Econ 220B PS 2

AC Watt

2021-12-03

Contents

Part a

Construct sample statistics similar to Tabel IIa and IIb on page 1003.

Table 1: Summary of Replacement Data (Subsample of buses for which at least 1 replacement occurred)

Bus Group	Mileage at Replacement				Elapsed Time (Months)				Number of Observations
	Max	Min	Mean	Standard Deviation	Max	Min	Mean	Standard Deviation	
1	0	0	0.0	0.0	0	0	0.0	0.0	0
2	0	0	0.0	0.0	0	0	0.0	0.0	0
3	273400	124800	199733.3	37459.4	73	37	58.1	10.9	27
4	387300	121300	257336.4	65477.0	115	27	72.7	23.3	33
5	322500	118000	245290.9	60257.9	126	30	84.4	29.7	11
6	237200	82400	150785.7	61006.8	126	48	73.7	35.2	7
7	331800	121000	208963.0	48980.9	103	40	67.3	16.9	27
8	297500	132000	186700.0	43956.1	103	35	57.4	22.2	19
Full Sample	387300	82400	216354.0	60475.4	126	27	67.1	22.4	124

Part b

Estimate his model using a two-step procedure. Please briefly document your estimation procedure in words. Create a table similar to the first five columns of Table IX on page 1021. If you forward simulate the value function, see Pakes Pollard (1989) for reference on how to draw errors and compute confidence intervals.

Part c

Estimate his model using a nested fixed point procedure. Create a table similar to the first five columns of Table IX on page 1021.

Appendix A: R Code

```r
rm(list=ls())
knitr::opts_chunk$set(echo = F)
# stargazer table type (html, latex, or text)
# Change to latex when outputting to PDF, html when outputting to html
table_type = "latex"
Cache = TRUE

# Packages
library(stargazer)
library(ggplot2)
library(readxl)
library(tidyverse)
library(lubridate)
library(kableExtra)
# Load data from PS2a with previous log variables
f = 'rust_data_clean.xlsx'
headers = read_excel(f, sheet='header_names', col_names='header')
sheets = excel_sheets(f)[1:8]

replacement_dates = read_excel(f, sheet=sheets[1], n_max=11, col_names=F, col_types="numeric") %>%
    t() %>%
    data.frame(stringsAsFactors = F) %>%
    mutate(sheet = sheets[1],
           group = 1)

i = 2
for (sheet in sheets[2:8]) {
    df_ = read_excel(f, sheet=sheet, n_max=11, col_names=F, col_types="numeric") %>%
        t() %>%
        data.frame(stringsAsFactors = F) %>%
        mutate(sheet = sheet,
               group = i)
    replacement_dates = rbind(replacement_dates, df_)
    i = i + 1
}

replacement_dates = replacement_dates %>%
    rename(!!set_names(paste0("X", 1:11), str_replace_all(headers$header, " ", "_"))) %>%
    mutate(purchase_date = ymd(str_c(purchase_year, purchase_month, "-01")),
           replace_date1 = ymd(str_c(engine_replacement_year_1, engine_replacement_month_1, "-01")),
           replace_date2 = ymd(str_c(engine_replacement_year_2, engine_replacement_month_2, "-01")),
           pivot_months_to_replace1 = interval(purchase_date, replace_date1) %/% months(1),
           pivot_months_to_replace2 = interval(replace_date1, replace_date2) %/% months(1),
           pivot_replace_mileage1 = ifelse(engine_replacement_odometer_reading_1 == 0, NA, engine_repla
           pivot_replace_mileage2 = ifelse(engine_replacement_odometer_reading_2 == 0, NA,
                                           engine_replacement_odometer_reading_2 - engine_replacement_o
    select(bus_number, purchase_date, replace_date1, pivot_months_to_replace1, replace_date2, pivot_mon

# summary2a =
    replacement_dates %>%
```

```r
    # filter(!is.na(replacement_dates$pivot_months_to_replace1)) %>%
    pivot_longer(starts_with("pivot_"),
                 names_to = c(".value", "replacement"),
                 names_pattern = "pivot_(.*)(.)",
        values_drop_na = T
    ) %>%
    mutate(group = as.character(group)) %>%
    bind_rows(mutate(., group = "Full Sample")) %>%
    group_by(group) %>%
    summarize(max_mileage = max(replace_mileage, na.rm=T),
              min_mileage = min(replace_mileage, na.rm=T),
              mean_mileage = mean(replace_mileage, na.rm=T),
              sd_mileage = sd(replace_mileage, na.rm=T),
              max_months = max(months_to_replace, na.rm=T),
              min_months = min(months_to_replace, na.rm=T),
              mean_months = mean(months_to_replace, na.rm=T),
              sd_months = sd(months_to_replace, na.rm=T),
              n = n()) %>%
  complete(group = c(as.character(1:8), "Full Sample")) %>%
    mutate(across(everything(), ~replace_na(.x, 0))) %>%
    # stargazer(digits=2, summary=F, type='text',
    #           title="Summary of Replacement Data\\\\(Subsample of buses for which at least 1 replacem
    #           column.labels = c('Bus\\\\Group', 'Max', 'Min', 'Mean', 'Standard\\\\Deviation',
    #                  'Max', 'Min', 'Mean', 'Standard\\\\Deviation', 'Number of\\\\Observations'),
    #           rownames = F)
  kbl(caption = "Summary of Replacement Data (Subsample of buses for which at least 1 replacement occur
      col.names = c('Bus Group', 'Max', 'Min', 'Mean', 'Standard Deviation',
                  'Max', 'Min', 'Mean', 'Standard Deviation', 'Number of Observations'),
      align = 'cc', digits = 1, booktabs = T,
      ) %>%
  kable_styling(latex_options = "HOLD_position") %>%
column_spec(1, width = "1.5cm") %>%
column_spec(c(2:4, 6:8), width = "1cm") %>%
column_spec(c(5, 9), width = "1.5cm") %>%
column_spec(10, width = "2cm") %>%
add_header_above(c(" " = 1, "Mileage at Replacement" = 4, "Elapsed Time (Months)" = 4, " " = 1)) %>%
      row_spec(8, hline_after = T)




# summary2b = replacement_dates %>%
#     filter(is.na(replacement_dates$months_to_replace1))

# need to convert to long first, then combine because there's different number of months for each set
# month_odo = lapply(setNames(sheets, sheets), function(x) read_excel(f, sheet=x, skip=11)) %>%
#     bind_cols(., .id="Sheet")
```