State Dependence in a Multi-state Model of Employment Dynamics

Victoria Prowse *†

Abstract

A multinomial choice framework is used to investigate the nature of women's transitions between full-time employment, part-time employment and non-employment. The stochastic framework allows time varying and time invariant unobserved preferences, and also controls for the possible endogenity of education, fertility and non-labor income. Significant positive true state dependence is found in both full-time and part-time employment. This finding is robust to the specification of unobserved preferences. The results are used the assess the dynamic effects of three temporary wage subsidies. All three policies have substantial effects on employment behavior for up to 6 years. However, obtaining a permanent increase in employment requires sustained or repeated interventions.

Key Words: Dynamic labor supply, Heterogeneity, Multinomial choice, State dependence.
JEL Classification: C15, C35, J62.

*Address for correspondence: Nuffield College, Oxford, OX1 1NF, UK. Email: victoria.prowse@nuffield.ox.ac.uk. Telephone: +44 01865 278560.

†I would like to thank Steve Bond, Valérie Lechene, Tuomas Pekkarinen and participants at a seminar in Oxford. This work has been supported by the E.S.R.C, grant number PAT-030-2003-00229.

1 Introduction

This paper uses a multinomial choice framework to explore the nature of women's transitions between full-time employment, part-time employment and non-employment. Within the multinomial choice framework, particular care is taken to distinguish between the effects of past employment experience and persistent unobservables on current employment behavior. The results are used to investigate the dynamic effects of three temporary wage subsidies.

The literature contains several studies of dynamic labor force participation behavior (see, for example, Booth, Jenkins and Serrano 1999, Eckstein and Wolpin 1989, Heckman and Borjas 1980, Heckman and Willis 1977, Hyslop 1999, Knights, Harris and Loundes 2002, Narendranathan and Elias 1993). However studies of employment dynamics which differentiate between full-time employment and part-time employment are far less common. Exceptions include Blank (1989) and Burdett and Taylor (1994) who use competing risk duration models to study movements between several different labor market states. While these studies are informative about the nature of transitions between various employment states, in both cases, the treatment of unobserved individual specific heterogeneity is constrained by the duration framework.

Nevertheless, determining how individuals combine part-time and full-time employment over time is curtail to understanding individuals' life course employment decisions. Previous research has indicated that part-time employment plays several important roles in individuals' dynamic employment behavior, especially for women. Blank (1989) suggests that part-time employment may provide a stepping stone, facilitating the transition between non-employment and full-time employment. Alternatively, part-time employment may play a maintenance role, whereby part-time and full-time employment are interchanged to allow an individual to combine domestic responsibilities and employment (see Corral and Isusi 2004). Finally, it has been claimed that part-time employment may be exclusionary: part-time jobs are often insecure, low wage jobs, offering little opportunity for career progression. Thus, individuals who choose part-time employment may become trapped in an exclusionary cycle of low wage, part-time jobs and non-employment (see Fagan and Burchell 2002, Martin and Roberts 1984). Moreover, an understanding of the nature of individuals' transitions between full-time employment and part-time employment is central to evaluating the dynamic effects of policy interventions, such as wage subsidies, minimum wage legislation and job creation schemes.

There are several reasons to suspect that, after controlling for observed individual characteristics, there might be a dynamic structure to individuals' employment behavior. For example,

one might suspect that an individual's presence in a particular state at time t will increase the probability, conditional on the individual's observed characteristics, that they are in the same state at time $t+1$. This type of behavior has been called state dependence. Heckman and Willis (1977) discuss two possible explanations for state dependence. Firstly, state dependence will be observed if an individual's presence in a state at time t changes prices, preferences or constraints which are relevant to their future behavior. This could take the form of past employment experience increasing an individual's stock of human capital, which, in turn, increases their future wage (see Mincer and Polachek 1974). Alternatively, fixed costs related to job search can make employment more attractive if the individual is already employed than if they are currently non-employed (see, for example, Heckman and Borjas 1980, Hyslop 1999, Layard and Bean 1989). Secondly, state dependence will be observed if there is intertemporally correlated, unobserved, individual specific heterogeneity. This heterogeneity can be time varying or time invariant, or some combination thereof. Heckman and Willis term the two cases true state dependence and spurious state dependence respectively. As noted by Heckman and Borjas (1980), inadequate controls for unmeasured variables gives rise to a conditional relationship between future and past employment behavior that is due entirely to uncontrolled heterogeneity.

For the purpose of policy evaluation, it is important to correctly distinguish between true and spurious state dependence. Consider a policy intervention which has the effect of temporally moving non-employed individuals into full-time jobs. If there is positive true state dependence in full-time employment, the policy intervention will cause a persistent increase in the number of individuals who are in full-time employment. Consequently the intervention is likely to reduce the number of individuals who are dependent on benefits or living on low incomes. In contrast, if there is only spurious state dependence, the policy intervention will not have a lasting effect on employment behavior.

With the importance of correctly distinguishing between true and spurious state dependence in mind, the model is estimated allowing several different and increasingly flexible distributions of unobservables. In the most general of the specifications, autocorrelated and time invariant unobserved preferences are permitted. Furthermore, the possible endogenity of education, fertility and non-labor income is incorporated by using the procedure described in Chamberlain (1984). Tests for true state dependence in the presence of various forms of spurious state dependence are conducted.

The data used in this application are taken from waves 1-12 of the British Household Panel Survey (BHPS). Attention is focused on a sub-sample of married or cohabiting, non-retired

women aged between 16 and 65 years. The BHPS and the sample used in this application are discussed in more detail below. The model is estimated using Maximum Simulated Likelihood (MSL) estimation, with the GHK simulator (see Geweke 1991, Hajivassiliou and Rudd 1994, Keane 1994) used to evaluate the likelihood.

The results indicate that unobserved preferences contain both time invariant and autocorrelated elements. Also, there is some evidence of preference endogenity, that is unobserved heterogeneity the is correlated with observed individual characteristics. Irrespective of the assumed structure of unobserved preferences, there is significant positive true state dependence in both full-time and part-time employment.

The presence of significant positive true state dependence in employment behavior suggests that policy interventions aiming to reduce non-employment might have prolonged effects. In order to assess this possibility further, the effects of three temporary wage subsidies are simulated and compared. The first policy is a one year wage subsidy of 5%, paid regardless of hours of work. The second and third policies subsidize the wages of individuals in full-time employment and part-time employment respectively, again by 5% and for the period of one year. All three policy interventions are found to substantially reduce non-employment for up to 6 years. However, over the longer term, the effects of all three wage subsidy policies are negligible. This suggests that persistent or sustained interventions are required in order to obtain a permanent reduction in non-employment.

This paper proceeds as follows. Section 2 introduces the economic model and the econometric specification. Section 3 discusses the data, and Section 4 presents the results. Section 5 compares the effects of the three wage subsidies, and Section 6 concludes. Appendices contain a Monte Carlo study of the performance of the estimator used in this application, and variable descriptions.

2 Model

An individual's labor supply problem can be written as follows:

$$Max_j \; U_j(w_{i,j,t}, x_{i,t}, Z_{i,t-1}, \varepsilon_{i,j,t}) \qquad \text{subject to } j \in B. \tag{1}$$

In Equation (1), $U_j(w_{i,j,t}, x_{i,t}, Z_{i,t-1}, \varepsilon_{i,j,t})$ is individual i's utility if they choose alternative j at time t. $w_{i,j,t}$ is the wage the individual receives if they choose alternative j at time t. Thus the specification allows the wage to vary across employment states. This is important as wages have often been found to vary with hours of work (see Metcalf 1999, Robson, Dex, Wilkinson

and Salido Cortes 1999). $x_{i,t}$ is a k by 1 vector of observed individual characteristics at time t and $Z_{i,t-1}$ represents individual i's employment history up to and including time $t-1$. $\varepsilon_{i,j,t}$ is a scalar random variable representing the unobserved component of individual i's preference for employment state j at time t. B denotes the budget set of available alternatives. The budget set is determined by income and prices, and also by the tax and benefit system and institutional constraints, such as restrictions on hours of work.

In the current application, the budget set is assumed to comprise of three states denoted $j = n, p, f$. State n is non-employment, corresponding to zero hours of work. States p and f correspond to part-time employment and full-time employment respectively. In this model, individuals who are observed in state n are assumed to be voluntarily unemployed.[1]

Let $y_{i,j,t}$ be an indicator variable taking the value one if individual i chooses state j at time t and zero otherwise. Utility maximizing behavior implies:

$$y_{i,j,t} = \begin{cases} 1 & \text{if } U_j(w_{i,j,t}, x_{i,t}, Z_{i,t-1}, \varepsilon_{i,j,t}) > U_k(w_{i,k,t}, x_{i,t}, Z_{i,t-1}, \varepsilon_{i,k,t}) \text{ for all } k \neq j,\ k \in B \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

The model can be implemented by choosing a functional form for $U_j(w_{i,j,t}, x_{i,t}, Z_{i,t-1}, \varepsilon_{i,j,t})$. In keeping with the literature on random utility models, $U_j(w_{i,j,t}, x_{i,t}, Z_{i,t-1}, \varepsilon_{i,j,t})$ is assumed to comprise of an observed component and an unobserved component. Specifically, the utility function is defined as:

$$U_j(w_{i,j,t}, x_{i,t}, Z_{i,t-1}, \varepsilon_{i,j,t}) = \beta_{2,j} x_{i,t} + \eta_2 w_{i,j,t} + \gamma_j y_{i,t-1} + \varepsilon_{i,j,t},\ j = n, p, f \text{ and } t = 2, ..., T. \tag{3}$$

In the above, $y_{i,t-1}$ is 2 by 1 a vector of lagged employment state indictors given by $y_{i,t-1} = (y_{i,f,t-1}, y_{i,p,t-1})'$. $w_{i,j,t}$ is now taken to be the log wage. $\beta_{2,j}$ for $j = n, p, f$ are 1 by k dimensional vectors of parameters and $\gamma_j = (\gamma_{f,j}, \gamma_{p,j})'$ for $j = n, p, f$. There is positive true state dependence in full-time employment if $\gamma_{f,f} > 0$, and, similarly, there is positive true state dependence in part-time employment if $\gamma_{p,p} > 0$.[2]

According to this implementation, the only element of an individual's employment history which is relevant to their current employment behavior is their employment state in the immediately proceeding period. Thus true state dependence is assumed to be Markovian. Such a specification has been motivated by the presence of search or transition costs (see, for example, Heckman and Borjas 1980, Hyslop 1999, Layard and Bean 1989). Heckman and Borjas (1980)

[1] The omission of involuntary unemployment as a labor market state does not represent a major oversimplification in the current context: in any year less than 1% of the sampled individuals are involuntarily unemployed.

[2] Previous employment behavior can also influence current utility through the cross-state coefficients $\gamma_{f,p}$ and $\gamma_{p,f}$.

discuss several other forms of true state dependence. These include occurrence dependence, where the number of previous spells in each state affects current employment behavior. Alternatively, if the time spent in the current state affects current employment behavior then duration dependence is present. Similarly, the employment process may exhibit lagged duration dependence, where current employment behavior depends on the length of time spent in each previous employment state. In this study, attention is restricted to Markovian true state dependence as the other forms of state dependence pose additional complications when dealing with the initial conditions problem, discussed below.

Likelihood contributions take the form of the joint density of each individual's employment outcomes over the sample period. Given that the data used in this application are taken from a panel survey, for most individuals the first employment state which is observed is part way through their life-time employment period. Moreover, the first observed employment state for an individual will depend the individual's previous employment behavior, which is unobserved by the econometrician. Treating the first observed employment state as predetermined or exogenous will, in the presence of unobserved, intertemporally correlated heterogeneity, lead to inconsistent parameter estimates (see Heckman 1981*a*). Alternatively, the first observations could be treated as equilibrium values of the employment process. However, this approach is problematic in the presence of non-stationary covariates, such as age or income, which are well established determinants of employment behavior. Here, the initial conditions problem is dealt with by using the most general of the methods suggested in Heckman (1981*a*). In particular, the first period utility function is approximated as follows:

$$U_j(w_{i,j,1}, x_{i,1}, \varepsilon_{i,j,1}) = \beta_{1,j} x_{i,1} + \eta_1 w_{i,j,1} + \varepsilon_{i,j,1}, \;\; j = n, p, f, \tag{4}$$

and the unobserved element of preferences at $t = 1$, $\varepsilon_{i,j,1}$, is allowed to be correlated with future unobserved preferences.

More generally, this approach to the initial conditions problem requires the econometrician to model the relevant elements of individuals' employment histories at t=1. If, for example, the employment process exhibits duration dependence one would have to model the time spent in the initial state prior to the start of the survey. This is more challenging than modelling the initial state itself, and in many cases such information is unavailable or unreliable.

Examining Equation (2), it is clear that individuals' behavior is determined by the relative utility of the available alternatives. A normalization is required as the level of an individual's utility does not affect their behavior. For what follows, the utility of non-employment is normalized to zero for all individuals. With this normalization imposed, an individual's utility

if they choose state p or state f is their utility from choosing each of the respective states, relative to their utility if they were to choose to be non-employed. Scale normalizations must also be made. These are explained below.

Attention in now turned to the specification of the unobserved component of individuals' preferences. Define a $l \times 1$ dimensional vector z_i, where the elements of z_i correspond to the average over t of selected time varying elements from x_i. Let $\varepsilon_{i,j,t} = \pi_{i,j,t} + \epsilon_{i,j,t}$, where $\pi_{i,j,1} = \lambda_{1,j} z_i$ and $\pi_{i,j,t} = \lambda_{2,j} z_i$ for $t > 1$. Here $\lambda_{1,j}$ and $\lambda_{2,j}$ for $j = p, f$ are $1 \times l$ dimensional vectors of parameters.

Define $\epsilon_{i,t} = (\epsilon_{i,f,t}, \epsilon_{i,p,t})$, and let ϵ_i be $\epsilon_{i,t}$ stacked over t. Similarly, x_i, y_i and w_i denote $x_{i,t}$, $y_{i,t}$ and $w_{i,t}$ stacked over t. The following distributional assumption is made:

$$\epsilon_i | x_i, w_i \sim N(0, \Sigma), \tag{5}$$

where Σ is an unrestricted covariance matrix. This specification of unobserved preferences, which follows Chamberlain (1984), allows unobserved preferences to contain both time varying and time invariant elements, and, through π, allows individuals' unobserved preferences to be correlated with their observed characteristics. Thus this specification allows, for example, education, fertility and non-labor income to be endogenous.

As mentioned above, the scale of some of the parameters is not identified. Consider an individual's choice problem at $t = 1$. Multiplying the utility of each alternative at $t = 1$ by a positive constant does not change the individual's problem. Thus, the variance of one element of $\epsilon_{i,1}$ must be normalized to some positive value. The same applies at $t = 2$. Given that β_2 is assumed to be time invariant, no normalizations are necessary at subsequent time periods. Let $\widetilde{\Sigma}$ and $\widetilde{\epsilon}_i$ denote Σ and ϵ_i with these two normalizations imposed.

The importance of including alternative specific covariates, such as the wage in this model, in multinomial choice models was first noted by Keane (1992). Keane found identification in the single period multinomial probit with only individual specific covariates to be "extremely tenuous". In particular, distinguishing between the effects of the slope coefficients and parameters of the covariance matrix was found to be difficult in the absence of alternative specific covariates, despite such covariates being unnecessary for formal identification. Rendtel and Kaltenborn (2004) extend Keane's results by considering a multiperiod multinomial probit model, again without alternative specific covariates. The authors find that the multiperiod model suffers from fragile identification problems similar to those encountered in the single period model.[3]

[3]While an individual's wage any state not chosen by the individual is not observed by the econometrician, it is possible to predict alternative specific wages based on sample information. The procedure for constructing

In order to derive the likelihood, some further definitions must be made. Let x, y and w denote the vectors x_i, y_i and w_i stacked over i. Also, let θ be a vector containing all the parameters in the model. Assuming independence over i, the likelihood can be written as follows:

$$L_{y|x,w}(\theta) = \prod_{i=1}^{N} L_{y_i|x_i,w_i}(\theta). \tag{6}$$

Individual contributions to the likelihood are given by:

$$\begin{aligned} L_{y_i|x_i,w_i}(\theta) &= Prob(y_{i,1}, y_{i,2}, ..., y_{i,T}|x_i, w_i) & (7)\\ &= \int_{\widetilde{\epsilon}_i \in A_i} \phi(\widetilde{\epsilon}_i) d\widetilde{\epsilon}_i, & (8) \end{aligned}$$

where $\phi(\widetilde{\epsilon}_i)$ is the density of $\widetilde{\epsilon}_i$ and A_i is a set containing the values of $\widetilde{\epsilon}_i$ such that Equation (2) implies the observed sequence of employment behavior, y_i.

Two problems hinder maximum likelihood estimation of this model. Firstly, the model contains high dimensional integrals which are computationally demanding to evaluate. With 3 alternatives and T time periods evaluating the likelihood requires one to evaluate a $2T$ dimensional integral. Numerical approaches to this problem are infeasibly slow. However, simulation methods exist which are both fast and accurate. Here the GHK or Smooth Recursive Conditioning (SRC) simulator is used to evaluate the likelihood (see Geweke 1991, Hajivassiliou and Rudd 1994, Keane 1994).

Briefly, the GHK simulator is explained as follows. Suppose that one wishes to evaluate $P(\epsilon \leq \mu)$ where ϵ and μ are K dimensional vectors, and $\epsilon \sim N(0, \Omega)$. The parameters contained in μ and Ω are assumed to be known. Let L be a lower triangular matrix such that $LL' = \Omega$. Denote the $(k,j)^{\text{th}}$ element of L by $L_{k,j}$. $P(\epsilon \leq \mu)$ can be approximated by $\widetilde{P} = \frac{1}{R}\Phi\left(\frac{\mu_{11}}{L_{11}}\right)\sum_{r=1}^{R}\prod_{k=2}^{K}\Phi\left(\frac{\mu_k - \sum_{j=1}^{k-1} L_{kj}\epsilon_j^r}{L_{kk}}\right)$ where $\epsilon_1^r = \Phi^{-1}\left(u_1^r \Phi\left(\frac{\mu_1}{L_{11}}\right)\right)$ and $\epsilon_k^r = \Phi^{-1}\left(u_k^r \Phi\left(\frac{\mu_k - \sum_{j=1}^{k-1} L_{kj}\epsilon_j^r}{L_{kk}}\right)\right)$ for $k = 2, ..., K$ and where u_j^r for $j = 1, ..., K$ are independent standard uniform random variables.[4] Maximizing the simulated likelihood produces the Maximum Simulated Likelihood (MSL) estimator. Using the GHK simulator, the simulated likelihood is unbiased for a finite number of replications, however the log simulated likelihood is biased. Thus, for a finite number of replications, the MSL estimator is biased. However, Hajivassiliou and Rudd (1994) show that the MSL estimator is consistent if $R \longrightarrow \infty$

alternative specific wages is explained in Section 3.1.

[4] Hajivassiliou, McFadden and Ruud (1996) provide a comparison of several different methods for evaluating multivariate normal probabilities. The authors conclude that the GHK simulator is overall the most reliable method.

as $N \longrightarrow \infty$, and is asymptotically efficient and asymptotically equivalent to the Maximum Likelihood Estimator if $R/\sqrt{N} \longrightarrow \infty$ as $N \longrightarrow \infty$.

There are two alternative simulation methods that could be applied to this problem. The Method of Simulated Moments (MSM) estimator expresses the score of the likelihood as a set of moment conditions. These moment conditions are then simulated (see McFadden 1989).[5] The Method of Simulated Scores (MSS) solves for the root the the simulated scores directly (see Hajivassiliou and McFadden 1998). Unlike MSL, both of these methods yield consistent estimators for a finite number of replications, as long as an unbiased simulator of the moment conditions or the score function can be obtained. However, as discussed in Hyslop (1999) and elsewhere, MSL is simple to implement. In contrast, implementing MSM or MSS often requires substantial manipulation of the problem. Moreover, MSL is computationally robust whereas MSM can be numerically unstable (see Geweke, Keane and Runkle 1997, Hajivassiliou and Rudd 1994).

The second problem concerning maximum likelihood estimation of this model is the large number of parameters in the model, especially the large number of parameters in the covariance matrix. With 3 alternatives and T time periods the covariance matrix contains $(2T(2T+1)/2-2)$ free parameters. Without further restrictions on the nature of unobserved preferences, maximizing the likelihood is computationally intensive, and possibly prohibitive. For this reason, further restrictions are placed on structure of unobserved preferences. It is well known that mis-specification of the unobserved element of preferences in dynamic, discrete choice models leads to misleading inferences regarding the effects of lagged dependant variables, and consequently incorrect conclusions concerning the extent of true state dependence (see Heckman 1981*b*). Here, the model is estimated with several different specifications of unobserved preferences. The most general specification includes time invariant and autocorrelated unobservables and also allows preference endogentiy, thus this specification is quite flexible. The sensitivity of the results to the specification of unobservables is considered, and simulations based on the estimated models are used to determine the preferred specification of unobserved preferences.

The specification of unobservables is now considered in more detail. Σ is assumed to have a components of variance structure. Denote $var(\epsilon_{i,1}) = u$ where u is a 2 by 2 symmetric matrix with both diagonal elements equal to 1. Also, denote $cov(\epsilon_{i,1}, \epsilon_{i,t}) = c$ for $t = 2, ..., T$. Let $\epsilon_{i,t} = \xi_{i,t} + \nu_i$ for $t = 2, ..., T$, where $\xi_{i,t}$ and ν_i are 2 by 1 vectors. Here, $\xi_{i,t}$ and ν_i represent

[5] Keane (1994) introduced a computationally practical MSM estimator for discrete panel data problems such as the model in hand.

respectively the time varying and time invariant components of individuals' preferences. Denote $var(\xi_{i,t}) = v$ for $t = 2, ..., T$ and $var(\nu_i) = \mu$. v and μ are such that the diagonal elements of $v + \mu$ are equal to 1. $\xi_{i,t}$ may or may not be intertemporally correlated. Specifically, let $\xi_{i,t} = \rho\xi_{i,t-1} + e_{i,t}$, where ρ is a scalar lying in the interval [-1,1] and $e_{i,t}$ is independent over time: when $\rho = 0$ the time varying individual effects are intertemporally uncorrelated. The following models, corresponding to different specifications of unobserved preferences, are estimated:

Model 1 Time invariant unobserved preferences, uncorrelated with x_i: $\rho = 0$, $\mu \neq \mathbf{0}$ and $\lambda = \mathbf{0}$.

Model 2 Autocorrelated unobserved preferences, uncorrelated with x_i: $\rho \neq 0$, $\mu = \mathbf{0}$ and $\lambda = \mathbf{0}$.

Model 3 Time invariant and autocorrelated unobserved preferences, uncorrelated with x_i: $\rho \neq 0$, $\mu \neq \mathbf{0}$ and $\lambda = \mathbf{0}$.

Model 4 Time invariant unobserved preferences, correlated with x_i: $\rho = 0$, $\mu \neq \mathbf{0}$ and $\lambda \neq \mathbf{0}$.

Model 5 Autocorrelated unobserved preferences, correlated with x_i: $\rho \neq 0$, $\mu = \mathbf{0}$ and $\lambda \neq \mathbf{0}$.

Model 6 Time invariant and autocorrelated unobserved preferences, correlated with x_i: $\rho \neq 0$, $\mu \neq \mathbf{0}$ and $\lambda \neq \mathbf{0}$.

Appendix I contains a Monte Carlo study of the performance of the MSL estimator in this context. The results indicate that for a small number of replications the MSL estimator is substantially biased. However, for a sufficiently large number of replications, the estimator performs well.

3 Data

The data used in this application are taken from the BHPS. The BHPS commenced in 1991, surveying a representative sample of approximately 5500 households in Great Britain, containing about 10000 persons.[6] The original survey respondents, together with their co-residents have been re-interviewed annually. See Taylor, Brice, Buck and Prentice (2001) for a complete description of the BHPS.

The sample used here is a balanced panel covering the first 12 waves of the BHPS. In this study, attention is restricted to married or cohabiting, non-retired women aged between

[6]The BHPS also includes additional households surveyed for the European Community Household Panel (waves 7-11), the Scotland and Wales Extension samples (wave 9 onwards) and the Northern Ireland Household Panel Survey (wave 11 onwards). Since this study uses a balanced panel, individuals in these households are not included.

18 and 65 years. This sample contains 8784 person-wave observations. Due to attrition, the individuals in this sample will not be representative of the corresponding population. However, this sample can be used to estimate structural parameters provided that attrition, conditional on observed individual characteristics, is not related to the employment status of the individual, or in other words, if there is no selectivity problem.[7]

At each wave, all individuals are assigned to either full-time employment, part-time employment or non-employment on the basis of their reported usual weekly hours of work. Non-employment corresponds to zero usual weekly hours of work. Individuals reporting usual weekly hours of work of between zero and 30 hours are classified as part-time employed, and individuals reporting usual weekly hours of work over 30 hours are classified as full-time employed. Table 1 shows the proportion of individuals observed in each state. On average, approximately one third of individuals were in each state. Over the sample period, the proportion of individuals who were non-employed fell from 37% at wave 1 to 33% by wave 12. The proportions of individuals in full-time and part-time employment rose slightly over the sample period.

	WAVE												
STATE	ALL	1	2	3	4	5	6	7	8	9	10	11	12
n	0.34	0.37	0.36	0.35	0.34	0.33	0.33	0.33	0.34	0.31	0.32	0.32	0.33
p	0.33	0.30	0.33	0.34	0.35	0.35	0.35	0.34	0.34	0.34	0.32	0.31	0.31
f	0.33	0.34	0.31	0.31	0.31	0.32	0.32	0.33	0.32	0.35	0.36	0.37	0.36

Table 1: Proportion of individuals in each state: All waves and waves 1-12 separately.

Table 2 shows the proportion of individuals in each state according to the age of the youngest child in the household. Unsurprisingly, the presence of a child aged under 3 years in the household substantially increases the probability of non-employment and decreases the probabilities of both full-time and part-time employment. Women in households where the youngest child is aged 3-4 years are more likely to work part-time and less likely to be non-employed than women in households where the youngest child is aged under 3 years. Women in households where the youngest child is aged 5 years or over have a relatively high probability of being in employment, either full-time or part-time. Table 3 show the proportion of individuals in each employment state according to the level of qualifications. Amongst individuals with academic qualifications, individuals with qualification of A-levels or above are less likely to be non-employed and are more likely to be full-time employed than individuals with qualifications below A-levels. Individuals with vocational qualifications have similar employment patterns to individuals with academic qualifications of A-levels or above, except they are slightly more

[7] Hausman and Wise (1979) discuss the problems posed by attrition in panel data.

likely to work part-time, and are less likely to work full-time.

STATE	YOUNGEST CHILD AGED UNDER 3 YEARS	YOUNGEST CHILD AGED 3-4 YEARS	YOUNGEST CHILD AGED 5 YEARS OR OVER
n	0.55	0.49	0.31
p	0.27	0.36	0.42
f	0.18	0.16	0.27

Table 2: Proportion of individuals in each state according to the age of the youngest child in the household.

STATE	QUALIFICATIONS BELOW A-LEVELS	QUALIFICATIONS OF A-LEVELS OR ABOVE	VOCATIONAL QUALIFICATIONS
n	0.33	0.28	0.28
p	0.37	0.28	0.33
f	0.31	0.44	0.40

Table 3: Proportion of individuals in each state according to the level of qualifications.

Table 4 shows the transition matrix. As expected there is a substantial amount of state dependence in employment behavior. 87% of individuals who are non-employed at time t are non-employed at time $t+1$. Similarly, 82% of individuals who are in part-time employment at time t and 88% of individuals who are in full-time employment at time t are in the same employment state one year later. Thus, part-time employment appears to be a less absorbing state than either full-time employment or non-employment. The transition matrix also shows that individuals are more likely to move to an adjacent state than to a non-adjacent state. For example, individuals who are non-employed at time t have a 10% probability of being in part-time employment at time $t+1$ but only a 3% probability of being in full-time employment at time $t+1$.

		STATE AT TIME $t+1$		
		n	p	f
STATE AT TIME t	n	**0.87**	0.10	0.03
	p	0.08	**0.82**	0.10
	f	0.04	0.09	**0.88**

Table 4: Transition matrix.

Table shows 5 the frequencies of the different combinations of employment states. 130 individuals are non-employed at all 12 waves, and 13 and 31 individuals are part-time employed and full-time employed respectively at all 12 waves. These figures again suggest that part-time employment is a less absorbing state than either full-time employment or non-employment.

STATES OBSERVED	FREQUENCY
Only n	130
Only p	13
Only f	31
n and p	124
n and f	48
p and f	120
n, p and f	88

Table 5: Frequencies of combinations of states.

Amongst individuals observed in more than one employment state over the 12 waves, combinations of non-employment and part-time employment and part-time employment and full-time employment are more common than combinations involving both non-employment and full-time employment. This is evidence against the stepping stone pattern of employment transitions. Indeed, it appears that most instances part-time employment fall into either the exclusionary or maintenance categories.

Appendix II contains definitions and descriptive statistics of the explanatory variables used in this study.

3.1 Wage Equations

As noted above, multinomial choice models with only individual specific covariates suffer from fragile identification problems (see Keane 1992, Rendtel and Kaltenborn 2004). To avoid the problems associated with fragile identification, alternative specific wages are included in the model. However, at any wave, an individual's potential wage in any employment state which they did not choose is not observed by the econometrician. In order to obtain alternative specific wages for all individuals and all alternatives, separate wage equations are estimated for part-time wages and full-time wages. Heckman selection models are used to correct for any selectivity in observed wages. Each wage equation is estimated using the relevant log wage as the dependent variable and pooling all 12 waves of data. The regressors in each of the wage equations are an intercept, indicators of high and low academic qualifications, an indicator of vocational qualifications, age and age squared and an indicator of union membership. The selection equations contain these regressors and also the number of children in the household aged 0-2 years, 3-4 years, 5-11 years and 12-15 years, and log non-labor income.[8]

[8] When estimating the wage equations, AGE, AGE^2 and log non-labor income (LOTHERY) have been transformed to have zero mean and unit variance.

	FULL-TIME WAGES	PART-TIME WAGES
	8784 OBSERVATIONS, 5859 CENSORED	8784 OBSERVATIONS, 5868 CENSORED
VARIABLE	COEFFICIENT	COEFFICIENT
WAGE EQUATION		
EDUC1	0.12^{**} (0.02)	0.09^{**} (0.02)
EDUC2	0.34^{**} (0.02)	0.35^{**} (0.02)
VOC	0.02 (0.01)	0.03 (0.02)
AGE	0.07 (0.06)	0.08 (0.07)
AGE^2	−0.05 (0.06)	−0.08 (0.07)
UNION	0.16^{**} (0.03)	0.16^{**} (0.04)
INTERCEPT	1.11^{**} (0.04)	1.09^{**} (0.05)
SELECTION EQUATION		
EDUC1	0.34^{**} (0.05)	-0.08^{*} (0.04)
EDUC2	0.54^{**} (0.05)	-0.41^{**} (0.05)
VOC	0.13^{**} (0.04)	0.13^{**} (0.03)
AGE	0.16 (0.16)	−0.12 (0.15)
AGE^2	-0.31^{*} (0.15)	0.14 (0.15)
UNION	2.14^{**} (0.05)	1.93^{**} (0.05)
NCH02	-0.74^{**} (0.07)	-0.19^{**} (0.06)
NCH34	-0.76^{**} (0.06)	0.04 (0.05)
NCH511	-0.45^{**} (0.03)	0.18^{**} (0.02)
NCH1215	-0.20^{**} (0.03)	0.18^{**} (0.03)
LOTHERY	-0.07^{**} (0.02)	−0.01 (0.02)
INTERCEPT	-1.17^{**} (0.04)	-0.92^{**} (0.04)
ρ	0.01 (0.06)	−0.02 (0.09)
σ	0.35 (0.00)	0.39 (0.01)
Log likelihood	-4867.45	-5950.73
LR test ($\rho = 0$)	0.03	0.04

Table 6: Wage Equations: Heckman selection models for full-time and part-time wages. Standard errors in parenthesis. * indicates significance at the 0.05 level and ** indicates significance at the 0.01 level.

Table 6 shows the results of the Heckman selection models. The effects of the variables included in the wage equation are as expected, and similar for part-time and full-time wages.[9]

[9]Clearly, for the predicted wage to be an alternative specific covariate it must be that predicted wages in part-

Specifically, education and vocational qualifications increase the wage, the wage is quadratic in age, and union membership tends to increase the wage. Interestingly, for both wage equations the null hypothesis that ρ, the correlation between the error in the wage equation and the error in the selection equation, is equal to zero can not be rejected.

In order to predict full-time and part-time wages for all individuals, it is necessary to know each individual's union status in both full-time and part-time employment. When an individual's union status in a state is not observed, it is assumed to be equal to the average level of union membership amongst individuals in the state. This can be interpreted as predicting an individual's state specific wage based on the individual's expected union status if they were to choose the state, which in turn is their current union status in the state, when this is known, or otherwise the average union status of the individuals in the state.

4 Results

The results for each of the six models described in Section 2 are shown in Table 7.[10,11] The vector $x_{i,t}$ consists of an intercept, indicators of high and low academic qualifications, age and age squared, the number of children in the household aged 0-2 years, 3-4 years, 5-11 years and 12-15 years, and log non-labor income. z_i consists of the average over the 12 waves of the indicators of high and low academic qualifications, the numbers of children in the household aged 0-2 years, 3-4 years, 5-11 years and 12-15 years, and log non-labor income.[12,13]

All six models show significant positive true state dependence in both full-time and part-time employment. The results also indicate a higher level of true state dependence in full-time employment than in part-time employment. In the models without correlated preferences, ρ, the parameter governing the nature of the autocorrelated element of individuals' unobserved preferences, is significantly positive in Model 2, where time invariant unobserved preferences are absent, and significantly negative in Model 3, where time invariant unobserved preferences are

time employment and in full-time employment differ for at least some individuals. The estimation results in Table 6 show some small differences in the coefficients for part-time and full-time wages. Additional differences in predicted wages occur as an individual's union status will differ between full-time and part-time employment.

[10] All numerical calculations were preformed using MATLAB.

[11] The likelihood was evaluated using 60 replications of the GHK simulator.

[12] The sample means are denoted $\overline{\text{NCH02}}$, $\overline{\text{NCH03}}$, $\overline{\text{NCH511}}$, $\overline{\text{NCH1215}}$, $\overline{\text{LOTHERY}}$, $\overline{\text{EDUC1}}$ and $\overline{\text{EDUC2}}$.

[13] When estimating the model, AGE, AGE^2 and LOTHERY have been scaled to have zero mean and unit variance. Predicted log full-time wages and predicted log part-time wages have been adjusted by subtracting the mean of predicted part-time wages and dividing by the standard deviation of predicted part-time wages. These normalizations improve the numerical performance of the MSL estimator.

present. This suggests that in Model 2 ρ is proxying for the absence of time invariant unobserved preferences. In contrast, in the models with correlated preferences, ρ is significantly negative in both specifications which allow autocorrelated preferences.[14]

VARIABLE	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5	MODEL 6
$\text{INTERCEPT}_{1,f}$	1.69^{**} (0.20)	1.37^{**} (0.19)	1.75^{**} (0.20)	1.57^{**} (0.23)	1.97^{**} (0.24)	1.63^{**} (0.24)
$\text{NCH02}_{1,f}$	-1.16^{**} (0.18)	-0.98^{**} (0.17)	-1.19^{**} (0.19)	-0.97^{**} (0.23)	-0.99^{**} (0.26)	-0.97^{**} (0.24)
$\text{NCH34}_{1,f}$	-0.80^{**} (0.17)	-0.76^{**} (0.16)	-0.82^{**} (0.17)	-0.48 (0.27)	-0.53 (0.30)	-0.47 (0.27)
$\text{NCH511}_{1,f}$	-0.40^{**} (0.08)	-0.40^{**} (0.07)	-0.41^{**} (0.08)	-0.05 (0.20)	-0.04 (0.23)	-0.03 (0.21)
$\text{NCH1215}_{1,f}$	-0.13 (0.12)	-0.17 (0.12)	-0.13 (0.12)	-0.10 (0.19)	-0.18 (0.22)	-0.11 (0.20)
$\text{EDUC1}_{1,f}$	-0.91^{**} (0.19)	-0.62^{**} (0.18)	-0.95^{**} (0.19)	-1.35^{**} (0.39)	-1.55^{**} (0.45)	-1.40^{**} (0.40)
$\text{EDUC2}_{1,f}$	-3.12^{**} (0.35)	-2.28^{**} (0.33)	-3.23^{**} (0.36)	-3.73^{**} (0.48)	-4.66^{**} (0.55)	-3.88^{**} (0.49)
$\text{AGE}_{1,f}$	0.10 (0.68)	0.47 (0.67)	0.04 (0.69)	0.33 (0.71)	-0.54 (0.78)	0.28 (0.71)
$\text{AGE}^2{}_{1,f}$	-0.51 (0.77)	-0.80 (0.76)	-0.46 (0.78)	-0.77 (0.79)	0.07 (0.86)	-0.73 (0.80)
$\text{LOTHERY}_{1,f}$	-0.12^{*} (0.06)	-0.11 (0.06)	-0.12^{*} (0.06)	-0.19^{*} (0.08)	-0.20^{*} (0.09)	-0.19^{*} (0.08)
$\text{INTERCEPT}_{1,p}$	1.83^{**} (0.21)	1.43^{**} (0.20)	1.90^{**} (0.21)	1.93^{**} (0.25)	2.38^{**} (0.25)	2.00^{**} (0.25)
$\text{NCH02}_{1,p}$	-0.48^{**} (0.15)	-0.48^{**} (0.14)	-0.51^{**} (0.15)	-0.09 (0.21)	-0.06 (0.23)	-0.08 (0.21)
$\text{NCH34}_{1,p}$	-0.18 (0.14)	-0.22 (0.13)	-0.20 (0.14)	-0.04 (0.23)	-0.13 (0.25)	-0.06 (0.23)
$\text{NCH511}_{1,p}$	0.05 (0.07)	-0.03 (0.07)	0.05 (0.07)	0.19 (0.18)	0.22 (0.20)	0.20 (0.18)
$\text{NCH1215}_{1,p}$	0.16 (0.11)	0.15 (0.11)	0.17 (0.12)	0.09 (0.18)	0.06 (0.19)	0.08 (0.18)
$\text{EDUC1}_{1,p}$	-0.91^{**} (0.17)	-0.63^{**} (0.17)	-0.95^{**} (0.18)	-0.63 (0.42)	-0.76 (0.46)	-0.64 (0.43)
$\text{EDUC2}_{1,p}$	-3.56^{**} (0.36)	-2.64^{**} (0.34)	-3.69^{**} (0.37)	-3.31^{**} (0.49)	-4.26^{**} (0.55)	-3.44^{**} (0.50)
$\text{AGE}_{1,p}$	-0.63 (0.68)	-0.34 (0.67)	-0.71 (0.69)	-0.98 (0.72)	-1.97^{**} (0.76)	-1.13 (0.73)
$\text{AGE}^2{}_{1,p}$	0.68 (0.77)	0.36 (0.76)	0.76 (0.78)	0.90 (0.80)	1.95^{*} (0.84)	1.06 (0.82)
$\text{LOTHERY}_{1,p}$	-0.13^{*} (0.06)	-0.14^{*} (0.06)	-0.13^{*} (0.06)	-0.16 (0.09)	-0.11 (0.09)	-0.15 (0.09)
$\text{INTERCEPT}_{2,f}$	0.31^{**} (0.12)	0.35^{**} (0.10)	0.15 (0.12)	0.36^{**} (0.13)	-0.68^{**} (0.09)	0.20 (0.13)
$\text{NCH02}_{2,f}$	-0.73^{**} (0.09)	-0.57^{**} (0.07)	-0.76^{**} (0.09)	-0.70^{**} (0.09)	-0.74^{**} (0.10)	-0.75^{**} (0.10)
$\text{NCH34}_{2,f}$	-0.46^{**} (0.08)	-0.38^{**} (0.06)	-0.44^{**} (0.08)	-0.43^{**} (0.08)	-0.23^{**} (0.09)	-0.42^{**} (0.08)
$\text{NCH511}_{2,f}$	-0.22^{**} (0.04)	-0.26^{**} (0.04)	-0.21^{**} (0.04)	-0.22^{**} (0.04)	-0.12^{**} (0.04)	-0.21^{**} (0.04)
$\text{NCH1215}_{2,f}$	-0.06 (0.04)	-0.09 (0.04)	-0.05 (0.04)	-0.06 (0.05)	0.00 (0.05)	-0.06 (0.05)
$\text{EDUC1}_{2,f}$	-0.47^{**} (0.10)	-0.20 (0.11)	-0.49^{**} (0.10)	-0.48^{**} (0.19)	-0.37 (0.22)	-0.45^{*} (0.19)

[14] In a dynamic model of labor force participation including time invariant and autocorrelated unobserved preferences, Hyslop (1999) also finds negative autocorrelation in the time varying element of individuals' unobserved preferences.

VARIABLE	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5	MODEL 6
$\text{EDUC2}_{2,f}$	-2.03^{**} (0.15)	-1.33^{**} (0.17)	-2.07^{**} (0.15)	-2.06^{**} (0.21)	-1.75^{**} (0.22)	-2.08^{**} (0.21)
$\text{AGE}_{2,f}$	0.54 (0.27)	0.49 (0.29)	0.50 (0.27)	0.50 (0.28)	−0.08 (0.23)	0.46 (0.28)
$\text{AGE}^2{}_{2,f}$	-0.80^{**} (0.27)	-0.66^{**} (0.28)	-0.75^{**} (0.27)	-0.76^{**} (0.27)	−0.15 (0.23)	-0.72^{**} (0.27)
$\text{LOTHERY}_{2,f}$	-0.16^{**} (0.02)	-0.13^{**} (0.02)	-0.17^{**} (0.02)	-0.17^{**} (0.03)	-0.21^{**} (0.03)	-0.18^{**} (0.03)
$\gamma_{f,f}$	1.58^{**} (0.15)	0.64^{**} (0.09)	1.91^{**} (0.19)	1.54^{**} (0.15)	3.34^{**} (0.14)	1.88^{**} (0.19)
$\gamma_{f,p}$	0.91^{**} (0.10)	0.36^{**} (0.07)	1.14^{**} (0.13)	0.88^{**} (0.11)	2.25^{**} (0.06)	1.10^{**} (0.13)
$\text{INTERCEPT}_{2,p}$	0.83^{**} (0.10)	0.93^{**} (0.10)	0.69^{**} (0.10)	0.86^{**} (0.11)	−0.10 (0.07)	0.74^{**} (0.11)
$\text{NCH02}_{2,p}$	-0.32^{**} (0.07)	-0.29^{**} (0.05)	-0.32^{**} (0.07)	-0.31^{**} (0.07)	-0.33^{**} (0.08)	-0.33^{**} (0.07)
$\text{NCH34}_{2,p}$	-0.14^{*} (0.06)	-0.14^{**} (0.04)	−0.12 (0.06)	-0.14^{*} (0.06)	−0.06 (0.07)	-0.13^{*} (0.06)
$\text{NCH511}_{2,p}$	−0.01 (0.03)	0.02 (0.03)	−0.01 (0.03)	−0.03 (0.03)	−0.02 (0.03)	−0.03 (0.03)
$\text{NCH1215}_{2,p}$	0.05 (0.04)	0.06 (0.04)	0.05 (0.04)	0.02 (0.04)	0.03 (0.04)	0.02 (0.04)
$\text{EDUC1}_{2,p}$	-0.56^{**} (0.09)	-0.40^{**} (0.10)	-0.57^{**} (0.09)	-0.58^{**} (0.10)	-0.46^{**} (0.08)	-0.58^{**} (0.10)
$\text{EDUC2}_{2,p}$	-2.49^{**} (0.15)	-1.89^{**} (0.19)	−2.51 (0.15)	-2.48^{**} (0.16)	-2.13^{**} (0.11)	-2.52^{**} (0.15)
$\text{AGE}_{2,p}$	0.01 (0.25)	0.13 (0.27)	−0.08 (0.25)	−0.03 (0.26)	-0.47^{*} (0.21)	−0.13 (0.25)
$\text{AGE}^2{}_{2,p}$	−0.05 (0.25)	−0.15 (0.27)	0.03 (0.25)	−0.01 (0.25)	0.42^{*} (0.21)	0.08 (0.25)
$\text{LOTHERY}_{2,p}$	-0.11^{**} (0.02)	-0.09^{**} (0.02)	-0.11^{**} (0.02)	-0.11^{**} (0.02)	-0.13^{**} (0.03)	-0.11^{**} (0.02)
$\gamma_{p,f}$	0.79^{**} (0.11)	0.27^{**} (0.06)	1.00^{**} (0.14)	0.73^{**} (0.11)	2.11^{**} (0.06)	0.93^{**} (0.13)
$\gamma_{p,p}$	1.22^{**} (0.10)	0.14^{**} (0.05)	1.53^{**} (0.12)	1.22^{**} (0.10)	2.61^{**} (0.08)	1.52^{**} (0.12)
η_1	1.74^{**} (0.14)	1.37^{**} (0.13)	1.79^{**} (0.15)	1.69^{**} (0.15)	2.06^{**} (0.16)	1.75^{**} (0.16)
η_2	1.18^{**} (0.06)	0.93^{**} (0.07)	1.18^{**} (0.05)	1.19^{**} (0.06)	0.99^{**} (0.05)	1.20^{**} (0.06)
$v_{1,2}$	0.28^{**} (0.04)	0.56^{**} (0.05)	0.35^{**} (0.05)	0.26^{**} (0.04)	0.81^{**} (0.04)	0.31^{**} (0.05)
$\mu_{1,1}$	0.58 (0.06)	-	0.51 (0.07)	0.60 (0.06)	-	0.53 (0.06)
$\mu_{2,2}$	0.47 (0.04)	-	0.40 (0.05)	0.48 (0.05)	-	0.42 (0.05)
$\mu_{1,2}$	0.44^{**} (0.04)	-	0.38^{**} (0.05)	0.46^{**} (0.04)	-	0.40^{**} (0.05)
$u_{1,2}$	0.50^{**} (0.09)	0.60^{**} (0.09)	0.51^{**} (0.09)	0.58^{**} (0.09)	0.42^{**} (0.02)	0.57^{**} (0.09)
$c_{1,1}$	0.52 (0.05)	0.59 (0.04)	0.47 (0.06)	0.52 (0.05)	0.09 (0.02)	0.48 (0.05)
$c_{1,2}$	0.28^{**} (0.06)	0.25^{**} (0.05)	0.25^{**} (0.06)	0.31^{**} (0.05)	0.03 (0.03)	0.28^{**} (0.05)
$c_{2,1}$	0.36^{**} (0.06)	0.34^{**} (0.05)	0.32^{**} (0.06)	0.36^{**} (0.06)	0.04 (0.31)	0.30^{**} (0.06)
$c_{2,2}$	0.43 (0.04)	0.53 (0.04)	0.38 (0.05)	0.42 (0.04)	0.07 (0.37)	0.36 (0.05)
ρ	-	0.88^{**} (0.01)	-0.24^{**} (0.03)	-	-0.33^{**} (0.07)	-0.25^{**} (0.03)
$\overline{\text{NCH02}}_{1,f}$	-	-	-	0.06 (0.49)	0.07 (0.11)	0.10 (0.47)
$\overline{\text{NCH34}}_{1,f}$	-	-	-	−0.05 (0.58)	−0.02 (0.04)	−0.07 (0.56)

VARIABLE	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5	MODEL 6
$\overline{\text{NCH511}}_{1,f}$	-	-	-	−0.10 (0.13)	0.05 (0.22)	−0.05 (0.12)
$\overline{\text{NCH1215}}_{1,f}$	-	-	-	0.16 (0.16)	0.10 (0.26)	0.15 (0.16)
$\overline{\text{LOTHERY}}_{1,f}$	-	-	-	0.03 (0.05)	0.13 (0.30)	0.05 (0.05)
$\overline{\text{EDUC1}}_{1,f}$	-	-	-	−0.07 (0.20)	−0.19 (0.36)	−0.13 (0.20)
$\overline{\text{EDUC2}}_{1,f}$	-	-	-	−0.01 (0.20)	−0.21** (0.08)	−0.07 (0.20)
$\overline{\text{NCH02}}_{1,p}$	-	-	-	0.26 (0.41)	0.09 (0.11)	0.23 (0.40)
$\overline{\text{NCH34}}_{1,p}$	-	-	-	−0.25 (0.49)	−0.15** (0.03)	−0.23 (0.48)
$\overline{\text{NCH511}}_{1,p}$	-	-	-	−0.01 (0.11)	0.09 (0.07)	0.02 (0.10)
$\overline{\text{NCH1215}}_{1,p}$	-	-	-	0.16 (0.14)	0.09 (0.08)	0.15 (0.13)
$\overline{\text{LOTHERY}}_{1,p}$	-	-	-	0.03 (0.04)	0.09 (1.31)	0.04 (0.04)
$\overline{\text{EDUC1}}_{1,p}$	-	-	-	−0.05 (0.07)	−0.09 (1.56)	−0.07 (0.07)
$\overline{\text{EDUC2}}_{1,p}$	-	-	-	−0.07 (0.07)	−0.09 (0.38)	−0.07 (0.07)
$\overline{\text{NCH02}}_{2,f}$	-	-	-	0.17 (1.17)	−0.35 (0.80)	0.13 (1.19)
$\overline{\text{NCH34}}_{2,f}$	-	-	-	0.99 (1.39)	1.78** (0.12)	1.09 (1.41)
$\overline{\text{NCH511}}_{2,f}$	-	-	-	−0.60* (0.30)	−0.88 (0.50)	−0.67** (0.33)
$\overline{\text{NCH1215}}_{2,f}$	-	-	-	−0.37 (0.71)	−0.24 (0.53)	−0.37 (0.72)
$\overline{\text{LOTHERY}}_{2,f}$	-	-	-	0.15 (0.11)	0.16 (1.33)	0.15 (0.11)
$\overline{\text{EDUC1}}_{2,f}$	-	-	-	0.55 (0.45)	0.44 (1.44)	0.56 (0.46)
$\overline{\text{EDUC2}}_{2,f}$	-	-	-	0.81* (0.40)	0.87** (0.33)	0.83* (0.39)
$\overline{\text{NCH02}}_{2,p}$	-	-	-	−1.80 (1.20)	−2.80** (0.68)	−2.00 (1.22)
$\overline{\text{NCH34}}_{2,p}$	-	-	-	1.27 (1.32)	1.92** (0.12)	1.40 (1.34)
$\overline{\text{NCH511}}_{2,p}$	-	-	-	−0.71* (0.30)	−0.83 (0.51)	−0.74** (0.31)
$\overline{\text{NCH1215}}_{2,p}$	-	-	-	0.13 (0.62)	0.32 (0.54)	0.15 (0.63)
$\overline{\text{LOTHERY}}_{2,p}$	-	-	-	0.08 (0.11)	0.04 (0.12)	0.07 (0.11)
$\overline{\text{EDUC1}}_{2,p}$	-	-	-	−0.29 (0.49)	−0.43 (0.51)	−0.31 (0.50)
$\overline{\text{EDUC2}}_{2,p}$	-	-	-	−0.15 (0.52)	−0.11 (0.54)	−0.15 (0.52)
Log likelihood	-3665.90	-3944.70	-3640.50	-3646.80	-3724.00	-3619.40
Pseudo R^2	0.62	0.59	0.62	0.62	0.61	0.62

Table 7: Results for models 1-6: Standard errors in parenthesis. * indicates significance at the 0.05 level and ** indicates significance at the 0.01 level.

It is clear that young children reduce the utility of full-time employment, and to a lesser extent young children also reduce the utility of part-time employment. Conditional on the wage, education reduces the utility of both full-time and part-time employment, with the effect being greater for a high level of education than for a low level of education. The results

also show a small yet significant negative effect of non-labor income on the utility of full-time employment and also on the utility of part-time employment. Non-labor income has a greater effect, in absolute terms, on the utility of full-time employment than on the utility of part-time employment. Thus as an individual's non-labor income increases, they are increasingly likely to prefer part-time employment to full-time employment.

Table 8 shows the total marginal effect of each demographic variable on the probability of being in each employment state. Furthermore, the total marginal effects are decomposed into wage effects and preference effects. The wage effect of a variable is defined as the change in the employment probabilities due to the effect the variable has on wages, holding preferences fixed. Similarly, the preference effect of a variable is defined as the change in the employment probabilities due to the effect the variable has on preferences, holding wages fixed.[15,16,17]

First, the results that are common across the six models are discussed. Table 8 shows that a low level of education increases the probability of full-time employment and reduces the probability of part-time employment. Older individuals have a higher probability of being in full-time employment than younger individuals, and non-labor income reduces the probabilities of both full-time and part-time employment. The birth of a child at wave 1 reduces the probability of full-time employment and increases the probability of non-employment. Vocational qualifications increase the probabilities of both full-time employment and part-time employment.

Amongst the models without correlated preferences, Model 2, which does not have time invariant unobserved preferences, produces somewhat different results than either Model 1 or Model 3, which both include time invariant unobserved preferences. Similarly, Models 4 and 6 produce similar marginal effects, but these differ somewhat from the marginal effects implied by Model 5. This suggests that the estimated marginal effects are sensitive to whether or not time invariant unobserved preferences are permitted.

The decomposition of the total marginal effects into wage effects and preference effects reveals some interesting results. Consider the results for Model 6. Although the total effect of a high level of education is to increase the probabilities of full-time and part-time employment, this effect is due to the large wage effect associated with a high level of education. Individuals

[15] All marginal effects have been averaged over the 12 waves and refer to a women who, at wave 1, is aged 20 years. At each wave the women has no children, no educational or vocational qualifications, is not a member of a union and has a non-labor income of £10000 per year.

[16] The marginal effect of a child refers to the effect of a child who is aged 1 year at wave 1, and ages one year per wave.

[17] The marginal effect of income refers to the effect of a £500 per year increase in non-labor income.

who have a high level of education have, ceteris paribus, a lower preference for both full-time and part-time employment than individuals with no academic qualifications.[18]

	WAGE EFFECT			PREFERENCE EFFECT			TOTAL EFFECT		
	f	p	n	f	p	n	f	p	n
Model 1									
EDUC1	0.302	0.074	-0.376	-0.098	-0.132	0.231	0.204	-0.058	-0.146
EDUC2	0.241	0.352	-0.593	-0.174	-0.239	0.413	0.067	0.113	-0.180
AGE	0.009	0.006	-0.015	0.002	-0.003	0.001	0.010	0.003	-0.013
INCOME	-	-	-	-0.002	-0.002	0.004	-0.002	-0.002	0.004
CHILD	-	-	-	-0.098	0.011	0.087	-0.098	0.011	0.087
VOC	0.020	0.078	-0.099	-	-	-	0.020	0.078	-0.099
MODEL 2									
EDUC1	0.187	0.087	-0.275	-0.030	-0.108	0.138	0.158	-0.021	-0.137
EDUC2	0.243	0.313	-0.555	-0.158	-0.241	0.399	0.085	0.072	-0.157
AGE	0.006	0.005	-0.011	0.002	0.000	-0.003	0.008	0.005	-0.013
INCOME	-	-	-	-0.002	-0.002	0.004	-0.002	-0.002	0.004
CHILD	-	-	-	-0.075	0.015	0.060	-0.075	0.015	0.060
VOC	0.018	0.047	-0.065	-	-	-	0.018	0.047	-0.065
MODEL 3									
EDUC1	0.298	0.101	-0.399	-0.107	-0.161	0.268	0.191	-0.060	-0.131
EDUC2	0.208	0.368	-0.576	-0.174	-0.236	0.410	0.034	0.132	-0.166
AGE	0.011	0.008	-0.019	0.002	-0.007	0.005	0.013	0.001	-0.014
INCOME	-	-	-	-0.003	-0.001	0.004	-0.003	-0.001	0.004
CHILD	-	-	-	-0.099	0.008	0.091	-0.099	0.008	0.091
VOC	0.016	0.087	-0.103	-	-	-	0.016	0.087	-0.103
MODEL 4									
EDUC1	0.313	0.043	-0.356	-0.094	-0.177	0.272	0.218	-0.134	-0.084
EDUC2	0.312	0.276	-0.588	-0.156	-0.274	0.430	0.156	0.002	-0.158
AGE	0.008	0.007	-0.015	0.003	-0.007	0.004	0.011	0.000	-0.012
INCOME	-	-	-	-0.005	-0.001	0.006	-0.005	-0.001	0.006
CHILD	-	-	-	-0.088	0.009	0.080	-0.088	0.009	0.080
VOC	0.018	0.090	-0.109	-	-	-	0.018	0.090	-0.109
MODEL 5									
EDUC1	0.394	-0.005	-0.389	-0.160	-0.247	0.407	0.234	-0.252	0.018
EDUC2	0.159	0.321	-0.480	-0.189	-0.331	0.520	-0.030	-0.010	0.040
AGE	0.012	0.013	-0.025	-0.006	-0.022	0.028	0.006	-0.009	0.003
INCOME	-	-	-	-0.001	-0.003	0.004	-0.001	-0.003	0.004
CHILD	-	-	-	-0.115	0.010	0.105	-0.115	0.010	0.105
VOC	0.015	0.134	-0.149	-	-	-	0.015	0.134	-0.149
MODEL 6									
EDUC1	0.303	0.055	-0.358	-0.086	-0.199	0.285	0.218	-0.145	-0.073
EDUC2	0.249	0.328	-0.577	-0.145	-0.279	0.424	0.104	0.049	-0.153
AGE	0.007	0.006	-0.013	0.002	-0.009	0.008	0.008	-0.003	-0.005
INCOME	-	-	-	-0.003	-0.002	0.005	-0.003	-0.002	0.005
CHILD	-	-	-	-0.073	-0.001	0.074	-0.073	-0.001	0.074
VOC	0.017	0.100	-0.117	-	-	-	0.017	0.100	-0.117

Table 8: Marginal effects of demographic variables on employment probabilities: Total effects are decomposed into wage effects and preference effects.

[18] In Model 6, a low level of education also has a positive wage effect and a negative preference effect. In the case of full-time employment, the wage effect dominates, as for a high level of education, and in the case of part-time employment, the preference effect dominates.

Table 9 shows that, for all six models, the hypothesis that the initial conditions are exogenous is strongly rejected. Thus treating the initial conditions as exogenous would produce inconsistent parameter estimates. In the models which allow correlated unobserved preferences, the hypothesis that education, fertility or non-labor income are exogenous can be tested. Table 9 shows that, for all three of the relevant models, the hypothesis that children are exogenous is rejected. Non-labor income is exogenous is Models 4 and 6, but there is significant evidence that non-labor income is not exogenous in Model 5. Education is exogenous in all three models with correlated unobserved preferences.

VARIABLE	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5	MODEL 6
Income	-	-	-	2.08 (0.72)	**12.15** (0.02)	2.52 (0.64)
NCH02	-	-	-	3.87 (0.42)	4.71 (0.32)	4.02 (0.40)
NCH04	-	-	-	1.60 (0.81)	2.72 (0.61)	1.75 (0.78)
NCH511	-	-	-	6.96 (0.14)	**10.30** (0.04)	7.18 (0.13)
NCH1215	-	-	-	2.15 (0.71)	1.47 (0.83)	1.99 (0.74)
All Children	-	-	-	**26.73** (0.04)	**32.79** (0.01)	**27.81** (0.03)
EDUC1	-	-	-	4.10 (0.39)	4.47 (0.35)	4.77 (0.31)
EDUC2	-	-	-	5.78 (0.22)	4.89 (0.30)	5.85 (0.21)
All Education	-	-	-	6.55 (0.59)	6.96 (0.54)	6.93 (0.54)
Exogenity of Initial Conditions	**104.57** (0.00)	**156.95** (0.00)	**76.94** (0.00)	**94.43** (0.00)	**38.41** (0.00)	**73.16** (0.00)

Table 9: Wald tests for exogenity: p values in parenthesis.

In order to determine if one particular model is preferred to the other models, simulations, based on the data set used for estimation, are conducted to assess the fit of each model. The steps of each simulation are as follows:

1. For each individual in the sample, draw a 1 by $2T$ vector of standard normal variables.
2. Given these draws and the estimated parameters, construct each individual's employment behavior over the 12 waves of the survey.
3. Repeat steps 1-2 100 times.

The results of these simulations are presented in Tables 10 and 11. For all six models, the predicted proportions of individuals in each state at each wave are similar to the actual proportions, shown in Table 1. However, only Models 3 and 6 which include both time invariant and autocorrelated unobserved preferences generate a decrease then an increase in full-time employment and non-employment, as is observed in the sample.

The transition matrices for the simulated data, given in Table 11, show that all six models capture a large amount of the observed state dependence. Furthermore, all of the models capture the greater tendency of individuals to move to adjacent states than to non-adjacent states. Models 4-6, which permit correlated unobserved preferences, produce marginally higher levels of state dependence than Models 1-3, which do not include correlated unobserved preferences. Based on these simulation results, Model 6 appears to be the preferred model. This model generates employment probabilities which closely match the sample proportions, and also predicts a large amount of the observed state dependence.

	WAVE												
STATE	ALL	1	2	3	4	5	6	7	8	9	10	11	12
						MODEL 1							
n	0.32	0.35	0.34	0.33	0.33	0.32	0.32	0.32	0.31	0.31	0.31	0.31	0.31
p	0.33	0.30	0.33	0.34	0.34	0.34	0.35	0.34	0.34	0.34	0.33	0.33	0.33
f	0.34	0.35	0.33	0.33	0.34	0.34	0.34	0.34	0.34	0.35	0.35	0.36	0.36
						MODEL 2							
n	0.33	0.35	0.34	0.34	0.34	0.33	0.33	0.33	0.32	0.32	0.32	0.32	0.32
p	0.33	0.30	0.34	0.34	0.33	0.34	0.35	0.34	0.34	0.34	0.33	0.33	0.33
f	0.34	0.35	0.32	0.32	0.33	0.33	0.33	0.33	0.33	0.35	0.35	0.35	0.35
						MODEL 3							
n	0.32	0.35	0.34	0.33	0.33	0.32	0.32	0.32	0.31	0.31	0.32	0.32	0.32
p	0.33	0.30	0.33	0.34	0.34	0.34	0.35	0.34	0.34	0.34	0.33	0.33	0.33
f	0.34	0.35	0.33	0.33	0.34	0.34	0.34	0.34	0.34	0.35	0.35	0.36	0.36
						MODEL 4							
n	0.34	0.34	0.34	0.34	0.34	0.34	0.33	0.33	0.33	0.33	0.33	0.33	0.33
p	0.35	0.33	0.35	0.35	0.35	0.35	0.36	0.35	0.35	0.35	0.35	0.34	0.34
f	0.32	0.33	0.31	0.31	0.31	0.31	0.31	0.32	0.32	0.32	0.32	0.33	0.33
						MODEL 5							
n	0.31	0.35	0.34	0.33	0.32	0.32	0.31	0.30	0.30	0.29	0.30	0.30	0.30
p	0.33	0.31	0.33	0.33	0.33	0.34	0.34	0.34	0.34	0.34	0.33	0.33	0.33
f	0.35	0.34	0.34	0.34	0.34	0.35	0.35	0.36	0.36	0.37	0.37	0.38	0.38
						MODEL 6							
n	0.32	0.34	0.33	0.33	0.33	0.32	0.32	0.32	0.32	0.31	0.32	0.31	0.32
p	0.33	0.31	0.34	0.34	0.34	0.34	0.34	0.34	0.34	0.33	0.33	0.33	0.33
f	0.34	0.35	0.33	0.33	0.34	0.34	0.34	0.35	0.34	0.35	0.35	0.36	0.36

Table 10: Simulation results for Models 1-6: Predicted proportion of individuals in each state for all waves combined and waves 1-12 separately.

TRANSITION	MODEL 1	MODEL 2	MODEL 3	MODEL 4	MODEL 5	MODEL 6
$n_t n_{t+1}$	0.85	0.82	0.86	0.86	0.86	0.86
$n_t p_{t+1}$	0.11	0.13	0.11	0.11	0.10	0.11
$n_t f_{t+1}$	0.03	0.05	0.03	0.03	0.03	0.04
$p_t n_{t+1}$	0.09	0.11	0.09	0.09	0.08	0.09
$p_t p_{t+1}$	0.80	0.77	0.81	0.81	0.82	0.80
$p_t f_{t+1}$	0.11	0.12	0.11	0.10	0.10	0.11
$f_t n_{t+1}$	0.04	0.06	0.04	0.05	0.03	0.04
$f_t p_{t+1}$	0.09	0.11	0.09	0.10	0.08	0.09
$f_t f_{t+1}$	0.86	0.83	0.87	0.86	0.89	0.86

Table 11: Simulation results for models 1-6: Transition matrices for simulated data.

5 Simulating the Effect of a Wage Subsidy

As noted in the introduction, true and spurious state dependence have very different policy implications. If all state dependence is spurious, any temporary policy intervention which has the effect of moving individuals between employment states will only have a single period effect on employment behavior. In contrast, if true dependence is present, such a policy intervention will have a lasting effect on employment behavior.

Focusing on Model 6, the results in Table 7 indicate significant positive true state dependence in full-time and part-time employment behavior. However, time invariant and autocorrelated elements to individuals' preferences are also present, implying some degree of spurious state dependence. In order to assess the relative importance of true and spurious state dependence, the effects of three temporary wage subsidies are simulated and compared. The first policy subsidizes the wages of all workers by 5%. The second policy subsidizes the wages of full-time workers by 5%, and the third policy subsidizes the wages of part-time worker by 5%. All three policies last for one year. The first intervention is equivalent to a 5% reduction in income tax, whereas the other two interventions correspond to incentives for individuals to work full-time and part-time respectively.

Figures 1-3 illustrate the dynamic effects of the three different policies. The results of the simulations suggest that wage subsidies aimed solely at individuals in full-time employment will attract individuals from both part-time employment and non-employment. Similarly, a wage subsidy aimed at increasing part-time employment has an adverse effect on the number of individuals in full-time employment. Only the policy of subsidizing all wages increases both full-time and part-time employment. Indeed, it is this policy that produces the largest contemporaneous reduction in non-employment. Examining Figures 1-3 reveals that there is

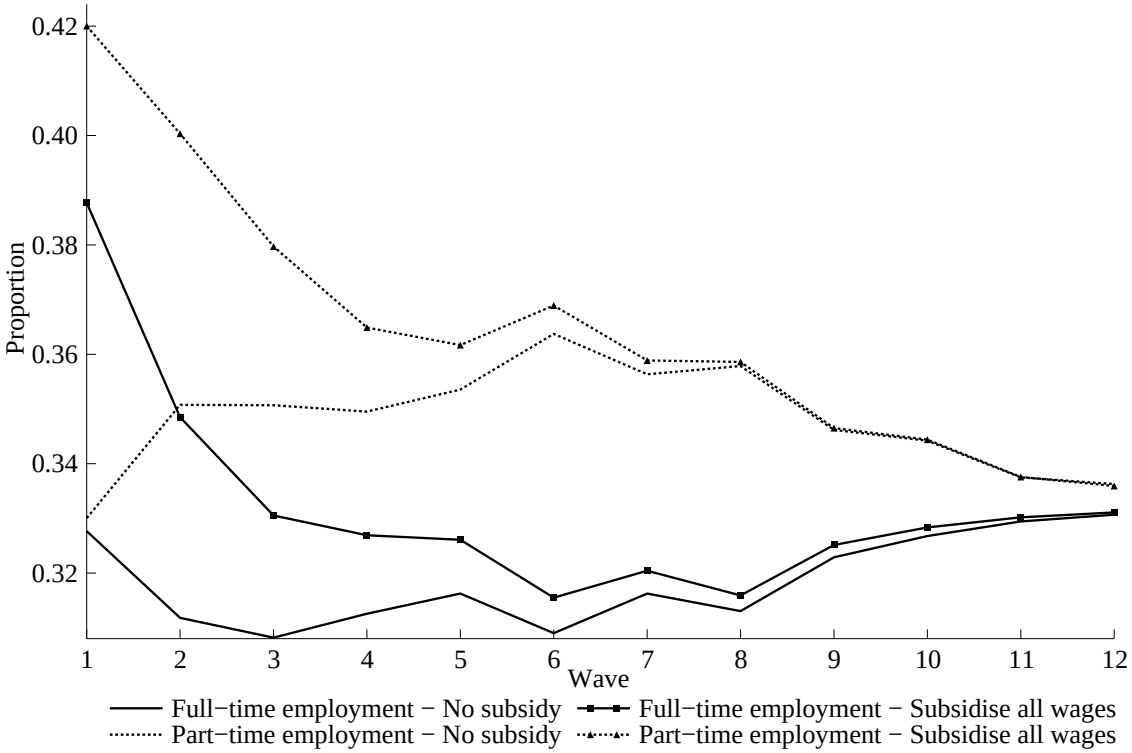

Figure 1: Dynamic effects of temporally subsidizing all wages.

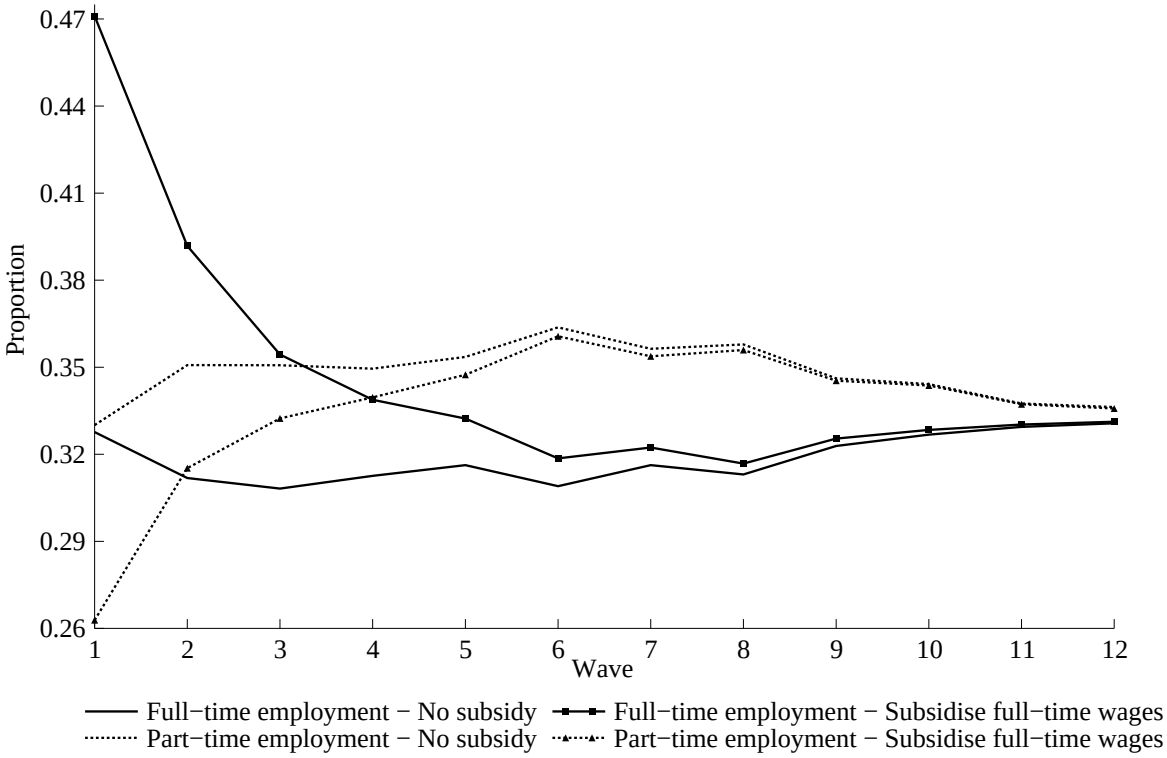

Figure 2: Dynamic effects of temporally subsidizing full-time wages.

somewhat more true state dependence in full-time employment than in part-time employment. This is consistent with the parameter estimates presented above. Also, despite the presence of

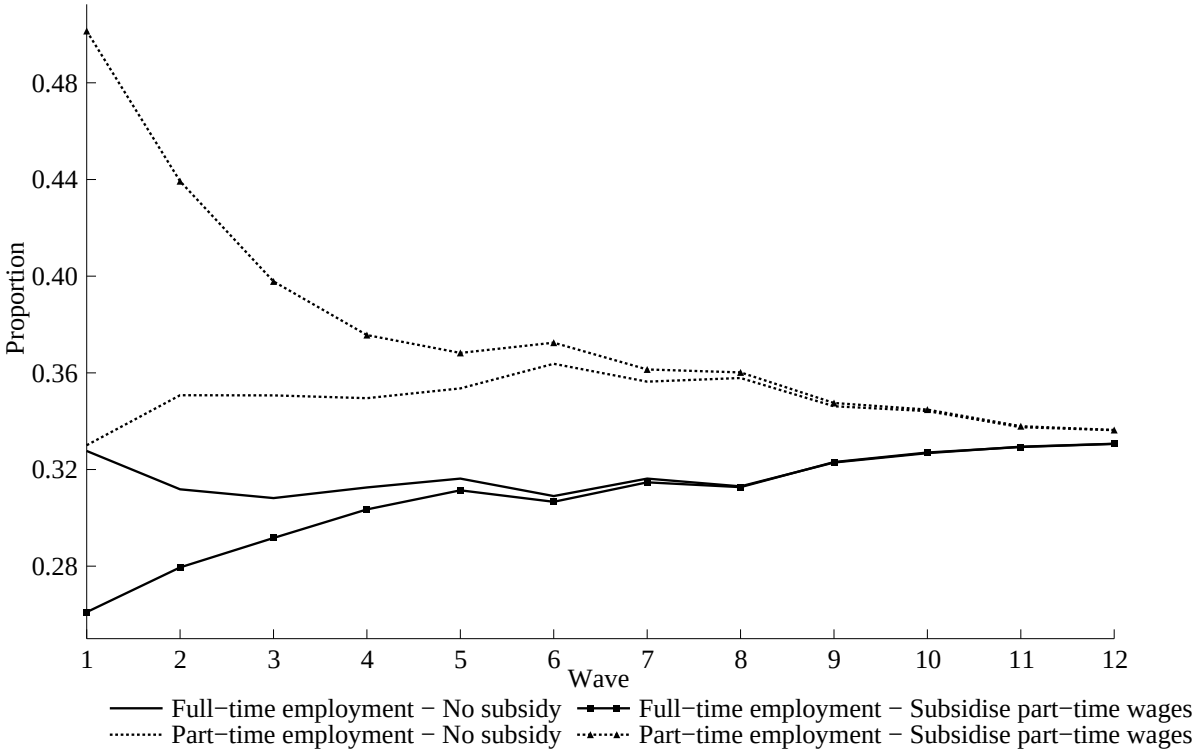

Figure 3: Dynamic effects of temporally subsidizing part-time wages.

significant true state dependence in both full-time employment and part-time employment, the effects of each of the three policies under consideration gradually decrease over time. Indeed, around six years after the interventions cease, the remaining effects are negligible.

In order to determine which wage subsidy a policy maker may prefer, the following payoff function is specified for the policy maker:

$$\Pi = \sum_{t=1}^{T} \delta^{t-1} \left[\alpha P_{f,t} + (1-\alpha) P_{p,t} \right], \tag{9}$$

where $\delta \in [0,1]$ is the policy maker's discount factor, $P_{f,t}$ and $P_{p,t}$ are the proportions of individuals in full-time and part-time employment at time t and $\alpha \in [0,1]$ is the weight the policy maker attaches to full-time employment relative to part-time employment.[19]

Table 12 shows the policy maker's payoffs from each of the three wage subsidy policies, relative to their payoff if no intervention is made, and ignoring any costs of the policy interventions. The policy maker's payoff is shown for α=1/2 and α=2/3. For these two values of α the policy maker's payoff is computed for $\delta = 1$, corresponding to the case where there is no discounting, and when δ=1/2, corresponding to substantial discounting. A policy maker

[19]Given the effects of all three wage subsidy policies are essentially zero by wave 12, there is no loss in only considering the policy maker's preferences over the 12 waves following the intervention, instead of over a longer time horizon.

	α=$\frac{1}{2}$, δ=1	α=$\frac{1}{2}$, δ=$\frac{1}{2}$	α=$\frac{2}{3}$, δ=1	α=$\frac{2}{3}$, δ=$\frac{1}{2}$
POLICY 1	0.181	0.106	0.175	0.099
POLICY 2	0.095	0.054	0.176	0.103
POLICY 3	0.116	0.072	0.033	0.018

Table 12: Policy maker's payoff from wage subsidy policies 1, 2 and 3.

who cares equally about full-time employment and part-time employment prefers to subsidize all wages as opposed to subsidizing either full-time wages or part-time wages. This is true for $\delta = 1$ and δ=1/2. Interestingly, if such a policy were unavailable, this policy maker prefers to subsidize the wages of individuals working part-time as opposed to subsidizing the wages of individuals working full-time. This is because, at each wave, the former policy produces a larger increase in total employment than the latter policy. In contrast, a policy maker who places twice as much weight on full-time employment than on part-time employment prefers to subsidize the wages of individuals working full-time, rather then subsidizing all wages or only subsidizing the wages of individuals working part-time. Again this is true for $\delta = 1$ and δ=1/2.

6 Conclusion

The above results mirror many of the well established findings from the labor force participation literature. In particular, children reduce the likelihood of employment, and educational and vocational qualifications both increase the probability of employment. Non-labor income has a small negative effect on the probability of employment. The multinomial framework adopted here also allows one to characterize how these variables affect individuals' choices between full-time and part-time employment. As expected, the income effect is larger for full-time employment than for part-time employment, and high levels of education increase the likelihood of full-time employment relative to part-time employment. Children, especially young children, tend to make full-time employment less attractive relative to part-time employment.

However, the most interesting results presented above relate to the dynamic nature of individuals' employment behavior. Observed employment behavior exhibits substantial positive state dependence, which is greater for full-time employment and non-employment than for part-time employment. The results above attribute the observed state dependence to a combination of true state dependence, which is greater for full-time employment than for part-time employment, and spurious state dependence.

The presence of significant positive true state dependence in employment behavior implies that any temporary policy intervention will have more than a transitory effect on employment

behavior. To explore this possibility further the effects of three temporary wage subsidies have been investigated. All three wage subsidies have substantial contemporaneous effects on employment behavior. Moreover, the effects of the subsidies persist for several years after the interventions cease. However, after 6 years the effect of any policy intervention is negligible: the element of true state dependence in employment behavior to not large enough to allow policy interventions to generate longer lasting effects. Thus, in order to generate a permanent reduction in non-employment, persistent or sustained policy interventions are required.

References

Blank, R. (1989), 'The role of part-time work in women's labor market choices over time', *American Economic Review* **79**(2), 295–99.

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* **61**(2), 167–97.

Burdett, K. and Taylor, M. (1994), Semi-markov and markov labour histories, ISER working papers 1994-27, Institute for Social and Economic Research.

Chamberlain, G. (1984), Panel data, *in* Griliches and Intriligator, eds, 'Handbook of Econometrics', Vol. 2, Amsterdam: North-Holland.

Corral, A. and Isusi, I. (2004), 'Part-time work in europe', *European Foundation for the Improvement of Living and Working Conditions* .

Eckstein, Z. and Wolpin, K. (1989), 'Dynamic labour force participation of married women and endogenous work experience', *Review of Economic Studies* **56**(3), 375–90.

Fagan, C. and Burchell, B. (2002), 'Gender, jobs and working conditions in the european union', *European Foundation for the Improvement of Living and Working Conditions* .

Geweke, J. (1991), 'Efficient simulation from the multivariate normal and student-t distributions subject to linear constraints', *Computer Sciencesand Statistics Proceedings of the 23rd Symposium on the Interface* pp. 571–578.

Geweke, J. F., Keane, M. P. and Runkle, D. E. (1997), 'Statistical inference in the multinomial multiperiod probit model', *Journal of Econometrics* **80**(1), 125–165.

Geweke, J., Keane, M. and Runkle, D. (1994), 'Alternative computational approaches to inference in the multinomial probit model', *The Review of Economics and Statistics* **76**(4), 609–32.

Hajivassiliou, V. (1999), Some practical issues in maximum simulated maximum likelihood, *in* S. T. Mariano, R and M. Weeks, eds, 'Simulation-Based Inference in Econometrics: Methods and Applications', Cambridge: University Press, chapter 3, pp. 71–99.

Hajivassiliou, V. and McFadden, D. (1998), 'The method of simulated scores for the estimation of ldv models', *Econometrica* **66**(4), 863–896.

Hajivassiliou, V., McFadden, D. and Ruud, P. (1996), 'Simulation of multivariate normal rectangle probabilities and their derivatives theoretical and computational results', *Journal of Econometrics* **72**(1), 85–134.

Hajivassiliou, V. and Rudd, P. A. (1994), Classical estimation methods for ldv models using simulation, *in* C. Engle and D. McFadden, eds, 'Handbook of Econometrics', Amsterdam: North-Holland, pp. 2383–41.

Hausman, J. and Wise, D. (1979), 'Attrition bias in experimental and panel data: The gary income maintenance experiment', *Econometrica* **47**(2), 455–73.

Heckman, J. (1981*a*), The incidental parameters problem and the problem of initial condition in estimating a discrete time-discrete data stochastic process, *in* C. Manski and D. McFadden, eds, 'Structural Analysis of Discrete Data and Econometric Applications', Cambridge: The MIT Press, chapter 4, pp. 179–197.

Heckman, J. (1981*b*), Statistical models for discrete panel data, *in* C. Manski and D. McFadden, eds, 'Structural Analysis of Discrete Data and Econometric Applications', Cambridge: The MIT Press, chapter 3, pp. 115–178.

Heckman, J. J. and Borjas, G. J. (1980), 'Does unemployment cause future unemployment? definitions, questions and answers from a continuous time model of heterogeneity and state dependence', *Economica* **47**(127), 247–83.

Heckman, J. and Willis, R. (1977), 'A beta-logistic model for the analysis of sequential labor force participation by married women', *The Journal of Political Economy* **85**(3), 27–58.

Hyslop, D. (1999), 'State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women', *Econometrica* **67**(6), 1255–1294.

Keane, M. (1994), 'A computationally practical simulation estimator for panel data', *Econometrica* **62**(1), 95–116.

Keane, M. P. (1992), 'A note on identification in the multinomial probit model', *Journal of Business Economic Statistics* **10**(2), 193–200.

Knights, S., Harris, M. and Loundes, J. (2002), 'Dynamic relationships in the australian labour market: Heterogeneity and state dependence', *The Economic Record* **78**(127), 284–98.

Layard, R. and Bean, C. (1989), 'Why does unemployment persist?', *Scandinavian Journal of Economics* **91**(2), 371–96.

Martin, J. and Roberts, C. (1984), 'Women and employment: a lifetime perspective', London: HMSO.

McFadden, D. (1989), 'A method of simulated moments for estimation of discrete response models without numerical integration', *Econometrica* **57**(5), 995–1026.

Metcalf, D. (1999), 'The low pay commission and the national minimum wage', *Economic Journal* **109**(127), 46–66.

Mincer, J. and Polachek, S. (1974), 'Family investments in human capital: Earnings of women', *The Journal of Political Economy* **82**(2), 76–108.

Narendranathan, W. and Elias, P. (1993), 'Influences of past history on the incidence of youth unemployment: Empirical findings for the uk', *Oxford Bulletin of Economics and Statistics* **55**(2), 161–85.

Rendtel, U. and Kaltenborn, U. (2004), The stability of simulation based estimation of the multiperiod multinominal probit model with individual specific covariates. Diskussionsbeiträge des Fachbereichs Wirtschaftswissenschaft der Freien Universität Berlin; 2004,5: Volkswirtschaftliche Reihe.

Robson, P., Dex, S., Wilkinson, F. and Salido Cortes, O. (1999), 'Low pay, labour market institutions, gender and part-time work: Cross-national comparisons', *European Journal of Industrial Relations* **5**(2), 187–207.

Taylor, M., Brice, J., Buck, N. and Prentice, E. (2001), *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices*, Colchester: University of Essex.

Appendix I: Monte Carlo Study

The literature contains several simulation studies of the small sample properties of estimators for multiperiod, multinomial choice models (see, for example, Geweke, Keane and Runkle 1994, Hyslop 1999, Keane 1992, Keane 1994, Rendtel and Kaltenborn 2004). However, none of these studies examine the properties of estimators for multiperiod, multinomial choice models with more than two alternatives and lagged dependant variables, as is the case in the above model.

Here, a Monte Carlo study is conducted to examine the performance of the MSL estimator using the GHK simulator in the context of the current application. In particular, the sensitivity of the results to the number of replications of the GHK simulator, R, is examined.

The model used in this simulation study consists of 3 alternatives and 5 time periods. The alternatives are labelled 0, 1 and 2 and the utility of alternative 0 is normalized to zero. The utility of alternative j at $t = 1$ is given by:

$$U_{i,j,1} = \alpha_{1,j} + \beta_{1,j} x_{i,1} + \overline{\beta}_{1,j} \overline{x}_i + \eta_1 w_{i,j,1} + \epsilon_{i,j,1}, \ j = 1, 2, \tag{10}$$

and at subsequent periods the utility function takes the following form:

$$U_{i,j,t} = \alpha_{2,j} + \beta_{2,j} x_{i,t} + \overline{\beta}_{2,j} \overline{x}_i + \eta_2 w_{i,j,t} + \gamma_{1,j} y_{i,1,t-1} + \gamma_{2,j} y_{i,2,t-1} + \epsilon_{i,j,t}, \ j = 1, 2 \text{ and } t = 2, ..., 5, \tag{11}$$

where $y_{i,j,t-1}$ for $j = 1, 2$ is as defined in Section 2. $x_{i,t}$ is a scalar random variable generated such that:

$$x_{i,t} \sim Uniform[0, 1], \tag{12}$$

and $\overline{x}_i$ is the average of $x_{i,t}$ over t. $w_{i,j,t}$ is an alternative specific covariate with the following distribution:

$$w_{i,j,t} \sim N(0, 1). \tag{13}$$

As in Section 2, $\epsilon_i | x_i, z_i \sim N(0, \Sigma)$. This simulation study is based on the most general form of Σ considered above i.e. allowing both time invariant and autocorrelated unobservables. Parameter values have been chosen such that the proportions of individuals in states 0, 1 and 2 are approximately 0.2, 0.5 and 0.3 respectively.

Simulations are conducted for R=10, 30 and 60. For each simulation, 50 data sets are generated, each containing 1000 individuals. Table 13 summarizes the results of this study. $\overline{\widehat{\theta}}$ is the mean value of the estimated parameters over the 50 replications, RMSE is the root mean squared error of the estimated parameters and $\overline{ASE}$ is the mean asymptotic standard error.

The Monte Carlo results shows substantial biases in some of parameters when R=10. In particular the γ's are substantially biased upwards, and the autocorrelation parameter ρ is

biased downwards. Thus when R=10 the MSL estimator suggests more true state dependence and less spurious state dependence than is actually the case. When R=30 the magnitude of the bias of the MSL estimator is reduced, although some bias still remains. When R is increased to 60, most of the bias is eliminated. As expected, as R increases the discrepancy between the mean asymptotic standard errors and the root mean squared errors decreases. These results indicate that a large number of replications are required to eliminate the bias introduced by simulating the likelihood.

		R=10			R=30			R=60		
PARAMETER	TRUTH	$\bar{\hat{\theta}}$	RMSE	$\overline{ASE}$	$\bar{\hat{\theta}}$	RMSE	$\overline{ASE}$	$\bar{\hat{\theta}}$	RMSE	$\overline{ASE}$
$\alpha_{2,1}$	-1	-1.09	0.19	0.16	-1.04	0.18	0.17	-1.02	0.17	0.17
$\beta_{2,1}$	0.6	0.61	0.03	0.04	0.61	0.03	0.04	0.61	0.03	0.04
$\gamma_{1,1}$	1.5	1.70	0.11	0.12	1.58	0.12	0.12	1.58	0.12	0.12
$\gamma_{1,2}$	0.5	0.60	0.11	0.10	0.54	0.13	0.11	0.52	0.11	0.11
$\alpha_{2,2}$	-0.8	-0.92	0.18	0.17	-0.87	0.16	0.18	-0.83	0.18	0.18
$\beta_{2,2}$	0.4	0.41	0.04	0.04	0.40	0.04	0.04	0.41	0.04	0.04
$\gamma_{2,1}$	0.5	0.60	0.10	0.11	0.56	0.11	0.11	0.55	0.12	0.11
$\gamma_{2,2}$	1.8	2.03	0.12	0.14	1.90	0.14	0.14	1.87	0.14	0.14
$\alpha_{1,1}$	-1	-1.02	0.24	0.28	-1.00	0.29	0.27	-1.01	0.30	0.27
$\beta_{1,1}$	-0.4	-0.45	0.06	0.08	-0.43	0.07	0.07	-0.41	0.06	0.07
$\alpha_{1,2}$	-0.8	-0.82	0.21	0.22	-0.85	0.25	0.22	-0.82	0.22	0.22
$\beta_{1,2}$	0.6	0.61	0.06	0.06	0.61	0.06	0.06	0.62	0.07	0.06
η_1	1.5	1.55	0.08	0.08	1.53	0.07	0.08	1.54	0.07	0.08
η_2	1	1.01	0.04	0.04	1.00	0.03	0.04	1.01	0.04	0.05
$v_{2,1}$	0.2	0.26	0.08	0.07	0.23	0.08	0.07	0.21	0.07	0.08
$\mu_{1,1}$	0.5	0.51	0.02	0.07	0.51	0.03	0.08	0.50	0.03	0.08
$\mu_{2,2}$	0.5	0.50	0.02	0.09	0.50	0.02	0.09	0.49	0.02	0.10
$\mu_{1,2}$	0.1	0.06	0.08	0.03	0.07	0.08	0.04	0.08	0.08	0.04
$c_{1,1}$	0.5	0.31	0.08	0.09	0.44	0.09	0.11	0.44	0.09	0.12
$c_{1,2}$	0	0.02	0.09	0.09	0.02	0.11	0.10	0.00	0.13	0.11
$c_{2,1}$	0	-0.02	0.08	0.07	0.00	0.09	0.07	-0.03	0.08	0.08
$c_{2,2}$	0.5	0.38	0.08	0.09	0.47	0.07	0.10	0.47	0.08	0.10
$u_{1,2}$	0.5	0.38	0.11	0.15	0.42	0.11	0.15	0.45	0.12	0.15
ρ	0.2	0.08	0.11	0.09	0.12	0.11	0.12	0.18	0.13	0.13
$\overline{\beta}_{1,1}$	0.2	0.21	0.12	0.15	0.20	0.17	0.15	0.19	0.16	0.15
$\overline{\beta}_{1,2}$	-0.2	-0.21	0.13	0.12	-0.19	0.15	0.12	-0.22	0.11	0.12
$\overline{\beta}_{2,1}$	0.1	0.08	0.10	0.08	0.10	0.08	0.08	0.09	0.08	0.08
$\overline{\beta}_{2,2}$	-0.3	-0.31	0.09	0.09	-0.29	0.08	0.09	-0.31	0.10	0.09

Table 13: Monte Carlo results.

The test proposed by Hajivassiliou (1999) is used to confirm that the bias in the MSL estimator is due to simulation noise. The test can be motivated by noting that, in the absence of simulation noise, the expectation of the score function evaluated at the truth is zero. Thus the null hypothesis is:

$$H_0 : E\left[\frac{\partial l(\theta_0; y)}{\partial \theta}\right] = 0, \tag{14}$$

where l and y denote the log likelihood and data respectively, and θ_0 is the true parameter vector. Rejection of the null hypothesis implies significant simulation noise, and consequently, the MSL estimator will be biased.

The test is implemented by constructing multiple data sets based on the data generation process $y(\theta_0)$. The empirical mean (m) and variance (v) of the score vector evaluated at θ_0 are then calculated. Under the null hypothesis:

$$NSm'v^{-1}m \sim \chi^2(k), \tag{15}$$

where S is the number artificial data sets, N is the number observations per data set, and k is the dimension of θ.

This test is constructed for R=10, 30, 60, 100 and 500, with S=100 throughout. Table (14) summarizes the results. For $R \leq 100$ the null hypothesis of no simulation noise is rejected. Comparing the test statistics for different values of R revels that there is a larger reduction in simulation noise when R is increased from 10 to 30, and also when R is increased from 30 to 60. Further increases in R lead to less dramatic reductions in the value of the test statistic. Only when R increases to 500 can the null hypothesis of no simulation noise not be rejected at the 1% level.

	R=10	R=30	R=60	R=100	R=500
χ^2 statistic (Score=0)	2221.10	945.35	238.77	145.47	47.36
p value	0.000	0.000	0.000	0.000	0.014

Table 14: Test statistics and p values for Hajivassiliou (1999) test of the score function.

Thus complete elimination of simulation noise requires a very large number of replications to be used when simulating the likelihood. Given that the computational speed of the GHK simulator is approximately linear in the number of replications, MSL estimation is very time intensive at high values of R. Since the bias of the MSL estimator is small, albeit significant, when R=60 it seems reasonable to use this value of R for empirical work.

Appendix II: Description of Variables

VARIABLE	DEFINITION
EDUC1	Indicator variable taking the value 1 if the individual has academic qualifications below A-levels, and zero otherwise.
EDUC2	Indicator variable taking the value 1 if the individual has academic qualifications of A-levels or above, and zero otherwise.
VOC	Indicator variable taking the value 1 if the individual has vocational qualifications and zero otherwise.
UNION	Indicator variable taking the value one if the individual's workplace is covered by a union that negotiates wages and zero otherwise.
AGE	Age in years.
AGE^2	AGE squared.
NCH02	Number of children aged under 3 years in the household.
NCH34	Number of children aged 3-4 years in the household.
NCH511	Number of children aged 5-11 years in household.
NCH1215	Number of children aged 12-15 years in household.
INCOME	Household income, excluding any labor market income of the individual under study, expressed in 1991 prices in thousands of British pounds.[a]
LOTHERY	Log of INCOME.
WAGE-f	Hourly full-time wage expressed 1991 prices in British pounds.
WAGE-p	Hourly part-time wage expressed in 1991 prices in British pounds.

[a]Income and wages have been deflated using the Retail Price Index (RPI).

Table 15: Definitions of variables.

	WAVE												
VARIABLE	ALL	1	2	3	4	5	6	7	8	9	10	11	12
EDUC1	0.36	0.40	0.40	0.39	0.39	0.38	0.37	0.36	0.35	0.34	0.33	0.32	0.32
EDUC2	0.43	0.36	0.37	0.39	0.40	0.41	0.42	0.43	0.44	0.46	0.47	0.49	0.50
VOC	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.38	0.39
UNION	0.55	0.58	0.55	0.54	0.53	0.53	0.55	0.56	0.54	0.55	0.54	0.58	0.57
AGE	42.10	36.65	37.60	38.60	39.63	40.64	41.60	42.59	43.59	44.57	45.60	46.59	47.59
NCH02	0.09	0.18	0.16	0.14	0.14	0.11	0.09	0.09	0.06	0.04	0.03	0.03	0.02
NCH34	0.11	0.17	0.18	0.19	0.15	0.14	0.14	0.11	0.10	0.08	0.06	0.03	0.03
NCH511	0.54	0.52	0.53	0.55	0.57	0.58	0.58	0.58	0.58	0.54	0.52	0.47	0.42
NCH1215	0.30	0.25	0.28	0.27	0.27	0.30	0.30	0.31	0.32	0.33	0.32	0.34	0.35
INCOME	21.87	19.13	19.99	20.41	20.64	20.91	21.57	22.03	22.30	22.81	23.60	24.33	24.68
WAGE-f	5.40	4.31	4.47	4.71	4.85	4.91	5.18	5.45	5.57	5.88	5.98	6.30	6.81
WAGE-p	4.89	4.11	4.28	4.24	4.34	4.46	4.72	4.68	5.16	5.32	5.49	6.08	5.89

Table 16: Sample means of variables.