Dynamic Discrete Choice Structural Models: A Survey, JoE, 2010

Wenzhi Wang *

March 20, 2023

1 Models and Examples

1.1 Single-Agent Models

Time is discrete and indexed by t. We index agents by i. Agents have preferences defined over a sequence of states of the world from period $t = 0$ until period $t = T$. The time horizon T can be either finite or infinite. The state of the world at period t for individual i has two components: a vector of state variables s_{it} that is known at period t; and a decision a_{it} chosen at period t that belongs to the discrete set $A = \{0, 1, ..., J\}$. The time index t can be a component of the state vector s_{it}, which may also contain time-invariant individual characteristics. Agents' preferences over possible sequences of states of the world can be represented by a utility function $\sum_{j=0}^{T} \beta^j U(a_{i,t+j}, s_{i,t+j})$, where $\beta \in (0,1)$ is the discount factor and $U(a_{i,t}, s_{i,t})$ is the current utility function. The decision at period t affects the evolution of future values of the state variables, but the agent faces uncertainty about these future values. The agent's beliefs about future states can be represented by a Markov transition distribution function $F(s_{i,t+1} \mid a_{it}, s_{it})$. These beliefs are rational in the sense that they are the true transition probabilities of the state variables. If the only data available are (longitudinal) data on choices and states, then preferences, beliefs and the actual transition probabilities cannot be separately identified in general. In this sense, rational expectations is an identification assumption. Every period t the agent observes the vector of state variables s_{it} and chooses his action $a_{it} \in A$ to maximize the expected utility

$$E\left(\sum_{j=0}^{T-t} \beta^j U\left(a_{i,t+j}, s_{i,t+j}\right) \mid a_{it}, s_{it}\right). \tag{1}$$

This is the agent's dynamic programming (DP) problem. Let $\alpha(s_{it})$ and $V(s_{it})$ be the optimal decision rule and the value function of this DP problem, respectively. By Bellman's principle of optimality the value function can be obtained using the recursive expression:

$$V(s_{it}) = \max_{a \in A}\left\{U(a, s_{it}) + \beta \int V(s_{i,t+1}) dF(s_{i,t+1} \mid a, s_{it})\right\}. \tag{2}$$

*This note is written down during my M.phil. period at the University of Oxford.

and the optimal decision rule is then $\alpha(s_{it}) = \arg\max_{a\in A}\{v(a, s_{it})\}$ where, for every $a \in A$,

$$v(a, s_{it}) \equiv U(a, s_{it}) + \beta \int V(s_{i,t+1}) dF(s_{i,t+1} \mid a, s_{it}) \tag{3}$$

is a choice-specific value function.

We are interested in the estimation of the structural parameters in preferences, transition probabilities, and the discount factor β [1]. Suppose that a research has panel data for N individuals who behave according to this decision model. For every observation (i, t) in this panel dataset, the researcher observes the individual's action a_{it} and a sub-vector of the state variable s_{it}. Therefore, we can distinguish two subsets of state variables: $s_{it} = (x_{it}, \varepsilon_{it})$, where the sub-vector ε_{it} is observed by the agent but not by the researcher. Note that ε_{it} is a source of variation in the decisions of agents conditional on the variables observed by the researcher. In some applications, the researcher also observes one or more payoff variables. We define a payoff variable as a variable y_{it} which contains information about utility but is not one of the model's actions or state variables. That is, we can write $U(a_{it}, s_{it})$ as $\tilde{U}(y_{it}, a_{it}, s_{it})$. For instance, in a model of firm behaviour the researcher may observe firms' output, revenue or the wage bill; or in a model of individual behaviour the econometrician may observe individual earnings. Payoff variables depend on current action and state variables. We specify this relationship as $y_{it} = Y(a_{it}, x_{it}, \varepsilon_{it})$, where $Y(\cdot)$ is the payoff function, e.g., an earnings function, a production function, etc. In summary, the researcher's dataset is :

$$\text{Data} = \left\{a_{it}, x_{it}, y_{it} : i = 1, ..., N; t = 1, ..., T_i\right\} \tag{4}$$

where T_i is the number of periods over which we observe individual i.

Let θ be the vector of structural parameters and let $g_N(\theta)$ be an estimation criterion for this model and data, such as a likelihood or a GMM criterion. For instance, if the data are a random sample over individuals and the criterion is a log-likelihood, then $g_N(\theta) = \sum_{i=1}^{N} \ell_i(\theta)$, where $\ell_i(\theta)$ is the contribution to the log-likelihood function of individual i's history:

$$\begin{aligned} \ell_i(\theta) &= \log \Pr\left\{a_{it}, y_{it}, x_{it} : t = 1, 2, \ldots, T_i \mid \theta\right\} \\ &= \log \Pr\left\{\alpha\left(x_{it}, \varepsilon_{it}, \theta\right) = a_{it}, Y\left(a_{it}, x_{it}, \varepsilon_{it}, \theta\right) = y_{it}, x_{it} : t = 1, 2, \ldots, T_i \mid \theta\right\} \end{aligned} \tag{5}$$

Whatever the estimation criterion, in order to evaluate it for a particular value of it is necessary to know the optimal decision rules $\alpha\left(x_{it}, \varepsilon_{it}, \theta\right)$. Therefore, for each trial value of θ the DP problem needs to be solved exactly, or its solution approximated in some way.

Assumption 1. AS (Additive Separability). The one-period utility function is additively separable in the observable and unobservable components: $U(a, x_{it}, \varepsilon_{it}) = u(a, x_{it}) + \varepsilon_{it}(a)$, where $\varepsilon_{it}(a)$ is a zero mean real random variable with unbounded support. That is, there is one unobservable state variable for each choice alternative, so the dimension of ε_{it} is $(J + 1) \times 1$.

Assumption 2. IID (i.i.d. Unobservables). The unobserved state variables in ε_{it} are independently and identically distributed over agents and over time with CDF $G_\varepsilon(\varepsilon_{it})$ich has finite first moments and is continuous and twice differentiable in ε_{it}.

[1] The discount factor is assumed constant across agents. In most applications this parameter is not estimated because it is poorly identified.

Assumption 3. CI-X (Conditional Independence of Future x). Conditional on the current values of the decision and the observable state variables, next period observable state variables do not depend on current ε: i.e., $\text{CDF}(x_{i,t+1} \mid a_{it}, x_{it}, \varepsilon_{it}) = F_x(x_{i,t+1} \mid a_{it}, x_{it})$. We use θ_f to represent the vector of parameters that describe the transition probability function F_x.

Assumption 4. CI-Y (Conditional Independence of y). Conditional on the values of the decision and the observable state variables, the value of the payoff variable y is independent of ε: i.e., $Y(a_{it}, x_{it}, \varepsilon_{it}) = Y(a_{it}, x_{it})$. The vector of parameters that describe Y is θ_Y.

Assumption 5. CLOGIT. The unobserved state variables $\{\varepsilon_{it}(a) : a = 0, 1, ..., J\}$ are independent across alternatives and have an extreme value type 1 distribution.

Assumption 6. DIS (Discrete Support of x). The support of x_{it} is discrete and finite: $x_{it} \in X = \{x^{(1)}, x^{(2)}, ..., x^{(|X|)}\}$ with $|X| < \infty$.

Assumptions 2 IID and 3 CIX imply that $F(x_{i,t+1}, \varepsilon_{i,t+1} \mid a_{it}, x_{it}, \varepsilon_{it}) = G_\varepsilon(\varepsilon_{i,t+1}) F_x(x_{i,t+1} \mid a_{it}, x_{it})$, which corresponds to the conditional independence assumption in Rust (1987). It is convenient to distinguish several components in the vector of structural parameters: $\theta = \{\theta_u, \theta_Y, \theta_f\}$, where θ_Y and θ_f have been defined above and θ_u represents all the parameters in the utility function which are not in θ_Y as well as the parameters in the distribution of G_ε.

Example 7. (Retirement from the Labour Force).
Reference:
Rust, J., Phelan, C., 1997. How social security and medicare affect retirement behavior in a world of incomplete markets. Econometrica 65, 781-832.
Karlstrom, A., Palme, M., Svensson, I., 2004. A dynamic programming approach to model the retirement behaviour of blue-collar workers in Sweden. Journal of Applied Econometrics 19, 795-807.

Every period the individual decides whether to continue working ($a_{it} = 1$) or to retire and start collecing social security pension benefits ($a_{it} = 0$). This is an optimal stopping problem with a finite horizon, where $t = 1$ is the earliest age at which the individual can retire and T is age at death. Letra_{it} denote the individual's retirement status. If he has not retired yet, $ra_{it} = 0$. If retired, ra_{it} is the age at which he retired. The utility functiion is additively separable in consumption and leisure. More specifically,

$$U(a_{it}, x_{it}, \varepsilon_{it}) = E\left(c_{it}^{\theta_{u1}} \mid a_{it}, x_{it}\right) \times \exp\left\{\theta_{u2} + \theta_{u3} h_{it} + \theta_{u4} m_{it} + \theta_{u5} \frac{t_{it}}{1 + t_{it}}\right\} - \theta_{ut} a_{it} + \varepsilon_{it}(a_{it}) \quad (6)$$

c_{it} is current consumption. θ_{u1} is the coefficient of relative risk aversion. The expression in the exponential term captures individual heterogeneity in the marginal utility of consumption. In particular, h_{it} is an indicator of good health status, m_{it} is an indicator of marital status, and t_{it} is age. Finally, the last term is associated with the utility of leisure. θ_{u6} is the disutility of working. In our notation, $\theta_u = (\theta_{u1}, \ldots, \theta_{u6})$. The unobservable state variables $\varepsilon_{it}(1)$ and $\varepsilon_{it}(0)$ enter additively in the utilities of working and not working, respectively, and they can be interpreted as transitory and idiosyncratic shocks to the utility from leisure. These random variables are independently distributed over time and over individuals with an extreme value distribution.

Consumption is equal to current income (y_{it}) minus health care expenditures net of insurance reimbursements (hc_{it}): $c_{it} = y_{it} - hc_{it}$. If the individual works, his income is equal to stochastic labor earnings w_{it}, hence the expectation $E\left(c_{it}^{\theta_{u1}} \mid a_{it}, x_{it}\right)$. If the individual decides to retire, then his earnings are equal to social security pension benefits b_{it}. The econometrician observes earnings y_{it} and $y_{it} = a_{it}w_{it} + (1 - a_{it})b_{it}$. Labor earnings depend on age, health status, arital status, past earnings history through pension points and an unobservable shock ξ_{it}. In particular,

$$w_{it} = \exp\left\{\theta_{w1} + \theta_{w2}h_{it} + \theta_{w3}m_{it} + \theta_{w4}\frac{t_{it}}{1+t_{it}} + \theta_{w5}pp_{it} + \xi_{it}\right\}. \tag{7}$$

Earnings y_{it} are a payoff variable and $\theta_Y = (\theta_{w1}, ..., \theta_{w5}, \sigma_\xi^2)$ is the vector of parameters in the corresponding payoff function, where σ_ξ^2 is the variance of ξ_{it}. Retirement benefits depend on retirement age (ra_{it}) and on pension points or social security wealth $pp_{it} : b_{it} = b(ra_{it}, pp_{it})$. The form of the function $b(\cdot,\cdot)$ depends on the rules of the pension system, which are known to the econometrician. Health care expenditures are stochastic, with a Pareto distribution conditional on health and marital status.

The vector of observable state variables is $x_{it} = \{h_{it}, hc_{it}, m_{it}, t_{it}, ra_{it}, pp_{it}\}$. Age and retirement age have obvious deterministic transition rules. Pension points are a deterministic function of past earnings history. However, for many pension systems, the transition rule of pension points can be very closely approximated by a Markov process with transition probability function $F_{pp}(pp_{i,t+1} \mid w_{it}, pp_{it})$. This transition probability is nonparametrically specified. Health status and marital status follow first order Markov processes with transition probabilities $F_h(h_{i,t+1} \mid h_{it})$ and $F_m(m_{i,t+1} \mid m_{it})$ which are nonparametrically specified. Though health expenditures and pension points are continuous variables, we can discretize these variables. An important feature of this model is that the shocck to wages ξ_{it} is assumed serially uncorrelated, independent of the state variables x_{it} and ε_{it}, and unknown to the individual at the time he makes his period t decision. Therefore, ξ_{it} determines the transition of labor earnings but it is not a state variable and Assumption 4 CIY holds, i.e., conditional of (a_{it}, x_{it}) observed earnings are independent of the unobserved state variables in ε_{it}.

Marginal additive separability implies that the marginal utility of observable state variables, today and in future periods, does not depend on unobservables. For instance, in Example 7 the decision to retire or to continue working determines the current and future level of consumption. However, additive separability in the econometric model implies that observed variation in retirement age cannot be linked to unobserved heterogeneity in the marginal utility of consumption. Relatedly, an individual considering early retirement in this model weighs uncertainty about the value of her marginal utility of consumption in the future. Additive separability implies that unobserved state variables induce no uncertainty about the marginal utility of consumption, which may have an effect on the patterns of behaviour the model can explain and on estimation of structural parameters such as the coefficient of relative risk aversion.

Assumptions 2 IID and 3 CIX restrict the joint transition probability of state variables. These restrictions have two main implications which can be illustrated in Example 7. First, the unobserved shocks to the utility of leisure are serially uncorrelated, i.e., transitory. And second, the probability that a person's health or marital status will change between the current period and the next one could depend on whether this person is currently working, but once we take this into account it does not depend on the current value of the shocks to the utility of leisure. Suppose we interpret these unobservable shocks as health shocks not included in the observable h_{it}. The shocks are important enough because they can explain individuals' changes in labour supply decisions, but they should be "transitory" since according to Assumption 3 CIX they cannot have any direct impact on next period's health.

An important implication of Assumptions 2 IID and 3 CIX is that the solution to the DP problem is fully characterized by the integrated value function or $Emax$ function, $\overline{V}(x_{it})$, which is the expectation of the value function over the distribution of unobservable state variables, conditional on the observable state variables: $\overline{V}(x_{it}) \equiv \int V(x_{it}, \varepsilon_{it}) dG_{\varepsilon}(\varepsilon_{it})$. This function is the unique solution to the integrated Bellman equation:

$$\overline{V}(x_{it}) = \int \max_{a \in A} \left\{ u(a, x_{it}) + \varepsilon_{it}(a) + \beta \sum_{x_{i,t+1}} \overline{V}(x_{i,t+1}) f_x(x_{i,t+1} \mid a, x_{it}) \right\} dG_{\varepsilon}(\varepsilon_{it}) \tag{8}$$

Under these assumptions, the size of the state space X is the relevant measure of computational complexity, and given that X is discrete and finite the DP problem can be solved exactly. The choice-specific value function in (3) can be decomposed as $v(a, x_{i,t}) + \varepsilon_{it}(a)$ as in static random utility models, where

$$v(a, x_{it}) = u(a, x_{it}) + \beta \sum_{x_{i,t+1}} \overline{V}\left(x_{i,t+1}\right) f_x\left(x_{i,t+1} \mid a, x_{it}\right). \tag{9}$$

Another important implication of Assumptions 2 IID and 3 CIX is that the observable state vector x_{it} is a sufficient statistic for the current choice. The contribution of individual i to the log-likelihood function can be factored as

$$\begin{aligned} \ell_i(\theta) = &\sum_{t=1}^{T_i} \log P(a_{it} \mid x_{it}, \theta) + \sum_{t=1}^{T_i} \log f_Y\left(y_{it} \mid a_{it}, x_{it}, \theta_Y\right) \\ &+ \sum_{t=1}^{T_i - 1} \log f_x\left(x_{i,t+1} \mid a_{it}, x_{it}, \theta_f\right) + \log \Pr(x_{i1} \mid \theta) \end{aligned} \tag{10}$$

f_Y is the density of the payoff variable conditional on (a_{it}, x_{it}). f_x is the transition density function associated with F_Y. The term $\log \Pr(x_{i1} \mid \theta)$ is the contribution of initial conditions to the likelihood of individual i. In most applications a conditional likelihood approach is followed and this term is ignored. No bias is incurred as long as unobservables are serially independent. Accordingly, hereafter the term "likelihood" by default refers to the conditional likelihood. The term $P(a_{it} \mid x_{it}, \theta)$ is the Conditional Choice Probability (CCP) obtained by integrating the optimal decision rule over the unobservable state variables. The optimal decision rule is $\alpha(x_{it}, \varepsilon_{it}) = \arg\max_{a \in A}\{v(a, x_{it}) + \varepsilon_{it}(a)\}$. Therefore, for any $(a, x) \in A \times X$ and $\theta \in \Theta$, the conditional choice probability is

$$\begin{aligned} P(a \mid x, \theta) &\equiv \int I\{\alpha(x, \varepsilon; \theta) = a\} dG_{\varepsilon}(\varepsilon) \\ &= \int I\left\{v(a, x_{it}) + \varepsilon_{it}(a) > v\left(a', x_{it}\right) + \varepsilon_{it}\left(a'\right) \text{ for all } a'\right\} \mathrm{d}G_{\varepsilon}(\varepsilon_{it}) \end{aligned} \tag{11}$$

When $\varepsilon_{it}(a)$ are i.i.d. type 1 extreme value random variables, as in Example 7, the multi-dimension integrals in the integrated Bellman equation and in the definition of CCPs have closed form analytical expressions. This is the DP conditional logit model with Bellman equaiton:

$$\overline{V}(x_{it}) = \log\left(\sum_{a=0}^{J} \exp\left\{u(a, x_{it}) + \beta \sum_{x_{i,t+1}} \overline{V}\left(x_{i,t+1}\right) f_x\left(x_{i,t+1} \mid a, x_{it}\right)\right\}\right) \tag{12}$$

and choice probabilities:

$$P(a \mid x_{it}, \theta) = \frac{\exp\{v(a, x_{it})\}}{\sum_{j=0}^{J} \exp\{v(j, x_{it})\}}. \tag{13}$$

If $v(a, x_{it})$ were a linear function of the parameters θ, these expressions would be familiar as the choice probability of binary probit, logit or multinomial logit models. In general, $v(a, x_{it})$ is a complex nonlinear function of which has to be computed from the Bellman equation in Equation 12.

We then highlight four departures from the previous model: (1) unobservables which do not satisfy Assumption 1 AS; (2) observable payoff variables which are choice censored and do not satisfy Assumption 4 CI-Y; (3) permanent unobserved heterogeneity (a departure of Assumption 2 IID); and (4) unobservables which are correlated across choice alternatives (implying no Assumption 5 CLOGIT). Many applications have included at least one of these four features, and in this survey we group all of them under the label Eckstein-Keane-Wolpin (EKW) after the authors who are the main contributors.

Example 8. (Occupational Choice and the Career Decisions of Young Men).
Reference:
Keane, M., Wolpin, K., 1997. The career decisions of young men. Journal of Political Economy 105, 473-522.

Each period (year), starting at age 16 through a maximum age T, an individual chooses between staying at home ($a_{it} = 0$), attending school ($a_{it} = 4$), or working at one of three occupations: white collar ($a_{it} = 1$), blue collar ($a_{it} = 2$), or the military ($a_{it} = 3$). The specification of the one-period utility function is:

$$\begin{aligned} U(0, s_{it}) &= \omega_i(0) + \varepsilon_{it}(0) \\ U(4, s_{it}) &= \omega_i(4) - \theta_{tc1} I(h_{it} \geq 12) - \theta_{tc2} I(h_{it} \geq 16) + \varepsilon_{it}(4) \\ U(a, s_{it}) &= W_{it}(a) \quad \text{for } a = 1,2,3 \end{aligned} \tag{14}$$

h_{it} is schooling (in years). θ_{tc1} and θ_{tc2} are parameters that represent tuition costs in college and graduate school, respectively. $W_{it}(a)$ is the wage of individual i at period t in occupation a. Wages in occupation a are the product of the skill price in that occupation, r_a, and the individual's skill level for that occupation, which is an exponential function of an individual-specific endowment at age 16, schooling, experience and a transitory shock. That is:

$$W_{it}(a) = r_a \exp\Big\{\omega_i(a) + \theta_{a1} h_{it} + \theta_{a2} k_{it}(a) - \theta_{a3} (k_{it}(a))^2 + \varepsilon_{it}(a)\Big\} \quad \text{for } a = 1,2,3 \tag{15}$$

where $k_{it}(a)$ is cumulated work experience (in years) in occupation a. The vector $\varepsilon_{it} = \{\varepsilon_{it}(a) : a \in A\}$ contains choice-specific shocks to skill levels and to the monetary values of a year at school or at home. They are assumed to be serially uncorrelated with a joint normal distribution with zero means and unrestricted variance matrix. The vector $\omega_i = \{\omega_i(a) : a \in A\}$ contains occupation and individual-specific endowments which are fixed from age 16. This vector has a discrete support, and its probability distribution is nonparametrically specified. Both ε_{it} and ω_i are unobservable to the econometrician but observable to the individual when he makes his decision at period t. The vector of observable state variables is $x_{it} = \{h_{it}, t_{it}, k_{it}(a) : a = 1,2,3\}$, where t_{it} represents age. All the variables in x_{it} have a discrete and finite support and their transitions conditional on choices are deterministic so $f_x(x_{i,t+1} \mid a, x_{it})$ is degenerate. Labour earnings are observable. This payoff variable is equal to zero when $a_{it} \in \{0,4\}$ and equal to $W_{it}(a_{it})$ when $a_{it} \in \{1,2,3\}$. Therefore, the payoff function is

$$I\left\{a_{it} \in \{1,2,3\}\right\} r_a \exp\left\{\omega_i(a) + \theta_{a1} h_{it} + \theta_{a2} k_{it}(a) - \theta_{a3} \left(k_{it}(a)\right)^2 + \varepsilon_{it}(a)\right\}.$$

Assumption 4 CIY does not hold because the unobserved state variables $\left\{\varepsilon_{it}(1), \varepsilon_{it}(2), \varepsilon_{it}(3)\right\}$ have a direct effect on observed labour earnings.

The opportunity cost of investing in human capital by attending school is the value of foregone earnings and work experience, or the utility of staying home. Working also has an investment value since it increases occupation-specific skills and future earnings. An individual's optimal career path is partly determined by comparative advantage embedded in endowments at age 16.

We now use Example 8 to illustrate the practical implications for estimation of relaxing some of Rust's assumptions. The sum of the permanent and transitory unobserved skill components, $\varepsilon_{it}(a) + \omega_i(a)$, is a serially correlated state variable. The presence of autocorrelated unobservables has important implications for the estimation of structural parameters. Individuals self-select into occupations and schooling classes on the basis of persistent differences in skills which are unobserved by the researcher. Ignoring this source of self-selection can result into an overestimation of the returns to schooling and occupation-specific experience. Also, persistence in occupational choices in the data may arise because there is a disutility of switching occupations (state dependence), or because there are persistent differences in skills (unobserved heterogeneity). Failure to control for the latter, if present, would lead to biased estimates of switching disutilities. With serially correlated unobservables, the probability of an individual's sequence of choices cannot be factored into a product of conditional choice probabilities as in Equation 10. The observable state variables x_{it} is not a sufficient statistic for a_{it} because lagged choices contain information about the permanent components ω_i. However, conditional on ω_i the transitory components $\{\varepsilon_i(a)\}$ do satisfy Assumption 2 IID. Since ω_i has discrete support, each individual's likelihood contribution can be obtained as a finite mixture of likelihoods, each of which has the same form as in Equation 10. If the support of ω_i is $\Omega = \{\omega^1, \ldots, \omega^L\} \in \mathbb{R}^{JL}$, then we have that

$$\ell_i(\theta, \Omega, \pi) = \log\left(\sum_{l=1}^{L} \mathscr{L}_i(\theta, \omega^l)\pi_{l|x_{i1}}\right) \tag{16}$$

where $\pi_{l|x} \equiv \Pr(\omega_i = \omega^l \mid x_{i1} = x)$; π is the vector of parameters $\{\pi_{l|x} : l = 1, \ldots, L; x \in X\}$; and [2]

$$\begin{aligned} \mathscr{L}_i\left(\theta, \omega^l\right) \equiv & \prod_{t=1}^{T_i} P\left(a_{it} \mid x_{it}, \theta, \omega^l\right) f_Y\left(y_{it} \mid a_{it}, x_{it}, \theta, \omega^l\right) \\ & \times \prod_{t=1}^{T_i - 1} f_x\left(x_{i,t+1} \mid a_{it}, x_{it}, \theta_f, \omega^l\right) \end{aligned} \tag{17}$$

Note that the conditional density of the payoff variable y_{it} depends not only on θ_Y, as in Rust's model, but on the whole vector of structural parameters θ. The reason for this is that the payoff

[2]In Example 2, the transition of the state variables are deterministic. Recall the state variables are $x_{it} = \left\{h_{it}, t_{it}, k_{it}(a) : a = 1,2,3\right\}$. Therefore, in this example, θ_f is empty and the likelihood does not include $f_x\left(x_{i,t+1} \mid a_{it}, x_{it}, \theta_f, \omega^l\right)$.

functions (the earnings equations) do not satisfy Assumption 4 CIY and depend on unobservable state variables $\varepsilon_{it}(a)$. Conditioning on the current action a_{it} introduces selection or censoring in the payoff variable and this selection effect depends on all the parameters of the model.

Several issues are worth highlighting here. First, in general permanent unobserved heterogeneity poses an initial conditions problem because the initial state x_{i1} is correlated with permanent unobserved components. This problem is avoided if the structural model has an initial period in which all the individuals have the same value of the vector x, and if this initial period is observed in the sample. Under these conditions $x_{i1} = x_1$ for every i and then $\Pr(\omega^l \mid x_{i1})$ is constant over individuals and equal to π_i, i.e., the unconditional mass probability of type l individuals which is a primitive structural parameter. Their model is an example of this in that all individuals start life in the model at age 16 with zero experience in all occupations and the NLSY data provide histories as of age 16. If x can take different values in the behavioural model's first decision period, then the researcher needs to specify how the distribution of permanent unobserved heterogeneity varies with x_{i1}. In their data there is variation in schooling measured at age 16. It seems likely that initial schooling is correlated with other age 16 endowments, and they allow for this as our mixture likelihood in Equation 16 suggests. A more difficult case arises if individual histories are left-censored. We will return to this issue later in Section **??**. The focus there is on a structure with (discrete) permanent unobserved heterogeneity and (continuous) i.i.d. transitory components, which is the simplest way of allowing for autocorrelation in unobservable state variables.

Second, in order to evaluate the mixture of likelihoods the DP problem needs to be solved as many times as the number of components in the mixture. This is the reason why permanent unobserved heterogeneity is almost always introduced with a discrete and finite support. Third, the integrated value function or $Emax$ function (conditional on individual's type) $\overline{V}_l(x_{it})$ still fully characterizes the solution to the DP problem. The integrated Bellman equation for individual type l is:

$$\overline{V}_l(x_{it}) = \int \max_{a \in A} \left\{ U(a, x_{it}, \omega^l, \varepsilon_{it}) + \beta \sum_{x_{i,t+1}} \overline{V}_l(x_{i,t+1}) f_x(x_{i,t+1} \mid a, x_{it}, \omega^l) \right\} dG_\varepsilon(\varepsilon_{it}). \tag{18}$$

The optimal decision rule is $\alpha(x_{it}, \omega^l, \varepsilon_{it}) = \arg\max_{a \in A} \{ v(a, x_{it}, \omega^l, \varepsilon_{it}) \}$, where $v(a, x_{it}, \omega^l, \varepsilon_{it})$ is the term in brackets in Equation 18. When ε's are additively separable and extreme value distributed, then we still have closed form expressions for the integrated Bellman equation and for CCPs in terms of choice specific value functions. In the occupational choice model of Example 8 the transitory shocks to skills are correlated across choices or occupations. This seems like a realistic (and even necessary) assumption to make in this application; e.g., if the transitory component of labour market skills reflects an unobservable health shock, the shock would likely have an effect on both white collar and blue collar skills. But correlation across choices implies that integrated value functions and conditional choice probabilities do not have the convenient closed forms of McFadden's conditional logit. In this case, as well as in models with more general autocorrelation structures, repeated calculation of multi-dimensional integrals in the solution and/or estimation of the model is unavoidable. Section **??** will review their simulation and interpolation method, developed for the occupational choice model. Simulation is used in order to compute multi-dimensional integrals, interpolation in order to handle problems with very large state spaces.

Fourth, in Example 8 transitory shocks to wages are observable to the decision maker and determine the optimal occupational choice. This feature of the model implies that, in contrast with

Example 1, there is a self-selection bias if we estimate the wage equation for an occupation by OLS using the sub-sample of individuals whose wages are observed because they chose that occupation. And fifth, in Example 8 the shocks $\varepsilon_{it}(1), \varepsilon_{it}(2), \varepsilon_{it}(3)$ do not satisfy Assumption 1 AS. Having one shock per alternative which is additively separable and has unbounded support is an expedient way of making sure that the optimal choice probabilities $P(a \mid x, \omega^l, \theta)$ are strictly positive for every value of (a, x, ω^l, θ). When this property fails the joint probability distribution over actions may be degenerate and the model is said to be saturated. This leads to well-known complications in estimation. However, this particular problem does not arise in Example 2 [3].

1.2 Competitive Equilibrium Models

The single-agent models of Examples 7 and 8 are partial equilibrium models. They study one side of the market (i.e., labour supply) taking prices and aggregate quantities as exogenously given. Though useful for policy evaluation, it is well known that partial equilibrium can give misleading results if the assumption that prices are invariant to changes in the policy variables of interest is not a good approximation. Equilibrium models are also better suited to improve our understanding of economy-wide trends. Despite the limitations of partial equilibrium, there have been very few studies that specify and estimate dynamic general equilibrium (GE) models with heterogenous agents using microdata.

References:
Heckman, J., Lochner, L., Taber, C., 1998. Explaining rising wage inequality: Explanations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. Review of Economic Dynamics 1, 1-58.
Lee, D., 2005. An estimable dynamic general equilibrium model of work, schooling and occupational choice. International Economic Review 46, 1-34.
Lee, D., Wolpin, K., 2006. Intersectoral labor mobility and the growth of the service sector. Econometrica 74, 1-46.

Example 9. (Occupational Choice in Equilibrium). In order to make it more suitable for the analysis of economy-wide aggregate data, consider the following changes in the model of Example 2: (a) Exclude the "military" occupational choice. (b) Allow the utility parameters to vary with gender and make the utility of the "home" alternative dependent on the number of children of pre-school age $n_{it}: U(0, s_{it}) = \omega_i(0) + \theta_c n_{it} + \varepsilon_{it}(0)$. The observable state variable n_{it} which is now part of x_{it} follows an exogenous Markov process conditional on age, gender, education and cohort. In Example 2, the skill rental prices r_a in the wage equations 15 were assumed constants. In this example, the state vector is augmented to include skill prices and, most important, the law of motion of skill prices is endogenously determined in the model as an equilibrium outcome. The labour market is

[3] In other words, Assumption 1 AS is sufficient but not necessary to avoid degeneracy. In order to illustrate how multiplicative errors can lead to degeneracy, even when their support is unbounded, consider a binary choice model with choice alternatives 0 and 1 and choice-specific utilities $U(0) = W(0)\exp(\varepsilon(0))$ and $U(1) = W(1)\exp(\varepsilon(01))$ where $W(0)$ and $W(1)$ are functions of observable covariates and parameters. Alternative 1 is chosen iff $\exp(\varepsilon(0)) < [W(1)/W(0)]\exp(\varepsilon(1))$. If $W(1)/W(0) < 0$ for some configuration of observable covariates and parameters, then the right hand side of the inequality is always negative and the left hand side is always positive for every value of the shocks, and alternative 1 is never chosen.

assumed to be competitive. The supply side consists of overlapping generations of individuals ages 16 through 65, whose behaviour corresponds to the occupational choice model. The demand side can be characterized by following Cobb-Douglas, constant returns to scale aggregate production function:

$$Y_t = z_t S_{1t}^{\alpha_{1t}} S_{2t}^{\alpha_{2t}} K_t^{1-\alpha_{1t}-\alpha_{2t}} \tag{19}$$

where Y_t is aggregate output, S_{1t} and S_{2t} are aggregate quantities of white collar and blue collar skills, and z_t and K_t are total factor productivity and the aggregate capital stock, which follow exogenously determined processes. Technical change is captured deterministically in time time-varying factor shares $\{\alpha_{1t}, \alpha_{2t}\}$ as well as total factor productivity. Demand side competitive behaviour implies that skill rental prices satisfy the value-of-marginal-product conditions:

$$r_{at} = (\frac{\alpha_{at}}{S_{at}}) \left(z_t S_{1t}^{\alpha_{1t}} S_{2t}^{\alpha_{2t}} K_t^{1-\alpha_{1t}-\alpha_{2t}} \right) \quad \text{for } a = 1, 2. \tag{20}$$

Let $\tilde{X}_t$ denote aggregate state variables relevant to the individual's occupational choice problem, i.e., current skill rental prices and other aggregate variables which agents use to predict future skill rental prices. The state vector of the occupational choice model is augmented with $\tilde{X}_t$ and individuals are assumed to know its law of motion. The aggregate supplies of skills are obtained by adding individual skill supplies over all individuals in the economy:

$$S_{at} = \int_{x,\omega,\varepsilon} k_{it}(a) I\left\{ a = \alpha(x_{it}, \omega_{it}, \varepsilon_{it}, \tilde{X}_t) \right\} \quad \text{for } a = 1, 2 \tag{21}$$

where $k_{it}(a)$ is individual i's stock of skill in occupation a, $\alpha()$ is the optimal decision rule, and the integral represents the appropriately weighted sum over the distribution of individual state variables.

In a rational expectation competitive equilibrium, the sequence of skill rental prices $\{r_{1t}, r_{2t} : t = 1, 2\}$ and the law of motion of $\tilde{X}_t$ satisfy the following conditions. First, individuals solve the occupational choice problem taking the law of motion of $\tilde{X}_t$ as given. Second, labour markets for both white and blue collar skills clear, i.e., the resulting aggregate skill supply functions, together with skill prices, satisfy the labour demand conditions 20. And, third, the actual law of motion of $\tilde{X}_t$ is consistent with individual's beliefs. The equilibrium need not be stationary: besides dependence on initial conditions $\tilde{X}_1$, there are other sources of non-stationarity such as time variation in factor shares in the aggregate production function, changes in cohort size, cohort effects in fertility, etc.

An important question the equilibrium model can address is the extent to which "feedback" effects offset the impact of policies such as the college tuition subsidy evaluated by Keane and Wolpin in partial equilibrium. That is, the subsidy may increase the supply of white collar skills to the extent that it induces higher college enrolment. If the relative price of white collar skills falls as a result, this reduces the returns to schooling and college enrolment.

The estimation is more demanding than that of its partial equilibrium version in Example 2 for two reasons. First, imposing the equilibrium restrictions increases the computational burden of estimation by an order of magnitude. Second, estimation of the equilibrium model requires additional data which can only be obtained from different sources and having to combine multiple data sources poses some complications for estimation and inference.

Recall that the state space of the individual agent's occupational choice problem is augmented with aggregate variables $\tilde{X}_t$, e.g. current and past values of skill prices, total factor productivity, cohort sizes, the distributions of schooling and occupation-specific experience, and other variables which predict future skill prices. Note that any such variable omitted by the econometrician is a potential source of endogeneity bias because it is likely to be correlated with current skill prices. The dimensionality of $\tilde{X}_t$ is potentially so large as to make solution - let alone estimation infeasible without some further simplification.

For any given values of the parameters to be estimated, the equilibrium laws of motion solve a fixed point problem which uses as input the solutions to the individual DP problems of all agents who interact in the economy. This implies that a single evaluation of the estimation criterion nests two layers of fixed point problems. The estimation criterion is based on a set of moment conditions which summarize occupational choices and wages obtained from micro-survey data.

1.3 Dynamic Discrete Games

The analysis of many economic and social phenomena requires the consideration of dynamic strategic interactions between a relatively small number of agents. The study of the dynamics of oligopoly industries is perhaps the most notorious example of these dynamic strategic interactions. Competition in oligopoly industries involves important investment decisions which are irreversible to a certain extent. Market entry in the presence of sunk entry costs is a simple example of this type of investment. Dynamic games are powerful tools for the analysis of these dynamic strategic interactions.

Until very recently, econometric models of discrete games had been limited to relatively simple static games. Two main econometric issues explain this limited range of applications: the computational burden in the solution of dynamic discrete games, and the indeterminacy problem associated with the existence of multiple equilibria. The existence of multiple equilibria is a prevalent feature in most empirical games where best response functions are nonlinear in other player's actions. Models with multiple equilibria do not have a unique reduced form and this incompleteness may pose practical and theoretical problems in the estimation of structural parameters. Furthermore, the computational burden in the structural estimation of games is especially severe. The dimension of the state space, and the cost of computing an equilibrium, increases exponentially with the number of heterogeneous players. An equilibrium is a fixed point of a system of best response operators and each player's best response is itself the solution to a dynamic programming problem.

In Section **??** we review recent papers which have taken Rust's framework for single-agent problems, combined with Hotz-Miller's estimation approach, as a starting point in the development of estimable dynamic discrete games of incomplete information. Private information state variables are a convenient way of introducing unobservables in the econometric model. Furthermore, under certain regularity conditions dynamic games of incomplete information have at least one equilibrium while that is not always the case in the corresponding complete information versions. We provide here a description of the basic framework with no payoff variables, based on suitably extended versions of the AS, IID and CIX assumptions.

Consider a game that is played by N players that we index by $i \in I = \{1, 2, \ldots, N\}$. Every period t these players decide simultaneously a discrete action. Let $a_{it} \in A = \{1, 2, \ldots, J\}$ be the action of player i at period t. At the beginning of period t a player is characterized by two vectors of state variables which affect her current utility: x_{it} and ε_{it}. Variables in x_{it} are common knowledge for all players in the game, but the vector ε_{it} is private information of player i. Let $x_t \equiv (x_{1t}, \ldots, x_{Nt})$

be the vector of common knowledge state variables, and similarly define $a_t = (a_{1t}, \ldots, a_{Nt})$, and let ε_t be the vector of players' private information. Let $U_i(a_t, x_t, \varepsilon_{it})$ player i's payoff function, that depends on the actions of all the players, the common knowledge state variables, and his own private information ε_{it}. A player chooses his action to maximize expected discounted intertemporal utility $E_t\left[\sum_{j=0}^{T-t} \beta^j U_i(a_{t+j}, x_{t+j}, \varepsilon_{i,t+j})\right]$, where $\beta \in (0,1)$ is the discount factor. Players have uncertainty about other player's current and future actions, about future common knowledge state variables, and about their own future private information shocks. We assume that $\{x_t, \varepsilon_t\}$ follows a controlled Markov process with transition probability function $F(x_{t+1}, \varepsilon_{t+1} \mid a_t, x_t, \varepsilon_t)$. This transition probability is common knowledge.

Let $\alpha = \{\alpha_i(x_t, \varepsilon_{it} : i \in I)\}$ be a set of strategies functions, one for each player. Player i's strategy does not depend on $\{\varepsilon_{jt} : j \neq t\}$ because the shocks are private information. Taking as given the strategies of all players other than i, the decision problem of player i is just a standard single-agent DP problem. Let $V_i^\alpha(x_t, \varepsilon_{it})$ be the value function of this DP problem. The Bellman equation is $V_i^\alpha(x_t, \varepsilon_{it}) = \max_{a_i \in A}\left\{v_i^\alpha(a_i, x_t, \varepsilon_{it})\right\}$ where, for every $a_i \in A$,

$$
\begin{aligned}
v_i^\alpha(a_i, x_t, \varepsilon_{it}) \equiv & E_{\varepsilon_{-it}}\{U_i(a_i, \alpha_{-i}(x_t, \varepsilon_{-it}), x_t, \varepsilon_{it})\} + \\
& \beta E_{\varepsilon_{-it}}\left\{\int V_i^\alpha(x_{t+1}, \varepsilon_{t+1}) \mathrm{d}F(x_{t+1}, \varepsilon_{t+1} \mid a_i, \alpha_{-i}(x_t, \varepsilon_{-it}), x_t, \varepsilon_t)\right\}
\end{aligned}
\tag{22}
$$

and $E_{\varepsilon_{-it}}$ represents the expectation over other player's private information shocks. The best response function of player i is $b_i(x_t, \varepsilon_{it}, \alpha_{-i}) = \arg\max_{a_i \in A}\left\{v_i^\alpha(a_i, x_t, \varepsilon_{it})\right\}$. This best response function gives the optimal strategy of player i if the other players behave, now and in the future, according to their respective strategies in α_{-i}. A Markov perfect equilibrium (MPE) in this game is a set of strategy functions α^* such that for each player i and for any (x_t, ε_{it}), we have that $\alpha_i^*(x_t, \varepsilon_{it}) = b_i(x_t, \varepsilon_{it}, \alpha_{-t}^*)$.

We now formulate the Assumptions AS, IID and CI-X the context of this game.

Assumption 10. AS-Game: The one-period utility function is additively separable in common knowledge and private information components: $U_i(a_t, x_t, \varepsilon_{it}) = u_i(a_t, x_t) + \varepsilon_{it}(a_{it})$, where $\varepsilon_{it}(a)$ is the ath component of vector ε_{it}. The support of $\varepsilon_{it}(a)$ is the real line for all a.

Assumption 11. IID-Game: Private information shocks ε_{it} are independently and identically distributed over agents and over time with CDF $G_\varepsilon(\varepsilon_{it})$ which has finite first moments and is continuous and twice differentiable in ε_{it}.

Assumption 12. CI-X-Game: Conditional on the current values of player's actions and common knowledge state variables, next period common knowledge state variables do not depend on current private information shocks: i.e., CDF$(x_{t+1} \mid a_t, x_t, \varepsilon_t) = F_x(x_{t+1}\ mid a_t, x_t)$.

As in the case of single-agent models, under Assumption 10 AS, 12 CIX and 11 IID the integrated value function $\overline{V}_i^\alpha \equiv \int \max_{a_i \in A}\left\{v_i^\alpha(a_i, x_t) + \varepsilon_{it}(a_i)\right\} dG_\varepsilon(\varepsilon_{it})$ where

$$
v_i^\alpha(a_i, x_t) = E_{\varepsilon_{-it}}\left\{u_i(a_i, \alpha_{-i}(x_t, \varepsilon_{-it}), x_t) + \beta \int \bar{V}_i^\alpha(x_{t+1}) \mathrm{d}F_x(x_{t+1} \mid a_i, \alpha_{-i}(x_t, \varepsilon_{-it}), x_t)\right\} \tag{23}
$$

Another implication of these assumptions is that a MPE can be described as a fixed point of a mapping in the space of CCPs which 'integrate out' player's private information variables. Given a set of strategy functions $\alpha = \{\alpha_i(x_t, \varepsilon_{it} : i \in I)\}$ we define a set of conditional choice probabilities $\mathbf{P}^\alpha = \left\{P_i^\alpha(a_i \mid x) : (i, a_i, x) \in I \times A \times X\right\}$ such that

$$P_i^\alpha(a_i \mid x) \equiv \Pr(\alpha_i(x_t, \varepsilon_{it}) = a_i \mid x_t = x) = \int I\left\{a_i = \arg\max_{j \in A}\left\{v_i^\alpha(j, x_t) + \varepsilon_{it}(j)\right\}\right\} \mathrm{dG}G_\varepsilon(\varepsilon_{it}). \tag{24}$$

These probabilities represent the expected behaviour of firm i from the point of view of the rest of the firms when firm i follows its strategy in α. The value functions $v_i^\alpha(a_i, x_t)$ depend on other players' strategies only through other players' choice probabilities. That is, in Equation 23, we can replace the expectation $E_{\varepsilon_{-it}}\{\ldots \alpha_{-i}(x_t, \varepsilon_{-it} \ldots)\}$ by $\sum_{a_{-i}} \Pr(a_{-i} \mid x_t; \alpha)$ where $\sum_{a_{-i}}$ represents the sum over all possible values in A^{N-1}, and $\Pr(a_{-i} \mid x_t; \alpha) = \prod_{j \neq i} P_i^\alpha(a_j \mid x_t)$. To emphasize this point we will use notation $v_i^{\mathbf{P}}$ instead of v_i^α to represent these value functions. Then, we can define player i's best response probability function as:

$$\Lambda\left(a_i \mid v_i^{\mathbf{P}}(\cdot, x_t)\right) \equiv \int I\left\{a_i = \arg\max_{j \in A}\left\{v_i^{\mathbf{P}}(j, x_t) + \varepsilon_{it}(j)\right\}\right\} \mathrm{d}G_\varepsilon(\varepsilon_{it}). \tag{25}$$

2 Estimation Methods for Sigle-Agent Models

2.1 Methods for Rust's Model

We introduced Rust's model as a single-agent model with the additive separability, i.i.d., conditional independence and logistic assumptions. We describe four methods/algorithms which have been applied to the estimation of this class of models: (1) Rust's nested fixed point algorithm; (2) Hotz-Miller's CCP method; (3) The NPL algorithm, a recursive CCP method; and (4) A simulation-based CCP method. The first of these is a full solution method, i.e., the DP problem is solved for every trial value of the parameters. Methods (2) - (4) avoid repeated full solutions of the DP problem, taking advantage of the existence of an invertible mapping between conditional choice probabilities and differences in choice-specific value functions.

Almost all applications in models with more than two choice alternatives have imposed the CLOGIT assumption on unobservables. Although the CLOGIT assumption per se is not essential for the statistical properties of Rust's model and the methods which we review here, it leads to closed form expressions for several of the econometric model's key objects. This is an advantage which all four of these methods fully exploit, and CCP methods are specially dependent on it.

2.1.1 Rust's nested fixed point algorithm

The nested fixed point algorithm (NFXP) is a gradient iterative search method to obtain the MLE of the structural parameters. More specifically, this algorithm combines a BHHH method (outer algorithm), that searches a root of the likelihood equations, with a value function or policy iteration method (inner algorithm), that solves the dynamic programming problem for each trial value of the structural parameters. The algorithm is initialized with an arbitrary vector of structural parameters, say $\hat{\theta}_0$. A BHHH iteration is defined as:

$$\hat{\theta}_{k+1} = \hat{\theta}_k + \left(\sum_{i=1}^N \frac{\partial \ell_i(\hat{\theta}_k)}{\partial \theta} \frac{\partial \ell_i(\hat{\theta}_k)}{\partial \theta'}\right)^{-1} \left(\sum_{i=1}^N \frac{\partial \ell_i(\hat{\theta}_k)}{\partial \theta}\right) \tag{26}$$

where $\ell_i(\theta)$ is the log-likelihood function for individual i. Given the form of the likelihood in Equation 10, and that $\theta = (\theta'_u, \theta'_Y, \theta'_f)$, we have:

$$\frac{\partial \ell_i(\theta)}{\partial \theta} = \left[\begin{array}{c} \sum_{t=1}^{T_i} \frac{\partial}{\partial \theta_u} \log P\left(a_{it} \mid x_{it}, \theta\right) \\ \sum_{t=1}^{T_i} \frac{\partial}{\partial \theta_Y} \log P\left(a_{it} \mid x_{it}, \theta\right) + \sum_{t=1}^{T_i} \frac{\partial}{\partial \theta_Y} \log f_Y\left(y_{it} \mid a_{it}, x_{it}, \theta_Y\right) \\ \sum_{t=1}^{T_i} \frac{\partial}{\partial \theta_f} \log P\left(a_{it} \mid x_{it}, \theta\right) + \sum_{t=1}^{T_i - 1} \frac{\partial}{\partial \theta_f} \log f_x\left(x_{i,t+1} \mid a_{it}, x_{it}, \theta_f\right) \end{array} \right]. \tag{27}$$

Note that $\frac{\partial \ell_i(\theta)}{\partial \theta}$ is a column vector. The terms $\frac{\partial}{\partial \theta_Y} \log f_Y$ and $\frac{\partial}{\partial \theta_f} \log f_x$ are standard because the transition probability function and the payoff function are primitives of the model. However, to obtain $\frac{\partial}{\partial \theta} \log P$ we need to solve the DP problem for $\theta = \hat{\theta}_k$ in order to compute the conditional choice probabilities and their derivatives with respect to the components of $\hat{\theta}_k$. There are different ways to solve the DP problem. When the model has finite horizon (i.e., T is finite) the standard approach is to use backward induction. For infinite horizon models, one can use either value function iterations (described below) or policy function iterations, or a hybrid of both.

Consider a version of Rust's DP conditional logit model with infinite horizon and discrete observable state variables x. Let $\overline{\mathbf{V}}(\theta)$ be the column vector of values $\{\overline{V}(x,\theta) : x \in X\}$. Following Equation 12, this vector of values can be obtained as the unique fixed point in $\overline{\mathbf{V}}$ of the following Bellman equation in vector form:

$$\overline{\mathbf{V}} = \log\left(\sum_{a=0}^{J} \exp\left\{ \mathbf{u}(a,\theta) + \beta \mathbf{F}_x(a) \overline{\mathbf{V}} \right\} \right) \tag{28}$$

where $\mathbf{u}(a,\theta)$ is the column vector of utilities $\{u(a,x,\theta) : x \in X\}$ and $\mathbf{F}_x(a)$ is the transition probability matrix with elements $f_x(x' \mid a, x)$ for all x and x' in X, where x and x' denote the current state and next period's state, respectively. The choice probabilities $P(a \mid x, \theta)$ have the conditional logit form:

$$P(a \mid x, \theta) = \frac{\exp\left\{ u\left(a, x_{it}, \theta\right) + \beta \mathbf{F}_x(a,x)' \overline{\mathbf{V}}(\theta) \right\}}{\sum_{j=0}^{J} \exp\left\{ u\left(j, x_{it}, \theta\right) + \beta \mathbf{F}_x(j,x)' \overline{\mathbf{V}}(\theta) \right\}} \tag{29}$$

where $\mathbf{F}_x(a,x)$ is the column vector $\left\{ f_x(x' \mid a, x) : x' \in X \right\}$. To obtain the gradient $\frac{\partial}{\partial \theta} \log P_{it}$ of this DP conditional logit model, it is useful to take into account that the denominator in Equation 29 is equal to the exponential of $\overline{V}(x_{it}, \theta)$. It can be shown that this gradient has the following analytic form:

$$\begin{aligned} \frac{\partial \log P_{it}}{\partial \theta_u} &= \frac{\partial u\left(a_{it}, x_{it}\right)}{\partial \theta_u} + \beta \frac{\partial \overline{\mathbf{V}}'}{\partial \theta_u} \mathbf{F}_x\left(a_{it}, x_{it}\right) - \frac{\partial \overline{V}\left(x_{it}\right)}{\partial \theta_u} \\ \frac{\partial \log P_{it}}{\partial \theta_f} &= \beta \left(\frac{\partial \mathbf{F}_x\left(a_{it}, x_{it}\right)'}{\partial \theta_f} \overline{\mathbf{V}} + \frac{\partial \overline{\mathbf{V}}'}{\partial \theta_f} \mathbf{F}_x\left(a_{it}, x_{it}\right) \right) - \frac{\partial \overline{V}\left(x_{it}\right)}{\partial \theta_f}. \end{aligned} \tag{30}$$

The expression for $\frac{\partial}{\partial \theta_Y} \log P_{it}$ is equivalent to that of $\frac{\partial}{\partial \theta_u} \log P_{it}$. And given the Bellman equation in Equation 28, the Jacobian matrix $\frac{\overline{\mathbf{V}}(\theta)}{\partial \theta'}$ is:

$$\begin{aligned} \frac{\partial \overline{\mathbf{V}}(\theta)}{\partial \theta'_u} &= \left(I - \beta \sum_{a=0}^{J} \mathbf{P}(a \mid \theta) * \mathbf{F}_x(a) \right)^{-1} \times \left(\sum_{a=0}^{J} \mathbf{P}(a \mid \theta) * \frac{\partial \mathbf{u}(a,\theta)}{\partial \theta'_u} \right) \\ \frac{\partial \overline{\mathbf{V}}(\theta)}{\partial \theta'_f} &= \beta \left(I - \beta \sum_{a=0}^{J} \mathbf{P}(a \mid \theta) * \mathbf{F}_x(a) \right)^{-1} \times \left(\sum_{a=0}^{J} \mathbf{P}(a \mid \theta) * \frac{\partial \mathbf{F}_x(a)}{\partial \theta'_f} \overline{\mathbf{V}}(\theta) \right) \end{aligned} \tag{31}$$

where $\mathbf{P}(a \mid \theta)$ is the column vector of choice probabilities $\{P(a \mid x, \theta) : x \in X\}$, and $*$ represents the element-by-element product. Note that we obtain the gradient of the value functions in a relatively simple manner as a by-product of the iterative DP solution method. This is an important additional advantage of the DP conditional logit specification. Without it, it would be necessary to "perturb" each element of θ and obtain new solutions to the DP model for each perturbation in order to compute numerical derivatives, which is more costly.

The NFXP algorithm proceeds as follows. We start with an arbitrary value of θ, say θ_0. Given θ_0, in the inner algorithm we obtain the vector $\overline{\mathbf{V}}(\hat{\theta}_0)$ by successive iterations in the Bellman equation 28: starting with some guess $\overline{\mathbf{V}}_0$, we iterate in $\overline{\mathbf{V}}_{h+1} = \log\left(\sum_{a=0}^{J} \mathbf{u}(a, \hat{\theta}_0) + \beta \mathbf{F}_x(a) \mathbf{V}_h\right)$ until convergence. Then, given θ_0 and $\overline{\mathbf{V}}(\hat{\theta}_0)$, we construct the choice probabilities $P(a \mid x, \hat{\theta}_0)$ using the formula in Equation 29, the matrix $\frac{\overline{\mathbf{V}}(\hat{\theta}_0)}{\theta'}$ using Equation 31, and the gradient $\frac{\partial \ell_i(\hat{\theta}_0)}{\partial \theta}$ using Equation 30. Finally, in the "outer" algorithm we use the gradient $\frac{\partial \ell_i(\hat{\theta}_0)}{\partial \theta}$ to make a new BHHH iteration to obtain θ_1. We proceed in this way until the distance between $\hat{\theta}_{k+1}$ and $\hat{\theta}_k$ or the difference in the likelihoods is smaller than a pre-specified convergence constant.

When the model has finite horizon, we can solve for the value function, its gradient and choice probabilities using backward induction in the inner algorithm of the NFXP. That is, the sequence of value vectors at ages T, $T-1$, etc, can be obtained starting with $\overline{\mathbf{V}}_T(\hat{\theta}) = \log\left(\sum_{a=0}^{J} \exp\{\mathbf{u}_T(a, \hat{\theta})\}\right)$, and then using the recursive formula

$$\overline{\mathbf{V}}_t(\hat{\theta}) = \log\left(\sum_{a=0}^{J} \exp\{\mathbf{u}_t(a, \hat{\theta}) + \beta \mathbf{F}_{x,t}(a) \overline{\mathbf{V}}_{t+1}(\hat{\theta})\}\right) \quad \text{for } t \leq T-1.$$

At each iteration the choice probabilities are

$$P_t(a \mid \hat{\theta}) = \frac{\exp\{u_t(a, \hat{\theta}) + \beta \mathbf{F}'_{x,t}(a) \overline{\mathbf{V}}_{t+1}(\hat{\theta})\}}{\sum_{j=0}^{J} \exp\{u_t(j, \hat{\theta}) + \beta \mathbf{F}'_{x,t}(j) \overline{\mathbf{V}}_{t+1}(\hat{\theta})\}},$$

and the gradients have the following recursive forms

$$\frac{\partial \overline{\mathbf{V}}_t(\hat{\theta})}{\partial \theta'_u} = \sum_{a=0}^{J} \mathbf{P}_t(a \mid \hat{\theta}) * \left\{ \frac{\partial \mathbf{u}_t(a, \hat{\theta})}{\partial \theta'_u} + \beta \mathbf{F}_{x,t}(a) \frac{\partial \overline{\mathbf{V}}_{t+1}(\hat{\theta})}{\partial \theta'_u} \right\}$$

and

$$\frac{\partial \overline{\mathbf{V}}_t(\hat{\theta})}{\partial \theta'_f} = \beta \sum_{a=0}^{J} \mathbf{P}_t(a \mid \hat{\theta}) * \left\{ \frac{\partial \mathbf{F}_{x,t}(a)}{\partial \theta'_f} \overline{\mathbf{V}}_{t+1}(\hat{\theta}) + \beta \widehat{\mathbf{F}}_{x,t}(a) \frac{\partial \overline{\mathbf{V}}_{t+1}(\hat{\theta})}{\partial \theta'_f} \right\}.$$

As any other gradient method, the NFXP algorithm returns a solution to the likelihood equations. In general, the likelihood function of this class of models is not globally concave. Therefore, some global search is necessary to check whether the root of the likelihood equations that has been found is actually the global maximum and not just a local optimum.

We have described the NFXP algorithm in the context of full information MLE (FIML). However, most applications of this algorithm have considered a sequential partial likelihood approach. In the first step, the partial likelihoods $\sum_{i,t} \log f_Y(y_{it} \mid a_{it}, x_{it}, \theta_Y)$ and $\sum_{i,t} \log f_x(x_{i,t+1} \mid a_{it}, x_{it}, \theta_f)$ are separately maximized to obtain estimators of parameters θ_Y and θ_f, respectively. This step does not involve solving the DP problem. Given these estimators, in the second step the parameters in θ_u are estimated using the NFXP algorithm and the partial likelihood $\sum_{i,t} \log P(a_{it} \mid x_{it}, \theta_u, \hat{\theta}_Y, \hat{\theta}_f)$. Multiple-step strategies such as this are widely used to alleviate the computational burden of estimation of dynamic programming structural models. The two-step partial likelihood approach of Rust's model relies on Assumption 3 CIX and 4 CIY as well as rational expectations. It can greatly simplify the estimation problem in models with many parameters in the transition probabilities. In Example 7, the state variables with stochastic transitions were health status, health expenses, marital status and public pension points. The payoff function is the labour earnings equation. Nevertheless, it should be noted that in the sequential partial likelihood approach the calculation of standard errors in the second step correcting for estimation error in the first step is not a trivial task. It is often no more costly to consider a third estimation step consisting of a single FIML-BHHH iteration which, as it is well known, delivers an asymptotically efficient estimator as well as the correct standard errors.

References:

Kennan, J., Walker, J.R., 2005. The Effect Income on Individual Migration Decisions. Manuscript, University of Wisconsin-Madison.

2.1.2 Hotz-Miller's CCP method

The main advantages of the NFXP algorithm are its conceptual simplicity and, more importantly, that it gives the MLE which is asymptotically efficient under the assumptions of the model. The main limitation of this algorithm is its computational cost. In particular, the DP problem has to be solved exactly for each trial value of the structural parameters. Given the cost of solving some DP problems, this characteristic of the algorithm limits the range of applications that are feasible, even in the DP conditional logit case.

Hotz and Miller (1993) observed that, under the assumptions of Rust model, it is not necessary to solve the DP problem even once in order to estimate the structural parameters. A key idea in their methods is that, using nonparametric estimates of choice and transition probabilities, it is possible to obtain a simple representation of the choice-specific value functions $v(a,x,\theta)$ for values of θ around the true vector of structural parameters. This representation is particularly simple and useful for estimation in the DP conditional logit model with linear-in-parameters utility as in Example **??**. We describe Hotz-Miller's method in detail for this case but we also include a brief discussion of alternative specifications including the nonlinear, infinite horizon case. Furthermore, we assume that the dataset does not include payoff variables or, if it does, Assumption 4 CIY holds and θ_Y has been estimated and subsumed into the known z functions.

Suppose that $u(a,x_{it},\theta_u) = z(a,x_{it})'\theta_u$ where $z(a,x_{it})$ is a vector of know basis functions. An example of this is the entry/exit model with no strategic interactions, Example **??** in Section 1. In that model, the decision is binary, the vector of observable state variables is $(S_t, a_{i,t-1})$, and the one-period payoffs are $u(0,x_{it},\theta_u) = 0$ and $u(1,x_{it},\theta_u) = \theta_{RS}\log(S_t) - \theta_{FC} - \theta_{EC}(1 - a_{i,t-1})$. Therefore, we can write these payoffs as $u(a,x_{it},\theta_u) = z(a,x_{it})'\theta_u$ with $z(0,x_{it})' = (0,0,0)$, $z(1,x_{it})' = \log(S_t), -1, -(1-a_{i,t-1})$, and $\theta_u = (\theta_{RS},\theta_{FC},\theta_{EC})'$. In Hotz and Miller's representation the choice-specific value functions are written as follows:

$$v(a,x_t,\theta) = \tilde{z}(a,x_t,\theta)\theta_u + \tilde{e}(a,x_t,\theta) \tag{32}$$

where $\tilde{z}(a, x_t, \theta)$ is the expected and discounted sum of current and future z vectors $\{z(a_{t+j}, x_{t+j}) : j = 0, 1, 2, \ldots\}$ which may occur along all possible histories originating from the choice of a in state x_t, if the individual behaves optimally in the future. More formally,

$$\begin{aligned} \tilde{z}(a, x_t, \theta) &= z(a, x_t) + \sum_{j=1}^{T-t} \beta^j E_{(x_{t+j}, \varepsilon_{t+j}) \mid a_t = a, x_t} \left[z\left(\alpha\left(x_{t+j}, \varepsilon_{t+j}, \theta\right), x_{t+j}\right)\right] \\ &= z(a, x_t) + \sum_{j=1}^{T-t} \beta^j E_{x_{t+j} \mid a_t = a, x_t} \left[\sum_{a_{t+j}=0}^{J} P\left(a_{t+j} \mid x_{t+j}, \theta\right) z\left(a_{t+j}, x_{t+j}\right)\right] \end{aligned} \tag{33}$$

where $\alpha(\cdot)$ is the optimal decision rule as defined in Section 1 and $E_{x_{t+j} \mid a_t = a, x_t}(\cdot)$ represents the expectation over the distribution of x_{t+j} condition on $(a_t = a, x_t)$, given that state variables evolve as a controlled stochastic process defined by the optimal decision rule and transition probability functions. Likewise, $\tilde{e}(a, x_t, \theta)$ is the expected and discounted sum of the stream $\{\varepsilon(a_{t+j}) : j = 1, 2, \ldots\}$ along the histories originating at $(a_t = a, x_t)$ if the individual behaves optimally in the future.

$$\begin{aligned} \tilde{e}(a, x_t, \theta) &= \sum_{j=1}^{T-t} \beta^j E_{(x_{t+j}, \varepsilon_{t+j}) \mid a_t = a, x_t} \left[\varepsilon\left(\alpha\left(x_{t+j}, \varepsilon_{t+j}, \theta\right)\right)\right] \\ &= \sum_{j=1}^{T-t} \beta^j E_{x_{t+j} \mid a_t = a, x_t} \left[\sum_{a_{t+j}=0}^{J} P\left(a_{t+j} \mid x_{t+j}, \theta\right) e\left(a_{t+j}, x_{t+j}\right)\right]. \end{aligned} \tag{34}$$

The function $e(a, x_t)$ is the expectation of $\varepsilon_t(a)$ conditional on x_t and on alternative a being optimal, i.e., $e(a, x_t) \equiv E(\varepsilon_t(a \mid x_t, \alpha(x_t, \varepsilon_t) = a)$. Hotz and Miller showed that the conditional expectation $e(a, x_t)$ is a function of choice probabilities and the distribution G_ε only. To see why, first note that the event $\{\alpha(x_t, \varepsilon_t) = a\}$ is equivalent to $\left\{v(a, x_t) + \varepsilon_t(a) \geq v(a', x_t) + \varepsilon_t(a') \text{ for any } a' \neq a\right\}$. Therefore,

$$\begin{aligned} e(a, x_t) &= E\left(\varepsilon_t(a) \mid x_t, v(a, x_t) + \varepsilon_t(a) \geq v\left(a', x_t\right) + \varepsilon_t\left(a'\right) \text{ for any } a' \neq a\right) \\ &= \frac{1}{P(a \mid x_t)} \int \varepsilon_t(a) I\left\{\varepsilon_t\left(a'\right) - \varepsilon_t(a) \leq v(a, x_t) - v\left(a', x_t\right) \forall a' \neq a\right\} \mathrm{dG}_\varepsilon(\varepsilon_t) \end{aligned} \tag{35}$$

This expression shows that $e(a, x_t)$ depends on the choice probability $P(a \mid x_t)$, the distribution of ε_t, and the vector of value fifference $\tilde{v}(x_t) \equiv \{v(a, x_t) - v(0, x_t) : a \in A\}$. Second, choice probabilities depend on the primitives of the model only through the probability distribution G_ε and the vector of value differences $\tilde{v}(x_t)$: i.e., $P(a \mid x_t) = \int I\left\{\varepsilon_t(a') - \varepsilon_t(a) \leq v(a, x_t) - v(a', x_t) \forall a' \neq a\right\} dG_\varepsilon(\varepsilon_t)$. Hotz and Miller proved that this mapping relating choice probabilities and value differences is invertible. Thus, solving this inverse mapping into Equation 35, we have that $e(a, x_t)$ is a function of choice probabilities and the distribution G_ε only. The particular form of the mapping that relates CCPs and value differences depends on the probability distribution G_ε. For instance, in the model of Example **??**, $v(1, x_t)$ is the firm's expected stream of profits if it decides to be active in state x_t and behave optimally thereafter. For this binary choice model, $\tilde{v}(x_t) = v(1, x_t) - v(0, x_t)$. Under assumption CLOGIT on G_ε the probability that the firm will be active is $P(a \mid x_t) = \frac{\exp(\tilde{v}(x_t))}{1+\exp(\tilde(v)(x_t))}$. Inverting this relationship, we get $\tilde{v} = \log(P(1 \mid x_t)) - \log(P(0 \mid x_t))$. Also, under the CLOGIT assumption, the conditional expectation $e(1, x_t)$ is equal to $\gamma - \log(\frac{\exp(\tilde{v}(x_t))}{1+\exp(\tilde(v)(x_t))})$, where γ is Euler's constant, and similarly $e(0, x_t) = \gamma - \log(\frac{1}{1+\exp(\tilde{(}v)(x_t))})$. Using the inverse mapping above, it turns out that $e(a, x_t) = \gamma - \log P(a \mid x_t, \theta)$.

Note that $\tilde{z}(a, x_t, \theta)$ and $\tilde{e}(a, x_t, \theta)$ depend on θ only through the parameters in the transition probabilities, θ_f, and the vector of CCPs, $\mathbf{P}(\theta) = \{P(a \mid x, \theta) : (a, x) \in A \times X\}$. In fact, we can define $\tilde{z}$ and $\tilde{e}$ for an arbitrary vector of CCPs (optimal or not). To emphasize this point, we use the notation $\tilde{z}^{\mathbf{P}}(a, x_t)$ to represent $\sum_{j=1}^{T-t} \beta^j E_{x_{t+j}|a_t=a,x_t} \left[\sum_{a\prime=0}^{J} P\left(a' \mid x_{t+j}, \theta\right) z\left(a', x_{t+j}\right) \right]$ and $\tilde{e}^{\mathbf{P}}(a, x_t)$ to represent $\sum_{j=1}^{T-t} \beta^j E_{x_{t+j}|a_t=a,x_t} \left[\sum_{a\prime=0}^{J} P\left(a' \mid x_{t+j}, \theta\right) e\left(a', x_{t+j}\right) \right]$.

Let $\theta^0 = (\theta_u^0, \theta_f^0)$ be the true value of θ in the population of individuals under study. And let $\mathbf{P}^0$ be the conditional choice probabilities in the population. If we knew $(\mathbf{P}^0, \theta_f^0)$, we could construct the values $\tilde{z}^{\mathbf{P}^0}(a, x)$ and $\tilde{e}^{\mathbf{P}^0}(a, x)$ and then use Hotz and Miller's representation to approximate the choice-specific values $v(a, x, \theta)$ as simple linear functions of θ_u close to its true value. Although we do not know $(\mathbf{P}^0, \theta_f^0)$, we can estimate them consistently without having to sole the DP problem. Consistent estimates of transition probabilities can be obtained using a (partial) MLE of θ_f^0 that maximizes the (partial) likelihood $\sum_{i,t} \log f_x(x_{i,t+1} \mid a_{it}, x_{it}, \theta_f)$. Conditional choice probabilities can be estimated using nonparametric regression methods (i.e., $P^0(a \mid x) = E(I\{a_{it} = a\} \mid x_{it} = x)$) such as a Nadaraya-Watson kernel estimator or a simple frequency estimator. Let $\hat{\mathbf{P}}$ and $\hat{\theta}_f$ be the estimators of $\mathbf{P}^0, \theta_f^0$, respectively. Based on these estimates, Hotz and Miller's idea is to approximate $v(a, x_{it}, \theta)$ by $\tilde{z}^{\hat{\mathbf{P}}}(a, x_i t)\theta_u + \tilde{e}^{\hat{\mathbf{P}}}$ in order to obtain the CCP's in Equation 29 They propose the GMM estimator that solves in θ_u the sample moment conditions:

$$\sum_{i=1}^{N} \sum_{t=1}^{T_i} H(x_{it}) \times \begin{bmatrix} I\{a_{it} = 1\} - \frac{\exp\left\{\tilde{z}^{\hat{\mathbf{P}}}(1, x_{it})\theta_u + \tilde{e}^{\hat{\mathbf{P}}}(1, x_{it})\right\}}{\sum_{a=0}^{J} \exp\left\{\tilde{z}^{\hat{P}}(a, x_{it})\theta_u + \tilde{e}^{\hat{P}}(a, x_{it})\right\}} \\ \vdots \\ I\{a_{it} = J\} - \frac{\exp\left\{\tilde{z}^{\hat{\mathbf{P}}}(J, x_{it})\theta_u + \tilde{e}^{\hat{\mathbb{P}}}(J, x_{it})\right\}}{\sum_{a=0}^{J} \exp\left\{\tilde{z}^{\hat{P}}(a, x_{it})\theta_u + \tilde{e}^{\hat{P}}(a, x_{it})\right\}} \end{bmatrix} = 0 \tag{36}$$

where $H(x_{it})$ is a matrix with dimension $\dim(\theta_u) \times J$ with functions of x_{it} which are used as instruments. For instance, in the entry-exit model of Example **??** the vector θ_u has three parameters, the decision is binary ($J = 1$), and the vector of observable state variables is $(S_t, a_{i,t-1})$. Therefore, a possible choic

e for $H(x_i t)$ is the vector $(1, S_t, a_{i,t-1})'$.

The main advantage of this estimator is its computational simplicity. Nonparametric estimation of choice probabilities is a (relatively) simple task. The main task is the computation of the values $\tilde{z}^{\hat{\mathbf{P}}}(a, x_{it})$ and $\tilde{e}^{\hat{\mathbf{P}}}(a, x_{it})$, based on a recursive version of expressions 33 and 34. However, these values are calculated just once and remain fixed in the search for the Hotz-Miller estimator. In contrast, in a full solution method such as the NFXP algorithm these values are recomputed exactly for each trial value of . Thus, Hotz-Miller's method greatly reduces the computational burden of NFXP's 'inner' algorithm. Another important advantage of Hotz-Miller's CCP estimator is that, for the DP conditional logit model with linear-in-parameters utility, the system of equations 36 that defines the estimator has a unique solution. Therefore, a global search is not needed.

Previous conventional wisdom was that Hotz-Miller's estimator achieved a significant computational gain at the expense of efficiency, both in finite samples and asymptotically. Thus, researchers had the choice between two extremes: a full solution NFXP-ML estimator with the attendant computational burden, or the much faster but less efficient Hotz-Miller estimator. However, there is a pseudo-maximum likelihood version of Hotz-Miller's estimator is asymptotically equivalent to partial MLE. The 'two-step' pseudo-maximum likelihood (PML) estimator is defined as the value of θ_u that maximizes the pseudo-likelihood function

$$Q\left(\theta_u, \hat{\mathbf{P}}, \hat{\theta}_f\right) = \sum_{i=1}^{N} \sum_{t=1}^{T_i} \log \frac{\exp\left\{\tilde{z}^{\hat{\mathbf{P}}}\left(a_{it}, x_{it}\right)\theta_u + \tilde{e}^{\hat{\mathbf{P}}}\left(a_{it}, x_{it}\right)\right\}}{\sum_{a=0}^{J} \exp\left\{\tilde{z}^{\hat{\mathbf{P}}}\left(a, x_{it}\right)\theta_u + \tilde{e}^{\hat{\mathbf{P}}}\left(a, x_{it}\right)\right\}}. \tag{37}$$

The asymptotic variance of this two-step PML estimator is just equal to the variance of the partial MLE of θ_u described at the end of Section **??**. That is, the initial nonparametric estimator of $\mathbf{P}^0$ and the PML estimator of θ_u^0 are asymptotically independent and therefore there is no asymptotic efficiency loss from using an inefficient initial estimator of $\mathbf{P}^0$. But its finite sample bias an be much larger. Imprecise initial estimates of choice probabilities do not affect the asymptotic properties of the estimator, but they can generate serious small sample biases in the two-step PML estimator and, more generally, in the whole class of Hotz-Miller's CCP estimators. A recursive extension of the two-step method, which we describe in the next subsection, deals with this problem.

We now describe in some detail on the computation of the values $\tilde{z}^{\hat{\mathbf{P}}}(a, x_{it})$ and $\tilde{e}^{\hat{\mathbf{P}}}(a, x_{it})$. As defined Equations 33 and 34, these values must satisfy the following recursive expressions:

$$\begin{aligned} \tilde{z}_t^{\mathbf{P}}\left(a, x_t\right) &= z\left(a, x_t\right) + \beta \sum_{x_{t+1} \in X} f_x\left(x_{t+1} \mid a, x_t\right) \times \left[\sum_{a'=0}^{J} P_{t+1}\left(a' \mid x_{t+1}\right) \tilde{z}_t^{\mathbf{P}} \right. \\ \tilde{e}_t^{\mathbf{P}}\left(a, x_t\right) &= \beta \sum_{x_{t+1} \in X} f_x\left(x_{t+1} \mid a, x_t\right) \times \left[\sum_{a'=0}^{J} P_{t+1}\left(a' \mid x_{t+1}\right)\left(e_{t+1}\left(a', x_{t+1}\right) + \tilde{e}_{t+1}^{\mathbf{P}} \right.\right. \end{aligned} \tag{38}$$

For a finite horizon model, we can obtain the sequence of values $\left\{\tilde{z}^{\mathbf{P}}(a, x_{it}) \text{ and } \tilde{e}^{\mathbf{P}}(a, x_{it})\right\} : t = 1, 2, \ldots, T$ using the backward induction. Starting at the last period, we have $\tilde{z}^{\hat{\mathbf{P}}}(a, x_T) = z(a, x_T)$ and $\tilde{e}^{\hat{\mathbf{P}}}(a, x_T) = 0$. For every $t < T$ we iterate on Equation 38. The computational burden incurred is equivalent to that of solving the finite horizon DP problem once.

For stationary infinite horizon models succh as Example **??**, we can rewrite Equation 38 as follows:

Reference: Murphy, A., 2008. A Dynamic Model of Housing Supply Manuscript. Department of Economics. Duke University.