Lecture 22: Dynamic Discrete Choice II

ResEcon 703: Topics in Advanced Econometrics

Matt Woerman
University of Massachusetts Amherst

Agenda

Last time

- Dynamic Discrete Choice

Today

- Rust (1987)

Upcoming

- Reading for next time
  - Train textbook, Chapter 13
  - Nevo and Whinston (2010)
- Problem sets
  - Problem Set 4 is due
- Final exam
  - More information next time

Dynamic Discrete Choice

Many optimization problems have a dynamic component

- A choice made today affects the choice set or utility of choices in the future
- When making a choice today, a decision maker considers not only the current utility of each alternative but also future effects

Example: College

- Attending college may generate less utility than working, joining the Peace Corps, hiking the Trans-Canada Trail, or whatever else an 18-year-old might want to do
- But attending college opens up future career opportunities and salaries that may not exist after these other alternatives

Dynamic models have an additional layer of complexity that often require more advanced estimation methods

Rust (1987)

An Empirical Model of Harold Zurcher

This paper develops a dynamic discrete choice model to describe the behavior of Harold Zurcher

- Superintendent of maintenance at the Madison (Wisconsin) Metropolitan Bus Company
- Responsible for determining when to replace municipal bus engines

Why do we care about Harold Zurcher and bus engine replacement?

- We don't, *per se*, but this application provides a framework to illustrate two ideas
    - Bottom-up (or microfounded) model of a dynamic investment/replacement decision
    - Nested fixed point algorithm for estimating dynamic discrete choice models

Bus Engine Replacement Decision

Each month and for each city bus, Harold Zurcher must choose between two alternatives

- Perform routine maintenance on the bus engine
- Replace the bus engine

Harold Zurcher faces a tradeoff between these two choices

- Routine engine maintenance is relatively inexpensive
- Engine replacement is more expensive but reduces the probability of a costly breakdown

Harold considers a bus engine's mileage and other factors to determine when it is time to replace the engine

Data

Data on bus mileage and maintenance

- 162 buses in the Madison Metro fleet
- Data for the period December, 1974 to May, 1985
- Monthly observations of mileage (odometer readings) for each bus
- Maintenance log with date, mileage, and repair description

Model: Variable Definitions

State variable: x_t, a bus engine's mileage in month t

- A state variable describes the state of the dynamic system
- Mileage is a continuous variable that Rust discretizes into 90 bins for tractability

Choice variable: $i_t \in \{0, 1\}$ defines bus engine replacement

$$\begin{aligned} i_t = 0 \quad &\Rightarrow \quad \text{perform routine engine maintenance in month } t \\ i_t = 1 \quad &\Rightarrow \quad \text{replace the bus engine in month } t \end{aligned}$$

Transition probability: $p(x_{t+1} \mid x_t, i_t, \theta_3)$ governs the evolution of $\{x_t\}$

$$p(x_{t+1} \mid x_t, i_t, \theta_3) = \begin{cases} \theta_3 e^{\theta_3 (x_{t+1} - x_t)} & \text{if } i_t = 0 \\ \theta_3 e^{\theta_3 x_{t+1}} & \text{if } i_t = 1 \end{cases}$$

- Probability of mileage next month conditional on mileage this month, the maintenance/replacement decision, and a parameter

Model: Utility and Cost Functions

Harold Zurcher's utility function (or the Madison Metro profit function) for each bus in the fleet is

$$u(x_t, i_t, \theta_1) = \begin{cases} -c(x_t, \theta_1) & \text{if } i_t = 0 \\ -[RC + c(0, \theta_1)] & \text{if } i_t = 1 \end{cases}$$

where

- $c(\cdot, \theta_1)$ is the cost of routine engine maintenance (including the expected cost of a breakdown)
- θ_1 is a parameter that defines the cost function
- RC is the engine replacement cost

Model: Value Function

The value function is defined as the unique solution to the Bellman Equation

$$V_\theta(x_t) = \max_{i_t}[u(x_t, i_t, \theta_1) + \beta EV_\theta(x_t, i_t)]$$

where

$$EV_\theta(x_t, i_t) = \int V_\theta(y)p(dy \mid x_t, i_t, \theta_3)$$

Approach 1: Solution

There is an optimal replacement policy that solves this Bellman Equation

$$i_t = f(x_t, \theta) = \begin{cases} 0 & \text{if } x_t \leq \gamma(\theta) \\ 1 & \text{if } x_t > \gamma(\theta) \end{cases}$$

where $\gamma(\theta)$ is the unique solution to

$$(1-\beta)RC = \int_0^{\gamma(\theta)} \left[1 - \beta e^{-\theta_3(1-\beta)y}\right] \frac{\partial c(y, \theta_1)}{\partial y} dy$$

$\gamma(\theta)$ is a constant (conditional on θ_1 and θ_3) that defines the threshold for optimal bus engine replacement

- Whenever mileage exceeds $\gamma(\theta)$, it is optimal to replace the engine

Approach 1: Problems

The data do not reflect this result that all bus engines are replaced when they cross a constant mileage threshold

- Mileage at replacement varies from 82,400 miles to 387,300 miles

We can rationalize these data by assuming Harold Zucher has additional knowledge of the state of a bus engine, ε_t, that he uses to determine if a bus engine should be replaced

$$i_t = f(x_t, \theta) + \varepsilon_t$$

This equation reflects a "reduced-form" error term

- This error term is internally inconsistent with our structural model
- The error term should enter into Harold Zucher's optimization, not simply be tacked on at the end

Approach 2: Model

The second approach to solving this model adds the error term into the Bellman Equation

$$V_\theta(x_t, \varepsilon_t) = \max_{i_t}[u(x_t, i_t, \theta_1) + \varepsilon_t(i_t) + \beta EV_\theta(x_t, \varepsilon_t, i_t)]$$

where

$$EV_\theta(x_t, \varepsilon_t, i_t) = \int_y \int_\eta V_\theta(y, \eta) p(dy, d\eta \mid x_t, \varepsilon_t, i_t, \theta_2, \theta_3)$$

- $p(x_{t+1}, \varepsilon_{t+1} \mid x_t, \varepsilon_t, i_t, \theta_2, \theta_3)$ is the transition density for (x_t, ε_t)
- θ_2 is a new parameter that governs the joint evolution of x_t and ε_t

ε_t is an unobserved state variable, which adds complexity to estimation

Conditional Independence Assumption

The transition density of the controlled process $\{x_t, \varepsilon_t\}$ factors as

$$p(x_{t+1}, \varepsilon_{t+1} \mid x_t, \varepsilon_t, i_t, \theta_2, \theta_3) = q(\varepsilon_{t+1} \mid x_{t+1}, \theta_2) p(x_{t+1} \mid x_t, i_t, \theta_3)$$

Any correlation between ε_t and ε_{t+1} can be captured by x_{t+1}

- The $\{\varepsilon_t\}$ is essentially noise superimposed on the $\{x_t\}$ process

We can further simplify the problem by assuming ε_t is i.i.d. extreme value

- Any additional information known to Harold Zurcher (but unobserved by us) can be treated as random noise

These assumptions greatly reduce the curse of dimensionality and yield closed-form expressions for choice probabilities

Theorems 1 and 2

The choice probability of the bus engine replacement decision is

$$P(i_t \mid x_t, \theta) = \frac{e^{u(x_t, i_t, \theta_1) + \beta EV_\theta(x_t, i_t)}}{\sum_{j_t \in \{0,1\}} e^{u(x_t, j_t, \theta_1) + \beta EV_\theta(x_t, j_t)}}$$

where EV_θ is the unique solution to

$$EV_\theta(x_t, i_t) = \int \ln \left[\sum_{j_t \in \{0,1\}} e^{u(y, j_t, \theta_1) + \beta EV_\theta(y, j_t)} \right] p(dy \mid x_t, i_t, \theta_3)$$

The likelihood function is given by

$$\ell^f(x_1, \ldots, x_T, i_1, \ldots, i_T \mid x_0, i_0, \theta) = \prod_{t=1}^{T} P(i_t \mid x_t, \theta) p(x_t \mid x_{t-1}, i_{t-1}, \theta_3)$$

Estimating the Model

This model is estimated by maximum likelihood using a nested fixed point algorithm

- The likelihood function depends on choice probabilities
- Choice probabilities depend on the value function, which is unknown
- Conditional on a set of parameters, a nested fixed point algorithm solves for the value function

Two loops for estimation

- Outer loop find the values of θ that maximize the likelihood function
- Within each iteration of the outer loop, an inner loop uses a nested fixed point algorithm to find the value function EV_θ

Value Function Fixed Point

We need to find the value function for a set of parameters

$$EV_\theta(x_t, i_t) = \int \ln \left[\sum_{j_t \in \{0,1\}} e^{u(y,j_t,\theta_1)+\beta EV_\theta(y,j_t)} \right] p(dy \mid x_t, i_t, \theta_3)$$

- Rust discretizes x_t into 90 mileage bins, so we want to find a value of EV_θ for each of those bins

EV_θ is also on the right-hand side of this equation, so we want to find the fixed point

- We want to find the EV_θ function that, when put into the right-hand side of the equation, is also what is returned on the left-hand side of the equation

Nested Fixed Point Algorithm

Three steps to estimate this model

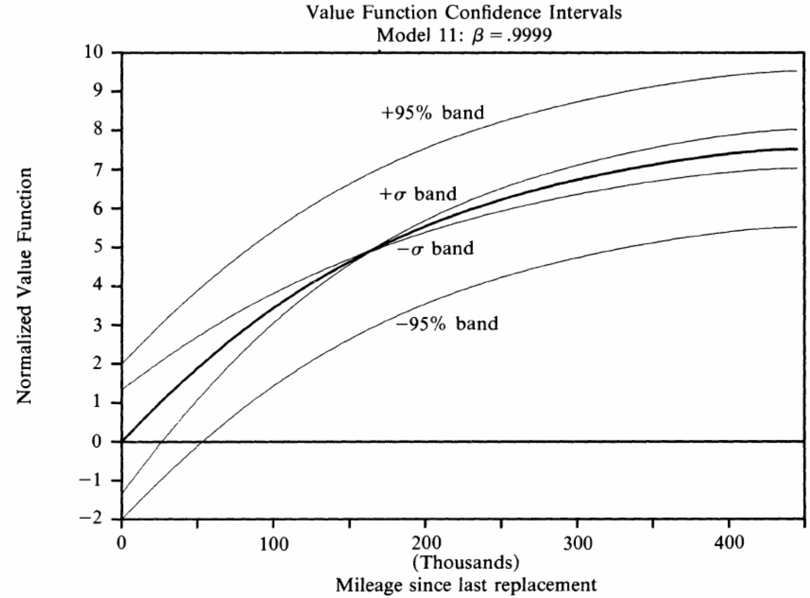
1. Impute a value of the discount factor, β
2. Estimate θ_3, the parameters of the transition probabilities, which can be done without the behavioral model of Harold Zurcher
3. Outer loop: search over (θ_1, RC) to maximize the likelihood function
    1. Inner loop: find the fixed point of the value function conditional on $(\beta, \theta_1, \theta_3, RC)$
    2. Use this value function to calculate choice probabilities and then likelihood

Structural Estimates

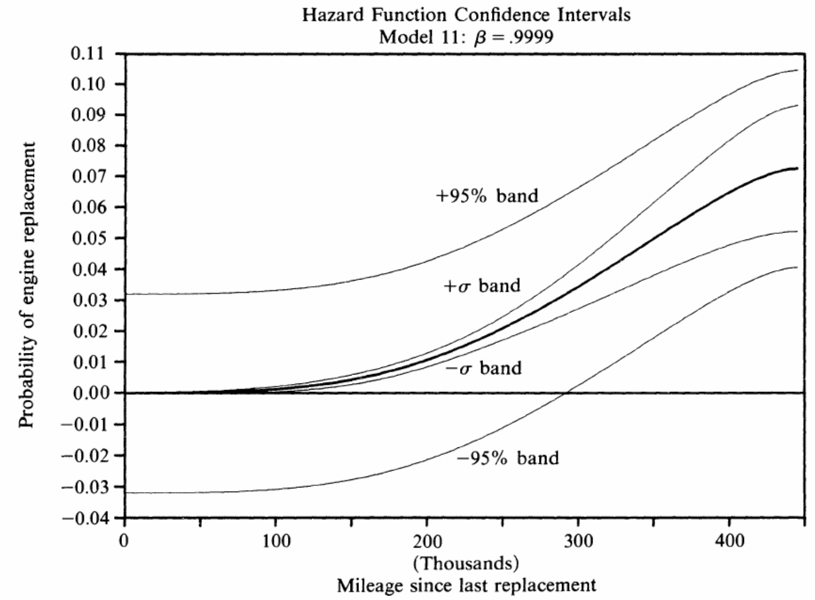
TABLE IX
STRUCTURAL ESTIMATES FOR COST FUNCTION $c(x, \theta_1) = .001\theta_{11}x$
FIXED POINT DIMENSION = 90
(Standard errors in parentheses)

Parameter		Data Sample			Heterogeneity Test	
Discount Factor	Estimates/ Log-Likelihood	Groups 1, 2, 3 3864 Observations	Group 4 4292 Observations	Groups 1, 2, 3, 4 8156 Observations	LR Statistic ($df = 4$)	Marginal Significance Level
$\beta = .9999$	RC	11.7270 (2.602)	10.0750 (1.582)	9.7558 (1.227)	85.46	1.2E−17
	θ_{11}	4.8259 (1.792)	2.2930 (0.639)	2.6275 (0.618)		
	θ_{30}	.3010 (.0074)	.3919 (.0075)	.3489 (.0052)		
	θ_{31}	.6884 (.0075)	.5953 (.0075)	.6394 (.0053)		
	LL	−2708.366	−3304.155	−6055.250		
$\beta = 0$	RC	8.2985 (1.0417)	7.6358 (0.7197)	7.3055 (0.5067)	89.73	1.5E−18
	θ_{11}	109.9031 (26.163)	71.5133 (13.778)	70.2769 (10.750)		
	θ_{30}	.3010 (.0074)	.3919 (.0075)	.3488 (.0052)		
	θ_{31}	.6884 (.0075)	.5953 (.0075)	.6394 (.0053)		
	LL	−2710.746	−3306.028	−6061.641		
Myopia test:	LR Statistic ($df = 1$)	4.760	3.746	12.782		
$\beta = 0$ vs. $\beta = .9999$	Marginal Significance Level	0.0292	0.0529	0.0035		

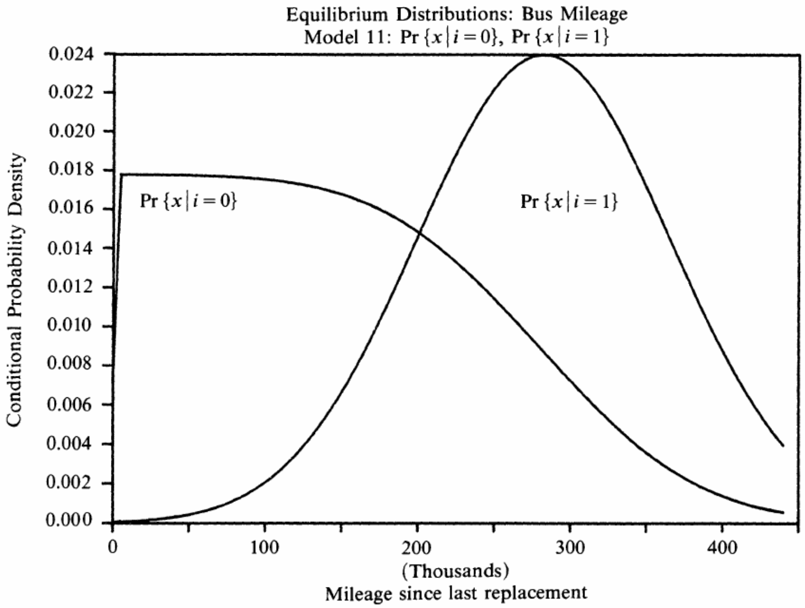
Value Function



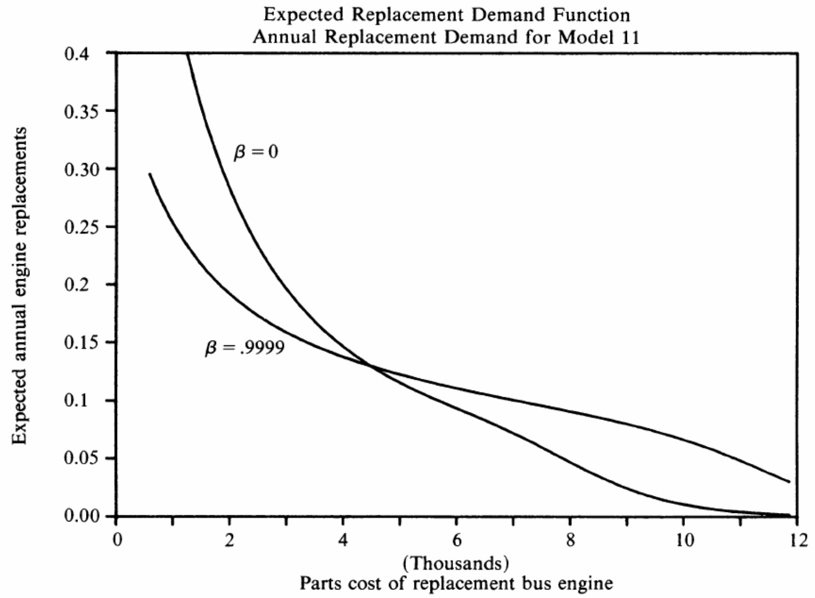
Probability of Engine Replacement



Mileage Distributions



Engine Replacement Demand



Announcements

Reading for next time

- Train textbook, Chapter 13
- Nevo and Whinston (2010)

Upcoming

- Final exam information next time