Supplementary Material

Supplement to "Effects of parental leave policies on female career and fertility choices"

(*Quantitative Economics*, Vol. 10, No. 3, July 2019, 1195–1232)

Shintaro Yamaguchi
Faculty of Economics, University of Tokyo

Appendix A: Details of data

A.1 *Variable definitions*

A.1.1 *Eligibility status for parental leave* Eligibility for mandated PL is determined by the age of the youngest child, lagged employment sector, and calendar year. Specifically, the legal eligibility status ELG_{it} is given by

$$ELG_{it} = \begin{cases} 1 & \text{if } (a_{k,it} = 0, e_{r,it-1} = 1) \text{ or } (a_{k,it} = 0, e_{n,it-1} = 1, t \geq 2005), \\ 0 & \text{otherwise.} \end{cases}$$

Remember that one can take PL if the child is less than age 1 and the PL taker was employed in the eligible sector in the prior year. The regular sector is the eligible sector throughout the period of analysis, while the nonregular sector became an eligible sector since 2005.

In the counterfactual simulations in Section 7, one can take PL for 3 years in both regular and nonregular sectors. In the simulation of 3-year job protection, the legal eligibility status ELG_{it} is redefined as

$$ELG_{it} = \begin{cases} 1 & \text{if } (a_{k,it} \leq 2, e_{r,it-1} = 1) \text{ or } (a_{k,it} \leq 2, e_{n,it-1} = 1), \\ 0 & \text{otherwise.} \end{cases}$$

A.1.2 *Labor market status* The choice variable for labor market status has four possible, mutually exclusive states. It is determined by the following hierarchical rule. First, I determine if a woman is on parental leave. If not, I examine whether she works in the regular or nonregular sector. If she is not on leave or does not work, I consider she stayed at home.

Parental leave take-up For those who report childbearing, JPSC asks whether an individual took a PL or not. If yes, she is considered on PL for the year. If not, I check her employment status as of October and whether she had a baby. The employment status as of October includes information on whether the respondent is on PL or not, but this answer alone does not seem reliable. Women are considered on PL, if they (1) give birth

Shintaro Yamaguchi: syamaguchi@e.u-tokyo.ac.jp

 https://doi.org/10.3982/QE965

and (2) are on PL or leave other than parental, caregiving, and medical leave as of October.

A woman may be on PL even when she does not deliver a baby, because the leave can be for older children. To determine if their reported PL is correct, I check if they have a child and the age of the youngest child. Ten women report PL as of October, but they have no child. These respondents seem to be on pregnancy leave, because they had a baby in the next year. They are not considered to be on PL.

For those who have a child and report PL, the age of the youngest child is 4 or less. For those who have a child aged 4, the reported PL seems false, because they have a baby in the following year and the child is too old for a PL. They are likely to be on pregnancy leave, not on PL. For two out of three women who have a child age 3, the reported PL seems false for the same reason as above. One exception is the woman with ID number 766, who does not deliver a baby in the following year and works full-time for the whole year. I consider her PL is true. For those who have children aged 1 or 2, I consider their reported PL is all true, because the child is reasonably young for PL and they report PL in the previous year.

Work in the regular and nonregular employment sectors If a woman is considered not on PL according to the criteria above, I determine if she works in the regular or nonregular sector. If a woman works as a regular or nonregular employee as of October, I consider that she works in the reported employment sector for the year. If a woman is employed, reports PL or leave other than parental (caregiving,or medical) and gives birth in the next year, she is considered to work in her reported employment sector. This is because she is likely to be on short pregnancy leave in October and work most of the year.

Stay at home If a woman is considered not on PL and not at work according to the criteria above, I determine if she stays at home. If a woman was on some kind of leave, a homemaker, or did not do any work as of October, she is considered to stay at home.

A.1.3 *Sector-specific experiences* Retrospective labor market status from age 18 is available for the 1997 and newer cohorts in the year they first appear in the survey. It is also available for the 1993 cohort in 1997. Part-time job, dispatched work, and minor paid-work at home are all considered as the nonregular work. The labor market status constructed Section A.1.2 is used to construct the sector specific experiences for years when individuals are surveyed. Years stayed at home is topcoded at ten.

A.1.4 *Other variables* Childbearing is identified if an individual reports that she had a baby or if the reported age of the youngest child is zero. In constructing the number of children and the age of the youngest child, I count all children regardless of whether they live with the survey respondent. This is relatively innocuous, because most children age ten or younger live with their mothers. Years of education is constructed from the completed education level. Junior high-school is 9 years, high school is 12 years, 2-year college and vocational school are 14 years, 4-year university is 16 years, and advanced degree is 18 years. Finally, own and husband's labor income are deflated by the 2010 CPI.

Appendix B: Details of model estimation

This subsection explains the estimation of the structural model. In Section B.1, I define the likelihood function. In Section B.2, I describe the estimation algorithm.

B.1 *The likelihood*

Define d_i as a sequence of choices made by individual i from $\tau(i)+1$ to T_i where $\tau(i)$ and T_i are the first and last years when individual i is observed in the data, respectively, that is, $d_i = (d_{i\tau(i)+1}, d_{i\tau(i)+2}, \dots, d_{iT_i})$. Sequences of own and husband's (male spouse's) earnings are similarly defined and given by y_i and $y_{m,i}$, respectively. Let $\theta = (\theta_1, \dots, \theta_K)$ be a vector of parameters for all K types where θ_k is a vector of parameters for type k. Let $\pi = (\pi_1, \dots, \pi_K)$ be a vector of parameters for type probability. Define $z_{i\tau(i)}$ as a vector of observed characteristics and choice in year $t = \tau(i)$: $z_{i\tau(i)} = (d_{i\tau(i)}, S_{i\tau(i)}, edu_i)$ where edu_i is years of education. The likelihood of observed sequences of choices, own earnings, and husband's earnings conditional on $z_{i\tau(i)}$ is

$$\begin{aligned}
&\mathcal{L}(d_i, y_{m,i}, y_i | z_{i\tau(i)}; \theta, \pi) \\
&\quad = \sum_{k=1}^{K} p_k(z_{i\tau(i)}; \pi) L(d_i, y_{m,i}, y_i | d_{i\tau(i)}, S_{i\tau(i)}; \theta_k),
\end{aligned} \tag{12}$$

where $p_k(\cdot)$ is the probability of being type k and $L(\cdot|\cdot; \theta_k)$ is the conditional likelihood of the sequences given being type k and the observed choices and state variables in the first year (i.e., $t = \tau(i)$).

Given the first-order Markov structure of the model, the likelihood of the observed sequences can be rewritten as a product of probability functions. The parameter vector for type k consists of the subparameter vectors such that $\theta_k = (\theta_k^y, \theta_k^{ym}, \theta_k^u)$, where θ_k^y is a parameter vector for own earnings functions, θ_k^{ym} is a parameter vector for husband's earnings function, and θ_k^u is a parameter vector for the utility function

$$\begin{aligned}
&L(d_i, y_{m,i}, y_i | d_{i\tau(i)}, S_{i\tau(i)}; \theta_k) \\
&\quad = \prod_{t=\tau(i)+1}^{T_i} l(d_{it}, y_{m,it}, y_{it} | S_{it-1}, d_{it-1}; \theta_k) && (13) \\
&\quad = \prod_{t=\tau(i)+1}^{T_i} l^d\big(d_{it} | S_{it}; \theta_k^d, \theta_k^y, \theta_k^{ym}\big) \cdot l^y\big(y_{it} | S_{it}, d_{it}; \theta_k^y\big) \\
&\qquad \cdot l^{ym}\big(y_{m,it} | S_{it-1}, d_{it-1}; \theta_k^{ym}\big), && (14)
\end{aligned}$$

where $l^d(\cdot)$ is the conditional choice probability given the structural model and state variables in year t, $l^y(\cdot)$ is the likelihood of earnings given the state variables and choice in year t, and $l^{ym}(\cdot)$ is the conditional likelihood for earnings of husband in year t given the choice and state variables in the previous year $t-1$, respectively.

The likelihood for individual's own and her husband's earnings is straightforward. Let $\hat{y}_{it}$ and $\hat{y}_{m,it}$ be the predicted values for y_{it} and $y_{m,it}$, respectively. The likelihood for y_{it} and $y_{m,it}$ is given by

$$l^y\big(\ln y_{it}|S_{it}, d_{it}; \theta^y\big) = \phi\big((\ln y_{it} - \widehat{\ln y}_{it})/\sigma_y\big)/\sigma_y, \tag{15}$$

$$l^{ym}\big(y_{m,it}|S_{it}, d_{it}; \theta^{ym}\big) = \phi\big((y_{m,it} - \hat{y}_{m,it})/\sigma_m\big)/\sigma_m, \tag{16}$$

where $\phi(\cdot)$ is the density function for the standard normal distribution and σ_y and σ_m are standard deviations. Note that I model the level, not log, of the husband's earnings to allow for the value zero.

Next, consider the likelihood for employment and fertility choices. The choice-specific error term $\varepsilon^f_{j,it}$ follows a generalized extreme distribution so that error terms may be correlated with each other. Specifically, I use the generalized nested logit model that allows for overlapping nests (see Wen and Koppelman (2001)). There are four nests of alternatives labeled, $B_1, \ldots, B_4$. Nest B_1 includes alternatives for nonconception ($d_{f,it} = 0$) regardless of labor supply choices, nest B_2 includes alternatives for conception ($d_{f,it} = 1$) regardless of labor supply choices, nest B_3 includes alternatives for work ($d_{r,it} = 1$ or $d_{n,it} = 1$) regardless of fertility choices, and nest B_4 includes alternatives for nonwork ($d_{h,it} = 1$ or $d_{l,it} = 1$) regardless of fertility choices. Formally, the nests are defined as

$$\begin{aligned} B_1 = \big\{&(d_{h,it} = 1, d_{f,it} = 0), (d_{r,it} = 1, d_{f,it} = 0), (d_{n,it} = 1, d_{f,it} = 0),\\ &(d_{l,it} = 1, d_{f,it} = 0)\big\}, \end{aligned} \tag{17}$$

$$\begin{aligned} B_2 = \big\{&(d_{h,it} = 1, d_{f,it} = 1), (d_{r,it} = 1, d_{f,it} = 1), (d_{n,it} = 1, d_{f,it} = 1),\\ &(d_{l,it} = 1, d_{f,it} = 1)\big\}, \end{aligned} \tag{18}$$

$$\begin{aligned} B_3 = \big\{&(d_{r,it} = 1, d_{f,it} = 0), (d_{n,it} = 1, d_{f,it} = 0), (d_{r,it} = 1, d_{f,it} = 1),\\ &(d_{n,it} = 1, d_{f,it} = 1)\big\}, \end{aligned} \tag{19}$$

$$\begin{aligned} B_4 = \big\{&(d_{h,it} = 1, d_{f,it} = 0), (d_{l,it} = 1, d_{f,it} = 0), (d_{h,it} = 1, d_{f,it} = 1),\\ &(d_{l,it} = 1, d_{f,it} = 1)\big\}. \end{aligned} \tag{20}$$

Denote by $\bar{V}^f_j(S_{it})$ the choice-specific value less the preference shock $\varepsilon^f_{j,it}$ such that

$$\bar{V}^f_j(S_{it}; \theta) = U^f_j(S_{it}; \theta) + \beta E\big[V(S_{it+1}, \varepsilon_{it+1})|S_{it}, d_{it}\big]. \tag{21}$$

Let $b = 1, \dots, 4$ be an index for a nest. The likelihood of choosing labor supply choice $j \in \{h, r, n, l\}$ and fertility choice $f \in \{0, 1\}$ is given by

$$\begin{aligned} & l^d\big(d_{j,it} = 1, d_{it}^f = 1 | S_{it}; \theta_k^d, \theta_k^y, \theta_k^{ym}\big) \\ & = \frac{\displaystyle\sum_b \big(\mu_b \exp(\bar{V}_j^f)\big)^{1/\lambda_b} \Bigg(\sum_{j,f \in B_b} \big(\mu_b \exp(\bar{V}_j^f)\big)^{1/\lambda_b} \Bigg)^{\lambda_b - 1}}{\displaystyle\sum_{b'=1}^{4} \Bigg(\sum_{j,f \in B_{b'}} \big(\mu_{b'} \exp(\bar{V}_j^f)\big)^{1/\lambda_{b'}} \Bigg)^{\lambda_{b'}}}, \end{aligned} \tag{22}$$

where $\bar{V}_j^f(S_{it}; \theta)$ is denoted as $\bar{V}_j^f$ for brevity. The parameter λ_b is a dissimilarity parameter and indicates the degree of independence among alternatives within the nest. It takes the value between zero and one, and a higher value of λ_b implies greater independence and less correlation. The parameter μ_b is an allocation parameter that reflects the extent to which an alternative is a member of nest b. To facilitate interpretation, it is assumed that $\mu_1 = \mu_2$, $\mu_3 = \mu_4$, and $1 - \mu_1 = \mu_3$.

B.2 *The algorithm*

I first describe the estimation algorithm for the model in which individuals are homogeneous, which is based on Kasahara and Shimotsu (2011). I then explain how this estimation algorithm can be applied to the model in which individuals are heterogeneous, using the ESM algorithm proposed by Arcidiacono and Jones (2003). Following Arcidiacono, Bayer, Bugni, and James (2013), the value function is approximated based on sieves in both cases.

B.2.1 *Homogeneous individuals*

When individuals are homogeneous, the log-likelihood is given by

$$\begin{aligned} & \ln L\big(\{d_i, y_{m,i}, y_i\}_{i=1}^N | \{d_{i\tau(i)}, S_{i\tau(i)}\}_{i=1}^N; \theta\big) \\ & = \sum_{i=1}^N \sum_{t=\tau(i)+1}^{T_i} \ln l^d\big(d_{it} | S_{it}; \theta^d, \theta^y, \theta^{ym}\big) + \ln l^y\big(y_{it} | S_{it}, d_{it}; \theta^y\big) \\ & \quad + \ln l^{ym}\big(y_{m,it} | S_{it-1}, d_{it-1}; \theta^{ym}\big). \end{aligned} \tag{23}$$

Consistent estimates for the parameter vectors θ^y and θ^{ym} are given by

$$\hat{\theta}^y \equiv \arg\max_{\theta^y} \sum_{i=1}^N \sum_{t=\tau(i)+1}^{T_i} \ln l^y\big(y_{it} | S_{it}, d_{it}; \theta^y\big), \tag{24}$$

$$\hat{\theta}^{ym} \equiv \arg\max_{\theta^{ym}} \sum_{i=1}^N \sum_{t=\tau(i)+1}^{T_i} \ln l^{ym}\big(y_{m,it} | S_{it}, d_{it}; \theta^{ym}\big). \tag{25}$$

Note that the consistent estimates for the parameters $\hat{\theta}^y$ and $\hat{\theta}^{ym}$ can be obtained separately from the parameters in the utility function. Because estimation of these parameters $\hat{\theta}^y$ and $\hat{\theta}^{ym}$ is straightforward, I focus on the algorithm for estimating θ^d in the following.

The Bellman equation (10) can be rewritten in terms of the expectation of the value function

$$EV(S_{it}) = E\Big[\max_{j,f} U_j^f(S_{it}) + \varepsilon_{j,it}^f + \beta E\big[V(S_{it+1}, \varepsilon_{it+1})|S_{it}, d_{it}\big]\Big] \tag{26}$$

$$= E\Big[\max_{j,f} U_j^f(S_{it}) + \varepsilon_{j,it}^f + \beta \int EV(S_{it+1})\, dF(S_{it+1}|S_{it}, d_{it})\Big], \tag{27}$$

where expectation is taken over $\varepsilon_{j,it}^f$ and $F(\cdot|\cdot)$ is the cumulative distribution function for S_{it+1}. The Bellman operator is defined by the right-hand side of the above equation so that

$$\big[\Gamma(\theta, EV)\big](S_{it}) \equiv E\Big[\max_{j,f} U_j^f(S_{it}) + \varepsilon_{j,it}^f + \beta \int EV(S_{it+1})\, dF(S_{it+1}|S_{it}, d_{it})\Big]. \tag{28}$$

The Bellman equation (27) is compactly rewritten as $EV = \Gamma(\theta, EV)$. I also define the mapping $\Lambda(\theta, EV)$ as

$$\big[\Lambda(\theta, EV)\big]\big(d_{j,it} = 1, d_{it}^f = 1|S_{it}\big) \equiv l^d\big(d_{j,it} = 1, d_{it}^f = 1|S_{it}; EV, \theta^d, \hat{\theta}^y, \hat{\theta}^{ym}\big). \tag{29}$$

The consistent estimate for the parameter vector θ^d is given by

$$\hat{\theta}^d = \arg\max_{\theta^d} \frac{1}{N} \sum_{i=1}^{N} \ln \Lambda\big(\theta^d, \hat{\theta}^y, \hat{\theta}^{ym}, EV\big) \quad \text{subject to } EV = \Gamma(\theta, EV). \tag{30}$$

Computation of the likelihood function by the nested fixed point algorithm by Rust (1987) requires solving the fixed points of $EV = \Gamma(\theta, EV)$ at each trial parameter value in maximizing the objective function with respect to θ^d. The q-NPL algorithm proposed by Kasahara and Shimotsu (2011) iterates the Bellman operator for only q times rather than finding fixed points.

Define a q-fold operator of Γ as $\Gamma^q(\theta, EV)$. Denote by $\widetilde{EV}(M)$ the estimates for the expected value function in the Mth iteration. Starting from an initial estimate $\widetilde{EV}(0)$ for the expectation of the value function, the q-NPL algorithm iterates the following steps until $\widetilde{EV}$ and $\tilde{\theta}^d$ converge:

1. Given $\widetilde{EV}(M-1)$, update $\tilde{\theta}^d$ by

$$\tilde{\theta}^d(M) = \arg\max_{\theta^d} \frac{1}{N} \sum_{i=1}^{N} \ln \Lambda\big(\theta^d, \hat{\theta}^y, \hat{\theta}^{ym}, \Gamma^q\big(\theta, \widetilde{EV}(M-1)\big)\big). \tag{31}$$

2. Update $\widetilde{EV}$ using the obtained estimate $\tilde{\theta}^d(M)$

$$\widetilde{EV}(M) = \Gamma^q\big(\tilde{\theta}(M), \widetilde{EV}(M-1)\big), \tag{32}$$

where $\tilde{\theta}(M) = (\tilde{\theta}^d(M), \hat{\theta}^y, \hat{\theta}^{ym})$.

Kasahara and Shimotsu (2011) proved that this sequence converges when q is large enough and yields a consistent estimate for θ^d. I tried different values for q and find that $q = 6$ is a good choice in terms of the total computational time for the model and data in this paper.

To further accelerate computation of a model with a large state space, I approximate the Bellman operator by a higher order polynomial function, which is proposed by Arcidiacono et al. (2013). Let $W(S_{it})$ be a vector of polynomials of the state variables. Let ρ be a vector of parameters that approximates the value function. For any state variable S_{it}, the sieve approximation satisfies

$$W(S_{it})'\rho \approx EV(S_{it}). \tag{33}$$

Because the error terms in the utility function follow a generalized extreme value distribution, the closed form solution to $EV(S_{it})$ is given by

$$EV(S_{it}) = \ln\left[\sum_{b'=1}^{4}\left(\sum_{j,f\in B_{b'}}\left(\mu_{b'}\exp\left(\bar{V}_j^f\right)\right)^{1/\lambda_{b'}}\right)^{\lambda_{b'}}\right], \tag{34}$$

which implies that

$$W(S_{it})'\rho \approx EV(S_{it}) \tag{35}$$

$$= \ln\left[\sum_{b'=1}^{4}\left(\sum_{j,f\in B_{b'}}\left(\mu_{b'}\exp\left(U_j^f(S_{it}) + \beta E\left[W(S_{it+1})'\rho|S_{it}, d_{it}\right]\right)\right)^{1/\lambda_{b'}}\right)^{\lambda_{b'}}\right] \tag{36}$$

$$= \ln\left[\sum_{b'=1}^{4}\left(\sum_{j,f\in B_{b'}}\left(\mu_{b'}\exp\left(U_j^f(S_{it}) + \beta E\left[W(S_{it+1})|S_{it}, d_{it}\right]'\rho\right)\right)^{1/\lambda_{b'}}\right)^{\lambda_{b'}}\right]. \tag{37}$$

A key convenience of this approach based on a polynomial function is that the parameter ρ can be taken out of the expectation operator $E(\cdot)$ as it can be seen in the last equality. This can save the computational time, because the expectation of $E[W(S_{it+1})|S_{it}, d_{it}]$ needs to be calculated only once as long as the parameters for transition probabilities remain the same.

B.2.2 *Heterogeneous individuals* In this subsection, I describe the algorithm for the case of heterogeneous individuals. The method described in the last subsection is combined with the EM algorithm developed by Arcidiacono and Jones (2003).

Expectation step In the expectation step, I calculate the conditional probability of being in each unobserved type given the values of the parameters, choices, earnings, and observed state variables. Let $\tilde{\theta}(M-1)$ and $\tilde{\pi}(M-1)$ be the vectors of parameters obtained from the $(M-1)$th iteration. The estimates for the expectation of the value function is denoted by $\widetilde{EV}(M-1)$. The likelihood of the observations on individual i given the parameters at the $(M-1)$th iteration is

$$L_i^{(M-1)} = \mathcal{L}\left(d_i, y_{m,i}, y_i|z_{i\tau(i)}; \widetilde{EV}(M-1), \tilde{\theta}(M-1), \tilde{\pi}(M-1)\right). \tag{38}$$

Similarly, I denote by $L_{ik}^{(M-1)}$ the likelihood of the observations and being type k for individual i so that $L_i^{(M-1)} = \sum_k L_{ik}^{(M-1)}$. At iteration M, following from the Bayes rule, the probability of individual i being type k, $q_{ik}(M)$ is given by

$$q_{ik}(M) = \frac{L_{ik}^{(M-1)}}{L_i^{(M-1)}}. \tag{39}$$

Maximization step The parameter vector is updated to $\tilde{\theta}(M)$ by choosing θ and π to maximize

$$\begin{aligned} &\sum_{i=1}^{N}\sum_{k=1}^{K} q_{ik}(M) \ln \mathcal{L}\big(d_i, y_{m,i}, y_i | d_{i\tau(i)}, S_{i\tau(i)}; \widetilde{EV}(M-1), \theta, \pi\big) \\ &\quad = \sum_{i=1}^{N}\sum_{k=1}^{K} q_{ik}(M) \Bigg(\ln p_k(z_{i\tau(i)}; \pi) + \sum_{t=\tau(i)+1}^{T_i} \ln l^d\big(d_{it} | S_{it}; \widetilde{EV}(M-1), \theta_k\big) \\ &\qquad + \ln l^y\big(y_{it} | S_{it}, d_{it}; \theta_k^y\big) + \ln l^{ym}\big(y_{m,it} | S_{it-1}, d_{it-1}; \theta_k^{ym}\big)\Bigg). \end{aligned} \tag{40}$$

Because of the additive separability, I can maximize the objective function sequentially. Specifically, the updated parameter vectors are given by

$$\tilde{\pi}(M) = \arg\max_{\pi} \frac{1}{N} \sum_{i=1}^{N}\sum_{k=1}^{K} q_{ik}(M) \ln p_k(z_{i\tau(i)}; \pi), \tag{41}$$

$$\tilde{\theta}^y(M) = \arg\max_{\theta^y} \frac{1}{N} \sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{t=\tau(i)+1}^{T_i} q_{ik}(M) \ln l_k^y\big(y_{it} | S_{it}, d_{it}; \theta^y\big), \tag{42}$$

$$\tilde{\theta}^{ym}(M) = \arg\max_{\theta^{ym}} \frac{1}{N} \sum_{i=1}^{N}\sum_{k=1}^{K}\sum_{t=\tau(i)+1}^{T_i} q_{ik}(M) \ln l_k^{ym}\big(y_{m,it} | S_{it-1}, d_{it-1}; \theta^{ym}\big), \tag{43}$$

$$\tilde{\theta}^d(M) = \arg\max_{\theta^d} \frac{1}{N} \sum_{i=1}^{N} q_{ik}(M) \ln \Lambda\big(\theta^d, \tilde{\theta}^y(M), \tilde{\theta}^{ym}(M), \Gamma^q\big(\theta, \widetilde{EV}(M-1)\big)\big). \tag{44}$$

In updating θ^d, the Bellman operator Γ is approximated by a higher order polynomial function as outlined above for the case of homogeneous individuals. Finally, the estimate of the expectation of the value function is updated by

$$\widetilde{EV}(M) = \Gamma^q\big(\tilde{\theta}(M), \widetilde{EV}(M-1)\big). \tag{45}$$

Appendix C: Additional tables

See Tables 20–23.

TABLE 20. Earnings function for husband.

	Estimate	S.E.
Intercept (Type 1)	−0.425	0.350
Intercept (Type 2)	−0.102	0.349
Intercept (Type 3)	−0.394	0.347
Intercept (Type 4)	−0.360	0.348
Husband's Earnings	0.832	0.003
Age	0.065	0.019
Age-sq	−0.069	0.025
Sqrt. of No. Children	−0.008	0.023
Age of Youngest Child	−0.001	0.003
Reg.	−0.152	0.034
Non-Reg.	−0.103	0.031
PL	−0.098	0.099
Conception	0.055	0.039
Unempl. Rate	−0.009	0.017
Std. Dev. of Error Term	1.083	0.002

Note: The dependent variable is the level of earnings so that zero earnings can be included.

TABLE 21. Type probability function and share.

	Type 2		Type 3		Type 4	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Intercept	−1.431	3.501	3.110	2.506	6.059	3.114
Some College	1.455	0.383	0.125	0.286	0.579	0.333
4-Yr College	3.014	0.999	0.329	0.984	1.621	0.993
Age	−0.071	0.131	−0.128	0.087	−0.206	0.115
Years in Home	0.627	0.472	−0.168	0.285	−0.261	0.325
Years in Reg.	0.278	0.350	0.169	0.285	−0.092	0.322
Years in Non-Reg.	−0.164	0.440	−0.271	0.289	−0.423	0.408
Age × Years in Home	−2.066	1.433	0.282	0.814	0.969	0.938
Age × Years in Reg.	−0.579	1.049	−0.169	0.853	0.407	0.981
Age × Years in Non-Reg.	0.592	1.300	1.011	0.818	1.064	1.232
Husband's Earnings	0.433	0.098	0.017	0.085	−0.013	0.101
Sqrt. of No. Children	0.319	0.325	0.621	0.264	0.255	0.304
Age of Youngest Child	−0.011	0.075	0.045	0.049	−0.007	0.064
Reg. in 1st Year	1.993	0.566	1.356	0.498	1.702	0.530
Non-Reg. in 1st Year	0.988	0.502	1.655	0.392	1.243	0.452
Conceived in 1st Year	−0.370	0.395	−0.591	0.350	−1.181	0.446

	Type 1	Type 2	Type 3	Type 4
Share	0.131	0.247	0.331	0.291

Table 22. Effects of parental leave policies on labor market outcomes.

	Years since birth						
	−1	0	1	2	3	5	10
Work							
JP:0, RR:0%	1.00	0.19	0.33	0.34	0.39	0.46	0.58
JP:1, RR:0%	1.00	0.13	0.54	0.50	0.53	0.57	0.66
JP:3, RR:0%	1.00	0.14	0.54	0.48	0.55	0.59	0.67
JP:0, RR:50%	1.00	0.17	0.32	0.34	0.38	0.45	0.58
JP:1, RR:50%	1.00	0.11	0.55	0.50	0.53	0.58	0.67
JP:3, RR:50%	1.00	0.12	0.55	0.49	0.55	0.59	0.68
JP:0, RR:50% + Need to Take PL	1.00	0.19	0.35	0.36	0.40	0.47	0.59
JP:1, RR:50% + Need to Take PL	1.00	0.12	0.57	0.52	0.55	0.59	0.68
JP:3, RR:50% + Need to Take PL	1.00	0.12	0.56	0.50	0.56	0.60	0.69
On PL							
JP:0, RR:0%	0.00	0.12	0.02	0.01	0.01	0.00	0.00
JP:1, RR:0%	0.00	0.50	0.06	0.07	0.05	0.02	0.00
JP:3, RR:0%	0.00	0.54	0.12	0.15	0.09	0.05	0.01
JP:0, RR:50%	0.00	0.13	0.02	0.01	0.01	0.00	0.00
JP:1, RR:50%	0.00	0.52	0.06	0.08	0.06	0.03	0.00
JP:3, RR:50%	0.00	0.56	0.12	0.16	0.10	0.06	0.01
JP:0, RR:50% + Need to Take PL	0.00	0.15	0.02	0.02	0.01	0.00	0.00
JP:1, RR:50% + Need to Take PL	0.00	0.54	0.06	0.09	0.06	0.03	0.00
JP:3, RR:50% + Need to Take PL	0.00	0.57	0.13	0.17	0.10	0.06	0.01
Reg. Work							
JP:0, RR:0%	0.59	0.12	0.19	0.18	0.17	0.17	0.19
JP:1, RR:0%	0.59	0.07	0.34	0.31	0.30	0.29	0.29
JP:3, RR:0%	0.59	0.08	0.33	0.29	0.31	0.30	0.30
JP:0, RR:50%	0.59	0.10	0.18	0.17	0.16	0.16	0.19
JP:1, RR:50%	0.59	0.06	0.36	0.31	0.31	0.30	0.30
JP:3, RR:50%	0.59	0.06	0.34	0.29	0.31	0.31	0.32
JP:0, RR:50% + Need to Take PL	0.59	0.12	0.21	0.19	0.19	0.18	0.20
JP:1, RR:50% + Need to Take PL	0.59	0.07	0.37	0.33	0.32	0.31	0.32
JP:3, RR:50% + Need to Take PL	0.59	0.07	0.35	0.30	0.32	0.32	0.33
Non-Reg. Work							
JP:0, RR:0%	0.41	0.07	0.14	0.17	0.22	0.29	0.39
JP:1, RR:0%	0.41	0.05	0.20	0.19	0.23	0.28	0.36
JP:3, RR:0%	0.41	0.06	0.21	0.20	0.24	0.29	0.36
JP:0, RR:50%	0.41	0.07	0.14	0.17	0.22	0.29	0.39
JP:1, RR:50%	0.41	0.05	0.20	0.19	0.23	0.28	0.36
JP:3, RR:50%	0.41	0.06	0.21	0.20	0.24	0.28	0.36
JP:0, RR:50% + Need to Take PL	0.41	0.07	0.14	0.17	0.22	0.29	0.39
JP:1, RR:50% + Need to Take PL	0.41	0.05	0.20	0.19	0.22	0.28	0.36
JP:3, RR:50% + Need to Take PL	0.41	0.06	0.21	0.20	0.24	0.28	0.36

(*Continues*)

TABLE 22. *Continued.*

	Years Since Birth						
	−1	0	1	2	3	5	10
Earnings (mil. JPY)							
JP:0, RR:0%	2.62	0.54	0.71	0.80	0.81	0.87	1.11
JP:1, RR:0%	2.62	0.30	1.05	1.32	1.30	1.38	1.59
JP:3, RR:0%	2.62	0.32	1.01	1.19	1.26	1.38	1.63
JP:0, RR:50%	2.62	0.46	0.66	0.77	0.78	0.84	1.09
JP:1, RR:50%	2.62	0.26	1.06	1.34	1.31	1.41	1.64
JP:3, RR:50%	2.62	0.27	1.02	1.21	1.28	1.42	1.70
JP:0, RR:50% + Need to Take PL	2.62	0.51	0.76	0.87	0.88	0.94	1.16
JP:1, RR:50% + Need to Take PL	2.62	0.28	1.11	1.40	1.37	1.48	1.70
JP:3, RR:50% + Need to Take PL	2.62	0.29	1.05	1.25	1.32	1.45	1.74

TABLE 23. Effects of parental leave policies on fertility.

	Years Since Birth						
	−1	0	1	2	3	5	10
Pregnancy							
JP:0, RR:0%	1.00	0.04	0.13	0.10	0.06	0.03	0.00
JP:1, RR:0%	1.00	0.04	0.15	0.12	0.07	0.04	0.00
JP:3, RR:0%	1.00	0.04	0.16	0.13	0.08	0.04	0.00
JP:0, RR:50%	1.00	0.04	0.13	0.10	0.06	0.03	0.00
JP:1, RR:50%	1.00	0.04	0.16	0.12	0.07	0.04	0.00
JP:3, RR:50%	1.00	0.04	0.16	0.14	0.08	0.05	0.00
JP:0, RR:50% + Need to Take PL	1.00	0.04	0.13	0.10	0.06	0.03	0.00
JP:1, RR:50% + Need to Take PL	1.00	0.04	0.16	0.12	0.07	0.04	0.00
JP:3, RR:50% + Need to Take PL	1.00	0.04	0.16	0.14	0.08	0.05	0.00
Number of Children							
JP:0, RR:0%	0.59	1.64	1.69	1.82	1.92	2.02	2.10
JP:1, RR:0%	0.59	1.64	1.68	1.83	1.95	2.07	2.15
JP:3, RR:0%	0.59	1.64	1.68	1.84	1.97	2.10	2.20
JP:0, RR:50%	0.59	1.64	1.69	1.82	1.92	2.03	2.10
JP:1, RR:50%	0.59	1.64	1.68	1.84	1.96	2.09	2.18
JP:3, RR:50%	0.59	1.64	1.68	1.84	1.98	2.12	2.23
JP:0, RR:50% + Need to Take PL	0.59	1.64	1.69	1.81	1.91	2.02	2.09
JP:1, RR:50% + Need to Take PL	0.59	1.64	1.68	1.84	1.96	2.09	2.18
JP:3, RR:50% + Need to Take PL	0.59	1.64	1.68	1.84	1.98	2.12	2.23

REFERENCES

Arcidiacono, P., P. Bayer, F. A. Bugni, and J. James (2013), "Approximating high-dimensional dynamic models: Sieve value function iteration." *Structural Econometric Models (Advances in Econometrics)*, 31, 45–95. [5, 7]

Arcidiacono, P. and J. B. Jones (2003), "Finite mixture distributions, sequential likelihood and the EM algorithm." *Econometrica*, 71 (3), 933–946. [5, 7]

Kasahara, H. and K. Shimotsu (2011), "Sequential estimation of dynamic programming models with unobserved heterogeneity." University of British Columbia. [5, 6, 7]

Rust, J. (1987), "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher." *Econometrica*, 55 (5), 999–1033. [6]

Wen, C.-H. and F. S. Koppelman (2001), "The generalized nested logit model." *Transportation Research Part B: Methodological*, 35 (7), 627–641. [4]

Co-editor Peter Arcidiacono handled this manuscript.

Manuscript received 4 September, 2017; final version accepted 2 October, 2018; available online 7 November, 2018.