Online Appendix to Estimating a Dynamic Adverse-Selection Model: Labor-Force Experience and the Changing Gender Earnings Gap 1968–97.

George-Levi Gayle and Limor Golan
Tepper School of Business, Carnegie Mellon University.

January 30, 2011

Abstract

This Appendix contains details on identification, estimation, and asymptotic properties of the estimator.

1 Identification

Given panel data, $\{a_{nt}, z_{nt}, c_{nt}, z_{nt}^{\mathcal{P}}, S_{nt}\}_{n=1,\tau\in\{\mathrm{P,NP}\},t=1}^{N,T}$, where n indexes individual and t indexes the year, we outline our identification strategy below.

1. We show that under standard regularity conditions on $u_{i2}(c_t; x_t, \varepsilon_{2t})$, if it is multiplicatively separable in ε_{2t}, then the standard independence assumption between z_t and ε_{2t} allows us to identify $\eta\lambda_t$ up to a proportionality constant.

2. Assuming that $z_t^{\mathcal{P}}$ is fully observed by the econometrician, we show that, under standard regularity conditions, the equilibrium salary schedule and β are identified.

3. Let $F_1(\varepsilon_{0t}, \varepsilon_{1t})$ be the marginal of $F_1(\varepsilon_{0t}, \varepsilon_{1t}, \varepsilon_{2t})$. Given that $S_{i\tau t}(h_t; \omega_t)$, β, and $\eta\lambda_t$ are identified, we show that $u_{0i}(z_t, \zeta_t)$ and $u_{i1}(l_t; z_t)$ are identified up to $F_1(\varepsilon_{0t}, \varepsilon_{1t})$ and two additive constants. Putting 1, 2, and 3 together, we conclude that our model is identified up to $F_1(\varepsilon_{0t}, \varepsilon_{1t})$ and two additive constants.

The following assumption is used in the proof of all the steps above:

Assumption 1.1 *The econometrician observes all the worker's state variables except for the idiosyncratic components ε_{nt}, the unobserved heterogeneity (in production) ν_n, and the worker's marginal utility of wealth, η_n .*

1.0.1 Identification of the Marginal Utility of Wealth

It is well known in the literature on the estimation of consumption functions that the general form of utility with risk aversion is not identified without quantity and price data, which we do not have. Therefore, we follow the literature and state sufficient conditions to obtain identification of the marginal utility of wealth. We state the conditions below.

Assumption 1.2 *The marginal utility of consumption has the following form.*

$$\frac{\partial u_{i2}(c_{nt}, x_{nt}, \varepsilon_{2t})}{\partial c_{nt}} = \exp(u_{2c}(c_{nt})) \exp(-u_{2x}(x_{nt})) \exp(-\varepsilon_{2nt}). \tag{1}$$

Assumption 1.3 *1) $E[\varepsilon_{2t}|x_t] = 0$ for all n and t. 2) $E_n[\log(\eta_n)|x_{nt}] = 0$.*

Assumption 1.4 *1) x_{nt} has a continuous element x_{cnt} with continuous variation on its support $[\underline{x_c}, \overline{x_c}]$. 2) $exp(-u_{2x}(\underline{x_c}, .)) = 0$.*

Assumption (1.2) states that the marginal utility of consumption is multiplicatively separable. For example, both the class of constant absolute risk aversion and the class of constant relative risk aversion satisfy this assumption. Assumption 1.3(1) formally states that the error is mean independent of x_{nt} with expectation zero. Assumption 1.3(2) is the standard normalization needed in a panel data model in order to recover the level of the time component. Finally, Assumption 1.4(1) states that at least one variable with continuous variation on its support is required, and Assumption 1.4(2) is a boundary condition. Assumption 1.4 can be replaced with a parametric assumption on the function $u_{2x}(x_{nt})$.

Lemma 1.1 *If $u_{2c}(c_{nt})$ is known, and assumptions 1.2–1.4 are satisfied. Then $\eta_n \lambda_t$ is identified.*

Proof. Using the functional form assumption in (1.2) the Euler equation is therefore:

$$\exp(u_{2c}(c_{nt})) \exp(-u_{2z}(x_{nt})) \exp(-\varepsilon_{2nt}) = \eta_n \lambda_t. \tag{2}$$

Taking the log and then first difference of equation (2) and rearranging gives us

$$\Delta u_{2c}(c_{nt}) = \Delta u_{2x}(x_{nt}) + \Delta \log(\lambda_t) + \Delta \varepsilon_{2nt} \tag{3}$$

By assumption 1.3(1), then,

$$E[\Delta u_{2c}(c_{nt})|x_{nt}, x_{nt-1}] = \Delta u_{2x}(x_{nt}) + \log(\lambda_t) \tag{4}$$

Taking the derivative of $E[\Delta u_{2c}(c_{nt})|x_{nt}, x_{nt-1}]$ with respect to x_{cnt} and x_{cnt-1}, respectively, and integrating back up to x_{cnt} and x_{cnt-1}, respectively, gives

$$u_{2xi}(x_{nt}) = u_{2x}(\underline{x_c}, x_{c'nt}) + \int_{x_c}^{x_{cnt}} \left\{ \frac{\partial E[\Delta u_{2c}(c_{nt})|x_{nt}, x_{nt-1}]}{\partial x_c} \right\} \mathrm{d}x_c \tag{5}$$

$$u_{2xi}(x_{nt-1}) = u_{2z}(\underline{x_c}, x_{c'nt-1}) + \int_{x_c}^{x_{cnt-1}} \left\{ \frac{\partial E[\Delta u_{2c}(c_{nt})|x_{nt}, x_{nt-1}]}{\partial x_{c-1}} \right\} \mathrm{d}x_{c-1}, \tag{6}$$

which by Assumption 1.4(2) and from Chesher's (2007) results is identified. Therefore

$$\begin{aligned} \Delta \log(\lambda_t) &= E[\Delta u_{2c}(c_{nt})|x_{nt}, x_{nt-1}] - \int_{x_c}^{x_{cnt}} \left\{ \frac{\partial E[\Delta u_{2c}(c_{nt})|x_{nt}, x_{nt-1}]}{\partial x_c} \right\} \mathrm{d}z_c \\ &\quad + \int_{x_c}^{x_{cnt-1}} \left\{ \frac{\partial E[\Delta u_{2c}(c_{nt})|x_{nt}, x_{nt-1}]}{\partial x_{c-1}} \right\} \mathrm{d}x_{c-1} \end{aligned} \tag{7}$$

and, by Assumption 1.3(1),

$$\log(\lambda_1) = E[\Delta u_{2c}(c_{n1})|x_{n1}] - \int_{x_c}^{x_{cn1}} \left\{ \frac{\partial E[\Delta u_{2c}(c_{nt})|x_{n1}]}{\partial x_c} \right\} \mathrm{d}x_c. \tag{8}$$

Hence, λ_t is identified. Finally, by Assumption 1.3(2), we have

$$\log(\eta_n) = E_t\{u_{2c}(c_{nt}) - \log(\lambda_t) - u_{2x}(x_{nt})|x_{nt}\}. \tag{9}$$

Using Chesher's (2007) result and the fact that $u_{2c}()$ is assumed known, we use the results from equations (7), (8), and (9). ■

1.0.2 Identification of the Equilibrium Salary Schedule

Next, we establish the identification of the equilibrium salary schedule. First we will carry out the analysis assuming that all the element of $z_{nt}^{\mathcal{P}}$ are observed by the econometrician. The modification to include unobserved (to the econometrician) individual specific effect requires additional functional form assumptions. These assumptions are made in the estimation section where the estimation and identification are illustrated using standard linear panel data methods. We will derive the result under symmetric information, but the proof goes through under asymmetric information by replacing ω_t with ω_t^*.

Assumption 1.5 *There exist observable characteristics, ω_t, on a set of positive measure, such that*

1. $\triangle\, \widetilde{p}_{\tau t+1}(h_t, \omega_t) = \widetilde{p}_{m\tau t+1}(h_t, \omega_t) - \widetilde{p}_{w\tau t+1}(h_t, \omega_t) \neq 0 \quad \vee\tau$
2. $Y_\tau(0, z_t^{\mathcal{P}}, K_{\tau t}) = 0 \ \vee\tau, z_t^{\mathcal{P}}, K_{\tau t}.$

Assumption 1.5(1) states that in each occupation there is a range of hours, labor market experience and individual characteristic for which the employers hold different belief about men and women, Assumption 1.5(2) states that an input of zero hours produces zero output.

Lemma 1.2 *Under Assumption(1.5), $y_{\tau t}(h_t, H_{t-1}, z_t^p)$, β, and γ_τ are identified, and there are at least two over-identifying restrictions.*

Proof. of Lemma 1.2.

This result is established by proving the following (all the differences are taken with respect to gender).

$$\beta\gamma_\tau = \frac{\triangle\ E_t[d_{nt}I_{n\tau t}S_{n\tau t}|h_{nt},\omega_{nt}]}{\triangle\ \widetilde{p}_{\tau t+1}(h_{nt},\omega_{nt})}, \tag{10}$$

$$\begin{aligned} y_{\tau t}(h_{nt}, z_{nt}^{\mathcal{P}}, K_{\tau t}) \quad = \quad & \int_0^{h_{nt}} \left\{\partial\left[\frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|h_{nt},\omega_{nt}]}{E_t[d_{nt}I_{n\tau t}|h_{nt},\omega_{nt},i]}\right]\backslash\partial h\right\} \mathrm{d}h \\ & - \int_0^{h_{nt}} \left\{\partial\left[\frac{\triangle\ E_t[d_{nt}I_{n\tau t}S_{n\tau t}|h_{nt},\omega_{nt}]}{\triangle\ \widetilde{p}_{\tau t+1}(h_{nt},\omega_{nt})}\right]\backslash\partial h\right\} \mathrm{d}h \end{aligned} \tag{11}$$

for $i \in \{m, w\}$ and

$$\begin{aligned} \gamma_\tau \quad = \quad & \left\{\int_0^{h_{nt}} \left\{\partial\left[\frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|h_{nt},\omega_{nt}]}{E_t[d_{nt}I_{n\tau t}|h_{nt},\omega_{nt},i]}\right]\backslash\partial h\right\} \mathrm{d}h - \frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|h_{nt},\omega_{nt}]}{E_t[d_{nt}I_{n\tau t}|h_{nt},\omega_{nt},i]}\right\} \\ & - \left\{\int_0^{h_{nt}} \left\{\partial\left[\frac{\triangle\ E_t[d_{nt}I_{n\tau t}S_{n\tau t}|h_{nt},\omega_{nt}]}{\triangle\ \widetilde{p}_{\tau,t+1}(h_{nt},\omega_{nt})}\right]\backslash\partial h\right\} \mathrm{d}h - \frac{\triangle\ E_t[d_{nt}I_{n\tau t}S_{n\tau t}|h_{nt},\omega_{nt}]}{\triangle\ \widetilde{p}_{\tau,t+1}(h_{nt},\omega_{nt})}\right\} \end{aligned} \tag{12}$$

for $i \in \{m, w\}$ and where $\triangle\ E_t[d_{nt}I_{n\tau t}S_{n\tau t}|h_{nt},\omega_{nt}] = E_t\left[\frac{E_t[d_{nt}I_{n\tau t}S_{mn\tau t}|h_{nt},\omega_{nt}]}{E_t[d_{nt}I_{n\tau t}|h_{nt},\omega_{nt},i=m]} - \frac{E_t[d_{nt}I_{n\tau t}S_{fn\tau t}|h_{nt},\omega_{nt}]}{E_t[d_{nt}I_{n\tau t}|h_{nt},\omega_{nt},i=f]}\right]$.

Applying the results from Chesher (2007), all these parameters are identified because data is informative about $\widetilde{p}_{i\tau,t+1}(h_{nt},\omega_{nt})$ by part (3) of Proposition 3.1. *Note that the two over-identifying restrictions come from equations(11) and(12). There is one parameter to identify for each occupation, however, there are two equations (one for each gender) for each occupation.*

From part one of Proposition 3.1, the zero-profit condition implies that

$$E_t[d_{nt}I_{n\tau t}(S_{in\tau t} - y_{\tau t}(h_{nt}, z_{nt}^{\mathcal{P}}, K_{\tau t}) + \gamma_\tau - d_{nt+1}I_{n\tau t+1}\beta\gamma_\tau)|h_{nt},\omega_{nt},i] = 0. \tag{13}$$

Rearranging and noting that

$$\widetilde{p}_{i\tau,t+1}(h_{nt},\omega_{nt}) = \frac{E_t[d_{nt}I_{n\tau t}d_{nt+1}I_{n\tau t+1}|h_{nt},\omega_{nt},i]}{E_t[d_{nt}I_{n\tau t}|h_{nt},\omega_{nt},i]}, \tag{14}$$

we can write the zero-profit condition as

$$\frac{E_t[d_{nt}I_{n\tau t}S_{in\tau t}|h_{nt},\omega_{nt}]}{E_t[d_{nt}I_{n\tau t}|h_{nt},\omega_{nt},i]} = y_{\tau t}(h_{nt}, z_{nt}^{\mathcal{P}}, K_{\tau t}) - \gamma_\tau + \beta\gamma_\tau\widetilde{p}_{i\tau,t+1}(h_{nt},\omega_{nt}). \tag{15}$$

Taking the difference between equation (15) for men and women and rearranging gives equation which is well defined by Assumption 1.5(1). Substituting equation (10) into (15) for men and women gives the following system of equations.

$$(16)\quad y_{\tau t}(h_{nt}, z^{\mathcal{P}}_{nt}, K_{\tau t}) - \gamma_\tau = \frac{E_t[d_{nt} I_{n\tau t} S_{in\tau t}|h_{nt}, \omega_{nt}]}{E_t[d_{nt} I_{n\tau t}|h_{nt}, \omega_{nt}, i]} - \frac{\triangle\ E_t[d_{nt} I_{n\tau t} S_{n\tau t}|h_{nt}, \omega_{nt}] \times \widetilde{p}_{i\tau t+1}(h_{nt}, \omega_{nt})}{\triangle\ \widetilde{p}_{\tau t+1}(h_{nt}, \omega_{nt})},\ i = \{m, w\}$$

Differentiating equation (16) with respect to hours and then integrating,

$$\begin{aligned} y_{\tau t}(h_{nt}, z^{\mathcal{P}}_{nt}, K_{\tau t}) &= y_{\tau t}(0, z^{\mathcal{P}}_{nt}, K_{\tau t}) + \int_0^{h_{nt}} \left\{ \partial \left[\frac{E_t[d_{nt} I_{n\tau t} S_{in\tau t}|h_{nt}, \omega_{nt}]}{E_t[d_{nt} I_{n\tau t}|h_{nt}, \omega_{nt}, i]} \right] \backslash \partial h \right\} \mathrm{d}h \\ (17)\qquad & \quad - \int_0^{h_{nt}} \left\{ \partial \left[\frac{\triangle\ E_t[d_{nt} I_{n\tau t} S_{n\tau t}|h_{nt}, \omega_{nt}]}{\triangle\ \widetilde{p}_{\tau t+1}(h_{nt}, \omega_{nt})} \right] \backslash \partial h \right\} \mathrm{d}h \end{aligned}$$

and by Assumption 1.5(2), we have the lemma's results. By substituting (17) into (16) and rearranging, we obtain the final equation in the lemma. ■

The two over-identifying restrictions allow to test the restriction that there is no difference in productivity between men and women. The main identifying assumption in Lemma (1.2) is that there is a difference in future participation probability between men and women. An alternative assumption is the there is a difference in future participation probability of different cohorts for a given gender.

1.0.3 Identification of the Utility of Non-market Time

In the literature on the estimation of dynamic Markovian games, it is standard to use time-series data to estimate and identify models. We extend this approach to a panel data setting, considering age–education cohort partition of data generated by a single path of play, exploiting, therefore, the information contained in the repeated observation of the same players in the cohort partition along the path of play. Because different cohorts may be playing different equilibria, we also have variation across cohort partitions. Below, we formalize the equilibrium selection discussed in the previous paragraph.

Assumption 1.6 (Equilibrium Selection) *Conditional on the time invariant component of* $z^{\mathcal{P}}_t$*, the data for each age–education cohort is generated by only one equilibrium.*

This assumption rules out the possibility that for any given age–education cohort and the time invariant component of $z^{\mathcal{P}}_t$, the time series data is generated by a mixture of two or more equilibria.

Let us redefine the primitives of our problem as follows. The per-period utility is

$$(18)\qquad U_{kit}(\omega_{nt}) = \begin{cases} u_{i1}(1, z_{nt}) & \text{for } k = 0 \\ u_{i0}(z_t, \zeta_t) + u_{i1}(l_t, z_t) + \eta\lambda_t \sum\limits_{\tau \in \{\text{P,NP}\}} I_{\tau t} S_{i\tau t}(h_{nt}, \omega^*_{nt}) & \text{for } k = 1 \end{cases}$$

We can write the ex-ante value functions more concisely as

(19)

$$V_{kit}(\omega_{nt}) \equiv \max_{\{h_s;\{I_{\tau s}\}_{\tau=1}^{\Upsilon}\}_{s=t}^{T}} E_t \left[\sum_{s=t}^{T} \beta^{s-t} \{ d_s [U_{1is}(\omega_{ns}) + \varepsilon_{1ns}] + (1-d_s)[U_{0is}(\omega_{ns}) + \varepsilon_{0ns}] \} \mid d_t = k \right]$$

Our game differs in two important dimensions from the games typically estimated. First, in our model, employers learn and update beliefs based on the complete history of workers' behavior. Second, the state variables do not have discrete support because labor-market experience has continuous components. Using the stochastic finite state dependence property of model, we show that our model is identified.

Equation (15, paper) and Lemma 1 of Hotz and Miller (1993) imply that:

$$V_{1it}(\omega_{nt}) - V_{0it}(\omega_{nt}) = Q^{-1}(p_{it}(\omega_{nt})). \tag{20}$$

Proposition 1 of Hotz and Miller (1993) also states that there exists a mapping $\varphi_k : [0,1] \to R$, that measures the expected value of the unobservable in the current utility, conditional on action $k \in \{0,1\}$. That is,

$$\varphi_k(p_{it}(\omega_{nt})) \equiv E[\varepsilon_{knt} \mid \omega_{nt}, d_{nt}^{o} = k]. \tag{21}$$

To set some notation, let $\omega_{kt}^{(s)}$ denote the state in period $t+s$ if, at time t, the k^{th} option is taken—that is, $d_t = k$—and the sequence of decisions for the next s periods is $d_{kt+1}^{\rho(\omega_0)}(\omega_{t+1}), \ldots, d_{kt+s}^{\rho(\omega_0)}(\omega_{t+s})$. Denote by $p_{kit}^{(s)}$, the probability that $d_{t+s} = 1$ conditional on $\omega_{kt}^{(s)}$, i.e. $p_{kit}^{(s)} = E\left[d_{t+s} | \omega_{kt}^{(s)}\right]$.

Equations (14, paper) and (25, paper) are necessary conditions and therefore must hold in all equilibria. $\left\{ p_{it}(\omega_{nt}), \left\{ p_{0it}^{(s)}, p_{1it}^{(s)} \right\}_{s=1}^{\rho(\omega_{nt})} \right\}_{i=w,m}$, and the distribution over which E_t is taken, are the only elements which differ across the different equilibria. These elements are conditional expectation functions and, given Assumption (1.6), can be recovered from the data; therefore, they are identified.[1] For the purpose of the identification of the utility of non-market hours, we can treat $\left\{ p_{it}(\omega_{nt}), \left\{ p_{0it}^{(s)}, p_{1it}^{(s)} \right\}_{s=1}^{\rho(\omega_{nt})} \right\}_{i=w,m}$ as known. Denote by $\eta_n^o \lambda_t^o$ and $S_{i\tau t}^o(h_{nt}, \omega_{nt}^*)$ the shadow prices and salary schedule under the true equilibrium in the data, respectively. The true probabilities under the true equilibrium in the data are denoted by $\left\{ p_{it}^{o}(\omega_{nt}), \left\{ p_{0it}^{o(s)}, p_{1it}^{o(s)} \right\}_{s=1}^{\rho(\omega_{nt})} \right\}_{i=w,m}$, and define

$$\begin{aligned} Y_{i1nt} \equiv\ & \eta_n^o \lambda_t^o \sum_{\tau=1}^{\Upsilon} I_{n\tau t} S_{i\tau t}^o(h_{nt}, \omega_{nt}^*) + \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left\{ \varphi_0\left(p_{0it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right) \right. \\ & + p_{0it}^{o(s)} \left[Q^{-1}\left(p_{0it}^{o(s)}\right) + \varphi_1\left(p_{0it}^{o(s)}\right) - \varphi_0\left(p_{0it}^{o(s)}\right) \right] \\ & \left. - p_{1it}^{o(s)} \left\{ Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right) \right\} \right\} - Q^{-1}(p_{it}^{0}(\omega_{nt})) \end{aligned} \tag{22}$$

[1] See Haavelmo (1944) for detail.

and

$$
\begin{aligned}
Y_{i2nt} \equiv\ & \eta_n^o \lambda_t^o \sum_{\tau=1}^{\Upsilon} I_{in\tau t} \frac{\partial S_{i\tau t}^o(h_{nt}, \omega_{nt}^*)}{\partial h_t} - \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \Bigg\{ \frac{\partial \varphi_0 \left(p_{1it}^{o(s)}\right)}{\partial h_t} \\
& + \frac{\partial p_{1it}^{o(s)}}{\partial h_t} \left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right) \right] \\
& + p_{1it}^{o(s)} \frac{\partial \left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right) \right]}{\partial h_t} \Bigg\}
\end{aligned}
\tag{23}
$$

Because $\eta_n^o \lambda_t^o$, $S_{i\tau t}^o(h_{nt}, \omega_{nt}^*)$, $F(\varepsilon_{0t}, \varepsilon_{1t})$, and β are treated as known, Y_{i1nt} and Y_{i2nt} can be treated as observed data.

The following Lemma establishes necessary conditions for equilibrium,

Lemma 1.3 *In all the equilibria, the following system of equations holds.*

$$
\begin{aligned}
Y_{i1nt} =\ & u_{i1}(1, z_{nt}) - u_{i0}(z_t, \zeta_t) - u_{i1}(l_{nt}, z_{nt}) \\
& + \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[u_{i1}\left(1, z_{1nt}^{(s)}\right) - u_{i1}\left(1, z_{0nt}^{(s)}\right) \right] + \xi_{i1nt}
\end{aligned}
\tag{24}
$$

$$
Y_{i2nt} = -\frac{\partial u_{i1}(l_{nt}, z_{nt})}{\partial h_t} - \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[\frac{u_{i1}\left(1, z_{1nt}^{(s)}\right)}{\partial h_t} \right] + \xi_{i2nt}
\tag{25}
$$

where $z_{knt}^{(s)}$ denote the worker's type and labor experience in period $t+s$, if, at time t, $d_t = k$, and the sequence of decisions for the next s periods is $d_{kt+1}^{\rho(\omega_0)}(\omega_{t+1})$, …, $d_{kt+s}^{\rho(\omega_0)}(\omega_{t+s})$: $E_t^o[\xi_{i1nt}|\omega_{nt}] = 0$ and $E_t^o[\xi_{i2nt}|\omega_{nt}, d_{nt} = 1] = 0$ for $i = \{w, m\}$, and E_t^o is taken over the actual equilibrium played. The formal definition of the residual ξ_{i1nt} and ξ_{i2nt} are in the proof .

Combining equation (14, paper), equation (20), and equation (21) with the ex-ante value function (13) allows us to write the ex-ante equilibrium value function for any initial state ω_0:

$$
\begin{aligned}
V_{kit}(\omega_0) =\ & U_{kit}(\omega_0) + E_t \Bigg\{ \sum_{s=1}^{\rho(\omega_0)} \beta^s \Big[U_{0it+s}\left(\omega_{kt}^{(s)}\right) + \varphi_0\left(p_{kit}^{(s)}\right) \\
& + p_{kit}^{(s)} \left\{ Q^{-1}\left(p_{kit}^{(s)}\right) + \varphi_1\left(p_{kit}^{(s)}\right) - \varphi_0\left(p_{kit}^{(s)}\right) \right\} \Big] \\
& + \beta^{\rho(\omega_0)+1} \Big[V_{0it+\rho(\omega_0)+1}(\omega_{\rho(\omega_0)+1}) + \varphi_0\left(p_{kit}^{(\rho(\omega_0)+1)}\right) \\
& + p_{kit}^{(\rho(\omega_0)+1)} \left\{ Q^{-1}\left(p_{kit}^{(\rho(\omega_0)+1)}\right) + \varphi_1\left(p_{kit}^{(\rho(\omega_0)+1)}\right) - \varphi_0\left(p_{kit}^{(\rho(\omega_0)+1)}\right) \right\} \Big] \Bigg\}
\end{aligned}
\tag{26}
$$

A proof of this representation can be found in Altug and Miller (1998).

Next, we characterize, using 20 and 26, the necessary conditions for equilibrium (participation and hours) in equations (14, paper) and (25, paper). First, we characterize the equilibrium relationship from (14). Substituting (26) into (20) gives equilibrium:

$$\begin{aligned} Q^{-1}(p_{it}(\omega_{nt})) &= U_{1it}(\omega_{nt}) - U_{0it}(\omega_{nt}) + E_t \left\{ \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[U_{0it+s}\left(\omega_{0t}^{(s)}\right) - U_{0it+s}\left(\omega_{1t}^{(s)}\right) \right.\right. \\ & \left. +\varphi_0\left(p_{0it}^{(s)}\right) - \varphi_0\left(p_{1it}^{(s)}\right)\right] + p_{0it}^{(s)}\left[Q^{-1}\left(p_{0it}^{(s)}\right) + \varphi_1\left(p_{0it}^{(s)}\right) - \varphi_0\left(p_{0it}^{(s)}\right)\right] \\ & \left. - p_{1it}^{(s)}\left[Q^{-1}\left(p_{1it}^{(s)}\right) + \varphi_1\left(p_{1it}^{(s)}\right) - \varphi_0\left(p_{1it}^{(s)}\right)\right]\right\} \end{aligned} \tag{27}$$

Note that all the elements from period $\rho(\omega_{nt})+1$ onward are the same irrespective of whether action 1 or 0 is taken today by finite state dependence. Hence, they fall out of the above equation. Similarly, using equation(26), the necessary condition for equilibrium hours can be rewritten as

$$\begin{aligned} -\frac{\partial U_{1it}(\omega_{nt})}{\partial h_t} &= E_t \left\{ \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[\frac{\partial U_{0it+s}\left(\omega_{1t}^{(s)}\right)}{\partial h_t} + \frac{\partial \varphi_0\left(p_{1it}^{(s)}\right)}{\partial h_t} \right] \right. \\ & + \frac{\partial p_{1it}^{(s)}}{\partial h_t}\left[Q^{-1}\left(p_{1it}^{(s)}\right) + \varphi_1\left(p_{1it}^{(s)}\right) - \varphi_0\left(p_{1it}^{(s)}\right)\right] \\ & \left. + p_{1it}^{(s)} \frac{\partial\left[Q^{-1}\left(p_{1it}^{(s)}\right) + \varphi_1\left(p_{1it}^{(s)}\right) - \varphi_0\left(p_{1it}^{(s)}\right)\right]}{\partial h_t} \right\} \end{aligned} \tag{28}$$

Note that again, by finite state dependence, all the elements from period $\rho(\omega_{nt})+1$ onward fall out of the above equations.

Proof. Define the errors as

$$\begin{aligned} \xi_{i1nt} &= \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[U_{0it+s}\left(\omega_{0t}^{(s)}\right) - U_{0it+s}\left(\omega_{1t}^{(s)}\right)\right] + p_{0it}^{o(s)}\left[Q^{-1}\left(p_{0it}^{o(s)}\right) + \varphi_1\left(p_{0it}^{o(s)}\right) - \varphi_0\left(p_{0it}^{o(s)}\right)\right] \\ & - p_{1it}^{o(s)}\left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right] - E_t^0\left\{\sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[U_{0it+s}\left(\omega_{0t}^{(s)}\right) - U_{0it+1s}\left(\omega_{1t}^{(s)}\right)\right]\right. \\ & \left. + p_{0it}^{o(s)}\left[Q^{-1}\left(p_{0it}^{o(s)}\right) + \varphi_1\left(p_{0it}^{o(s)}\right) - \varphi_0\left(p_{0it}^{o(s)}\right)\right] - p_{1it}^{o(s)}\left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right]\right\} \end{aligned}$$

and

$$\begin{aligned} \xi_{i2nt} &= \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[\frac{\partial U_{0it+s}\left(\omega_{1t}^{(s)}\right)}{\partial h_t} + \frac{\partial \varphi_0\left(p_{1it}^{o(s)}\right)}{\partial h_t}\right] + \frac{\partial p_{1it}^{o(s)}}{\partial h_t}\left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right] \\ & + p_{1it}^{(s)} \frac{\partial\left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right]}{\partial h_t} - E_t^o\left\{\sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[\frac{\partial U_{0it+s}\left(\omega_{1t}^{(s)}\right)}{\partial h_t} + \frac{\partial \varphi_0\left(p_{1it}^{(s)}\right)}{\partial h_t}\right]\right. \\ & \left. + \frac{\partial p_{1it}^{o(s)}}{\partial h_t}\left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right] + p_{1it}^{o(s)} \frac{\partial\left[Q^{-1}\left(p_{1it}^{o(s)}\right) + \varphi_1\left(p_{1it}^{o(s)}\right) - \varphi_0\left(p_{1it}^{o(s)}\right)\right]}{\partial h_t}\right\} \end{aligned}$$

Given these definitions, the result follows immediately. ■

Note that the data is informative about E_t^o under Assumption (1.6). The identifiability of the structural functions depends on whether these functions can be deduced from knowledge of $E_t^o[Y_{i1nt}|\omega_{nt}]$ and $E_t^o[Y_{i2nt}|\omega_{nt}, d_{nt} = 1]$. The following lemma establishes the main identification result.

Lemma 1.4 *Under Assumptions (1.2)–(1.6), $u_{0i}(z_{nt}, \zeta_t)$ is identified up to an additive constant, and $u_{i1}(l_{nt}, z_{nt})$ is identified up to an additive function of z_{nt}.*

The above Lemma implies that the levels of non-market-hours utility are not identified, but the marginal utility of non-market hours is identified.

Proof of Lemma 1.4.

To establish the results, we prove that

$$u_{0i}(z_{nt}, \zeta_t) = C_{1it} - E_t^o[Y_{i1nt}|\omega_{nt}] + \frac{1}{2}\int_0^{h_{nt}} \left\{ E_t^o[Y_{i2nt}|\omega_{nt}] + \frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h} \right\} \mathrm{d}h \tag{29}$$

and

$$u_{1i}(l_{nt}, z_{nt}) = C_{2i}(z_{nt}) + \frac{1}{2}\int_0^{h_{nt}} \left\{ E_t^o[Y_{i2nt}|\omega_{nt}] + \frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h} \right\} \mathrm{d}h, \tag{30}$$

where

$$C_{1it} = E_t^o \left\{ \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[u_{i1}(1, z_{1nt}^{(s)}) - u_{i1}(1, z_{0nt}^{(s)}) \right] \right\}$$

and $C_{2i}(z_{nt}) = u_{i1}(1, z_{nt})$. By applying the results from Chesher (2007) and using the above results, we obtain our functional $\mathcal{F}^{-1}(F_{Y|X})$.

Taking expectations of equations (24) and (25) gives

$$\begin{aligned} E_t^o[Y_{i1nt}|\omega_{nt}] &= u_{i1}(1, z_{nt}) - u_{i0}(z_{nt}, \zeta_t) - u_{1i}(l_{nt}^*, z_{nt}) \\ &+ E_t^o\{ \sum_{s=1}^{\rho(\omega_{nt})} \beta^s [u_{i1}(1, z_{1nt}^{(s)}) - u_{1i}(1, z_{0nt}^{(s)})]\} \end{aligned} \tag{31}$$

$$E_t^o[Y_{i2nt}|\omega_{nt}] = -\frac{\partial u_{1i}(l_{nt}^*, z_{nt}, H_{nt-1})}{\partial h_t} - E_t^o \left\{ \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[\frac{u_{i1}(1, z_{1nt}^{(s)}, H_{1ns-1}^{(s)})}{\partial h_t} \right] \right\} . \tag{32}$$

Note that $z_{1ns-1}^{(s)}$ is a function of h_{nt} while $z_{0ns-1}^{(s)}$ is not. Hence, taking the derivative of (31) with respect to h_{nt} gives

$$\frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h_{nt}} = -\frac{\partial u_{i1}(l_{nt}^*, z_{nt})}{\partial h_t} + E_t^o \left\{ \sum_{s=1}^{\rho(\omega_{nt})} \beta^s \left[\frac{u_{1i}(1, z_{1nt}^{(s)})}{\partial h_t} \right] \right\}, \tag{33}$$

which implies that

$$E_t^o\left\{\sum_{s=1}^{\rho(\omega_{nt})}\beta^s\left[\frac{u_{1i}(1,z_{1nt}^{(s)})}{h_t}\right]\right\}=\frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h_{nt}}+\frac{\partial u_{i1}(l_{nt}^*,z_{nt})}{h_t}. \tag{34}$$

Substituting (34) into (25) gives

$$E_t^o[Y_{i2nt}|\omega_{nt}]=-2\frac{\partial u_{i1}(l_{nt}^*,z_{nt})}{\partial h_t}-\frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h_{nt}}. \tag{35}$$

Rearranging, we get

$$\frac{\partial u_{i1}(l_{nt}^*,z_{nt})}{\partial h_t}=-\frac{1}{2}\left\{\frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h_{nt}}+E_t^o[Y_{i2nt}|\omega_{nt}]\right\}. \tag{36}$$

Integrating up to h_{nt} gives

$$u_{i1}(l_{nt}^*,z_{nt})=u_{i1}(1,z_{nt})-\frac{1}{2}\int_0^{h_{nt}}\left\{\frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h}+E_t^o[Y_{i2nt}|\omega_{nt}]\right\}\mathrm{d}h \tag{37}$$

Let

$$C_{1it}=E_t^o\left\{\sum_{s=1}^{\rho(\omega_{nt})}\beta^s\left[u_{1i}(1,z_{1nt}^{(s)})-u_{i1}(1,z_{0nt}^{(s)})\right]\right\},$$

then substituting (37) into (31) and rearranging gives

$$u_{0i}(z_{nt},\zeta_t)=C_{1it}-E_t^o[Y_{i1nt}|\omega_{nt}]+\frac{1}{2}\int_0^{h_{nt}}\left\{E_t^o[Y_{i2nt}|\omega_{nt}]+\frac{\partial E_t^o[Y_{i1nt}|\omega_{nt}]}{\partial h}\right\}\mathrm{d}h \tag{38}$$

and we obtain the result in the Lemma. ■

2 Estimation

2.1 Estimation of Consumption and Earnings Equations

In the first step, we use the Euler equation for consumption to form the moment condition:

$$\frac{\partial u_{2i}(c_{nt},x_{nt},\varepsilon_{2nt},\theta_c)}{\partial c_{nt}}=\eta_n\lambda_t. \tag{39}$$

Here, we are assuming that the functional form of $u_2()$ is known up to a finite-dimensional parameter vector, θ_c. Recall that we assume that

$$u_{2i}(c_{nt},x_{nt},\varepsilon_{2nt},\theta_c)=\exp(x_{nt}'B_4+\varepsilon_{2nt})c_{nt}^{\alpha}/\alpha.$$

Let $\triangle$ denote the first-difference operator. Taking the logarithm of each side of this expression, differencing, and rearranging implies

$$(40) \qquad (1-\alpha)^{-1} \triangle \varepsilon_{2nt} = \triangle \ln(c_{nt}) - (1-\alpha)^{-1} \triangle x'_{nt} B_4 + \triangle (1-\alpha)^{-1} \ln(\lambda_t).$$

Let Θ_c denote the $(K+T-1)$-dimensional vector of parameters to be estimated, defined as

$$\Theta_c = \begin{pmatrix} (1-\alpha)^{-1} B_4 \\ \triangle(1-\alpha)^{-1}\ln(\lambda_2) \\ \vdots \\ \triangle(1-\alpha)^{-1}\ln(\lambda_T) \end{pmatrix}.$$

We also define $Y_n = (\triangle \ln(c_{n2}), \ldots, \triangle \ln(c_{nT}))'$ as a vector of endogenous variables and Z_n^c as the exogenous variables:

$$Z_n^c = \begin{bmatrix} \triangle x'_{n2} & D_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \triangle x'_{nT} & 0 & \ldots & D_T \end{bmatrix},$$

where D_t denotes a time dummy for $t \in \{2, \ldots, T\}$. The assumptions in Section 2 imply that the unobserved variable ε_{5nt} is independent of individual-specific characteristics. Therefore $E((1-\alpha)^{-1} \triangle \varepsilon_{2nt} \mid x_{nt}) = 0$. Using equation (40), one can obtain a set of orthogonality conditions,

$$E\left[(Y_n - Z_n^c \Theta_c)\, Z_n^c\right] = 0,$$

that can be exploited to estimate Θ_c using an optimal instrumental-variable estimation technique.

We use a traditional fixed-effect estimator to estimate $(1-\alpha)^{-1}\ln(\eta_n)$. Let T_1 be the number of time periods for which the marginal utility of consumption equation is estimated. Let

$$(41) \qquad (1-\alpha)^{-1}\ln(\eta_n) \equiv \sum_{t \in T_1} \left[\ln(c_{nt}) - (1-\alpha)^{-1} x'_{nt} B_4 + (1-\alpha)^{-1}\ln(\lambda_t)\right] / T_1.$$

The fixed-effects estimates of $(1-\alpha)^{-1}\ln(\eta_n)$ are obtained as the simple time averages of the estimated residuals of the consumption equation, which correspond to the sample counterparts of $(1-\alpha)^{-1}\ln(\eta_n)$ defined above. In order to form the sample counterpart of (41), we need an estimate of $\left\{(1-\alpha)^{-1}\ln(\lambda_t)\right\}_{t=1}^{T_1}$. From the estimate of Θ_c, however, we can only obtain estimates of $\left\{\triangle(1-\alpha)^{-1}\ln(\lambda_2)\right\}_{t=2}^{T_1}$. This requires us to make the additional assumption that $E_n[\eta_n \mid x_{nt}] = 0$, where $E_n[.]$ is the expectation operator over individuals. This assumption enables us to obtain an estimate of $(1-\alpha)^{-1}\ln(\lambda_1)$ as the sample analogue of

$$(1-\alpha)^{-1}\ln(\lambda_1) = -E_n\left[\ln(c_{n1}) - (1-\alpha)^{-1} x'_{n1} B_4\right].$$

We now have estimates of $\left\{(1-\alpha)^{-1}\ln(\lambda_t)\right\}_{t=1}^{T_1}$ and $(1-\alpha)^{-1}\ln(\eta_n)$, enabling us to recover α in the third step of our estimation.

Next, we turn our attention to the estimation of the earnings equations. Let $d_{n\tau t} = I_{n\tau t} \times d_{nt}$. The following transformed zero profit condition holds (see Appendix B.2 in the paper for details):

$$E_t[\widetilde{S}_{in\tau t} - K_{\tau t} - b_{1\tau}\widetilde{h}_{\tau nt} - b_{2\tau}\widetilde{h^2_{\tau nt}} - \sum_{r=1}^{\rho} b_{3r\tau}\widetilde{h}_{\tau nt-r} - \textstyle\sum_{r=1}^{\rho} b_{\tau 4r}\widetilde{d}_{\tau nt-r} - \widetilde{Z}_{\tau nt}{}' B_{5\tau} \\ - \widetilde{d_{nt+1}I_{n\tau t+1}}\beta\gamma_\tau | \{\omega^*_{nt} \backslash \nu_n\}, i, d_{nt}I_{n\tau t} = 1] = 0 \tag{42}$$

where $\widetilde{}$ denotes *deviations-from-time means* and let $b_\tau = (b_{\tau 1}, b_{\tau 2}, b_{\tau 31}, \ldots, b_{\tau 4\rho})$.

Let $\Theta_{w\tau}$ define the vector of parameters to be estimated as ,

$$\Theta_{w\tau} = \begin{pmatrix} b_\tau \\ B_{\tau 5} \\ \beta\gamma_\tau \\ K_{\tau 2} \\ \vdots \\ K_{\tau T} \end{pmatrix}.$$

We also define $Y_{n\tau} = (d_{n\tau 1}\widetilde{S}_{\tau n1}, \ldots, d_{n\tau T}\widetilde{S}_{\tau nT})'$ as a vector of endogenous variables and $X_{\tau n}$ as the exogenous variables,

$$X_{n\tau} = \begin{bmatrix} \widetilde{x}'_{\tau 2} & D_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{x}'_{\tau T} & 0 & \ldots & D_T \end{bmatrix},$$

where $\widetilde{x}'_{\tau nt} = d_{n\tau t}(\widetilde{h}_{\tau nt}, \widetilde{h}^2_{\tau nt}, \widetilde{h}_{\tau nt-1}, \ldots, \widetilde{h}_{\tau nt-\rho}, \widetilde{d}_{\tau nt-1}, \ldots, \widetilde{d}_{\tau nt-\rho}, \widetilde{Z}_{\tau nt}, \widetilde{d_{nt+1}I_{n\tau t+1}})$. Letting Z_n be the matrix of conditioning variables and using equation (42), one can obtain a set of orthogonality conditions:

$$E\left[(Y_{n\tau} - X_{n\tau}\Theta_{w\tau}) Z_n\right] = 0,$$

which can be exploited to estimate $\Theta_{w\tau}$ using an optimal instrumental-variable technique. The aggregate effect and fixed effect in the earnings equation are estimated in a similar way to those in the consumption equation.

2.1.1 Conditional Choice Probabilities and Beliefs

There are five inputs of equations (24)–(25) to be estimated before we can form the empirical counterparts of Y_{i1nt} and Y_{i2nt}. First, Y_{i1nt} is a function of the equilibrium salary schedule, which is a function of the employers' beliefs. These beliefs will be estimated nonparametrically. Second, Y_{i2nt} is a function of the derivative of the equilibrium salary schedule with respect to current hours; we estimate this derivative nonparametrically. Third, Y_{i1nt} is a function of the current conditional choice probabilities, $p^0_{int}(\omega_{nt})$, which we will also estimate nonparametrically. Finally, Y_{i1nt} and Y_{i2nt} are functions of $p^{o(s)}_{kint}$ and their derivatives, respectively, which will also be estimated nonparametrically.

Estimation of the Equilibrium Beliefs and their Derivatives The equilibrium beliefs for each occupation, $\widetilde{p}_{in\tau t+1}$, are computed as a nonlinear regression of the product of next-period participation and occupation choice index, $d_{nt+1} \times I_{n\tau t+1}$ on today's public information variables, z_{nt}^{p}, work histories, H_{nt-1}, and hours worked, h_{nt}, conditional on working today in occupation τ. Let $X_{nt} = (z_{nt}^{\mathcal{P}}, H_{nt-1}, h_{nt}, \nu_n, Gender_n)$ and $NY_{n\tau t-1}$ be the total number of years worked in occupation τ up to period $t-1$. Only two occupations are used in the estimation, so $\tau \in \{1,2\}$. The labor-market history used in this paper is defined as

$$(43) \quad H_{nt-1} = (NY_{n1t-1}, NY_{n2t-1}, d_{nt-3}I_{n1t-3}, d_{nt-3}I_{n2t-3}, \ldots, d_{nt-1}I_{n1t-1}, d_{nt-1}I_{n2t-1}, h_{nt-3}, \ldots, h_{nt-1})$$

Let $J_1[\delta_{1N}^{-1}(X_{nt} - X_{n's})]$ denote a kernel where δ_N is the bandwidth associated with each argument. The nonparametric estimate of $\widetilde{p}_{in\tau t+1}$, denoted $\widetilde{p}_{in\tau t+1}^{N}$, is computed using the kernel estimator:

$$(44) \qquad \widetilde{p}_{in\tau t+1}^{N} = \frac{\sum_{n'=1, n'\neq n}^{N} \sum_{s=1}^{T-1} i_{n'} d_{n's+1} I_{n'\tau s+1} d_{ns} I_{n'\tau s} J_1[\delta_{1N}^{-1}(X_{nt} - X_{n's})]}{\sum_{n'=1, n'\neq n}^{N} \sum_{s=1}^{T-1} i_{n'} d_{n's} I_{n'\tau s} J_1[\delta_{1N}^{-1}(X_{nt} - X_{n's})]}.$$

The derivative is then estimated using the standard nonparametric derivative kernel estimator (see Pagan and Ullah, 1999).

Estimation of the Conditional Choice Probabilities The estimation of the conditional choice probabilities requires us to be more specific about the state variables. In contrast to the beliefs, the conditional choice probabilities are defined from the workers' perspective and not the firms perspective. From the estimation of the consumption equation, $\eta_n \lambda_t$ is known up to a proportionality constant. The elements included in z_{nt} are *number of individuals in the family unit*, *number of children younger than three*, *number of children between three and fourteen*, *age*, *years of completed education*, *marital status*, *spouse's years of education (if married)*, and *gender*.

The conditional choice probabilities, p_{int}, are computed as nonlinear regressions of the participation index, d_{nt}, on the current state, $\omega_{nt}^{N} \equiv (z'_{nt,} H_{nt-1}, \eta_n^N \lambda_t^N)'$, where the N superscript denotes an estimated quantity. We denote by $J\left[\delta_N\left(\omega_{nt}^N - \omega_{n's}^N\right)\right]$ the kernel and by δ_N the bandwidth associated with each argument. The nonparametric estimate of p_{int}, denoted by p_{int}^N, is computed using the kernel estimator:

$$(45) \qquad p_{int}^{N} = \frac{\sum_{n'=1, n'\neq n}^{N} \sum_{s=1}^{T} i_{n'} d_{n's} J\left[\delta_N^{-1}\left(\omega_{nt}^N - \omega_{n's}^N\right)\right]}{\sum_{n'=1, n'\neq n}^{N} \sum_{s=1}^{T} i_{n'} J\left[\delta_N^{-1}\left(\omega_{nt}^N - \omega_{n's}^N\right)\right]}.$$

2.2 Estimation of the Finite-State Path Probabilities and their Derivatives

Recall that $p_{kit}^{(s)} = E[d_{t+s}|\omega_{kt}^{(s)}]$, hence it can be estimated as nonlinear regressions of the participation index, d_{nt}, on the hypothetical state, $\omega_{knt}^{(s)\mathcal{N}}$, conditional on $d_{knt}^{(s)} = 1$. Specifically,

$$p_{iknt}^{(s,N)} = \frac{\sum_{n'=1,n'\neq n}^{N} \sum_{r=1}^{T} i_{n'} d_{n'r} d_{kn'rs}^{(s)} J\left[\delta_N^{-1}\left(\omega_{knt}^{(s)\mathcal{N}} - \omega_{n'r}^{\mathcal{N}}\right)\right]}{\sum_{n'=1,n'\neq n}^{N} \sum_{r=1}^{T} i_{n'} d_{kn'r}^{(s)} J\left[\delta_N^{-1}\left(\omega_{knt}^{(s)\mathcal{N}} - \omega_{n'r}^{\mathcal{N}}\right)\right]}, \tag{46}$$

To evaluate the term $\partial p_{i1nt}^{(s)}/\partial h_{nt}$, which appears in the definition of Y_{i2nt}, define

$$f_{i1nt}^{(s)} \equiv f_{i1}\left(\omega_{1nt}^{(s)} \mid d_{nt+s} = 1\right) \tag{47}$$

to be the probability density function for $\omega_{1nt}^{(s)}$, conditional on participating at date $t+s$. Likewise, let $f_{int}^{(s)} \equiv f_i\left(\omega_{1nt}^{(s)}\right)$ be the related probability density that is not conditioned on participating in period $t+s$ for $s = 1,\ldots,3$. Denote their derivatives with respect to h_{nt}^* by $f_{i1nt}^{'(s)}$ and $f_{int}^{'(s)}$, respectively. We can then show that

$$\frac{\partial p_{i1nt}^{(s)}}{\partial h_{nt}} = \left[\frac{f_{i1nt}^{'(s)}}{f_{i1nt}^{(s)}} - \frac{f_{int}^{'(s)}}{f_{int}^{(s)}}\right] p_{1nt}^{(s)}, \quad s = 1,\ldots,3. \tag{48}$$

We derive this expression using the representation of $p_{i1nt}^{(s)}$ as $p_{i1nt}^{(s)} = \Pr\left(d_{nt+s} = 1 \mid \omega_{1nt}^{(s)}\right) = \Pr(d_{nt+s} = 1) f_{i1nt}^{(s)}/f_{int}^{(s)}$. Differentiating this expression with respect to h_{nt} yields the above expression. The nonparametric estimates of $f_{i1nt}^{(s)}$ and $f_{int}^{(s)}$ are defined, respectively, as the numerators and denominators of $p_{iknt}^{(s)N}$ in equation (48). The estimates of $f_{i1nt}^{'(s)}$ and $f_{int}^{'(s)}$ are obtained from the derivatives of the estimates, $f_{i1nt}^{(s)N}$ and $f_{int}^{(s),N}$, with respect to h_{nt} (Pagan and Ullah, 1999).

2.3 Estimation of the Final Stage

Note that from the second step, we have estimates of $b_{\tau 1}$, $b_{\tau 2}$, β,γ_τ, and all the other parameters of the production function. In addition, from the first step, we have an estimate of ϕ_{nt},

$$\phi_{nt} = (1-\alpha)^{-1} \ln(\eta_n \lambda_t).$$

The third step yields estimates of p_{nt}, $p_{1nt}^{(s)}$, $\widetilde{p}_{n\tau t+1}$, $\frac{\partial p_{1nt}^{(s)}}{\partial h_{nt}}$, and $\frac{\partial \widetilde{p}_{n\tau t+1}}{\partial h_{nt}}$. Substituting these into equations (34, paper) and (35, paper), we can form the moment conditions:

$$m_{2nt}\left(\Theta u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right) = \exp\left((1-\alpha)\phi_{nt}^{(N)}\right)\sum_{\tau=1}^{M} I_{n\tau t}[y_{\tau t}(h_{nt}, z_{nt}^{\mathcal{P}}, \theta_e^{(N)}) - \gamma_\tau^{(N)} + \beta\gamma_\tau^{(N)}\widetilde{p}_{n\tau t+1}^{(N)}]$$

$$+\sigma \ln\left[p_{nt}^{(N)} / \left(1-p_{nt}^{(N)}\right)\right] + \sigma\sum_{s=1}^{\rho}\beta^s \ln\left(\frac{1-p_{1nt}^{(s)(N)}}{1-p_{0nt}^{(s)(N)}}\right)$$

$$+\zeta_t + \sum_{s=1}^{2}\kappa_{is}d_{t-s} + x_t' B_{i1} - x_t' h_t B_{i2} - \theta_{0i}\left(1-l_t^2\right) - \sum_{s=1}^{2}\theta_{si}h_t(l_{t-s}+\beta^s) \tag{49}$$

and

$$m_{3nt}\left(\Theta u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right) = d_{nt}\left\{\sigma\sum_{s=1}^{\rho}\beta^s\left(1-p_{1nt}^{(s)(N)}\right)^{-1}\frac{\partial p_{1nt}^{(s)(N)}}{\partial h_{nt}} - x_t' B_{i2} - 2\theta_{0i}l_t + \sum_{s=1}^{2}\theta_{si}(l_{t-s}+\beta^s)\right.$$

$$\left.\exp\left((1-\alpha)\phi_{nt}^{(N)}\right)\sum_{\tau=1}^{M} d_{n\tau t}[b_{\tau 1}^{(N)} + 2b_{\tau 2}^{(N)}h_{nt} + \beta\gamma_\tau^{(N)}\frac{\partial\widetilde{p}_{n\tau t+1}^{(N)}}{\partial h_{nt}}]\right\}, \tag{50}$$

where $\psi^{(N)} = \left(p_{nt}^{(N)}, p_{0nt}^{(s)(N)}, p_{1nt}^{(s)(N)}, \widetilde{p}_{n\tau t+1}^{(N)}\right)$ are the nonparametric second-step estimates and $\Theta_u = (\sigma, \alpha, \beta, \zeta_1, \ldots, \zeta_T, B_1, B_2, \theta_0, \ldots, \theta_\rho, \kappa_1, ..., \kappa_{i\rho})$ are the structural parameters left to be estimated.

There are now two sources of errors in evaluating the sample counterparts of (49) and (50). The first is the forecast errors from replacing the expectations of future variables with their realizations. The second is the approximation error that arises from replacing the true values of the conditional choice probabilities, conditional expectation, and time-invariant individual-specific effects with their estimates. Let us define the 2×1 vector

$$m_{4nt}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right) \equiv \left[m_{2nt}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right), m_{3nt}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)\right]'$$

and let T_3 denote the set of periods for which the hours and participation equations are valid. Define the vector

$$m_{4n}^{(N)}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right) \equiv \left(m_{4n1}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)', \ldots, m_{4nT_3}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)'\right)'$$

as the vector of the idiosyncratic errors for a given individual over time. Define

$$\Omega_{nt}^{(N)} \equiv E_t\left[m_{4nt}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right) m_{4nt}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)'\right]$$

. The off-diagonal elements of $\Omega_{nt}^{(N)}$ are zero because

$$E_t\left[m_{4nt}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right) m_{4nr}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)'\right] = 0 \text{ for } r \neq t, r < t.$$

The 2×2 conditional heteroskedasticity matrix $\Omega_{nt}^{(N)}$ associated with the individual-specific errors, $m_{4nt}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)$, is evaluated using a nonparametric estimator based on the estimated moments, $m_{4nt}\left(\Theta_{1u}^{(N)}, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)$, derived from an initial consistent estimate of $\Theta_{1u}^{(N)}$. The optimal instrumental-variables estimator for $\Theta_u^{(N)}$ is

$$(51)\quad \Theta_u^{(N)} \equiv \arg\min_{\Theta u} \frac{\sum_{n=1}^{N} m_{4n}^{(N)}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)\left(\Omega_n^{(N)}\right)^{-1} m_{4n}^{(N)}\left(\Theta_u, \Theta_c^{(N)}, \Theta_e^{(N)}, \psi^{(N)}\right)}{N}.$$

2.4 Asymptotic Properties

It is well known in the econometric literature that under certain regularity conditions, pre-estimation does not have any impact on the consistency of the parameters in the subsequent steps of a multistage estimation (Newey, 1984; Newey and McFadden, 1994; Newey, 1994). The asymptotic variance, however, is affected by the pre-estimation. In order to conduct inference in this type of estimation, one has to correct the asymptotic variance for the pre-estimation. The method used for correcting the variance in the final step of estimation depends on whether the pre-estimation parameters are of finite or infinite dimension. Unfortunately, our estimation strategy combines both finite- and infinite-dimensional parameters. Combining results from two sources (Newey, 1984; Newey and McFadden, 1994), however, allows us to derive the corrected asymptotic variance for our estimator.

Following Newey (1984), we can write the sequential-moments conditions for the first- and third-step estimation as a set of joint moment conditions:

$$m_n(\Theta_u, \Theta_c, \Theta_e, \psi) = \begin{bmatrix} (Y_n - Z_n\Theta_c)\, Z_n^c \\ (Y_{n1} - X_{1n}\Theta_{wP})\, Z_n \\ (Y_{n2} - X_{Mn}\Theta_{eNP}) Z_n \\ m_{4n}(\Theta_u, \Theta_c, \Theta_e, \psi) \end{bmatrix},$$

where $(Y_n - Z_n\Theta_c)\, Z_n^c$ is the orthogonality condition from the estimation of the consumption equation, $(Y_{n\tau} - X_{n\tau}\Theta_{w\tau})Z_n$ is the orthogonality condition from the estimation of the earnings equation, and $m_{4n}(\Theta_u, \Theta_c, \Theta_e, \psi)$ is the moment condition from the third-step estimation. Let $\Theta = (\Theta_u, \Theta_c, \Theta_e)'$, with the true value denoted by Θ_0. Note that each element of ψ is a conditional expectation. Redefine each element as $\psi^j(z^j) = f_{z^j}(z^j)E\left[\widetilde{d}_n^j \mid z^j\right]$, where $\widetilde{d}_{nt}^j = [1, d_{nt}]'$ for the estimation of p_{nt}, $\widetilde{d}_{nt}^j = [d_{knt}^{(r)}, d_{knt}^{(r)} d_{nt}]'$ for the estimation of $p_{knt}^{(r)}$, and $\widetilde{d}_{nt}^j = [d_{n\tau t}, d_{n\tau t} d_{n\tau t+1}]'$ for the estimation of $\widetilde{p}_{n\tau t+1}$. Therefore, $\psi^{j(N)}(z^j) = \frac{1}{N}\sum_{n=1}^{N} \widetilde{d}_n^j J_{\delta_N}(z^j - z_n^j)$. The conditions below ensure that $\psi^{(N)}$ is close enough to ψ_0 for N large enough, in particular that $\sqrt{N}\left\|\psi^{(N)} - \psi_0\right\|^2$ converges to zero.

> **A3:** *There is a version of $\psi_0(z)$ that is continuously differentiable of order κ, greater than the dimension of z and $\psi_{10}(z) = f_z(z)$ is bounded away from 0.*

A4: $\int J(u)\,\mathrm{d}u = 1$ *and for all* $j < \kappa$, $\int J(u)\left(\bigotimes_{s=1}^{j} u\right)\mathrm{d}u = 0$.

A5: *The bandwidth,* δ_N, *satisfies* $N\delta_N^{2\dim(z)}/(\ln(N))^2 \to \infty$ *and* $N\delta_N^{2\kappa} \to 0$.

A6: *There exists a* $\Psi(\omega)$, $\epsilon > 0$, *such that*

$$\|\nabla_\Theta m_n(\omega.\Theta,\psi) - \nabla_\Theta m_n(\omega.\Theta_0,\psi_0)\| \leq \Psi(\omega)\left[\|\Theta-\Theta_0\|^\epsilon + \|\psi-\psi_0\|^\epsilon\right]$$

and $E[\Psi(\omega)] < \infty$.

A7: $\Theta^{(N)} \to \Theta_0$ *with* Θ_0 *in the interior of its parameter space.*

A8: (Boundedness)

(i) Each element of $m_n(\Theta,\psi)$ is bounded almost surely: $E[\|m_n(\Theta,\psi)\|^2] < \infty$;

(ii) $E[Z'_{n'}Z_n] < \infty$, $E[X'_{\tau n}Z_n] < \infty$, $E[\exp((1-\alpha)\phi_{nt})] < \infty$, $E[z_{nt}] < \infty$, $E[y_{\tau t}(h_{nt}, H_{nt-1}, z^p_{nt}, \theta_e)] < \infty$, $\gamma_\tau < \infty$, $E[\nabla_{h_{nt}}\widetilde{p}_{n\tau t+1}] < \infty$, $E[X_{n\tau}] < \infty$ for $\tau = 1,2$;

(iii) $p_{nt}, p^{(r)}_{knt}, \widetilde{p}_{n\tau t+1}, \in (0,1)$, for $k \in \{0,1\}$, $r = 1,\ldots,\rho$, and $\tau = 1,2$;

(iv) $E[\nabla_h f_{z^j}(z^j)] < \infty$ and $E[\nabla_h E[\widetilde{d}^j_n \mid z^j]] < \infty$;

Theorem 1 *Under A1–A8 and* $\Phi(\omega)$, *defined below,*

$$\sqrt{N}\left(\Theta^{(N)} - \Theta_0\right) \Rightarrow N(0,\Sigma(\Theta_0)),$$

where

$$\begin{aligned}\Sigma(\Theta_0) &= E\left[\nabla_\Theta m_n(\omega)\Omega_n^{-1}\nabla_\Theta m_n(\omega)'\right]^{-1} \\ &\quad \times E\left[\nabla_\Theta m_n(\omega)\Omega_n^{-1}\left\{m_n(\omega)+\Phi(\omega)\right\}\left\{m_n(\omega)+\Phi(\omega)\right\}'\Omega_n^{-1}\nabla_\Theta m_n(\omega)'\right] \\ &\quad \times E\left[\nabla_\Theta m_n(\omega)\Omega_n^{-1}\nabla_\Theta m_n(\omega)'\right]^{-1}.\end{aligned}$$

Assumptions A3–A8 are standard in the semiparametric literature, see Newey and McFadden (1994) for details. One can now use Theorem 1 to calculate the standard errors for all the parameters in our estimation.

The proof of Theorem 1 follows from checking the conditions for Theorem 8.12 in Newey and McFadden (1994).

Proof of Theorem 1. We first check the various boundedness requirements of Theorem 8.12 in Newey and McFadden (1994). By assumption A8(i), we have that $E[\|m_n(\Theta,\psi)\|^2] < \infty$. It is obvious, from inspection, that $m_n(\Theta,\psi)$ is continuously differentiable in Θ and by A8(ii–iv) that $E[\nabla_\Theta m_n(\Theta,\psi)] < \infty$. Additionally, $\nabla_{\psi\psi} m_n(\Theta_0,\psi_0)$ is bounded: $E[\|\nabla_{\psi\psi} m_n(\Theta_0,\psi_0)\|] < \infty$.

Second, consider a point-wise Taylor expansion for the j^{th} element of m_n,

$$\begin{aligned}m^j(\omega,\psi) &= m^j(\omega,\psi_0) + \nabla_\psi m^j(\omega,\psi_0)(\psi(z)-\psi_0(z)) \\ &\quad + (\psi(z)-\psi_0(z))'\nabla_{\psi\psi} m^j(\omega,\psi_0)(\psi(z)-\psi_0(z)) + o(\|\psi(z)-\psi_0(z)\|^2),\end{aligned}$$

where the norm over ψ is the sup-norm. Next, note that

$$\begin{aligned}\left|m^j(\omega,\psi) - m^j(\omega,\psi_0)\nabla_\psi m^j(\omega,\psi_0)(\psi(z) \; - \; \psi_0(z))\right| & \\ \leq & \left\|(\psi(z)-\psi_0(z))'\nabla_{\psi\psi}m^j(\omega,\psi_0)(\psi(z)-\psi_0(z))\right\| \\ & + o(\|\psi(z)-\psi_0(z)\|^2) \\ \leq & \|\psi-\psi_0\|^2 \left\|\nabla_{\psi\psi}m^j(\omega,\psi_0)\right\| + o(\|\psi-\psi_0\|^2),\end{aligned}$$

using the triangle inequality and the Cauchy-Schwartz inequality. Therefore, for $\|\psi - \psi_0\|$ small enough,

$$\left|m^j(\omega,\psi) - m^j(\omega,\psi_0) - \nabla_\psi m^j(\omega,\psi_0)(\psi(z)-\psi_0(z))\right| \leq \|\psi-\psi_0\|^2 \left\|\nabla_{\psi\psi}m^j(\omega,\psi_0)\right\|.$$

So that

$$\begin{aligned}\|m(\omega,\psi) - m(\omega,\psi_0) - \nabla_\psi m(\omega,\psi_0)(\psi(z)-\psi_0(z))\| &\leq \|\psi-\psi_0\|^2 \|\nabla_{\psi\psi}m(\omega,\psi_0)\| \\ \|m(\omega,\psi) - m(\omega,\psi_0) - \nabla_\psi m(\omega,\psi_0)(\psi(z)-\psi_0(z))\| &\leq \|\psi-\psi_0\|^2 \|\nabla_{\psi\psi}m(\omega,\psi_0)\|\end{aligned}$$

Hence $\Gamma(\omega,\psi-\psi_0) = \nabla_\psi m(\omega,\psi_0)(\psi(z)-\psi_0(z))$ and $\Psi(\omega) = \|\nabla_{\psi\psi}m(\omega,\psi_0)\|$. It follows that both $\Gamma(\omega,\psi-\psi_0)$ and $\Psi(\omega)$ are bounded from the boundedness conditions established above.

Next we establish the form of the influence function. Note that we have

$$\begin{aligned}\int \Gamma(\omega,\psi)F_0(\,\mathrm{d}\omega) &= \int f_z(z)E[\nabla_\psi m(\omega,\psi_0) \mid z]\psi(z)\,\mathrm{d}z \\ &= \int \upsilon(z)\psi(z),\end{aligned}$$

where $\upsilon(z) = f_z(z)E[\nabla_\psi m(\omega,\psi_0) \mid z]$. So, by the arguments on page 2208 of Newey and McFadden (1994), we have the influence function for $m(\omega,\psi^{(N)})$:

$$\begin{aligned}\Phi(\omega) &= \upsilon(z) - E\left[\upsilon(z)\tilde{d}\right] \\ &= f_z(z)E\left[\nabla_\psi m(\omega,\psi_0) \mid z\right] - E\left[f_z(z)E[\nabla_\psi m(\omega,\psi_0) \mid z]\tilde{d}\right].\end{aligned}$$

Again by the boundedness of $\nabla_\psi m(\omega,\psi_0)$, it follows that $\int \|\upsilon(z)\|\,\mathrm{d}z < \infty$. Finally Assumption A7 guarantees that the Jacobian term converges. ■

References

[1] Altug, S., and R. A. Miller (1998): "The Effect of Work Experience on Female Wages and Labour Supply," *Review of Economic Studies*, 65, 45–85.

[2] Chesher, A. (2007): "Instrumental Values," *Journal of Econometrics*, 139, 15–34.

[3] Haavelmo, T. M. (1944): "The Probability Approach in Econometrics," *Econometrica*, 12, Supplement, iii–vi 1–115.

[4] Hotz, V. J., and R. A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, 60, 497–529.

[5] Newey, W. (1984): "A Method of Moments Interpretation of Sequential Estimators," *Economic Letters*, 14, 201–206.

[6] Newey, W. (1994): "The Asymptotic Variance of Semiparametric Estimators," *Econometrica*, 16, 1–32.

[7] Newey, W., and D. McFadden (1994): "Large Sample Estimation and Hypothesis Testing," in: R. F. Engle and D. L. McFadden eds. *Handbook of Econometrics*. North-Holland, Amsterdam, pp. 2111–2245.

[8] Pagan, A., and P. Ullah (1999): *Nonparametric Econometrics*. Cambridge University Press, London.

.