statistic calculated. The statistic is also shown in the table and is distributed χ_2^2, making the critical value for the test approximately 11 at the 5% level. Therefore, the test clearly rejects the null hypothesis of homoskedasticity with a value above 24. This may imply that the model is heteroskedastic.

However, rejecting homoskedasticity is also potentially compatible with inclusion of a regressor that is correlated with the error due to omitted variables or measurement error. Consider the following probit model in which the i and t subscripts have been supressed

$$y_1 = 1(z_1\eta_1 + \tau y_2 + u_1 > 0) \tag{46}$$

$$y_2 = z_1\eta_{21} + z_2\eta_{22} + v_2 \tag{47}$$

where y_2 is continuous and (u_1, v_2) has a zero mean, bivariate normal distribution and is independent of z. Thus, y_2 is endogenous if u_1 and v_2 are correlated. In this case, if only (46) is estimated the test statistic could reject the null hypothesis of constant variance even if u_1 and v_2 are homoskedastic. It is possible to estimate this model and test for exogeneity of y_2 by using a two-step procedure such as that of Rivers and Vuong (1988). Given the

[3] For additional details on firm dividend policy theory and data set construction see sections 1.2 and A.

Table 3.2: Restricted Estimates for the Dynamic Fractional Response Model

Share Repurchase Ratio	Coefficient Estimate
Lagged Repurchase Ratio	**0.483**
	(0.025)
Market-to-book Value	-0.002
	(0.001)
Operating Income Ratio	**0.644**
	(0.120)
Non-operating Income Ratio	0.185
	(0.243)
Volatility of Earnings	1.032
	(0.666)
Test Statistic	24.33
df	5
5% critical value	11.07

Number of observations = 4530
Number of firms = 453

Note: Bold type indicates significance at the 1% level, italics 5%.
Quantity in parentheses is standard error.

joint normality of (u_1, v_2) and $Var(u_1) = 1$

$$u_1 = \omega_1 v_2 + e_1 \tag{48}$$

where e_1 is normally distributed and $E(e_1) = 0$. Now

$$y_1 = 1(z_1\eta_1 + \tau y_2 + \omega_1 v_2 + e_1 > 0) \tag{49}$$

and $e_1|z, y_2, v_2$ is normally distributed with zero mean and constant variance. Since v_2 is not observed, this model is estimated in two steps: 1) Perform an OLS regression of y_2 on z and save the residuals, $\hat{v_2}$; and 2) Estimate the probit model of y_1 on z_1, y_2, and $\hat{v_2}$. The exogeneity of y_2 is then tested with the usual t-statistic for the hypothesis $\omega_1 = 0$. However, if y_2 is endogenous the usual standard errors and t-statistics will not be valid, and the asymptotic variance of the two-step estimator can be dervied using standard results for maxumum likelihood estimation.

3.5 Discussion

This paper presents a score test for heteroskedasticity in the dynamic probit and tobit panel models with unobserved heterogeneity through the use of exclusion restrictions. While the test is derived under the assumption of normally distributed unobserved heterogeneity and requires specification of the conditional mean, it leads to a computationally simple test that is straightforward to implement. This is demonstrated through an application of the test to a model of firm dividend policy using the dynamic, fractional response model.

There are several avenues open for additional research. The power properties of the test need to be studied and an evaluatation of its performance against different types of heteroskedasticity should be performed. In addition, a test for non-normality of the conditional distribution of the unobserved effect might also be developed along similar lines, which would allow testing for this underlying assumption.

APPENDIX A

Chapter 1 Summary Statistics & Variable Definitions

Table A.1: Compustat Quarterly & Annual Data Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
Share Repurchase Ratio	0.281	0.362	0.000	1.000
Market-to-book Value	6.191	154.095	-533.895	10,473.960
Operating Income Ratio	0.148	0.173	-1.419	4.864
Non-operating Income Ratio	0.001	0.026	-0.388	1.080
Volatility of Earnings	0.013	0.021	0.000	0.541

Table A.2: Industry Dummy Variable Definitions

Industry Dummy	4-digit SIC Code
Agriculture	0000-1000
Mining & Construction	1000-1099, 1200-1799
Manufacturing	2000-3999
Transportation	4000-4799
Communications	4800-4899
Utility	4900-4999
Wholesale	5000-5199
Retail	5200-5999
Financial	6000-6799
Services	7000-8999

APPENDIX B

Chapter 2 Summary Statistics & Parameter Estimates

Table B.1: A.C. Nielsen Yogurt Data Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
Brand	1.23	1.24	0	3
Yoplait price per oz	0.10	0.01	0.04	0.13
Yoplait special	0.01	0.11	0	1
Yoplait featured	0.03	0.18	0	1
Dannon price per oz	0.07	0.01	0.04	0.08
Dannon special	0.04	0.19	0	1
Dannon featured	0.17	0.38	0	1
Nordica price per oz	0.08	0.01	0.06	0.10
Nordica special	0.00	0.00	0	0
Nordica featured	0.03	0.18	0	1

Table B.2: Parameter estimates for Table 2.1

Estimated Process	Coefficient Estimates	
(A) T=5, N=100	x_1	x_2
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	**0.178** **(0.079)**	1.075 (0.101)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2)$	**0.172** **(0.075)**	1.035 (0.098)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.171** **(0.077)**	1.022 (0.102)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	*0.177* *(0.093)*	1.071 (0.147)
(B) T=5, N=500	x_1	x_2
$c_i \| x_i \sim N(\psi + \bar{x}_i, \sigma_a^2)$	0.178 (0.033)	1.064 (0.045)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2)$	0.171 (0.034)	1.018 (0.044)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.173** **(0.050)**	1.014 (0.055)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	*0.171* *(0.041)*	1.052 (0.044)
(C) T=25, N=100	x_1	x_2
$c_i \| x_i \sim N(\psi + \bar{x}_i, \sigma_a^2)$	0.178 (0.032)	1.064 (0.042)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2)$	0.178 (0.031)	1.065 (0.044)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.176 (0.033)	1.054 (0.043)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	*0.171* *(0.041)*	1.052 (0.044)

continued on next page

Table B.2 (continued)

Estimated Process	Coefficient Estimates	
(D) T=25, N=500	x_1	x_2
$c_i \| x_i \sim N(\psi + \bar{x}_i, \sigma_a^2)$	0.177 (0.014)	1.061 (0.020)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2)$	0.176 (0.014)	1.062 (0.019)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.175 (0.015)	1.051 (0.020)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.175 (0.016)	1.051 (0.020)

Quantity in parentheses is standard error.

Estimates in bold are significant at the 5% level & $\hat{\beta}$ equals its true value (italics 10%).

Table B.3: Parameter estimates for Table 2.2

Estimated Process	Coefficient Estimates		
(i)	$P(y_{i0}\|x_i, c_i) = P(y_{i0})$		
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.172** **(0.080)**	1.028 (0.100)	0.159 (0.164)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2)$	**0.174** **(0.081)**	1.046 (0.104)	0.152 (0.162)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.187 (0.026)	**1.106** **(0.214)**	0.187 (0.319)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.245 (0.174)	**1.284** **(0.327)**	0.253 (0.448)
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.172** **(0.056)**	1.018 (0.044)	**0.176** **(0.083)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2)$	**0.174** **(0.079)**	1.033 (0.105)	0.157 (0.167)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.215** **(0.032)**	*1.123* *(0.207)*	0.205 (0.316)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.261 (0.160)	**1.348** **(0.309)**	0.252 (0.520)
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	0.176 (0.032)	1.054 (0.042)	**0.174** **(0.071)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2)$	0.175 (0.033)	1.056 (0.047)	**0.173** **(0.075)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.176 (0.031)	1.056 (0.044)	**0.175** **(0.070)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i{}^2 + \psi_3\bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*0.179* *(0.039)*	1.069 (0.061)	**0.173** **(0.085)**

continued on next page

Table B.3 (continued)

Estimated Process	Coefficient Estimates		
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	0.174 (0.014)	1.041 (0.019)	0.174 (0.030)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2)$	*0.177* *(0.041)*	1.052 (0.025)	**0.173** **(0.054)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.174 (0.015)	1.052 (0.020)	0.175 (0.032)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.183** **(0.052)**	1.070 (0.045)	**0.179** **(0.079)**

continued on next page

Table B.3 (continued)

Estimated Process	Coefficient Estimates		
(ii)	$P(y_{i0} = 1 \mid x_i, c_i) = \Phi(x_{i0}\beta)$		
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	**0.161** **(0.067)**	0.977 (0.091)	0.156 (0.150)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2)$	**0.166** **(0.072)**	0.990 (0.090)	0.143 (0.155)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	*0.180* *(0.104)*	1.040 (0.196)	0.194 (0.315)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.204 (0.144)	**1.161** **(0.245)**	0.205 (0.466)
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	0.160 (0.039)	0.970 (0.040)	**0.163** **(0.074)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2)$	**0.163** **(0.068)**	0.985 (0.097)	0.142 (0.161)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	*0.190* *(0.103)*	1.099 (0.245)	0.217 (0.406)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.212 (0.139)	**1.196** **(0.241)**	0.242 (0.426)
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	0.175 (0.031)	1.045 (0.043)	**0.172** **(0.067)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2)$	0.176 (0.033)	1.046 (0.045)	**0.171** **(0.073)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.176 (0.031)	1.043 (0.043)	**0.172** **(0.070)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.181** **(0.058)**	1.058 (0.057)	*0.180* *(0.097)*

continued on next page

Table B.3 (continued)

Estimated Process	Coefficient Estimates		
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	0.173 (0.014)	1.040 (0.019)	0.173 (0.031)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2)$	**0.176** **(0.048)**	1.041 (0.024)	**0.173** **(0.059)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.173 (0.014)	1.040 (0.019)	0.173 (0.030)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.172 (0.026)	1.059 (0.046)	**0.182** **(0.061)**

continued on next page

Table B.3 (continued)

Estimated Process	Coefficient Estimates		
(iii)	$P(y_{i0} = 1 \mid x_i, c_i) = \Phi(\eta_0 + \eta_1 \bar{x}_i + \eta_2 c_i)$		
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	**0.176** **(0.080)**	1.033 (0.110)	0.206 (0.205)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2)$	0.190 (0.121)	**1.118** **(0.195)**	0.201 (0.402)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.251** **(0.178)**	**1.234** **(0.312)**	0.481 (0.512)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.297 (0.208)	**1.435** **(0.373)**	0.482 (0.560)
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	**0.180** **(0.062)**	1.017 (0.055)	**0.223** **(0.156)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2)$	*0.196* *(0.103)*	*1.150* *(0.200)*	0.218 (0.395)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.250 (0.190)	**1.277** **(0.305)**	0.537 (0.481)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	*0.259* *(0.144)*	**1.177** **(0.253)**	0.374 (0.402)
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	0.179 (0.034)	1.056 (0.048)	**0.185** **(0.083)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2)$	*0.182* *(0.037)*	1.072 (0.048)	**0.196** **(0.097)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	*0.182* *(0.038)*	1.066 (0.062)	*0.201* *(0.114)*
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.189** **(0.067)**	1.083 (0.074)	0.202 (0.147)

continued on next page

Table B.3 (continued)

Estimated Process	Coefficient Estimates		
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	0.178 (0.020)	1.052 (0.026)	0.182 (0.057)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2)$	**0.182** **(0.042)**	1.066 (0.031)	**0.196** **(0.080)**
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.182 (0.024)	1.062 (0.037)	0.202 (0.091)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.184 (0.030)	1.086 (0.058)	0.194 (0.118)

Quantity in parentheses is standard error.
Estimates in bold are significant at the 5% level & $\hat{\beta}$ equals its true value (italics 10%).

Table B.4: Parameter estimates for Table 2.3

Estimated Process	Coefficient Estimates		
(i)	$P(y_{i0}\|x_i, c_i) = P(y_{i0})$		
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	**0.174** **(0.085)**	1.310 (0.109)	**0.503** **(0.178)**
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2)$	**0.174** **(0.082)**	1.044 (0.107)	*0.499* *(0.171)*
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*0.193* *(0.114)*	*1.118* *(0.212)*	*0.562* *(0.333)*
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.238 (0.172)	**1.269** **(0.300)**	0.593 (0.452)
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	0.169 (0.035)	1.014 (0.048)	0.509 (0.075)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2)$	*0.174* *(0.040)*	1.023 (0.059)	0.504 (0.087)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*0.187* *(0.113)*	**1.148** **(0.220)**	*0.671* *(0.364)*
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*0.259* *(0.148)*	**1.323** **(0.291)**	0.640 (0.429)
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2)$	0.178 (0.035)	1.054 (0.046)	0.524 (0.077)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2)$	0.174 (0.035)	1.056 (0.048)	0.523 (0.078)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.177 (0.035)	1.054 (0.047)	0.526 (0.074)
$c_i\|x_i \sim N(\psi_0 + \psi_1'\bar{x}_i + \psi_2\bar{x}_i^2 + \psi_3\bar{x}_i^3, \sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.179** **(0.044)**	1.061 (0.049)	0.522 (0.084)

continued on next page

Table B.4 (continued)

Estimated Process	Coefficient Estimates		
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \vert x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	0.176 (0.015)	1.051 (0.021)	0.525 (0.034)
$c_i \vert x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2)$	**0.177** **(0.049)**	1.052 (0.026)	0.531 (0.051)
$c_i \vert x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.175 (0.015)	1.051 (0.021)	0.525 (0.036)
$c_i \vert x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.179** **(0.049)**	1.058 (0.027)	0.526 (0.058)

continued on next page

Table B.4 (continued)

Estimated Process	Coefficient Estimates		
(ii)	$P(y_{i0}=1\|x_i,c_i)=\Phi(x_{i0}\beta)$		
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2)$	**0.163** **(0.070)**	0.975 (0.091)	**0.493** **(0.167)**
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2)$	**0.162** **(0.072)**	0.985 (0.094)	0.479 (0.169)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*0.175* *(0.102)*	**1.030** **(0.174)**	0.513 (0.318)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*(0.204)* *(0.118)*	**(1.139)** **(0.241)**	(0.575) (0.451)
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2)$	0.159 (0.030)	0.963 (0.040)	0.489 (0.074)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2)$	0.162 (0.036)	0.973 (0.059)	0.487 (0.087)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	*0.193* *(0.100)*	1.068 (0.193)	*0.582* *(0.343)*
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	**0.205** **(0.098)**	**1.193** **(0.239)**	0.660 (0.475)
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2)$	0.174 (0.035)	1.040 (0.046)	0.521 (0.076)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2)$	0.176 (0.035)	1.044 (0.047)	0.520 (0.081)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.174 (0.033)	1.040 (0.044)	0.518 (0.074)
$c_i\|x_i \sim N(\psi_0+\psi_1'\bar{x}_i+\psi_2\bar{x}_i^2+\psi_3\bar{x}_i^3,\sigma_a^2 exp(\lambda_1'\bar{x}_i))$	0.178 (0.036)	1.050 (0.052)	0.523 (0.089)

continued on next page

Table B.4 (continued)

Estimated Process	Coefficient Estimates		
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	0.172 (0.014)	1.038 (0.020)	0.521 (0.034)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2)$	0.173 (0.031)	1.039 (0.024)	0.523 (0.045)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.173 (0.015)	1.038 (0.020)	0.522 (0.034)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.176 (0.031)	1.044 (0.029)	0.523 (0.056)

continued on next page

Table B.4 (continued)

Estimated Process	Coefficient Estimates		
(iii)	$P(y_{i0} = 1 \| x_i, c_i) = \Phi(\eta_0 + \eta_1 \bar{x}_i + \eta_2 c_i)$		
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	**0.175** **(0.086)**	1.029 (0.114)	**0.528** **(0.199)**
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2)$	0.181 (0.124)	1.144 (0.220)	0.508 (0.380)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.255 (0.187)	**1.283** **(0.335)**	*0.827* *(0.495)*
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.288 (0.201)	**1.420** **(0.356)**	0.791 (0.508)
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	*0.174* *(0.041)*	1.015 (0.056)	**0.537** **(0.133)**
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2)$	0.186 (0.120)	*1.131* *(0.219)*	0.581 (0.375)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.277 (0.190)	**1.343** **(0.342)**	*0.874* *(0.488)*
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.313 (0.193)	**1.572** **(0.320)**	0.763 (0.551)
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	0.179 (0.036)	1.055 (0.049)	0.528 (0.087)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2)$	0.176 (0.036)	1.059 (0.049)	0.536 (0.091)
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	*0.181* *(0.036)*	1.063 (0.056)	*0.555* *(0.115)*
$c_i \| x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i^2 + \psi_3 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	*0.182* *(0.039)*	1.069 (0.061)	**0.558** **(0.135)**

continued on next page

Table B.4 (continued)

Estimated Process	Coefficient Estimates		
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2)$	0.176 (0.020)	1.052 (0.026)	0.531 (0.060)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2)$	0.178 (0.031)	1.054 (0.025)	0.534 (0.048)
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	0.179 (0.025)	1.056 (0.038)	*0.558* *(0.103)*
$c_i \mid x_i \sim N(\psi_0 + \psi_1' \bar{x}_i + \psi_2 \bar{x}_i{}^2 + \psi_3 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i))$	**0.185** **(0.052)**	1.072 (0.048)	*0.557* *(0.116)*

Quantity in parentheses is standard error.

Estimates in bold are significant at the 5% level & $\hat{\beta}$ equals its true value (italics 10%).

Table B.5: Parameter Estimates for Table 2.4

Estimated Process	Coefficient Estimates		
		$\rho = .25$	
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	**0.176**	1.037	0.163
	(0.084)	(0.119)	(0.184)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$	0.241	**1.237**	0.245
	(0.160)	**(0.311)**	(0.375)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x_i}^2))$	0.299	**1.469**	0.307
	(0.240)	**(0.449)**	(0.628)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x_i}^2))$	0.384	**1.593**	0.423
	(0.269)	**(0.445)**	(0.611)

(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	0.178	1.031	0.164
	(0.029)	(0.044)	(0.075)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$	**0.190**	1.090	0.181
	(0.046)	(0.082)	(0.133)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x_i}^2))$	0.178	1.028	0.164
	(0.029)	(0.045)	(0.074)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x_i}^2))$	0.178	1.028	0.164
	(0.029)	(0.045)	(0.074)

continued on next page

Table B.5 (continued)

Estimated Process	Coefficient Estimates		
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	0.180 (0.039)	1.060 (0.058)	0.174 (0.089)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i{}^2 + \alpha_4 \bar{x}_i{}^3, \sigma_a^2)$	**0.190** **(0.046)**	1.090 (0.082)	0.181 (0.133)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x_i}^2))$	**0.194** **(0.076)**	1.102 (0.140)	0.181 (0.206)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i{}^2 + \alpha_4 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x_i}^2))$	**0.189** **(0.064)**	1.114 (0.113)	0.197 (0.217)
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	0.178 (0.031)	1.058 (0.041)	0.170 (0.082)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i{}^2 + \alpha_4 \bar{x}_i{}^3, \sigma_a^2)$	**0.191** **(0.053)**	*1.093* *(0.086)*	0.183 (0.127)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x_i}^2))$	**0.196** **(0.078)**	**1.115** **(0.155)**	0.189 (0.101)
$c_i\|x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i{}^2 + \alpha_4 \bar{x}_i{}^3, \sigma_a^2 exp(\lambda_1' \bar{x_i}^2))$	**0.187** **(0.063)**	*1.106* *(0.106)*	0.196 (0.088)

continued on next page

Table B.5 (continued)

Estimated Process	Coefficient Estimates		
		$\rho = .75$	
(A) T=5, N=100	x_1	x_2	y_{t-1}
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	**0.178**	1.035	0.497
	(0.088)	(0.118)	(0.181)
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$	0.232	**1.207**	0.596
	(0.142)	**(0.258)**	(0.361)
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	0.289	**1.403**	0.691
	(0.244)	**(0.411)**	(0.605)
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	0.344	**1.549**	0.749
	(0.237)	**(0.390)**	(0.594)
(B) T=5, N=500	x_1	x_2	y_{t-1}
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	0.173	1.029	0.519
	(0.035)	(0.051)	(0.085)
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$	**0.204**	1.058	0.511
	(0.041)	(0.061)	(0.099)
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	0.168	1.029	0.516
	(0.035)	(0.054)	(0.092)
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	0.168	1.029	0.516
	(0.035)	(0.054)	(0.092)

continued on next page

Table B.5 (continued)

Estimated Process	Coefficient Estimates		
(C) T=25, N=100	x_1	x_2	y_{t-1}
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	0.177	1.058	0.527
	(0.38)	(0.051)	(0.084)
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$	**0.191**	1.095	*0.535*
	(0.059)	(0.090)	*(0.117)*
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	**0.198**	1.128	**0.540**
	(0.079)	(0.170)	**(0.208)**
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	**0.185**	1.102	0.534
	(0.061)	(0.098)	(0.195)
(D) T=25, N=500	x_1	x_2	y_{t-1}
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2)$	0.175	1.056	0.509
	(0.011)	(0.023)	(0.027)
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2)$	**0.190**	1.090	**0.537**
	(0.048)	(0.081)	**(0.107)**
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	**0.196**	1.200	**0.551**
	(0.063)	(0.044)	**(0.099)**
$c_i \mid x_i, y_{i0} \sim N(\alpha_0 + \alpha_1 y_{i0} + \alpha_2 \bar{x}_i + \alpha_3 \bar{x}_i^2 + \alpha_4 \bar{x}_i^3, \sigma_a^2 exp(\lambda_1' \bar{x}_i^2))$	**0.187**	1.117	*0.544*
	(0.044)	(0.036)	*(0.109)*

Quantity in parentheses is standard error.

Estimates in bold are significant at the 5% level & $\hat{\beta}$ equals its true value (italics 10%).

APPENDIX C

Chapter 3 Summary Statistics

Table C.1: Union Membership Data Summary Statistics

Variable	Mean	Std. Dev.	Min	Max
Union Membership	0.244	0.430	0	1
Marital Staus	0.439	0.496	0	1

BIBLIOGRAPHY

[1] Arellano, M. and S. Bond, 1991, Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *The Review of Economic Studies*, 58, 277-297.

[2] Arellano, M. and R. Carrasco, 2003, Binary Choice Panel Data Models with Predetermined Variables, *Journal of Econometrics*, 115, 125-157.

[3] Baltagi, B.H., 2001, *Econometric Analysis of Panel Data, 2nd Edition*, New York, NY: John Wiley & Sons.

[4] Bera, A.K. and C.R. McKenzie, 1986, Alternative Forms and Properties of the Score Test, *Journal of Applied Statistics*, 13, 13-25.

[5] Bera, A.K., Jarque, C.M., and L.F. Lee, 1984, Testing the Normality Assumption in Limited Dependent Variable Models, *International Economic Review*, 25, 563-578.

[6] Breusch, T.S. and A.R. Pagan, 1979, A Simple Test for Heteroskedasticity and Random Coefficient Variation, *Econometrica*, 47, 1287-1294.

[7] Breusch and Pagan, 1980, The Lagrange Multiplier Test and its Applications to Model Specification in Econometrics, *Review of Economic Studies*, 47, 239-253.

[8] Chamberlain, G., 1980, Analysis of Covariance with Qualitative Data, *The Review of Economic Studies*, 47, 225-238.

[9] Chay, K. and D. Hyslop, 2000, Identification and Estimation of Binary Response Panel data Models: Empirical Evidence using Alternative Approaches, mimeo., University of California at Berkley Department of Economics.

[10] Chintagunta, P., Kyriazidou, E. and J. Perktold, 2001, Panel Data Analysis of Household Brand Choices, *Journal of Econometrics*, 103, 111-153.

[11] Davidson, R. and J.G. MacKinnon, 1984, Model Specification Tests Based on Artificial Linear Regressions, *International Economic Review*, 25, 485-502.

[12] Engle, R.F., 1984, Wald, Likelihood Ratio, and Lagrange Multiplier Tests in Econometrics, *Handbook of Econometrics*, 2, ed. Z. Griliches and M.D. Intriligator, Elsevier Science Publishers: New York, 775-826.

[13] Fenn, G.W. and N. Liang, 2001, Corporate Payout Policy and Managerial Stock Incentives, *Journal of Financial Economics*, 60, 45-72.

[14] Godfrey, L.G., 1981, On the Invariance of the Lagrange Multiplier Test with Respect to Certain Changes in the Alternative Hypothesis, *Econometrica*, 49, 1443-1455.

[15] Hahn, J. and W. Newey, 2004, Jackknife and Analytical Bias Reduction for Nonlinear Panel Models, *Econometrica*, 72, 1295-1319.

[16] Ham, J.C. and R.J. Lalonde, 1996, The Effect of Sample Selection and Initial Conditions in Duration Models: Evidence from Experimental Data on Training, *Econometrica*, 64, 175-205.

[17] Hausman, J.A. and G.K. Leonard, 1997, Superstars in the National Basketball Association: Economic Value and Policy, *Journal of Labor Economics*, 15, 586-624.

[18] Heckman, J.J., 1981, The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process, in *Structural Analysis of Discrete Data with Econometric Applications*, ed. C.F. Manski and D. McFadden, MIT Press: Cambridge, MA, 179-195.

[19] Heckman, J.J., 2001, Micro Data Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture, *Journal of Political Economy*, 109, 673-748.

[20] Honore, B.E., 1993, Orthogonality Conditions for Tobit Models with Fixed Effects and Lagged Dependent Variables, *Journal of Econometrics*, 59, 36-61.

[21] Honore, B.E. and E. Kyriazidou, 2000a, Estimation of Tobit Type Models with Individual Specific Effects, *Econometric Reviews*, 19, 341-366.

[22] Honore, B.E. and E. Kyriazidou, 2000b, Panel Data Discrete Choice Models with Lagged Dependent Variables, *Econometrica*, 68, 839-874.

[23] Honore, B.E. and A. Lewbel, 2002, Semiparametric Choice Panel Data Models Without Stictly Exogenous Regressors, *Econometrica*, 70, 2053-2063.

[24] Honore, B.E. and E. Tamer, 2003, Bounds on Parameters in Dynamic Discrete Choice Models, Working Paper, Princeton University.

[25] Horowitz, J.L., 1997, Bootstrap Methods in Econometrics: Theory and Numerical Performance, in *Advances in Economics and Econometrics: Theory and Applications*, 3, ed. D.M. Kreps and K.F. Wallis, Cambridge: Cambridge University Press, 188-222.

[26] Hsiao, C., 1986, *Analysis of Panel Data*, Cambridge University Press: Cambridge.

[27] Hyslop, D.R., 1999, State Dependence, Serial Correlation, and Heterogeneity in Intertemporal Labor Force Participation of Married Women, *Econometrica*, 67, 1255-1294.

[28] Jagannathan, M., Stephans, C., and M. Weisbach, 2000, Financial Flexibility and the Choice between Dividends and Stock Repurchases, *Jounral of Financial Economics*, 57, 355-384.

[29] Lechner, M., 1995, Some Specification Tests for Probit Models Estimated on Panel Data, *Journal of Business & Economic Statistics*, 13, 475-488.

[30] Lee, C.J., 2000, Dynamic Unobserved Effects Model for Continuous and Binary Response, Dissertation, Michigan State University.

[31] Lintner, J., 1956, Distribution of Incomes of Corporations among Dividends, Retained Earnings, and Taxes, *American Economic Review*, 46, 97-113.

[32] MacKinnon, J.G., 1992, Model Specification Tests and Artificial Regressions, *Journal of Economic Literature*, 30, 102-146.

[33] Maddala, G.S., 1983, *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press: Cambridge.

[34] Manos, R., 2002, Dividend Policy and Agency Theory: Evidence on Indian Firms, University of Manchester Institute for Development Policy and Management, Finance and Development Research Programme Working Paper No.41

[35] Manski, C.F., 1987, Semiparametric Analysis of Random Effects Linear Models from Binary Panel Data, *Econometrica*, 55, 357-362.

[36] Moh'd, M.A., Perry, L.G. and J.N. Rimbley, 1995, An Investigation of the Dynamic Relationship between Agency Theory and Dividend Policy, *The Financial Review*, 30, 367-385.

[37] Moon, C.G., 1988, Simultaneous Specification Test in a Binary Logit Model: Skewness and Heteroskedasticity, 17, 3361-3387.

[38] Mundlak, Y., 1978, On the Pooling of Time Series and Cross Section Data, *Econometrica*, 46, 69-85.

[39] Papke, L.E., 1994, Tax Policy and Urban Development: Evidence from the Indiana Enterprise Zone Program, *Journal of Public Economics*, 54, 37-49.

[40] Papke, L.E. and J.M. Wooldridge, 1996, Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates, *Journal of Applied Econometrics*, 11, 619-632.

[41] Rivers, D. and Q.H. Vuong, 1988, Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models, *Journal of Econometrics*, 39, 347-366.

[42] Savin, N.E. and A.H. Wurtz, 1999, Power Tests in Binary Response Models, *Econometrica*, 67, 413-421.

[43] Wooldridge, J.M. (1991), On the Application of Robust Regression-Based Diagnostic Models of Conditional Means and Conditional Variances, *Journal of Econometrics*, 47, 5-46.

[44] Wooldridge, J.M. (1991), Specification Testing and Quasi-Maximum-Likelihood Estimation, *Journal of Econometrics*, 48, 29-55.

[45] Wooldridge, J.M., 1995, Selection Corrections for Panel Data Models under Conditional Mean Independence Assumptions, *Journal of Econometrics*, 68, 115-132.

[46] Wooldridge, J.M., 1997, Multiplicative Panel Data Models without the Strict Exogeneity Assumption, *Econometric Theory* 13, 667-678.

[47] Wooldridge, J.M., 2000, A Framework for Estimating Dynamic, Unobserved Effects Panel Data Models with Possible Feedback to Future Explanatory Variables, *Economics Letters*, 68, 245-250.

[48] Wooldridge, J.M., 2002, *Econometric Analysis of Cross Section and Panel Data*, MIT Press: Cambridge, MA.

[49] Wooldridge, J.M., 2002, Inverse Probability Weighted M-Estimators for Sample Selection, Attrition, and Stratification, *Portuguese Economic Journal*, 1, 117-139.

[50] Wooldridge, J.M. (2005a), Simple Solutions to the Initial Conditions Problem in Dynamic, Nonlinear, Panel Data Models With Unobserved Heterogeneity, *Journal of Applied Econometrics*, 20, 39-54.

[51] Wooldridge, J.M. (2005b), "Unobserved Heterogeneity and Estimation of Average Partial Effects," in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, D.W.K. Andrews and J.H. Stock (eds.), Cambridge: Cambridge University Press.

[52] Wooldridge, J.M. (2005c), Fixed-Effects and Related Estimators for Correlated Random-Coefficient and Treatment-Effect Panel Data Models, *The Review of Economics and Statistics*, 87, 385-390.