Recent Macro Trends: Data Analysis

Rodolphe Aurèle Zumbrunnen

2025-10-27

0. Set-up

Work description

- Following Karahan et al. (2024), proposing an overview of the evolution of the startup rate in Japan and it's potential link with the labour force growth rate.

Data Description

- Population demographics for Japan are retrieved retrieved from OECD data.
  - **WAP:** OECD. *Working-age population (15-64) – Quarterly growth rate* [Data set]. OECD Data Explorer. Retrieved October 28, 2025, from https://data-explorer.oecd.org/?lc=en
  - **CLF:** OECD. *Labour force – Quarterly level for 15-64 age group* [Data set]. OECD Data Explorer. Retrieved October 28, 2025, from https://data-explorer.oecd.org/?lc=en
- Japanese firms demographics are not available in any digital format. The values are thus retrieved from PDF reports and compiled using ChatGPT (also used for traduction).
  - **BEFORE 2006:** Japan Institute for Labour Policy and Training (JILPT). (2009). *Chapter 12: Trends in establishments,* pp. 139–147. Tokyo: JILPT. https://www.jil.go.jp/kokunai/statistics/kako/2009/documents/12_p139-147.pdf
  - **AFTER 2006:** Small and Medium Enterprise Agency. (2020). *Whte Paper on SMEs in Japan, 2020, Part 1, Chapter 3, §1: Changes in number of enterprises and openings/closings.* Tokyo. https://www.chusho.meti.go.jp/pamflet/hakusyo/2020/PDF/chusho/03Hakusyo_part1_chap3_web.pdf

Packages

```
library(dplyr)
library(tidyr)
library(zoo)
library(ggplot2)
```

1. Population and labour force

Loading Data

```
df_pop <- read.csv("WAP_OECD.csv", header = T )
df_labour <- read.csv("labour_force.csv", header = T )
```

WAP Cleaning

```
df_pop <- df_pop %>%
    filter(REF_AREA == "JPN") %>%
    select(REF_AREA, TIME_PERIOD, OBS_VALUE) %>%
    rename(WAP = OBS_VALUE)
```

CLF Cleaning

```
df_labour <- df_labour %>%
    filter(REF_AREA == "JPN") %>%
    select(REF_AREA, TIME_PERIOD, OBS_VALUE) %>%
    rename(CLF = OBS_VALUE) %>%
    arrange(TIME_PERIOD)
```

Merging

- Only keeping overlapping years.

```
df_population <- df_labour %>%
    inner_join(df_pop, by = c("REF_AREA", "TIME_PERIOD"))
```

Estimating Rates

- Simple calculation of WAP and CLF growth rates.

```
df_population <- df_population %>%
  group_by(REF_AREA) %>%
  mutate(CLF_growth = ((CLF - lag(CLF)) / lag(CLF)*100),
         WAP_growth = ((WAP - lag(WAP)) / lag(WAP)*100) )
```

Long Formatting

```
df_population_long <- pivot_longer(df_population,
                                   cols = c(CLF_growth, WAP_growth),
                                   names_to = "Series",
                                   values_to = "Growth")
```

2. Firm demographics

Loading Data

```
df_jap <- read.csv("japan_data.csv", header=T)
```

Data Cleaning

- Keeping only the starting years, and the rates.
- Reformatting to a long format for the plot.

```
df_jap <- df_jap %>%
    select(Start_year, Opening_rate_percent, Closure_rate_percent) %>%
    pivot_longer(cols = c(Opening_rate_percent, Closure_rate_percent),
                 names_to = "Variable", values_to = "Rate") %>%
    mutate(Variable = recode(Variable,
                             Opening_rate_percent = "Entry_rate",
                             Closure_rate_percent = "Exit_rate"))
```

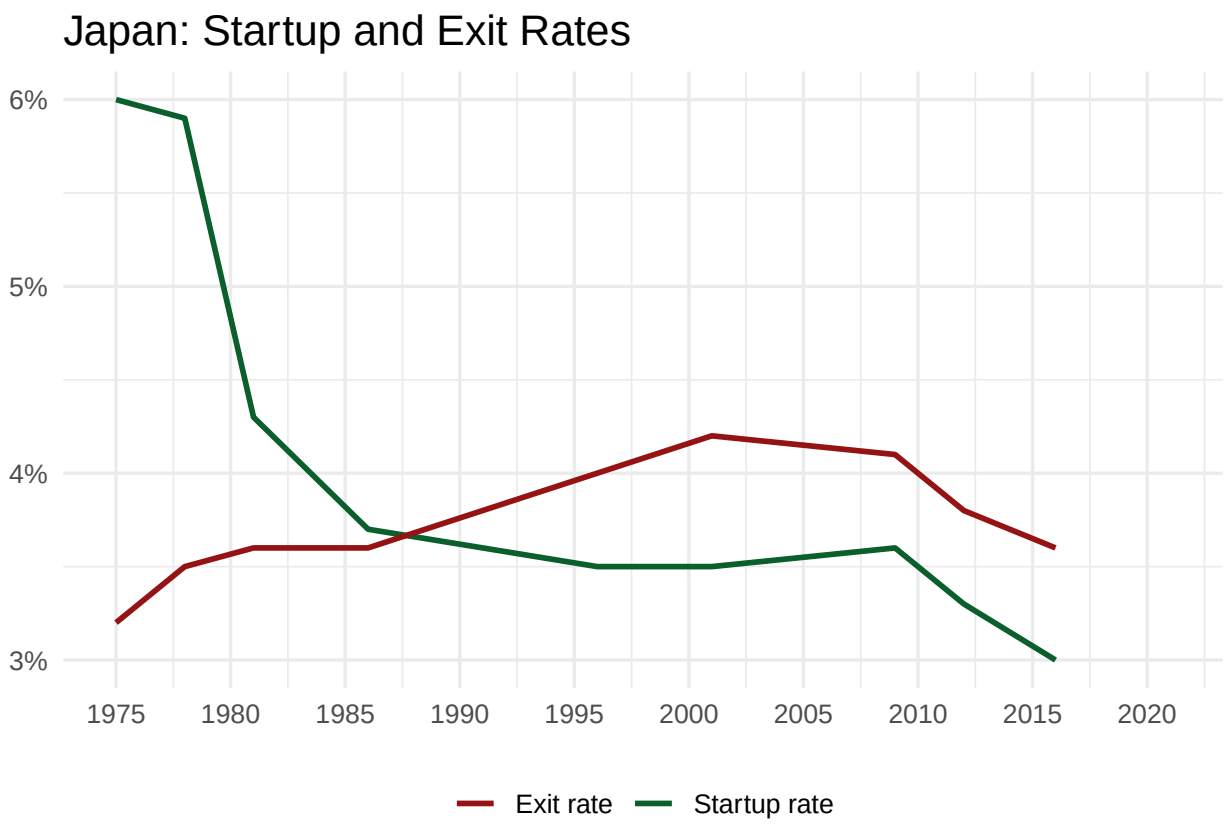
3. Graphical analysis

Labour Force Evolution

```
ggplot(df_population, aes(x = as.integer(TIME_PERIOD))) +
  geom_line(aes(y = CLF_growth / 100,
                color = "Civilian Labour Force"), linewidth = 1) +
  geom_line(aes(y = WAP_growth / 100,
                color = "Working-age Population"), linewidth = 1) +
  geom_hline(yintercept = 0, color = "black", size = 1) +
  labs(title = "Japan: Labour Supply Growth Rates",
       x = NULL, y = NULL, color = NULL) +
  scale_color_manual(values = c("Civilian Labour Force" = "#0A5F2D",
                                "Working-age Population" = "#961314")) +
  scale_y_continuous(labels = scales::percent_format(accuracy = 1)) +
  scale_x_continuous(breaks = seq(1975, 2021, by = 5), limits = c(1975, 2021))+
  theme_minimal(base_size = 13) +
  theme(legend.position = "bottom") +
  coord_cartesian(ylim = range(c(df_population$CLF_growth,
                                 df_population$WAP_growth), na.rm = TRUE) /100)
```

Firm Demograhics Evolution

```
ggplot(df_jap, aes(x = as.integer(Start_year))) +
  geom_line(data = subset(df_jap, Variable == "Entry_rate"),
            aes(y = Rate/100, color = "Startup rate"),
            linewidth = 1) +
  geom_line(data = subset(df_jap, Variable == "Exit_rate"),
            aes(y = Rate/100, color = "Exit rate"),
            linewidth = 1) +
  geom_hline(yintercept = 0, color = "black", size = 1) +
  labs(title = "Japan: Startup and Exit Rates",
       x = NULL, y = NULL, color = NULL) +
  scale_color_manual(values = c("Startup rate" = "#0A5F2D",
                                "Exit rate" = "#961314")) +
  theme_minimal(base_size = 13) +
  theme(legend.position = "bottom") +
  coord_cartesian(ylim = range(df_jap$Rate/100, na.rm = TRUE)) +
  scale_x_continuous(breaks = seq(1975, 2021, by = 5), limits = c(1975, 2021)) +
  scale_y_continuous(labels = scales::percent_format(accuracy = 1))
```



4. Elasticity evaluation

Merging dataframes

```
df_merged <- df_jap %>%
  filter(Variable == "Entry_rate") %>%
  left_join(df_population %>%
              select(TIME_PERIOD, WAP_growth),
            by = c("Start_year" = "TIME_PERIOD")) %>%
  select(Start_year, Rate, WAP_growth) %>%
  rename(Year = Start_year, Entry_rate = Rate)
```

Estimating Elasticity

- Log-log technique is not producing realistic results due to the rates being close to zero.
- Here a simple level difference is used as a proxy for the elasticity, making use of the first order Taylor expansion of log(1+x) = x.

```
df_elast <- df_merged %>%
  arrange(Year) %>%
  mutate(d_entry = Entry_rate - lag(Entry_rate),
         d_labour = WAP_growth - lag(WAP_growth)) %>%
  filter(!is.na(d_entry), !is.na(d_labour), abs(d_labour) > 0.001)
```

```r
elasticity <- sum(df_elast$d_entry) / sum(df_elast$d_labour)

cat("Elasticity of the Startup rate WRT labour force:", round(elasticity, 2))
```

```
## Elasticity of the Startup rate WRT labour force: 1.69
```