# A reversal of fortunes: cities and the gender wage gap

César Garro-Marín

November 19, 2020

Boston University

# Introduction

# Introduction

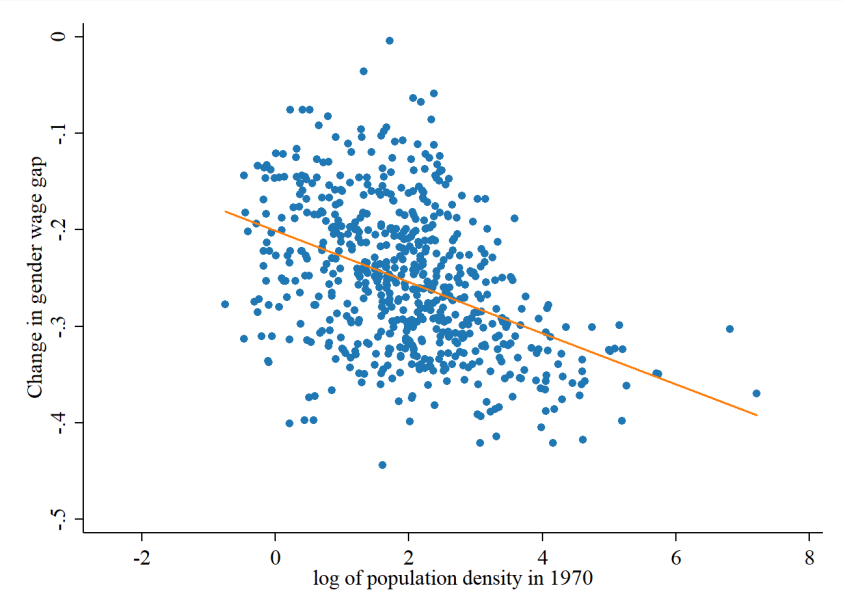
- There is a wage premium from living in a large labor market [citations here].
- But little attention has been paid as to how this premium varies by gender.
  In this paper I document three facts that link the gender wage-gap with the urban wage premium:
  1. Today, women **gain relatively** more from being in a denser labor market $\Longrightarrow$ denser labor markets have **a narrower gender-wage gap**.
  2. Women relative gain from being in a denser labor market has increased since 1970 $\Longrightarrow$ the gender gap-density elasticity went from **0.04** to **-0.01**.
  3. sth about possible channels here

# Literature

- Gender gap literature
- Decline in the urban wage premium
- Urban literature.

# Today the gender gap is narrower in denser labor markets

**Figure 1:** Change in gender wage gap, 1970-2020




**Note:** figure restricts to CZ with more than 1 people per $km^2$. Figure generated on 17 Nov 2020 at 13:59:44. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

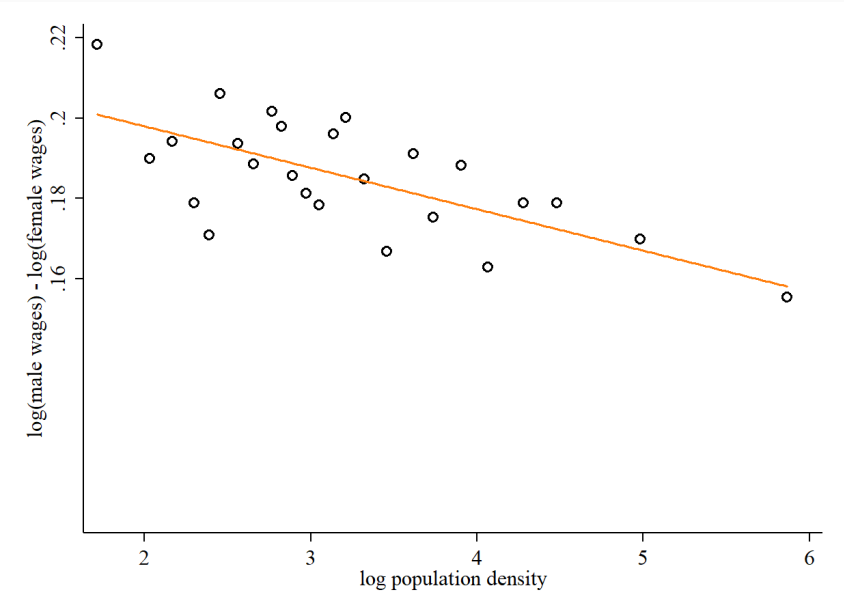
The negative gradient is robust to:

- Limiting to the largest CZ Largest CZ.
- Weighting scheme Weighted regression.
- Controlling for individual characteristics Residualized wages.

# The negative correlation is robust

## Limiting to the largest CZ

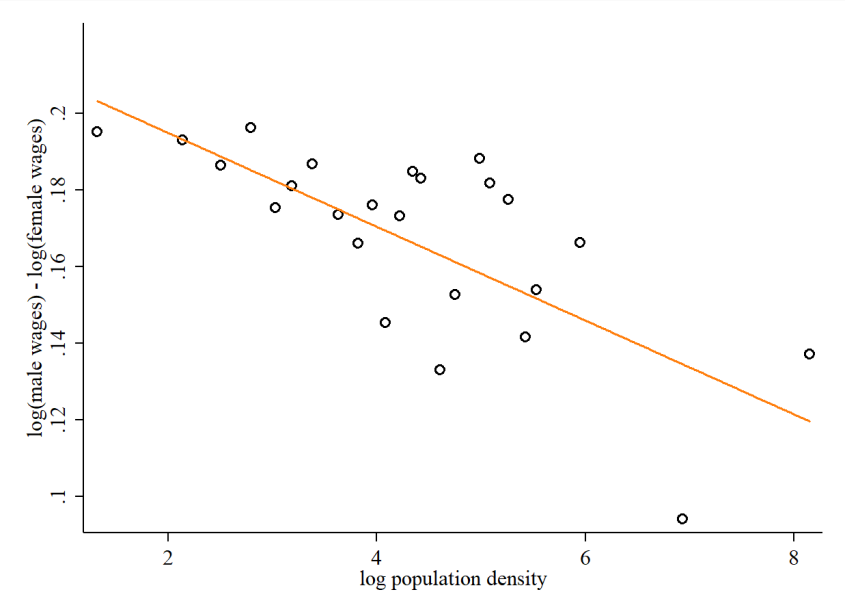
**Figure 2:** Gender wage gap and population density in the largest CZ




**Note:** figure restricts to CZ with more than 1 people per $km^2$. Each point represents about 13 CZ. Figure generated on 17 Nov 2020 at 13:42:00. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

## Weighting by population

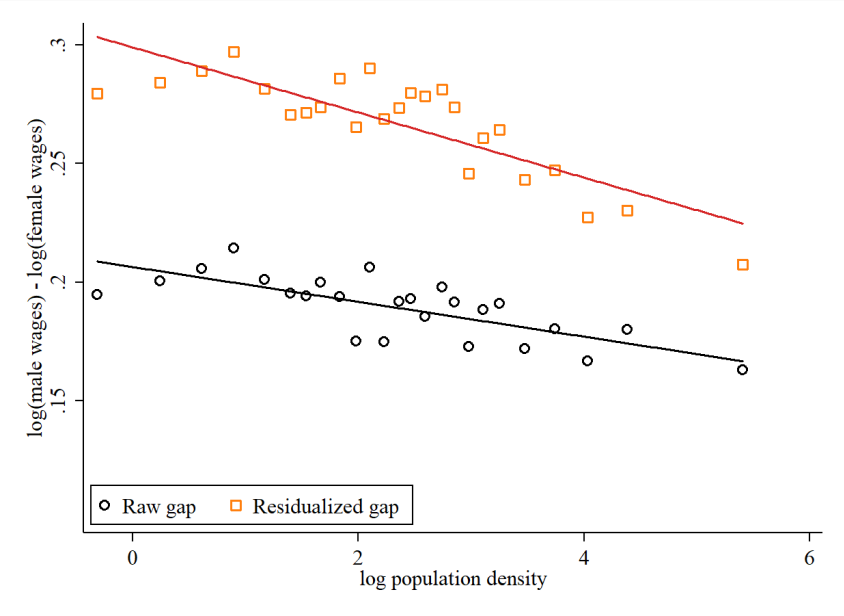
**Figure 3:** Gender wage gap and population density in 2020 (population weighted)




**Note:** figure restricts to CZ with more than 1 people per $km^2$. Each point represents about 4 percent of the working age population. Figure generated on 17 Nov 2020 at 13:42:00. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

# It is robust to accounting for individual characteristics

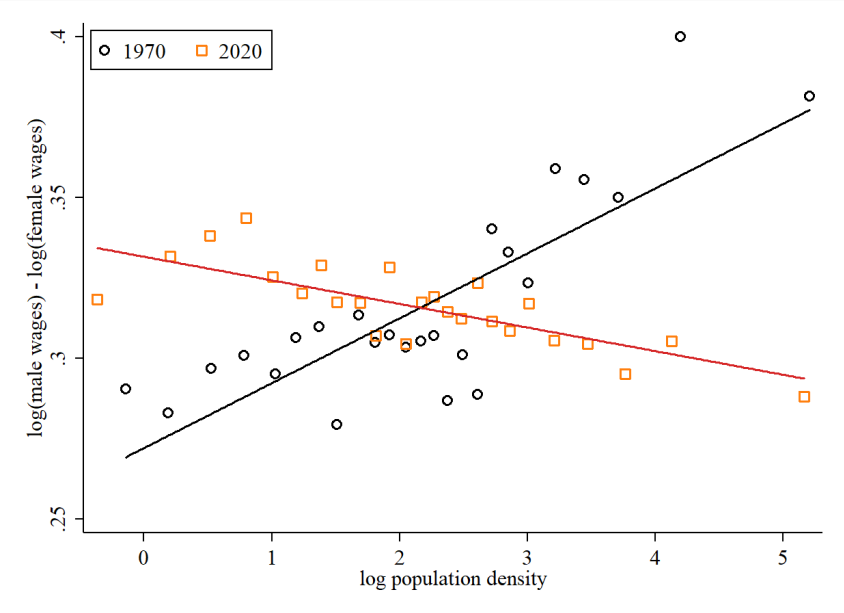
**Figure 4:** Gender wage gap and population density in 2020 (population weighted)




**Note:** figure restricts to CZ with more than 1 people per $km^2$. Each point represents about 4 percent of the working age population. Figure generated on 17 Nov 2020 at 13:42:01. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

# The gender-gap density gradient has declined since 1970

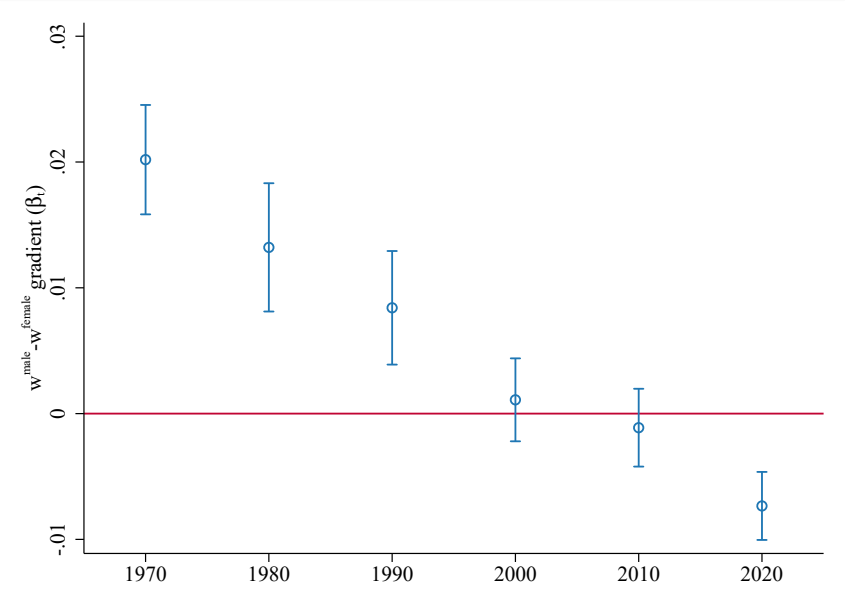
**Figure 5:** Gender wage gap and population density




**Note:** figure restricts to CZ with more than 1 people per km$^2$. Each point represents about 25 CZ. Year fixed effects are absorbed. Figure generated on 17 Nov 2020 at 13:42:02. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

Graph without absorb

**Figure 6:** Coefficient on population density $\beta_t$




**Note:** figure restricts to CZ with more than 1 people per km$^2$. Bars show 95% confidence intervals. Standard errors clustered at the CZ level. Figure generated on 17 Nov 2020 at 16:23:36. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

With individual controls

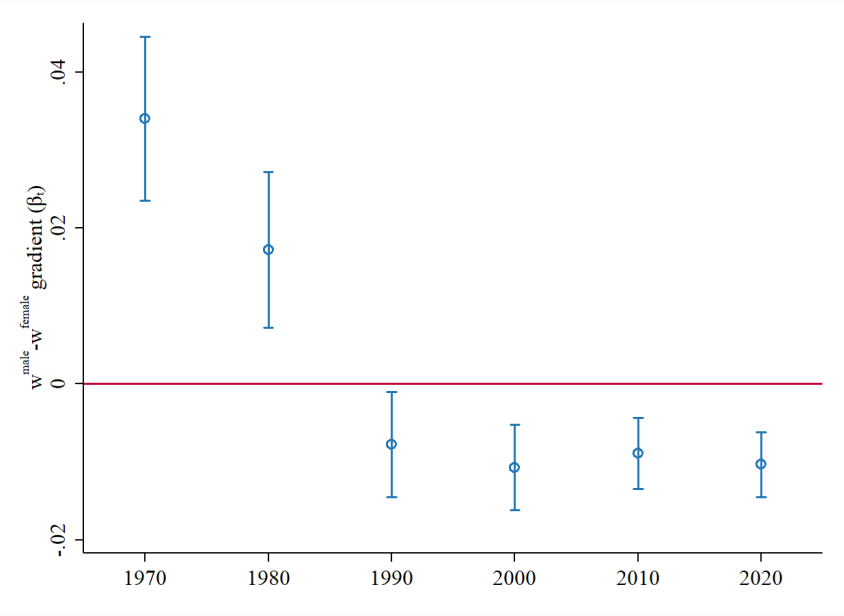
# Robustness

The change in the gradient is robust to:

- Limiting regressions to the biggest CZ biggest CZ.
- Using weighting schemes commonly used in the literature (estimates become much more imprecise) population weighting
- Using only within-region / within-state variation adding FE
- Adjusting wages by: age, education, foreign born dummy, and education residualizing wages.
- Using log of CZ population as independent variable population as density measure .

# Robustness

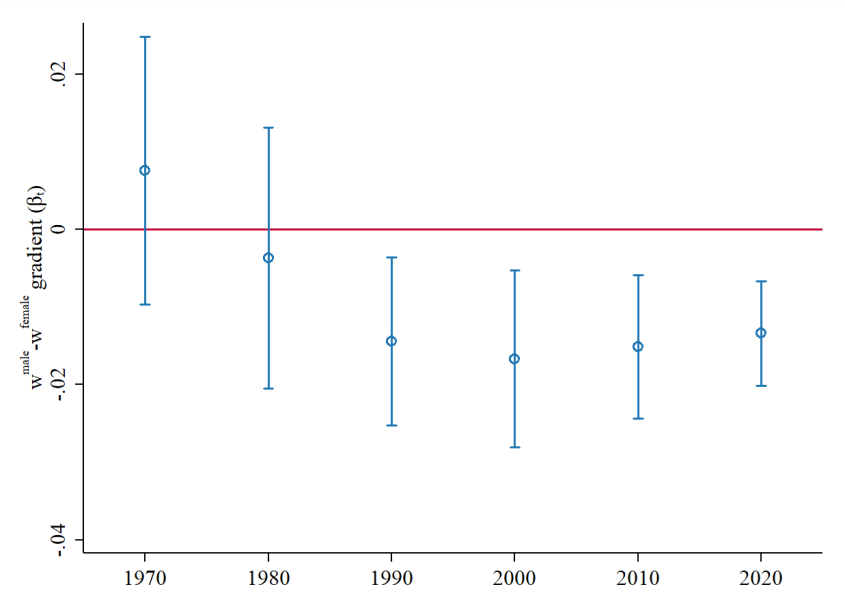
**Figure 7:** Coefficient on population density $\beta_t$ for largest CZ




**Note:** figure restricts to more than 2 people per km$^2$ in 1950. Bars show 95% confidence intervals. Standard errors clustered at the CZ level. Figure generated on 17 Nov 2020 at 13:42:04. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

With individual controls

**Figure 8:** Coefficient on population density $\beta_t$ (population weighted)

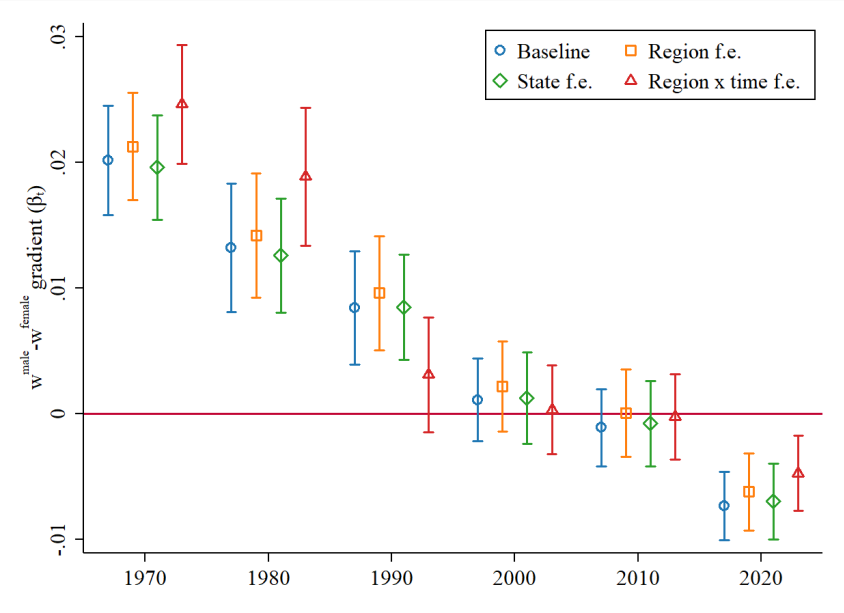



**Note:** figure restricts to more than 1 people per km$^2$ in 1950. Regressions weighted by the CZ population in 1970. Bars show 95% confidence intervals. Standard errors clustered at the CZ level. Figure generated on 17 Nov 2020 at 13:42:03. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

Graph without absorb

# Robustness

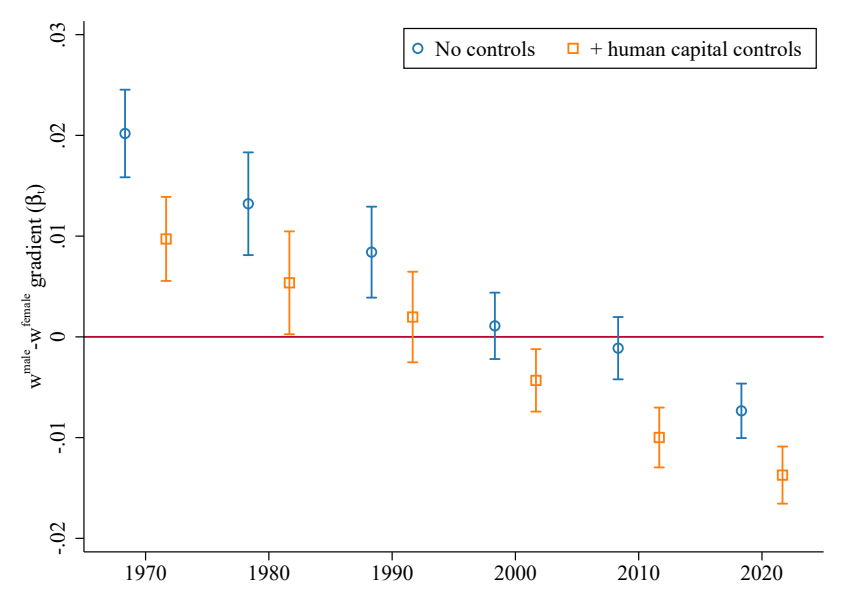
## **Figure 9:** Coefficient on population density $\beta_t$ adding fixed effects




**Note:** figure restricts to more than 1 people per km$^2$ in 1950. Bars show 95% confidence intervals. Standard errors clustered at the CZ level. Figure generated on 17 Nov 2020 at 13:42:05. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

With individual controls

## **Figure 10:** Coefficient on population density $\beta_t$ controlling for worker characteristics




**Note:** figure restricts to CZ with more than 1 people per km$^2$. The regressions are done on data aggregated at the CZ level. Bars show 95% robust confidence intervals. Standard errors clustered at the CZ level. Figure generated on 17 Nov 2020 at 13:42:05. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

Graph without absorb

# The gradient decline in perspective: several benchmarks

- Translate coefficients in terms of gap from small vs densest CZ
- Coefficients are percent of gender gap sd.
- Change in coefficients in terms of urban wage premium.

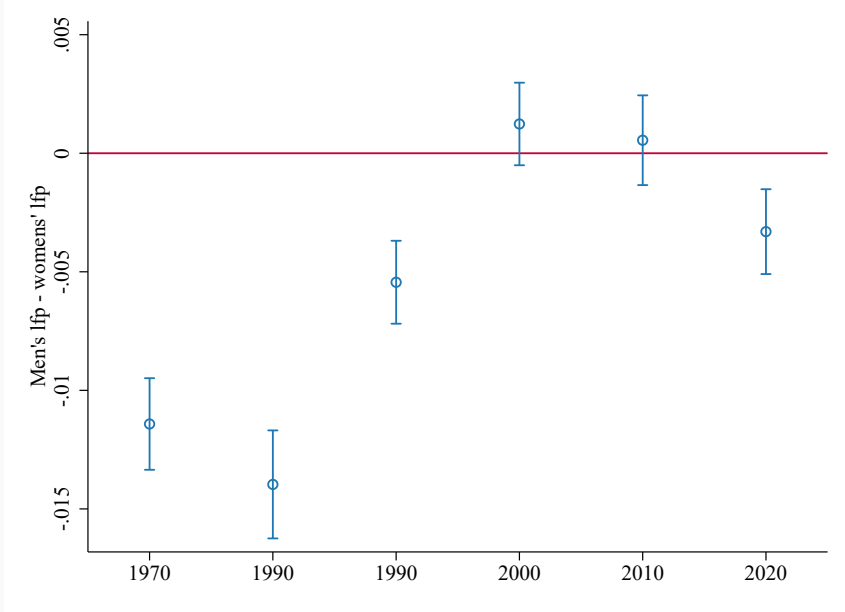
# A tale of women's success and male decline in denser labor markets

**Figure 11:** Coefficient on population density $\beta_t$


**Note:** figure restricts to CZ with more than 1 people per km$^2$. Bars show 95% confidence intervals. Standard errors clustered at the CZ level. Figure generated on 17 Nov 2020 at 09:25:19. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

Graph without absorb

**Figure 12:** Coefficient on population density $\beta_t$

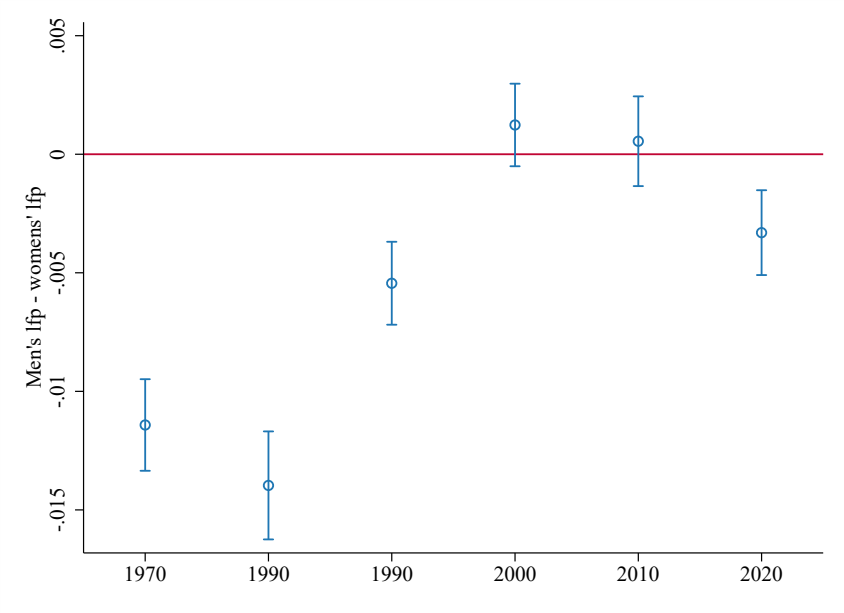



**Note:** figure restricts to CZ with more than 1 people per km$^2$. Bars show 95% confidence intervals. Standard errors clustered at the CZ level. Figure generated on 17 Nov 2020 at 09:25:20. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

Weighted regressions | With individual controls

# A tale of women's success and male decline in denser labor markets

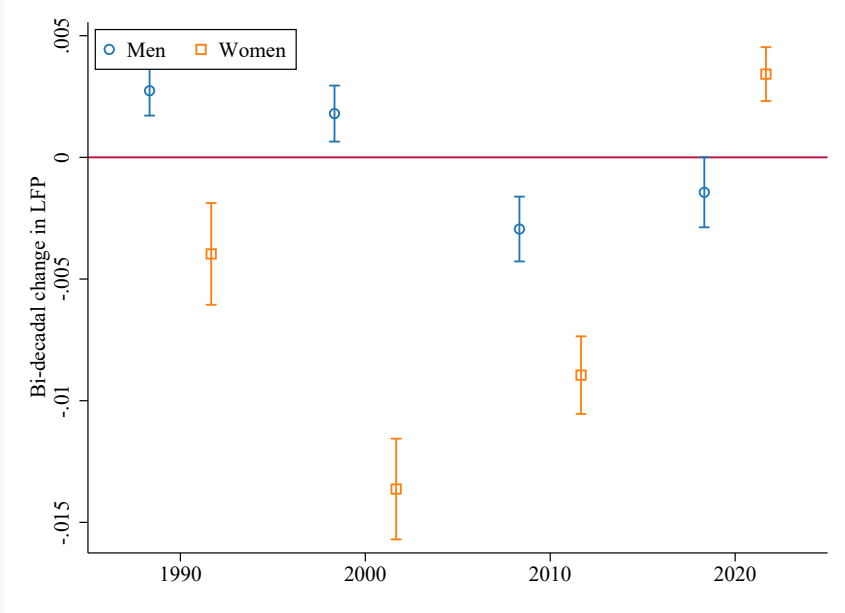
**Figure 13:** Coefficient on population density $\beta_t$




**Note:** figure restricts to CZ with more than 1 people per km$^2$. Bars show 95% confidence intervals. Standard errors clustered at the CZ level. Figure generated on 8 Nov 2020 at 13:34:16. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

Graph without absorb

**Figure 14:** Coefficient on population density $\beta_t$




**Note:** figure restricts to CZ with more than 1 people per km$^2$. Bars show 95% confidence intervals. Standard errors clustered at the CZ level. Figure generated on 17 Nov 2020 at 09:25:21. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

Weighted regressions | With individual controls

# What can explain the decline in the gradient?

- Increased sorting of high-skill women in the densest CZ.
- Rise of women college education
- CZ industrial structure.

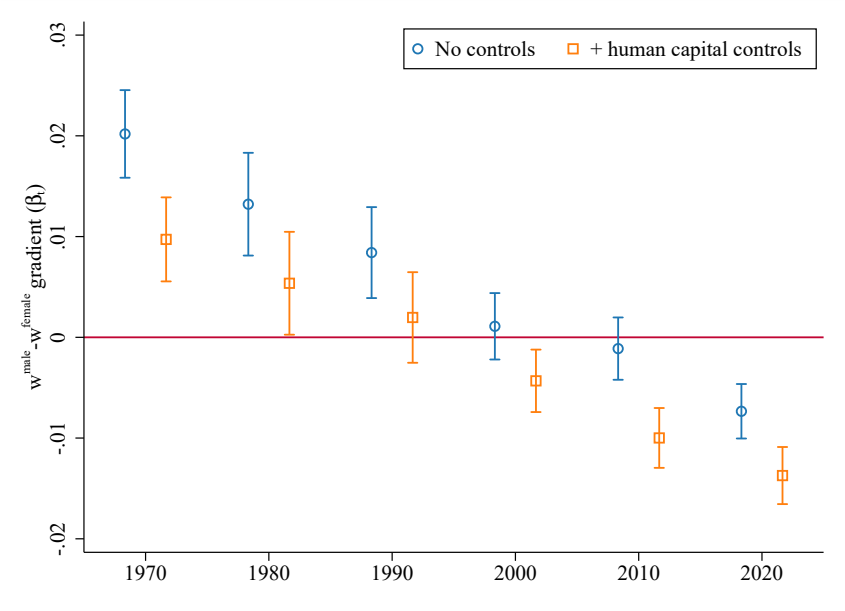
# Industrial structure

- Increased sorting of high-skill women in the densest CZ.
- Rise of women college education
- CZ industrial structure.

# Not readily explained by sorting on human-capital variables

- If high-skill women increasingly sort into into denser labor markets (relative to men) $\Longrightarrow$ gender gaps in denser CZ will decrease faster over time.
  - Decrease in the gradient should disappear once education is taken into account.
  - Not supported by the data.

**Figure 15:** Coefficient on population density $\beta_t$ controlling for worker characteristics



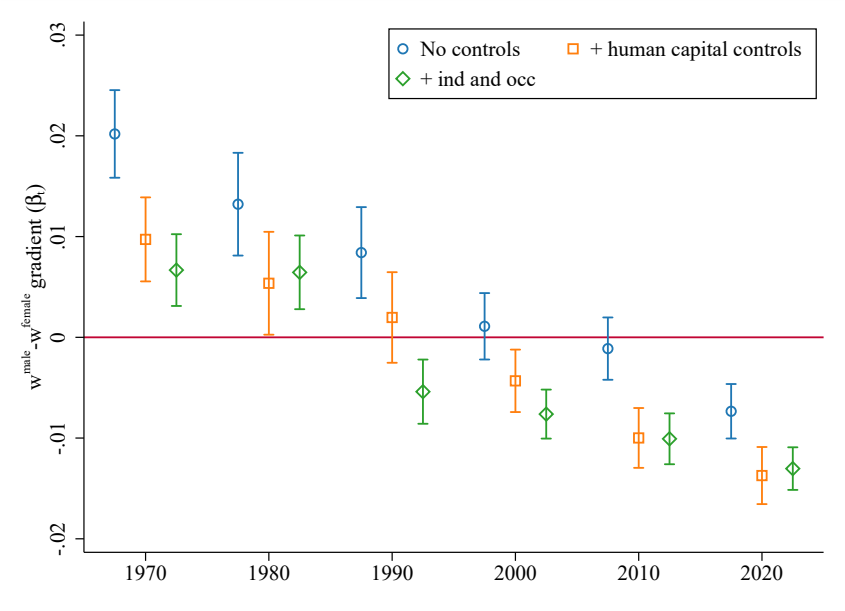

**Note:** figure restricts to CZ with more than 1 people per km$^2$. The regressions are done on data aggregated at the CZ level. Bars show 95% robust confidence intervals. Standard errors clustered at the CZ level. Figure generated on 17 Nov 2020 at 13:42:05. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

# Change in the gender gap gradient is concentrated in workers without a college degree

# The role of industries

- My main takeaway: changes between 1970-1980 accounted by industry f.e. $\Longrightarrow$ women were more able to enter into high paying occupations in denser labor markets.

**Figure 16:** Coefficient on population density $\beta_t$ controlling for worker characteristics




**Note:** figure restricts to CZ with more than 1 people per $km^2$. The regressions are done on data aggregated at the CZ level. Bars show 95% robust confidence intervals. Standard errors clustered at the CZ level. Figure generated on 17 Nov 2020 at 09:33:10. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

# Is it polarization of the occupational structure?

# Not explained by increase in inequality

# Fact 1: there are persistent differences in the level of the gender gap across CZ

**Table 1:** CZ-level gender gap statistics

| | Census year | | | | | |
|---|---|---|---|---|---|---|
| **Statistic** | **1970** | **1980** | **1990** | **2000** | **2010** | **2020** |
| Average gap | 0.44 | 0.41 | 0.33 | 0.26 | 0.21 | 0.19 |
| Standard deviation | 0.07 | 0.08 | 0.06 | 0.05 | 0.05 | 0.05 |
| **Distribution** | | | | | | |
| p90 | 0.53 | 0.51 | 0.40 | 0.32 | 0.26 | 0.25 |
| p75 | 0.49 | 0.47 | 0.37 | 0.29 | 0.24 | 0.22 |
| Median | 0.44 | 0.41 | 0.33 | 0.26 | 0.21 | 0.19 |
| p25 | 0.40 | 0.36 | 0.29 | 0.23 | 0.17 | 0.16 |
| p10 | 0.35 | 0.32 | 0.26 | 0.20 | 0.15 | 0.13 |

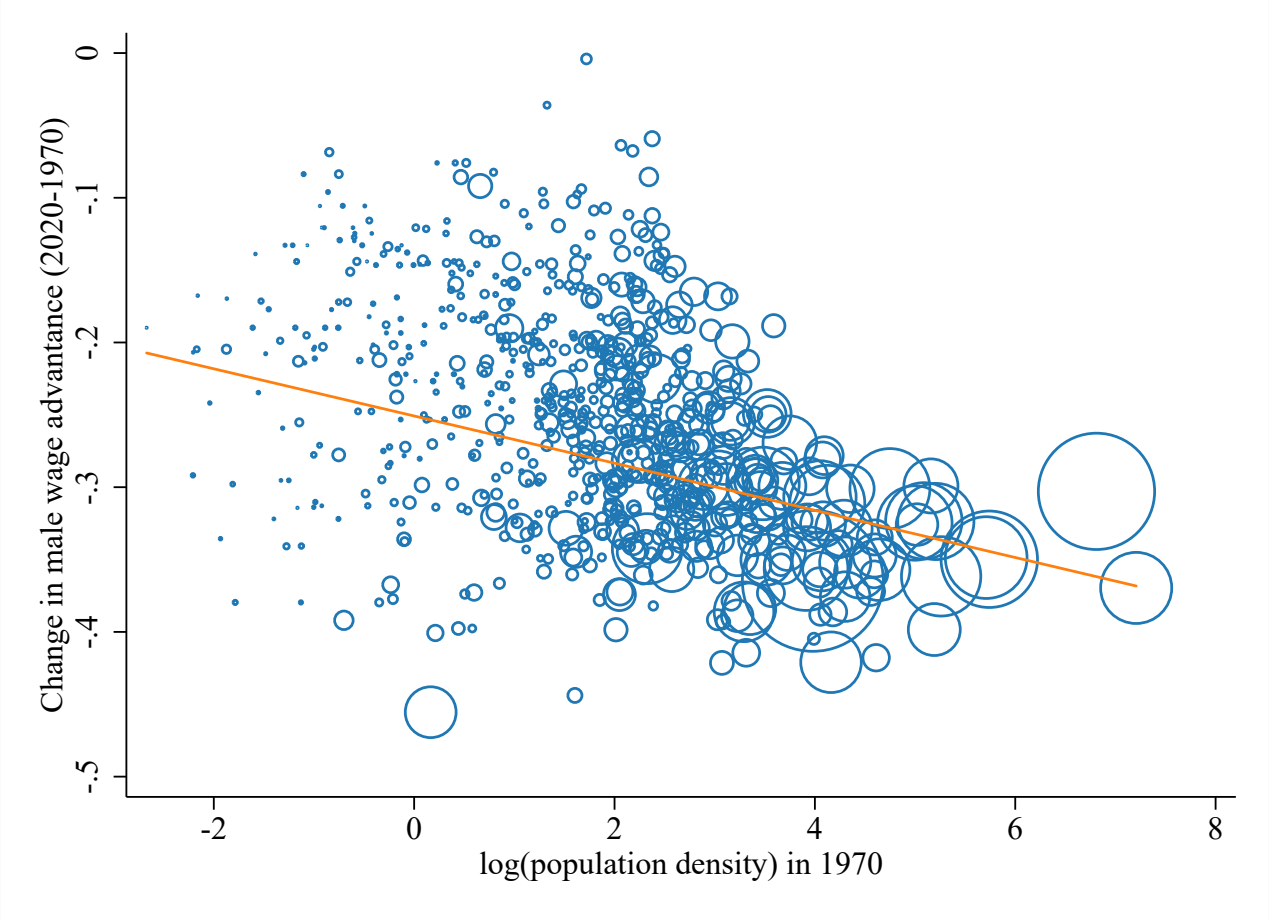
**Persistence:** 20-year auto-correlation coefficient $> 50\%$.

Geographical pattern

# Fact 2: there is wide variation in the change of gender wage gap across CZ

# Fact 3: decline of the gender gap was faster in denser CZ

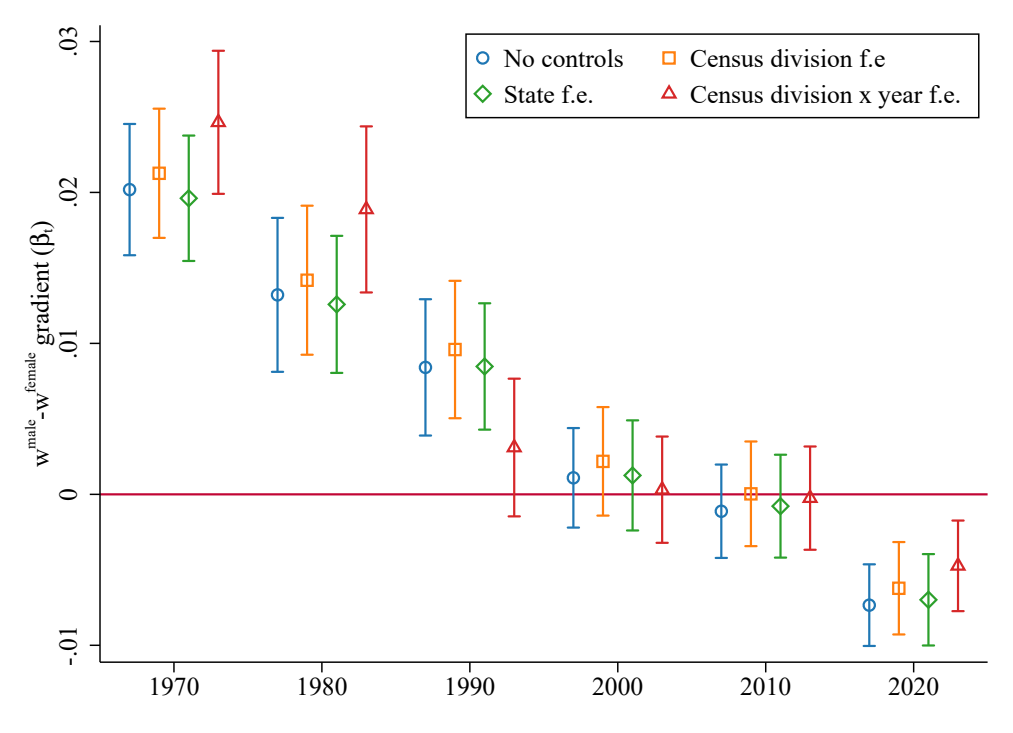
**Figure 17:** Change in male wage advantage in US CZ



# Fact 4: the gender gap - density gradient has inverted

**Regression specification:** $w_{rt}^{men} - w_{rt}^{women} = \alpha_{rt} + \beta_t \ln(density)_{rt} + \ldots$

**Figure 18:** Coefficient on population density $\beta_t$




**Note:** figure restricts to CZ with more than 1 people per km$^2$. Bars show 95% confidence intervals. Standard errors clustered at the CZ level. Figure generated on 19 Oct 2020 at 19:41:25. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

# The change in the gender-gap gradient is big

- 
- Men vs women urban wage premium

This change in the gradient:

- is not driven **permanent** regional differences across the US [graph adding the fixed effects].
- is not driven any single CZ [1970 vs 2020 graph].
- is not present when looking at race race graphs.
- is also present when I zoom in on the 240 most dense labor markets [limit to the 40% most dense labor markets]
- also appears when including part-time workers in the sample

On coefficient size  Distribution illustration  Within-group graphs

# What is driving these pattern? A mix of women's progress and men decline

There are two distinct periods:

- **1970-1990:** both sexes gain in denser CZ, but women's gains are larger.
- **1990-2010:** changes in the gradient are driven by men's decline in denser labor markets.

# How big are these coefficients?

**Table 2:** Male advantange changes implied by estimated elasticities

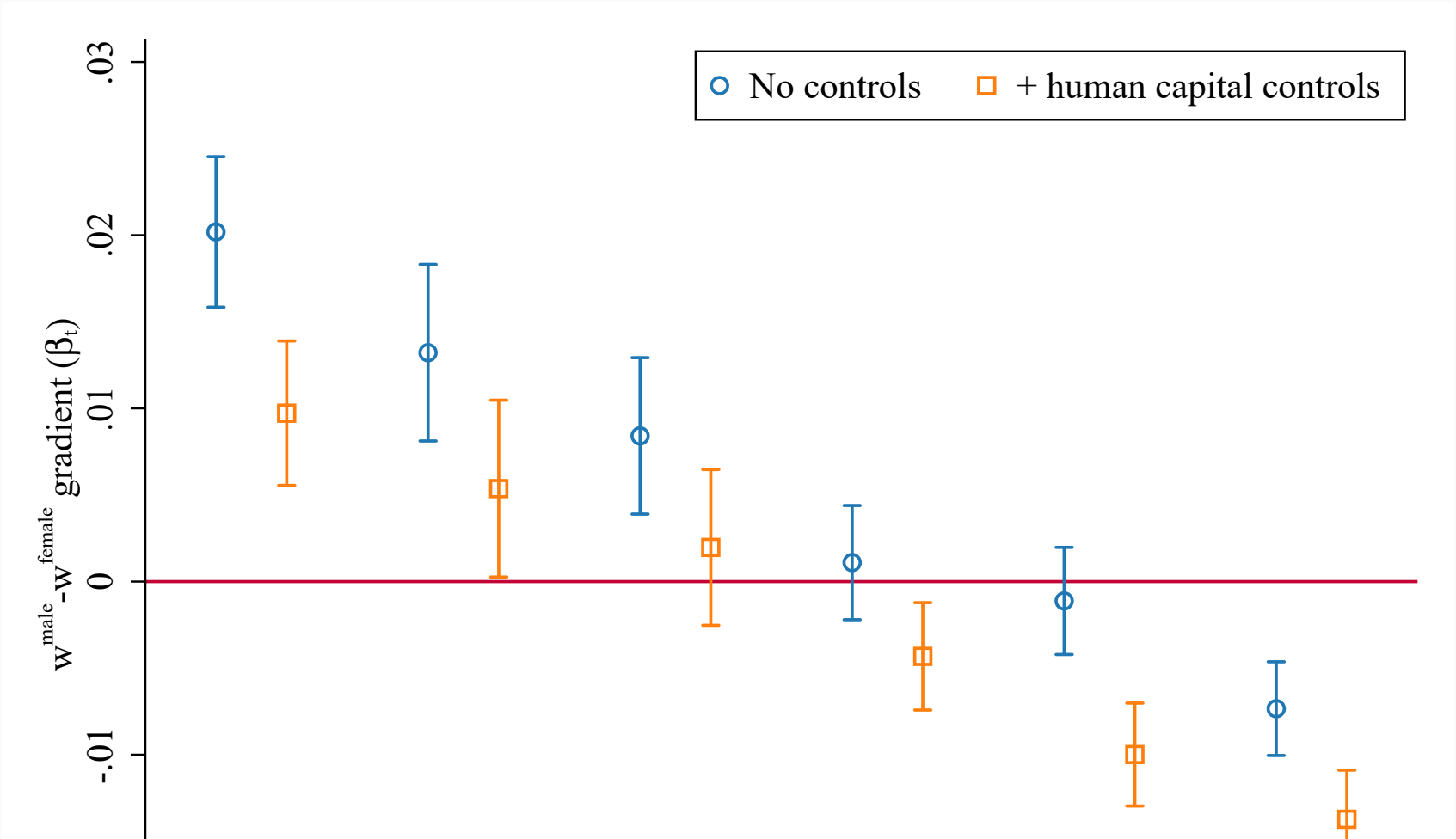
| | 1970 | 1980 | 1990 | 2000 | 2010 | 2020 |
|---|---|---|---|---|---|---|
| Density elasticity ($\beta$) | 0.020 | 0.013 | 0.008 | 0.001 | -0.001 | -0.007 |
| s.d. wage gap | 0.073 | 0.077 | 0.060 | 0.049 | 0.049 | 0.050 |
| $\beta/sd$ | 0.278 | 0.173 | 0.141 | 0.022 | -0.023 | -0.146 |
| IC range | 0.029 | 0.019 | 0.013 | 0.002 | -0.002 | -0.012 |
| (% mean gap) | 0.065 | 0.047 | 0.040 | 0.007 | -0.009 | -0.064 |
| 90 - 10 pctile range | 0.061 | 0.040 | 0.027 | 0.004 | -0.004 | -0.025 |
| (% mean gap) | 0.137 | 0.097 | 0.082 | 0.014 | -0.018 | -0.133 |

*Note:* changes based on unweighted estimated elasticities. Sample restricted to full-time year-round workers. Table generated on 28 Sep 2020 at 15:15:18.

# What can account for the change in the density-gradient?

**Regression specification:** $w_{rt}^{men} - w_{rt}^{women} = \alpha_{rt} + \beta_t \ln(density)_t$

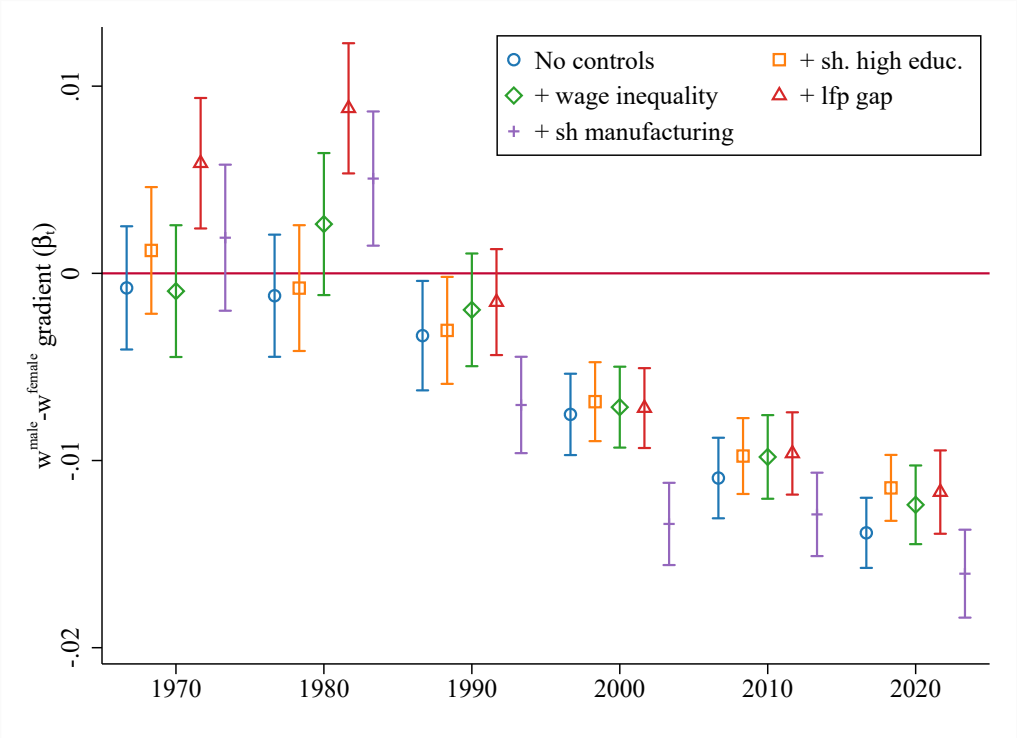
**Figure 19:** Coefficient on population density $\beta_t$ controlling for worker characteristics



# Adding czone-level variables

**Regression specification:** $w_{rt}^{men} - w_{rt}^{women} = \alpha_{rt} + \beta_t \ln(density)_t + X_{rt}\gamma_t$

**Figure 20:** Coefficient on population density $\beta_t$ controlling for worker characteristics




**Note:** figure restricts to CZ with more than 1 people per km$^2$. Regression includes census division $\times$ year fixed-effects. Additional controls include number of children, marital status and being a female head of household. The regressions are done on data aggregated at the CZ level. Bars show 95% robust confidence intervals. Standard errors clustered at the CZ level. Figure generated on 19 Oct 2020 at 19:41:29.