Decomposing Trends in U.S. Wages: Inequality Within and Between Migrants and Stayers by Levels of Education

Bent Sorensen [1], Fan Wang [1] Adi Zytek [2]

[1]University of Houston [2]Rice University

Background & Contribution

- We propose a framework for documenting the contribution of a demographic group to changes in income inequality over time
- Two methods have been predominant in recent income inequality literature: **variance decomposition** (VD) and **unconditional quantile regression** (UQR) (Firpo 2009)
- We show that variance decomposition *underestimates* a group's contribution to inequality when within-group income distributions are heavy tailed; and that UQR *overestimates* the contribution of heavy tailed group's and derive a formula for the exact marginal effect
- We apply our framework to document the contribution of six education-migration groups on US income inequality from 2001-2019, taking wages each year as exogenous

Proposed Framework

- Total Effect: total impact on inequality of a group; sensitive to within-group and between-group variance and relative proportion of a group
- Mean Effect: impact of a group's mean wage on inequality; sensitive to between-group variance and relative proportion, insensitive to within group variance
- Marginal Effect: impact of marginally increasing population share of a particular group; sensitive to within and between group variance, insensitive to relative proportion

We use the **interdecile range**—the gap between the 90th and 10th percentile of the distribution — as a measure of inequality

Connection to Variance Decomposition

For given year, the variance of wages nationally, σ^2, can be decomposed as:

$$\sigma^2 = \sum_{m\in M}\sum_{e\in E}\omega_{e,m}\cdot\sigma^2_{e,m} + \sum_{m\in M}\sum_{e\in E}\omega_{e,m}\cdot(\mu_{e,m}-\mu)^2 \ ,$$

where $\sigma^2_{e,m}$ is the within group var. of edu-mig group $\{e,m\}$ and $\mu_{e,m}$ is the mean wage for $\{e,m\}$.

We can then derive the total and mean contributions of group e^*, m^* to overall variance as:

$$R^{total}_{e^*,m^*} = \omega_{e^*,m^*}\cdot(\sigma^2_{e^*,m^*}-\overline{\sigma}^2) + \omega_{e^*,m^*}\cdot(\mu_{e^*,m^*}-\mu)^2 \ , R^{mean}_{e^*,m^*} = \omega_{e^*,m^*}\cdot(\mu_{e^*,m^*}-\mu)^2 \ .$$

where $\overline{\sigma}^2$ is the total weighted within-group variance

Data

- ACS Census Survey, taken annually from 2001-2019; includes individuals residing in the US, excluding territories
- Analyze inequality on earned-income (wages) coming from wage, salary, commissions, & tips; All analysis is on age-controlled log-income figures, adjusted for inflation
- Limit sample to individuals working at least 36 hours per week
- Interested in two education groups (college , non-college) and three migration groups (domestic non-movers, domestic-movers, foreign-born)

Total, Mean, & Marginal Effects

Total & Mean Effects

- The variance decomposition accurately quantifies a group's mean effect only when within-group distributions are normally distributed. The two college educated-mover groups in particular have heavy tailed distributions
- The total effect can be calculated precisely as the effect on inequality of removing one group from the sample, i.e the effect of shifting $\omega_{e,m}$ to 0
- The mean effect can be calculated precisely by shifting one group's mean to the national mean, i.e. setting $\mu_{e,m} = \mu$.

Marginal Effects: Consider the case of two subgroups:

- F_1 is the distribution function of our subgroup of interest; F_2 to is a mixture of the other (five) subgroups; ω is the population weight of group 1
- The distribution of the entire population, F, is equivalent to: $F = \omega F_1 + (1-\omega)F_2$

The marginal effect is equivalent to the derivative of ω on the inverse distribution function, F^{-1}:

$$R^{marg}_{e^*,m^*} = \frac{dF^{-1}(\cdot)}{d\omega_{e^*,m^*}} = \frac{F_2(q_\tau)-F_1(q_\tau)}{\omega f_1(q_\tau)+(1-\omega)f_2(q_\tau)} \ . \quad (1)$$

This formula can be applied to any distribution for which f_1 and f_2 can be estimated

Unconditional Quantile Regression

Firpo (2009) provides one method for estimating this marginal effect for a generic distribution. For quantile τ and individual log wages Y_i, the re-centered influence function is

$$\text{RIF}(Y_i, q_\tau) = q_\tau + \frac{\tau - \mathbf{1}\{Y_i \le q_\tau\}}{f(q_\tau)} \ , \quad (2)$$

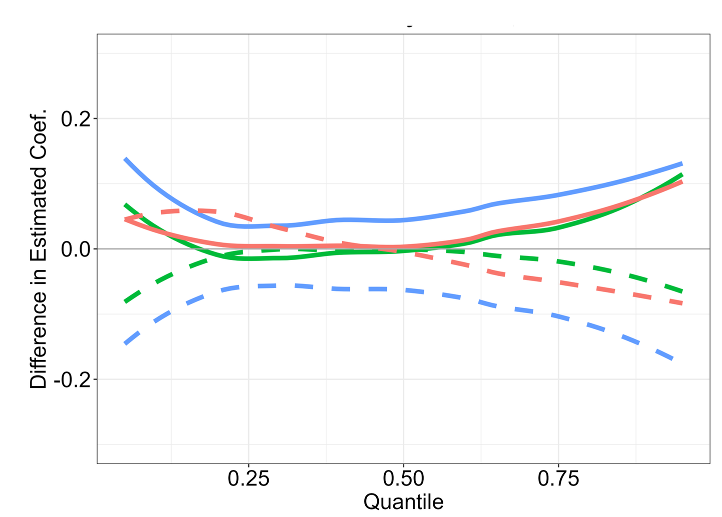
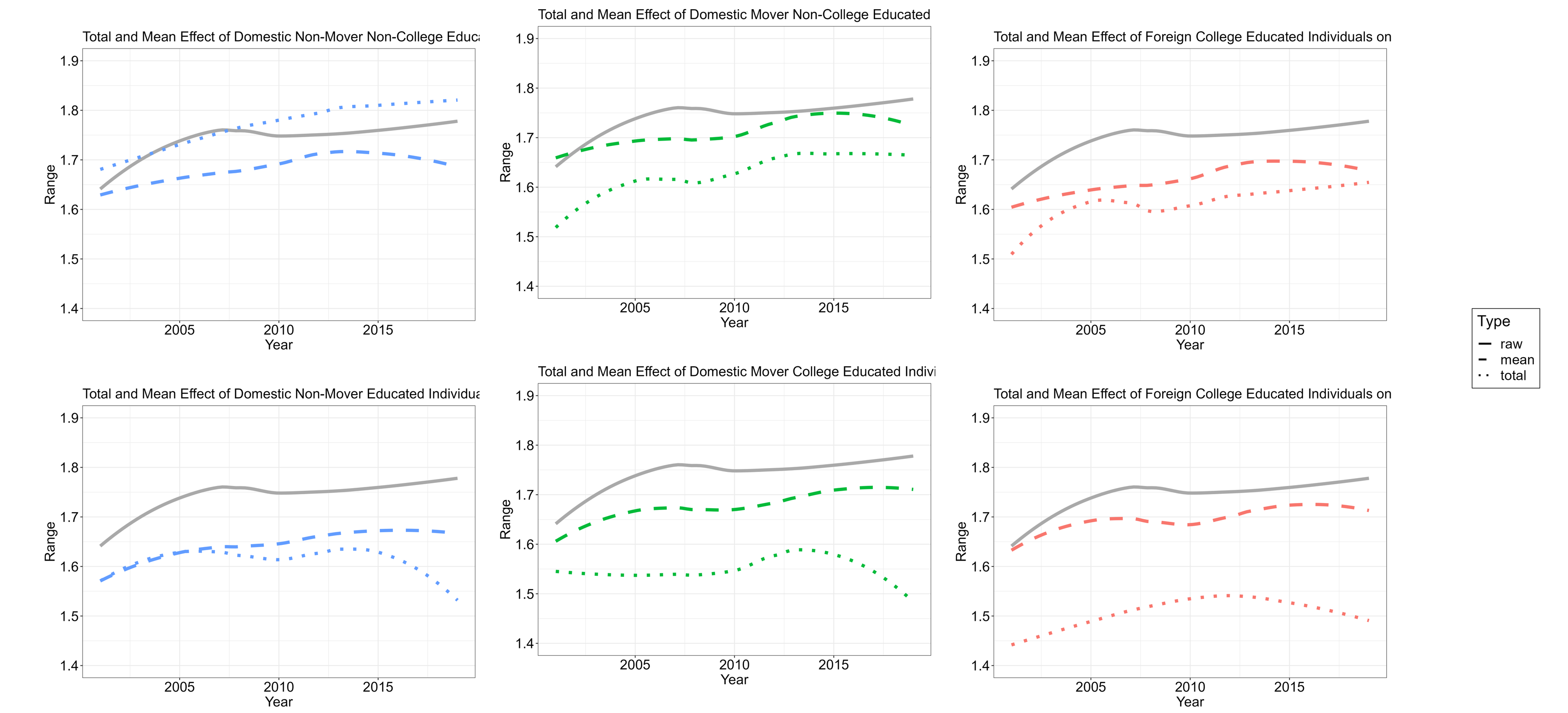


Figure 1. Difference Between Firpo RIF and Exact Estimate (2019)

- $\mathbf{1}\{Y_i \le q_\tau\} = 1$ if $Y_i \le \tau$th quantile q_τ.
- $f(q_\tau)$ is density function of wages. Estimated non-parametrically.
- Marginal effect obtained by regressing RIF(Y_i, q_τ) on group dummies and covariates
- We find that unconditional quantile regression overestimates the contribution of some groups to inequality for the inter-decile range
- Our method only requires distributional (quantile) information, whereas RIF requires large sample of points

Results

Total & Mean Effects:



Marginal Effects:

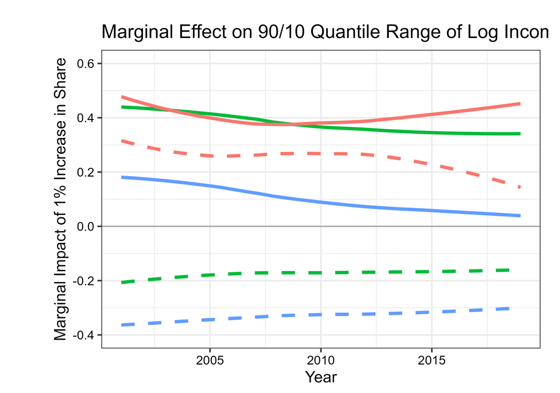


Figure 2. Marginal Effect of Each Group on 90/10 Range

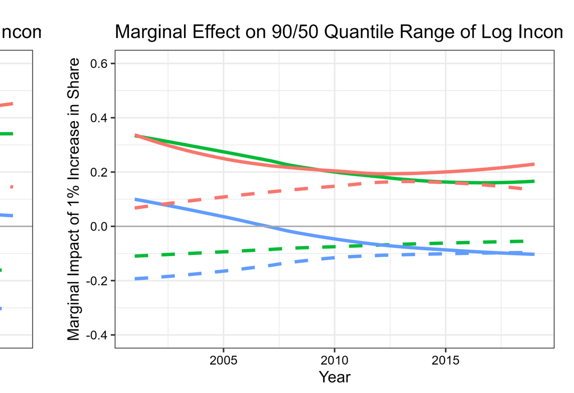


Figure 3. Marginal Effect of Each Group on 90/50 Range

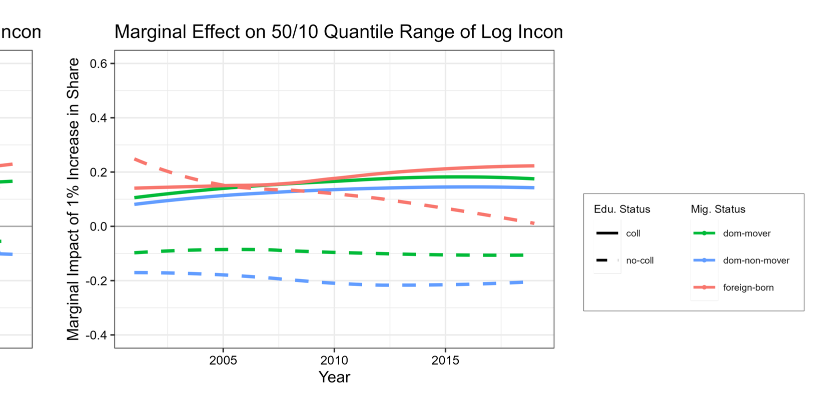


Figure 4. Marginal Effect of Each Group on 50/10 Range

Discussion

- Commonly used methods in inequality research, such as variance decomposition and unconditional quantile regression have limitations, especially when investigated inequality at very high and low quantiles
- College educated movers, domestic and foreign-born, had the greatest total and marginal effects on inequality, despite having the lowest relative population shares
- The marginal contribution of non-college educated foreigners has fallen over time, as their representation in the population has risen
- All education-migration groups have contributed to rising inequality

References

adizytek@gmail.com