Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher:

John Rust

- "Top down" approach to investment: estimate investment function that is based on variables that are aggregated within industries, such as aggregate capital, aggregate investment.

- "Bottoms up" approach to investment: estimate a disaggregated investment behavior.

- "Structural estimation": estimate the structural parameters of a model: parameters of the utility function, cost function, etc.

- Harold Zurcher: superintendent of maintenance at the Madison Metropolitan Bus company.

- Makes engine replacement decision.

- Model: he tries to minimize the overall cost. He compares the current replacement cost with the future maintenance cost, cost of consumer disutility from sudden engine failure.

- Data: monthly data over 10 years on Mileage and engine replacement records for a subsample of 104 buses.

Model:

- State variable:
  x_t: mileage of a bus
  i_t: replacement record of a bus: 1 if replace, 0 if otherwise.

- $c(x_t, \theta_1)$: operating costs
  These cost items are not separately identifiable, since there are no cost data.

$m(x_1, \theta_{11})$: Maintenance, fuel and insurance costs.
$\mu(x, \theta_{12})b(x, \theta_{13})$: Potential cost of engine breakdown.

$$c(x_t, \theta_1) = m(x_1, \theta_{11}) + \mu(x, \theta_{12})b(x, \theta_{13})$$

Per period deterministic utility function:

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

RC: replacement cost.

Per period utility function:

$$u(x_t, i_t, \theta_1) + \epsilon_t(i_t)$$

Transition probability function:

- If not replaced, ($i_t = 0$) mileage piles up.
- If replaced, ($i_t = 1$) mileage goes back to 0.

$$p(x_{t+1} \mid x_t, i_t, \theta_2) = \begin{cases} \theta_2 exp[\theta_2(x_{t+1} - x_t)] & \textit{if } i_t = 0 \textit{and } x_{t+1} \geq x_t \\ \theta_2 exp[\theta_2 x_{t+1}] & \textit{if } i_t = 1 \textit{and } x_{t+1} \geq 0 \\ 0 & \textit{otherwise} \end{cases}$$

Harold Zurcher maximized the following utility function. Discounted sum of present and future utility functions.

$$\begin{aligned} V(x_t, \epsilon_t, \theta) &= max \{ u(x_t, i_t, \theta_1) + \epsilon_t(i_t) \\ &\quad + E\left[\sum_{j=t+1}^{\infty} \beta^{j-t} \left[u(x_j, i_j, \theta_1) + \epsilon_j(i_j) \right] \mid x_t \right] \Bigg\} \end{aligned}$$

where the choice is over whether to replace or not, i.e.

$$\left\{i_t, i_{t+1}, ...\right\}$$

One can express the above problem as the following dynamic programming format.

$$\begin{aligned} V(x_t, \epsilon_t, \theta) \;&=\; max_{i_t}\{u(x_t, i_t, \theta_1) + \epsilon_t(i_t) \\ &\quad + \beta \int_{\epsilon', y} V(y, \epsilon')q(d\epsilon')p(dy \mid x_t, i_t, \theta_2)\Big\} \end{aligned}$$

How to solve for the above optimization problem. Suppose for now, θ_1 is given.

1. Last period: iteration 1: For every x and every ϵ, calculate

$$V^{(1)}(x, \epsilon, \theta) = max \{u(x, i, \theta_1) + \epsilon(i);\ i \in \{r, mr\}\}$$

2. One period before the last period: iteration 2. For every x and every $i \in \{r, mr\}$, derive the integral

$$\int_{\epsilon', y} V^{(1)}(y, \epsilon') q(d\epsilon') p(dy \mid x, i, \theta_2)$$

Then, for every x, ϵ, derive the value function as follows.

$$\begin{aligned} V^{(2)}(x,\epsilon,\theta) &= max_{x;i\in\{r,nr\}}\{u(x,i,\theta_1)+\epsilon(i) \\ &\quad + \beta\int_{\epsilon',y} V^{(1)}(y,\epsilon')q(d\epsilon')p(dy \mid x,i,\theta_2)\Big\} \end{aligned}$$

3. n periods before the last period: iteration $n+1$. Suppose as above, value function of iteration has been computed. That is, for every x and ϵ, $V^{(n)}(x,\epsilon,\theta)$ is known. For every x and every $i \in \{r, mr\}$, derive the integral

$$\int_{\epsilon',y} V^{(n)}(y,\epsilon')q(d\epsilon')p(dy \mid x,i,\theta_2)$$

Then, for every x, ϵ, derive the value function as follows.

$$\begin{aligned} V^{(n+1)}(x,\epsilon,\theta) &= max_{x;i\in\{r,nr\}}\{u(x,i,\theta_1)+\epsilon(i) \\ &\quad + \beta\int_{\epsilon',y} V^{(n)}(y,\epsilon')q(d\epsilon')p(dy \mid x,i,\theta_2)\Big\} \end{aligned}$$

If for every x, ϵ

$$\left|V^{(n+1)}(x,\epsilon,\theta) - V^{(n)}(x,\epsilon,\theta)\right| \leq \delta$$

for a small δ (for example, $\delta = 0.000001$), then, we stop the algorithm and we let the value function as follows:

$$V(x,\epsilon,\theta) = V^{(n+1)}(x,\epsilon,\theta)$$

Likelihood Increment:

- Likelihood increment for mileage:

$$
\begin{aligned}
p(x_{t+1} \mid x_t, i_t, \theta_2) &= \theta_2 exp\left\{\theta_2(x_{t+1} - x_t)\right\} & \textit{if } i_t = 0 \\
p(x_{t+1} \mid x_t, i_t, \theta_2) &= \theta_2 exp\left\{\theta_2 x_{t+1}\right\} & \textit{if } i_t = 1
\end{aligned}
$$

- Engine replacement probability: Replace if

$$
\begin{aligned}
& u(x, r, \theta_1) + \beta V(y, \epsilon')q(d\epsilon')p(dy \mid x, r, \theta_2) + \epsilon(r) \\
& \geq u(x, nr, \theta_1) + \beta V(y, \epsilon')q(d\epsilon')p(dy \mid x, nr, \theta_2) + \epsilon(nr)
\end{aligned}
$$

That is, the replacement probability is

$$
\begin{aligned}
& Pr(replace|x, \theta) \\
= \; & Pr\left(\epsilon(nr) - \epsilon(r) \leq u(x, r, \theta_1) - u(x, nr, \theta_1)\right. \\
+ \; & \left.\beta V(y, \epsilon')q(d\epsilon')p(dy \mid x, r, \theta_2) - \beta V(y, \epsilon')q(d\epsilon')p(dy \mid x, nr, \theta_2)\right)
\end{aligned}
$$

$$Pr(\textit{no replace} \mid x, \theta) = 1 - Pr(\textit{replace} \mid x, \theta)$$

Log likelihood increment for bus i, time t: data is x_{it} ,i_{it}

$$\begin{aligned} l_{it} &= logp(x_{i,t+1} \mid x_{it}, i_{it}, \theta_2) \\ &\quad + logP(r|x_{it}, \theta) i_{it} + log[1 - P(r \mid x_{it}, \theta)](1 - i_{it}) \end{aligned}$$

Log likelihood:

$$\sum_{i=1}^{N} \sum_{t=1}^{T} l_{it}$$

Difficulty with Estimation

- Value function has to be solved at each state space point x, ϵ.

- While doing that, $\int_{\epsilon', y} V^{(n)}(y, \epsilon') q(d\epsilon') p(dy \mid x, i, \theta_2)$ has to be integrated with respect to y and ϵ.

- For each bus, at each period t

$$
\begin{aligned}
& Pr(replace|x, \theta) \\
= \; & Pr\left(\epsilon(nr) - \epsilon(r) \le u(x, r, \theta_1) - u(x, nr, \theta_1)\right. \\
+ \; & \left.\beta V(y, \epsilon') q(d\epsilon') p(dy \mid x, r, \theta_2) - \beta V(y, \epsilon') q(d\epsilon') p(dy \mid x, nr, \theta_2)\right)
\end{aligned}
$$

- has to be computed, which is an integral over $\epsilon(nr) - \epsilon(r)$.

- The above computation has to be repeated for every candidate θ that comes up during the ML routine.

Simplifications of the Solution/Estimation Algorithm

Assume $\epsilon(r)$,$\epsilon(nr)$ are i.i.d. extreme value distributed. Then, let

$$\overline{V}^{(n)}(x, i, \theta) = u(x, i, \theta_1) + \beta \int_{\epsilon', y} V^{(n-1)}(y, \epsilon') q(d\epsilon') p(dy \mid x_t, i_t, \theta_2)$$

That is, $\overline{V}^{(n)}(x, i, \theta)$ is the deterministic value of choice i. Then,

$$\int_{\epsilon'} V(x, \epsilon', \theta) q(d\epsilon') = log \left[\sum_i exp \left(\overline{V}(x, i, \theta) \right) \right]$$

Hence, the integration over ϵ has the closed form.
Also,

$$Pr\left(r \mid x, \theta\right) = \frac{exp\left[\overline{V}(x, r, \theta)\right]}{exp\left[\overline{V}(x, r, \theta)\right] + exp\left[\overline{V}(x, nr, \theta)\right]}$$

Do not calculate the expected value function for every x. Only calculate it for a finite grid points of x.

Nested fixed point algorithm for structural estimation.

- Outer algorithm: Maximum likelihood calculation Use Newton routine to seach for the parameters maximizing the likelihood.

- Inner algorithm: Dynamic Programming algorithm: Given parameter values, compute the expected value function $V(x, \epsilon, \theta)$ and the likelihood.

Estimation steps.

1) First estimate parameters of the transition probability densities of mileage: does not require solving for the value function.

2) Then, estimate the rest of the parameters: requires solving the value function.

Discretized transition probability distribution: 90 equally spaced grid points $x_0 = 0$, $x_1 = 5,000$, $x_{90} = 450,000$. $p\left(x_j \mid x = x_j\right) = \theta_{31}$, $p\left(x_{j+1} \mid x = x_j\right) = \theta_{32}$,$p\left(x' \geq x_{j+2} \mid x = x_j\right) = \theta_{33}$ and $p\left(x_{j+l} \mid x = x_j\right) = 0$ for $l < 0$.

Restriction: within group mileage is the same.

				Log likelihood	
	θ_{31}	θ_{32}	θ_{33}	Restricted	Unrestr.
1:83 Grumman	0.197	0.789	0.014	-203.99	-187.81
2:81 Chance	0.391	0.599	0.010	-138.57	-136.77
3:79 GMC	0.307	0.683	0.010	-2219.58	-2167.04
4:75 GMC	0.392	0.595	0.013	-3140.57	-3094.38
5:74 GMC (V8)	0.489	0.507	0.005	-1079.18	-1068.45
6:74 GMC (V6)	0.618	0.382	0.000	-831.05	-826.32
7:82 GMC (V8)	0.600	0.397	0.003	-1550.32	-1523.49
8:72 GMC (V6)	0.722	0.278	0.000	-1330.35	-1317.69

	Restricted	Unrestr.	L ratio	df	significanc
1,2,3	-2575.98	-2491.51	168.93	198	0.934
1,2,3,4	-5755.00	-5585.89	338.21	309	0.121
4,5	-4243.73	-4162.83	161.80	144	0.147
6,7	-2384.50	-2349.81	69.39	81	0.818
6,7,8	-3757.76	-3668.50	180.52	135	0.005
5,6,7,8	-4904.41	-4735.95	336.93	171	1.5E-17
Full sample	-11,237.68	-10,321.84	1,831.67	483	7.7E-10

Pool groups 1,2,3 and 6,7.

Parameter Estimates:

	G1,2,3	G4		LR	Sign.
			$\beta = 0.9999$		
RC	11.7270	10.0750	9.7558	85.46	1.2E-17
θ_{11}	4.8259	2.2930	2.6275		
θ_{30}	0.3010	0.3919	0.3489		
θ_{31}	0.6884	0.5953	0.6394		
LL	-2708.366	-3304.155	-6055.250		
			$\beta = 0.0$		
RC	8.2985	7.6358	7.3055	89.73	1.5E-18
θ_{11}	109.9031	71.5133	70.2769		
θ_{30}	0.3010	0.3919	0.3488		
θ_{31}	0.6884	0.5953	0.6394		
LL	-2710.746	-3306.028	-6061.641		
LL	4.760	3.746	12.782		
$Sign.$	0.0292	0.0529	0.0035		

- Myopic model is rejected.

- Homogeneity between two groups is rejected.

- In a myopic model, in order to justify frequent replacement, small replacement cost and large maintenance cost increase with mileage is necessary.

- Even though the likelihood is bigger in the non-parametric case, one cannot reject the parametric functional form.

Maintenance Cost Specification.

	G1,2,3	G4	
Linear $c(x, \theta_1) = 0.001\theta_{11}x$			
$\beta = 0.9999$	-132.389	-163.584	-300.250
$\beta = 0.0$	-134.747	-165.458	-306.641
nonparametric			
$\beta = 0.9999$	-110.832	-138.556	-261.642
$\beta = 0.0$	-110.832	-138.556	-261.642

Fundamental nonidentification problem of the dynamic discrete choice models: If you do not put any parametric restrictions to your model, you cannot know whether the individual is myopic or not.

Implied Demand for Machine Replacement.

1. Change the value of the replacement cost RC: price of replacement.

2. Given the replacement cost and other parameters that are estimated, solve for the Dynamic programming problem and derive the value function.

$$V(x, \epsilon, \theta) = V^{(n+1)}(x, \epsilon, \theta)$$

3. Derive the replacement probability.

$$\begin{aligned} & Pr(replace|x, \theta) \\ = & \; Pr\left(\epsilon(nr) - \epsilon(r) \leq u(x, r, \theta_1) - u(x, nr, \theta_1)\right. \\ + & \left. \beta V(y, \epsilon')q(d\epsilon')p(dy \mid x, r, \theta_2) - \beta V(y, \epsilon')q(d\epsilon')p(dy \mid x, nr, \theta_2)\right) \end{aligned}$$

Derive the mileage density.

$$p(x_{t+1} = x_j \mid x_t = x_i, i_t, \theta_2) = \begin{cases} \theta_{j-i} & \textit{if } i_t = 0 \textit{and } x_{t+1} \geq x_t \\ \theta_j & \textit{if } i_t = 1 \textit{and } x_{t+1} \geq 0 \\ 0 & \textit{otherwise} \end{cases}$$

They together are the transition probability of a Markov Chain from x_t to x_{t+1}.

4. Derive the stationary distribution of the Markov Chain of x:$\pi(x)$.

5. 5. Stationary replacement probability:

$$\int Pr(replace \mid x, \theta)\pi(x)dx$$

Notice that the replacement cost is not even observable. But if we solve and estimate the Dynamic Discrete choice models, we can derive the implied replacement demand function. Even though we have only two groups of buses with only two replacement cost possibilities, we can use the actual replacement data (sample for this is large) and derive the replacement demand function.

value of mean mileage given that replacement hasn't yet occurred is 159,305 which is also within half a standard deviation of the actual value of 134,862. Thus, use of a stationary distribution to compute replacement demand does not appear to be greatly at odds with the data.

By parametrically varying replacement costs, I can trace out the equilibrium distribution π_θ as a function of RC. In particular, using formula (6.3) I can compute the expected demand curve for replacement investment. Figure 7 presents the expected demand function $d(RC)$ for model 11 for a fleet containing a single bus, $M = 1$. For comparison, I also present the implied demand curve for the static model with $\beta = 0$. We can see significant differences in the predictions of the two models. As one might expect, the demand curve for the myopic model is much more sensitive to the cost of replacement bus engines, overpredicting demand at low prices, underpredicting demand at high prices. Notice, however, that the maximum likelihood procedure insures that both models generate the same predictions at the actual replacement cost of \$4343.

Figure 7 summarizes the value of the "bottom-up" approach to replacement investment. Since engine replacement costs have not varied much in the past, estimating replacement demand by a "reduced-form" approach which, for example, regresses engine replacements on replacement costs, is incapable of producing reliable estimates of the replacement demand function. In terms of Figure 7, all the data would be clustered in a small ball about the intersection of the two demand curves: obviously many different demand functions would

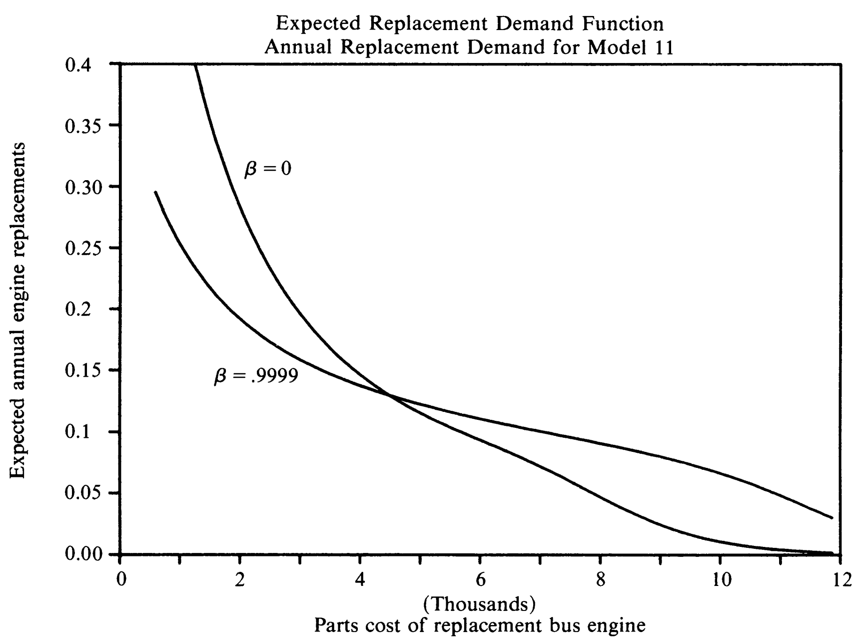

FIGURE 7