

IZA DP No. 4889

**Modeling Employment Dynamics with State
Dependence and Unobserved Heterogeneity**

Victoria Prowse

April 2010

Modeling Employment Dynamics with State Dependence and Unobserved Heterogeneity

Victoria Prowse

*University of Oxford
and IZA*

Discussion Paper No. 4889

April 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Modeling Employment Dynamics with State Dependence and Unobserved Heterogeneity*

This paper extends existing work on labor force participation dynamics by distinguishing between full-time and part-time employment and allowing unobserved heterogeneity in the effects of previous employment outcomes, children and education on employment dynamics. The results reveal significant autocorrelation in unobservables, and significant variation in the effects of children and education on labor supply preferences. Moreover, omission of random coefficients or autocorrelation can bias significantly estimates of policy effects. On average, policies temporarily incentivizing part-time and full-time employment are equally effective tools for reducing non-employment. However, non-employment among women with young children is more responsive to policies encouraging part-time rather than full-time work.

JEL Classification: C15, C25, J6, J22

Keywords: discrete labor supply, unobserved heterogeneity, state dependence, repeated multinomial choice

Corresponding author:

Victoria Prowse
University of Oxford
Department of Economics
Manor Road Building
Manor Road
Oxford, OX1 3UQ
United Kingdom
E-mail: victoria.prowse@economics.ox.ac.uk

* I express my thanks for useful comments on this paper and earlier drafts to the following people: Steve Bond; Richard Blundell, Martin Browning, Ian Crawford, Kathryn Graddy; Andrey Launov; Valérie Lechene; Robert Moffitt; Bent Nielsen; and Peter Haan. I have also benefited from discussions with participants in seminars at CAM, IZA, Oxford, Surrey, Sussex, and at EEA and Econometric Society conferences. Any remaining errors are my own.

1 Introduction

Labor supply behavior measured at the individual level displays a great deal of persistence.¹ Persistence is observed both in participation decisions and in the hours of work of those in employment. In other words, we observe persistence on the extensive margin and on the intensive margin. It is well established, see for example Heckman (1981a) and Heckman (1981c), that persistence in labor supply behavior can be generated by two different mechanisms. On one hand, individual characteristics may lead an individual to choose repeatedly the same employment state. Relevant characteristics consist of observables, such as educational qualifications and household structure variables, and unobservables including unobserved preferences and ability. Alternatively, persistence in labor supply behavior may arise from state dependencies whereby an individual's previous labor supply behavior has a causal effect on his or her current labor supply incentives. State dependencies may be generated by, for example, changes in preferences or constraints caused by previous working behavior. For the purpose of policy evaluation, it is critical to determine the relative contributions of state dependence and individual characteristics to the observed persistence in labor supply behavior. Indeed, if labor supply choices are driven entirely by observed or unobserved individual characteristics then the effect of a policy intervention, such as a wage subsidy or an in-work benefit, will cease the moment the policy is withdrawn. In contrast, if past labor market outcomes exert a causal effect on current labor supply behavior then the policy intervention will affect labor market outcomes beyond the duration of the policy.

There exist several studies of labor force participation dynamics. Notably, Heckman (1981a) studied the dynamics of women's labor force participation decisions, while controlling for persistent unobserved individual characteristics. The results showed that unobserved individual characteristics contribute significantly to the observed persistence in women's labor force participation behavior but causal effects, or state dependencies, were also found to be present. A number of other studies report similar results, see *inter alia*, Booth *et al.* (1999) and Heckman and Willis (1977). Keane (1993) provided the first model of labor force participation with autocorrelated unobservables, while Hyslop (1999) extended the literature further by allowing both autocorrelated unobservables and correlated random effects, operationalized by including non-contemporaneous measures of observed individuals characteristics, including measures of fertility at different points in the life-cycle. Keane and Sauer (–) in turn extend the work of Hyslop (1999) by including classification error in the dependent variable along with an alternative treatment of the initial conditions. The inclusion of classification error was found to change conclusions concerning the exogeneity of fertility in the labor supply equation. Specifically, when classification error in labor force participation was included fertility became endogenous in the labor supply equation.

This paper does not consider the implications of classification error or correlated random effects, but instead provides two alternative extensions of existing studies of labor supply dynamics. First, this paper analyzes the dynamics of individual labor supply in a multinomial choice framework, rather than the more common binary choice setting. While the generalization to a multinomial model introduces concerns pertaining to identification and furthers computational complexities, this extension provides additional insight as it allows a study of the nature of intertemporal dependencies associated with full-time and part-time employment. The results of this analysis are therefore informative about the

¹Based on a sample of American women, Francesconi (2002) reports that 87.6% of women who were non-participants last year are also non-participants this year. The corresponding figures for full-time and part-time work are 87.6% and 68.9% respectively. Booth *et al.* (1999) report similar figures for the United Kingdom.

relative long-run effectiveness of labor market policies facilitating full-time and part-time employment. This analysis would not be possible using a binary model of labor market participation, such as the reduced form approach of Heckman (1981a) or the structural approach proposed by Eckstein and Wolpin (1989).

The second generalization offered by this paper is to implement an econometric specification that permits more general structures of unobservables than in previous studies of labor supply dynamics. Specifically, in addition to autocorrelated and time invariant unobservables, as used by Hyslop (1999) in a binary setting, unobserved heterogeneity in the effects of previous employment outcomes, children and educational qualifications on employment dynamics is included. This second extension is valuable from an econometric perspective because it allows an exploration of the implications of different assumptions regarding the distribution of the unobservables for estimates of state dependencies and the long-run effects of policy interventions in the context of a dynamic multinomial labor supply model. Furthermore, determining the extent of any heterogeneity in the effects of demographic variables on women's labor supply behavior is intrinsically important as it is informative about the extent to which life-cycle labor supply behavior varies between women with identical observed characteristics. This is particularly true when the characteristic in question is children; our analysis allows us to determine whether the relatively high rates of non-employment observed among women with young children are due to a common, or systematic, effect of young children on labor supply or whether young children affect the labor supply behavior of some, but not all, women.

The central econometric framework is a dynamic mixed multinomial logit model including persistent unobservables which are assumed to follow particular distributions and to occur independently of observed individual characteristics, i.e., unobservables are assumed to be random effects.² Parameter estimates are obtained using Maximum Simulated Likelihood estimation. We do not attempt to estimate a fixed effects version of our model as bias-corrected fixed effects estimators applicable to this dynamic, non-stationary, non-linear problem are not readily available. Moreover, unlike in a fixed effects setting, within our random effects framework we can accommodate time-varying persistent unobservables in the form of autocorrelation and random coefficients on time-varying characteristics. These features of the model specification are found to be empirically important. Moreover, we provide evidence which supports the validity of the random effects assumptions.

In contrast to the binary case, for which identification is relatively straight forward, little is known about the appropriateness of assumptions concerning the distribution of the unobservables in our setting. Indeed, previous work on multinomial choice models with complex specifications of unobservables in other areas of economics has shown that identification concerns sometimes arise in this setting. Specifically, Keane (1992) showed that identification of the parameters describing the distribution of unobservables in cross-sectional multinomial choice models is often reliant on functional form restrictions. Harris and Keane (1998) explore this issue further and show that in a cross-sectional setting alternative specific covariates are required for robust identification.³

In the current setting employment state specific covariates are not available; for example there are no obvious variables that affect the utility from full-time employment but which do not enter the utilities associated with part-time employment and non-employment. Thus, given only cross-

²Correlations between the persistent unobservables and the previous employment outcomes are naturally present in this setting and are fully accounted for in the estimation. Note also that the inclusion of persistent unobservables breaks the Independence of Irrelevant Alternatives property which plagues the standard multinomial logit model.

³Indeed, the mixed multinomial logit model has been widely applied to problems in which there are alternative specific explanatory variables, for example in marketing, see Keane (1997).

sectional data, a multinomial labor supply model with a flexible specification of unobservables would be identified only by functional form restrictions. Similarly, in repeated multinomial choice models, including the dynamic multinomial labor supply model under consideration, some of the more obvious, fully parametric, specifications of unobservables generate models which are identified purely by the functional form of the distribution of unobserved individual characteristics. It is shown below that, given repeated observations of individuals' choices, there are specifications of unobservables for which identification is not reliant on functional form restrictions. The empirical analysis below is conducted using only specifications of unobservables for which identification is not driven by functional form restrictions.⁴ While some structures of unobserved heterogeneity are therefore ruled out, time invariant individual specific random effects, autocorrelated unobservables and random coefficients are permitted. Thus persistent unobservables may be time-varying with different distributions depending on the individual's observed characteristics and previous employment outcomes.

The empirical analysis is conducted using a fourteen year longitudinal sample taken from the British Household Panel Survey. The unbalanced panel sample comprises of married and cohabiting women and spans the years 1991-2005 inclusive. Three employment states are distinguished, namely full-time work, part-time work and non-employment. For the sample of women under consideration all three employment states are quantitatively important. Furthermore, there is a growing literature documenting the relatively poor status of part-time jobs in the United Kingdom: Connolly and Gregory (2008) and Manning and Petrongolo (2008) show that part-time jobs are typically poorly paid and are concentrated menial occupations. Within the context of this literature it seems important to establish whether part-time jobs are also associated with lower labor market attachment than full-time jobs.

We find significant autocorrelation in unobservables, and significant variation in the effects of children and education on labor supply preferences. Our analysis shows that, irrespective of the assumed distribution of the unobservables, significant positive own-state dependencies are present in both full-time and part-time work. Equivalently, temporary policies incentivizing women to move into either full-time or part-time work will affect women's employment behavior beyond the duration of the policy. This is in line with existing work on dynamic labor supply including Keane (1993) and Hyslop (1999). Additionally, and in contrast to existing dynamic binary models of labor force participation, our generalized model also allows us to exploit the multinomial structure to draw conclusions about the relative effectiveness of temporary policies incentivizing jobs with different hours of work. Considering the sample average, we find that policies incentivizing full-time employment are partly crowded out by a reduction in part-time work, while policies that encourage part-time work have a positive effect on the rate of full-time employment. These asymmetric cross-state effects mean that, although own-state dependence is higher for full-time employment than for part-time employment, policies temporarily incentivizing part-time and full-time employment are equally effective tools for reducing non-employment. However, we find that for specific subgroups, defined by observed individual characteristics, labor force participation is most effectively targeted by creating either full-time or part-time jobs. For example, for women with young children, policies incentivizing part-time employment lead to larger reductions in non-employment than policies incentivizing full-time employment.

Our results further show significant variation in preferences for full-time and part-time work, relative to non-employment, among women with young children. We explore the implications of this variation for employment dynamics following the birth of a child. We find that women with a high

⁴This restriction has the further advantage of generating models that are well behaved and means that parameters can be estimated without encountering numerical problems.

unobserved preference for full-time work when they have a young child are not more likely to be non-employed following the birth of a child. Thus, the high rates of non-work among women with children are due to changes in labor force participation behavior among a subset of women with children, specifically those women who have a low unobserved preference for full-time work in the event that they have a young child.

A comparison of the results across the different specifications of unobservables reveals that estimates of state dependencies are sensitive to the assumed distribution of the unobservables. As has been frequently found in studies of labor force participation and unemployment, state dependencies are overestimated if persistent unobservables are ignored. Less predictably, estimated state dependencies tend to increase as the distribution of the unobservables is generalized from a specification allowing time invariant random intercepts to more general specifications allowing autocorrelated unobservables and random coefficients. We conclude that estimating dynamic labor supply models and ignoring autocorrelation or variation in the effects of observed individual characteristics on labor supply behavior may bias significantly estimates of the long-term effectiveness of labor market policies. The biases induced by ignoring autocorrelation or variation in the effects of observed individual characteristic pertain predominantly to the long-run effects of policies that facilitate full-time, rather than part-time, work.

The next section outlines a model describing an individual's choice between full-time employment, part-time employment and non-employment. Section 3 introduces a dynamic mixed multinomial logit model of labor supply behavior. Section 4 provides an overview of the British Household Panel Survey, and summarizes the main features of the estimation sample. Section 5 contains the results, including comparisons of policy effects based on different assumptions concerning the distribution of the unobservables. Section 6 concludes. Appendices I and II are devoted to Monte Carlo simulations illustrating the performance of Maximum Simulated Likelihood estimators of the parameters of several dynamic mixed multinomial logit models.

2 A Dynamic Multi-state Labor Supply Model

In year t individual i chooses between full-time employment (f), part-time employment (p) and non-employment (n) so as to maximize her current payoff.^{5,6} The individual receives a payoff $V^j(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,j,t})$ if she chooses employment state j at time t for $j = f, p, n$. Payoffs are functions of the relevant elements of the individual's employment history, $\Omega_{i,t-1}$, individual characteristics observed by both the individual and the econometrician, denoted $X_{i,t}$ and henceforth referred to as explanatory variables, and individual characteristics that are known to the individual but which are unobserved to the econometrician, denoted $\varrho_{i,j,t}$ for $j = f, p, n$. The variables $\Omega_{i,t-1}$, $X_{i,t}$ and $\varrho_{i,j,t}$ may be vectors. Conditional on observed characteristics and the individual's employment history, optimizing behavior on the part

⁵Discrete choice labor supply models have three advantages over continuous or mixed discrete-continuous models of labor supply. First, a discrete opportunity set reflects that many individuals face a choice between a small number of wage-hours contracts, and consequently are unable to vary their hours of work continuously. Second, the grouping of reported hours into a small number of categories tends to reduce measurement error. Last, discrete approaches generate empirically tractable labor supply functions consistent with non-linearities or discontinuities in budget set generated by fixed costs of employment or the structure of the tax and benefit system (see van Soest, 1995).

⁶The term "payoff" in this context is taken to refer to the individual's utility associated with a particular employment state, taking into account any costs and benefits, as well as the income, associated with the employment state.

of the individual implies the following labor supply probabilities

$$P_{i,f,t}(\Omega_{i,t-1}, X_{i,t}) = P \left(\begin{array}{l} V^f(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,f,t}) \geq V^p(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,p,t}) \\ V^f(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,f,t}) \geq V^n(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,n,t}) \end{array} \middle| \Omega_{i,t-1}, X_{i,t} \right), \quad (1a)$$

$$P_{i,p,t}(\Omega_{i,t-1}, X_{i,t}) = P \left(\begin{array}{l} V^p(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,p,t}) > V^f(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,f,t}) \\ V^p(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,p,t}) \geq V^n(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,n,t}) \end{array} \middle| \Omega_{i,t-1}, X_{i,t} \right), \quad (1b)$$

$$P_{i,n,t}(\Omega_{i,t-1}, X_{i,t}) = P \left(\begin{array}{l} V^n(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,n,t}) > V^f(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,f,t}) \\ V^n(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,n,t}) > V^p(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,p,t}) \end{array} \middle| \Omega_{i,t-1}, X_{i,t} \right), \quad (1c)$$

where $P_{i,j,t}(\Omega_{i,t-1}, X_{i,t})$ is the probability of individual i choosing employment state j at time t and $P()$ denotes a probability.

The above formulation is sufficiently general so as to allow dependencies between an individual's past and current labor supply decisions due to habit formation in labor supply behavior (Bover, 1991; Kubin and Prinz, 2002; Woittiez and Kapteyn, 1998), wage based rewards for human capital accumulated via labor market experience (Altug and Miller, 1998; Eckstein and Wolpin, 1989; Imai and Keane, 2004; Wolpin, 1992) and job search costs (Heckman and MaCurdy, 1980; Hyslop, 1999). Habit formation gives rise to non-separabilities in the utility function, while accumulation of human capital and job search costs imply non-separabilities in the budget constraint. All three mechanisms create dependencies between past and current labor supply choices. Specifically, job search costs generate dependencies between labor supply choices in consecutive years, while habit formation and the accumulation of human capital have the potential to create dependencies in labor supply behavior spanning several years.

In all that follows reduced form, rather than structural, approaches to estimating the parameters of the above model are adopted. Reduced form approaches are adequate because, in this study, the quantities of interest are the nature of any intertemporal dependencies in labor supply behavior rather than the underlying structural parameters. Moreover, taking a reduced form perspective avoids having to solve the dynamic programming model implicit in the above, leading to computationally simpler models within which it is feasible to accommodate relatively general distributions of unobservables.

3 Estimation Strategy

Dynamic mixed multinomial logit models are obtained by adopting a specification for the payoff functions appearing in the above labor supply probabilities and then placing appropriate distributional assumptions on the unobserved individual characteristics. The specification of the payoffs may be interpreted as an approximation to the state specific value functions occurring in the underlying dynamic programming problem (see, for example, Eckstein and Wolpin, 1989; Francesconi, 2002). In this case, preference parameters are compounded with parameters appearing in the budget constraint. This section proceeds by discussing the specification of payoffs, the treatment of the initial conditions, and issues surrounding identification. Finally, the chosen empirical specification is presented together with the proposed Maximum Likelihood estimation method.

3.1 Specification of Payoffs

An examination of Equations (1a)-(1c) reveals that labor supply probabilities can be expressed in terms of the two indices $V^f(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,f,t}) - V^n(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,n,t})$ and $V^p(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,p,t}) -$

$V^n(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,n,t})$.⁷ The following specification is adopted

$$V^j(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,j,t}) - V^n(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,n,t}) = \Omega_{i,t-1}\gamma_j + X_{i,t}b_j + h_j(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,j,t}) \text{ for } j = f, p, \quad (2)$$

where b_j and γ_j for $j = f, p$ are suitably dimensioned vectors of unknown parameters and $h_j()$ for $j = f, p$ are functions describing the unobserved components of the individual's payoffs.⁸ The unobserved individual characteristics $\varrho_{i,j,t}$ for $j = f, p$ are assumed to occur independently of the individual's observed characteristics $X_{i,s}$ for all s and t , and thus are individual specific random effects. In Section 5.5 we present evidence supporting the robustness of our results to the random effects assumption. The coefficient vectors b_f and b_p measure the deterministic effect of the individual characteristics in $X_{i,t}$, such as education and household structure variables, on an individual's payoffs from, respectively, full-time work and part-time work relative to her payoff from non-employment. The explanatory variables $X_{i,t}$ do not include employment state specific variables, such as wages or incomes, because such quantities are unobserved for all employment states not chosen by the individual at time t .

The coefficient vectors γ_f and γ_p measure the deterministic effect of the individual's employment history on, respectively, her payoff from full-time employment relative to her payoff from non-employment and on her payoff from part-time employment relative to her payoff from non-employment. State dependencies are present if any elements of γ_f or γ_p are significantly different from zero. The econometric analysis is conducted using panel data where information about an individual's employment history is restricted to the duration of the individual's presence in the panel. Thus, prior to estimation, restrictions on the specification of $\Omega_{i,t-1}$ are required. In this study attention is restricted to the case where only the individual's labor market outcomes in the past two years affect her payoffs in the current year. Specifically $\Omega_{i,t-1} = [Y_{i,f,t-1}, Y_{i,p,t-1}, Y_{i,f,t-2}, Y_{i,p,t-2}]$, where $Y_{i,j,t}$ is an indicator taking the value one if individual i was in employment state j at time t and zero otherwise. Suppose labor market outcomes are observed in years $t = 1, \dots, T$. Equation (2) then holds for $t = 3, \dots, T$. This specification should not be overly restrictive as the strongest intertemporal dependencies in labor supply incentives are likely to occur over short time horizons. Further support for this specification of the relevant employment history is provided in Section 5.5.

3.2 The Initial Conditions Problem

An initial conditions problem arises when estimating this model. Given the dynamic structure of the model and the above described specification of the individual's employment history, $\Omega_{i,t-1}$, the individual's employment outcome in the year $t = 1$ depends on her employment outcomes in the years $t = 0$ and $t = -1$, which are unobserved to the econometrician. Likewise, the individual's employment outcome in the year $t = 2$ depends on her unobserved employment outcome in the year $t = 0$. Therefore, employment outcomes in the years $t = 1$ and $t = 2$, referred to as the initial conditions, cannot be modeled in the same way as subsequent employment outcomes. When estimating the parameters of the above model, the treatment of the initial conditions proposed by Wooldridge (2005)

⁷The third index $V^f(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,f,t}) - V^p(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,p,t})$ is redundant as it is equal to the difference between the other two indices.

⁸Without loss of generality, the unobserved characteristics $\varrho_{i,j,t}$ for $j = f, p$ are henceforth taken to represent unobserved characteristics that affect that the difference between the individual's payoff from employment state j and her payoff from non-employment.

is adopted. According to this approach, individual likelihood contributions are defined as the joint probability of an individual's observed employment outcomes at $t = 3, \dots, T$ conditional on explanatory variables and initial conditions.⁹

Implementation of the approach of Wooldridge (2005) requires the functions $h_j(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,j,t})$ for $j = f, p$ appearing in Equation (2) be modified to include the individual's employment outcomes in the years $t = 1$ and $t = 2$, denoted by IC_i . In all that follows, IC_i consists of five variables indicating if the woman worked full-time, worked part-time or was non-employed in both $t = 1$ and $t = 2$, worked both full-time and part-time in her first two years in the sample or worked full-time and was non-employed in her first two years in the sample. The modified functions describing the unobservables are denoted $\tilde{h}_j(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,j,t}, IC_i)$ for $j = f, p$. For $t = 3, \dots, T$, the unobserved individual characteristics $\varrho_{i,j,t}$ for $j = f, p$ are assumed to occur independently of IC_i and $\{X_{i,s}\}_{s=3}^T$.

The following definitions are required prior to deriving individual i 's contribution to the likelihood. Define $\varrho_{i,j}$ as $\varrho_{i,j,t}$ stacked over $t = 3, \dots, T$ for $j = f, p$ and let $G(\varrho_{i,f}, \varrho_{i,p})$ denote the distribution of $(\varrho_{i,f}, \varrho_{i,p})$. Define the one by three dimensional vectors $A_f = (1, 0, 0)$, $A_p = (0, 1, 0)$ and $A_n = (0, 0, 1)$ and let $A_{i,t} = A_j$ if individual i chose state j at time t for $j = f, p, n$. Additionally define the two by three dimensional matrix $B_{i,t}$ as follows

$$B_{i,t} = \begin{cases} (A'_p, A'_n)' & \text{if } Y_{i,f,t} = 1 \\ (A'_f, A'_n)' & \text{if } Y_{i,p,t} = 1 \\ (A'_f, A'_p)' & \text{if } Y_{i,n,t} = 1 \end{cases} \quad \text{for } t = 3, \dots, T. \quad (3)$$

Lastly, let $V_{i,t}$ denote $V^j(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,j,t})$ stacked vertically over $j = f, p, n$. Individual i 's contribution to the likelihood takes the following form

$$\mathcal{L}_i = \int_{\varrho_{i,f}, \varrho_{i,p}} \mathbf{I}(A_{i,3}V_{i,3} \geq \max\{B_{i,3}V_{i,3}\} \cap A_{i,4}V_{i,4} \geq \max\{B_{i,4}V_{i,4}\} \cap \dots \\ \dots \cap A_{i,T}V_{i,T} \geq \max\{B_{i,T}V_{i,T}\}) dG(\varrho_{i,f}, \varrho_{i,p}), \quad (4a)$$

$$= \int_{\varrho_{i,f}, \varrho_{i,p}} \prod_{t=3}^T \mathbf{I}(A_{i,t}V_{i,t} \geq \max\{B_{i,t}V_{i,t}\}) dG(\varrho_{i,f}, \varrho_{i,p}), \quad (4b)$$

where $\mathbf{I}(\cdot)$ is an indicator of whether the statement in parenthesis is true and the integral in the above equations is over the entire support of $(\varrho_{i,f}, \varrho_{i,p})$.

3.3 Identification and the Distribution of Unobservables

Identification of multinomial choice models requires well-known scale and location normalizations (see Ben-Akiva and Lerman, 1985; Bunch, 1991; Keane, 1992). By specifying the problem in terms of

⁹In contrast, Heckman (1981b) suggests modeling the joint distribution of an individual's employment outcomes over the entire sample period. This requires a specification of the joint distribution of the individual's employment outcomes in the years $t = 1$ and $t = 2$ and a specification of the distribution of unobserved heterogeneity. In general, the exact distribution of the initial conditional is impossible to derive. (In the absence of non-stationary explanatory variables it is possible to derive the equilibrium distribution of the process which can then be used as the distribution of the initial conditions. However the presence of a number of non-stationary explanatory variables, including age and the time dummies, in the current application renders using an exact specification of the initial conditions impossible.) Instead, Heckman (1981b) suggests approximating the distribution of the initial observations conditional exogenous explanatory variables. The Wooldridge approach does not require a model of the joint distribution of the individual's employment outcomes in the years $t = 1$ and $t = 2$ or a specification of the joint distribution of unobservables occurring at $t = 1$ and $t = 2$ and the unobservables occurring in future years. This has the computational advantage, relative to the Heckman (1981b) approach, of reducing the number of unknown parameters.

differences in payoffs the required location normalizations have been imposed. However, depending on the distribution of the unobservables, an identifying scale normalization might be required as multiplying all payoffs, including the components of payoffs attributable to unobservables, by a positive constant does not change the individual's behavior. In all that follows $G(\varrho_{i,f}, \varrho_{i,p})$ is taken to be the distribution of the unobserved individual characteristics after the minimum normalizations required to ensure identification have been imposed.

The functions $\tilde{h}_j(\cdot)$ for $j = f, p$ and the distribution function $G(\varrho_{i,f}, \varrho_{i,p})$ dictate both the structure of persistence in unobservables and the joint distribution of unobservables occurring in a particular year. As discussed above, allowing persistence in unobservables is necessary for determining correctly the nature of state dependencies in labor supply behavior. Meanwhile, Hausman and Wise (1978) show that estimates of marginal effects, substitution patterns and elasticities are not robust to the assumed intratemporal distribution of the unobservables. It is therefore desirable to work with a flexible distribution of unobservables. However, even after imposing all necessary identifying scale and location normalizations, care is required when working with flexible forms of the above described structure of the unobservables. Indeed, unlike in the binary case, in the current multinomial labor supply model some of the more obvious, fully parametric, specifications of $\tilde{h}_j(\cdot)$ for $j = f, p$ and distribution of the unobserved individual characteristics, $G(\varrho_{i,f}, \varrho_{i,p})$, generate models which are identified purely by the functional form of the distribution of the unobserved individual characteristics and therefore are nonparametrically unidentified.

The possibility of an absence of nonparametric identification can be understood by manipulating the individual likelihood contributions displayed above in Equation (4b). For given choices of $\tilde{h}_j(\cdot)$ for $j = f, p$, the unobserved individual characteristics, $(\varrho_{i,f}, \varrho_{i,p})$, can be decomposed into $(\varrho_{i,f}^a, \varrho_{i,p}^a)$ and $(\varrho_{i,f}^t, \varrho_{i,p}^t)$ for $t = 3, \dots, T$ where $(\varrho_{i,f}^t, \varrho_{i,p}^t)$ for $t = 3, \dots, T$ is the collection of all unobserved individual characteristics that affect payoffs at time t and occur independently of the unobserved individual characteristics affecting payoffs in all other years, while the unobserved individual characteristics $(\varrho_{i,f}^a, \varrho_{i,p}^a)$ affect payoffs in at least two years for some individuals. Let $G_a(\varrho_{i,f}^a, \varrho_{i,p}^a)$ denote the distribution of $(\varrho_{i,f}^a, \varrho_{i,p}^a)$ and let $G_t(\varrho_{i,f}^t, \varrho_{i,p}^t)$ denote the distribution of $(\varrho_{i,f}^t, \varrho_{i,p}^t)$ for $t = 3, \dots, T$. It follows that

$$\mathcal{L}_i = \int_{\varrho_{i,f}^a, \varrho_{i,p}^a} \prod_{t=3}^T \left(\int_{\varrho_{i,f}^t, \varrho_{i,p}^t} \mathbf{I}(A_{i,t}V_{i,t} \geq \max\{B_{i,t}V_{i,t}\} | \varrho_{i,f}^a, \varrho_{i,p}^a) dG_t(\varrho_{i,f}^t, \varrho_{i,p}^t) \right) dG_a(\varrho_{i,f}^a, \varrho_{i,p}^a). \quad (5)$$

In a nonparametric setting each of the distribution functions $G_t(\varrho_{i,f}^t, \varrho_{i,p}^t)$ for $t = 3, \dots, T$ can be varied independently of $G_a(\varrho_{i,f}^a, \varrho_{i,p}^a)$. In other words there are unknown parameters that can affect the conditional probability that $A_{i,t}V_{i,t} \geq \max\{B_{i,t}V_{i,t}\}$ but which do not enter $G_a(\varrho_{i,f}^a, \varrho_{i,p}^a)$. In this case the bivariate distribution functions, which appear in parenthesis in Equation (5), are nonparametrically unidentified as the same variables affect both the probability that $A_{i,t}V_{i,t} > B_{i,t,1}V_{i,t}$ and the probability that $A_{i,t}V_{i,t} > B_{i,t,2}V_{i,t}$, where $B_{i,t,k}$ for $k = 1, 2$ denotes the k^{th} row of $B_{i,t}$.^{10,11} Consequently, parametric distributions of unobservables that do not impose any structure, beyond their functional

¹⁰Equation (5) further shows that the lack of nonparametric identification is related to the absence of employment state specific explanatory variables. If, in contrast to the specification given in Equation (2), the payoffs included, for example, employment state specific incomes or other characteristics of the employment states, parameters would, under appropriate regularity conditions, be nonparametrically identified (see Harris and Keane, 1998; Keane, 1992).

¹¹The lack of nonparametric identification is specific to discrete choice models with three or more alternatives; in the corresponding binary model nonparametric identification is less problematic as choice probabilities depend on a single index.

form, on $G(\varrho_{i,f}, \varrho_{i,p})$ generate dynamic mixed multinomial choice models that are nonparametrically unidentified.

Keane (1992) presents Monte Carlo evidence illustrating the very poor empirical performance of cross-sectional multinomial choice models which are identified only by the functional form of the distribution of the unobservables. Specifically, Keane (1992) showed that the cross-sectional multinomial probit model with heteroscedastic and correlated unobservables suffers from possible non-convergence, highly biased parameter estimates and a close to singular Hessian which translates into huge standard errors. These problems arise because it is possible to adjust the intercepts and coefficients on explanatory variables so as to offset almost completely the effect on choice probabilities of changes in the parameters describing the distribution of the unobservables, namely variance and correlation parameters. Consequently, the criterion function, a likelihood in Keane's study, is almost completely flat over a large subset of the parameter space. Further Monte Carlo simulations, discussed below in Appendix I, illustrate the severity of the numerical problems afflicting a dynamic mixed multinomial logit model in which identification is reliant on the functional form of the distribution of the unobservables.

Given these problems, in this study attention is restricted to model specifications in which the distribution of unobservables is nonparametrically identified. This entails three requirements. Firstly, as illustrated above, the structure of the unobservables must be restricted such that $G_t(\varrho_{i,f}^t, \varrho_{i,p}^t)$ for $t = 3, \dots, T$ cannot be varied independently of $G_a(\varrho_{i,f}^a, \varrho_{i,p}^a)$. Secondly, the probability of a transition between each pair of employment states must be strictly positive for each value of $(X_{i,t}, IC_i)$. Indeed, with no change in employment status between $t = 3$ and $t = T$ repeated observations provide no additional identifying information concerning the distribution of the unobservables. Finally, $X_{i,t}$ for $t = 3, \dots, T$ must contain one variable that is continuously distributed over individuals and b_f and b_p must be such that both coefficients on this variable are non-zero. A continuously distributed variable is required because conditional on $(\{X_{i,s}\}_{s=3}^T, IC_i)$ identification is limited by the number of different combinations of transitions between employment states. Variation in $X_{i,t}$ across individuals creates variation in the probability associated with each possible combination of transitions between employment states. Additional Monte Carlo simulations reported in Appendix I show that reliable estimates of the parameters of a dynamic mixed multinomial logit model are obtained when the model specification satisfies these requirements.

3.4 Empirical Specification

The adopted specification of the unobservables allows time invariant individual specific random effects, autocorrelated unobservables and time invariant random coefficients. Mathematically, $\tilde{h}_j(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,j,t}, IC_i)$ for $j = f, p$ take the following form

$$\begin{aligned} \tilde{h}_j(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,j,t}, IC_i) = \\ \Omega_{i,t-1}\omega_{i,j} + W_{i,t}\pi_{i,j} + IC_i(\vartheta_j + \psi_{i,j}) + \zeta_{i,j,t} + \nu_{i,j} + \xi_{i,f,t} \quad \text{for } j = f, p; t = 3, \dots, T, \end{aligned} \quad (6a)$$

where $W_{i,t}$ denotes selected elements of $X_{i,t}$, and $\zeta_{i,f,t}$ and $\zeta_{i,p,t}$ follow first order autoregressive processes

$$\zeta_{i,j,t} = \rho_j \zeta_{i,j,t-1} + \varsigma_{i,j,t} \quad \text{for } j = f, p; t = 3, \dots, T. \quad (6b)$$

In the above $(\{\xi_{i,j,t}, \varsigma_{i,j,t}\}_{t=3}^T, \omega_{i,j}, \pi_{i,j}, \psi_{i,j}, \nu_{i,j}, \zeta_{i,j,2})$ for $j = f, p$ are unobserved individual characteristics and ϑ_j for $j = f, p$ and suitably dimensioned vectors of unknown parameters.

The pairs $(\xi_{i,f,t}, \xi_{i,p,t})$ for $t = 3, \dots, T$ are assumed to occur independently over time and thus represent time-varying shocks to individuals' payoffs. In accordance with the above described requirements for nonparametric identification, the pairs $(\xi_{i,f,t}, \xi_{i,p,t})$ for $t = 3, \dots, T$ are assumed to have distributions that do not contain unknown parameters. In what follows, $\xi_{i,f,t}$ and $\xi_{i,p,t}$ are defined respectively as $\epsilon_{i,f,t} - \epsilon_{i,n,t}$ and $\epsilon_{i,p,t} - \epsilon_{i,n,t}$ for $t = 3, \dots, T$ where $\epsilon_{i,j,t}$ for $j = f, p, n$ and are mutually independent and independent of $\{X_{i,s}\}_{s=3}^T$ and IC_i . Furthermore $\epsilon_{i,j,t}$ for $j = f, p, n$ are assumed to have type I extreme value distributions.¹²

The remaining unobserved individual characteristics, denoted $(\bar{\varrho}_{i,f,t}, \bar{\varrho}_{i,p,t})$, therefore consist of four distinct components: (i) $\omega_{i,f}$ and $\omega_{i,p}$ are the random components of the coefficients on the individual's employment history; (ii) $\pi_{i,f}$ and $\pi_{i,p}$ represent the random components of the coefficients on the explanatory variables, $W_{i,t}$; (iii) $\psi_{i,f}$ and $\psi_{i,p}$ are the random components of the coefficients on the initial conditions; (iv) $\zeta_{i,f,t}$ and $\zeta_{i,p,t}$ represent the autocorrelated random components of the employment state specific intercepts while $\nu_{i,f}$ and $\nu_{i,p}$ are time invariant random components of the employment state specific intercepts. The random coefficients allow different amounts of unobserved variation in payoffs, depending on the individual's observed characteristics, initial employment state and previous working behavior, while the random components of the employment state specific intercepts capture persistent unobserved differences between individuals that occur irrespective of the observed characteristic and previous employment behavior.

Let $\bar{h}_{i,j,t} = \bar{h}_j(\Omega_{i,t-1}, X_{i,t}, \bar{\varrho}_{i,j,t}, IC_i)$ denote $\tilde{h}_j(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,j,t}, IC_i)$ net of $\xi_{i,j,t}$. It follows that, conditional on $(\Omega_{i,t-1}, X_{i,t}, \bar{\varrho}_{i,f,t}, \bar{\varrho}_{i,p,t}, IC_i)$, the individual's choice probabilities are independent over time and take the familiar multinomial logit form

$$P_{i,j,t}(\Omega_{i,t-1}, X_{i,t}, \bar{\varrho}_{i,f,t}, \bar{\varrho}_{i,p,t}, IC_i) = \frac{\exp(\Omega_{i,t-1}\gamma_j + X_{i,t}b_j + \bar{h}_j(\Omega_{i,t-1}, X_{i,t}, \bar{\varrho}_{i,j,t}, IC_i))}{1 + \sum_{k=f,p} \exp(\Omega_{i,t-1}\gamma_k + X_{i,t}b_k + \bar{h}_k(\Omega_{i,t-1}, X_{i,t}, \bar{\varrho}_{i,k,t}, IC_i))}$$

for $j = f, p; t = 3, \dots, T$, (7a)

$$P_{i,n,t}(\Omega_{i,t-1}, X_{i,t}, \bar{\varrho}_{i,f,t}, \bar{\varrho}_{i,p,t}, IC_i) = \frac{1}{1 + \sum_{k=f,p} \exp(\Omega_{i,t-1}\gamma_k + X_{i,t}b_k + \bar{h}_k(\Omega_{i,t-1}, X_{i,t}, \bar{\varrho}_{i,k,t}, IC_i))}$$

for $t = 3, \dots, T$. (7b)

Let $(\bar{\varrho}_{i,f}, \bar{\varrho}_{i,p})$ denote $(\bar{\varrho}_{i,f,t}, \bar{\varrho}_{i,p,t})$ stacked over $t = 3, \dots, T$ and let $F(\bar{\varrho}_{i,f}, \bar{\varrho}_{i,p})$ denote the distribution of $(\bar{\varrho}_{i,f}, \bar{\varrho}_{i,p})$. The above expression for individual i 's likelihood contribution can be rewritten as follows

$$\mathcal{L}_i = \int \prod_{\bar{\varrho}_{i,f}, \bar{\varrho}_{i,p}} \prod_{t=3}^T \prod_{j=f,p,n} P_{i,j,t}(\Omega_{i,t-1}, X_{i,t}, \bar{\varrho}_{i,f,t}, \bar{\varrho}_{i,p,t}, IC_i)^{Y_{i,j,t}} dF(\bar{\varrho}_{i,f}, \bar{\varrho}_{i,p}). \quad (8)$$

As described above in Section 3.3 nonparametric identification of $F(\bar{\varrho}_{i,f}, \bar{\varrho}_{i,p})$ requires a variable in $X_{i,t}$ that is continuously distributed across individuals. In this application non-labor income fulfills this role. It follows that if $\bar{h}_j(\Omega_{i,t-1}, X_{i,t}, \bar{\varrho}_{i,j,t}, IC_i)$ allows only time invariant individual specific random effects, and therefore excludes random coefficients and autocorrelation, then $F(\bar{\varrho}_{i,f}, \bar{\varrho}_{i,p})$ is nonparametrically identified if $T \geq 4$. If $\bar{h}_j(\Omega_{i,t-1}, X_{i,t}, \bar{\varrho}_{i,j,t}, IC_i)$ additionally allows autocorrelation in the employment state specific intercepts then $T \geq 5$ is required in order for $F(\bar{\varrho}_{i,f}, \bar{\varrho}_{i,p})$ to be nonparametrically identified. In this case, an extra year of observations is required in order to separate the autocorrelated and time invariant unobservables. The introduction of random coefficients does not require a longer panel for nonparametric identification provided that random coefficients on the time dummies and any other variables that, for all individuals, are non-zero in a maximum of one

¹²Assuming a type I extreme value distribution implies $P(\epsilon_{i,j,t} \leq q) = \exp(-\exp(-q))$.

year between $t = 3$ and T are excluded. This ensures that all random coefficients with a distribution containing unknown parameters affect payoffs in at least two years between $t = 3$ and $t = T$.

Six different specifications of the unobservables are considered. The first specification, presented primarily for comparative purposes, consists of a standard multinomial logit model. The second and third specifications allow the employment state specific intercepts to include time invariant individual effects. In the second specification these are jointly normally distributed with mean zero and an unrestricted covariance matrix while the third specification assumes a distribution generated by a mixture of two normal distributions with different means and covariance matrices. The fourth specification allows the employment state specific intercepts to contain time invariant components, assumed to be jointly normally distributed, and autocorrelated components, where the autocorrelation processes are jointly normal and the initial conditions of the autocorrelation processes ensure stationary. The fifth specification allows time invariant individual effects, assumed to be jointly normally distributed, and random coefficients on the individual's previous employment outcomes, the initial conditions and selected explanatory variables. The two random coefficients on a particular variable, for example the k^{th} elements of $\pi_{i,f}$ and $\pi_{i,p}$, are assumed to be jointly normally distributed with zero mean and an unrestricted covariance matrix, and all pairs of random coefficients are mutually independent and independent of the random components of the employment state specific intercepts.¹³ The sixth specification is the most general specification under consideration and additionally allows autocorrelation, as previously described, in the employment state specific intercepts.

3.5 Estimation Methodology and Performance

Given a sample of N individuals and assuming independence over individuals, the likelihood function is the product of the individual likelihood contributions for the sample members, given above in Equation (8). However, due to the integration with respect to the unobserved individual characteristics, analytic expressions for the individual likelihood contributions are unavailable for all but the simplest specifications of unobserved heterogeneity. Let Υ_i denote the $2(T - 2)$ by $2(T - 2)$ covariance matrix of the unobservables $\bar{h}_j(\Omega_{i,t-1}, X_{i,t}, \bar{\varrho}_{i,j,t}, IC_i)$ stacked over $j = f, p$ and then over $t = 3, \dots, T$. The dimension of the integral occurring in the individual's likelihood contribution is equal to the rank of Υ_i , which in turn depends on the assumed distribution of the unobservables. Specifications in which unobservables take the form of time invariant random intercepts require integration over two dimensions while each pair of random coefficients adds two to the dimension of the integral, up to a maximum of $2(T - 2)$. Specifications that include autocorrelation involve $2(T - 2)$ dimensional integrals. For two dimensional problems fast and accurate quadrature methods are available to evaluate the individual likelihood contributions (Geweke, 1996, provides a survey). However numerical methods are unable to evaluate the likelihood contributions with sufficient speed and accuracy to be effective in problems where the dimension of integration is greater than two (see Bhat, 2001; Hajivassiliou and Rudd, 1994). Consequently, numerical methods to evaluate the likelihood contributions are unavailable when unobservables feature random coefficients on several variables or when T is moderately large and unobservables are autocorrelated.

For models where an analytic expression for the likelihood is unavailable we use simulation techniques to evaluate the likelihood contributions. Simulation methods replace the intractable integral in the likelihood function by a sum over likelihood functions evaluated at different draws from the distribution of unobserved heterogeneity. Let $(\bar{\varrho}_{i,f}^r, \bar{\varrho}_{i,p}^r)$ denote the r^{th} draw from the distribution

¹³Allowing correlations between all pairs of random coefficients leads to a prohibitively large number of parameters.

$F(\bar{\varrho}_{i,f}, \bar{\varrho}_{i,p})$ for individual i . Individual i 's likelihood contribution is simulated as follows

$$\mathcal{L}_i^s = \frac{1}{R} \sum_{r=1}^R \prod_{t=3}^T \prod_{j=f,p,n} P_{i,j,t}(\Omega_{i,t-1}, X_{i,t}, \bar{\varrho}_{i,f,t}^r, \bar{\varrho}_{i,p,t}^r, IC_i)^{Y_{i,j,t}}. \quad (9)$$

Continuing to assume independence over individuals, the simulated likelihood is the product of the simulated individual likelihood contributions for the sample members. Maximum Simulated Likelihood estimates are obtained by maximizing the log simulated likelihood function. By the strong law of large numbers the Maximum Simulated Likelihood estimates converge almost surely to the true parameters as $R \rightarrow \infty$ and $N \rightarrow \infty$. Moreover, if R increases at a fast enough rate relative to N , Maximum Simulated Likelihood estimation is asymptotically equivalent to Maximum Likelihood estimation. In particular, with pseudo random draws, $\sqrt{N}/R \rightarrow 0$ as $N \rightarrow \infty$ is required (Hajivassiliou and Rudd, 1994).

In this application, the likelihood is simulated using antithetic variates rather than pseudo random draws. Antithetic variates are a variance reduction technique which reduces simulation noise by using draws from the distribution of the unobservables with more even coverage than pseudo random draws. R antithetic draws are obtained by taking $R/2$ pseudo random draws from the distribution of $(\bar{\varrho}_{i,f}, \bar{\varrho}_{i,p})$, denoted $\{\bar{\varrho}_{i,f}^r, \bar{\varrho}_{i,p}^r\}_{r=1}^{R/2}$. Assuming $F(\bar{\varrho}_{i,f}, \bar{\varrho}_{i,p})$ is symmetric around zero, the remaining $R/2$ draws are given by $\{-\bar{\varrho}_{i,f}^r, -\bar{\varrho}_{i,p}^r\}_{r=1}^{R/2}$. Hajivassiliou (1999) presents Monte Carlo evidence which shows that the use of antithetic variates in Maximum Simulated Likelihood problems approximately halves the number of draws required to obtain a given level of accuracy.

Appendix II presents Monte Carlo evidence demonstrating the empirical properties of the Maximum Simulated Likelihood estimator of the parameters of the two most complex specifications of unobservables under consideration. In summary, for a specification in which unobservables include random coefficients but exclude autocorrelated unobservables evaluation of the likelihood using 500 antithetic draws yields parameter estimates with tolerably small amounts of bias. A specification including autocorrelation displays a moderate amount of simulation bias when 500 or 2000 antithetic draws are used, but biases are relatively small when estimation uses 5000 antithetic draws. Therefore all of the empirical analysis conducted in this paper uses 5000 antithetic draws.

4 Data and Sample

The data source used for the empirical analysis is the British Household Panel Survey (BHPS). The BHPS is an ongoing annual panel survey that started in 1991 with a nationally representative sample of approximately 5,500 households in Great Britain. The sample used for analysis is an unbalanced panel covering the fourteen years 1991-2004. The last year of data for each individual is used purely to construct variables measuring fertility expectations and therefore a maximum of thirteen observations of a woman's employment behavior are available. Attention is restricted to married or cohabiting, non-retired women aged between 18 and 65 years and hence single mothers and single adult households are excluded from the sample. Women enter that sample in the first year in which they responded to the BHPS and satisfied the sample criteria. A woman remains in the sample unless she failed to respond to the BHPS, ceased to be married or cohabiting, retired or reached age 65 years. Furthermore, only women who provided at least four consecutive years of data are included in the sample. The final sample consists of 4,663 different women.

Table 1 shows number of women joining the sample in each year and the distribution of durations in the sample for each cohort of entrants. Fewer than half of the sampled women entered at the start

of the BHPS in 1991; entry was observed every year with substantial additional numbers of women entering in 1997, 1999 and 2001, when additional samples were added to the BHPS. While there is a high level of attrition, a large number of women remained in the sample for five or more years and over eight hundred women were present in all fourteen years.¹⁴

First Year in Sample	Number of Years in the Sample											Total
	4	5	6	7	8	9	10	11	12	13	14	
1991	202	126	109	106	101	90	107	87	92	67	808	1895
1992	14	18	11	10	8	12	6	7	15	54	-	155
1993	14	11	7	13	3	3	6	5	64	-	-	126
1994	17	11	6	8	5	3	6	48	-	-	-	104
1995	9	10	6	9	3	4	60	-	-	-	-	101
1996	10	9	14	6	5	59	-	-	-	-	-	103
1997	37	227	13	8	56	-	-	-	-	-	-	341
1998	34	7	6	72	-	-	-	-	-	-	-	119
1999	94	75	675	-	-	-	-	-	-	-	-	844
2000	42	160	-	-	-	-	-	-	-	-	-	202
2001	673	-	-	-	-	-	-	-	-	-	-	673
Any	1146	654	847	232	181	171	185	147	171	121	808	4663

Table 1: The number of women entering the sample in each year 1991-2001 and the distribution of durations in the sample for each cohort of entrants.

The measure of employment status is based on reported usual weekly hours of work. Figure 1(a) shows the density of the observed usual hours of work of the sampled women in employment, that is those with strictly positive usual hours of work. There are pronounced peaks at around 20 and 38 hours of work per week representing the hours of work frequently associated with, respectively, full-time and part-time work. For the purpose of the empirical analysis, and in accordance with the conventional British definitions of full-time and part-time work, women reporting usual weekly hours of work of between zero and 30 hours are classified as part-time employed, and women reporting usual weekly hours of work of over 30 hours are classified as full-time employed. Non-employment corresponds to zero usual weekly hours of work.¹⁵ Classification error in employment status should be minimal as observations of usual hours of work refer to usual working hours at the exact time of the annual survey, rather than being a retrospective report of usual working hours at some previous date. The top panel of Table 2 shows the percentage of women observed in each employment state in each of the years 1991-2003. On average, approximately a quarter of women were non-employed, 30% were working part-time and 45% were working full-time. There were no pronounced changes over the sample period in the proportions of women in each employment state.

Figures 1(b)- 1(d) illustrate the high level of persistence in women's employment outcomes. Around 85% of women who were working full-time one year previously are in full-time employment this year. Similarly, approximately 80% of women who were working part-time or who were non-employed one year previously are in the same employment state this year. There is also evidence of persistence over a longer time horizon. For example, around 55% of women who were working full-time 12 years previously are currently in full-time work. The corresponding figures for part-time work and non-employment are 50% and 39% respectively.

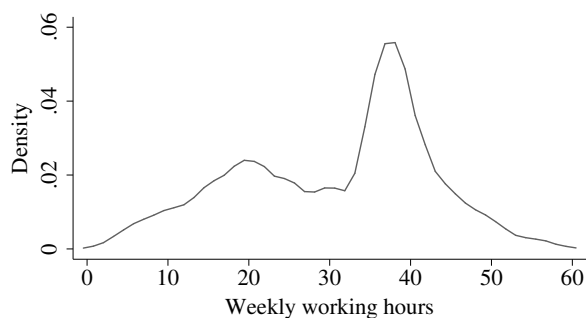
¹⁴Due to attrition the women in this sample will not be representative of the corresponding population. However this sample can be used to estimate parameters of interest provided that, conditional on observed characteristics, attrition is unrelated to employment status or, in other words, if there is no selectivity problem.

¹⁵Manning and Petrongolo (2008) discuss the relative merits of various definitions of part-time work.

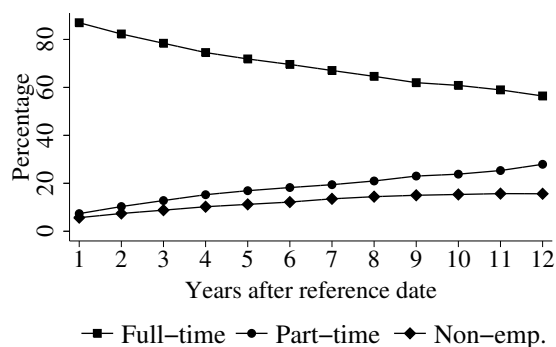
Variable	Year											
	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2003
Full-time Employment (%)	43.27	42.00	42.46	43.55	44.38	45.43	45.24	45.82	46.00	45.68	46.16	44.27
Part-time Employment (%)	29.71	29.37	30.19	29.60	29.23	29.84	29.72	28.70	29.84	29.16	28.44	31.00
Non-emp. Employment (%)	27.02	28.63	27.34	26.85	26.39	24.74	25.03	25.48	24.16	25.16	25.40	24.73
DEGREE	0.11	0.12	0.12	0.12	0.12	0.13	0.13	0.13	0.15	0.15	0.17	0.18
HIGH QUAL.	0.13	0.14	0.15	0.17	0.17	0.18	0.18	0.20	0.21	0.23	0.23	0.26
MEDIUM QUAL.	0.09	0.09	0.10	0.10	0.11	0.12	0.12	0.12	0.13	0.13	0.13	0.12
LOW QUAL.	0.28	0.27	0.27	0.27	0.26	0.25	0.25	0.24	0.24	0.22	0.22	0.20
INCOME	12.62	12.11	11.89	11.93	11.89	12.08	11.72	11.99	12.42	12.35	12.48	13.56
AGE	37.80	38.53	38.96	39.08	39.21	39.20	39.25	39.44	39.65	39.84	40.17	41.82
PRE SCH. CH.	0.28	0.26	0.25	0.24	0.22	0.22	0.22	0.22	0.23	0.24	0.25	0.24
YOUNGEST CH. 0-1 YEAR	0.07	0.05	0.06	0.06	0.06	0.05	0.06	0.05	0.06	0.07	0.07	0.05
YOUNGEST CH. 1-2 YEARS	0.11	0.11	0.10	0.09	0.10	0.10	0.10	0.10	0.09	0.09	0.10	0.10
YOUNGEST CH. 3-4 YEARS	0.06	0.08	0.08	0.08	0.07	0.06	0.08	0.07	0.07	0.07	0.07	0.08
YOUNGEST CH. 5-6 YEARS	0.05	0.05	0.05	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.05	0.06
YOUNGEST CH. 7-11 YEARS	0.12	0.12	0.11	0.12	0.12	0.12	0.12	0.12	0.13	0.14	0.14	0.13
YOUNGEST CH. 12-15 YEARS	0.10	0.10	0.09	0.10	0.09	0.08	0.09	0.09	0.09	0.09	0.10	0.11
EXP. CH. IN 0-3 MONTHS	0.02	0.01	0.02	0.02	0.01	0.01	0.01	0.02	0.02	0.01	0.01	0.01
EXP. CH. IN 4-6 MONTHS	0.01	0.02	0.01	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01

Notes: DEGREE, HIGH QUAL., MEDIUM QUAL. and LOW QUAL. are indicators of a woman's highest qualification being, respectively, a university degree, a non-degree level higher qualification, A levels or equivalent qualifications (A Levels are usually taken at age 18 years) or O levels or grade A-C GCSEs or equivalent qualifications (O levels were replaced by GCSEs, which are usually taken at age 16 years, in 1988). INCOME refers to a woman's annual net non-labor income including her partner's income expressed in pounds sterling and deflated to year 1991 prices using the Retail Price Index. AGE is the woman's age in years. PRE SCH. CH. is the number of children aged under five years in the household. YOUNGEST CH. $a - b$ YEARS is an indicator of the youngest child in the household being aged between a and b years. EXP. CH. IN 0-3 MONTHS and EXP. CH. IN 4-6 MONTHS are indicators of the woman expecting a child in the next three months and in 4-6 months time.

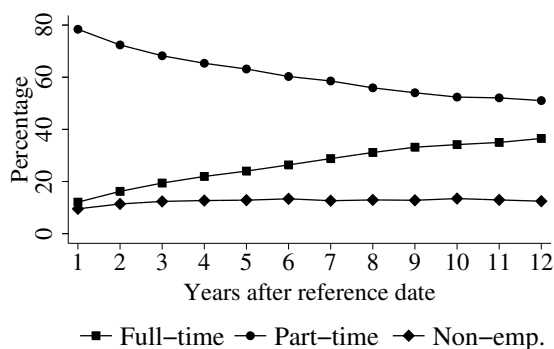
Table 2: Percentage of women in each employment state and sample means of the explanatory variables.



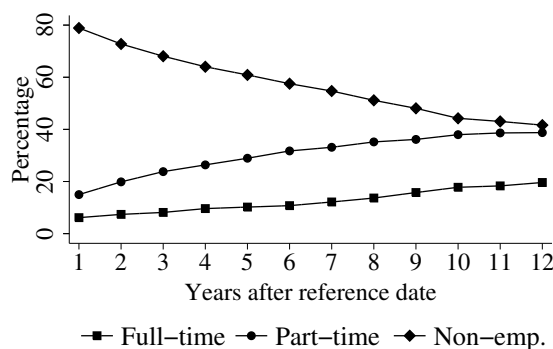
(a) Density of observed usual weekly working hours for women in employment.



(b) Labor market outcomes of women who were working full-time t years previously.



(c) Labor market outcomes of women who were working part-time t years previously.



(d) Labor market outcomes of women who were non-employed t years previously.

Notes: The sample used to construct Figure 1(a) has been truncated at 60 hours per week which excludes 0.5% of the observations.

Figure 1: Density of observed hours and observed persistence in employment outcomes.

The explanatory variables used in the empirical analysis are the conventional variables used in studies of women's labor supply behavior: education; age; child related variables; and non-labor income. Additionally, measures of short-term fertility expectations are constructed in order to capture any adjustments in women's labor supply behavior shortly before the arrival of a child. The bottom panel of Table 2 reports the sample means of the explanatory variables and further details concerning the definition of these variables are given in the accompanying notes.

5 Results

The dynamic mixed multinomial logit model is estimated with six different specifications of unobserved individual characteristics, as described above in Section 3.4. The parameter estimates and average marginal effects obtained from Specification VI are discussed in Section 5.1. In sections 5.2 and 5.3 we explore respectively the long-term labor supply effects of job creation policies and the extent of any heterogeneity in labor supply dynamics following the birth of a child, again based on Specification VI. In Section 5.4 we investigate the importance of allowing autocorrelation and random coefficients by making comparisons with the results obtained from Specifications I-V, which impose more restrictive distributions of unobservables.

5.1 Parameter Estimates and Average Marginal Effects

Specification VI is the most general specification under consideration. This specification allows random intercepts with both time invariant and autocorrelated components, and time invariant random coefficients on the indicators of having a degree and the woman's youngest child being aged under one year. Time invariant random coefficients on previous employment outcomes and the initial conditions are also included. Experimentation with various specifications of the random coefficients revealed that there are no random coefficients with significant amounts of variation on any other explanatory variables.

The last two columns of Table 3 show the deterministic components of the coefficients on the previous employment outcomes and explanatory variables. The coefficient estimates are as expected and are not discussed. Instead we focus on Table 5 which shows how the coefficients translate into average marginal effects. The results for Specification VI show that any increase in qualifications from no qualifications significantly increases the probability full-time work, but has no significant effect the probability of working part-time. There is a small but significant negative income effect for full-time work. In contrast, changes in non-labor income do not significantly affect the probability of part-time work. Young children have a very strong negative effect on working full-time. Specifically, women whose youngest child is aged one year or under are, on average, 33.87(1.82) percentage points less likely to be working full-time than otherwise identical women without children. The effect of children on a woman's probability of engaging in full-time work decreases quickly as the age of the woman's youngest child increases. Indeed, a woman whose youngest child is aged between 12 and 16 years has the same probability of working full-time as an otherwise identical woman without children. A youngest child aged between 1 and 7 years has a large positive effect the probability of working part-time. Expecting a child in 4-6 months time significantly decreases the probability of full-time work while, *ceteris paribus*, women who are expecting a child in the next three month have a significantly lower probability of working full-time or part-time than women who are not expecting a child.

The results for Specification VI in Table 5 further show that, on average over the sampled women, working full-time rather than being non-employed in the previous year increases the probability of working full-time in the current year by 44.24(2.43) percentage points. Similarly, working part-time rather than being non-employed increases the probability of working part-time in the current year by 24.62(2.41) percentage points. These results confirm the presence of significant state dependence on women's labor supply behavior. Intertemporal dependencies are explored in more detail below in Section 5.2.

In terms of the distribution of the unobservables, the results in Table 4 pertaining to Specification VI reveal negative first order autocorrelation in the unobservables affecting payoffs from full-time employment and positive first order autocorrelation in the unobservables affecting payoffs from part-time employment.¹⁶ Women with young children have very large amounts of variation in their unobserved payoffs from working full-time. There is also significant variation in women's unobserved preferences for part-time employment if they have a young child, but far less than for full-time employment. This variation in payoffs might reflect unobserved variation in child-care costs or productivity in home production. Similarly, women with a degree level qualification have a significantly higher level of unobserved variation in their payoffs from working full-time than women with other levels of qualifications. This is consistent with a relatively high level of heterogeneity in the labor market

¹⁶A likelihood ratio test for the joint significance of ρ_f and ρ_p reveals that the null hypothesis of no autocorrelation in the time-varying component of unobservables is rejected at all conventional significance levels.

Variable	Spec. I		Spec. II		Spec. III		Spec. IV		Spec. V		Spec. VI	
	<i>f</i>	<i>p</i>	<i>f</i>	<i>p</i>	<i>f</i>	<i>p</i>	<i>f</i>	<i>p</i>	<i>f</i>	<i>p</i>	<i>f</i>	<i>p</i>
$Y_{i,f,t-2}$	2.19 (0.11)	1.04 (0.09)	1.59 (0.13)	0.66 (0.11)	1.59 (0.13)	0.67 (0.11)	1.98 (0.26)	0.82 (0.16)	1.72 (0.18)	0.78 (0.16)	2.13 (0.29)	0.83 (0.18)
$Y_{i,p,t-2}$	1.09 (0.11)	1.37 (0.07)	0.72 (0.13)	0.95 (0.09)	0.72 (0.13)	0.95 (0.09)	0.88 (0.17)	1.02 (0.12)	0.77 (0.16)	1.01 (0.12)	0.92 (0.21)	1.14 (0.15)
$Y_{i,f,t-1}$	3.81 (0.10)	1.50 (0.09)	2.90 (0.13)	0.98 (0.11)	2.93 (0.13)	1.01 (0.11)	4.27 (0.46)	1.01 (0.21)	3.50 (0.22)	1.16 (0.16)	5.28 (0.57)	1.37 (0.27)
$Y_{i,p,t-1}$	2.09 (0.11)	2.85 (0.07)	1.55 (0.13)	2.34 (0.09)	1.55 (0.13)	2.34 (0.09)	1.87 (0.26)	2.38 (0.13)	1.83 (0.19)	2.43 (0.10)	2.28 (0.31)	2.52 (0.16)
DEGREE	0.66 (0.10)	0.30 (0.09)	1.00 (0.16)	0.54 (0.13)	0.98 (0.16)	0.53 (0.13)	1.23 (0.21)	0.63 (0.15)	1.04 (0.17)	0.60 (0.14)	1.24 (0.21)	0.66 (0.16)
HIGH QUAL.	0.67 (0.09)	0.34 (0.08)	1.00 (0.13)	0.57 (0.11)	0.98 (0.13)	0.55 (0.11)	1.24 (0.18)	0.66 (0.13)	1.03 (0.14)	0.64 (0.12)	1.24 (0.19)	0.72 (0.14)
MEDIUM QUAL.	0.41 (0.11)	0.12 (0.09)	0.62 (0.16)	0.26 (0.14)	0.62 (0.16)	0.26 (0.14)	0.77 (0.21)	0.32 (0.15)	0.59 (0.17)	0.27 (0.15)	0.73 (0.21)	0.31 (0.16)
LOW QUAL.	0.34 (0.08)	0.14 (0.07)	0.50 (0.13)	0.25 (0.10)	0.48 (0.13)	0.24 (0.10)	0.63 (0.16)	0.30 (0.11)	0.58 (0.15)	0.35 (0.13)	0.69 (0.18)	0.41 (0.14)
INCOME/100	-1.21 (0.15)	-0.28 (0.13)	-1.32 (0.21)	-0.35 (0.16)	-1.36 (0.22)	-0.36 (0.16)	-1.66 (0.30)	-0.47 (0.17)	-1.34 (0.24)	-0.40 (0.16)	-1.73 (0.34)	-0.52 (0.18)
AGE/10	0.42 (0.25)	0.34 (0.24)	0.69 (0.35)	0.43 (0.31)	0.54 (0.35)	0.30 (0.31)	0.99 (0.43)	0.57 (0.34)	0.46 (0.34)	0.33 (0.34)	0.31 (0.46)	0.28 (0.37)
AGE ² /100	-0.09 (0.03)	-0.04 (0.03)	-0.14 (0.04)	-0.05 (0.04)	-0.12 (0.04)	-0.03 (0.04)	-0.19 (0.05)	-0.07 (0.04)	-0.11 (0.05)	-0.04 (0.04)	-0.11 (0.06)	-0.04 (0.05)
PRE SCH. CH.	0.45 (0.12)	0.13 (0.09)	0.23 (0.15)	0.06 (0.11)	0.23 (0.15)	0.06 (0.11)	0.35 (0.19)	0.07 (0.12)	0.01 (0.17)	-0.03 (0.12)	0.09 (0.22)	0.01 (0.13)
YOUNGEST CH. 0-1 YEAR	-4.46 (0.19)	-1.57 (0.17)	-5.19 (0.24)	-2.00 (0.20)	-5.12 (0.24)	-1.96 (0.20)	-6.73 (0.63)	-2.38 (0.31)	-5.80 (0.39)	-1.92 (0.26)	-7.79 (0.85)	-2.31 (0.35)
YOUNGEST CH. 1-2 YEARS	-1.30 (0.17)	0.41 (0.15)	-2.07 (0.22)	0.05 (0.18)	-1.99 (0.22)	0.10 (0.18)	-2.72 (0.38)	-0.13 (0.21)	-1.73 (0.25)	0.14 (0.20)	-2.27 (0.39)	-0.06 (0.23)
YOUNGEST CH. 3-4 YEARS	-0.91 (0.15)	0.34 (0.12)	-1.49 (0.15)	0.15 (0.15)	-1.42 (0.19)	0.21 (0.15)	-1.98 (0.30)	0.06 (0.17)	-1.27 (0.20)	0.23 (0.16)	-1.68 (0.31)	0.11 (0.18)
YOUNGEST CH. 5-6 YEARS	-0.23 (0.14)	0.56 (0.12)	-0.69 (0.18)	0.49 (0.14)	-0.63 (0.18)	0.54 (0.15)	-0.98 (0.25)	0.44 (0.16)	-0.57 (0.19)	0.56 (0.16)	-0.75 (0.26)	0.54 (0.17)
YOUNGEST CH. 7-11 YEARS	-0.10 (0.11)	0.35 (0.10)	-0.29 (0.15)	0.41 (0.13)	-0.23 (0.14)	0.46 (0.13)	-0.46 (0.19)	0.40 (0.14)	-0.19 (0.15)	0.47 (0.14)	-0.28 (0.20)	0.48 (0.15)
YOUNGEST CH. 12-15 YEAR	0.30 (0.13)	0.27 (0.12)	0.27 (0.16)	0.32 (0.15)	0.33 (0.16)	0.36 (0.15)	0.28 (0.19)	0.34 (0.16)	0.34 (0.17)	0.37 (0.16)	0.39 (0.20)	0.37 (0.17)
EXP. CH. IN 0-3 MONTHS	-1.44 (0.19)	-1.14 (0.21)	-1.81 (0.22)	-1.39 (0.22)	-1.82 (0.22)	-1.39 (0.22)	-2.32 (0.33)	-1.58 (0.24)	-2.31 (0.26)	-1.68 (0.23)	-2.96 (0.41)	-1.83 (0.26)
EXP. CH. IN 4-6 MONTHS	-0.61 (0.26)	-0.38 (0.20)	-0.86 (0.22)	-0.49 (0.22)	-0.86 (0.31)	-0.50 (0.22)	-1.08 (0.40)	-0.53 (0.23)	-0.96 (0.33)	-0.55 (0.24)	-1.23 (0.43)	-0.61 (0.25)
INTERCEPT	-3.59 (0.48)	-3.20 (0.47)	-4.11 (0.68)	-3.41 (0.63)	-4.61 (0.79)	-3.66 (0.78)	-5.69 (0.95)	-3.72 (0.69)	-4.10 (0.78)	-3.28 (0.69)	-4.97 (1.05)	-3.22 (0.75)
Log Likelihood	-17524.43	-10697.87	-10690.95	-10683.49	-10622.91	-10607.21						
Pseudo R ²	47.57%	54.57%	54.60%	54.88%	54.88%	54.95%						
AIC	36849.94	21557.75	21555.90	21538.98	21473.82	21452.43						
BIC	36885.00	22205.51	22251.64	22226.73	22385.49	22404.08						

Notes: Standard errors in parenthesis. Estimates of the coefficients on the time dummies and the initial conditions are omitted. Columns headed *f* and *p* contain the coefficients describing payoffs from respective full-time employment and part-time employment. AIC=-2Log likelihood+2Parameters; BIC=-2Log likelihood+ ln(NT)Parameters; Pseudo R²=1-Restricted Log likelihood/Unrestricted Log likelihood.

Table 3: Estimates of the deterministic components of the coefficients for Specifications I-VI of the dynamic mixed multinomial logit model.

	Spec. II	Spec. III	Spec. IV	Spec. V	Spec. VI
$\Sigma_{\text{Intercept } 1}$	$\begin{pmatrix} 2.78 & . \\ (0.32) & \end{pmatrix}$	$\begin{pmatrix} 1.52 & . \\ (0.62) & \end{pmatrix}$	$\begin{pmatrix} 4.37 & . \\ (0.74) & \end{pmatrix}$	$\begin{pmatrix} 0.18 & . \\ (0.39) & \end{pmatrix}$	$\begin{pmatrix} 0.09 & . \\ (0.25) & \end{pmatrix}$
$\Sigma_{\text{Intercept } 2}$	$\begin{pmatrix} 1.62 & 1.39 \\ (0.22) & (0.20) \end{pmatrix}$	$\begin{pmatrix} 1.00 & 1.18 \\ (0.98) & (3.13) \end{pmatrix}$	$\begin{pmatrix} 2.24 & 1.70 \\ (0.38) & (0.30) \end{pmatrix}$	$\begin{pmatrix} 0.32 & 0.62 \\ (0.40) & (0.36) \end{pmatrix}$	$\begin{pmatrix} 0.20 & 0.58 \\ (0.35) & (0.32) \end{pmatrix}$
μ_2		$\begin{pmatrix} 3.30 & . \\ (1.08) & \end{pmatrix}$			
α		$\begin{pmatrix} 1.48 & 0.83 \\ (0.56) & (0.38) \end{pmatrix}$			
ρ_f		$\begin{pmatrix} 2.08 \\ (2.13) \end{pmatrix}$			
ρ_p		$\begin{pmatrix} 1.47 \\ (1.10) \end{pmatrix}$			
ρ_f		0.70 (0.02)			
ρ_p			-0.17 (0.10)		-0.12 (0.10)
ρ_p			0.45 (0.29)		0.43 (0.26)
Σ_{ζ}			$\begin{pmatrix} 3.35 & . \\ (1.33) & \end{pmatrix}$		$\begin{pmatrix} 4.08 & . \\ (1.59) & \end{pmatrix}$
$\Sigma_{Y_{i,f,t-2}}$			$\begin{pmatrix} 0.93 & . \\ (0.32) & \end{pmatrix}$	$\begin{pmatrix} 0.93 & . \\ (0.43) & \end{pmatrix}$	$\begin{pmatrix} 1.18 & 0.47 \\ (0.53) & (0.37) \end{pmatrix}$
$\Sigma_{Y_{i,p,t-2}}$			$\begin{pmatrix} 0.61 & 0.43 \\ (0.34) & (0.40) \end{pmatrix}$	$\begin{pmatrix} 0.61 & 0.43 \\ (0.34) & (0.40) \end{pmatrix}$	$\begin{pmatrix} 0.00 & 0.17 \\ (0.00) & (0.29) \end{pmatrix}$
$\Sigma_{Y_{i,f,t-1}}$			$\begin{pmatrix} 0.74 & . \\ (0.48) & \end{pmatrix}$	$\begin{pmatrix} 0.74 & . \\ (0.48) & \end{pmatrix}$	$\begin{pmatrix} 1.32 & . \\ (0.76) & \end{pmatrix}$
$\Sigma_{Y_{i,p,t-1}}$			$\begin{pmatrix} 0.63 & 0.56 \\ (0.32) & (0.28) \end{pmatrix}$	$\begin{pmatrix} 0.63 & 0.56 \\ (0.32) & (0.28) \end{pmatrix}$	$\begin{pmatrix} 0.93 & 0.66 \\ (0.43) & (0.33) \end{pmatrix}$
$\Sigma_{Y_{i,f,t-1}}$			$\begin{pmatrix} 0.36 & . \\ (0.38) & \end{pmatrix}$	$\begin{pmatrix} 0.36 & . \\ (0.38) & \end{pmatrix}$	$\begin{pmatrix} 0.75 & . \\ (0.58) & \end{pmatrix}$
$\Sigma_{Y_{i,p,t-1}}$			$\begin{pmatrix} 0.35 & 0.71 \\ (0.38) & (0.55) \end{pmatrix}$	$\begin{pmatrix} 0.35 & 0.71 \\ (0.38) & (0.55) \end{pmatrix}$	$\begin{pmatrix} 0.51 & 0.58 \\ (0.45) & (0.55) \end{pmatrix}$
Σ_{DEGREE}			$\begin{pmatrix} 0.77 & . \\ (0.61) & \end{pmatrix}$	$\begin{pmatrix} 0.77 & . \\ (0.61) & \end{pmatrix}$	$\begin{pmatrix} 0.92 & . \\ (0.99) & \end{pmatrix}$
Σ_{DEGREE}			$\begin{pmatrix} 0.26 & 0.09 \\ (0.27) & (0.16) \end{pmatrix}$	$\begin{pmatrix} 0.26 & 0.09 \\ (0.27) & (0.16) \end{pmatrix}$	$\begin{pmatrix} 0.39 & 0.17 \\ (0.37) & (0.25) \end{pmatrix}$
Σ_{DEGREE}			$\begin{pmatrix} 1.02 & . \\ (0.56) & \end{pmatrix}$	$\begin{pmatrix} 1.02 & . \\ (0.56) & \end{pmatrix}$	$\begin{pmatrix} 1.96 & . \\ (0.91) & \end{pmatrix}$
Σ_{DEGREE}			$\begin{pmatrix} 0.61 & 0.48 \\ (0.39) & (0.35) \end{pmatrix}$	$\begin{pmatrix} 0.61 & 0.48 \\ (0.39) & (0.35) \end{pmatrix}$	$\begin{pmatrix} 0.97 & 0.52 \\ (0.54) & (0.40) \end{pmatrix}$
Σ_{CHILD}			$\begin{pmatrix} 14.31 & . \\ (3.04) & \end{pmatrix}$	$\begin{pmatrix} 14.31 & . \\ (3.04) & \end{pmatrix}$	$\begin{pmatrix} 27.15 & . \\ (7.22) & \end{pmatrix}$
Σ_{CHILD}			$\begin{pmatrix} 6.41 & 3.03 \\ (1.57) & (0.96) \end{pmatrix}$	$\begin{pmatrix} 6.41 & 3.03 \\ (1.57) & (0.96) \end{pmatrix}$	$\begin{pmatrix} 10.33 & 3.97 \\ (2.59) & (1.29) \end{pmatrix}$

Notes: Standard errors in parenthesis. Specification I has no unknown parameters in the distribution of the unobservables. In Specifications II and IV-VI, $\Sigma_{\text{Intercept } 1}$ is the covariance matrix of the time invariant components of the random intercepts. Specification III has time invariant random intercepts with a distribution obtained from the mixture of two bivariate normal distributions: with probability α the random intercepts have mean zero and variance $\Sigma_{\text{Intercept } 1}$ and with probability $(1 - \alpha)$ the random intercepts have mean μ_2 and variance $\Sigma_{\text{Intercept } 2}$. In specifications allowing autocorrelation in the random intercepts, ρ_f and ρ_p are the first order autocorrelation coefficients and Σ_{ζ} is the covariance matrix of the innovations in the autoregressive processes. Σ_{DEGREE} and Σ_{CHILD} are the covariance matrices of the random coefficients on the indicated variables. The covariance matrices of the random coefficients on the initial conditions in Specifications V and VI are not reported.

Table 4: Estimates of parameters appearing in the distribution of unobservables for Specifications II-VI of the dynamic mixed multinomial logit model.

returns to a university education. The coefficients on lagged employment behavior display relatively small and generally insignificant amounts of variation. In contrast, the random coefficients on the initial conditions, not reported, display significant variation indicating differences in the distribution of unobserved individual characteristics according the individual's initial employment status.

5.2 Employment Dynamic following Job Creation Policies

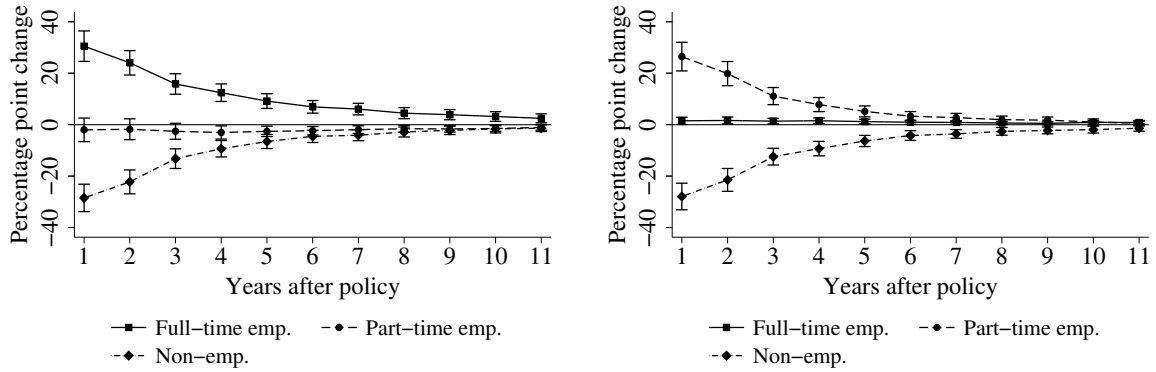
In this section we explore the short and long-run effectiveness of labor market policies that temporarily affect women's incentives to work full-time or to work part-time. The design of these "job creation

Variable	Spec. I		Spec. II		Spec. III		Spec. IV		Spec. V		Spec. VI	
	<i>f</i>	<i>p</i>	<i>f</i>	<i>p</i>	<i>f</i>	<i>p</i>	<i>f</i>	<i>p</i>	<i>f</i>	<i>p</i>	<i>f</i>	<i>p</i>
$n_{t-1} \rightarrow f_{t-1}$	50.80 (1.63)	-7.79 (1.39)	31.37 (1.75)	-9.45 (1.71)	31.48 (3.54)	-8.98 (3.02)	37.95 (2.48)	-15.61 (2.05)	36.00 (2.72)	-11.76 (2.89)	44.24 (2.43)	-18.03 (2.38)
$n_{t-1} \rightarrow p_{t-1}$	3.47 (1.09)	39.97 (1.31)	-2.37 (1.95)	27.77 (2.27)	-1.88 (1.97)	27.61 (1.75)	-0.22 (1.96)	25.02 (1.94)	-0.61 (1.94)	26.93 (3.47)	2.04 (1.71)	24.62 (2.41)
NO EDUC. \rightarrow LOW EDUC.	3.83 (0.67)	-0.40 (0.80)	5.81 (1.14)	0.11 (1.09)	5.67 (0.94)	-0.05 (1.07)	5.50 (1.14)	0.28 (1.20)	5.29 (0.98)	0.45 (0.92)	4.62 (1.06)	0.88 (0.98)
NO EDUC. \rightarrow MED. EDUC.	3.66 (0.59)	-0.06 (0.70)	5.63 (0.91)	0.41 (0.97)	5.53 (0.93)	0.20 (1.19)	5.38 (0.82)	0.56 (0.83)	4.98 (0.90)	0.94 (0.88)	4.36 (0.87)	1.43 (0.95)
NO EDUC. \rightarrow HIGH EDUC.	2.70 (0.61)	-0.91 (0.75)	4.02 (0.99)	-0.67 (0.91)	4.11 (1.09)	-0.82 (1.30)	3.75 (1.07)	-0.41 (0.89)	3.49 (0.98)	-0.53 (1.20)	3.09 (0.95)	-0.10 (1.16)
NO EDUC. \rightarrow DEGREE	2.05 (0.48)	-0.33 (0.62)	3.04 (0.99)	-0.10 (0.84)	2.92 (0.93)	-0.13 (0.96)	2.88 (0.76)	0.09 (0.84)	2.83 (0.71)	0.21 (0.63)	2.49 (0.68)	0.44 (0.86)
INCOME+£1000	-0.08 (0.02)	0.03 (0.02)	-0.08 (0.02)	0.04 (0.02)	-0.10 (0.02)	0.03 (0.02)	-0.11 (0.02)	0.05 (0.02)	-0.10 (0.02)	0.03 (0.02)	-0.08 (0.02)	0.03 (0.02)
AGE + 1 YEAR	-0.28 (0.04)	0.22 (0.04)	-0.42 (0.04)	0.27 (0.05)	-0.44 (0.05)	0.29 (0.05)	-0.40 (0.05)	0.27 (0.06)	-0.40 (0.04)	0.25 (0.04)	-0.36 (0.04)	0.23 (0.05)
PRE. SCH. CH. + 1	3.08 (0.98)	-1.26 (0.92)	1.73 (1.18)	-0.83 (0.96)	1.67 (1.27)	-0.74 (1.16)	2.05 (1.04)	-1.05 (1.15)	0.22 (1.17)	-0.35 (1.05)	0.52 (1.13)	-0.34 (1.08)
NO CHILDREN \rightarrow CHILD AGED \leq 1 YEAR	-36.73 (1.72)	2.31 (1.94)	-38.15 (1.61)	1.99 (2.31)	-37.99 (4.29)	1.84 (3.18)	-37.25 (1.47)	0.96 (2.11)	-34.61 (1.78)	1.30 (2.34)	-33.87 (1.82)	-0.21 (2.46)
NO CHILDREN \rightarrow 1 YEAR <CHILD AGED \leq 2 YEARS	-15.61 (1.77)	14.08 (1.54)	-21.51 (1.80)	15.58 (2.03)	-20.93 (2.67)	15.52 (2.55)	-20.23 (1.77)	13.98 (2.00)	-17.83 (1.91)	13.96 (2.03)	-15.83 (2.05)	11.78 (2.08)
NO CHILDREN \rightarrow 2 YEAR <CHILD AGED \leq 4 YEARS	-11.19 (1.27)	10.37 (1.22)	-16.44 (1.63)	12.98 (1.70)	-16.13 (1.97)	13.11 (1.98)	-15.77 (1.45)	12.18 (1.70)	-14.14 (1.54)	12.00 (1.66)	-12.77 (1.51)	10.70 (1.75)
NO CHILDREN \rightarrow 4 YEAR <CHILD AGED \leq 7 YEARS	-6.33 (1.10)	8.49 (0.99)	-11.04 (1.26)	11.71 (1.40)	-10.51 (1.49)	11.61 (1.51)	-10.52 (1.30)	11.03 (1.39)	-9.77 (1.39)	11.24 (1.44)	-8.68 (1.12)	10.26 (1.29)
NO CHILDREN \rightarrow 7 YEAR <CHILD AGED \leq 11 YEARS	-3.78 (0.82)	5.25 (0.88)	-6.41 (1.07)	7.69 (1.23)	-6.04 (0.98)	7.68 (1.06)	-6.37 (0.96)	7.53 (0.97)	-5.57 (0.82)	7.42 (0.95)	-5.11 (0.99)	7.13 (1.05)
NO CHILDREN \rightarrow 11 YEAR <CHILD AGED \leq 16 YEARS	0.34 (1.17)	1.89 (1.04)	-0.09 (1.03)	2.53 (1.09)	0.22 (1.23)	2.56 (1.10)	-0.37 (0.92)	2.58 (1.09)	0.15 (0.90)	2.62 (1.02)	0.08 (1.11)	2.47 (1.10)
NOT EXP. CHILD \rightarrow EXP. CHILD IN 0-3 MONTHS	-6.42 (1.54)	-5.37 (2.12)	-8.72 (1.75)	-6.05 (2.30)	-8.96 (2.02)	-6.09 (2.60)	-8.74 (1.82)	-6.41 (1.97)	-10.61 (1.83)	-6.89 (2.38)	-10.59 (2.10)	-6.43 (2.81)
NOT EXP. CHILD \rightarrow EXP. CHILD IN 4-6 MONTHS	-2.91 (2.04)	-0.94 (1.93)	-4.81 (2.36)	-0.49 (2.07)	-4.79 (2.29)	-4.65 (2.60)	-4.68 (2.17)	-0.53 (1.94)	-4.85 (2.14)	-0.82 ()	-4.78 (2.46)	-0.55 (2.20)

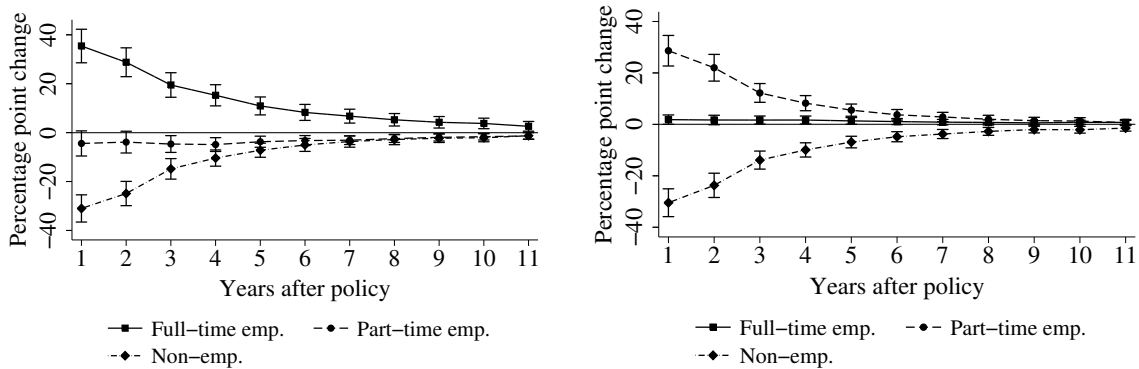
Notes: Standard errors in parenthesis. Marginal effects, expressed in percentage points, were computed by simulating, for each individual in the sample, employment behavior before and after a change in the relevant characteristic, and then averaging the responses over time and over individuals. Columns headed *f* contain the marginal effects for full-time employment and columns headed *p* contain marginal effects for part-time employment.

Table 5: Marginal effects for Specifications I-VI of the dynamic mixed multinomial logit model.

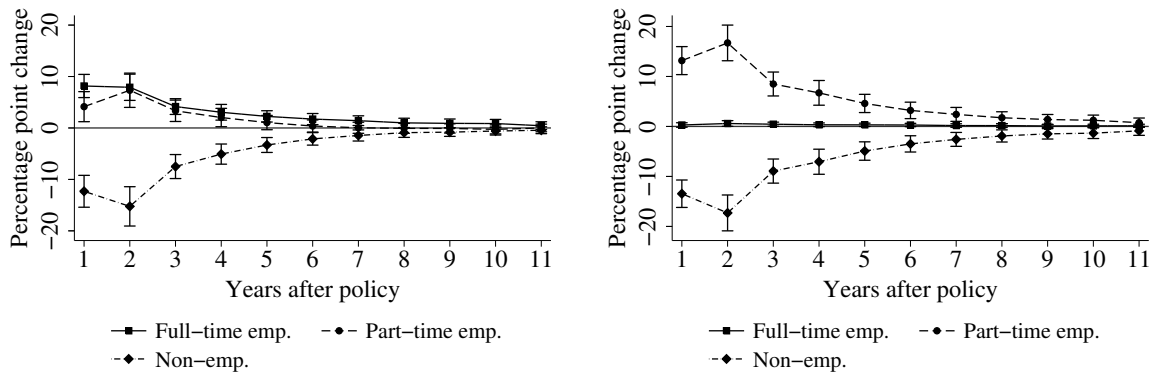
policies” is such that they cause women who chose non-employment in both their first and second years in the sample to move into, depending on the policy, either full-time or part-time work. The interventions themselves last only one year and therefore employment outcomes subsequent to the policies are affected only via the effect of the individual’s previous employment outcome on her current payoffs. We look first at the effect of these policies averaged over the sample and then explore how employment behavior following temporary job creation policies varies by education and household structure.



(a) Non-employed women moved into full-time work - Sample average. (b) Non-employed women moved into part-time work - Sample average.



(c) Non-employed women moved into full-time work - Highly educated. (d) Non-employed women moved into part-time work - Highly educated.



(e) Non-employed women moved into full-time work - Young child. (f) Non-employed women moved into part-time work - Young child.

Notes: Vertical bars represent 95% confidence intervals.

Figure 2: Simulated employment effects of policy interventions based on Specification VI.

Figure 2(a) shows the average long-run effect on employment behavior of a temporary policy causing non-employed individuals to move into full-time work. We see that this policy causes a significant increase in full-time work of 30.5 percentage points in the year immediately after the policy is removed. This policy continues to have a significant positive, but smaller, effect on employment in subsequent years. This increase in full-time employment is balanced by reductions in non-employment and, to a small extent, part-time work. The reduction in part-time work implies that temporarily incentivizing non-employed individuals to work full-time has a small crowding out effect on part-time work. For example, one year after the policy intervention, labor force participation is 28.5 percentage points higher than in the absence of the policy intervention, as the policy leads to a 2 percentage point reduction in part-time work.

Similarly, Figure 2(b) shows the average long-run effect on employment behavior of a temporary policy causing non-employed individuals to move into part-time work. This policy causes a significant increase in part-time work, and a slight increase in full-time work, together balanced by a reduction in non-employment. The one year own-state dependence effect for part-time employment is 26.5 percentage points, which is smaller than the corresponding own-state dependence effect for full-time employment. However, as the policy incentivizing part-time employment causes a slight increase in full-time employment while incentivizing full-time employment causes a reduction in subsequent part-time employment, the effects of the two policies on the rate non-employment are very similar. For example, one year after the policy intervention incentivizing part-time employment, non-employment is 28 percentage points lower, as compared to 28.5 percentage points if instead the policy targeting full-time employment was implemented. The asymmetric cross-state effects of the two policies highlight the additional insight gained from analyzing labor supply using a multinomial, rather than binary, framework; modeling only the decision as to work full-time would overstate the gain in terms of reduced non-employment obtainable from a policy incentivizing full-time work, and we learn that although part-time employment can provide a stepping-stone into full-time work, the magnitude of this effect is small.

The remaining panels in Figure 2 show how employment dynamics following job creation schemes vary between demographic groups.¹⁷ Figures 2(c) and 2(d) reveal that the state dependence effects for highly educated individuals, defined as those having a university degree, are similar to the effects for the sample average. Figures 2(e) and 2(f) meanwhile show the dynamic responses to the two policy interventions among women who have a child one year after the policy intervention. We see that the policy incentivizing full-time employment causes a far smaller increase in full-time employment among women with children than for the sample average and, in contrast to the results for the sample average, part-time employment increases for this group of women. In fact, subsequent to one year after the policy intervention, part-time employment increases by almost as much as full-time employment. Thus, for women with children, full-time employment provides a stepping stone into part-time employment. For women who experience a birth in the year after the policy intervention, the policy incentivizing part-time employment increases significantly part-time employment and has a positive, but tiny, effect on full-time work. Importantly, for this group of women, incentivizing part-time employment is more effective at reducing non-employment than incentivizing full-time employment; the reduction in non-employment between 1 and 7 years after the policy intervention is between 1 and 2 percentage points higher if the policy incentivized part-time rather than full-time employment. This results implies

¹⁷The illustrated employment effects in Figures 2(c)-2(f) are obtained by averaging over the sample distribution of all other individual characteristics.

that maximizing the long-term effectiveness of policies incentivizing employment requires tailoring of policies according to demographic characteristics. Again, it should be noted that a binary model of labor force participation is uninformative about the relative merits for different demographic groups of labor market policies facilitating full-time or part-time employment.

5.3 Heterogeneity in Labor Supply Dynamics after Child Birth

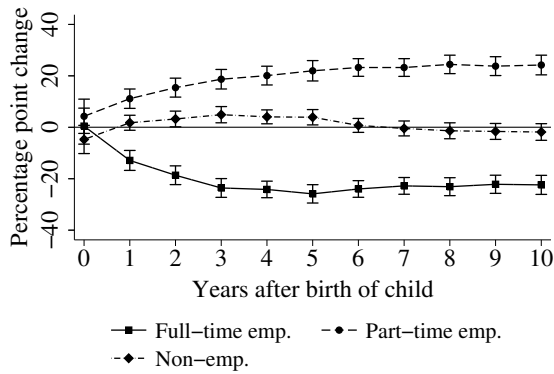
We use the parameter estimates for Specification VI to explore the extent of heterogeneity in labor supply dynamics following the birth of a child. Table 4 shows that there is significant variation in the effects of having a child aged under one year on a woman's payoffs from full-time and, to a lesser but still significant extent, part-time work, relative to non-employment. Together with the significant state dependence effects documented above, heterogeneity in the effects of a young child suggests that there may be persistent differences in labor supply behavior following the birth of a child.

Figure 3 shows the estimated effect a having a child on subsequent employment behavior for women at different points in the distribution of unobserved preferences for full-time and part-time work in the event that they have a young child. Figure 3(a) shows the for women who have a high unobserved preference for full-time work, having a young child has very little immediate effect on labor supply behavior. As the child becomes older these women become more likely to work part-time and less likely to work full-time, as compared to if they had not had a child. Non-employment increases slightly 3-6 years following the birth of the child. Thus, we conclude, that for women with a very strong preference for full-time work in the event that they have a child, there is a substitution away from full-time work, but no pronounced movement away from employment more generally. We see from Figure 3(b) that the picture is dramatically different for women who have a relatively low preference for full-time work if they have a young child. For such women, the birth of a child is accompanied by a large substitution away from both full-time and part-time work and into non-employment. After 2 years, part-time employment is higher than if the women had not had a child, however it takes many years before the labor supply behavior of women with a low preference for full-time work is similar to that of women with a high preference for full-time work. Indeed, 8 years after having a child the proportion of high preference women working full-time is still 6 percentage points higher than that of low preference women.

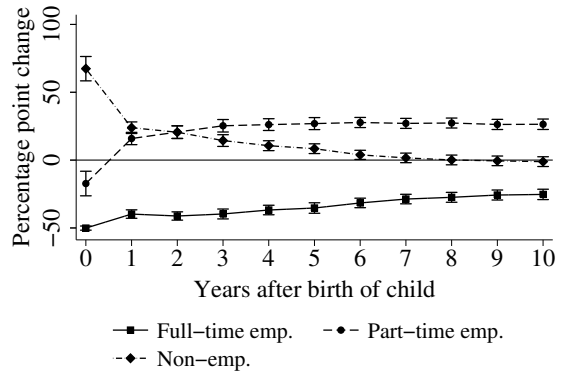
Figures 3(c) and 3(d) illustrate the change in labor supply behavior over time caused by the birth of a child for women with high and low preferences for part-time work in the event they have a child aged under one year. These figures attract two comments. First we see very little difference in labor supply behavior between women with high and with low unobserved preferences for part-time work when they have a young child. Thus, although we find significant heterogeneity in payoffs from part-time work among women who have a child aged under one year, this does not translate into appreciable differences in labor supply behavior following the birth of a child. Second, the confidence intervals at $t = 0$ are very large, which reflects the large amounts of unobserved variation in preferences for full-time employment among women who have a young child.

5.4 Comparisons with More Restrictive Specifications

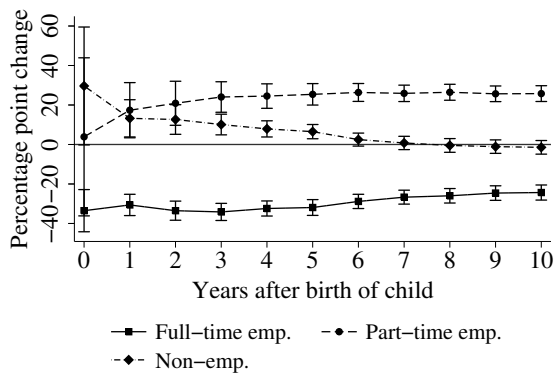
Comparisons are now made with specifications that impose more restrictive distributions of unobservables than Specification VI. Recall that the primary motivation for allowing a generality, in the form of autocorrelation and random coefficients, in the distribution of unobservables was that imposing an overly restrictive distribution of persistent unobservables would likely lead to inconsistent estimates of



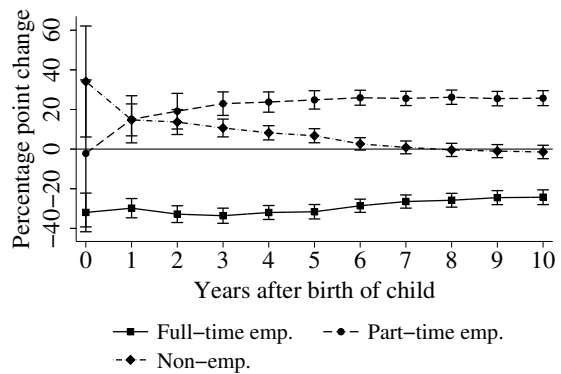
(a) Effect of having a child at $t = 1$ - High unobserved preference for full-time work.



(b) Effect of having a child at $t = 1$ - Low unobserved preference for full-time work.



(c) Effect of having a child at $t = 1$ - High unobserved preference for part-time work.



(d) Effect of having a child at $t = 1$ - Low unobserved preference for part-time work.

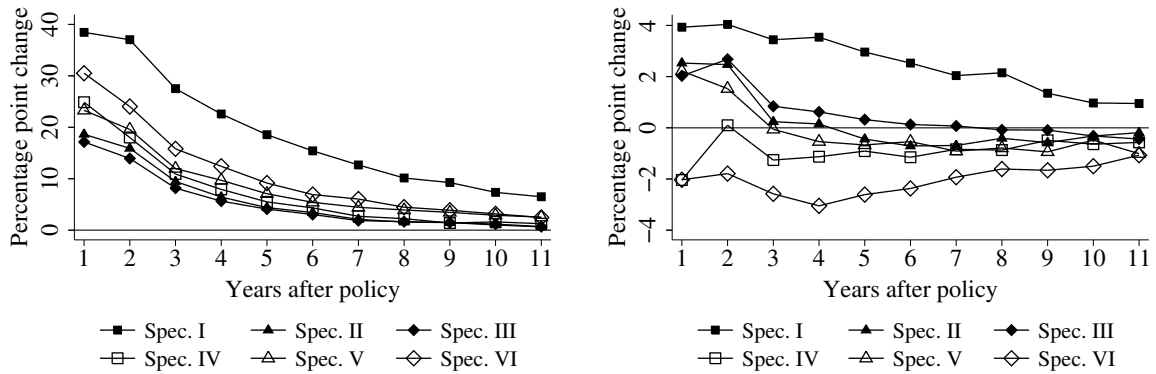
Notes: Vertical bars represent 95% confidence intervals. High and low unobservables refer to the 90th and 10th percentiles of the distribution of unobservables. Other unobservables are drawn from the appropriate conditional distribution. Effects were estimated by averaging over the sample distribution of all observed individual characteristics, except children.

Figure 3: Heterogeneity in labor supply dynamics after child birth.

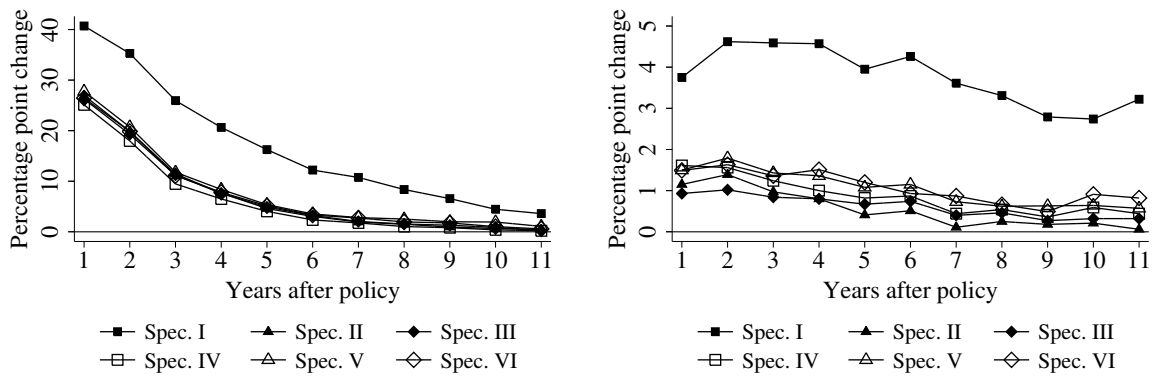
state dependence effects and the associated policy responses. Therefore, when comparing the various specifications, attention focused on differences in estimates of intertemporal dependencies. However, for completeness, at the end of this subsection we discuss briefly model selection criteria and average marginal effects.

Figures 4 and 5 show the effect of temporary job creation policies on subsequent employment behavior as implied by each of the six model specifications under consideration. As in Section 5.2 above, the policies under consideration cause women who chose non-employment in both their first and second years in the sample to move into, depending on the policy, either full-time or part-time work in their second year in the sample. We present state dependence effects averaged over the sample and for women who gave birth to a child one year after the policy intervention. Table 6 details the significance of the difference between the predications based on Specification VI and those based on more restrictive specifications.

Figure 4 shows that on average Specification I implies, for both full-time and part-time work, substantially larger own-state and cross-state dependencies than Specification VI; as expected completely ignoring persistent unobservables leads to an overestimate of the state dependencies in labor supply



(a) Non-employed women moved into full-time work - own-state effect - Sample average. (b) Non-employed women moved into full-time work - cross-state effect - Sample average.

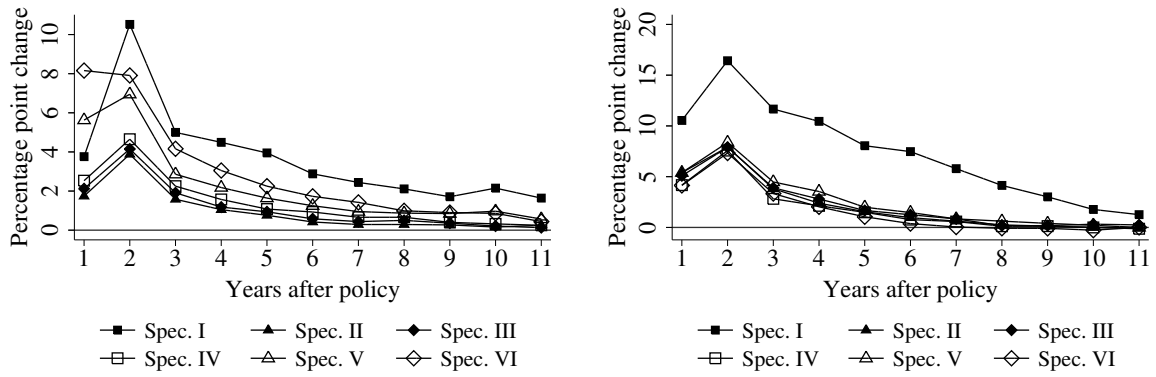


(c) Non-employed women moved into part-time work - own-state effect - Sample average. (d) Non-employed women moved into part-time work - cross-state effect - Sample average.

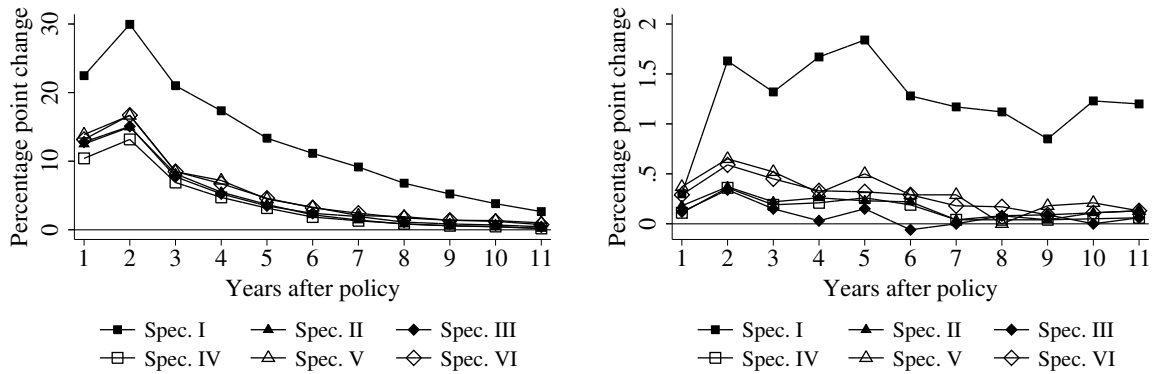
Figure 4: Comparison of the dynamic effects of policy interventions based on Specifications I-VI.

behavior. Looking across all of the own-state and cross-state dependence effects, there is very little difference between the results implied by Specifications II and III. In other words, despite of finding significant evidence for non-normally distributed random intercepts, generalizing the distribution of the time invariant random intercepts to be non-normal does not impact of estimates of intertemporal dependencies.

Focusing first on the effects of the policy incentivizing full-time employment and considering the sample average, we see that the own-state dependencies implied by Specification VI are larger than those implied by Specifications II-V. Specifically, according to Specification VI the one year own-state dependence effect for full-time employment is 30.5 percentage points, while Specifications IV and V, which exclude random coefficients and autocorrelation respectively, suggest a one year own-state dependence effect of around 25 percentage points. Specifications II and III, which allow only random intercepts, meanwhile suggest even lower own-state dependence effects. Furthermore, differences in the estimated own-state dependence effect for full-time employment are evident for several years. Table 6 shows that many of these differences are significant. Thus we conclude that permitting both random coefficients and autocorrelated unobservables is necessary to estimate accurately the degree of own-state dependence in full-time employment. There are also significant differences in the estimated cross-state effects of the policy incentivizing full-time employment on subsequent part-time employment. Specifically, Specification VI suggests a larger negative, or crowding out, effect on part-time employment than the other specifications. For part-time work, the own-state dependencies



(a) Non-employed women moved into full-time work - own-state effect - Young child. (b) Non-employed women moved into full-time work - cross-state effect - Young child.



(c) Non-employed women moved into part-time work - own-state effect - Young child. (d) Non-employed women moved into part-time work - cross-state effect - Young child.

Figure 5: Comparison of the dynamic effects of policy interventions based on Specifications I-VI.

implied by Specifications II and VI, as well as the three intermediate specifications, are very similar. The cross-state dependencies also show little variation. In summary, the modeling of unobserved heterogeneity has greater implications for estimation of own-state and cross-state dependencies for full-time employment than for part-time employment.

Figure 5 shows the own-state and cross-state dependence effect induced by the two policies for women who gave birth to a child one year after the policy intervention.¹⁸ Figure 5(a) reveals that one year after the policy incentivizing non-employed women to move into full-time employment the estimate rate of full-time employment is significantly higher according to Specification VI than according to any of the other specification. Specifically, according to Specification VI, this policy increases the rate of full-time employment one year after the policy by 8 percentage points, while Specification V, which excludes autocorrelation in the random intercepts, implies an effect of 6 percentage points. The remaining specifications suggest one year own-state dependence effects for full-time employment of between 2 and 4 percentage points. For the subgroup of women with children it appears therefore that random coefficients are more important than autocorrelation in the random intercepts, which is in contrast to the results for the sample average where specifications IV and V performed similarly. Table 6 shows that significant differences between the estimated own-state dependence effects for full-time

¹⁸As in Section 5.2, the illustrated employment effects in Figures 2(c)-2(f) are obtained by averaging over the sample distribution of all other individual characteristics.

Policy Effect	Years since Policy Intervention										
	1	2	3	4	5	6	7	8	9	10	11
Own-state Effects											
Difference between Spec. II and Spec. VI											
Δ Full-time emp. (Sample average)	3.93	3.36	3.12	3.44	3.27	2.79	3.43	2.64	2.37	2.24	2.06
Δ Part-time emp. (Sample average)	-0.18	0.00	-0.21	0.11	0.50	0.29	0.95	0.66	1.31	0.62	0.24
Δ Full-time emp. (Young child)	5.58	3.10	3.31	2.56	2.66	2.38	2.35	1.52	1.47	1.71	0.49
Δ Part-time emp. (Young child)	0.46	0.96	0.28	0.98	0.98	1.23	1.39	1.49	1.59	1.51	1.21
Difference between Spec. IV and Spec. VI											
Δ Full-time emp. (Sample average)	1.87	2.49	2.41	2.58	2.51	2.07	2.91	2.03	2.46	1.63	1.35
Δ Part-time emp. (Sample average)	0.47	0.79	0.95	0.94	1.02	1.05	1.17	1.31	1.52	1.05	0.69
Δ Full-time emp. (Young child)	4.90	2.51	2.45	1.90	2.09	1.45	1.58	0.72	1.16	1.32	0.49
Δ Part-time emp. (Young child)	1.93	1.95	1.31	1.59	1.55	1.61	1.55	1.43	1.59	1.42	1.21
Difference between Spec. V and Spec. VI											
Δ Full-time emp. (Sample average)	2.39	1.84	1.89	1.56	1.42	1.20	1.39	0.49	0.41	0.28	-0.08
Δ Part-time emp. (Sample average)	-0.46	-0.37	-0.41	-0.41	-0.19	-0.25	-0.10	-0.81	-0.37	-1.49	-0.91
Δ Full-time emp. (Young child)	2.20	0.74	1.67	1.12	1.09	0.87	0.98	0.26	0.12	-0.27	-0.33
Δ Part-time emp. (Young child)	-0.50	0.03	0.02	-0.37	0.13	-0.11	0.46	-0.26	0.00	-0.21	-0.53
Cross-state Effects											
Difference between Spec. II and Spec. VI											
Δ Part-time emp. (Sample average)	-1.95	-2.06	-1.78	-2.44	-2.11	-1.98	-1.63	-1.69	-1.69	-2.07	-1.54
Δ Full-time emp. (Sample average)	0.47	0.35	0.71	1.29	1.72	1.02	2.17	1.24	1.03	2.19	2.81
Δ Part-time emp. (Young child)	-0.85	-0.34	-0.44	-0.40	-0.70	-0.97	-1.05	-0.67	-0.48	-0.60	0.16
Δ Full-time emp. (Young child)	0.48	0.76	1.10	0.37	0.53	0.41	0.78	0.45	0.29	0.00	0.00
Difference between Spec. IV and Spec. VI											
Δ Part-time emp. (Sample average)	0.00	-0.91	-0.84	-1.47	-1.65	-1.45	-1.43	-1.04	-1.83	-1.51	-0.88
Δ Full-time emp. (Sample average)	-0.17	0.10	0.22	0.93	0.83	0.15	1.23	0.36	0.43	1.00	1.41
Δ Part-time emp. (Young child)	-0.02	-0.14	0.49	-0.11	-0.62	-0.65	-0.98	-0.41	-0.48	-0.60	0.34
Δ Full-time emp. (Young child)	0.78	0.79	1.24	0.63	0.35	0.59	0.78	0.65	0.29	0.35	0.44
Difference between Spec. V and Spec. VI											
Δ Part-time emp. (Sample average)	-1.82	-1.60	-1.59	-1.92	-1.88	-2.18	-1.34	-1.15	-1.13	-1.79	-0.11
Δ Full-time emp. (Sample average)	-0.01	-0.24	-0.13	0.27	0.26	-0.54	0.40	0.12	-0.47	0.84	0.93
Δ Part-time emp. (Young child)	-0.85	-0.61	-1.08	-1.66	-1.32	-1.82	-1.45	-1.43	-1.07	-1.07	0.00
Δ Full-time emp. (Young child)	-0.35	-0.21	-0.33	0.11	-1.06	0.00	-0.61	0.85	-0.53	-0.59	0.00

Notes: Standard errors are bootstrapped. Δ denotes the change relative to the baseline case, where no policy intervention took place. Young child refers to women who had a young child one year after the policy intervention. Specification I is omitted because predictions are always significantly different to those from Specification VI. Specification III is omitted because results are almost identical to those from Specification II.

Table 6: t tests for significance of differences in the own and cross-state effects of policies incentivizing full-time and part-time employment.

employment based on the different specifications persist for up to 7 years. Panels (c)- (d) in Figure 5 show little variation across specifications II-VI in the predicted employment behavior following the policy incentivizing part-time work. Additionally, cross-state effects of previous full-time employment on current part-time employment, shown in Figure 5(b) are also similar across specifications II-VI.

Table 5 shows that estimates of average marginal effects of changes in individual characteristics are rather robust to the assumed distribution of the unobservables. Thus the sensitivity of our results to the assumed distribution of unobservables does not extend to estimates of the average effect of observed individual characteristics on labor supply behavior. Finally, we note that model selection criteria, presented in Table 3, are inconclusive regarding the preferred specification of unobservables: the Akaike Information Criterion (Akaike, 1973) suggests that Specification VI is preferred, while the Bayesian Information Criterion (Schwartz, 1978), which imposes a greater penalty for model

complexity, selects Specification II.

5.5 Robustness to Random Effects Assumption

The robustness of the results to the random effects assumption common to all of these models is investigated by estimating dynamic linear probability models in which persistent unobserved individual characteristics take the form of fixed effects rather than random effects. The linearity of these models means that results will not necessarily be consistent with an interpretation as choice probabilities, in particular predicted probabilities may lie outside the unit interval and estimated marginal effects may be greater than one in absolute terms. However, the linearity of the resulting equations allows the inclusion of individual specific fixed effects and therefore persistent unobservables have an arbitrary distribution and an unrestricted relationship with the explanatory variables. A comparison of the estimation results from these models with the results obtained from a dynamic mixed multinomial logit model in which persistent unobservables take the form of time invariant individual specific random effects, i.e., Specifications II and III, therefore provides a robustness check of the random effects assumption used in the dynamic mixed multinomial logit models.

The dynamic linear probability models considered herein are described by the following two equations

$$Y_{i,f,t} = \Omega_{i,t-1}\lambda_f + X_{i,t}\beta_f + \theta_{i,f} + \varepsilon_{i,f,t} \text{ for } t = 3, \dots, T, \quad (10a)$$

$$Y_{i,p,t} = \Omega_{i,t-1}\lambda_p + X_{i,t}\beta_p + \theta_{i,p} + \varepsilon_{i,p,t} \text{ for } t = 3, \dots, T. \quad (10b)$$

In the above β_j and λ_j for $j = f, p$ are suitably dimensioned vectors of unknown parameters. The specification of the individual's employment history, $\Omega_{i,t-1}$, is as described above in Section 3. Total unobservables have been decomposed into time invariant components, $\theta_{i,f}$ and $\theta_{i,p}$, and time-varying components $\varepsilon_{i,f,t}$ and $\varepsilon_{i,p,t}$. If $\theta_{i,f}$ and $\theta_{i,p}$ are allowed to be fixed effects, consistent estimates of $[\beta_f, \lambda_f]$ and $[\beta_p, \lambda_p]$ can be obtained by using appropriate lagged levels of the woman's employment history and time-varying individual characteristics as instruments for first difference versions of the above equations (see Anderson and Hsiao, 1982; Arellano and Bond, 1991; Hotz *et al.*, 1988).

Table 7 reports the relevant results from the dynamic linear probability models of full-time and part-time work obtained from Generalized Method of Moments (GMM) estimation of the first difference version of Equations (10a) and (10b). While the dynamic mixed multinomial logit models and the dynamic linear probability models draw on different distributional assumptions, GMM estimation of the first difference linear probability models for full-time and part-time work implies own-state dependencies that are broadly comparable to the average marginal effects of previous employment outcomes based on the results from Specifications II and III of the dynamic mixed multinomial logit model, reported in Table 5. For example, Specification II of the dynamic mixed multinomial logit model implies that working full-time or part-time rather than being non-employed in the previous year increases a woman's probability of being in the same employment state in the current year by 31.37(1.74) and 27.77(2.27) percentage points respectively. The corresponding figures based on the dynamic linear probability models are 35.97(2.89) and 33.91(2.47) percentage points.

The one year cross-state effects implied by the dynamic mixed multinomial logit model and the dynamic linear probability models also are comparable. According to the dynamic linear probability models, working part-time rather than being non-employed decreases a woman's probability of working full-time in the current year by 3.83(1.64) percentage points, while previous full-time employment reduces the probability of current part-time employment by 7.21(2.81) percentage points. The cor-

responding figures based in Specification II of the dynamic mixed multinomial logit model, reported in Table 5, are 2.37(1.95) and 9.45(1.71) percentage points. These similarities in the own-state and cross-state dependencies suggest that the random effects assumption imposed when estimating the dynamic mixed multinomial logit models is not overly restrictive.

	Full-time Work	Part-time Work
$Y_{i,f,t-1}$	35.97 (2.89)	-7.21 (2.81)
$Y_{i,p,t-1}$	-3.83 (1.64)	33.91 (2.47)
$Y_{i,f,t-2}$	9.03 (1.53)	-0.44 (1.80)
$Y_{i,p,t-2}$	-1.15 (1.11)	8.17 (1.69)
Sargan Test	282.99 $p=0.373$	295.44 $p=0.293$
Test for AR(1) in residuals	-18.37 $p=0.000$	-21.25 $p=0.000$
Test for AR(2) in residuals	-1.07 $p=0.283$	-0.67 $p=0.500$
Test for AR(3) in residuals	0.64 $p=0.524$	0.30 $p=0.765$
F test for joint significance of $Y_{i,f,t-3}$ and $Y_{i,p,t-3}$	1.84 $p=0.159$	1.55 $p=0.213$
# Parameters	26	26
# Instruments	250	250
# Observations - $\sum_{i=1}^N T$	17299	17299

Notes: Coefficient estimates have been multiplied by 100. Standard errors in parenthesis. Both equations further include the household structure variables, fertility expectations, non-labor income, year dummies and age terms. Instruments consist of employment outcomes, income and child variables in the years $t - 2$, $t - 3$ and $t - 4$, and current dated values of age and the time dummies. While deeper lags of employment outcomes, income and child variables should also provide valid instruments, further moment conditions are not exploited as the Sargan statistics obtained when utilizing a large number of over-identifying restrictions have been shown to under reject when testing the validity of the instrument set or other relevant hypotheses (see Bowsher, 2002).

Table 7: GMM estimation of dynamic linear probability models of full-time and part-time work.

The results from GMM estimation of the dynamic linear probability models further reveal significant own state effects over a two year time horizon. However, for both full-time and part-time work, F tests for the joint significance of employment outcomes dated three years previously do not reject the null hypothesis that these additional variables are insignificant. Thus we conclude that, as imposed in the dynamic mixed multinomial logit models, there are state dependencies in women’s labor supply behavior spanning two years, but not longer.

6 Conclusion

This paper has extended the literature on binary models of labor force participation dynamics by including a distinction between full-time and part-time work and by allowing more general distribution of unobserved individual characteristics. Within this setting, we have found significant autocorrelation and significant variation in the effects of education and children on labor supply behavior. We have shown that excluding either of these two features of the distribution of unobservables impacts significantly on estimates of the long-term effects of temporary policy interventions. In particular, working with a specification of the unobservables allowing time invariant individual specific random effects, but no further generality in the distribution of unobservables, results in significant downward biases in the estimated effect of a woman’s previous employment behavior on her current choice between full-time work, part-time work and non-employment. More general specifications, allowing either autocorrela-

tion in the employment state specific intercepts or variation in the effects of children and education on labor supply preferences, perform better. However, there remains a downward bias relative to when both autocorrelation and random coefficient are permitted. The biases caused by imposing overly restrictive distributions of unobservables are large enough to make the choice of distribution of unobservables important when conducting policy evaluation.

Leveraging the multinomial nature of our model, we have investigated the relative effectiveness of policies facilitating full-time and part-time work. On average, over our sample of married or cohabiting women, we have shown that part-time employment and full-time employment are equally effective routes via which non-employed women can remain in work for a number of years. This result is notable in the context of growing evidence for the United Kingdom showing that women in part-time jobs tend to be poorly paid and are more likely to be working in manual occupations than their full-time counterparts. Moreover, we have shown that for women with young children incentivizing part-time work is more effective in obtaining a lasting reduction in non-employment than incentivizing full-time work.

Appendix I: Monte Carlo Simulations I

Monte Carlo simulations are used to illustrate the poor numerical properties of the Maximum Likelihood estimator of the parameters of a dynamic mixed multinomial logit model in which there are unobserved individual characteristics that affect payoffs in only one year and have distributions containing unknown parameters. Further simulations show that reliable parameter estimates are obtained if additional structure is imposed on the unobservables.

To maintain consistency, attention is restricted to the three state model of employment dynamics described above, however similar results are obtained for static models and models with more than three alternatives. The following specification of payoffs is adopted for $t = 3, \dots, T$

$$V^f(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,f,t}) - V^n(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,n,t}) = \gamma_{f,f-1}Y_{i,f,t-1} + \gamma_{f,p-1}Y_{i,p,t-1} + \gamma_{f,f-2}Y_{i,f,t-2} + \gamma_{f,p-2}Y_{i,p,t-2} + \beta_{f,0} + \beta_{f,1}X1_{i,t} + \beta_{f,2}X2_{i,t} + \eta_{i,f,t} + \xi_{i,f,t}, \quad (11a)$$

$$V^p(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,p,t}) - V^n(\Omega_{i,t-1}, X_{i,t}, \varrho_{i,n,t}) = \gamma_{p,f-1}Y_{i,f,t-1} + \gamma_{p,p-1}Y_{i,p,t-1} + \gamma_{p,f-2}Y_{i,f,t-2} + \gamma_{p,p-2}Y_{i,p,t-2} + \beta_{p,0} + \beta_{p,1}X1_{i,t} + \beta_{p,2}X2_{i,t} + \eta_{i,p,t} + \xi_{i,p,t}. \quad (11b)$$

In the above $Y_{i,j,t}$ for $j = f, p$ are indicators of employment outcomes and $X1_{i,t}$ and $X2_{i,t}$ are individual specific variables, constructed to be mutually independent, independent over time and individuals and to have standard normal distributions. Individuals' employment outcomes at $t = 1$ and $t = 2$ are determined randomly and are constructed to be independent of subsequent employment outcomes thus allowing the initial conditions to be ignored. The unobservables $\xi_{i,f,t}$ and $\xi_{i,p,t}$ are assumed to be mutually independent, independent over time and individuals and to have type I extreme value distributions. The first component of the unobservables ($\eta_{i,f,t}, \eta_{i,p,t}$) is assumed to be formed as follows

$$\eta_{i,f,t} = \nu_{i,f} + \sum_{t=3}^T \pi_{i,f,t} I_t \quad \text{for } t = 3, \dots, T, \quad (12a)$$

$$\eta_{i,p,t} = \nu_{i,p} + \sum_{t=3}^T \pi_{i,p,t} I_t \quad \text{for } t = 3, \dots, T, \quad (12b)$$

where $(\nu_{i,f}, \nu_{i,p})' \sim N(0, \Sigma)$, I_t for $t = 3, \dots, T$ are time dummies and $(\pi_{i,f,t}, \pi_{i,p,t})$ for $t = 3, \dots, T$ are random coefficients that are independent over time and individuals with $(\pi_{i,f,t}, \pi_{i,p,t})' \sim N(0, \Xi^t)$ for $t = 3, \dots, T$. This specification of the unobservables allows the employment state specific intercepts to include time invariant individual effects and additionally, via the random coefficients on the time

dummies, allows the time-varying components of the unobservables to be correlated or heteroscedastic. When estimating this model, normalizations are imposed on $\Xi_{1,1}^t$ for $t = 3, \dots, T$. Without such normalizations, scale identification relies on the slight difference in the shapes of the logistic and normal distributions (see Ben-Akiva *et al.*, 2001). However, as explained in Section 3.3, even following these normalizations identification remains reliant on the functional form of the distribution of the unobservables. Excluding the random coefficients on the time dummies leads to a model which is nonparametrically identified provided $T \geq 4$.

Parameter	Truth	Random Coef. on Time Dummies Excluded			Random Coef. on Time Dummies Permitted		
		E(parameter)	E(σ)	σ (parameter)	E(parameter)	E(σ)	σ (parameter)
$\gamma_{f,f-2}$	1	0.99	0.14	0.14	0.96	0.17	0.15
$\gamma_{f,p-2}$	0.5	0.48	0.14	0.13	0.49	0.22	0.25
$\gamma_{f,f-1}$	2	2.02	0.15	0.15	2.12	0.20	0.22
$\gamma_{f,p-1}$	1	1.00	0.14	0.14	1.11	0.32	0.45
$\beta_{f,0}$	-1	-1.00	0.17	0.18	-1.03	0.51	0.68
$\beta_{f,1}$	-0.8	-0.80	0.09	0.09	-0.78	0.23	0.31
$\beta_{f,2}$	0.5	0.50	0.07	0.07	0.48	0.14	0.18
$\gamma_{p,f-2}$	0.5	0.51	0.12	0.11	0.34	0.63	0.56
$\gamma_{p,p-2}$	1	0.99	0.13	0.11	1.71	2.19	1.76
$\gamma_{p,f-1}$	1	1.02	0.14	0.12	0.91	0.68	0.51
$\gamma_{p,p-1}$	2	2.01	0.12	0.13	3.60	4.63	3.82
$\beta_{p,0}$	0.5	0.50	0.13	0.13	0.41	0.49	0.43
$\beta_{p,1}$	1	1.01	0.08	0.08	2.59	4.63	3.85
$\beta_{p,2}$	-0.5	-0.51	0.06	0.06	-1.39	2.58	2.18
$\Sigma_{1,1}$	1	1.01	0.40	0.39	0.97	0.57	0.56
$\Sigma_{2,1}$	0.5	0.51	0.27	0.27	0.49	1.15	0.82
$\Sigma_{2,2}$	1	1.06	0.33	0.33	11.81	63.40	40.72
$\Xi_{1,1}^3$	4 [Fixed]	-	-	-	4	-	-
$\Xi_{2,1}^3$	1	-	-	-	-0.83	7.46	7.70
$\Xi_{2,2}^3$	2	-	-	-	59.19	314.23	171.98
$\Xi_{1,1}^4$	4 [Fixed]	-	-	-	4	-	-
$\Xi_{2,1}^4$	1	-	-	-	-0.40	6.46	6.10
$\Xi_{2,2}^4$	2	-	-	-	57.53	313.81	178.98
Average Iterations			4.18			38.41	
Maximum Iterations			10			200	

Notes: E(parameter) is the mean parameter estimate, E(σ) is the mean estimated standard error and σ (parameter) is the standard deviation of the parameter estimates over the 200 Monte Carlo replications. Maximum Simulated Likelihood estimation used 5000 antithetic draws. The number of iterations is limited to 200.

Table 8: Monte Carlo simulations illustrating the properties of the Maximum Likelihood estimator of the parameters of a dynamic mixed multinomial logit model with and without random coefficients on time dummies.

Monte Carlo simulations are conducted, first excluding random coefficients on the time dummies and then allowing random coefficient on the time dummies. For each of these two Monte Carlo experiments, the sample size is fixed at 3000 individuals and $T = 4$. For each of the specifications, 200 data sets were generated and Maximum Simulated Likelihood estimates obtained for each data set. The results are summarized in Table 8. In the simulations in which random coefficients on time dummies are excluded, average parameter estimates correspond closely to their true values and average standard errors are almost identical to the standard deviation of the parameter estimates. Convergence was obtained for all of the 200 Monte Carlo replications, and took an average of 4.18 iterations starting from the true parameter values. In contrast, the Monte Carlo results for the

specification in which random coefficients on the time dummies are permitted reveal major problems. In many cases, the average coefficients on the explanatory variables differ substantially from their true values, and average standard errors bear little resemblance to the standard deviation of the parameter estimates. The estimates of the parameters of the covariance matrices reveal even greater problems: in many cases average variances are several times greater than their true values and average standard errors are huge. Furthermore, in around 10% of the Monte Carlo replications, convergence was not obtained within the first 200 iterations.

Appendix II: Monte Carlo Simulations II

Two further Monte Carlo simulations are conducted in order to establish the empirical properties of the Maximum Simulated Likelihood estimator in the context of dynamic mixed multinomial logit models in which the deterministic components of payoffs are as described by Equations (11a) and (11b) and the unobservables as in Specifications V and VI, detailed above in Section 3.4. For each of the specifications of unobservables, 200 data sets were generated each with the same sample size, attrition pattern and distribution of the initial conditions as observed in the BHPS sample. In order to explore the how the simulation bias varies with R , the number of antithetic draws to evaluate the likelihood function, all simulations are conducted using $R = 500, 2000$ and 5000 .

Tables 9-10 summarize the estimates of the deterministic components of coefficients. For Specification V, which permits random coefficients but excludes autocorrelated unobservables, there is a close correspondence between the average estimates of the deterministic components of coefficients and the true values and the average standard errors are close to the standard deviation of the parameter estimates. This is true for $R = 500$ as well as for higher values of R . However, when $R = 500$ there is evidence of biases in some of the parameters appearing in the distribution of the unobservables. In particular, some of the estimates of the variances of the random coefficients appear to be biased downwards. These biases are substantially reduced when R is increased to 2000 and all but eliminated by using $R = 5000$. The result for Specification VI, which features autocorrelated unobservables in addition to random coefficients, show that there are small biases, up to 6% of the true parameter value, in the deterministic components of coefficients appearing in the payoffs when $R = 5000$ is used. Similarly, with $R = 5000$, there are downwards biases in many of the variance parameters appearing in the distribution of the unobservables. For both sets of parameters, lower values of R are associated with substantially greater biases.

Tables 13 and 14 show, for Specifications V and VI respectively, the average dynamic responses to policy interventions evaluated at the estimated parameter values and at the true parameter values. For Specification V, which excludes autocorrelated unobservables, the average estimated policy responses obtained using 500 antithetic draws are never more than 0.4 of a percentage point away from the true policy responses. Therefore moderately large biases in the parameter estimates translate into very small biases in the estimated policy responses. Increasing the number of antithetic draws to 2000 tends to reduce the difference between the estimated and true policy responses, while an increase to 5000 leads to a further, albeit small, decrease in the difference between the estimated and true policy responses.

The Monte Carlo simulations for Specification VI, which includes autocorrelated unobservables, show that relying on only 500 antithetic draws when evaluating the likelihood functions leads to average estimated policy responses that diverge by up to 2.2 percentage points from the true policy responses. For example, a policy that temporarily moves non-employed individuals into full-time work decreases

the rate of non-employment by 11.65 percentage points one year after the policy while the estimated effect is 13.84 percentage points. Increasing the number of antithetic draws to 2000 approximately halves the magnitude of the difference between the estimated and simulated policy responses and an increase to 5000 antithetic draws leads to a further reduction in the bias in the estimated policy responses. However, even using 5000 antithetic draws, which would generally be considered a large number of draws, there are some biases in the estimated policy responses. However, using $R = 5000$, the maximum bias in the estimated policy responses is only 0.6 of a percentage point, and in relative terms the biases are around 3-6% of the corresponding true quantity.

VARIABLE	TRUTH		$R = 500$		$R = 2000$		$R = 5000$	
	f	p	f	p	f	p	f	p
$Y_{i,f,t-2}$	1.00	0.50	0.98 (0.11)[0.15]	0.47 (0.12)[0.10]	0.97 (0.12)[0.14]	0.47 (0.12)[0.13]	0.98 (0.13)[0.12]	0.48 (0.12)[0.12]
$Y_{i,p,t-2}$	0.50	1.00	0.46 (0.10)[0.12]	0.95 (0.14)[0.09]	0.49 (0.11)[0.11]	0.98 (0.11)[0.10]	0.49 (0.12)[0.11]	0.99 (0.11)[0.10]
$Y_{i,f,t-1}$	2.00	1.00	1.94 (0.12)[0.14]	0.94 (0.13)[0.11]	1.97 (0.13)[0.15]	0.98 (0.13)[0.14]	2.00 (0.14)[0.15]	1.00 (0.13)[0.15]
$Y_{i,p,t-1}$	1.00	2.00	0.97 (0.10)[0.16]	1.93 (0.15)[0.10]	0.99 (0.12)[0.13]	1.98 (0.11)[0.12]	1.00 (0.12)[0.13]	1.99 (0.12)[0.13]
$X1_{i,t}$	-0.80	1.00	-0.78 (0.04)[0.05]	0.97 (0.04)[0.04]	-0.79 (0.05)[0.05]	1.00 (0.05)[0.05]	-0.79 (0.05)[0.05]	1.01 (0.05)[0.05]
$X2_{i,t}$	0.50	-0.50	0.48 (0.04)[0.04]	-0.49 (0.04)[0.04]	0.50 (0.04)[0.04]	-0.50 (0.04)[0.04]	0.50 (0.04)[0.04]	-0.50 (0.04)[0.04]
INTERCEPT	-1.00	0.50	-1.01 (0.14)[0.14]	0.48 (0.11)[0.12]	-1.01 (0.15)[0.16]	0.49 (0.13)[0.14]	-1.00 (0.16)[0.15]	0.50 (0.13)[0.12]

Notes: Average standard errors are given in round brackets and the standard deviation of the parameter estimates is given in square brackets. Estimates of the parameters on the initial conditions are omitted. Columns headed f contain the coefficient describing payoffs from full-time employment and columns headed p contain the coefficients describing payoffs from part-time employment. Results are based on 200 Monte Carlo replications.

Table 9: Results of Monte Carlo simulations for Specification V: Average estimates of the deterministic components of coefficients.

VARIABLE	TRUTH		$R = 500$		$R = 2000$		$R = 5000$	
	f	p	f	p	f	p	f	p
$Y_{i,f,t-2}$	1.00	0.50	0.96 (0.10)[0.12]	0.48 (0.10)[0.12]	0.95 (0.12)[0.13]	0.48 (0.11)[0.12]	0.98 (0.12)[0.15]	0.49 (0.12)[0.14]
$Y_{i,p,t-2}$	0.50	1.00	0.44 (0.09)[0.11]	0.92 (0.09)[0.11]	0.47 (0.11)[0.11]	0.96 (0.10)[0.10]	0.48 (0.11)[0.13]	0.97 (0.11)[0.12]
$Y_{i,f,t-1}$	2.00	1.00	1.98 (0.11)[0.13]	1.01 (0.10)[0.12]	1.97 (0.13)[0.14]	1.00 (0.12)[0.12]	2.00 (0.14)[0.15]	1.00 (0.13)[0.14]
$Y_{i,p,t-1}$	1.00	2.00	0.99 (0.10)[0.12]	1.96 (0.09)[0.12]	1.00 (0.11)[0.13]	1.97 (0.11)[0.13]	1.00 (0.12)[0.13]	1.98 (0.12)[0.13]
$X1_{i,t}$	-0.80	1.00	-0.70 (0.04)[0.05]	0.94 (0.04)[0.05]	-0.75 (0.05)[0.07]	0.96 (0.05)[0.05]	-0.76 (0.06)[0.07]	0.98 (0.05)[0.06]
$X2_{i,t}$	0.50	-0.50	0.43 (0.04)[0.04]	-0.47 (0.04)[0.04]	0.46 (0.04)[0.05]	-0.48 (0.04)[0.04]	0.48 (0.04)[0.05]	-0.49 (0.04)[0.04]
INTERCEPT	-1.00	0.50	-1.02 (0.14)[0.15]	0.26 (0.13)[0.15]	-1.00 (0.16)[0.19]	0.37 (0.14)[0.15]	-0.99 (0.18)[0.19]	0.44 (0.15)[0.16]

Notes: See Table 9.

Table 10: Results of Monte Carlo simulations for Specification VI: Average estimates of the deterministic components of coefficients.

	TRUTH	$R = 500$	$R = 2000$	$R = 5000$
$\Sigma_{Intercept\ 1}$	$\begin{pmatrix} 1 & . \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.95 & . \\ (0.19)[0.44] & . \\ 0.45 & 0.92 \\ (0.13)[0.29] & (0.17)[0.33] \end{pmatrix}$	$\begin{pmatrix} 0.99 & . \\ (0.26)[0.38] & . \\ 0.49 & 0.99 \\ (0.19)[0.27] & (0.23)[0.33] \end{pmatrix}$	$\begin{pmatrix} 0.99 & . \\ (0.30)[0.40] & . \\ 0.50 & 1.00 \\ (0.22)[0.29] & (0.27)[0.31] \end{pmatrix}$
$\Sigma_{Y_{i,f,t-2}}$	$\begin{pmatrix} 1 & . \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.72 & . \\ (0.26)[0.42] & . \\ 0.28 & 0.72 \\ (0.19)[0.32] & (0.24)[0.40] \end{pmatrix}$	$\begin{pmatrix} 0.83 & . \\ (0.32)[0.38] & . \\ 0.35 & 0.85 \\ (0.26)[0.33] & (0.32)[0.42] \end{pmatrix}$	$\begin{pmatrix} 0.95 & . \\ (0.36)[0.39] & . \\ 0.44 & 0.92 \\ (0.30)[0.33] & (0.37)[0.40] \end{pmatrix}$
$\Sigma_{Y_{i,p,t-2}}$	$\begin{pmatrix} 1 & . \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.66 & . \\ (0.23)[0.43] & . \\ 0.28 & 0.78 \\ (0.17)[0.28] & (0.20)[0.29] \end{pmatrix}$	$\begin{pmatrix} 0.87 & . \\ (0.32)[0.33] & . \\ 0.42 & 0.92 \\ (0.23)[0.26] & (0.25)[0.29] \end{pmatrix}$	$\begin{pmatrix} 0.98 & . \\ (0.35)[0.38] & . \\ 0.48 & 0.97 \\ (0.25)[0.26] & (0.27)[0.28] \end{pmatrix}$
$\Sigma_{Y_{i,f,t-1}}$	$\begin{pmatrix} 1 & . \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.69 & . \\ (0.25)[0.46] & . \\ 0.28 & 0.72 \\ (0.20)[0.35] & (0.25)[0.40] \end{pmatrix}$	$\begin{pmatrix} 0.91 & . \\ (0.34)[0.49] & . \\ 0.42 & 0.91 \\ (0.28)[0.38] & (0.35)[0.43] \end{pmatrix}$	$\begin{pmatrix} 0.98 & . \\ (0.39)[0.41] & . \\ 0.50 & 1.00 \\ (0.33)[0.39] & (0.40)[0.47] \end{pmatrix}$
$\Sigma_{Y_{i,p,t-1}}$	$\begin{pmatrix} 1 & . \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.68 & . \\ (0.24)[0.45] & . \\ 0.30 & 0.79 \\ (0.17)[0.33] & (0.21)[0.32] \end{pmatrix}$	$\begin{pmatrix} 0.89 & . \\ (0.32)[0.39] & . \\ 0.42 & 0.93 \\ (0.24)[0.27] & (0.26)[0.32] \end{pmatrix}$	$\begin{pmatrix} 0.96 & . \\ (0.36)[0.37] & . \\ 0.49 & 1.00 \\ (0.26)[0.28] & (0.28)[0.29] \end{pmatrix}$
$\Sigma_{X1_{i,t}}$	$\begin{pmatrix} 1 & . \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.89 & . \\ (0.12)[0.13] & . \\ 0.42 & 0.88 \\ (0.08)[0.09] & (0.10)[0.12] \end{pmatrix}$	$\begin{pmatrix} 0.95 & . \\ (0.13)[0.13] & . \\ 0.47 & 0.97 \\ (0.09)[0.09] & (0.11)[0.12] \end{pmatrix}$	$\begin{pmatrix} 0.96 & . \\ (0.13)[0.14] & . \\ 0.47 & 0.99 \\ (0.09)[0.10] & (0.12)[0.12] \end{pmatrix}$
$\Sigma_{X2_{i,t}}$	$\begin{pmatrix} 1 & . \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.88 & . \\ (0.11)[0.14] & . \\ 0.42 & 0.89 \\ (0.08)[0.08] & (0.09)[0.10] \end{pmatrix}$	$\begin{pmatrix} 0.97 & . \\ (0.13)[0.13] & . \\ 0.47 & 0.97 \\ (0.09)[0.09] & (0.11)[0.11] \end{pmatrix}$	$\begin{pmatrix} 0.99 & . \\ (0.13)[0.13] & . \\ 0.49 & 0.99 \\ (0.09)[0.08] & (0.11)[0.11] \end{pmatrix}$

Notes: Average standard errors are given in round brackets and the standard deviation of the parameter estimates is given in square brackets. Results are based on 200 Monte Carlo replications.

Table 11: Results of Monte Carlo simulations for Specification V: Average estimates of parameters in the distribution of unobservables.

	TRUTH	$R = 500$	$R = 2000$	$R = 5000$
$\Sigma_{Intercept\ 1}$	$\begin{pmatrix} 1 & . \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.79 & . \\ (0.17)[0.49] & . \\ 0.37 & 0.68 \\ (0.13)[0.32] & (0.16)[0.41] \end{pmatrix}$	$\begin{pmatrix} 0.79 & . \\ (0.24)[0.53] & . \\ 0.39 & 0.71 \\ (0.18)[0.35] & (0.24)[0.47] \end{pmatrix}$	$\begin{pmatrix} 0.82 & . \\ (0.31)[0.53] & . \\ 0.41 & 0.76 \\ (0.23)[0.35] & (0.31)[0.49] \end{pmatrix}$
$\Sigma_{Y_{i,f,t-2}}$	$\begin{pmatrix} 1 & . \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.53 & . \\ (0.21)[0.38] & . \\ 0.24 & 0.68 \\ (0.16)[0.32] & (0.23)[0.40] \end{pmatrix}$	$\begin{pmatrix} 0.79 & . \\ (0.31)[0.36] & . \\ 0.35 & 0.80 \\ (0.25)[0.31] & (0.31)[0.42] \end{pmatrix}$	$\begin{pmatrix} 0.89 & . \\ (0.35)[0.43] & . \\ 0.45 & 0.94 \\ (0.30)[0.38] & (0.37)[0.47] \end{pmatrix}$
$\Sigma_{Y_{i,p,t-2}}$	$\begin{pmatrix} 1 & . \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.53 & . \\ (0.20)[0.37] & . \\ 0.23 & 0.65 \\ (0.15)[0.24] & (0.18)[0.28] \end{pmatrix}$	$\begin{pmatrix} 0.77 & . \\ (0.30)[0.44] & . \\ 0.37 & 0.85 \\ (0.22)[0.30] & (0.25)[0.28] \end{pmatrix}$	$\begin{pmatrix} 0.43 & . \\ (0.25)[0.29] & . \\ 0.91 & 0.85 \\ (0.28)[0.31] & (0.37)[0.43] \end{pmatrix}$
$\Sigma_{Y_{i,f,t-1}}$	$\begin{pmatrix} 1 & . \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.48 & . \\ (0.21)[0.35] & . \\ 0.17 & 0.54 \\ (0.17)[0.28] & (0.23)[0.37] \end{pmatrix}$	$\begin{pmatrix} 0.71 & . \\ (0.30)[0.43] & . \\ 0.29 & 0.75 \\ (0.25)[0.37] & (0.32)[0.48] \end{pmatrix}$	$\begin{pmatrix} 0.83 & . \\ (0.34)[0.43] & . \\ 0.41 & 0.90 \\ (0.25)[0.33] & (0.29)[0.34] \end{pmatrix}$
$\Sigma_{Y_{i,p,t-1}}$	$\begin{pmatrix} 1 & . \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.43 & . \\ (0.19)[0.36] & . \\ 0.17 & 0.64 \\ (0.14)[0.27] & (0.19)[0.30] \end{pmatrix}$	$\begin{pmatrix} 0.73 & . \\ (0.30)[0.43] & . \\ 0.37 & 0.84 \\ (0.22)[0.32] & (0.26)[0.35] \end{pmatrix}$	$\begin{pmatrix} 0.90 & . \\ (0.14)[0.15] & . \\ 0.45 & 0.94 \\ (0.09)[0.10] & (0.12)[0.13] \end{pmatrix}$
$\Sigma_{x_{i,1,t}}$	$\begin{pmatrix} 1 & . \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.71 & . \\ (0.10)[0.12] & . \\ 0.36 & 0.80 \\ (0.07)[0.08] & (0.09)[0.11] \end{pmatrix}$	$\begin{pmatrix} 0.84 & . \\ (0.13)[0.12] & . \\ 0.42 & 0.88 \\ (0.08)[0.09] & (0.11)[0.12] \end{pmatrix}$	$\begin{pmatrix} 0.91 & . \\ (0.13)[0.14] & . \\ 0.45 & 0.94 \\ (0.09)[0.10] & (0.12)[0.13] \end{pmatrix}$
$\Sigma_{x_{i,2,t}}$	$\begin{pmatrix} 1 & . \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 0.71 & . \\ (0.10)[0.10] & . \\ 0.35 & 0.77 \\ (0.07)[0.08] & (0.09)[0.10] \end{pmatrix}$	$\begin{pmatrix} 0.85 & . \\ (0.12)[0.12] & . \\ 0.42 & 0.88 \\ (0.08)[0.08] & (0.11)[0.09] \end{pmatrix}$	$\begin{pmatrix} 0.99 & . \\ (0.62)[0.75] & . \\ 0.45 & 0.98 \\ (0.50)[0.54] & (0.63)[0.64] \end{pmatrix}$
ρ_f	0.7	0.80 (0.08)[0.23]	0.75 (0.08)[0.21]	0.74 (0.09)[0.13]
ρ_p	0.8	0.90 (0.03)[0.07]	0.87 (0.04)[0.08]	0.83 (0.05)[0.09]
$\Sigma_{\zeta 1,1}$	2	0.67 (0.10)[0.26]	1.30 (0.36)[0.74]	1.56 (0.48)[0.68]
$\Sigma_{\zeta 2,1}$	0.7	0.65 (0.18)[0.50]	0.67 (0.10)[0.22]	0.68 (0.11)[0.19]
$\Sigma_{\zeta 2,2}$	2	1.10 (0.22)[0.55]	1.63 (0.36)[0.60]	1.87 (0.45)[0.61]

Notes: See Table 11.

Table 12: Results of Monte Carlo simulations for Specification VI: Average estimates of parameters in the distribution of unobservables.

EMPLOYMENT STATE	YEARS SINCE POLICY INTERVENTION										
	1	2	3	4	5	6	7	8	9	10	11
True dynamic responses											
Non-employed moved into full-time work at $t = 2$											
Full-time	14.93	10.87	3.94	3.36	1.77	1.42	0.88	0.71	0.49	0.41	0.32
Part-time	-3.41	-4.60	-2.17	-2.32	-1.34	-1.14	-0.73	-0.61	-0.42	-0.36	-0.27
Non-employment	-11.51	-6.27	-1.77	-1.04	-0.43	-0.28	-0.15	-0.10	-0.06	-0.05	-0.05
Non-employed moved into part-time work at $t = 2$											
Full-time	-3.56	-3.41	-1.47	-1.66	-0.90	-0.77	-0.45	-0.39	-0.29	-0.23	-0.18
Part-time	16.43	10.84	3.67	3.12	1.53	1.17	0.68	0.54	0.38	0.31	0.23
Non-employment	-12.87	-7.43	-2.20	-1.46	-0.63	-0.41	-0.23	-0.15	-0.09	-0.08	-0.05
Estimated Dynamic responses $R = 500$											
Non-employed moved into full-time work at $t = 2$											
Full-time	15.31	11.08	4.05	3.29	1.74	1.35	0.84	0.66	0.47	0.38	0.29
Part-time	-3.49	-4.56	-2.11	-2.24	-1.27	-1.10	-0.70	-0.58	-0.41	-0.35	-0.27
Non-employment	-11.82	-6.52	-1.94	-1.05	-0.46	-0.25	-0.14	-0.08	-0.05	-0.04	-0.02
Non-employed moved into part-time work at $t = 2$											
Full-time	-3.28	-3.25	-1.36	-1.56	-0.83	-0.72	-0.43	-0.37	-0.27	-0.22	-0.16
Part-time	16.46	10.88	3.72	3.02	1.49	1.14	0.67	0.53	0.36	0.30	0.20
Non-employment	-13.18	-7.63	-2.36	-1.46	-0.66	-0.42	-0.23	-0.16	-0.09	-0.07	-0.04
Estimated Dynamic responses $R = 2000$											
Non-employed moved into full-time work at $t = 2$											
Full-time	14.98	10.90	3.97	3.36	1.77	1.41	0.87	0.70	0.48	0.40	0.29
Part-time	-3.37	-4.59	-2.14	-2.34	-1.33	-1.14	-0.73	-0.62	-0.43	-0.37	-0.27
Non-employment	-11.61	-6.31	-1.83	-1.01	-0.45	-0.27	-0.14	-0.08	-0.05	-0.03	-0.02
Non-employed moved into part-time work at $t = 2$											
Full-time	-3.47	-3.31	-1.41	-1.59	-0.85	-0.72	-0.44	-0.37	-0.28	-0.23	-0.17
Part-time	16.47	10.80	3.68	3.03	1.49	1.14	0.66	0.53	0.36	0.30	0.21
Non-employment	-13.00	-7.49	-2.26	-1.44	-0.63	-0.42	-0.22	-0.15	-0.09	-0.06	-0.04
Estimated Dynamic responses $R = 5000$											
Non-employed moved into full-time work at $t = 2$											
Full-time	14.98	10.89	3.94	3.34	1.78	1.42	0.88	0.71	0.49	0.42	0.32
Part-time	-3.41	-4.61	-2.13	-2.30	-1.33	-1.13	-0.73	-0.61	-0.43	-0.37	-0.28
Non-employment	-11.57	-6.28	-1.81	-1.03	-0.45	-0.28	-0.15	-0.10	-0.06	-0.05	-0.04
Non-employed moved into part-time work at $t = 2$											
Full-time	-3.52	-3.32	-1.42	-1.61	-0.86	-0.75	-0.44	-0.39	-0.29	-0.23	-0.18
Part-time	16.41	10.78	3.66	3.07	1.50	1.16	0.68	0.54	0.38	0.31	0.24
Non-employment	-12.90	-7.46	-2.24	-1.46	-0.64	-0.41	-0.23	-0.15	-0.09	-0.08	-0.05

Notes: Based on 200 Monte Carlo replications. All figures are percentage point changes for women affected by the policy.

Table 13: Average policy responses for Specification V using $R=500$, 2000 and 5000.

EMPLOYMENT STATE	YEARS SINCE POLICY INTERVENTION										
	1	2	3	4	5	6	7	8	9	10	11
True dynamic responses											
Non-employed moved to full-time work at $t = 2$											
Full-time	14.40	10.25	3.45	2.96	1.48	1.23	0.71	0.55	0.38	0.31	0.25
Part-time	-2.75	-3.59	-1.55	-1.79	-1.03	-0.96	-0.64	-0.48	-0.36	-0.28	-0.21
Non-employment	-11.65	-6.66	-1.92	-1.18	-0.45	-0.27	-0.07	-0.07	-0.02	-0.03	-0.04
Non-employed moved to part-time work at $t = 2$											
Full-time	-3.10	-2.07	-1.35	-1.40	-0.75	-0.62	-0.41	-0.33	-0.23	-0.16	-0.12
Part-time	16.00	10.98	3.76	3.09	1.47	1.04	0.63	0.45	0.28	0.23	0.18
Non-employment	-12.90	-7.91	-2.41	-1.69	-0.72	-0.43	-0.21	-0.12	-0.05	-0.07	-0.06
Estimated Dynamic responses $R = 500$											
Non-employed moved to full-time work at $t = 2$											
Full-time	16.16	11.34	4.21	3.16	1.65	1.23	0.77	0.59	0.40	0.30	0.23
Part-time	-2.32	-3.34	-1.53	-1.71	-0.96	-0.84	-0.54	-0.45	-0.32	-0.24	-0.18
Non-employment	-13.84	-8.00	-2.68	-1.44	-0.69	-0.39	-0.23	-0.14	-0.08	-0.06	-0.05
Non-employed moved to part-time work at $t = 2$											
Full-time	-2.88	-2.91	-1.22	-1.36	-0.72	-0.63	-0.37	-0.31	-0.22	-0.18	-0.13
Part-time	17.66	11.66	4.18	3.17	1.60	1.16	0.68	0.51	0.35	0.28	0.19
Non-employment	-14.78	-8.75	-2.97	-1.81	-0.88	-0.53	-0.32	-0.20	-0.13	-0.09	-0.06
Estimated Dynamic responses $R = 2000$											
Non-employed moved to full-time work at $t = 2$											
Full-time	15.16	10.57	3.84	3.06	1.60	1.21	0.73	0.57	0.37	0.32	0.23
Part-time	-2.53	-3.36	-1.53	-1.77	-1.00	-0.87	-0.55	-0.47	-0.31	-0.27	-0.19
Non-employment	-12.62	-7.21	-2.31	-1.29	-0.60	-0.35	-0.19	-0.11	-0.07	-0.05	-0.04
Non-employed moved to part-time work at $t = 2$											
Full-time	-3.11	-2.92	-1.22	-1.40	-0.76	-0.64	-0.39	-0.32	-0.22	-0.18	-0.14
Part-time	16.88	11.26	3.96	3.13	1.55	1.14	0.67	0.49	0.34	0.26	0.19
Non-employment	-13.77	-8.34	-2.73	-1.73	-0.79	-0.50	-0.28	-0.18	-0.11	-0.08	-0.05
Estimated Dynamic responses $R = 5000$											
Non-employed moved to full-time work at $t = 2$											
Full-time	14.89	10.57	3.77	3.06	1.56	1.20	0.73	0.56	0.38	0.31	0.22
Part-time	-2.65	-3.55	-1.57	-1.79	-0.99	-0.86	-0.55	-0.46	-0.31	-0.26	-0.19
Non-employment	-12.24	-7.02	-2.20	-1.27	-0.57	-0.34	-0.18	-0.10	-0.07	-0.05	-0.03
Non-employed moved to part-time work at $t = 2$											
Full-time	-3.08	-2.88	-1.23	-1.39	-0.74	-0.64	-0.39	-0.32	-0.23	-0.18	-0.14
Part-time	16.48	11.01	3.84	3.08	1.51	1.13	0.65	0.49	0.34	0.26	0.19
Non-employment	-13.40	-8.14	-2.61	-1.69	-0.77	-0.49	-0.27	-0.17	-0.11	-0.07	-0.05

Notes: Based on 200 Monte Carlo replications. All figures are percentage point changes for women affected by the policy.

Table 14: Average policy responses for Specification VI using $R=500$, 2000 and 5000.

References

- AKAIKE, H. (1973), Information Theory as an Extension of the Maximum Likelihood Principle, in B. Petrov and F. Csaki (Editors), *Second International Symposium on Information Theory*, pp. 267–281, Akademiai Kiado, Budapest.
- ALTUG, S. AND MILLER, R. (1998), The Effect of Work Experience on Female Wages and Labour Supply, *Review of Economic Studies*, **vol. 65**(1): pp. 45–85.
- ANDERSON, T. W. AND HSIAO, C. (1982), Formulation and Estimation of Dynamic Models using Panel Data, *Journal of Econometrics*, **vol. 18**(1): pp. 47–82.
- ARELLANO, M. AND BOND, S. (1991), Some Tests of Specification for Panel Data, *Review of Economic Studies*, **vol. 58**(2): pp. 277–297.
- BEN-AKIVA, M., BOLDUC, D. AND WALKER, J. (2001), Specification, Identification and Estimation of the Logit Kernel (or Continuous Mixed Logit Model), *Department of Civil Engineering Manuscript, MIT*.
- BEN-AKIVA, M. AND LERMAN, S. R. (1985), *Discrete Choice Analysis*, Cambridge, Massachusetts: MIT Press.
- BHAT, C. R. (2001), Quasi-random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model, *Transportation Research B: Methodological*, **vol. 35**(7): pp. 677–693.
- BOOTH, A., JENKINS, S. AND SERRANO, C. (1999), New Men and New Women? A Comparison of Paid Work Propensities from a Panel Data Perspective, *Oxford Bulletin of Economics and Statistics*, **vol. 61**(2): pp. 167–197.
- BOVER, O. (1991), Relaxing Intertemporal Separability: A Rational Habits Model of Labor Supply Estimated from Panel Data, *Journal of Labor Economics*, **vol. 9**(1): pp. 85–100.
- BOWSER, C. G. (2002), On Testing Overidentifying Restrictions in Dynamic Panel Data Models, *Economics Letters*, **vol. 77**(2): pp. 211–220.
- BUNCH, D. S. (1991), Estimability in the Multinomial Probit Model, *Transportation Research B: Methodological*, **vol. 25**(1): pp. 1–12.
- CONNOLLY, S. AND GREGORY, M. (2008), Moving Down: Women’s Part-Time Work and Occupational Change in Britain 1991-2001, *Economic Journal*, **vol. 118**: pp. F52–F76.
- ECKSTEIN, Z. AND WOLPIN, K. (1989), Dynamic Labour Force Participation of Married Women and Endogenous Wage Growth, *Review of Economic Studies*, **vol. 56**(3): pp. 375–390.
- FRANCESCONI, M. (2002), A Joint Dynamic Model of Fertility and Work of Married Women, *Journal of Labor Economics*, **vol. 20**(2): pp. 336–380.
- GEWEKE, J. F. (1996), Monte Carlo Simulation and Numerical Integration, in K. D. Amman, H. and J. Rust (Editors), *Handbook of Computational Economics*, vol. 13, pp. 731–800, Amsterdam: North Holland.

- HAIJIVASSILIOU, V. (1999), Some Practical Issues in Maximum Simulated Maximum Likelihood, in R. Mariano, T. Schuermann and M. Weeks (Editors), *Simulation-Based Inference in Econometrics: Methods and Applications*, chap. 3, pp. 71–99, Cambridge University Press.
- HAIJIVASSILIOU, V. AND RUDD, P. A. (1994), Classical Estimation Methods for LDV Models Using Simulation, in C. Engle and D. McFadden (Editors), *Handbook of Econometrics*, pp. 2383–2441, Amsterdam: North Holland.
- HARRIS, K. M. AND KEANE, M. P. (1998), A Model of Health Plan Choice: Inferring Preferences and Perceptions from a Combination of Revealed Preference and Attitudinal Data, *Journal of Econometrics*, **vol. 89**(1-2): pp. 131–157.
- HAUSMAN, J. A. AND WISE, D. A. (1978), A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Preferences, *Econometrica*, **vol. 46**(2): pp. 403–26.
- HECKMAN, J. (1981a), Heterogeneity and State Dependence, in S. Rosen (Editor), *Studies in Labor Markets*, pp. 91–139, Chicago, IL: Chicago Press.
- HECKMAN, J. (1981b), The Incidental Parameter Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process, in C. Manski and D. McFadden (Editors), *Structural Analysis of Discrete Data with Econometric Applications*, pp. 179–195, Cambridge, MA: MIT Press.
- HECKMAN, J. (1981c), Statistical Models for Discrete Panel Data, in C. Manski and D. McFadden (Editors), *Structural Analysis of Discrete Data with Econometric Applications*, pp. 114–178, Cambridge, MA: MIT Press.
- HECKMAN, J. AND MACURDY, T. (1980), A Life-cycle Model of Female Labor Supply, *Review of Economic Studies*, **vol. 47**(1): pp. 47–74.
- HECKMAN, J. J. AND WILLIS, R. (1977), A Beta-logistic Model for the Analysis of Sequential Labor Force Participation by Married Women, *Journal of Political Economy*, **vol. 85**(1): pp. 27–58.
- HOTZ, V., KYDLAND, F. AND SEDLACEK, G. (1988), Intertemporal Preferences and Labor Supply, *Econometrica*, **vol. 56**(2): pp. 335–360.
- HYSLOP, D. (1999), State Dependence, Serial Correlation and Heterogeneity in Intertemporal Labor Force Participation of Married Women, *Econometrica*, **vol. 67**(6): pp. 1255–1294.
- IMAI, S. AND KEANE, M. P. (2004), Intertemporal Labor Supply and Human Capital Accumulation, *International Economic Review*, **vol. 45**(2): pp. 601–641.
- KEANE, M. (1993), Simulation Estimation for Panel Data Models with Limited Dependent Variables, in G. Maddala, C. Rao and H. Vinod (Editors), *Handbook of Statistics*, Amsterdam: North Holland.
- KEANE, M. AND SAUER, R. (–), A Computationally Practical Simulation Estimation Algorithm for Dynamic Panel Data Models with Unobserved Endogenous State Variables, *Forthcoming in International Economic Review*.
- KEANE, M. P. (1992), A Note on Identification in the Multinomial Probit Model, *Journal of Business and Economic Statistics*, **vol. 10**(2): pp. 193–200.

- KEANE, M. P. (1997), Modeling Heterogeneity and State Dependence in Consumer Choice Behavior, *Journal of Business & Economic Statistics*, **vol. 15**(3): pp. 310–27.
- KUBIN, I. AND PRINZ, A. (2002), Labour Supply with Habit Formation, *Economics Letters*, **vol. 75**(1): pp. 75–79.
- MANNING, A. AND PETRONGOLO, B. (2008), The Part-Time Pay Penalty for Women in Britain, *Economic Journal*, **vol. 118**(526): pp. F28–F51.
- SCHWARTZ, G. (1978), Estimating the Dimension of a Model, *Annals of Statistics*, **vol. 6**: pp. 461–464.
- VAN SOEST, A. (1995), Structural Models of Family Labor Supply: A Discrete Choice Approach, *Journal of Human Resources*, **vol. 30**(1): pp. 63–88.
- WOITTEZ, I. AND KAPTEYN, A. (1998), Social Interactions and Habit Formation in a Model of Female Labour Supply, *Journal of Public Economics*, **vol. 70**(2): pp. 185–205.
- WOLPIN, K. I. (1992), The Determinants of Black-White Differences in Early Employment Careers: Search, Layoffs, Quits, and Endogenous Wage Growth, *The Journal of Political Economy*, **vol. 100**(3): pp. 535–560.
- WOOLDRIDGE, J. (2005), Simple Solutions to the Initial Conditions Problem for Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity, *Journal of Applied Econometrics*, **vol. 20**(1): pp. 39–54.