14.273 Industrial Organization: Pset4

Dave Holtz, Jeremy Yang

May 18, 2017

1. Model setup.

   Following the notations in Rust (1987), HZ's flow utility is:

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

   where RC is the replacement cost, x_t is the observed state variable for mileage, $c(\cdot)$ is cost function and i_t is the decision to replace engine and $\epsilon_t(\cdot)$ is action specific and type I extreme value distributed structural error (or unobserved state variable).

   The state transition probability is given by:

$$\theta_{3j} = \mathbb{P}(x_{t+1} = x_t + j | x_t, i_t = 0)$$

   $j \in \{0, 1, 2\}$ and if $i_t = 1$ then $x_{t+1} = 0$ with probability 1.

   HZ chooses i_t in every period t to maximize an infinite sum of discounted flow utilities. The maximal value is defined as the value function (suppress the dependency on θ_1, θ_3):

$$V(x_1, \epsilon_1) := \max_{i_t, t \in \{1,2,..\}} \mathbb{E}[\sum_{t=1}^{\infty} \beta^{t-1}(u(x_t, i_t, \theta_1) + \epsilon_t(i_t))]$$

   Rewrite the value function as in the Bellman optimality form:

$$V(x_t, \epsilon_t) = \max_{i_t} \ (u(x_t, i_t, \theta_1) + \epsilon_t(i_t)) + \beta \mathbb{E}[V(x_{t+1}, \epsilon_{t+1}) | x_t, i_t]$$

   where the expectation is with respect to (conditional) state transition probability of both x and ϵ, see Rust (1987) equation (4.5). The Bellman equation breaks the dynamic optimization problem into an infinite series of static choices.

2. (1) The choice specific value function can be derived by plugging a specific action into the value function:

$$\tilde{V}(x_t, \epsilon_t, i_t) = \begin{cases} -RC - c(0, \theta_1) + \epsilon_t(1) + \beta\mathbb{E}[V(x_{t+1}, \epsilon_{t+1})|x_t, i_t = 1] \\ -c(x_t, \theta_1) + \epsilon_t(0) + \beta\mathbb{E}[V(x_{t+1}, \epsilon_{t+1})|x_t, i_t = 0] \end{cases}$$

$$V(x_t, \epsilon_t) = \max\{\tilde{V}(x_t, \epsilon_t, 1), \tilde{V}(x_t, \epsilon_t, 0)\}$$

HZ's decision is about trading off the total (future) cost of maintaining an old engine and the lump sum cost of replacing to a new one. The time to replace is the stopping time in this problem, so it can be thought as an optimal stopping time problem where the optimal policy is characterized by a cutoff in x, HZ would choose to replace the engine if x is above that threshold (the threshold depends on realized value of ϵ).

(2) It's clear from 2 (1) that the optimal stopping rule is:

$$\begin{aligned} &-RC - c(0, \theta_1) + \epsilon_t(1) + \beta\mathbb{E}[V(x_{t+1}, \epsilon_{t+1})|x_t, i_t = 1] > \\ &-c(x_t, \theta_1) + \epsilon_t(0) + \beta\mathbb{E}[V(x_{t+1}, \epsilon_{t+1})|x_t, i_t = 0] \end{aligned}$$

or,

$$\tilde{V}(x_t, \epsilon_t, 1) > \tilde{V}(x_t, \epsilon_t, 0)$$

therefore, because the errors are type I extreme value distributed:

$$\mathbb{P}(i_t = 1|x_t) = \frac{\exp(u(x_t, 1, \theta_1) + \beta\mathbb{E}[V_{t+1}|x_t, i_t = 1])}{\sum_{k=\{0,1\}} \exp(u(x_t, k, \theta_1) + \beta\mathbb{E}[V_{t+1}|x_t, i_t = k]} \tag{2.1}$$

where $u(x_t, i_t, \theta_1)$ is defined in 1 and for convenience:

$$V_{t+1} := V(x_{t+1}, \epsilon_{t+1})$$

(3) For discrete x, under the assumption that the errors are type I extreme value distributed, we have (Rust (1987) equation (4.14)):

$$EV(x, i) = \sum_y \log\{\sum_j \exp[u(y, j) + \beta EV(y, j)]\} \cdot p(y|x, i) \tag{2.2}$$

where

$$EV(x, i) := \mathbb{E}[V_{t+1}|x_t, i_t]$$

and x, i are the state and choice of current period and y, j are the state and choice of the next period. Also note that here the transition probability does not depend on x_t but only on j (or Δx). To compute expected value function, we first need to estimate transition probability from the data, this can be done simply by counting:

$$\hat{\theta}_{30} = \frac{\sum_b \sum_t 1_{\{x_{bt+1} - x_{bt} = 0, i_{bt} = 0\}}}{\sum_b \sum_t 1_{\{i_{bt} = 0\}}}$$

$$\hat{\theta}_{31} = \frac{\sum_b \sum_t 1_{\{x_{bt+1} - x_{bt} = 1, i_{bt} = 0\}}}{\sum_b \sum_t 1_{\{i_{bt} = 0\}}}$$

$$\hat{\theta}_{32} = \frac{\sum_b \sum_t 1_{\{x_{bt+1} - x_{bt} = 2, i_{bt} = 0\}}}{\sum_b \sum_t 1_{\{i_{bt} = 0\}}}$$

we compute the expected value function in the inner loop of the nested fixed point algorithm (holding the value of θ fixed), we first guess the initial values of $EV(x, i)$ for all possible values of x, i and use the equation (2.2) to iterate expected value function until it converges. The criterion is:

$$\max_{x,i} |EV^{T+1}(x, i) - EV^T(x, i)| < \eta$$

The plot for $x = 1 - 30$ at the true value of parameters are shown in Figure 1.

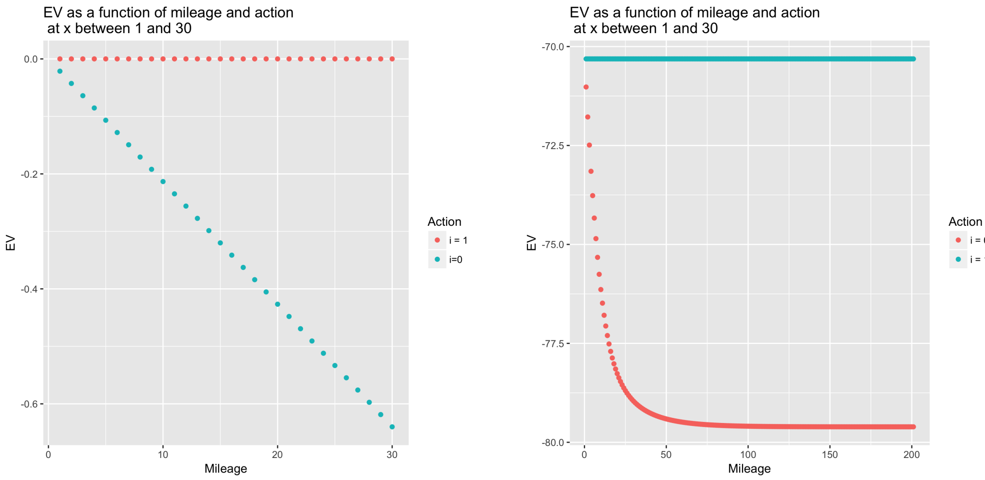

Figure 1: Expected Value Function for $i = 0$ and $i = 1$. Left panel shows results using iterative method, right panel shows provided Rust results.

Interestingly, our EV results are linear in mileage, which is probably not expected. Despite a good amount of debugging, we have been unable to identify a problem. However, it's also unclear how our calculated results for *just* 30 states should compare to the provided EV results, which provide information on 200 states. The first 30 states of the provided Rust EV estimates are decreasing in approximately linear fashion, suggesting our estimates might not be *so* bad. However, the order of magnitude of our EV values (e.g., 10^{-1}) is much smaller than the order of magnitude of EV values in the provided dataset (e.g., ~ 70), suggesting something is probably wrong. However, we don't have any more time to debug this, so we simply moved on.

(4) The provided dataset contains mileage and engine replacement information for 100 buses over 1,000 periods. The table below shows the mean mileage, maximum mileage, minimum mileage, standard deviation of the mileage, the average mileage at engine replacement across all buses and periods, and the

average number of engine replacements for a particular bus over the 1,000 periods.

avg miles	max miles	min miles	s.d. miles	avg replace miles	avg replacements
8.245	33.000	0.000	5.709	15.953	52.980

We might also be interested in understanding how each of these summary statistics vary across buses. For instance, maybe some buses have their engines replaced much more often. In order to study this, Figure 2 shows the distributions of average mileage, maximum mileage, s.d. mileage, avg miles at replacement, and number of replacements across the 100 buses in the sample. In general, these distributions are quite concentrated, suggesting that there are not systematic differences across buses.

3. (1) In the outer loop we search over a grid of values for (θ_1, β, RC), and compute the log likelihood function:

$$\log L = \sum_b \{\sum_t \log\mathbb{P}(i_{bt}|x_{bt}) + \sum_t \log\mathbb{P}(x_{bt}|x_{bt-1}, i_{t-1})\}$$

where b indexes for bus and t indexes for time period. We compute a log likelihood for each combination of values for (θ_1, β, RC) and choose the one that has the maximal value as our maximum likelihood estimation.

(2) In Hotz-Miller's approach, we will estimate the choice specific value function (as opposed to the expected value function as in Rust). We start by noting that conditional choice probability is observed directly from the data:

$$\hat{\mathbb{P}}(i=1|x) = \frac{\sum_b \sum_t 1_{\{i_{bt}=1, x_{bt}=x\}}}{\sum_b \sum_t 1_{\{x_{bt}=x\}}}$$

The choice-specific value function (minus the structural error, and suppressing the dependency on θ_1, θ_3) can be presented recursively in the following form:

$$\tilde{V}(x_t, i_t) = u(x_t, i_t) + \beta\mathbb{E}_{x_{t+1}}[\mathbb{E}_{i_{t+1}}[\mathbb{E}_{\epsilon_{t+1}}[u(x_{t+1}, i_{t+1}) + \epsilon_{t+1} + \beta(\cdots)|i_{t+1}, x_{t+1}]|x_{t+1}]|x_t, i_t]$$

where $(\cdots)$ represents higher (two and above) period forward expectations. In principle it's an infinite loop but in practice we need to stop at some T, for example, when $T=2$, $(\cdots)$ simplifies to:

$$(\cdots) = \mathbb{E}_{x_{t+2}}[\mathbb{E}_{i_{t+2}}[\mathbb{E}_{\epsilon_{t+2}}[u(x_{t+2}, i_{t+2}) + \epsilon_{t+2}|i_{t+2}, x_{t+2}]|x_{t+2}]|x_{t+1}, i_{t+1}]$$

For simplicity, in the code we use one-period forward simulation where:

$$\tilde{V}(x_t, i_t) = u(x_t, i_t) + \beta\mathbb{E}_{x_{t+1}}[\mathbb{E}_{i_{t+1}}[\mathbb{E}_{\epsilon_{t+1}}[u(x_{t+1}, i_{t+1}) + \epsilon_{t+1}|i_{t+1}, x_{t+1}]|x_{t+1}]|x_t, i_t]$$

it is estimated as:

$$\hat{\tilde{V}}(x_t, i_t) = \frac{1}{S}\sum_s [u(x_t, i_t) + \beta[u(x^s_{t+1}, i^s_{t+1}) + \gamma - \log(\hat{\mathbb{P}}(i^s_{t+1}|x^s_{t+1}))]]$$

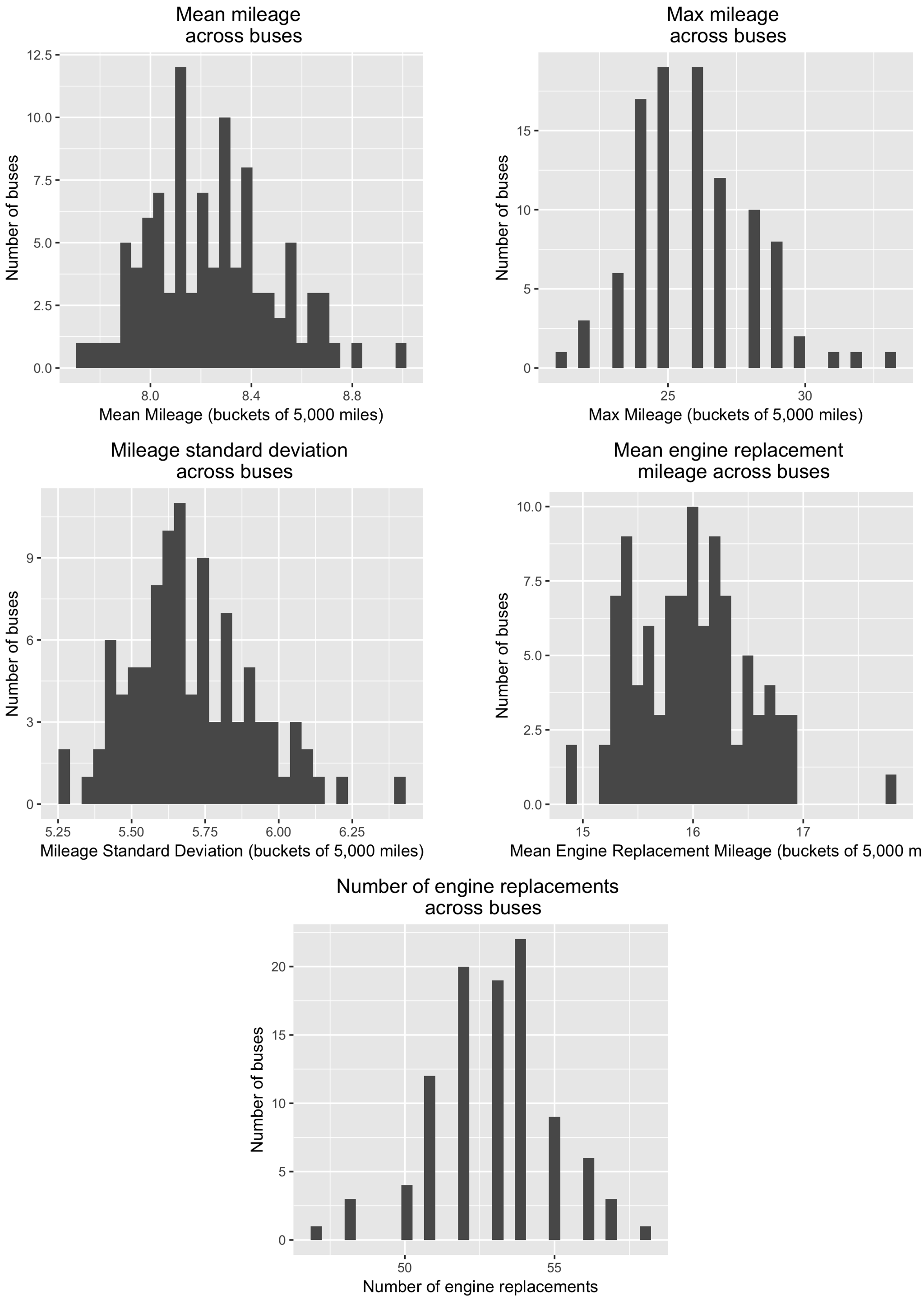

Figure 2: Distribution of various summary statistics across buses.

where x^s_{t+1} is drawn from the transition probability $\hat{\theta}_{30}, \hat{\theta}_{31}, \hat{\theta}_{32}$, and i^s_{t+1} is drawn from $\hat{\mathbb{P}}(i|x)$, γ is the Euler's constant.