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)

	Mileage at Replacement				Elapsed Time (Months)				
Bus Group	Max	Min	Mean	Standard Deviation	Max	Min	Mean	Standard Deviation	Number of Observations
1	0	0	0	0	0	0	0.0	0.0	0
2	0	0	0	0	0	0	0.0	0.0	0
3	273,400	124,800	199,733	37,459	74	38	59.1	10.9	27
4	387,300	121,300	257,336	65,477	116	28	73.7	23.3	33
5	322,500	118,000	245,290	60,257	127	31	85.4	29.7	11
6	237,200	82,400	150,785	61,006	127	49	74.7	35.2	7
7	331,800	121,000	208,962	48,980	104	41	68.3	16.9	27
8	297,500	132,000	186,700	43,956	104	36	58.4	22.2	19
Full Sample	387,300	82,400	216,354	60,475	127	28	68.1	22.4	124

Table 2: Summary of Censored Data (subsample of buses for which no replacements occurred)

	Mileage at May 1, 1985				Elapsed Time (Months)				
Bus Group	Max	Min	Mean	Standard Deviation	Max	Min	Mean	Standard Deviation	Number of Observations
1	120,151	65,643	100,116	12,929	25	25	25.0	0.0	15
2	161,748	142,009	151,182	8,529	49	49	49.0	0.0	4
3	280,802	199,626	250,766	21,324	70	70	70.0	0.0	21
4	352,450	310,910	337,221	17,802	117	117	117.0	0.0	5
5	326,843	326,843	326,843	0	126	126	126.0	0.0	1
6	299,040	232,395	265,263	33,331	126	126	126.0	0.0	3
7	0	0	0	0	0	0	0.0	0.0	0
8	0	0	0	0	0	0	0.0	0.0	0
Full Sample	352,450	65,643	207,781	85,207	126	25	63.9	33.5	49

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

```
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 top 11 rows only to get replacement dates and mileage
f = 'rust_data_clean.xlsx'
headers = read_excel(f, sheet='header_names', col_names='header')
sheets = excel_sheets(f)[1:8]

get_headers = function(f, sheets, i) {
    # Return dataframe of transformed i'th sheet
    read_excel(f, sheet=sheets[i], n_max=11, col_names=F, col_types="numeric") %>%
    t() %>%
    data.frame(stringsAsFactors = F) %>%
    mutate(sheet = sheets[i],
           group = i)
}

# Initialize dataframe
replacement_dates = get_headers(f, sheets, 1)

# Fill dataframe with other sheets
for (i in 2:8) {
    df_ = get_headers(f, sheets, i)
    replacement_dates = rbind(replacement_dates, df_)
}

# Pivot dataframe and create variables
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_1_year, engine_replacement_1_month, "-01")),
           replace_date2 = ymd(str_c(engine_replacement_2_year, engine_replacement_2_month, "-01")),
           pivot_months_to_replace1 = interval(purchase_date, replace_date1) %/% months(1) + 1,
           pivot_months_to_replace2 = interval(replace_date1, replace_date2) %/% months(1) + 1,
           pivot_replace_mileage1 = ifelse(engine_replacement_1_odometer_reading == 0, NA,
                                           engine_replacement_1_odometer_reading),
           pivot_replace_mileage2 = ifelse(engine_replacement_2_odometer_reading == 0, NA,
                                           engine_replacement_2_odometer_reading - engine_replacement_1
           initial_odometer_reading_date = ymd(str_c(initial_odometer_reading_year, initial_odometer_rea
    select(bus_number, purchase_date, replace_date1, pivot_months_to_replace1, replace_date2, pivot_mon
```

```
# Make Table IIa
replacement_dates %>%
    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 = as.integer(mean(replace_mileage, na.rm=T)),
              sd_mileage = as.integer(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))) %>%
    kbl(caption = "Summary of Replacement Data (Subsample of buses for which at least 1 replacement occu
        col.names = c('Bus Group', 'Max', 'Min', 'Mean', 'Standard Deviation',
                      'Max', 'Min', 'Mean', 'Standard Deviation', 'Number of Observations'),
        align = 'cc', digits = 1, booktabs = T, format.args = list(big.mark = ",")
    ) %>%
    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)

# Load rest of data below row 11 to get mileage readings
f = 'rust_data_clean.xlsx'
sheets = excel_sheets(f)[1:8]

get_data = function(f, sheet_) {
    # Return transformed dataframe from sheet
    read_excel(f, sheet=sheet_, skip=11,
               col_names=paste0('bus_', filter(replacement_dates, sheet==sheet_)$bus_number),
               col_types="numeric") %>%
        mutate(sheet = sheet_,
               m = row_number()) %>%
        pivot_longer(starts_with("bus_"),
                     names_to = "bus_number",
                     names_prefix = 'bus_') %>%
        mutate(bus_number = as.numeric(bus_number)) %>%
        rename(c('mileage' = 'value'))
}

# Initialize dataframe
```

```r
data = get_data(f, sheets[1])

# Get rest of sheets
for (i in 2:8) {
    df_ = get_data(f, sheets[i])
    data = rbind(data, df_)
}

# Join on variables from headers and create new variables
data = data %>%
    arrange(sheet, bus_number) %>%
    left_join(replacement_dates %>% select(bus_number,
                                           initial_odometer_reading_date,
                                           replace_date1,
                                           group),
              by = 'bus_number') %>%
    mutate(date = initial_odometer_reading_date + months(m - 1),
           group = as.character(group))

# Make Table IIa
data %>%
    filter(is.na(replace_date1)) %>%
    arrange(date) %>%
    group_by(group) %>%
    filter(date == last(date), .by_group = TRUE) %>%
    bind_rows(mutate(., group = "Full Sample")) %>%
    summarize(max_mileage = max(mileage, na.rm=T),
              min_mileage = min(mileage, na.rm=T),
              mean_mileage = as.integer(mean(mileage, na.rm=T)),
              sd_mileage = as.integer(sd(mileage, na.rm=T)),
              max_months = max(m, na.rm=T),
              min_months = min(m, na.rm=T),
              mean_months = mean(m, na.rm=T),
              sd_months = sd(m, na.rm=T),
              n = n()) %>%
    complete(group = c(as.character(1:8), "Full Sample")) %>%
    mutate(across(everything(), ~replace_na(.x, 0))) %>%
    kbl(caption = "Summary of Censored Data (subsample of buses for which no replacements occurred)",
        col.names = c('Bus Group', 'Max', 'Min', 'Mean', 'Standard Deviation',
                      'Max', 'Min', 'Mean', 'Standard Deviation', 'Number of Observations'),
        align = 'cc', digits = 1, booktabs = T, format.args = list(big.mark = ",")
    ) %>%
    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 May 1, 1985" = 4, "Elapsed Time (Months)" = 4, " " = 1)) %>%
    row_spec(8, hline_after = T)
```