# Local labor markets, population density and the gender gap

César Garro-Marín

September 18, 2020

Boston University

# Basic regressions

# Accounting for individual characteristics

**Route A**:

1. Run on individual-level data:

$$wage^g_{irt} = X^g_{iry}\gamma_t + \lambda^g_{rt} + \varepsilon^g_{irt}$$

2. In a second stage run:

$$\hat{\lambda}^{male}_{rt} - \hat{\lambda}^{female}_{rt} = \tau_t + \beta_t \log(density)_{rt} + \varepsilon^g_{irt}$$

# Accounting for individual characteristics

**Route B**:

If wages are determined at the individual level by the model:

$$w_{irt}^{g} = X_{irt}^{g}\gamma_t + \tau_t male_i + \varepsilon_t$$

By aggregating at the CZ level this model becomes:

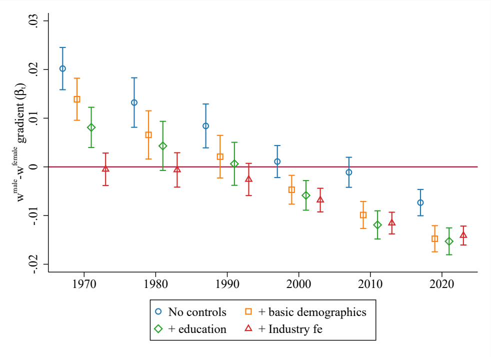
$$w_{rt}^{male} - w_{rt}^{female} = \tau_t + (\bar{X}_{rt}^{male} - \bar{X}_{rt}^{female})\gamma_t + u_t$$

Thus I run regressions of the form:

$$w_{rt}^{male} - w_{rt}^{female} = \tau_t + (\bar{X}_{rt}^{male} - \bar{X}_{rt}^{female})\gamma_t + \beta_t \log(density)_{rt} + u_t$$

# Route A

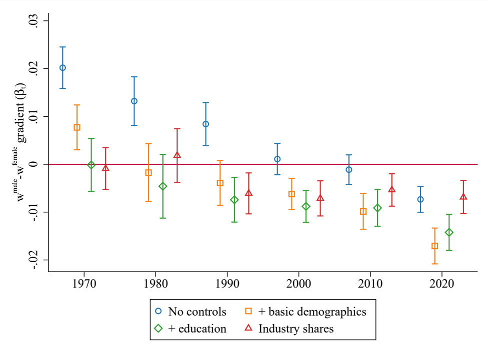
**Figure 1:** Coefficient on population density $\beta_t$ controlling for worker characteristics




**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.

# Route B

**Figure 2:** 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.