Local labor markets, population density and the gender gap

César Garro-Marín

October 12, 2020

Boston University

Introduction

Summary

In the next slides I document three main facts about the **gender gap** in the US for the period of 1970 and 2020:

1. There is a large dispersion in the **level** of the gender wage gap across labor markets in the US. The dispersion persists despite the general decrease in the level of the gap since 1970.
2. There are differences in the **change** of the gender wage gap. The largest reductions happened in densest labor markets.
3. The relationship between the **level** of gender wage gap and population density has inverted over the period. Today, the densest labor markets have a lower gender wage gap.

Data

Data

Source: IPUMS data for:

- 1950-2000 Decennial censuses.
- 5-year ACS for the years 2011 and 2018. For ease of presentation I label these datasets as 2010 and 2020 respectively.

Sample includes all full-time year-round workers whom:

- Aged 18-64.
- Not attending school.
- Not living in group quarters.
- For all graphs I limit the sample to people living in CZ with a population density of at least 1 person per-square kilometer in 1950.

Empirical facts

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

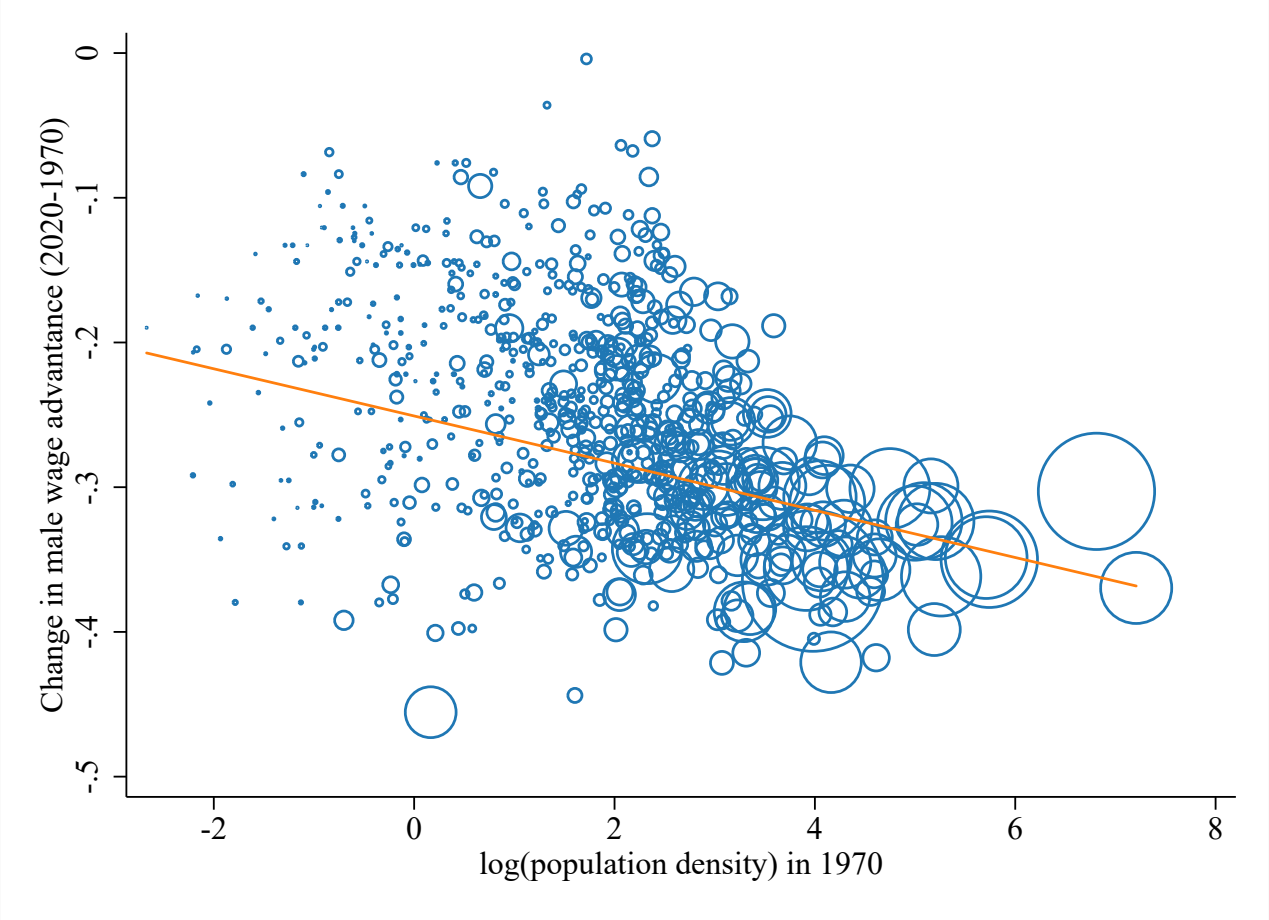
Persistence: 20-year auto-correlation coefficient $> 50\%$.

Geographical pattern Persistence

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

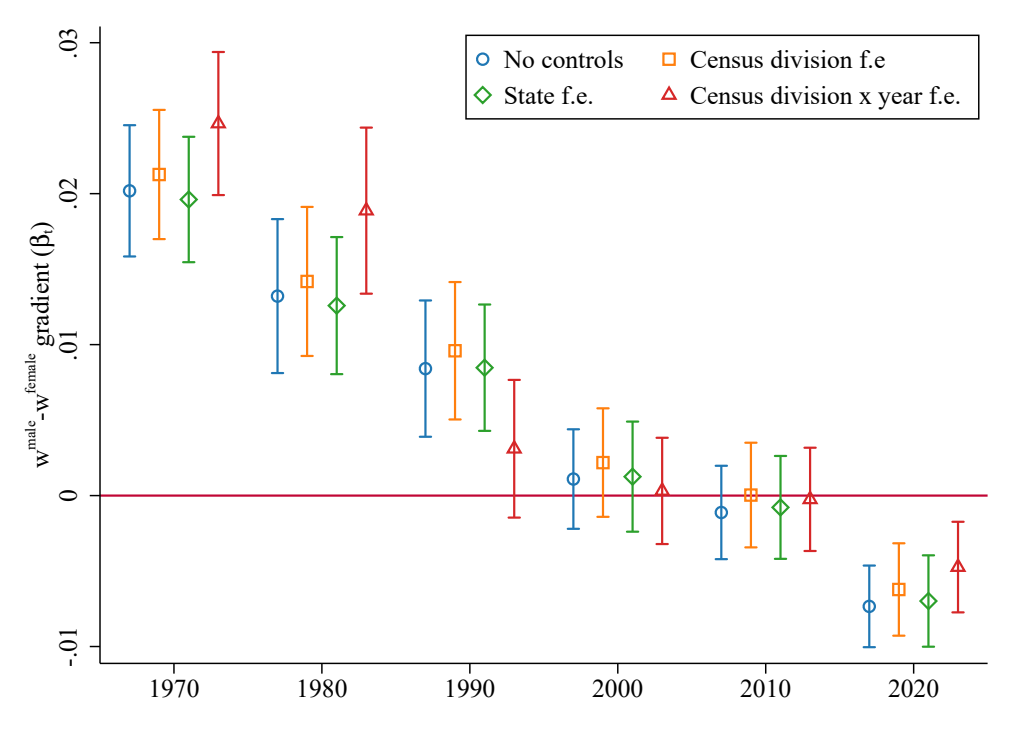
Figure 1: Change in male wage advantage in US CZ



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 2: Coefficient on population density β_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 12 Oct 2020 at 18:49:38. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

On coefficient size | Distribution illustration | Is this about gender? | Within-group graphs

How big are these coefficients?

Table 2: Male advantange changes implied by estimated elasticities

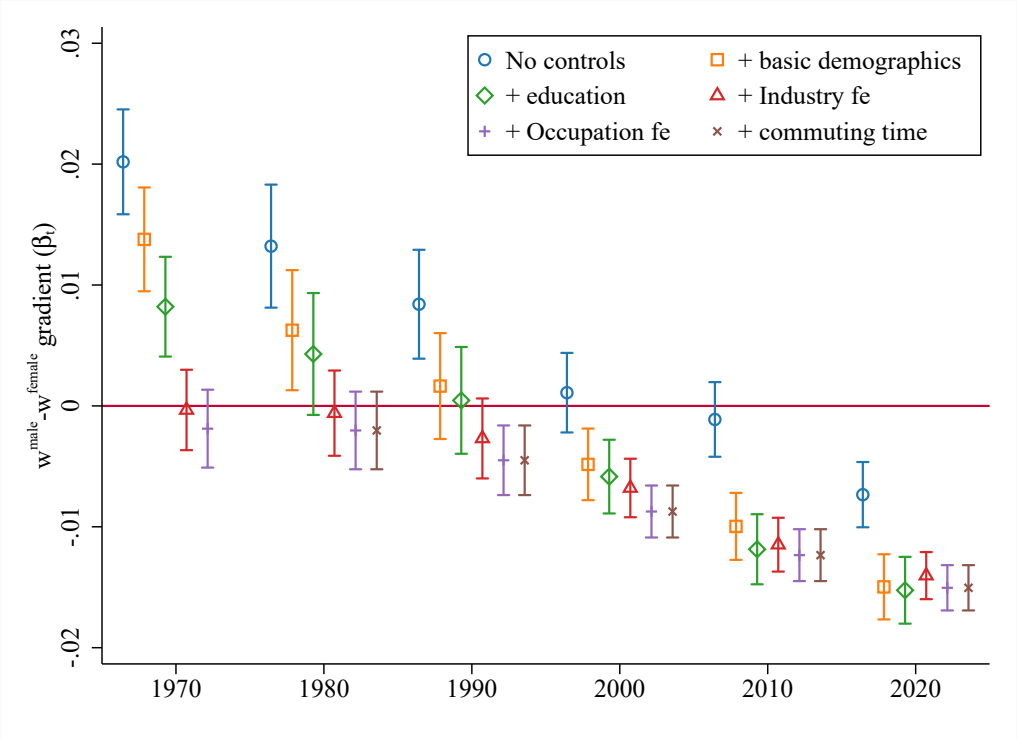
	1970	1980	1990	2000	2010	2020
Density elasticity (β)	0.020	0.013	0.008	0.001	-0.001	-0.007
s.d. wage gap	0.073	0.077	0.060	0.049	0.049	0.050
β/sd	0.278	0.173	0.141	0.022	-0.023	-0.146
IC range	0.029	0.019	0.013	0.002	-0.002	-0.012
(% mean gap)	0.065	0.047	0.040	0.007	-0.009	-0.064
90 - 10 pctile range	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: $w_{rt}^{men} - w_{rt}^{women} = \alpha_{rt} + \beta_t \ln(density)_t$

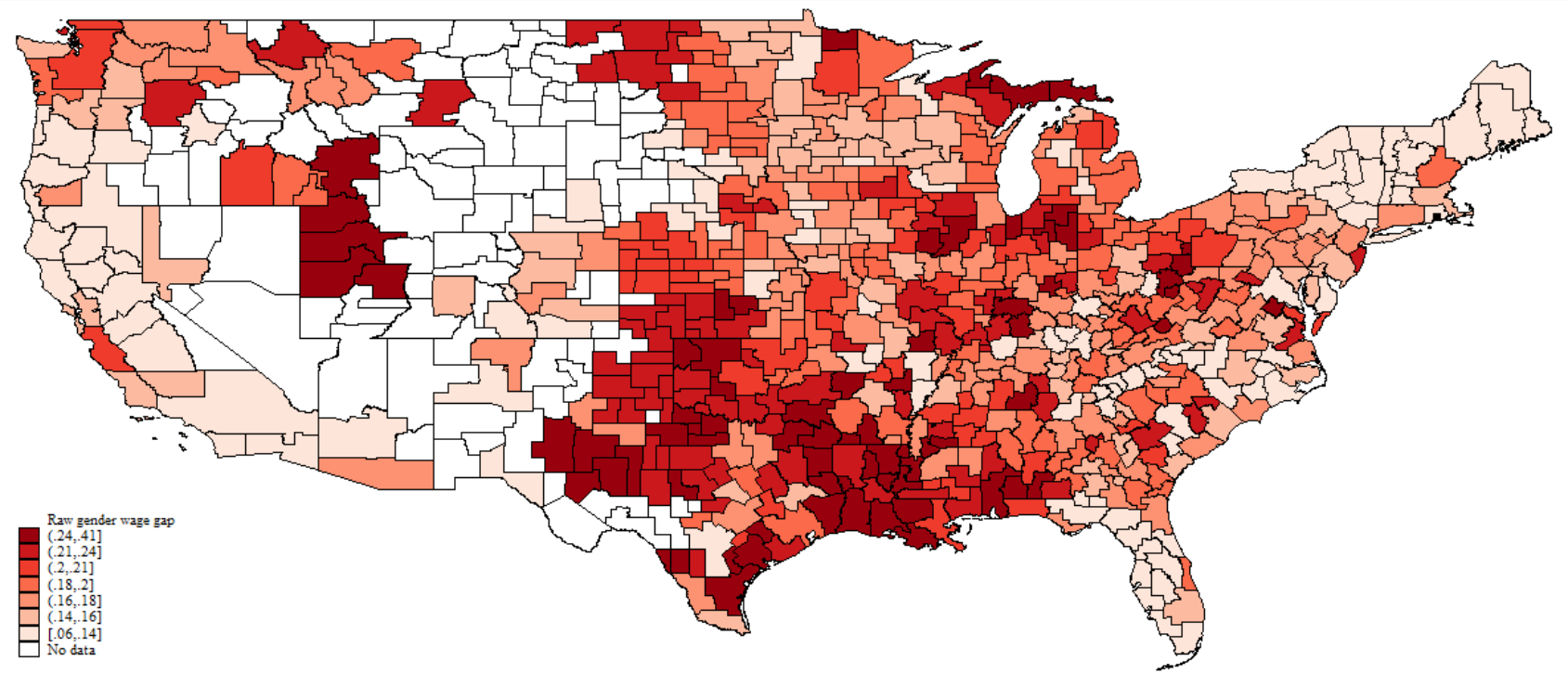
Figure 3: Coefficient on population density β_t controlling for worker characteristics



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 7 Oct 2020 at 10:48:33. Figure generated using the dofile 2_analysis/code_files/write_regression_coefplots.do.

The geography of the gender gap in 2020

Figure 4: The gender gap in the US in 2020

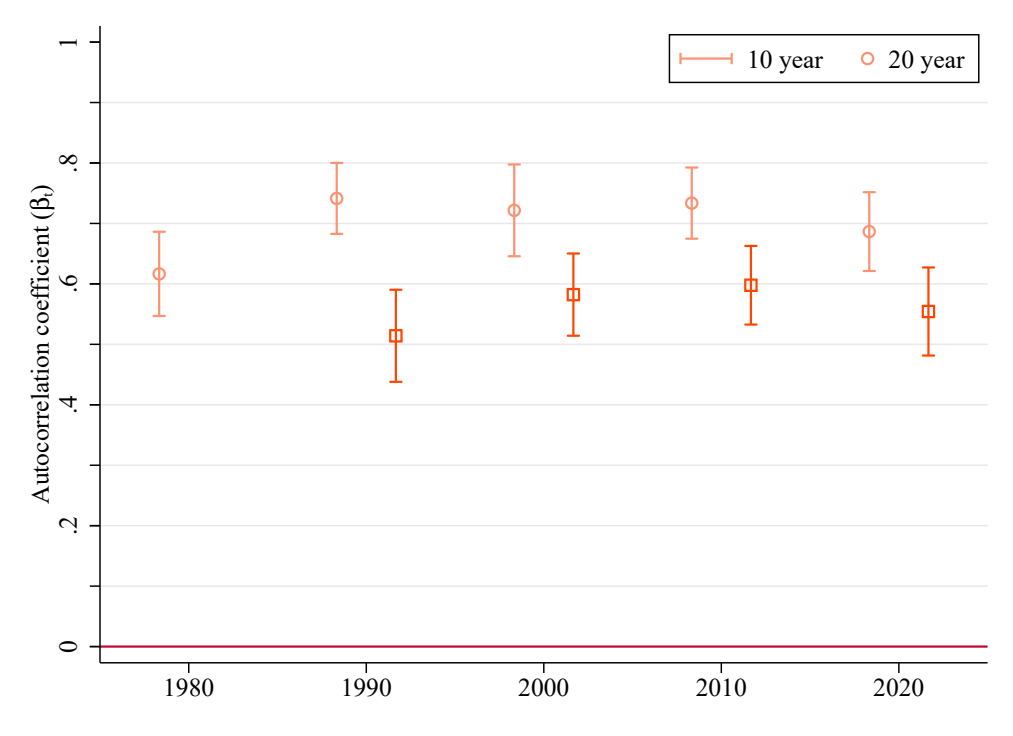


Note: darker colors denote higher relative wages for men. Figure restricts to czones with population densities above 1 person per km^2 and full-time year-round workers.

Return

20-year auto-correlation coefficient is above 50%

Regression specification: $w_{rt}^{men} - w_{rt}^{women} = \alpha_{rt} + \beta_t(w_{rt-j}^{men} - w_{rt-j}^{women})$



Note: figure restricts to CZ with more than people per km^2 and full-time year-round workers.. Bars show 95% robust confidence intervals. Standard errors are clustered at the CZ level. Dependent and independent variables are standardized

Return

Residualization procedure

1. Run the regression on individual level data:

$$wage_{igrt} = X_{igrt}\gamma_t + \lambda_{grt} + \varepsilon_{igrt}$$

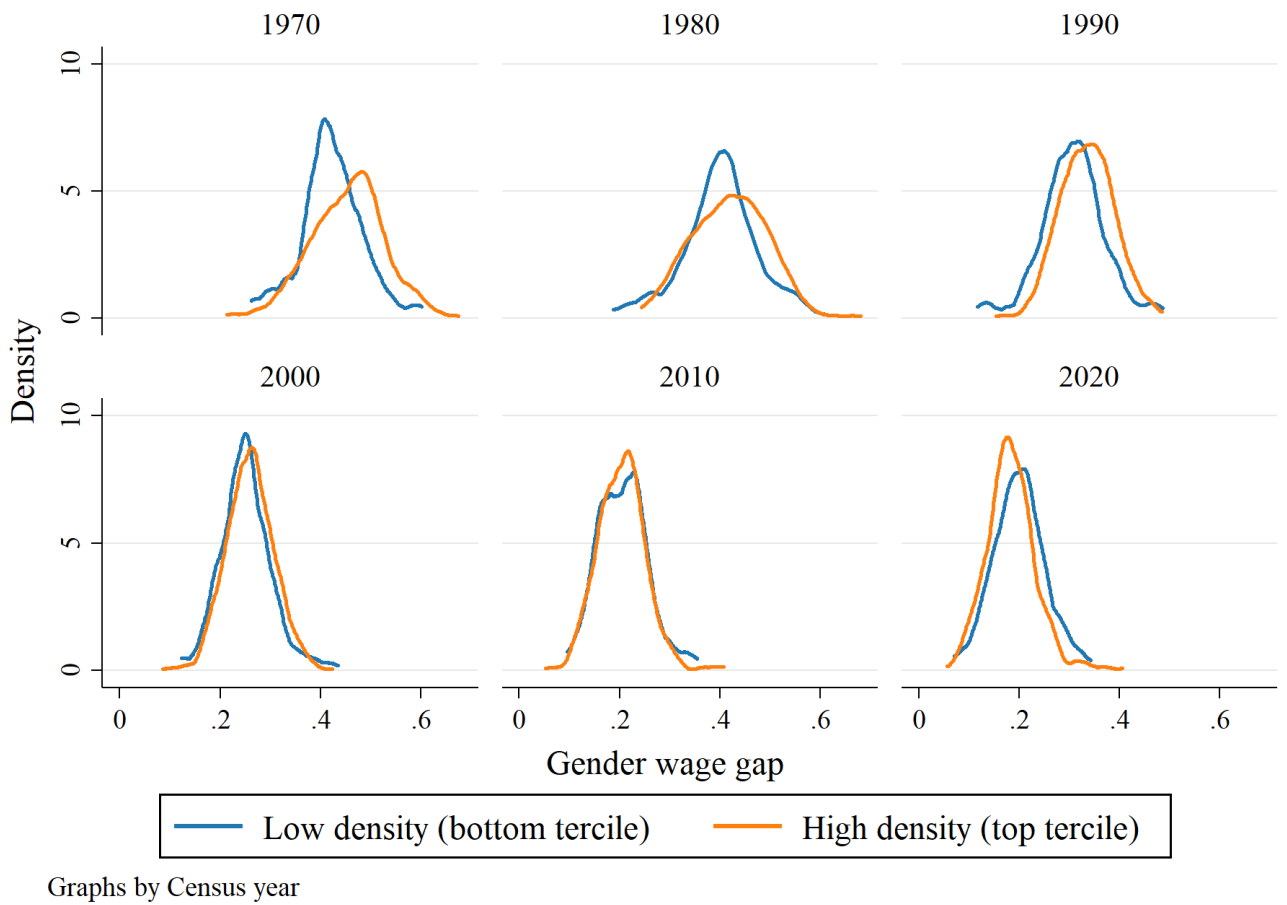
   where i, g, r, t index individual, sex, CZ and decade respectively. I impose the same return on individual level characteristics across sex and CZ.

2. Run the following regression at the CZ level:

$$\lambda_{mrt} - \lambda_{frt} = \alpha_t + \beta_t \ln(density)_{rt}$$

   no weight is imposed on the CZ-level regressions (Solon et al., 2015). Return

Low vs high density CZ



Note: figure restricts to CZ with more than 1 people per km^2. Figure generated on 28 Sep 2020 at 15:56:45. Figure generated using the dofile code_files/kernel_density_movement.do.

Return

References

Solon, G., Haider, S. J., and Wooldridge, J. M. (2015). What are we weighting for? *Journal of Human Resources*, 50(2):301–316.