Male-Female Differences

Henrik Schubert

2023-06-05

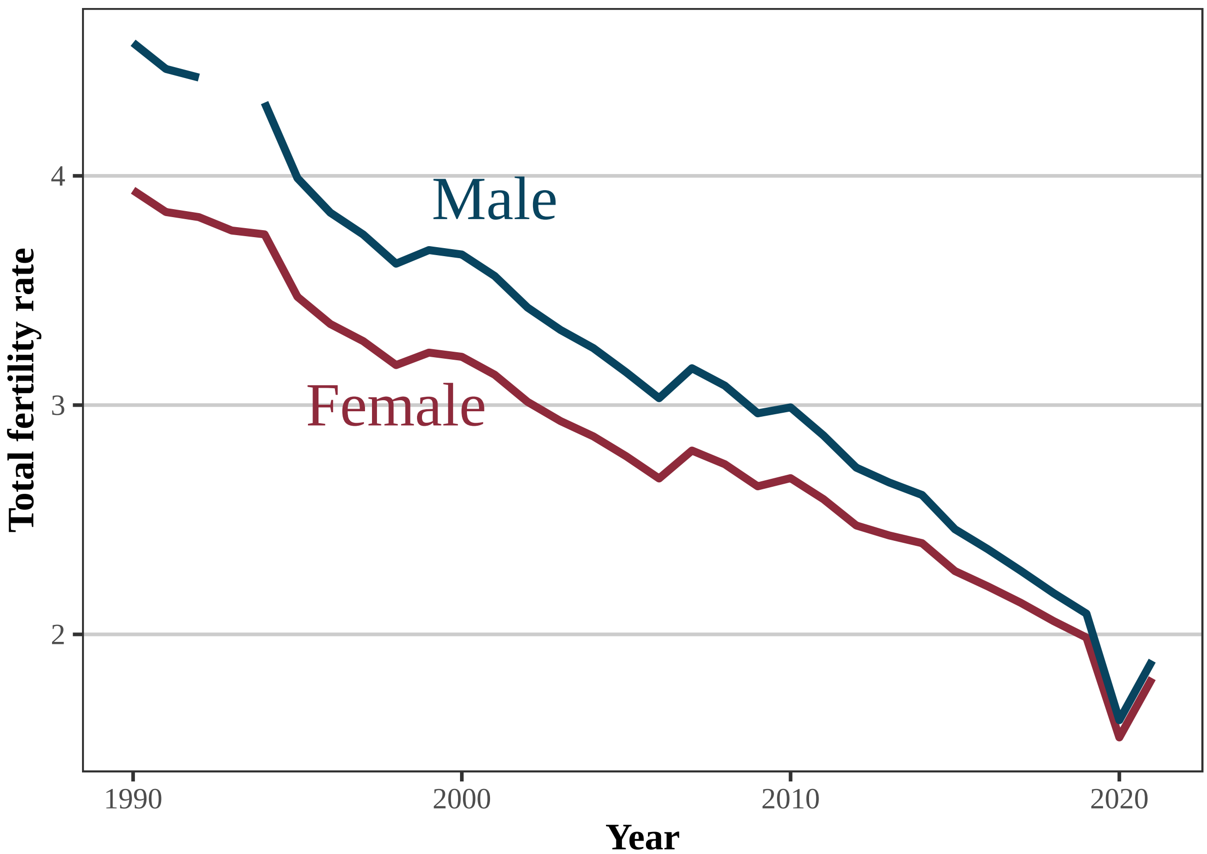
Abstract

This file is going to compare male and female fertility rates in Mexico. We will use graphical tools, distributional comparison measures, decomposition methods and statistical tests. The comparisons will be both over time, and at the national level

Time trend

First, we compare the time-trend of male and female total fertility rates in Mexico. We use the estimates based on vital statistics data from !(INEGI)[] and from !(the demographic group)[].

```
# What is the trend in the national tfr s
ggplot(tfr_nat, aes(year)) +
  geom_line(aes(y = tfr_f, colour = "female"), linewidth = 1.4) +
  geom_line(aes(y = tfr_m, colour = "male"), linewidth = 1.4) +
  annotate(geom = "text",
           x = 2001, y = 3.9,
           label = "Male",
           colour = MPIDRblue, size = 8, family = "serif") +
  annotate(geom = "text",
           x = 1998, y = 3,
           label = "Female",
           colour = MPIDRred, size = 8, family = "serif") +
  scale_colour_manual(values = c(MPIDRred, MPIDRblue)) +
  ylab("Total fertility rate") +
  xlab("Year") +
  guides(colour = "none")
```



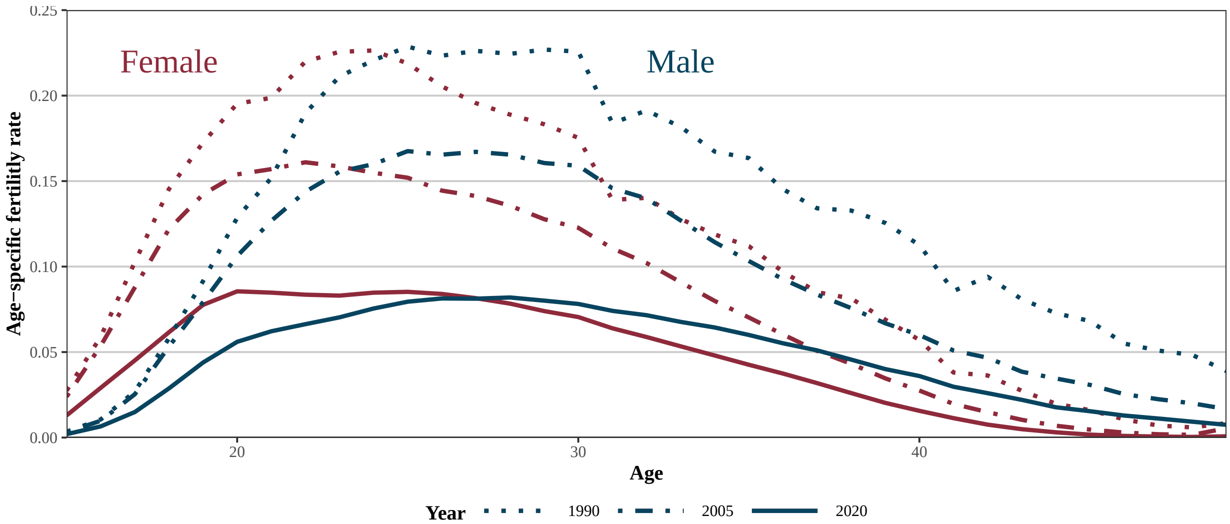
The graph shows several time trends. First, both male and female total fertility rates declined between 1990 and 2021 monolithic. Second, the male TFR declined stronger than the female TFR, which led to a closing of the sex difference in total fertility rates.

Age-Specific Differences

Now, we plot the age-specific differences in fertility between men and women.

```
# Load the national level data
load("Data/asfr_national_mexico.Rda")

# What is the trend in the national tfr s
ggplot(subset(asfr_nat, year %in% seq(1990, 2020, by = 15)), aes(x = age, linetype = as.factor(year))) +
  geom_line(aes(y = asfr_f, group = year),
            linewidth = 1.4, colour = MPIDRred) +
    geom_line(aes(y = asfr_m, group = year),
            linewidth = 1.4, colour = MPIDRblue) +
  annotate(geom = "text",
           x = 18, y = 0.22,
           label = "Female",
           colour = MPIDRred, size = 8, family = "serif") +
  annotate(geom = "text",
           x =33, y = 0.22,
           label = "Male",
           colour = MPIDRblue, size = 8, family = "serif") +
  scale_y_continuous(expand = c(0, 0), limits = c(0, 0.25)) +
  scale_x_continuous(expand = c(0, 0)) +
```

```
  scale_linetype_manual(name = "Year", values = c("dotted", "dotdash", "solid")) +
  ylab("Age-specific fertility rate") +
  xlab("Age") +
  guides(colour = "none", alpha = "none")
```



We find in contrast to other countries, that the fertilty decline is not largely driven by postponement, but rather by a decline in age-specific fertiltiy rates across all age-grouops. A fertility recuperation or a shift of childbearing at a later point is absent in Mexico. Regarding the differences between men and women, there are some noticeable differences. First, across the all years, childbearing for men takes place later in life. This is apparent in higher female age-specific fertility rates ($f(x)$) at younger ages and higher male $f(x)$ at higher ages. This life course behaviour produces a certain cross-over. Second, the spread of childbearing over the life course is wider. For women above age 45, the age-specific fertility rates are almost zero, while for men there is some childbearing at every age. This is consistent with findings from Schoumaker (2019) and Dudel and Klüsener (2019).

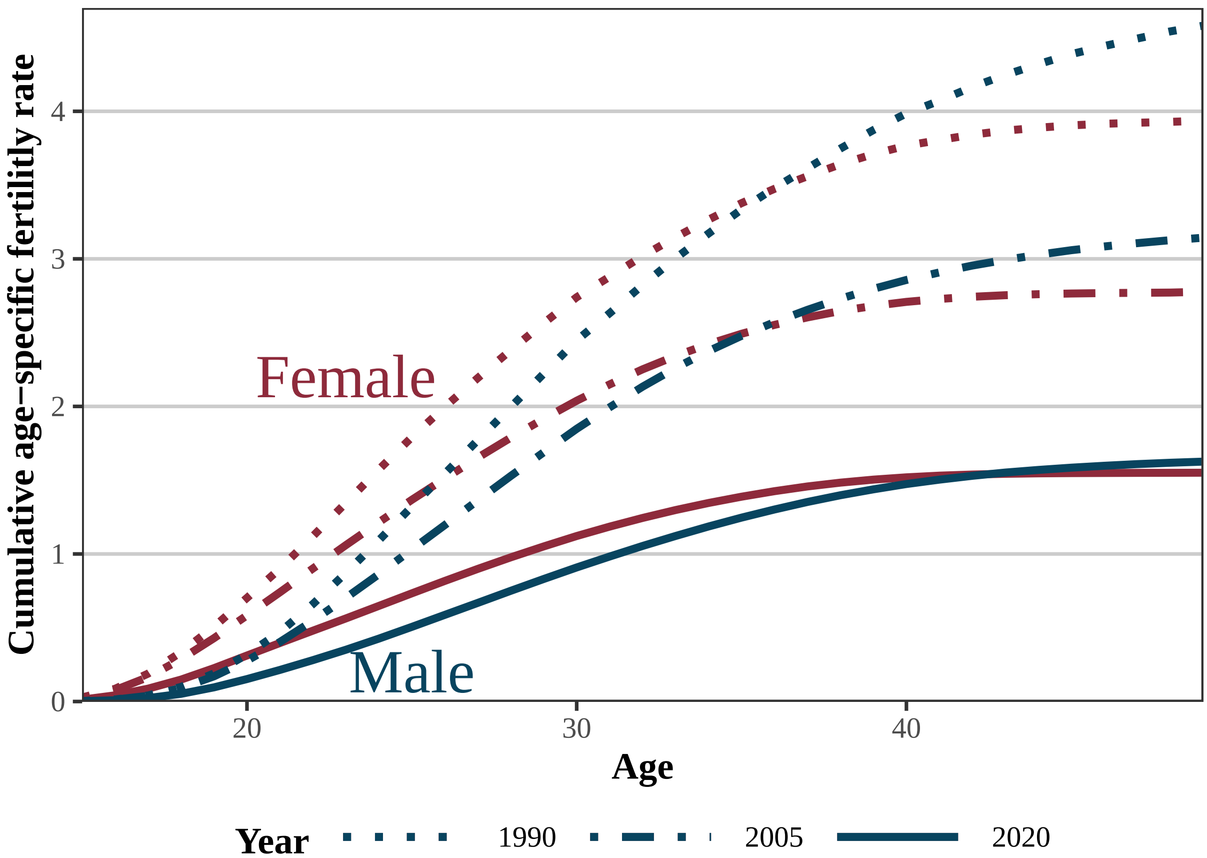
Cumulative distribution

Moreover, to understand better the postponement behaviour, we are going to estimate the cumulative age-specific fertility rate.

$$cumulative f(x) = \sum_{12}^{x} f(x)$$

```
# Estimate the cumulative distribution
asfr_nat <- asfr_nat %>%
  group_by(year) %>%
  arrange(age) %>%
  mutate(cum_f = cumsum(asfr_f),
         cum_m = cumsum(asfr_m))

# What is the trend in the national tfr s
ggplot(subset(asfr_nat, year %in% seq(1990, 2020, by = 15)), aes(x = age, linetype = factor(year))) +
  geom_line(aes(y = cum_f, group = year),
            linewidth = 1.4, colour = MPIDRred) +
    geom_line(aes(y = cum_m, group = year),
            linewidth = 1.4, colour = MPIDRblue) +
```

```
  annotate(geom = "text",
           x = 23, y = 2.2,
           label = "Female",
           colour = MPIDRred, size = 8, family = "serif") +
   annotate(geom = "text",
           x = 25, y = 0.2,
           label = "Male",
           colour = MPIDRblue, size = 8, family = "serif") +
  scale_linetype_manual(name = "Year", values = c("dotted", "dotdash", "solid")) +
  scale_y_continuous(expand = c(0, 0), limits = c(0, 4.7)) +
  scale_x_continuous(expand = c(0, 0)) +
  ylab("Cumulative age-specific fertilitly rate") +
  xlab("Age") +
  guides(colour = "none", alpha = "none")
```



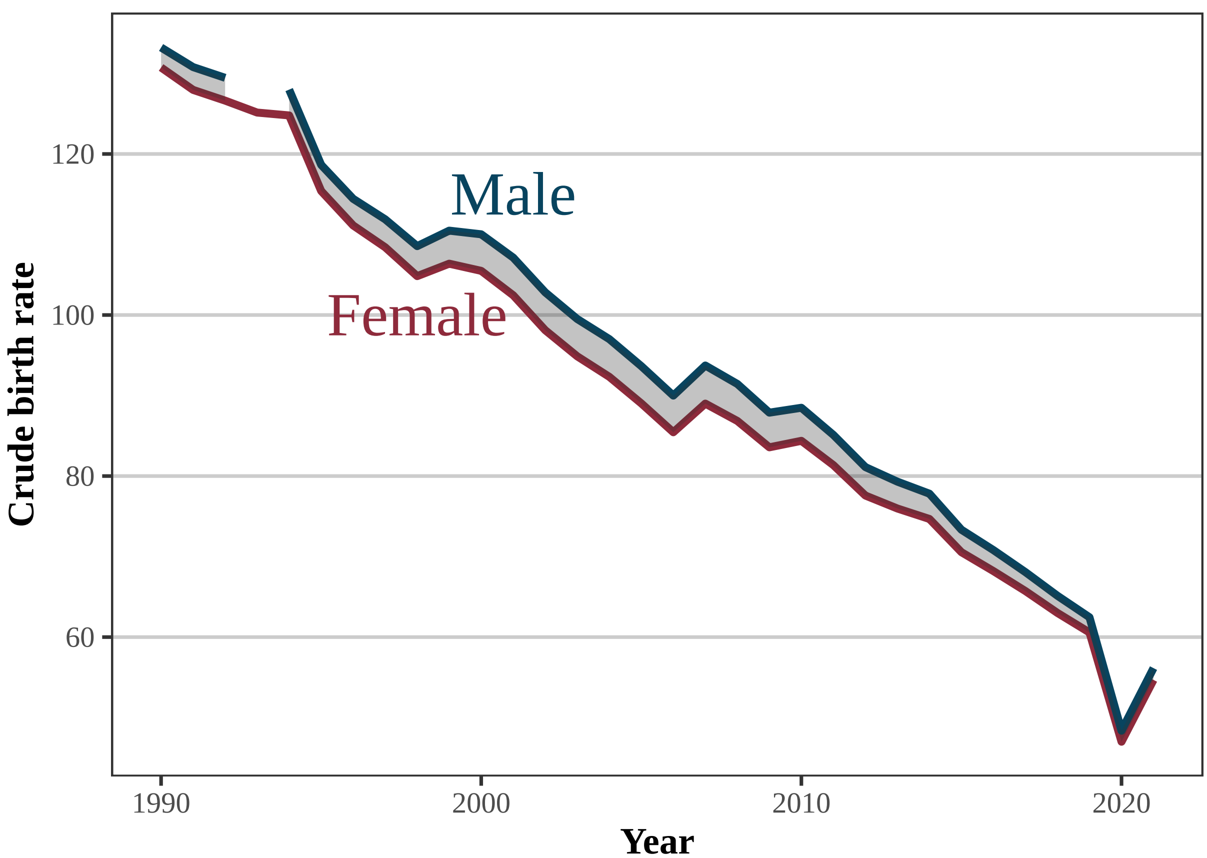
Again, we see that fertility declined over time. The cumulative fertility rate in recent years is much lower than the rates for 2005 and 1990. Moreover, we see that in every year the cross-over behavior of male and female fertility. Moreover, the cross-over for male and female fertiltiy rates occurs in all years. However, the cross-over shifted to later ages. In 1990, the male-female fertility cross-over occured at age 36, while in 2020, 43.

Decomposing the difference in the total fertility rate

Another tool to study the difference between the male and female is *decomposition*. The crude fertility rate can be decomposed into two components: *population structure* and *birth rates*. Because of the additive nature, it is feasible to use standard demographic decomposition methods (Preston, Heuveline, and Guillot (2008),

p. 28)

```r
# Plot the crude birth rate for men and women
cbr_nat <- asfr_nat |>
  group_by(year) |>
  summarise(cbr_f = 1000 * sum(births_f) / sum(exposure_f),
            cbr_m = 1000 * sum(births_m) / sum(exposure_m))

# Plot the crude birth rates
ggplot(data = cbr_nat, aes(x = year)) +
  geom_line(aes(y = cbr_f, colour = "female"), linewidth = 1.4) +
  geom_line(aes(y = cbr_m, colour = "male"), linewidth = 1.4) +
  geom_ribbon(aes(ymin = cbr_f, ymax = cbr_m), alpha = 0.3) +
  annotate(geom = "text",
           x = 2001, y = 115,
           label = "Male",
           colour = MPIDRblue, size = 8, family = "serif") +
  annotate(geom = "text",
           x = 1998, y = 100,
           label = "Female",
           colour = MPIDRred, size = 8, family = "serif") +
  scale_colour_manual(values = c(MPIDRred, MPIDRblue)) +
  ylab("Crude birth rate") +
  xlab("Year") +
  guides(colour = "none")
```

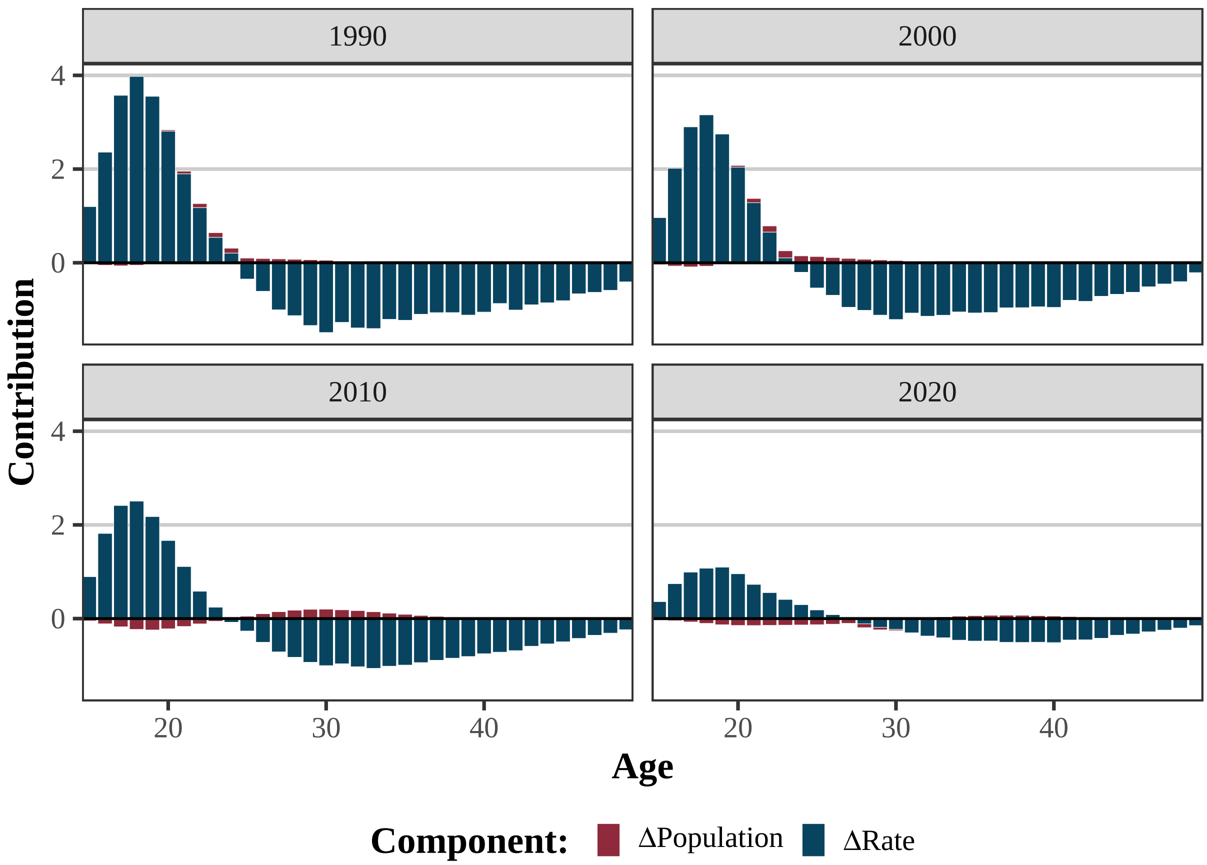


We see, that similar as for the total fertility rate, the gap between male and female crude fertility rates has diminished. The decomposition approach looks as follows:

$$\Delta CFR_{f,m} = \underbrace{\sum_i \left(C_i^f - C_i^m\right)\left[\frac{f_i^f + f_i^m}{2}\right]}_{\Delta Population} + \underbrace{\sum_i (f_I^f - f_i^m) \cdot \left[\frac{C_i^f + C_i^m}{2}\right]}_{\Delta Rate}$$

where f is the age-specific fertility rate, C is the population share, i indexes the different age-groups, and f and m are indicating the male or female part. The first term marks the weighted change in the rate ($\Delta Rate$) and the second term marks the weighted change in population ($\Delta Population$).

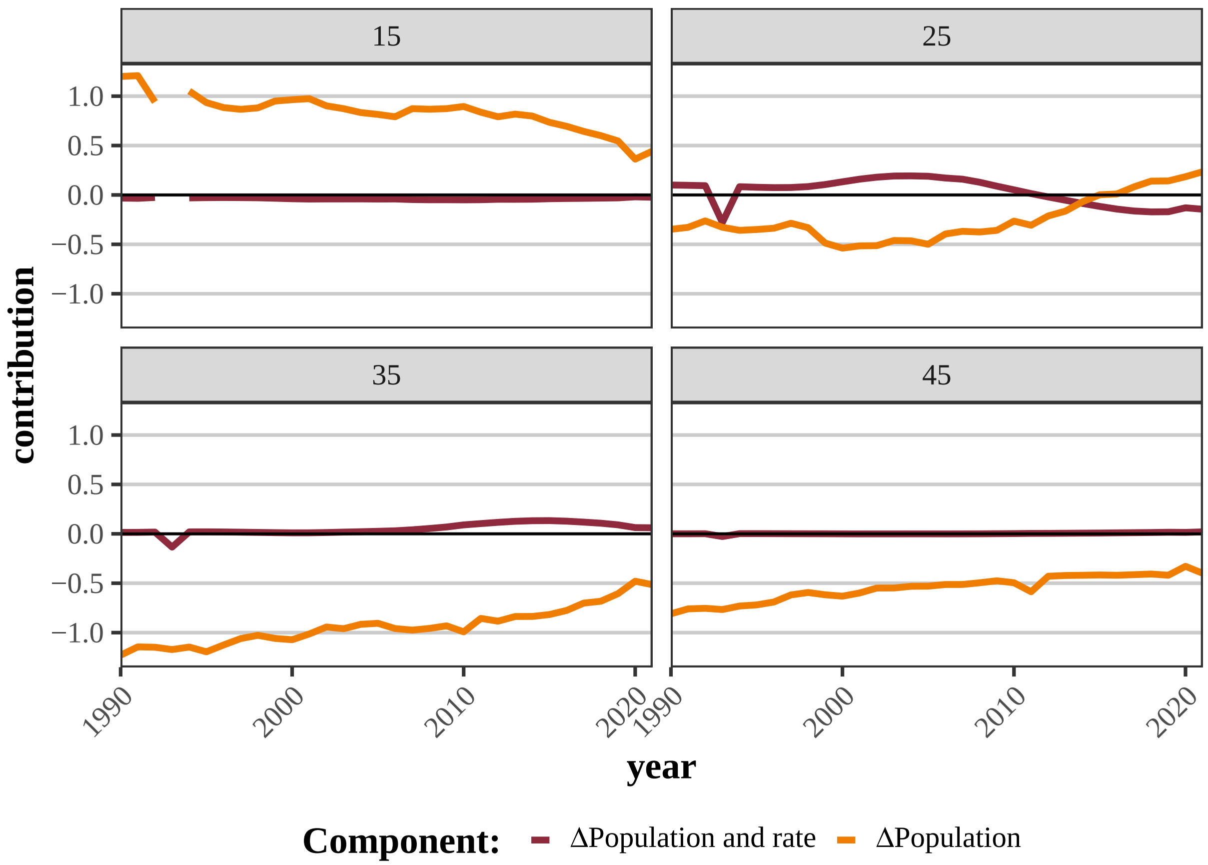
```
# Estimate the decomposition
comp_asfr <- asfr_nat |>
                group_by(year) |>
                mutate(across(starts_with("asfr"), ~ .x * 1000)) |>
                mutate(pop_share_f = pop_share(exposure_f),
                       pop_share_m = pop_share(exposure_m),
                       mean_pop    = averaging(pop_share_f, pop_share_m),
                       mean_rate   = averaging(asfr_f, asfr_m),
                       delta_rate  = (asfr_f - asfr_m) * mean_pop,
                       delta_pop   = (pop_share_f - pop_share_m) * mean_rate,
                       .groups     = "drop")

# Make the decomposition
decomp_nat <- comp_asfr %>%
  select(year, age, delta_rate, delta_pop) %>%
  pivot_longer(cols = starts_with("delta"),
               names_to = "component",
               values_to = "contribution")

# Plot the decomposition
decomp_nat %>%
  filter(year %in% c(1990, 2000, 2010, 2020)) %>%
  ggplot(aes(age, contribution, fill = component)) +
  geom_col(colour = "white", linewidth = 0.01) +
  geom_hline(yintercept = 0, colour = "black") +
  facet_wrap(~ year) +
  scale_fill_manual(values = c(MPIDRred,  MPIDRblue),
                    labels = c(expression(paste(Delta, "Population")),
                               expression(paste(Delta, "Rate"))),
                    name = "Component: ") +
  scale_x_continuous(expand = c(0, 0)) +
  ylab("Contribution") + xlab("Age") +
  theme(legend.key.width = unit(0.3, "cm"),
        legend.key.height = unit(0.2, "cm"))
```



In the figure below, we see that

Moreover, since we have the additive components for single years, we can also investigate how the different components changed over time in order to better explain the changes in the gap between male and female fertility rate.

```
# Plot the time trend for populatin estimates
decomp_nat %>%
  filter(age %in% c(15, 25, 35, 45)) %>%
  ggplot(aes(x = year, contribution, colour = component, group = component)) +
  geom_line(linewidth = 1.2) +
  geom_hline(yintercept = 0, colour = "black") +
  scale_colour_manual(values = c(MPIDRred, MPIDRorange, MPIDRblue),
                      labels = c(expression(paste(Delta, "Population and rate")),
                               expression(paste(Delta, "Population")),
                               expression(paste(Delta, "Rate"))),
                      name = "Component: ") +
  scale_y_continuous(labels = scales::label_number_si()) +
  scale_x_continuous(expand = c(0, 0)) +
  facet_wrap( ~ age) +
  theme(legend.key.width = unit(0.3, "cm"),
        legend.key.height = unit(0.2, "cm"),
        axis.text.x = element_text(angle = 45, hjust = 1) )
```



In order to better understand the size of the contribution, we create tables that display the components with the largest contribution.

```
# Estimate the relative contribution
relative_contribution <- left_join(decomp_nat, cbr_nat, by = c("year")) %>%
  mutate(diff = cbr_f - cbr_m,
         relative_contribution = contribution / diff, 2)


# Make a table for the relative contribution
relative_contribution %>%
  filter(year == 1990) %>%
  ungroup() %>%
  select(component, age, contribution,
         relative_contribution, cbr_f, cbr_m, diff) %>%
  arrange(desc(abs(relative_contribution))) %>%
  mutate(relative_contribution =  paste(round(100 * relative_contribution, 2), "%")) %>%
  slice_head(n = 5) %>%
  pander(col.names = c("Component", "Age", "Contribution", "Contribution (%)", "CBR (females)", "CBR (ma
```

Table 1: Table continues below

Component	Age	Contribution	Contribution (%)	CBR (females)
delta_rate	18	3.978	-160.21 %	130.7
delta_rate	17	3.574	-143.93 %	130.7
delta_rate	19	3.552	-143.07 %	130.7

Component	Age	Contribution	Contribution (%)	CBR (females)
delta_rate	20	2.811	-113.2 %	130.7
delta_rate	16	2.361	-95.11 %	130.7

CBR (males)	Difference
133.2	-2.483
133.2	-2.483
133.2	-2.483
133.2	-2.483
133.2	-2.483

```
# Cumulative contribution of components
relative_contribution |>
  group_by(year, component) |>
  summarise(contribution = sum(contribution),
            cbr_m = unique(cbr_m),
            cbr_f = unique(cbr_f),
            diff = unique(diff),
            .groups = "drop") |>
  filter(year %in% seq(1990, 2020, by = 15)) |>
  pivot_wider(names_from = "component", values_from = "contribution") |>
  pander(col.names = c("Year", "CBR (men)", "CBR (women)", "Difference", "Population ", "Rate"))
```

Year	CBR (men)	CBR (women)	Difference	Population	Rate
1990	133.2	130.7	-2.483	0.7684	-3.251
2005	93.65	89.02	-4.63	1.018	-5.648
2020	48.38	47	-1.371	-0.9146	-0.4562

References

Dudel, Christian, and Sebastian Klüsener. 2019. "Estimating Men's Fertility from Vital Registration Data with Missing Values." *Population Studies* 73 (3): 439–49. https://doi.org/10.1080/00324728.2018.1481992.

Preston, Samuel H., Patrick Heuveline, and Michel Guillot. 2008. *Demography: Measuring and Modeling Population Processes*. 9. [pr.]. Oxford: Blackwell.

Schoumaker, Bruno. 2019. "Male Fertility Around the World and Over Time: How Different Is It from Female Fertility?" *Population and Development Review* 45 (3): 459–87. https://doi.org/10.1111/padr.12273.