**Economics 696F: Econometrics of Dynamic Industrial Organization Models**
**Homework 1: Dynamic Single Agent Models**
**Due date: October 5, 2009**

We have included two datasets from a Rust-style optimal stopping problem for bus engines. The datasets, which are in comma-separated value format, will be e-mailed along with the assignment.

For the first dataset, the data-generating process is as follows:

- Mileage for any bus at any time can take on 11 values, $x = \{0,1,\ldots,10\}$.
- At any time period, there is a probability $\lambda$ that mileage will increase by 1, with a probability $1-\lambda$ that mileage remains constant. If mileage is already 10, then it does not increase.
- The per-period utility of having a bus engine with mileage x is $u^0(x,\theta) = -(\theta_1 x + \theta_2 x^2) + \varepsilon_0$. The per-period utility of replacing a bus engine with mileage x is $u^1(x,\theta) = -RC + \varepsilon_1$. Both the $\varepsilon$ terms are i.i.d. and distributed type 1 extreme value.
- Zurcher discounts the future at the rate of $\beta$ and chooses an optimal replacement policy given his objective function.

Since it is difficult to estimate $\beta$, use $\beta = .95$ throughout this assignment. Also, while it is reasonably straightforward to estimate $\lambda$, use $\lambda = .7$ throughout this assignment.

The dataset includes information at the level of the bus-time period. Each observation contains the following information: beginning-of-period mileage, replacement decision and end-of-period mileage.

**1) Prove cutoff rule**

Consider the above model, but without the $\varepsilon$ terms. Define a cutoff rule to mean that there exists a mileage level x such that Zurcher will always replace the bus engine if the engine has more miles than x and never replace the engine if it has less miles than x. Prove that the optimal decision for this model satisfies a cutoff rule. Explain in words why the full model does not satisfy a cutoff rule.

Together with your answers, please submit your computer code for questions 2–5.

**2) Estimation using Rust method**

a- Solve the dynamic programming problem for the model. Use $(\theta_1 = 0.3, \theta_2 = 0, RC = 4)$. Report the probability of replacement for any state.

b- Estimate the model using a maximum likelihood fixed point algorithm. Report your parameter estimates.

Let $\theta = (\theta_1, \theta_2, RC)$. You will want to perform a nonlinear search over $\theta$. For each $\theta$, perform the following process:

- Solve the dynamic programming problem
- Using the value function, derive the contingent value of each action at each state (this can be done at the same time as solving the dynamic programming problem if you want)
- Use the contingent valuation to calculate the probability of replacement at each state
- Use the probability of replacement by state to derive a likelihood

c - Calculate a variance/covariance matrix of the parameter estimates. We use the standard formula $\mathrm{Var}(\hat{\theta}) = \left[ \sum_j \frac{d \ln L_j(\theta)}{\theta} \frac{d \ln L_j(\theta)}{\theta}' \right]^{-1}$, calculating the derivatives numerically. Report standard errors of the parameter estimates.

**3) Counterfactual policies**

a- Give an example of a counterfactual policy for which it might be plausible to use the structural estimates from question 2. Explain why it is appropriate and how you would simulate the counterfactual policy.

b- Give an example of a counterfactual policy for which it is not appropriate to use the structural estimates from question 2. Explain why it is not appropriate to use these estimates. (Hint: it must be because the estimated parameters would change following the policy change.)

**4) Estimation using Aguirregabiria and Mira [AM]**

a- Generate the transition matrix: Start with the following estimate for the probability of replacement: replace with probability .3 if $x \in \{0,1,\ldots,5\}$; replace with probability .1 if $x \in \{6,\ldots,10\}$. Using these probabilities and $\lambda$, calculate a one-period transition matrix. Report the transition matrix. Use the transition matrix and a matrix inverse operator to generate the expected discounted distribution of future states as a function of current state. This is notated as $\left(I_M - \beta F^U(P)\right)^{-1}$ in AM. Report this matrix also. To save space, report only the first 5 rows and columns of both matrixes.

b- Estimate the parameters: Using the transition matrix from a-, search over the parameter space. For each parameter vector, find $V_\sigma$, the expected Bellman equation (8) by multiplying $\left(I_M - \beta F^U(P)\right)^{-1}$ as calculated in part (a) by the expected utility of each action times the probability of each action, using the estimated probabilities of replacement specified above. Remember to include the term for $E\left[\varepsilon_j \middle| \text{option } j\right]$. Now, express the probability of each action at each state using the Bellman formula in question 2a, but using $V_\sigma$ for the value next period, instead of iterating. As in question 2b, use the calculated probabilities to define a quasi-likelihood for this parameter vector. Maximize the quasi-likelihood over the parameter vector. Report your parameter estimates. (Don't worry about the standard errors for this question.)

c- Iterate on the estimator: Use the probability of each action at each state calculated at the estimated parameters from part b- to redo steps a- and b-. Repeat this process twice. Report both sets of parameter estimates. (Don't worry about the standard errors here either.)

d- True or false: the AM estimator from b- (without repetition) is a consistent but not efficient estimate of the parameter vector. Explain.

## 5) Unobserved heterogeneity

Consider now a modification of the model. There are two types of drivers, good and bad. Zurcher knows which driver is of which type. The union contract has randomly assigned a driver to each route and Zurcher cannot change the assignment. A fraction $\gamma$ of drivers are good and the remaining fraction bad. Bad drivers wear down the bus engine more quickly, so that Zurcher's utility function is now $u^{0,G}(x,\theta) = -\left(\theta_1^G x + \theta_2^G x^2\right) + \varepsilon_0$ and $u^{0,B}(x,\theta) = -\left(\theta_1^B x + \theta_2^B x^2\right) + \varepsilon_0$ for good and bad drivers respectively. Unfortunately, Zurcher can tell us neither $\gamma$ nor whether a route is driven by a good or bad driver.

Use the second dataset for this question. This dataset has the same variables as the first dataset but also contains an indicator for the route number.

Write down likelihood estimator for this model. The parameters are now $\left(\theta_1^G, \theta_1^B, \theta_0^G, \theta_0^B, RC, \gamma\right)$. Estimate the model using a nested fixed point maximum likelihood estimator. Report your parameter estimates.

## 6) Necessary conditions for AM estimator

a- Is the model in question 5 still estimable with the AM method? If not, explain why not, i.e. which assumption does it violate in the paper?

b- What if the bus drivers were randomly assigned each period to a new route? Could you now estimate the model with the AM method?