Replicating John Rust's Bus replacement engine paper

Dhananjay Ghei*

December 16, 2018

Abstract

This note replicates the estimation procedure of Rust [1987] paper in R. We have the data on 8 groups of buses. This is the same data set as used in the original paper from Rust [1987]. The paper presents the summary statistics to see if the data looks similar to the original paper and then proceeds to estimation for different groups of buses. I replicate the within and between group estimates for different bus groups and then estimate the full likelihood to get estimates for three different groups (Group 1,2,3, Group 4, and Group 1,2,3,4). Instead of using the BHHH algorithm, I use the inbuilt optimisation routines of R and the estimates are quite close to the ones in the original paper.

*Department of Economics, University of Minnesota. This replication paper was written as a part of Homework 4 for ECON 8602 - Industrial Organisation. The code for this exercise was written in R. Replication files are present on GitHub.

Contents

Table 1 Summary of replacement data

The table shows summary statistics of for the subsample of buses for which at least 1 replacement occurred. The table corresponds to the Table IIa in [Rust, 1987]

	Max	Min	Mean	Std. dev
Group 1	0	0	0	0
Group 2	0	0	0	0
Group 3	273400	124800	199733	37459
Group 4	387300	121300	261012	63929
Group 5	322500	118000	245291	60258
Group 6	237200	82400	150786	61007
Group 7	413100	170800	278752	78528
Group 8	334400	132000	194674	55463

1 Data

The data is available on Rust's website[1]. The data is taken from Rust [1987] paper. The data is available on 9 bus groups. However, in the estimation procedure, Rust [1987] uses only the maintenance records of 162 buses excluding the bus group "D309". I will drop this group as in the original paper and follow the procedure closely. I downloaded the zip file and read it in `R`.
The data is available from the period December 1974 until May 1985. The data is provided in a single column format which needs to be reshaped in the correct dimensions. Following this, the first 11 rows for each bus in each data set contains the data on the following items:

1. Bus Number
2. Month Purchased
3. Year Purchased
4. Month of 1st engine replacement
5. Year of 1st engine replacement
6. Odometer at replacement
7. Month of 2nd engine replacement
8. Year of 2nd engine replacement
9. Odometer at replacement
10. Month (Odometer data begins)
11. Year (Odometer data begins)

Subsequently, the remaining rows are the monthly mileage observations for each bus. Table 1 gives the summary statistics of the bus types included in the sample. This is consistent with Table IIa from Rust [1987] therefore, we have the correct data for estimation. One can see that a lot of these numbers are similar to the ones in the original paper thereby, guaranteeing that the data was read in correctly.

Finally, the data consists of $\{i_t, x_t\}$ where i_t^m is the engine replacement decision in month t for bus m and x_t^m is the mileage since last replacement of bus m in month t. In the next section, the goal is to estimate the parameters $\theta = (\beta, \theta_1, RC, \theta_3)$ by maximum likelihood using the nested fixed point algorithm.

[1] `https://editorialexpress.com/jrust/nfxp.html`

2 Methodology

The first step is to discretise the state space. In order to do so, I divide the mileage into equally sized ranges of length 5000 with an upper bound of 450,000 miles. Thus, $n = 90$ in this case. Following this, the state variable x will take only discrete integer values in the set $\{1, 2, 3, \ldots, n\}$. Using the discretised state variable x, the decision rule is then reduced to a simple multinomial distribution on the discretised space. In this case, I follow rust and construct an indicator for mileage which takes only 3 values $\{0, 1, 2\}$ corresponding to monthly mileage in the intervals $[0, 5000), [5000, 10,000)$ and $[10,000, \infty)$ respectively. The estimation strategy as laid out in Rust [1987] is in three stages corresponding to each of the likelihood functions l^1, l^2 and l^f, where l^f is the full likelihood function and l^1 and l^2 are "partial likelihood" functions given by:

$$l^1(x_1, \ldots, x_T, i_1, \ldots, i_T | x_0, i_0, \theta) = \Pi_{t=1}^T p(x_t | x_{t-1}, i_{t-1}, \theta_3)$$
$$l^2(x_1, \ldots, x_T, i_1, \ldots, i_T | \theta) = \Pi_{t=1}^T P(i_t | x_t, \theta)$$

The first step is to estimate $p(x)$ non-parametrically. I do this for within group and between group estimates as in the original paper. Given the discretised space, the probability π_j is then given by: $\pi_j = Pr\{x_{t+j} = x_t + j\}$ for $j = 0, 1, 2$. The non-parametric estimation then involves simply calculating the average number of times the bus was in a particular state. I calculate these averages for within group (i.e. for each bus group separately) and for between group (i.e. for combined bus groups). Standard errors are calculated based on the standard non-parametric estimation.
Once, we estimate these probabilities, we can construct a $n \times n$ transition matrix Π which is given by:

$$\Pi = \begin{bmatrix} \pi_0 & \pi_1 & \pi_2 & 0 & 0 & \ldots & 0 \\ 0 & \pi_0 & \pi_1 & \pi_2 & 0 & \ldots & 0 \\ 0 & 0 & \pi_0 & \pi_1 & \pi_2 & \ldots & 0 \\ \vdots & & & & & & \\ 0 & 0 & & \ldots & & \pi_0 & 1-\pi_1 \\ 0 & 0 & \ldots & 0 & \ldots & 0 & 1 \end{bmatrix}$$

This is the transition probability matrix for the uncontrolled Markov process. Now, we can proceed to the next step of estimation. Finally, I estimate the parameters θ_1 and RC by MLE using the nested fixed point algorithm. Set $\beta = .9999$. Instead of using the BHHH algorithm, I will use the inbuilt optimisation routines in `R`.

3 Results

Table 2 shows the within group estimates of θ_3 form 8 different bus groups. These are estimated non-parametrically by directly calculating the mean.

Table 2 Within group estimates

The table shows the within group estimates of θ_3 from 8 different groups as in Rust [1987]. Standard errors are in parantheses. These estimates match the ones in the original paper.

Parameters	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	Group 8
θ_{31}	0.197	0.391	0.307	0.392	0.489	0.618	0.6	0.722
	(0.021)	(0.035)	(0.008)	(0.007)	(0.013)	(0.014)	(0.01)	(0.009)
θ_{32}	0.789	0.599	0.683	0.595	0.507	0.382	0.397	0.278
	(0.022)	(0.035)	(0.008)	(0.007)	(0.013)	(0.014)	(0.01)	(0.009)
θ_{33}	0.014	0.01	0.01	0.013	0.005	0	0.003	0
	(0.006)	(0.007)	(0.002)	(0.002)	(0.002)	(0)	(0.001)	(0)

Table 3 shows the between group estimates of θ_3 from the same groups as taken in the paper. The estimates of probabilities match the ones in the paper. Consider, for example, the bus group 1,2,3,4 which is what we will be estimating later on. θ_{31} equals .349, θ_{32} equals .572 and θ_{33} equals .012.

Table 3 Between group estimates

The table shows the between group estimates of θ_3 from 8 different groups as in Rust [1987]. Standard errors are in parantheses. These estimates match the ones in the original paper.

Parameters	Group 1,2,3	Group 1,2,3,4	Group 4,5	Group 6,7	Group 6,7,8	Group 5,6,7,8	Full Sample
θ_{31}	0.301	0.349	0.417	0.607	0.652	0.618	0.486
	(0.007)	(0.004)	(0.006)	(0.008)	(0.006)	(0.006)	(0.003)
θ_{32}	0.688	0.639	0.572	0.392	0.347	0.38	0.507
	(0.007)	(0.004)	(0.006)	(0.008)	(0.006)	(0.006)	(0.003)
θ_{33}	0.011	0.012	0.011	0.002	0.001	0.002	0.007
	(0.002)	(0.001)	(0.001)	(0.001)	(0)	(0.001)	(0.001)

Table 4 shows the structural estimates for the full likelihood estimation where the cost function is assumed to be of linear form $c(x, \theta_1) = .001\theta_{11}x$ and β is set to .9999. These are the same specifications as used in the paper to get Table IX. Estimates for three different samples are presented as in the original paper. These samples are Group 1,2,3, Group 4, and Group 1,2,3,4. Consider for example the case of Group 1,2,3,4, my estimates are $RC = 9.792$ and $\theta_{11} = 2.6695$ which are close to the original estimates of $RC = 9.7558$ and $\theta_{11} = 2.6275$ from Table IX of Rust [1987]. The standard errors are calculated using the square root of the inverse of the Hessian matrix. The Hessian from the optimisation routine is the second-derivative of the likelihood function evaluated at the optimum. The standard errors estimated for the same group are $se(RC) = .909$ and $se(\theta_{11}) = .4781$ which are similar to the estimates from the Rust's paper of $se(RC) = 1.227$ and $se(\theta_{11}) = .618$. The difference in the standard errors comes from the fact that I am using the Fischer information matrix (Hessian) instead of the original method from the Rust paper.

Table 4 Structural Estimates for linear cost function

The table shows the structural estimates for cost function $c(x, \theta_1) = .001\theta_{11}x$. The fixed point dimension is 90 and $\beta = .9999$. Standard errors are in parantheses. This table replicates Table IX of [Rust, 1987].

	Group 1,2,3	Group 4	Group 1,2,3,4
RC	12.0328(2.006)	10.0933(1.359)	9.792(0.909)
θ_{11}	5.12(1.4359)	2.311(0.5584)	2.6695(0.4781)

4 Conclusion

In this note, I used the bus replacement data of Harlod Zurcher (which was used by Rust in his paper) to replicate the empirical model as estimated by John Rust. I feed in the data, get summary statistics to see if the data was indeed consistent with the paper. Then, I used the three step estimation as laid out in the paper to replicate the results as closely as possible. First, I estimate the within and between group estimates of the transition and then use it to estimate the full likelihood function assuming a linear cost function as in the original paper. I used the nested fixed point algorithm to estimate the results and follow as closely as possible to Rust's approach. The only difference being that I used the inbuilt optimisation algorithm instead of writing the BHHH algorithm for estimation.
Following this, I was able to replicate the results from Table IX of the paper. The point estimates are quite close to the original estimates.

5 Replication Files

The replication files are present on my Github repository. `https://www.github.com/dhananjayghei/io_estimation`. The code was written in `R`. The data set is present in `data/rust-data.zip`. The code is present in the `src/` folder. The `src/` folder has two files:

1. `rust_functions.R` - This file contains all the functions that I wrote for estimation which are called in the main file.

2. `rust.R` - This is the main file that reads in the data set from the zip file, performs estimation, and generates the LaTeXtable for results shown in the Results section.

Apart from the base packages that already come with `R`, you will need the following `R` packages to run the code: gtools, xtable

For people on Linux/Mac, there is also a `Makefile` present in the repository that you could use to run the code directly. Clone the Github repository to your local machine, open the terminal and go to the `src/` directory and run the following command: `make rust`.

References

John Rust. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, pages 999–1033, 1987.