Chaire de recherche Industrielle Alliance
sur les enjeux économiques
des changements démographiques
CEDIA.ca

18-11

Homeownership, Labour Market Transitions and Earnings

CAHIER DE RECHERCHE
WORKING PAPER

Thierry Kamionka and Guy Lacroix

Novembre / November 2018

Université Laval
Faculté des sciences sociales
HEC Montréal
ESG UQÀM

La Chaire de recherche Industrielle Alliance sur les enjeux économiques des changements démographiques est une chaire multi-institutionnelle qui s'appuie sur un partenariat avec les organisations suivantes :

- Centre interuniversitaire de recherche en analyse des organisations (CIRANO)
- iA Groupe financier
- Retraite Québec

Dépôt légal : Bibliothèque et Archives nationales du Québec et Bibliothèque et Archives Canada, 2018.
ISSN 2368-7207

Homeownership, Labour Market Transitions and Earnings[☆]

Thierry Kamionka[a], Guy Lacroix[b]

[a] *CNRS and CREST, Timbre J390. 15, Boulevard Gabriel Péri. 92245 Malakoff Cedex. France.*
[b] *Department of Applied Economics, HEC Montréal, IZA, CIRANO, Canada.*

Abstract

The paper investigates the links between homeownership, employment and earnings for which no consensus exists in the literature. Our analysis is cast within a dynamic setting and the endogeneity of each outcome is assessed through the estimation of a flexible panel multivariate model with random effects. The data we use are drawn from the French sample of the *EU Survey on Income and Living Conditions* for the years 2004–2013. The error terms are both correlated across equations and autocorrelated. Individual random effects are also correlated across equations. The model is estimated using a simulated maximum likelihood estimator and particular care is given to the initial conditions problem. Our results show that while homeowners have longer employment and unemployment spells, they must contend with lower earnings than tenants upon reemployment. They also stress the importance of unobserved heterogeneity in explaining the transitions on the labour and housing markets, and the relationship between earnings and the latter two. Failure to properly account for this is likely to yield biased parameter estimates.

Key words: Homeownership, Unemployment, Earnings, Heterogeneity, Simulation based estimation, Panel data.

JEL-codes: J21, J64, J31, C33, C35

[☆] The authors would like to thank Investissements d'Avenir (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047) for financial support. They are also grateful to Thierry Magnac and Julie Rochut for helpfull comments on the paper, as well as participants at the 2018 Meeting of the European Association of Labour Economists, the 2018 Econometric Society Australasian Meeting, the 2017 International Panel Data Conference, the 2017 Journées de Microéconomie Appliquée, and the Microeconometrics Seminars at CREST.

Email addresses: kamionka@ensae.fr (Thierry Kamionka), Guy.Lacroix@hec.ca (Guy Lacroix)

1. Introduction

As of 2014, approximately 70% of European households were homeowners[1] (INSEE, 2016). Homeownership rates vary from as little as 52.5% in Germany to as much as 90.3% in Slovakia. The United Kingdom (64.8%), France (65.1%), Italy (73.2%), and Spain (78.8%) are intermediate cases. The United Stated (63.5%, U.S. census 2016) would also qualify as an intermediate case.

It has been argued (Andrews and Sanchez, 2011) that such heterogeneity may partly stem from a wide array of programs and public policies that have been implemented over the years to foster access to homeownership. Programs such as subsidized loans, zero interest loans, reduced down payments, tax deductible mortgage interests, *etc.* are now widespread. In France, zero interest loans ("Prêts à Taux Zéro"), income tax incentives to stimulate investments in specific geographic areas (so called "Scellier, Duflot, Pinel" schemes) and capital tax abatements on the main residence were all designed to that end.

The rationale for subsidizing homeownership is manifold (Havet and Penot, 2010; Andrews and Sanchez, 2011). Positive externalities in the form of increased health and fertility, lower crime rates, and increased community involvement are often associated with higher rates of homeownership (see, e.g. Dietz and Haurin, 2003; DiPasquale and Glaeser, 1999; Glaeser and Sacerdote, 2000, for a summary of the literature). Haurin, Parcel and Haurin (2002) also underline the existence of a positive correlation between homeownership and children test scores, though such a correlation may simply reflect a better environment and geographic stability or the impact of some omitted heterogeneity.

Yet, another strand of the literature has emphasized the potentially negative effects of homeownership on the labor market. What is now conventionally referred to as "Oswald's hypothesis" or "Oswald's conjecture" suggests that higher homeownership may increase the unemployment rate.[2] Indeed, Oswald (1996) suggested that the high unemployment rates observed in OECD countries at the beginning of the nineties were due to increased homeownership. In his original paper, he concluded that an increase of 10 percentage points in homeownership was associated with an increase of 2 percentage points in the unemployment rates. Oswald (1997) additionally suggested that the differences in the unemployment rates across industrialized countries were mainly the consequence of the differences in the levels of homeownership. Finally, he also argued (Oswald, 1999) that reducing homeownership through more efficient rental markets would contribute to reducing the unemployment rates in Europe.

Oswald's conjecture was essentially based upon aggregate data on the labour and housing markets of several OECD countries (Oswald, 1996, 1997). More recently, Nickell et al. (2005) investigated the relation between unemployment rates and homeownership for a set OECD countries over the period 1961–1995. Their specification also included the lagged value of the unemployment rate, a set of variables that proxied the local employment protection laws, the unemployment benefits duration, country dummies, time dummies, and country specific trends. Their results showed no statistically significant relationship between homeownership and unemployment rate. Coulson and Fisher (2009) used aggregate metropolitan (MSA) data to address the potential endogeneity of homeownership rates. State marginal tax rates were used as an instrument since it is used to compute the mortgage interests deduction. Their results support Oswald's conjecture in that an increase in homeownership rates led to higher individual unemployment probability. The empirical evidence using aggregate data finds the opposite, *i.e.* higher ownership rates lead to lower unemployment rates.

Several papers have investigated the relationship between homeownership and unemployment using microeconomic data. The main difficulty in this literature arises from the potential endogeneity between homeownership and outcomes of interest. Indeed, individuals who self-select

[1] Including usufructuaries and free lodgers.

[2] In a recent paper, Beugnot et al. (2018) show that the relationship between unemployment and homeownership rates may be reversed under relatively innocuous assumptions.

into homeownership may have unobserved characteristics that jointly affect the outcome variable. In addition, homeownership may impact differently the employment probability, the wage rate, the duration of unemployment spells, the reemployment wage, the labor market mobility, *etc.* Recently, Munch et al. (2006) and Van Vuuren (2016) considered a model of unemployment duration with self-selection into homeownership while distinguishing the local and non-local labour markets. Van Vuuren (2016) considered a structural model in which becoming unemployed might involve a change of status from homeowner to tenant. Both authors estimated reduced form models with unobserved heterogeneity. In addition, Van Vuuren (2016) estimated the parameters of the structural model using indirect inference and a mixed single-spell proportional hazard model. His results showed that the exit rate from unemployment to the local labour market is significantly larger for home owners but not so to non-local markets. The main drawback of single-spells models is that the distribution of the unobserved heterogeneity may not be representative of its distribution in the entire population. This is not innocuous since individuals enter the labour market at different times. Consequently, using single spells to identify the parameters of the distribution of the unobserved heterogeneity may lead to a misspecified model.

In this paper we address some of the issues faced by most analyses based upon individual data. Hence we jointly model homeownership, labour market transitions and earnings in a dynamic framework. The model we estimate incorporates unobserved heterogeneity to account for self-selection into homeownership and employment, and earnings levels. The individual random effects are allowed to be correlated across all equations, just as are the idiosyncratic error terms. In addition, the latter are also allowed to be autocorrelated. The ownership status is instrumented by the going interest rate on the real estate market. The model is estimated using French panel data for the period ranging from 2004 to 2013. Because we observe many transitions both on the labour and housing markets, we model the past selection mechanisms that led to the initial status appropriately. These features allow to circumvent some of the caveats of the aforementioned single-spells studies. The model we use is dynamic and allows to take into account true and spurious state dependence. It also affords to distinguish very different situations in terms of ownership and labour market statuses. For instance, the behaviour of unemployed tenants and owners can be easily distinguished using the past realization of the stochastic processes. It also possible to isolate the behaviour of employed workers according to their residential status. Importatntly, the impact of homeownership on the reemployment earnings is consistently estimated.

The outline of the paper is as follows. The econometric model is presented in section 2. The data set we use is described in section 3. The estimation results are presented in section 4. The last section concludes.

2. Model Specification

Our empirical analysis focuses on the French labour and housing markets. Housing mobility in France is more or less halfway between Northern Europe (Sweden, Finland, Norway) and the USA, but higher than Germany and Great Britain, and certainly higher than Southern Europe (Spain and Italy) or countries such as Poland and Slovakia (Andrews and Sanchez, 2011). In 2014, as many as 7.3 million individuals, representing 11% of the entire population, moved into a new dwelling. For the same year, 74% of all moves took place within the same region, and of those 38.2% were within the same city (Levy and Dzikowski, 2017).

Mobility is intimately related to the occupational status. In 2013, for instance, among all households who had moved at least once during the past four years, 8.2% where homeowners, 20.8% were tenants in subsidized housing, and 48.5% were regular tenants (Delance and Vignolles, 2017). Average housing tenure was 27.2 years for outright homeowners, 7.5 years for mortgaged homeowners, 13.0 years for tenants in subsidized housing, and only 5.7 for regular tenants (Séverine et al., 2015). During a typical year, as many as 819,000 transactions are recorded which represents approximately 2.37% of the entire housing stock (Arnold, 2016).

The labour market in France is also "average" by European standards. Thus while the unemployment rate reached 9.4% in 2017, it was as low as 3.75% in Germany, 4.16% in Norway and 4.34% in the UK, but as high as 11.21% in Italy and 17,23% in Spain. The relatively dynamic French housing market coupled with a relatively high, but fluctuating, unemployment rate provides an adequate environment to investigate the relationship between ownership status, earnings and unemployment. Endogeneity issues require we do this through formal modelling.

2.1. The Model

Consider a dynamic model that encompasses (home) ownership status (h), employment (e) and wage rate (w). Let x_{jit}, $j \in \{h, e, w\}$, denote a vector of characteristics for individual $i = 1, \dots n$, and where $t = 1, \dots, T$ is a year index. Likewise, let β_j and δ_j be vectors of parameters associated with observed heterogeneity and past realizations of the endogenous and exogenous variables ($\delta_{jk} \in \mathbb{R}$), respectively, for $j \in E = \{h, e, w\}$.

The latent dependent variable y^*_{jit} is given by

$$y^*_{jit} = x'_{jit}\,\beta_j + z_j(y_{it-1}, x_{it-1})'\,\delta_j + r_{jit}, \tag{1}$$

for any $j \in \{h, e\}$, where $z_j(\cdot)$ is a vector containing the realizations of the lagged individual outcomes and characteristics. The observed values of the endogenous variables are denoted as $y_{it} = (y_{hit}, y_{eit}, y_{wit})' \in \{0, 1\} \times \mathbb{R}$.

For individual i at time t, the decision j, $j \in \{h, e\}$, is a binary variable that can be written as

$$y_{jit} = \mathbb{1}[y^*_{jit} > 0], \tag{2}$$

where $\mathbb{1}[.]$ is an indicator function equal to 1 if the event between brackets occurs and zero otherwise.

The log of the yearly earnings at time t is

$$y_{wit} = x'_{wit}\,\beta_w + z_w(y_{it-1}, y_{hit})'\,\delta_w + r_{wit}, \tag{3}$$

where δ_w, β_w and $z_w(\cdot)$ are defined similarly as above, save for $z_w(\cdot)$ which also depends on the contemporaneous value of the ownership status.

2.2. Stochastic Specification

The error term, r_{jit}, is written as the sum of a time-invariant outcome-specific unobserved individual component, α_{ij}, and a contemporaneous residual, u_{jit}:

$$r_{jit} = \alpha_{ij} + u_{jit}. \tag{4}$$

As is customary, the individual effects, α_{ij}, are assumed to be independent of the observable characteristics, x_i, to be normally distributed with mean zero and variance $\sigma^2_{\alpha_j}$, $j \in E$, and to be independent across $i \in \{1, \dots, n\}$ and $j \in \{h, e, w\}$.

The contemporaneous error term is further assumed to satisfy the following independence assumptions: $u_{jit} \perp\!\!\!\perp x_i$, and $u_{jit} \perp\!\!\!\perp u_{j'i't'}$. On the other hand, we allow u_{jit} to be autoregressive (but stationary, for $t > 0$).[3] Thus we write

$$u_{jit} = \rho_j\, u_{jit-1} + \epsilon_{jit}, \tag{5}$$

where $\alpha_{ij} \perp\!\!\!\perp \epsilon_{j'it}$, for all $j, j' \in E$, and $\epsilon_{jit} \perp\!\!\!\perp u_{j'it'}$, for all $t' < t$ and $j' \in E$. The time-dependency allows to measure the impact of a shock at t on individual outcomes at $t + 1$.

[3] Hyslop (1999) makes a similar assumption in the context of a single equation model.

In any dynamic panel data model with random effects and left censoring, the initial conditions much be modelled appropriately. We follow Heckman (1981) and write the initial conditions as a reduced-form specification which allows to correlate the error terms at $t = 0$ with those at $t > 0$. Thus write the system of equations at time t_0 as follows:

$$\begin{aligned} y_{hi0} &= \mathbb{1}[x'_{hi0}\,\beta^0_h + r_{hi0} > 0], \\ y_{ei0} &= \mathbb{1}[x'_{ei0}\,\beta^0_e + r_{ei0} > 0], \\ y_{wi0} &= x'_{wi0}\,\beta^0_w + r_{wi0}, \end{aligned} \tag{6}$$

where x_{ji0} is the vector of initial characteristics and r_{ji0} is an error term similarly defined as above ($j \in E = \{h, e, w\}$). Likewise β^0_j is an appropriately dimensioned vector of parameters ($\beta^0_j \in \mathbb{R}^{p_j}$, where $p_j \in \mathbb{N}^*$). The initial error terms are assumed to satisfy the following two assumptions:

$$r_{ji0} \sim \text{N}(0, \sigma^2_{j0}), \text{ where } j = h, e, w; \tag{7}$$

$$\epsilon_{jit} \sim \text{N}(0, \sigma^2_{\epsilon_j}). \tag{8}$$

As argued by Heckman (1981), it is likely that r_{ji0} is correlated with $r_{j'it}$, $j' \in E$, for $t = 1, \ldots,$ T. Hence, let $\rho_{\alpha_j\alpha_k}$ denote the correlation between the random effects α_{ij} and α_{ik}, specific to equations j and k, respectively, $j, k \in E$. Let ρ_{jk} denote the correlation between the idiosyncratic terms ϵ_{jit} and ϵ_{kit}, for all $t = 1, \ldots, T$ and $i \in \{1, \ldots, n\}$.

For the two binary equations, y^*_{hit} and y^*_{eit}, this specification is equivalent to a dynamic probit model with random effects. Additional assumptions must be made to insure the model is identified. Indeed, because the two dependent variables are dichotomous we must normalize the variance of the corresponding residuals as follows:

$$\sigma^2_{\alpha_j} + \sigma^2_{u_j} = 1,$$

and

$$\text{var}(r_{ji0}) = \sigma^2_{j0} = 1, \text{ for } j = h, e.$$

It can be shown that for $j \in E, t > 1$

$$var(u_{jit}) = \sigma^2_{u_j} = \frac{\sigma^2_{\epsilon_j}}{(1 - \rho^2_j)},$$

2.3. Model Estimation

For each individual and time period, we observe the realization of the variables $y_{jit} \in \{0; 1\}$, for $j = h, e$, as well as the log-earnings y_{wit} ($i = 1, \ldots, n$ and $t = 0, 1, \ldots, T$). The contribution of individual i to the likelihood function is [4] :

$$L_i(\theta) = \int_{A_i} \phi(r; \Omega)\, d\,r, \tag{9}$$

where $\phi(.; \Omega)$ is the probability density function of the normal distribution with mean zero and variance-covariance matrix Ω. The integration is computed over the error terms r that are compatible with the trajectories of the endogenous variables of individual i:

$$A_i = \left\{ r \in \mathbb{R}^{3(T+1)} : r = (r_{h0}, r_{e0}, r_{w0}, r_{h1}, ..., r_{hT}, ..., r_{w1}, ..., r_{wT}) \text{ and } a_{jit} \le r_{jt} \le b_{jit} \right\}.$$

[4] The order of integration is at most 30 because we use annual data over the period 2004 to 2013.

The domain of integration depends on the realizations of the dependent variables, the vector of explanatory variables, and the vector of parameters, θ (see Appendix D). The latter can be estimated by maximizing the logarithm of the simulated likelihood:

$$\hat{\ell}_{N,H}(\theta) = \sum_{i=1}^{N} \ln \left(\frac{1}{H} \sum_{h=1}^{H} \tilde{p}(x_i; u_i^h; \theta) \right), \tag{10}$$

where the drawings u_i^h are specific to the individual i and are $i.i.d.$ $(i = 1, \dots, n)$ (see Appendix E). These are drawn in such a way as to avoid rejection.[5] The main difficulty when making the draws is to account for the fact that the endogenous variables are both qualitative and continuous (Chang, 2009). The simulated maximum likelihood estimator of θ can be obtained maximizing the function (10). The SML estimator will be consistent and efficient as $\frac{\sqrt{N}}{H} \longrightarrow 0$ when $N \longrightarrow +\infty$ and $H \longrightarrow +\infty$ (cf., for instance, Gouriéroux and Monfort, 1991, 1993, 1997; Kamionka, 1998; Edon and Kamionka, 2008; Gilbert et al., 2011; Kamionka and Ngoc, 2016).

In practice, the number of draws is set to $H = 30$. Several authors have stressed that the SML estimator is near consistent even for relatively a small H (about 30). Indeed, some specifications were estimated for larger values of H without modifying significantly the estimation results (cf. Kamionka and G. Lacroix (2008)).

2.4. The Wooldridge [2005] Approach to the Initial Conditions Problem

Naturally, the initial observations at time t_0 do not correspond to the starting time of the data generating process. Hence, the initial state $y_{i0} = (y_{hi0}, y_{ei0}, y_{wi0})'$ is clearly not independent of the individual effects $\alpha_i = (\alpha_{ih}, \alpha_{ie}, \alpha_{iw})'$. Wooldridge (2005) suggests we consider the distribution of the random effects $\alpha_i' = (\alpha_{ih}, \alpha_{ie}, \alpha_{iw})$ conditionally to y_{i0} and, possibly, on a set of exogenous explanatory variables. When this conditional distribution is normally distributed, and given our previous assumptions about the error terms, the likelihood function boils down to the product of integrals defined over multivariate normal density functions.

In practice, it is reasonable to assume that ownership status, employment and earnings are generated by equations (1) to (3). As above, we assume that the error term u_{jit} follow the same autoregressive structure as in equations (5) and (8). We further assume that the conditional distribution of α_{ij} is a normally distribution:

$$\alpha_{ij} \mid y_{i0}, x_i \sim N(\lambda_{j0} + y_{hi0}\,\lambda_{jh} + y_{ei0}\,\lambda_{je}, \sigma^2_{\alpha_j}), \tag{11}$$

where λ_{j0}, λ_{jh} and λ_{je} are some real parameters to be estimated and y_{i0}=$(y_{hi0}, y_{ei0}, y_{wi0})'$. It turns out the constant λ_{j0} cannot be separately identified from the one embedded in β_j. Thus, without lost of generality, we set $\lambda_{j0} = 0$.

Let $\Sigma_1 = var(U_i)$, where $U_i = (U_{hi}', U_{ei}', U_{wi}')'$ and $U_{ji} = (u_{ji1}, \dots, u_{jiT})'$. Let $E_i = (E_{hi}', E_{ei}', E_{wi}')'$ denote the vector of unobserved heterogeneity terms, where $E_{ji} = \alpha_{ij}\,\mathbb{1}_T$ for $j \in E$. As per our previous assumptions, the variance-covariance matrix $\Omega_1 = var(U_i + E_i)$ is the same as that presented in the previous section. The contribution of individual i to the conditional likelihood function is

$$L_i(\theta) = \int_{A_i'} \phi(r; \Omega_1)\, dr,$$

where $\phi(.; \Omega_1)$ is the probability density function of a normal distribution with mean zero and variance-covariance matrix Ω_1. The integration is computed over the set

$$A_i' = \left\{ r \in \mathbb{R}^{3T} : r = (r_{h1}, \dots, r_{hT}, \dots, r_{w1}, \dots, r_{wT}) \text{ and } a_{jit} \le r_{jt} \le b_{jit} \right\}.$$

[5]The term $\tilde{p}(x_i; u_i^h; \theta)$ in equation (10) is defined in equation $(E.3)$ of Appendix E.

Once again, the domain of integration depends on the realizations of the dependent variables, the explanatory variables and the vector of parameters (see Appendix D).

We can estimate the vector of parameters, θ, by maximizing a simulated likelihood function similar to the one defined by the expression (10) using draws $u_i^h = (u_{i1}^h, \ldots, u_{i3T}^h)'$ constructed by a method similar to the one presented in Appendix E and by substituting Ω_1 for Ω.

3. Data

Our data are drawn from the French sample of the EU-SILC data set (European Union - Status on Income and Living Conditions). The French survey itself is based upon *L'enquête statistique sur les ressources et conditions de vie* (Dispositif SRCV). It is conducted from May to June every year since 2004 and is available until 2013. Over 9,091 households (26,353 individuals) were surveyed in 2004 and as many as 11,131 in 2013 (26,353 individuals).

The SRCV is a rotating panel. Each year since 2005, approximately 1/9 of all households are replaced by a new rotating group drawn from the list of all dwellings located in mainland France. Thus approximately 1/9 of those surveyed in 2004 were still in the panel in 2012. By virtue of its rotating design, the SRCV yields an unbalanced panel. Yet, it is representative of all the regular households living in metropolitan France and contains detailed information on income, living conditions, employment and ownership statuses, wealth, *etc.* While all individuals over sixteen years of age are surveyed, we restrict our sample to those between 20 and 56 since they are the most likely involved with employment and ownership decisions. This exclusion restriction yields a sample of over 30,077 individuals, and all are observed for at least two consecutive years.

Table 1 presents the main characteristics of the sample. The second column focuses on individual characteristics at entry into the panel (t_0) while the third column reports the statistics for the initial sample (*i.e.* in 2004). The second column is thus based on the incoming rotating groups. As time unfolds, the proportion of homeowners decreases by three percentage points (from 57.34% to 54.24%). The initial percentage of homeowners is marginally higher than in the entire population for the same year (56.6%, see Table 2). The difference is presumably due to the fact that we exclude seniors among whom there are proportionately fewer homeowners. The employment rate and the level of earnings are almost identical across columns. The incoming groups tend to be slightly more educated as they are proportionately more likely to hold a university degree. Likewise they are fewer to live in Paris, they are naturally younger, and fewer are married.

According to Table 2, the proportion of homeowners at the national level has remained fairly constant between 2004 and 2013. It is much lower than the European average of 65% (INSEE, 2010) but nevertheless much higher than that in Germany for those years (46%). Two of the main determinants of the ownership status are the going mortgage interest rates and housing prices. Figure 1a depicts the evolution of both variables over our sample period. As shown, the interest rate has decreased by nearly two percentage points, save for the 2006-2008 period during which they increased by one percentage point.[6] As of 2008, the interest rates have decreased steadily until 2015 and by as much as three percentage points from their peak of 5.7%. As expected, housing prices and interest rates are more or less inversely related, save for the period between 2006–2008 during both were increasing.[7] Housing prices have increased by as much as 35% between the first quarter of 2004 and their peak value in 2012. As shown in Figure 1b, the mean monthly duration of contracted mortgages has increased by over 19.9% (from 196 months to 235 months) and follows closely the fluctuations in the interest rate. The fluctuations on the housing market have occurred while the labour market was relatively depressed as the unemployment rates ranged between 7.4% to 10.2%. A lengthy period of observation provides much needed variations in the

[6]The figure plots the "Taux effectif des prêts immobiliers à taux fixes accordés aux particuliers" of the Banque de France.

[7]Indice des prix des logements anciens en France métropolitaine, Ensemble, Base 100 au premier trimestre de 2010, Série CVS, Insee.

exogenous variables which help identify the slope parameters of the model. In addition, observing individual households for up to ten years helps identify the nuisance parameters associated with the unobserved heterogeneity components. Finally, the yearly entry of rotating groups within our sample helps identify the parameter estimates of the initial conditions.

4. Results

Recall that the model we estimate includes three endogenous variables: ownership and employment statuses, equation (2), and earnings, equation (3). The stochastic specification is given by equations (4) and (5). The initial conditions in equation (6) are modelled according to two different approaches proposed by Heckman (1981) and more recently by Wooldridge (2005). The estimation results for each approach are presented in separate tables. Both tables run over four pages and are organised as follows: The first three pages report the parameter estimates of the ownership, employment and earnings equations, respectively. Each one is separated into three panels that report the slope parameters, the state dependence parameters, and the initial conditions parameters, respectively. The fourth page focuses on the stochastic specification.

4.1. Results Based on Wooldridge's Specification

We begin with the specification based on Wooldridge's approach in Table 3. The table presents three different specifications. Specification (1) includes lagged ownership and employment in the contemporaneous ownership and employment equations. Specification (2) adds the lagged yearly mean interest rate in the ownership equation. We estimate separate parameters for owners and tenants. We also include a lagged interaction term between unemployment and ownership in the employment equation. Finally, Specification (3) adds a series of lagged interaction terms between sex, unemployment and ownership in the employment equation.

4.1.1. Ownership Status

We begin our discussion of the results by focusing on each equation in turn. According to the parameter estimates, not surprisingly, married couples are more likely to own their home. This is consistent with a large literature that concludes likewise (see Haurin et al., 1996). Foreigners are less likely to be homeowners perhaps because they are less likely to inherit a property. More educated individuals are also more likely to be homeowners. This is not surprising since the highly educated have larger expected earnings and can thus more easily obtain a loan. Relative to the 30–39 age group, those in the 40–49 age group are more likely to own their home whereas those in the 50+ age group are less likely so. This is rather surprising but it may indicate that older individuals who do not yet own a house may find it hard to get a loan since retirement may be nearing.

The parameter estimates are consistent with sizeable state dependence: being employed and homeowner in the previous year increases the likelihood of being a homeowner the next. The higher the mortgage interest rate of the previous period the less likely a tenant will become a homeowner the next as expected. Conversely, those who already own a house are not affected in the short run by an increase in the interest rate.

The estimates of the last panel refer to equation (11), *i.e.* to λ_{jh} and λ_{je}. All estimates are significantly different from zero. In particular, they show that a homeowner at t_0 who becomes a tenant at t is less likely to be a homeowner at a later date. Likewise, an employed homeowner at t_0 who becomes unemployed at t is also less likely to become a homeowner at a later date if he/she became a tenant once unemployed.

4.1.2. Employment Status

The second part of Table 3 focuses on the employment dynamics. Most statistically significant slope parameters have the expected signs. Hence, women are less likely to work, married individuals are more likely so, and so are the more educated. Foreigners are found to be less likely to work, possibly the result of discrimination on the labour market. Employment increases with age until the age of 40–49 then it decreases significantly for those above 50 years of age. This is consistent with the evolution of the observed employment rates by age groups in France. As expected, the employment probability is inversely related to the contemporaneous unemployment rate.

According to the first specification, the employment status is largely state dependent. Indeed, employed individuals at time $t-1$ are much more likely to be employed at time t. Past ownership status is also positively linked to the contemporaneous conditional probability to be employed. Past ownership may be informative of the employability of a given individual. Specification (2) investigates the link between current employment and lagged unemployment by homeowners. Homeowners are less likely to be employed if they were unemployed the previous year. The data are thus consistent with previous empirical work according to which homeowners have lengthier unemployment spells. The last specification investigates this further by distinguishing between tenants and owners of different age groups. For a given age group, unemployed homeowners at $t-1$ are less likely to be employed than tenants. Thus, irrespective of age, homeowners have longer unemployment spells.

According to the last panel, the initial statuses are very informative of the unobserved heterogeneity distribution. Thus, employed individuals at time t_0 who become unemployed at t have unobserved characteristics that will impede an eventual transition into employment.

4.1.3. Earnings Equation

The next section of Table 3 focuses on the (log) earnings equation. All the slope parameters are consistent with standard human capital results: Women and foreigners earn less, presumably as a result of discrimination on the labour market in both cases, education increases earnings significantly, and age is positively related to earnings.

According to the first specification, homeowners have larger conditional expected earnings of approximately 16%. This is consistent with most of the empirical literature. This result indicates that ownership is informative of the earnings distribution. In other words, it is compatible with the fact that one is more likely to obtain a loan the larger the expected discounted value of the future stream of earnings.

State dependence is once again investigated by including a series of lagged (interacted) variables. Specifications (2) and (3) yield almost identical parameter estimates. These show that unemployed homeowners at $t-1$ must contend with lower earnings upon becoming employed at t. Recall that homeowners have longer unemployment spells. The earnings penalty of approximately 6% is consistent with the latter having revised their reservation wage downward. Finally, note that those who are employed at $t-1$ may expect an increase in earnings of approximately 7% the following year, irrespective of their ownership status.

In all three specifications, the initial conditions are informative of the unobserved heterogeneity. For instance, an individual who is initially employed and who eventually becomes unemployed faces lower conditional expected earnings relative to someone who was initially unemployed. Similarly, homeowners at t_0 who eventually become tenant will face lower conditional earnings when and if they eventually become homeowners anew.

4.1.4. Stochastic Specification

The last section of Table 3 reports the parameter estimates of the stochastic specification. According to the first panel, the variance of the random effect specific to the ownership status is relatively large, indicating important unobserved heterogeneity among homeowners. The variance of the random effect of the (log) earnings equation is relatively small which suggests there is little (conditional) heterogeneity. Except for Specification 3, there is little unobserved heterogeneity in the employment equation.

The unobserved heterogeneity components specific to the employment and earnings equations, not surprisingly, are positively correlated in all the specifications. Workers are thus a self-selected subset of the population. The same holds for the correlation between the random effects of the ownership and earnings equations in Specifications (2) and (3). Interestingly, we do not observed such correlation between the individual random components of the ownership and the employment equations.

Our specification is flexible enough to model the autocorrelation of the error terms of each endogenous variable. As shown, the error term of the ownership and employment equations are negatively autocorrelated, implying that a negative shock in a given year is likely to affect positively the outcome the following year. Conversely, the error term of the earnings equation is positively and significantly autocorrelated in all three specifications. Finally, our parameter estimates suggests the error terms of all three equations are weakly correlated at best.

4.2. Results Based on Heckman's Specification

The model was also estimated using Heckman (1981)'s approach to account for the initial conditions problems.[8] The results are reported in Table 4. Overall, both approaches yield qualitatively similar results although the slope parameters are more contrasted in the ownership and employment equations under Heckman's approach. For instance, females are found to be less likely to be homeowners, although the marginal effect is relatively small. More importantly, as in Wooldridge's specification, unemployed homeowners are much less likely to be employed next year and they have to contend with significantly lower re-employment earnings.

The parameter estimates of the initial equations are reported in the second panel of the table. Nearly all are statistically significant. Thus education has a positive impact on employment, earnings and ownership. Females and foreigners fare worse on the labour and housing markets. The differential effects of age on each equation is consistent with conventional results. According to our results, as individuals age they are more likely to own a house, reap higher earnings, and retire or face higher unemployment.

The parameters of the stochastic specification are reported in the last panel of the table. Interestingly, the correlations between the initial error terms are large, positive and statistically significant. The cross-correlations between the individual effects are also positive, large and significant. Moreover, the correlations between the initial error terms and the contemporaneous ones are also positive and statistically significant. Unfortunately, the proportion of the covariance of the initial error terms that is due to the unobserved heterogeneity cannot be identified. The results nevertheless suggest that it is rather large. This is confirmed by the estimates of the correlations of the idiosyncratic terms which are all weak and generally not statistically significant, save for the correlation between employment and earnings. Thus, given the observed individual characteristics, those who are observed in a given state at t are more likely to be observed in the same state at $t + 1$. This follows from the fact that the initial conditions are informative about the distribution of unobserved heterogeneity.

5. Conclusion

The motivation of this paper stems from the lack of consensus in the literature about the links between homeownership, earnings and unemployment. Conflicting results may arise as a result of failing to properly account for the endogeneity of homeownership and performances on the labour market. In this paper we jointly model homeownership, labour market transitions and earnings. The model incorporates unobserved heterogeneity to account for self-selection into homeownership and employment. Individual random effects are allowed to be correlated across all equations, just as are the idiosyncratic error terms. In addition, the latter are also allowed to be autocorrelated. Our model is very flexible as it does not rely on the usual conditional proportional hazard models (Munch et al., 2006; Van Vuuren, 2016).

[8] See Appendices B and D for the details.

Our results show that when unemployed, homeowners have longer spells than tenants. On the other hand, they also enjoy longer employment spells, presumably because they face higher mobility costs. While homeowners must also contend with lower reemployment earnings, the "penalty" is relatively small. Perhaps, firms reward them for their greater employment stability. Our results also stress the importance of unobserved heterogeneity in explaining the observed transitions on the labour and housing markets, and the relationship between earnings and the latter two. Failure to properly account for this is likely to yield biased parameter estimates.

Implicit in our analysis is the assumption that homeowners and tenants behave similarly on the labour market. Yet, the search behaviour of homeowners is presumably quite different due to their mobility costs (Beugnot et al., 2018). In addition, ownership status may be quite sensitive to the present value of future local property tax streams. We intend to address both issues in future research.

References

Andrews, D., Sanchez, A. C., 2011. The evolution of homeownership rates in selected OECD countries: Demographic and public policy influences. OECD Journal: Economic Studies 2011-1.

Arnold, C., 2016. Le parc de logements en france au 1er janvier 2016.

Beugnot, J., Charlot, O., Lacroix, G., Oct 2018. Does promoting homeownership always damage labour market performances? Journal of Economics.
URL `https://doi.org/10.1007/s00712-018-0637-x`

Chang, S.-K., 2009. Simulation estimation of two-tiered dynamic panel tobit models with an application to the labor supply of married women. Journal of Applied Econometrics 26 (5), 854–871.
URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.1141`

Coulson, N. E., Fisher, L. M., 2009. Housing tenure and labor market impacts: The search goes on'. Journal of Urban Economics 65, 252–264.

Delance, P., Vignolles, B., 2017. Ça déménage ? la mobilité résidentielle et ses déterminants.

Dietz, R. D., Haurin, D., 2003. The social and private micro-level consequences of homeownership. Journal of Urban Economics 54, 401–450.

DiPasquale, D., Glaeser, E., 1999. Incentives and social capital: Are homeowners better citizens? Journal of Urban Economics 45 (2), 354–384.

Edon, C., Kamionka, T., 2008. Modélisation dynamique de la participation au marché du travail des femmes en couple. Annales d'économie et de Statistique 86, 77–108.

Edon, C., Kamionka, T., 2014. Dynamic labor market behavior of married women with endogenous participation, unemployment, working time and wage. Mimeo, CREST.

Geweke, J., 1991. Efficient simulation from the multivariate normal and student-t distributions subject to linear constraints'. In: Computing Science and Statistics: Proceedings of the twenty-third symposium on the interface. Computing science and statistics, American statistical association, Alexandria, pp. 571–578.

Gilbert, L., Kamionka, T., Lacroix, G., 2011. The impact of government-sponsored training programs on the labor market transitions of disadvantaged men. In: Advances in Econometrics. Intech, pp. 953–307, iSBN 978.

Glaeser, E., Sacerdote, B., 2000. The social consequences of housing. Journal of Housing Economics 9, 1–23.

Gouriéroux, C., Monfort, A., 1991. Simulated based in econometric models with heterogeneity. Annales d'Économie et de Statistique 20-21, 69–107.

Gouriéroux, C., Monfort, A., 1993. Simulated based inference : A survey with special reference to panel data models. Journal of Econometrics 59, 5–33.

Gouriéroux, C., Monfort, A., 1997. Simulation-Based Econometric Methods. Oxford University Press, Core Lecture Series , Oxford.

Hajivassiliou, V., McFadden, D., Ruud, P., 1992. Simulation of the multivariate normal orthant probabilites : Methods and programs. Cowles Foundation, Discussion Paper 1021, Yale University.

Haurin, D., Hendershott, S., Wachter, S., 1996. Wealth accumulation and housing choices of young households. Journal of Housing Research 8, 137–154.

Havet, N., Penot, A., 2010. Does home ownership harm labour market performances? A survey. Working Paper GATE 2010-12.
URL https://halshs.archives-ouvertes.fr/halshs-00491074

Heckman, J. J., 1981. The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In: McFadden, C. M. D. (Ed.), Structural Analysis of Discrete Data. MIT Press.

Hyslop, D. R., 1999. State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women. Econometrica 67 (6), 1255–1294.
URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-0262.00080

INSEE, 2010. Tableaux de l'économie française, Paris.

INSEE, 2016. Tableaux de l'économie française, Paris.

Kamionka, T., 1998. SML estimation in transition models. The Econometrics Journal 1, C129–C153.

Kamionka, T., G. Lacroix, G., 2008. Assessing the external validity of an experimental wage subsidy'. Annales d'économie et de Statistique 91-92, 357–384.

Kamionka, T., Ngoc, X. V., 2016. Insertion des jeunes sur le marché du travail, diplôme et quartier d'origine: Une modélisation dynamique'. Revue économique 67 (3), 463–494.

Keane, M., 1994. A computationally practical simulation estimator for panel data. Econometrica 62, 95–116.

Levy, D., Dzikowski, C., Juin 2017. En 2014, un quart de la population qui déménage change de département. INSEE première.

Munch, J. R., Rosholm, M., Svarer, M., 2006. Are home owners really more unemployed? Economic Journal 116, 911–1013.

Nickell, S., Nunziata, L., Ochel, W., 2005. Unemployment in the OECD since the 1960s : What do we know? The Economic Journal 115, 1–27.

Oswald, A. J., 1996. A conjecture on the explanation of high unemployment in the industrialized nations: Part 1. Research Papers , University of Warwick, Warwick.

Oswald, A. J., 1997. The missing piece of the unemployment puzzle. mimeo, An inaugural Lecture.

Oswald, A. J., 1999. The housing market and Europe's unemployment: a non-technical paper. Working paper, University of Warwick, Warwick.

Séverine, A., Crusson, L., Donzeau, N., Rougerie, C., 2015. Les conditions de logement fin 2013 : Premiers résultats de l'enquête logement. Insee Première - Les conditions de logement fin 2013.

Van Vuuren, A., 2016. Using a structural-form model to analyze the impact of home ownership on unemployment duration. Journal of Applied Econometrics.

Wooldridge, J. M., 2005. Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. Journal of Applied Econometrics 20, 39–54.

Table 1: Sample Characteristics

	First Observation	Status in 2004
Dependent variables		
Homeowner	54.24	57.34
Employed	79.48	78.85
Earnings (euros)	19 020.68	18 723.26
Education level		
High-School or less	51.48	57.17
Post-Secondary	17.76	16.10
University	30.76	26.73
Gender		
Women	52.32	53.06
Men	47.68	46.94
Urban area		
Paris	14.57	16.69
$200000 \leq$ pop < 2 millions	23.80	22.99
$100000 \leq$ pop < 200000	5.74	5.53
$50000 \leq$ pop < 100000	7.07	7.02
$20000 \leq$ pop < 50000	5.78	5.24
$10000 \leq$ pop < 20000	5.25	4.66
$5000 \leq$ pop < 10000	4.23	4.56
pop < 5000	6.91	7.55
Rural township	26.65	25.77
Household Characteristics		
$20 \leq$ Age ≤ 29	20.81	14.97
$30 \leq$ Age ≤ 39	28.43	30.61
$40 \leq$ Age ≤ 49	30.23	32.48
$50 \leq$ Age	20.53	21.94
Married	53.18	61.72
Citizenship		
French	94.25	93.77
Other	5.75	6.23
Number of individuals	30,077	9,678

Note : SRCV 2004-2013. Percentages.

Table 2: Proportion of Homeowners (%)

2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
56.6	57.0	57.3	57.5	57.5	57.6	57.7	57.7	57.7	57.7	57.7

Excluding usufructuaries and free lodgers. Source: Crusson and Arnault, Insee (2015) and Insee (2010).

Table 3: Home Ownership, Employment and Earnings
Wooldridge Specification for Initial Conditions

Ownership Status

	(1)	(2)	(3)
	SLOPE PARAMETERS		
Intercept	-1.4447***	-1.1199***	-1.1310***
	(0.0731)	(0.1859)	(0.1861)
Female		-0.0089	-0.0087
		(0.0197)	(0.0201)
Married	0.2429***	0.2479***	0.2481***
	(0.0204)	(0.0204)	(0.0204)
Post-Secondary	0.1485***	0.1511***	0.1508***
	(0.0287)	(0.0275)	(0.0275)
University	0.2612***	0.2680***	0.2677***
	(0.0226)	(0.0227)	(0.0227)
Foreign-Born	-0.2423***	-0.2491***	-0.2491***
	(0.0474)	(0.0451)	(0.0451)
Age 20–29	-0.0410	-0.0426	-0.0430
	(0.0262)	(0.0283)	(0.0283)
Age 40–49	-0.1220***	-0.1220***	-0.1216***
	(0.0256)	(0.0250)	(0.0250)
Age 50+	-0.1350***	-0.1365***	-0.1363***
	(0.0332)	(0.0294)	(0.0294)
Unemployment rate		0.0134	0.0134
		(0.0122)	(0.0122)
	STATE DEPENDENCE		
$Employed_{t-1}$	0.2977***	0.1771***	0.1831***
	(0.0425)	(0.0398)	(0.0394)
$Owner_{t-1}$	3.5383***	2.9299***	2.9327***
	(0.0349)	(0.2003)	(0.2003)
$(Renter \times Interest)_{t-1}$		-0.0698**	-0.0693**
		(0.0291)	(0.0291)
$(Owner \times Interest)_{t-1}$		0.0819*	0.0818*
		(0.0429)	(0.0428)
	INITIAL CONDITIONS		
$Employed_0$	-0.0211	-0.0552***	-0.0532***
	(0.0159)	(0.0147)	(0.0146)
$Owner_0$	-0.1668***	-0.1763***	-0.1757***
	(0.0166)	(0.0160)	(0.0159)

(*) Significant at 10%. (**) Significant at 5%. (***) Significant at 1%.

Table 3: Home Ownership, Employment and Earnings (Continued)
Wooldridge Specification for Initial Conditions

Employment

	(1)	(2)	(3)
	SLOPE PARAMETERS		
Intercept	0.6709*** (0.0599)	1.0067*** (0.1009)	2.3180*** (0.0908)
Female		-0.2034*** (0.0183)	-0.2047*** (0.0174)
Married	0.0330* (0.0186)	0.0627*** (0.0178)	0.0542*** (0.0178)
Post-Secondary	0.1378*** (0.0239)	0.1629*** (0.0234)	0.1595*** (0.0232)
University	0.3128*** (0.0205)	0.3356*** (0.0204)	0.3352*** (0.0202)
Foreign-Born	-0.1832*** (0.0369)	-0.2098*** (0.0349)	-0.2124*** (0.0348)
Age 20–29	0.0413 (0.0256)	0.0037 (0.0259)	-0.0105 (0.0259)
Age 40–49	0.0940*** (0.0213)	0.1013*** (0.0211)	0.2935*** (0.0238)
Age 50+	-0.1984*** (0.0238)	-0.1812*** (0.0229)	-0.0048 (0.0251)
Unemployment rate		-0.0244** (0.0093)	-0.0227** (0.0094)
	STATE DEPENDENCE		
$Employed_{t-1}$	1.2979*** (0.0326))	1.4380*** (0.0393)	
$Owner_{t-1}$	0.1828*** (0.0307)		
$(Unemployed{\times}Owner)_{t-1}$		-0.3822*** (0.0479)	
$Female \times (Unemployed{\times}Owner)_{t-1}$		0.0916* (0.0515)	
$< 40\ yo \times (Unemployed{\times}Owner)_{t-1}$			-1.4513*** (0.0500)
$\geq 40\ yo \times (Unemployed{\times}Owner)_{t-1}$			-2.0713*** (0.0468)
$< 40\ yo \times (Unemployed{\times}Tenant)_{t-1}$			-1.2939*** (0.0409)
$\geq 40\ yo \times (Unemployed{\times}Tenant)_{t-1}$			-1.8649*** (0.0477)
	INITIAL CONDITIONS		
$Employed_0$	-0.4802*** (0.0114)	-0.3662*** (0.0132)	-0.3455*** (0.0128)
$Owner_0$	-0.0114 (0.0159)	-0.1207*** (0.0102)	-0.1019*** (0.0103)

(*) Significant at 10%. (**) Significant at 5%. (***) Significant at 1%.

Table 3: Home Ownership, Employment and Earnings (Continued)
Wooldridge Specification for Initial Conditions

Earnings

	(1)	(2)	(3)
	SLOPE PARAMETERS		
Intercept	10.0100****	9.0223***	9.0157***
	(0.0257)	(0.0234)	(0.0234)
Female		-0.3918***	-0.3917***
		(0.0096)	(0.0096)
Post-Secondary	0.2147****	0.2272***	0.2267***
	(0.0123)	(0.0127)	(0.0127)
University	0.5106***	0.5177***	0.5173***
	(0.0103)	(0.0105)	(0.0105)
Foreign-Born	-0.0908***	-0.0930***	-0.0927***
	(0.0226)	(0.0230)	(0.0230)
Age 20–29	-0.0537***	-0.0625***	-0.0631***
	(0.0138)	(0.0120)	(0.0120)
Age 40–49	0.0950***	0.0968***	0.0994***
	(0.0102)	(0.0092)	(0.0092)
Age 50+	0.1303***	0.1473***	0.1519***
	(0.0122)	(0.0113)	(0.0113)
$Owner_t$	0.1605***		
	(0.0149)		
	STATE DEPENDENCE		
$(Unemployed \times Owner)_{t-1}$		-0.0619**	-0.0616**
		(0.0307)	(0.0306)
$Female \times$			
$(Unemployed \times Owner)_{t-1}$		0.0341	0.0324
		(0.0336)	(0.0335)
$Employed_{t-1}$		1.0777***	1.0788***
		(0.0139)	(0.0139)
	INITIAL CONDITIONS		
$Employed_0$	-0.6257***	-0.2938***	-0.2903***
	(0.0051)	(0.0075)	(0.0075)
$Owner_0$	-0.0177**	-0.0714***	-0.0710***
	(0.0082)	(0.0056)	(0.0056)

(*) Significant at 10%. (**) Significant at 5%. (***) Significant at 1%.

Table 3: Home Ownership, Employment and Earnings (Continued)
Wooldridge Specification for Initial Conditions

Residuals

	(1)	(2)	(3)
Stochastic Specification $r_{jit} = \alpha_{ij} + u_{jit}$, $u_{jit} = \rho_j\, u_{jit-1} + \epsilon_{jit}$			
	Standard errors of individual effects (α_{ij})		
$\sigma_{\alpha_h} = \frac{\exp(v_h)}{1+\exp(v_h)}$	3.5084	2.3355***	2.3520***
	(10.7230)	(0.2287)	(0.2400)
$\sigma_{\alpha_e} = \frac{\exp(v_e)}{1+\exp(v_e)}$	-0.4155***	-0.0830	0.0325
	(0.0457)	(0.0603)	(0.0383)
$\sigma_{\alpha_w} = \exp(v_w)$	-0.2304***	-0.0669***	-0.0674***
	(0.0176)	(0.0229)	(0.0228)
	Correlations between individual effects (α_{ij})		
$\rho_{\alpha_h\alpha_e} = \tanh(c_{he})$	0.3840	3.4591	3.3884
	(4.3601)	(5.6455)	(5.2960)
$\rho_{\alpha_h\alpha_w} = \tanh(c_{hw})$	1.2511	0.5866***	0.5935***
	(31.4661)	(0.1506)	(0.1602)
$\rho_{\alpha_e\alpha_w} = \tanh(c_{ew})$	0.8416***	0.4777***	0.4857***
	(0.0233)	(0.0217)	(0.0227)
	Auto-Correlation of error terms (u_{jit})		
$\rho_h = \tanh(d_h)$	-0.1497***	-0.1438***	-0.1442***
	(0.0252)	(0.0227)	(0.0227)
$\rho_e = \tanh(d_e)$	-0.1187***	-0.1644***	-0.1605***
	(0.0212)	(0.0197)	(0.0196)
$\rho_w = \tanh(d_w)$	0.5193***	0.4200***	0.4192***
	(0.0047)	(0.0109)	(0.0109)
	Correlations between error terms (ϵ_{jit})		
$\rho_{he} = \tanh(f_{he})$	0.0536***	0.0098	0.0185
	(0.0208)	(0.0221)	(0.0206)
$\rho_{hw} = \tanh(f_{hw})$	-0.0553***	0.0056	0.0062
	(0.0114)	(0.0093)	(0.0091)
$\rho_{ew} = \tanh(f_{ew})$	-0.0108	0.0712***	0.0698***
	(0.0072)	(0.0074)	(0.0074)
	Standard error of (log) earnings (u_{wit})		
$\sigma_{u_w} = \exp(f)$	-0.4393***	-0.5799***	-0.5803***
	(0.0034)	(0.0071)	(0.0071)
Number of obs.	23 041	23 041	23 041

(*) Significant at 10%. (**) Significant at 5%. (***) Significant at 1%.

For instance, the estimated value of v_h is $\hat{v}_h$=2.3369, then the estimated value of σ_{α_h} is $\hat{\sigma}_{\alpha_h} = \frac{\exp(\hat{v}_h)}{1+\exp(\hat{v}_h)} = 0.9119$. We can remark that $var(\hat{\sigma}_{\alpha_h}) = (g'(\hat{v}_h)^2\, var(\hat{v}_h)$ where $g'(x) = e^x/(1+e^x)^2$. We obtain, finally, that $var(\hat{\sigma}_{\alpha_h}) = 0.0004$.

Table 4: Home Ownership, Employment and Earnings
Heckman's Specification for Initial Conditions

	Home ownership	Employment	(log) Earnings
	SLOPE PARAMETERS		
Intercept	-1.4227***	0.3622***	8.5348***
	(0.1775)	(0.0870)	(0.0160)
Female	-0.0449**	-0.3608***	-0.4720***
	(0.0197)	(0.0183)	(0.0098)
Married	0.3129***	0.0695***	
	(0.0232)	(0.0167)	
Post-Secondary	0.1854***	0.2143***	0.2387***
	(0.0273)	(0.0225)	(0.0128)
University	0.3418***	0.4273***	0.5637***
	(0.0234)	(0.0198)	(0.0106)
Foreign-Born	-0.3737***	-0.4340***	-0.2324***
	(0.0473)	(0.0338)	(0.0232)
Age 20–29	-0.1051***	-0.1001***	-0.0926***
	(0.0303)	(0.0244)	(0.0114)
Age 40–49	-0.0653**	0.1515***	0.1243***
	(0.0260)	(0.0198)	(0.0088)
Age 50+	-0.0588*	-0.1332***	0.1713***
	(0.0314)	(0.0216)	(0.0109)
Unemployment Rate	0.0114	-0.0211**	
	(0.0121)	(0.0085)	
	STATE DEPENDENCE		
$Employed_{t-1}$	0.1137***	1.2861***	1.0537***
	(0.0408)	(0.0523)	(0.0132)
$Owner_{t-1}$	2.7927***		
	(0.2043)		
$(Renter \times Interest)_{t-1}$	-0.0730**		
	(0.0285)		
$(Owner \times Interest)_{t-1}$	0.0768*		
	(0.0419)		
$(Unemployed \times Owner)_{t-1}$		-0.3042***	-0.0891***
		(0.0451)	(0.0293)
$Woman \times$			
$(Unemployed \times Owner)_{t-1}$		-0.0217	0.0211
		(0.0459)	(0.0320)

(*) Significant at 10%. (**) Significant at 5%. (***) Significant at 1%.

Table 4: Home Ownership, Employment and Earnings (Continued)
Heckman's Specification for Initial Conditions

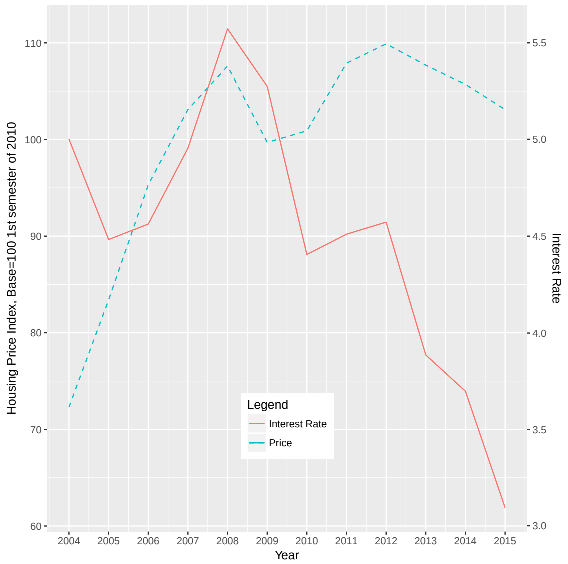
	Home ownership	Employment	(log) Earnings
	INITIAL EQUATIONS		
Intercept	-0.5520***	1.4640***	9.5025***
	(0.1050)	(0.1108)	(0.0123)
Female	-0.0158	-0.5367***	-0.5223***
	(0.0159)	(0.0176)	(0.0111)
Married	0.7330***	0.0358**	
	(0.0164)	(0.0179)	
Post-Secondary	0.2439***	0.2073***	0.2379***
	(0.0222)	(0.0234)	(0.0148)
University	0.3772***	0.4589***	0.5683***
	(0.0186)	(0.0208)	(0.0124)
Foreign-Born	-0.7345***	-0.6351***	-0.3553***
	(0.0337)	(0.0330)	(0.0259)
Age 20–29	-0.7453***	-0.3162***	-0.3731***
	(0.0236)	(0.0240)	(0.0136)
Age 40–49	0.3050***	0.1373***	0.1482***
	(0.0200)	(0.0227)	(0.0121)
Age 50+	0.4836***	-0.1280***	0.2174***
	(0.0229)	(0.0245)	(0.0147)
Unemployment Rate	0.0142	-0.0458***	
	(0.0115)	(0.0121)	

(*) Significant at 10%. (**) Significant at 5%. (***) Significant at 1%.

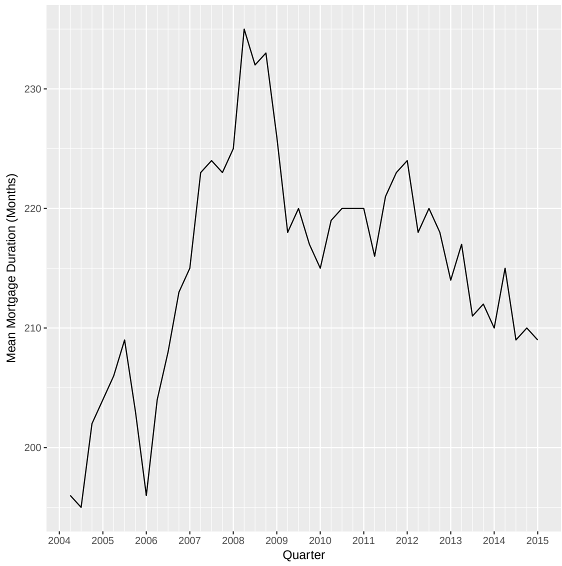
Table 4: Home Ownership, Employment and (log) Earnings (Continued)
Heckman's Specification for Initial Conditions

STOCHASTIC SPECIFICATION			
$r_{jit} = \alpha_{ij} + u_{jit}$			
$u_{jit} = \rho_j \, u_{jit-1} + \epsilon_{jit}$			
Standard errors of individual effects (α_{ij})		**Correlations between initial error terms** (r_{ji0})	
σ_{α_h}	0.1742** (0.0773)	ρ^0_{he}	0.2483*** (0.0110)
σ_{α_e}	0.6282*** (0.0175)	ρ^0_{hw}	0.1875*** (0.0080)
σ_{α_w}	0.6056*** (0.0055)	ρ^0_{ew}	0.4607*** (0.0070)
Correlations between individual effects (α_{ij})		**Correlations between initial error terms** (r_{ji0}) **and other error terms** ($r_{j'it}$)	
$\rho_{\alpha_h \alpha_e}$	0.9698*** (0.3637)	ρ^{00}_{hh}	0.2415*** (0.0293)
$\rho_{\alpha_h \alpha_w}$	0.6873** (0.2774)	ρ^{00}_{eh}	0.1432*** (0.0215)
$\rho_{\alpha_e \alpha_w}$	0.5535*** (0.0140)	ρ^{00}_{wh}	0.1141*** (0.0121)
Auto-Correlation of error terms (u_{jit})		ρ^{00}_{he}	0.2054*** (0.0115)
ρ_h	-0.0991*** (0.0260)	ρ^{00}_{ee}	0.4700*** (0.0183)
ρ_e	-0.1381*** (0.0203)	ρ^{00}_{we}	0.2764*** (0.0094)
ρ_w	0.3785*** (0.0086)	ρ^{00}_{hw}	0.1507*** (0.0085)
Correlations between error terms (ϵ_{jit})		ρ^{00}_{ew}	0.4007*** (0.0095)
ρ_{he}	0.0116 (0.0197)	ρ^{00}_{ww}	0.6596*** (0.0077)
ρ_{hw}	0.0038 (0.0086)	**Standard error of initial log of wage** (r_{wi0})	
ρ_{ew}	0.0779*** (0.0071)	σ_{w0}	0.8859*** (0.0043)
Standard error of log of wage (u_{wit})			
σ_{u_w}	0.5512*** (0.0035)		
Number of obsevations		30,077	

(*) Significant at 10%. (**) Significant at 5%. (***) Significant at 1%.



(a) Housing price index (dashed line) and interest rate (solid line)



(b) Mortgage Length (Data: Banque de France)

Figure 1: Housing Market

Appendix

A. Identification

Let $U_{ji} = (u_{ji1}, u_{ji2}, \cdots, u_{jiT})'$ denote the vector of the error terms for equation j and for periods 1 to T. Let $E_{ji} = \alpha_{ij}\ \mathbb{1}_T$ for $j \in E$. Then $\tilde{R}_i = (U'_{hi}, U'_{ei}, U'_{wi})' + (E'_{hi}, E'_{ei}, E'_{wi})'$ is a vector of residuals for individual i, for all $i = 1, \ldots, n$. Note that[9]

$$\begin{pmatrix} U_{hi} \\ U_{ei} \\ U_{wi} \end{pmatrix} \sim \text{ N}(0, \Sigma_1),$$

where

$$\Sigma_1 = \begin{pmatrix} \kappa_{hh}\ \Psi(\rho_h, \rho_h) & \kappa_{he}\ \Psi(\rho_h, \rho_e) & \kappa_{hw}\ \Psi(\rho_h, \rho_w) \\ \kappa_{he}\ \Psi(\rho_e, \rho_h) & \kappa_{ee}\ \Psi(\rho_e, \rho_e) & \kappa_{ew}\ \Psi(\rho_e, \rho_w) \\ \kappa_{hw}\ \Psi(\rho_w, \rho_h) & \kappa_{ew}\ \Psi(\rho_w, \rho_e) & \kappa_{ww}\ \Psi(\rho_w, \rho_w) \end{pmatrix},$$

and where $\kappa_{jk} = \frac{\rho_{\epsilon_j \epsilon_k}\ \sigma_{\epsilon_j}\sigma_{\epsilon_k}}{(1-\rho_j\ \rho_k)}$ and

$$\Psi(x, y) = \begin{pmatrix} 1 & y & y^2 & \ldots & y^{T-2} & y^{T-1} \\ x & 1 & y & \ldots & y^{T-3} & y^{T-2} \\ \vdots & & \ddots & & & \vdots \\ \vdots & & & \ddots & & \vdots \\ x^{T-2} & & & \ldots & 1 & y \\ x^{T-1} & & & \ldots & x & 1 \end{pmatrix}.$$

The variance of the vector $\tilde{R}_i$ is given by the following equation

$$var(\tilde{R}_i) = \Omega_1 = \Sigma_1 + \begin{pmatrix} \Sigma_{hh} & \Sigma_{he} & \Sigma_{hw} \\ \Sigma_{he} & \Sigma_{ee} & \Sigma_{ew} \\ \Sigma_{hw} & \Sigma_{ew} & \Sigma_{ww} \end{pmatrix},$$

where $\Sigma_{jk} = \rho_{\alpha_j\, \alpha_k}\sigma_{\alpha_j}\sigma_{\alpha_k}\ \mathbb{1}_T \mathbb{1}'_T$, $j, k \in E$.

We have necessarily that $0 < \sigma^2_{\alpha_j} < 1$ since $\sigma^2_{\alpha_j} + \sigma^2_{u_j} = 1$ for all $j = h, e$. Note that the parameters σ_{ϵ_j} and ρ_j are identified through the correlations between equation j and the time periods $1 \ldots T$ because the latter are identified in a multivariate probit ($T \geq 2$). Thus σ_{u_j} and σ_{α_j} are identified. Finally, κ_{jk} and $\rho_{\alpha_j \alpha_k}$ are identified through the correlations between equations j and k for $t = 1 \ldots T$. As κ_{jk} is identified, then so is the correlation $\rho_{\epsilon_j \epsilon_k}$.

Consider now $R_i = (r_{hi0}, r_{ei0}, r_{wi0}, r_{hi1}, \ldots, r_{hiT}, \ldots, r_{wi1} \ldots, r_{wiT})'$. Then $R_i \sim N(0, \Omega)$, where

$$\Omega = \begin{bmatrix} \Omega_{00} & \Omega_{0h} & \Omega_{0e} & \Omega_{0w} \\ \Omega'_{0h} & & & \\ \Omega'_{0e} & & \Omega_1 & \\ \Omega'_{0w} & & & \end{bmatrix}, \tag{A.1}$$

where $\Omega_{00} = Var \begin{pmatrix} r_{hi0} \\ r_{ei0} \\ r_{wi0} \end{pmatrix}$, and $\Omega_{0k} = Cov\left(\begin{pmatrix} r_{hi0} \\ r_{ei0} \\ r_{wi0} \end{pmatrix}, \begin{pmatrix} r_{ki1} \\ \vdots \\ r_{kiT} \end{pmatrix}' \right)$, $k \in E$. The expressions of these variance-covariance matrices are given in Appendix B. With state dependence it is necessary to consider the potential endogeneity of the initial variables save for the wage equation since it is

[9]See Appendix C.

not dynamic and lagged wages are not included in the other equations. These correlations are identified using similar arguments as above.

In principle, it would be possible to model the correlations between the error terms of the initial period of equations j (namely r_{ji0}) and the error terms specific to the periods t ($t > 0$) of equation j' ($j' \in E$). Such a specification would involve an unreasonably large number of nuisance parameters was thus not considered. Our specification of the variance-covariance matrix (A.1) is rather similar to the one used by Hyslop (1999) for the US in a single equation model.[10] Moreover, we consider another method to treat the initial conditions problem that was proposed by Wooldridge (2005, cf. section 2.4). If the two methods yield similar results, then it can legitimately be concluded that assumption of homogeneity of the correlations between r_{ji0} and $r_{j'it}$ is not restrictive.

B. Variance-covariance matrices (Heckman method)

$$\Omega_{00} = \begin{bmatrix} 1 & \rho^0_{he} & \rho^0_{hw}\sigma_{w0} \\ \rho^0_{he} & 1 & \rho^0_{ew}\sigma_{w0} \\ \rho^0_{hw}\sigma_{w0} & \rho^0_{ew}\sigma_{w0} & \sigma^2_{w0} \end{bmatrix},$$

$$\Omega_{0k} = \begin{bmatrix} \rho^{00}_{hk} & \dots & \rho^{00}_{hk} \\ \rho^{00}_{ek} & \dots & \rho^{00}_{ek} \\ \rho^{00}_{wk}\sigma_{w0} & \dots & \rho^{00}_{wk}\sigma_{w0} \end{bmatrix},$$

$k = h, e$.

$$\Omega_{0w} = \begin{bmatrix} \rho^{00}_{hw}\sqrt{\sigma^2_{\alpha_w} + \frac{\sigma^2_{\epsilon_w}}{1-\rho^2_w}} & \dots & \rho^{00}_{hw}\sqrt{\sigma^2_{\alpha_w} + \frac{\sigma^2_{\epsilon_w}}{1-\rho^2_w}} \\ \rho^{00}_{ew}\sqrt{\sigma^2_{\alpha_w} + \frac{\sigma^2_{\epsilon_w}}{1-\rho^2_w}} & \dots & \rho^{00}_{ew}\sqrt{\sigma^2_{\alpha_w} + \frac{\sigma^2_{\epsilon_w}}{1-\rho^2_w}} \\ \rho^{00}_{ww}\sigma_{w0}\sqrt{\sigma^2_{\alpha_w} + \frac{\sigma^2_{\epsilon_w}}{1-\rho^2_w}} & \dots & \rho^{00}_{ww}\sigma_{w0}\sqrt{\sigma^2_{\alpha_w} + \frac{\sigma^2_{\epsilon_w}}{1-\rho^2_w}} \end{bmatrix}$$

C. Autocorrelation (Edon and Kamionka, 2014)

1. Start from the following identity:

$$\begin{aligned} \kappa \equiv cov(u_{jt}, u_{kt}) &= cov(\rho_j\ u_{jt-1} + \epsilon_{jt}, \rho_k\ u_{kt-1} + \epsilon_{kt}) \\ &= \rho_{\epsilon_j\epsilon_k}\ \sigma_{\epsilon_j}\sigma_{\epsilon_k} + \rho_j\rho_k\ cov(u_{jt-1}, u_{kt-1}) \\ \text{so } \kappa &= \frac{\rho_{\epsilon_j\epsilon_k}\ \sigma_{\epsilon_j}\sigma_{\epsilon_k}}{1-\rho_j\rho_k} \end{aligned}$$

2. Likewise,

$$\begin{aligned} cov(u_{jt}, u_{kt-1}) &= cov(\rho_j\ u_{jt-1} + \epsilon_{jt}, u_{kt-1}) \\ &= \rho_j\ \kappa \\ &\dots \end{aligned}$$

3. Let us assume that $cov(u_{jt}, u_{kt-\ell+1}) = \rho_j^{\ell-1}\ \kappa$, where $\ell \leq t$.

[10] This matrix gives an account of the correlations between the individual effects of the initial observation period with the individual effects of the other periods of time. For the same reason, the matrix $\Omega_1 - \Sigma_1$ consists in the variance-covariance of the individual effects.

4. We now show that $cov(u_{jt}, u_{kt-\ell}) = \rho_j^\ell\,\kappa$. Indeed,

$$\begin{aligned} cov(u_{jt}, u_{kt-\ell}) &= cov(\rho_j\, u_{jt-1} + \epsilon_{jt}, u_{kt-\ell}) \\ &= \rho_j\, cov(u_{jt-1}, u_{kt-\ell}) \\ &= \rho_j\, \rho_j^{\ell-1}\, \kappa \\ &= \rho_j^\ell\, \kappa \end{aligned}$$

It can similarly be shown that $cov(u_{kt}, u_{jt-\ell}) = \rho_k^\ell\,\kappa = cov(u_{jt-\ell}, u_{kt})$.

D. Domain of Integration

Specification "à la Wooldridge"

The expressions of the boundaries a_{jit}, b_{jit}, for $j = h, e$ given by

$$\begin{cases} a_{jit} = -\infty, \text{ if } y_{jit} = 0 \text{ and } 0 \le t \le T, \\ b_{jit} = +\infty, \text{ if } y_{jit} = 1 \text{ and } 0 \le t \le T. \\ a_{jit} = -x'_{jit}\,\beta_j - z_j(y_{it-1}, x_{it-1})'\delta_j, \text{ if } y_{jit} = 1 \text{ and } 1 \le t \le T, \\ b_{jit} = -x'_{jit}\,\beta_j - z_j(y_{it-1}, x_{it-1})'\delta_j, \text{ if } y_{jit} = 0 \text{ and } 1 \le t \le T, \\ a_{ji0} = -x_{ji0}'\,\beta_j^0, \text{ if } y_{ji0} = 1, \\ b_{ji0} = -x_{ji0}'\,\beta_j^0, \text{ if } y_{ji0} = 0. \end{cases}$$

For earnings, as we have to consider a continuous variable, taking into account they cannot be observed when the individual is not employed ($y_{wit} = .$, say), the boundaries are the following ones

$$\begin{cases} a_{wit} = -\infty, \text{ if } y_{jit} = . \text{ and } 0 \le t \le T, \\ b_{wit} = +\infty, \text{ if } y_{jit} = . \text{ and } 0 \le t \le T, \\ a_{wit} = b_{wit} = y_{wit} - z_w(y_{it-1}, y_{hit})'\delta_w, \text{ if } y_{wit} \ne . \text{ and } 1 \le t \le T, \\ a_{wi0} = b_{wi0} = y_{wi0} - x_{wi0}'\beta_w^0, \text{ if } y_{wi0} \ne . \text{ and } t = 0. \end{cases}$$

Specification "à la Heckman"

The expressions of the boundaries a_{jit} and b_{jit} are fixed for $t = 1, \ldots, T$, and are given by:

$$\begin{cases} a_{jit} = -\infty, \text{ if } y_{jit} = 0, \\ b_{jit} = +\infty, \text{ if } y_{jit} = 1, \\ a_{jit} = -x'_{jit}\,\beta_j - z_j(y_{it-1}, x_{it-1})'\delta_j - y_{hi0}\,\lambda_{jh} - y_{ei0}\,\lambda_{je}, \text{ if } y_{jit} = 1, \\ b_{jit} = -x'_{jit}\,\beta_j - z_j(y_{it-1}, x_{it-1})'\delta_j - y_{hi0}\,\lambda_{jh} - y_{ei0}\,\lambda_{je}, \text{ if } y_{jit} = 0, \end{cases}$$

for $j = h, e$, and $1 \le t \le T$.

Earnings are continuous and are not observed when an individual is not employed ($y_{wit} = .$, say). The boundaries are thus given by:

$$\begin{cases} a_{wit} = -\infty, \text{ if } y_{wit} = ., \\ b_{wit} = +\infty, \text{ if } y_{wit} = ., \\ a_{wit} = b_{wit} = y_{wit} - z_w(y_{it-1}, y_{hit})'\delta_w - y_{hi0}\,\lambda_{wh} - y_{ei0}\,\lambda_{we}, \text{ if } y_{wit} \neq ., \end{cases}$$

where $1 \leq t \leq T$.

E. Simulation of a Contribution to the Likelihood Function

Assume that $r \in A_i$ (see section 2.3), where $A_i \subset \mathbb{R}^{3(T+1)}$ and $a_{ik} \leq r_k \leq b_{ik}$, $\forall\, k = 1, \ldots, 3(T+1)$. Let L denote the total number of observations per individual (L=$3(T$+$1)$).

The contribution of individual i to the likelihood function (9) can be estimated using the expression (see (Geweke, 1991; Hajivassiliou et al., 1992; Keane, 1994; Chang, 2009):

$$\hat{p}_i^S = \frac{1}{S}\sum_{s=1}^{S} \tilde{p}(x_i; u_i^s; \theta), \tag{E.2}$$

where $u_i^s = (u_{i1}^s, u_{i2}^s, \ldots, u_{iL}^s)'$ is a random draw. S is the number of draws used in the estimation and $\tilde{p}$ is an unbiased simulator of probability Prob$[r \in A_i \mid x_i; \theta]$, where r is the vector of error terms for a given individual.

Let $U = \Gamma\, u$ where $\Omega = \Gamma\,\Gamma'$ is the Cholesky decomposition of the matrix Ω. The random variable u is a drawn from the distribution $N(0_L, I_L)$ and Γ is a lower triangular matrix. Assume further that $\Gamma = [\Gamma_{jk}]$, so that Γ_{jk} is the j,k element of the matrix Γ ($j,k = 1,\ldots,L$). Moreover, let $\zeta_{ik} = \frac{a_{ik}}{\Gamma_{kk}}$, $\varphi_{ik} = \frac{b_{ik}}{\Gamma_{kk}}$ and $\Gamma_{jk}^0 = \frac{\Gamma_{jk}}{\Gamma_{jj}}$.

For individual i, the draw s is obtained using a vector u_i^s with length L such that $u_i^s = (u_{i1}^s, u_{i2}^s, \ldots, u_{iL}^s)'$.

In order to obtain the expression of the vector u_i^s each time the likelihood function is computed for a given value of the vector of parameters, we proceed iteratively. Hence, let $\delta_{xy} = 1$ if $x \neq y$ and $\delta_{xy} = 0$ if $x = y$. Note that $a_{ik} = b_{ik}$ if and only if the endogenous variable is not censored (we set in this case $\delta_{\zeta_{ik}\varphi_{ik}} = 0$).

The vector u_i^s is constructed as follows:

- Let $\tilde{u}_{i1}^s \sim U(0,1)$ and set

  $$u_{i1}^s = \Phi^{-1}\left[\left(\Phi(\varphi_{i1}) - \Phi(\zeta_{i1})\right)\tilde{u}_{i1}^s + \Phi(\zeta_{i1})\right]\delta_{\zeta_{i1}\varphi_{i1}} + (1-\delta_{\zeta_{i1}\varphi_{i1}})\,\zeta_{i1}$$

- Let $\tilde{u}_{i2}^s \sim U(0,1)$ and assume that

  $$\begin{aligned} u_{i2}^s = \Phi^{-1}\Big[\big(\,&\Phi(\varphi_{i2}-\Gamma_{21}^0\,u_{i1}^s) \\ &-\Phi(\zeta_{i2}-\Gamma_{21}^0\,u_{i1}^s)\,\big)\,\tilde{u}_{i2}^s + \Phi(\zeta_{i2}-\Gamma_{21}^0\,u_{i1}^s)\Big]\,\delta_{\zeta_{i2}\varphi_{i2}} \\ &+ (1-\delta_{\zeta_{i2}\varphi_{i2}})(\zeta_{i2}-\Gamma_{21}^0\,u_{i1}^s) \end{aligned}$$

- Let $\tilde{u}_{iL}^s \sim U(0,1)$ and assume further that

$$
\begin{aligned}
u^s_{iL} = \Phi^{-1} \Big[(\, & \Phi(\varphi_{iL} - \Gamma^0_{L\,(L-1)} \, u^s_{i\,(L-1)} - \ldots - \Gamma^0_{L\,1} \, u^s_{i1}) \\
& - \Phi(\zeta_{iL} - \Gamma^0_{L\,(L-1)} \, u^s_{i(L-1)} - \ldots - \Gamma^0_{L\,1} \, u^s_{i1}) \,) \, \tilde{u}^s_{iL} \\
& + \Phi(\zeta_{iL} - \Gamma^0_{L\,(L-1)} \, u^s_{i(L-1)} - \ldots - \Gamma^0_{L\,1} \, u^s_{i1}) \Big] \, \delta_{\zeta_{iL}\varphi_{iL}} \\
& + (1 - \delta_{\zeta_{iL}\varphi_{iL}})(\zeta_{iL} - \Gamma^0_{L\,(L-1)} \, u^s_{i(L-1)} - \ldots - \Gamma^0_{L\,1} \, u^s_{i1})),
\end{aligned}
$$

where Φ is the cumulative distribution function of the probability density function $N(0,1)$. We the sequentially obtain all the components of the random draw u^s_i.

The estimation of an individual contribution to the conditional likelihood function can be computed using the empirical mean of the following terms:

$$
\begin{aligned}
\tilde{p}(x_i; u^s_i; \theta) = & \left[[\Phi(\varphi_{i1} - \Phi(\zeta_{i1})] \delta_{\zeta_{i1}\varphi_{i1}} + (1 - \delta_{\zeta_{i1}\varphi_{i1}}) \frac{1}{\Gamma_{11}} \phi(u^s_{i1}) \right] \\
& \times \prod_{k=2}^{L} \Bigg[\Big[\Phi(\varphi_{ik} - \Gamma^0_{k\,(k-1)} \, u^s_{i(k-1)} - \ldots - \Gamma^0_{k\,1} \, u^s_{i1}) \\
& - \Phi(\zeta_{ik} - \Gamma^0_{k\,(k-1)} \, u^s_{i(k-1)} - \ldots - \Gamma^0_{k\,1} \, u^s_{i1}) \Big] \, \delta_{\zeta_{ik}\varphi_{ik}} \\
& + (1 - \delta_{\zeta_{ik}\varphi_{ik}}) \frac{1}{\Gamma_{kk}} \phi(u^s_{ik}) \Bigg],
\end{aligned}
\tag{E.3}
$$

where x_i is a vector of explanatory variables for individual i ($i = 1, \ldots, n$) and $h = 1, \ldots, S$. ϕ is the probability density function of the $N(0,1)$ distribution.