

1 Single-Agent Dynamic Models: The Basics

Notation

- As usual, agents $i = \{1, \dots, N\}$.
- Discrete time $t = \{1, \dots, T\}$ where $T \leq \infty$.
- a_{it} is agent i 's action in period t , where $a_{it} \in \mathcal{A} = \{1, \dots, A\}$.
- States: $s_{it} = (x_{it}, \epsilon_{it})$
 - public: $x_{it} \in \mathcal{X}$ with \mathcal{X} discrete.
 - private: $\epsilon_{it} = \{\epsilon_{i1t}, \dots, \epsilon_{iAt}\}$ iid $\sim G : \mathbf{R} \rightarrow [0, 1]$
- Discount factor β sometimes assumed to be known, sometimes estimated.
- Timing and information:
 1. Agents observe private shocks ϵ_{it} and x_{it} at start of period t ; they also know transition probabilities $f_a(x_{it+1}|x_{it})$ $[:=f_x(x_{it+1}|x_{it}, a_{it})]$;
 - Implicitly we're saying f is a function of a_{it} too. This is bad notation; the subscript is usually reserved for denoting the random variable for which the probability measure F applies. Aguirregabiria and Mira (2010) denote this $f_x(x_{it+1}|x_{it}, a_{it})$, and I will do the same for the remainder of these notes.
 - we've assumed $f_x(x_{it+1}|x_{it}, a_{it}, \epsilon_{it}) = f_x(x_{it+1}|x_{it}, a_{it})$
 - Sometimes we're also interested in estimating the parameters of f_x ; can be done non-parametrically or parametrically (Aguirregabiria and Mira, 2010).
 2. Agents choose action a_{it} after observing ϵ_{it} ;
 3. x_{it} updates according to transition probabilities; enter period $t + 1$.
- Conditional on x_{it} previous history does not matter – i.e. we've assumed x_{it} follows a **Markov process**, and given Bellman's principle of optimality this implies a_{it} depends exclusively on current and (expectations over) future values of the states.

The Payoff Structure

- Additive separability in flow payoff from a for period t :¹

$$u_a(x_{it}) + \epsilon_{iat}. \tag{1}$$

- Present value (ex-post) at time $t = 1$ of payoff from the sequence of actions $\{a_i\}_{t=0}^T$:

$$\sum_{t=1}^T \sum_{a=1}^A \beta^{(t-1)} \mathbb{1}_{\{a_{it}=a\}} [u_a(x_{it}) + \epsilon_{iat}] \tag{2}$$

These notes are based on Howard Smith's lectures from the 2018-2019 academic year.

Of course, this is our interpretation of the material Howard presented. Good content is his; mistakes are ours.

¹This is a simplifying assumption, relaxed e.g. in Keane and Wolpin (1997).

- The ex-ante value function is

$$V(x_{it}) = \int_{\epsilon_t} \max_{a \in \mathcal{A}} \{u_a(x_{it}) + \epsilon_{iat} + \beta \sum_{x_{it+1}} V(x_{it+1}) f_x(x_{it+1}|x_{it}, a)\} dG_{\epsilon_t} \quad (3)$$

i.e. we're integrating over the entire vector ϵ_t of length equal to the size of the action space, $|\mathcal{A}|$. $V(x_{it})$ is the expected maximum payoff given x_{it} before ϵ_{it} (and thus the optimal action) is known. The summation in the second term is because we're also taking an expectation over the (discrete) state space at $t + 1$.

Conditional Choice Probability (CCP)

- Plays the **same role as δ in static models**.
- For a given choice a the agent's value function is

$$u_a(x_{it}) + \epsilon_{iat} + \beta \sum_{x_{it+1}} V(x_{it+1}) f_x(x_{it+1}|x_{it}, a). \quad (4)$$

- We decompose this into this periods shock and the “mean choice specific value”

$$v_a(x_{it}) = u_a(x_{it}) + \beta \sum_{x_{it+1}} V(x_{it+1}) f_x(x_{it+1}|x_{it}, a) \quad (5)$$

which is deterministic (since we're going to solve for the value function, and in any event it's an expectation over the max of future payoffs.)

- Now we see the parallel with the static model, since

$$(i \text{ choose } a) \iff v_a(x_{it}) + \epsilon_{iat} > v_{a'}(x_{it}) + \epsilon_{ia't} \quad \forall a' \neq a. \quad (6)$$

- From this we obtain the *conditional choice probability* (CCP)

$$P_a(v(x_{it})) := P(a|x_{it}; \theta, \gamma) = \int_{\epsilon_t} \mathbb{1}\{v_a(x_{it}) + \epsilon_{iat} > v_{a'}(x_{it}) + \epsilon_{ia't} \quad \forall a' \neq a\} dG_{\epsilon_t} \quad (7)$$

the first identity makes it clear that the CCP is the conditional probability of choosing action a given the states and the parameter vector $(\theta', \gamma')'$, which describe the utility function and f_x respectively.

- Taking this one step further, if we are willing to assume that the unobserved states ϵ_{it} are type-one Extreme Value we can follow the exact same argument as with the static logit to obtain

$$P_a(v(x_{it})) = \frac{\exp v_a(x_{it})}{\sum_{j=0}^{|\mathcal{A}|} \exp v_j(x_{it})} \quad (8)$$

i.e. simple logit with the choice-specific value functions in place of mean utility.

- Without the Type-One EV assumption, we can still invert this bad boy to get

$$v(x_{it}) = P^{-1}(P(x_{it})) \quad (9)$$

where $\mathbf{v}(x_{it})$ and $\mathbf{P}(x_{it})$ are both vector-valued (i.e. $\mathbf{P}(x_{it})$ stacks the $P_a(\mathbf{v}(x_{it}))$'s over a , just like \mathbf{v}). This is possible for any absolutely continuous G_{ϵ_t} .

The Likelihood

- Data is agent level: $\{a_{it}, x_{it}; i = 1, \dots, N; t = 1, \dots, T_i\}$;
- θ, γ are parameters in $u(\cdot)$ and f_x respectively;
- In the notes, the likelihood for agent i is then given as

$$\ell_i(\theta, \gamma) = \sum_{t=1}^{T_i} \ln P_a(\theta|x_{it}) + \sum_{t=1}^{T_i-1} \ln f_x(x_{it+1}, \gamma|x_{it}, a_{it}) + \ln Pr(x_{i1}, \gamma) \quad (10)$$

This is very bad notation, since these probabilities should be conditional on the parameters, not the other way around. An alternative notation (similar to that of Aguirregabiria and Mira, 2010 and more representative of the underlying objects we're working with) is

$$\ell_i(\theta, \gamma) = \sum_{t=1}^{T_i} \ln P(a_{it}|x_{it}; \theta, \gamma) + \sum_{t=1}^{T_i-1} \ln f_x(x_{it+1}|x_{it}, a_{it}; \gamma) + \ln Pr(x_{i1}; \theta, \gamma) \quad (11)$$

As for how we get to this:

1. Start with $Pr(a_{iT_i}, x_{iT_i}, \dots, a_{i1}, x_{i1})$, which can be factorised into T_i terms with the form $Pr(a_{it}, x_{it}|\tilde{a}_{it-1}, \tilde{x}_{it-1})$, where \tilde{a}_{it-1} denotes the history of a_{it} up to $t-1$;
2. Diving joint likelihood in conditional and marginal, we have

$$Pr(a_{it}, x_{it}|\tilde{a}_{it-1}, \tilde{x}_{it-1}) = Pr(a_{it}|x_{it}, \tilde{a}_{it-1}, \tilde{x}_{it-1})Pr(x_{it}|\tilde{a}_{it-1}, \tilde{x}_{it-1})$$

3. By the Markovian structure, $Pr(x_{it}|\tilde{a}_{it-1}, \tilde{x}_{it-1}) = Pr(x_{it}|x_{it-1}, a_{it-1})$; the additively separable utility structure implies $Pr(a_{it}|x_{it}, \tilde{a}_{it-1}, \tilde{x}_{it-1}) = Pr(a_{it}|x_{it})$.
- We can also get consistent but inefficient estimates by omitting $\ln Pr(x_{i1}, \theta, \gamma)$, the initial condition. This gives us the conditional likelihood.

Estimation

- If we knew the value function, estimation would be exactly the same as in the static case. Trouble is we don't know it.
- Full solution method involves solving explicitly for the value function, either by backward induction (as implemented by Rust, 1987) or by value function iteration (i.e. iteration on (19)) or policy function iteration (see Aguirregabiria and Mira) to obtain $V(x)$, $\forall x \in \mathcal{X}$.
- Alternatively, we can implement one of the three inversion-based methods (inversion shown to hold by Hotz and Miller, 1993). The one we focus on is a two-step procedure wherein we **first estimate the choice probabilities** and **invert them to get estimates of the $v(x)$'s**, which are then used in the second step (as if they were data) to estimate the parameters of the model.² This is repeated until convergence is achieved.

²Note the similarities with the static case, in which δ is first estimated, then treated as data to obtain ξ , and conditions for ξ are finally used to identify model parameters.

2 Application: Optimal Stopping, Rust (1987)

2.1 Setup

- We ignore the subscript i for simplicity – the observations are buses.
- Agent has a product of age $x \in \mathcal{X}$
- Action set is $\mathcal{A} = (E, K)$, where E is “Exit”/scrap, and K is “Keep”.
- Product age is obviously deterministic, so no need to estimate transition probabilities: $x_t = \min(x_{t-1} + 1, X)$ where X is the maximum age of any bus (assumed to be something like 15 years – turns this into a finite horizon problem).
- Define the value function $V(x_{t+1})$ as $\sum_{x_{t+1}} V(x_{t+1})f_K(x_{t+1}|x_t)$.
- Period t flow payoff:

$$u_a(x_t) + \epsilon_{at} = \begin{cases} \theta_0 + \epsilon_{Et} & \text{if } a = E \\ \theta_1 - \theta_2 x_t + \epsilon_{Kt} & \text{if } a = K. \end{cases} \quad (12)$$

with ϵ_{at} iid Type-1 EV over a and t .

2.2 Solving and Estimating a Finite Horizon Version of the Model

- Solution method: Backward induction (only possible with finite number of periods).
- Let $t \in \{1, 2\}$. Product age $x < X$ in period $t = 1$.

In Period $t = 2$

- We only get to this stage if $a = K$ is period 1. Age is now $x + 1$.
- CCPs in period 2 given state $x + 1$ are the same as in a static model, since there is no continuation value in the final period:

$$P(a|x+1; \theta) = \begin{cases} \frac{1}{1+\exp(\theta_1 - \theta_2(x+1))} & \text{if } a = E \\ \frac{\exp(\theta_1 - \theta_2(x+1))}{1+\exp(\theta_1 - \theta_2(x+1))} & \text{if } a = K \end{cases} \quad (13)$$

where θ_0 is normalised to zero.

In Period $t = 1$

- In period $t = 1$, the period $t = 2$ shocks $\epsilon_2 \equiv (\epsilon_{E2}, \epsilon_{K2})$ are unknown. Hence the continuation value is

$$V(x+1) = E_{\epsilon_2} \max\{\epsilon_{E2}, \theta_1 - \theta_2(x+1) + \epsilon_{K2}\} \quad (14)$$

$$= \ln[1 + \exp(\theta_1 - \theta_2(x+1))] \quad (15)$$

by the usual log-sum result.

- Choice-specific value function (note that the state indicates the timing of the value function):

$$v_a(x) + \epsilon_{a1} = \begin{cases} \epsilon_{E1} & \text{if } a = E \\ \theta_1 - \theta_2 x + \epsilon_{K1} + \beta V(x+1) & \text{if } a = K \end{cases} \quad (16)$$

- CCPs in period $t = 1$ also have closed form:

$$P(a|x, \theta) = \begin{cases} \frac{1}{1 + \exp(\theta_1 - \theta_2 x + \beta V(x+1))} & \text{if } a = E \\ \frac{\exp(\theta_1 - \theta_2 x + \beta V(x+1))}{1 + \exp(\theta_1 - \theta_2 x + \beta V(x+1))} & \text{if } a = K \end{cases} \quad (17)$$

where of course $V(x+1) = \ln[1 + \exp(\theta_1 - \theta_2(x+1))]$. This problem is very easy to solve now. You just pick the path of actions that yields the highest expected payoff in period 1.

- Moreover, with the probabilities we've defined you could estimate this by MLE. The likelihood for any individual observation for which *an action is taken in both periods* is

$$\begin{aligned} \mathcal{L}(K_1, a_2, x, x+1) &= Pr(a_2, x+1, K_1, x, \theta) \\ &= Pr(a_2|x+1, K_1, x, \theta)Pr(x+1|K_1, x, \theta)Pr(K_1|x, \theta)Pr(x, \theta), \\ &= Pr(a_2|x+1; \theta)Pr(K_1|x, \theta)Pr(x, \theta), \\ &\quad (\text{Markov property \& } Pr(x+1|K_1, x; \theta) = 1) \\ \ell(\theta) &= \ln P(a_2|x+1, \theta) + \ln P(K_1|x, \theta) + \ln Pr(x, \theta) \\ &= \ln P(a_2|x+1, \theta) + \ln P(K_1|x, \theta) + \ln Pr(x, \theta) \end{aligned}$$

By similar reasoning, we can thus define the log-likelihood for any observation piecewise as

$$\ell(\theta) = \ln Pr(x, \theta) + \begin{cases} \ln P(a_2|x+1, \theta) + \ln P(K_1|x, \theta) & \text{if } a_1 = K_1 \\ \ln P(E_1|x, \theta) & \text{otherwise.} \end{cases} \quad (18)$$

2.3 Solving and Estimating by the Nested Fixed Point Algorithm (Rust, 1987)

- Although we work with finite- t data, the infinite-horizon treatment assumes an infinite process.
- We still have $x_{t+1} = \min(x_t + 1, X)$ if $a_t = K$. Given state $x_t \in \mathcal{X}$, the ex-ante value function is

$$V(x_t) = \mathbb{E}_{\epsilon_t} \max\{\epsilon_{Et}, \theta_1 - \theta_2 x_t + \epsilon_{Kt} + \beta V(x_{t+1})\} \quad (19)$$

i.e., for Type-1 EV errors,

$$V(x_t) = \ln[1 + \exp(\theta_1 - \theta_2 x_t + \epsilon_{Kt} + \beta V(x_{t+1}))], \quad \forall x_t \in \mathcal{X} = \{1, \dots, X\} \quad (20)$$

- **Estimation Algorithm - NFXP:**³

1. Guess some starting parameters θ .
2. Given θ , solve for $\{V(x)\}_{x \in \mathcal{X}}$ by value function iteration:

³The 'Nested' aspect comes from the inner loop in Step 2.

- (a) Set a feasible initial value $V(x_t)$ for all x_t (vector-valued **Bellman operator**);
 - (b) Substitute the initial value that maximises the Bellman equation into (20);
 - (c) Obtain $V(x_t)$ by maximising the value function on the RHS of (20);
 - (d) **Iterate**: put each update $V(x_t)$ from LHS into RHS until LHS no longer changes.
3. Use the $\{V(x)\}_{x \in \mathcal{X}}$ from step 2 and the factorization of $Pr(a_T, x_T, \dots, a_1, x_1, \theta)$ to set up the log-likelihood⁴ and solve for the parameters that maximize it.
 4. Repeat steps 2 and 3 until θ converges.
- Value function iteration is great because it can be shown to be a contraction mapping—hence we have guaranteed convergence—but it’s very slow, and only gets slower as the state space increases.

2.4 Solving and Estimating with Two-Step Methods: Inversion

Here we’re again using the result from Hotz and Miller (1993) which shows that any absolutely continuous probability measure admits an inverse. For the Dynamic Programming Conditional Logit (DPCL) – i.e. Rust’s (1987) model – this works in almost exact parallel with standard logit.

1. Given the logit CCPs

$$P(K|x_t, \theta) = \frac{\exp(v_K(x_t))}{1 + \exp(v_K(x_t))} \quad \text{and} \quad P(E|x_t, \theta) = \frac{1}{1 + \exp(v_K(x_t))} \quad (21)$$

by the usual inversion argument we have

$$v_K(x_t) = \ln P(K|x_t, \theta) - \ln P(E|x_t, \theta) \quad (22)$$

and, by the log-sum result, the ex-ante value function is

$$\begin{aligned} V(x_t) &\equiv E_{\epsilon_t}[\epsilon_{Et}, v_K(x_t) + \epsilon_{Kt}] \\ &= \ln[1 + \exp(\ln P(K|x_t, \theta) - \ln P(E|x_t, \theta))] \end{aligned} \quad (23)$$

2. We can then calculate the sample mean⁵ CCPs (21) for each $x \in \mathcal{X}$ as

$$\hat{P}(a|x) = \frac{\sum_t \mathbb{1}\{a_t = a, x_t = x\}}{\sum_t \mathbb{1}\{x_t = x\}} \quad (24)$$

from which we obtain sample mean choice-specific value function estimates

$$\hat{v}_K(x) = \ln \hat{P}(K|x) - \ln \hat{P}(E|x, \theta) \quad (25)$$

and mean ex-ante value function $\hat{V}(x) = \ln[1 + \exp(\hat{v}_K(x))]$; finally, we obtain

$$P(a|x_t, \theta) = \begin{cases} \frac{1}{1 + \exp(\theta_1 - \theta_2 x_t + \beta \hat{V}(x_{t+1}))} & \text{if } a = E \\ \frac{\exp(\theta_1 - \theta_2 x_t + \beta \hat{V}(x_{t+1}))}{1 + \exp(\theta_1 - \theta_2 x_t + \beta \hat{V}(x_{t+1}))} & \text{if } a = K \end{cases} \quad (26)$$

3. We estimate our parameters θ by GMM or PseudoML, both of which require data $\{a_t, x_t\}_{t=1}^T$:

⁴Recall that the value function enters the $Pr(\cdot)$, which is why we need to calculate it.

⁵Over i, t ; here just over t since we’ve omitted i for simplicity

(a) **GMM** - Hotz and Miller (1993)

Given (26), form prediction errors as $\xi_{at}(\theta) = \mathbb{1}_{\{a_t=a\}} - P(a|x_t; \theta)$ and, for instruments Z_t , moment conditions as

$$b(\theta) \equiv \frac{1}{T} \sum_t [Z_t \xi_{at}(\theta)] = 0$$

and, for some weighting matrix W , estimate θ by GMM as $\hat{\theta}^{GMM} = \arg \min_{\theta} b(\theta)^\top W b(\theta)$.

(b) **PseudoML** - Aguirregabiria and Mira (2002)

Given (26), we have sample log-likelihood as

$$\ell(\theta) = \sum_t \sum_a \mathbb{1}_{\{a_t=a\}} \ln P(a|x_t, \theta)$$

and can estimate θ as $\hat{\theta}^{PML} = \arg \min_{\theta} \ell(\theta)$. The likelihood depends on (24) and as such will suffer from measurement error bias.

The problem with inversion methods is that **they suffer from finite sample bias entering non-linearly in our moment conditions**.

2.5 Estimating with Two-Step Methods: Simulation

Inversion methods can be impractical when X is very large. Simulation method proposed by Hotz et al. (1994):

1. Set up a policy function, i.e. the choice rule that determines the **control variable** as a function of the **state variable**. It is a function of the mean choice-specific values $v_a(x)$ and errors – e.g. the simple choice rule **take action j iff** $v_j(x) + \varepsilon_{jt} \geq v_k(x) + \varepsilon_{kt}$, $\forall j \neq k$.
2. Non-parametrically estimate empirical CCPs as in (24);
3. Invert CCPs to obtain mean value function estimates as in (25);
4. **Forward Simulation:**
 - (a) Draw a vector of errors for each action from a chosen distribution for each t ;
 - (b) Obtain the action implied by the choice rule for observed x_t and drawn errors;
 - (c) Simulate x_{t+1} given the actions taken, from $t = 1$ onwards;
 - (d) Repeat until final period T ;
 - (e) Repeat for N_S matrices of simulated errors; obtain simulated *ex-ante* value functions $V^{sim}(x)$ by averaging over N_S , and $v^{sim}(x)$ by plugging $V^{sim}(x)$ in choice value functions;
5. Set up moment conditions matching empirical CCPs and simulated mean choice-specific values (which are functions of θ) and estimate θ by
 - (a) **GMM** or **PseudoML** (both as discussed above);
 - (b) **Contraction**: select the θ that best approximates the inverted CCP - in Rust's case

$$\tilde{\theta} = \arg \min_{\theta} || \ln \hat{P}_K - \ln \hat{P}_E - (\theta_0 - \theta_1 x_t + \beta V^{sim}(x_{t+1}; \theta)) || \quad (27)$$

in which simulation error enters linearly (finite sample bias still enters non-linearly as in the above.)