IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14701

Between-Group Inequality May Decline despite a Rising Skill Premium

Imran Aziz
Guido Matias Cortes

AUGUST 2021

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14701

Between-Group Inequality May Decline despite a Rising Skill Premium

Imran Aziz
Yorkville University

Guido Matias Cortes
York University and IZA

AUGUST 2021

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Between-Group Inequality May Decline despite a Rising Skill Premium*

A vast literature aimed at understanding the nature and causes of wage inequality focuses on the skill premium as a key object of interest. In an environment where both the skill premium and the share of skilled workers are changing, however, the between-skill-group component of inequality may fall even as the skill premium rises – a pattern that is indeed observed in the U.S. and in many local labor markets during the 2010s. Understanding the evolution of the skill premium is therefore not always useful in terms of understanding why broad inequality measures are changing.

JEL Classification: J31, J21, J24

Keywords: skill premium, skill-biased technical change, between-group inequality

Corresponding author:
Guido Matias Cortes
Department of Economics
York University
4700 Keele Street
Toronto, ON M3J 1P3
Canada

E-mail: gmcortes@yorku.ca

* This research has been supported by the Social Sciences and Humanities Research Council of Canada.

1 Introduction

Models of Skill-Biased Technical Change (SBTC) have been very influential in the analysis of wage inequality. These models posit that a key driver of inequality is the increasing demand for skilled workers due to the development of new technologies (Tinbergen, 1974, 1975; Katz & Murphy, 1992; Juhn et al., 1993; Autor et al., 1998). If the growth in the demand for skilled workers outstrips the growth in the supply of these workers induced by rising educational attainment, then the skill premium (i.e. the wage gap between skilled and unskilled workers) will grow. Understanding the drivers of the skill premium is viewed, through the lens of these models, as being critical to our understanding of broader wage inequality patterns. A wide range of empirical work has been inspired by these models and has focused on the skill premium as a key outcome of interest.

The skill premium is, however, not a comprehensive measure of wage inequality. In this paper we highlight the fact that wage inequality – and in particular the "between-skill-group" component of wage inequality – may fall even as the skill premium is rising. This is due to the fact that overall between-group inequality depends not only on the wage gap between the two skill groups, but also on the size of each group. All else equal, a rise in the skill premium increases between-group inequality. However, a rise in the skill share, all else equal, will decrease between-group inequality if the share of skilled workers is high enough. Hence, in an environment where both the skill premium and the share of skilled workers are growing (as is the case in practice), a rising skill premium may be observed alongside a decline in between-group inequality. We show that over recent years in the U.S. the between-group component of inequality has indeed declined in absolute terms, even as the skill premium has continued to rise. The same pattern is observed in many local labor markets. Increases in the skill premium therefore do not provide clear insights for understanding the rise in overall wage inequality in the U.S. economy over recent years.

Our analysis uses data for the period 1980–2019 from the U.S. Census and the American Community Survey. We begin by verifying the well-documented rise in the skill premium and in overall wage inequality in the U.S. economy over this time period. We then perform a decomposition of the variance of log wages into a component related to differences between skill groups (i.e. those with at least some college education and those with at most a high school degree), and a component related to wage differences within skill groups. Naturally, the skill premium is unrelated to the within-group component, but one may expect that it would provide useful guidance on the evolution of the between-group component. This is not however the case. Over the most recent decade (2010-2019), in spite of the continued increase in the skill premium, between-group inequality actually fell.

In order to clarify why between-group inequality can fall even as the skill premium is rising, we analyze the statistical link between these two measures. Between-group inequality depends not only on the skill premium, but also on the relative size of the two skill groups.[1] Holding the skill composition constant, a rise in the skill premium indeed leads to more between-group inequality. On the other hand, changes in the share of skilled workers (holding the skill premium constant) have an ambiguous effect on between-group inequality. A rise in the share of skilled workers will lead to a rise in between-group inequality if the share of skilled workers is below 50%, and will lead to a decline in between-group inequality if the share is above this threshold. This threshold is empirically relevant, given that it was crossed in the U.S. economy during the 1980s. Indeed, between 2010 and 2019, the downward pressure on between-group inequality exerted by the rising share of skilled workers empirically dominated the upward pressure on between-group inequality exerted by the rising skill premium.[2]

Our conclusions about the lack of a direct empirical link between the skill premium and overall between-group inequality (in an environment where skill shares are changing) are valid beyond a measure based on the variance of log wages, and are also observed when using other standard measures of inequality such as the Gini or the Theil indices. We also illustrate the potentially diverging signs of the change in between-group inequality and the change in the skill premium using data at the local labor market level. We find that locations where the skill premium went up but between-group inequality fell between 2010 and 2019 account for around one-fifth of the total workforce in 2019.

The SBTC framework is considered to be the "canonical model" for understanding changes in wage inequality (Acemoglu & Autor, 2011). The model is conceptually attractive and has proven to be empirically successful, at least over certain time periods and subject to some caveats (see for example Johnson, 1997; Card & Lemieux, 2001; Acemoglu, 2002; Goldin & Katz, 2008; Autor et al., 2008; Carneiro & Lee, 2011). Even though the predictions of the SBTC model relate solely to the skill premium, it is generally perceived to be a useful model in terms of understanding (at least some of) the reasons behind the rise in overall wage inequality. For example, as summarized by Acemoglu (2002), "the

[1] Elbers et al. (2008) also emphasize the fact that group sizes influence the computation of between-group inequality, and focus on how this hampers cross-country comparisons of the importance of between-group differences in accounting for overall inequality.

[2] Between-group inequality also depends on the absolute wage levels of unskilled workers, as we discuss in detail below. This factor, however, is not empirically relevant in explaining the diverging signs of the change in the skill premium and the change in between-group wage inequality observed in the data.

recent consensus is that technical change favors more skilled workers, replaces tasks previously performed by the unskilled, and exacerbates inequality". In a more policy-oriented setting, Mishel et al. (2013) state that "the influential "skill-biased technological change" (SBTC) explanation claims that technology raises demand for educated workers, thus allowing them to command higher wages – which in turn increases wage inequality", while Steelman & Weinberg (2005) argue that "overall, the best explanation for the increase in wage inequality appears to be skill-biased technical change". While many of the papers that study wage inequality provide more comprehensive analyses that go beyond the skill premium, the usefulness of the SBTC model as a framework to understand overall wage inequality relies on the premise that the skill premium is informative about broader inequality patterns.

Our paper makes two key contributions to the literature. The first is to show that, in an environment where both the skill premium and the share of skilled workers are changing, there is no direct empirical link between changes in the skill premium and changes in overall inequality patterns. As we indeed observe for the U.S. and for many local labor markets in the 2010s, between-group wage inequality may decline even as the skill premium is rising. In such cases, between-group differences are not contributing to the rise in overall inequality in spite of the increase in the skill premium. Understanding the evolution of the skill premium is therefore not necessarily in and of itself useful in order to understand why inequality is changing. Researchers and policymakers who are interested in understanding why aggregate measures of inequality are rising should be cautious about drawing inferences based on the rise in the skill premium alone.

The second key contribution of the paper is to show that educational expansions will lead to changes in wage inequality that go beyond the channel that has been highlighted in SBTC models, i.e. the idea that an increase in the supply of skills decreases inequality because it puts downward pressure on the skill premium. As we show, in addition to this indirect effect through the skill premium, a change in the supply of skills has a direct effect on inequality due to the statistical properties of standard inequality measures. This direct effect may amplify or offset the indirect effect operating through the change in the skill premium. It is important for researchers and policymakers to take this additional channel into account when considering the inequality-related implications of an educational expansion policy.

The key messages of this paper also apply to the more recent literature which presents a more nuanced view of the impacts of technological change by distinguishing between the skills that workers possess and the tasks that they perform (Autor et al., 2003; Goos

& Manning, 2007; Acemoglu & Autor, 2011). Recent work in this area, such as Acemoglu & Restrepo (2021), focuses on how the automation of certain tasks has led to changes in the wage structure across different demographic groups. Similarly to the SBTC literature, their framework assumes that the size of each demographic group is fixed. Hence, while the results in Acemoglu & Restrepo (2021) provide key new insights about how the wages of different groups have evolved, if we want to understand aggregate changes in between-group wage inequality (and hence the contribution of between-group differences to overall inequality), it is crucial to also take into account how the relative size of the different groups has changed. Our results show that overall between-group inequality may fall even if the relative wage gaps across demographic groups are rising. Moreover, educational expansions and other changes in the demographic composition of the economy may have first-order effects on aggregate measures of inequality that go beyond what is captured by the relative wage structure.

2 Data

We use data from the decennial Census and the American Community Survey (ACS), obtained through IPUMS (Ruggles et al., 2020). Specifically, we use Census data for 1980, 1990, and 2000, and ACS data for 2008-2010 (which we pool and refer to as 2010), and 2017-2019 (which we pool and refer to as 2019). We follow Acemoglu & Autor (2011) in imposing some sample restrictions. Specifically, we restrict our analysis to non-institutionalized individuals aged 16-64 who are not in the military and are not self-employed. We focus on log weekly earnings for full-time full-year workers, defined as those who worked at least 40 weeks in the previous year and usually work at least 35 hours a week. Nominal earnings are converted to real 2009 dollars using the Consumer Price Index from the Bureau of Labor Statistics. Also following Acemoglu & Autor (2011), top-coded earnings in 1980 are multiplied by 1.5,[3] and individuals earning less than $150 2009 dollars are dropped from the sample. All of the analysis uses the person weights provided in the data.

[3]See `https://usa.ipums.org/usa-action/variables/INCWAGE#codes_section` for details on how values exceeding the top code are already handled in the IPUMS data from 1990 onwards.

3 Skill Premium and Inequality

3.1 Aggregate Patterns

Panel A of Figure 1 plots the evolution of the skill premium over time. The solid blue line shows the ratio of average log real weekly wages for workers with at least some college education, relative to workers with no college experience in our sample. Meanwhile, the dashed red line displays the difference (rather than the ratio) between the average log wages of the two groups, and the dashed green line displays an analogous estimate for this differential that controls for other observable characteristics by estimating the following regression:

$$w_{i,t} = \alpha_t + \beta_t S_{i,t} + \theta_t X_{i,t} + \varepsilon_{i,t} \tag{1}$$

where $w_{i,t}$ represents the log real weekly wages of individual i in period t; $S_{i,t}$ is a dummy variable which is equal to one for individuals with at least some college education; and $X_{i,t}$ is a vector of controls which includes gender, nativity (a dummy for individuals who are foreign-born), two race/ethnicity categories (dummies for non-white and Hispanic individuals, respectively), and a quartic in age. Note that all of the coefficients in Equation (1) are allowed to vary across decades. The dashed green line in Panel A of Figure 1 presents estimates for the skill premium in each decade, $\widehat{\beta}_t$.

Regardless of whether one measures the skill premium based on the raw ratio, the raw differential or the regression-adjusted differential, the figure shows a sharp rise in the earnings gap between more and less educated workers during the 1980s (a rise of around 40% if we focus on the regression-adjusted differential). This gap continues to rise, albeit at a slower rate, between 1990 and 2010 (with a cumulative 26% increase over these two decades). Over the most recent decade, the gap continues to expand, although only slightly (around 3%).

As mentioned, models of Skill-Biased Technical Change (SBTC) focus on the skill premium as the key object of interest and often draw inferences about the evolution of inequality based on the evolution of the skill premium. Panel B of Figure 1 provides a direct measure of the evolution of wage inequality by computing the variance of log real weekly earnings. The solid blue line uses log individual real weekly wages directly, while the dashed red line plots the variance based on residual wages, obtained from a year-specific regression of log individual real weekly wages on the vector $X_{i,t}$ from Equation (1).

The wage variance series shown in Panel B of Figure 1 show a positive trend over time. This coincides with the positive time trend for the skill premium in Panel A, although the growth rates clearly differ across the two panels.

3.2 Within and Between Group Inequality

The fact that the skill premium and the total wage variance do not move in parallel is not surprising, given that the skill premium only captures differences between education groups, while the overall wage variance results from a combination of between-education and within-education group effects. Specifically, the variance of log wages at a given point in time can be decomposed as:

$$
\begin{aligned}
V &\equiv \frac{1}{N}\sum_i (w_i - \overline{w})^2 \\
&= \underbrace{\frac{1}{N}\sum_i \left(w_i - \overline{w}_{g(i)}\right)^2}_{\textit{within group}} + \underbrace{\frac{1}{N}\sum_i \left(\overline{w}_{g(i)} - \overline{w}\right)^2}_{\textit{between group}} \qquad (2)
\end{aligned}
$$

where N is the total sample size, $\overline{w}_{g(i)}$ is the average log wage for the group g (high or low skilled) that individual i belongs to, and $\overline{w}$ is the economy-wide average.[4] The first component captures within-group inequality (individual-level deviations from their group average), while the second component captures between-group inequality (group-level deviations from the economy-wide average).

One might expect that the skill premium would be closely linked to the evolution of between-group inequality. Panel C of Figure 1, however, shows that this is not the case. The figure shows the results of the decomposition from Equation (2), based on the two education groups used for the computation of the skill premium. At any given point in time, the vast majority of inequality is observed within education groups, a fact that is not surprising given the coarse grouping used (only two broad education categories). More surprisingly, the figure shows that the rise in inequality over the most recent decade is entirely driven by the rise in within-group inequality. In fact, even though the skill premium increased slightly during the 2010s, between-group inequality actually *decreased*.

From these figures we can draw an important conclusion: Despite its wide application as a tool to assess inequality, changes in the skill premium do not map directly to changes in between-group inequality. As we see for the most recent decade, these two measures

[4] Time subscripts have been omitted for notational simplicity.

can even move in opposite directions.

3.3 Statistical Link Between the Skill Premium and Between-Group Inequality

Why can the skill premium and the between-group component of the variance of log wages move in opposite directions? From Equation (2), the between-group component of the wage variance is given by:

$$\begin{aligned} V^B &= \frac{1}{N}\sum_i \left(\overline{w}_{g(i)} - \overline{w}\right)^2 \\ &= \frac{1}{N}\left[N_H\left(\overline{w}_H - \overline{w}\right)^2 + N_L\left(\overline{w}_L - \overline{w}\right)^2\right] \end{aligned} \tag{3}$$

where the subscripts H and L denote the high and low-skilled groups, respectively.

We can re-write this expression as:

$$\begin{aligned} V^B &= \overline{w}_L^2\left[\frac{N_H}{N}\left(\frac{\overline{w}_H - \overline{w}}{\overline{w}_L}\right)^2 + \frac{N_L}{N}\left(1 - \frac{\overline{w}}{\overline{w}_L}\right)^2\right] \\ &= \overline{w}_L^2\left[h\left(\omega - \frac{\overline{w}}{\overline{w}_L}\right)^2 + (1-h)\left(1 - \frac{\overline{w}}{\overline{w}_L}\right)^2\right] \end{aligned} \tag{4}$$

where h represents the share of skilled workers in the economy ($0 < h < 1$) and ω represents the skill premium ($\omega \equiv \overline{w}_H/\overline{w}_L > 1$). It can be shown that:

$$\frac{\overline{w}}{\overline{w}_L} = \omega h + (1-h)$$

Equation (4) therefore simplifies to:

$$V^B = \overline{w}_L^2\ h(1-h)(\omega-1)^2 \tag{5}$$

Equation (5) makes clear that the between-group variance of log wages depends not only on the skill premium, ω, but also on two additional factors: the average log wage level among unskilled workers, $\overline{w}_L$, and the fraction of skilled workers in the economy, h.

It is straightforward to observe that an increase in the skill premium, everything else equal, will lead to an increase in the between-group wage variance:

$$\frac{\partial V^B}{\partial \omega} = 2\,\overline{w}_L^2\; h(1-h)(\omega-1) > 0 \tag{6}$$

An increase in the unskilled wage level, everything else equal, also increases the wage variance:

$$\frac{\partial V^B}{\partial \overline{w}_L} = 2\,\overline{w}_L\; h(1-h)(\omega-1)^2 > 0 \tag{7}$$

This is essentially a scale effect: increasing the wage of unskilled workers holding the skill premium constant implies that the skilled wage is also increasing proportionately. In such a case, the whole distribution shifts up and, due to this higher scale, the measured variance increases. Panel A of Figure 2 shows that the mean unskilled log wage has actually shown a decreasing trend between 1980 and 2010, and then shows a slight uptick between 2010 and 2019 (see also Acemoglu & Autor, 2011; Acemoglu & Restrepo, 2021). Changes in the unskilled (log) wage level have therefore put downward pressure on between-group wage inequality between 1980 and 2010. Importantly, however, the (slight) increase in the unskilled wage level between 2010 and 2019 has put *upward* pressure on between-group wage inequality, thus compounding the effect of the increasing skill premium from Panel A of Figure 1. This implies that changes in unskilled wage levels cannot explain why between-group inequality is declining in the 2010s.

On the other hand, a change in the share of skilled workers, everything else equal, has an ambiguous effect on the wage variance:

$$\frac{\partial V^B}{\partial h} = \overline{w}_L^2\; (\omega-1)^2(1-2h) \tag{8}$$

The sign of the expression in Equation (8) depends on the value of h: it will be positive if h is below 0.5 and negative if h is above this threshold. Changes in the share of skilled workers (holding the skill premium and the wage level of unskilled workers constant) may therefore increase or decrease inequality, depending on whether the share of skilled workers is below or above 50%. This threshold is empirically relevant, as it was crossed in the U.S. in the 1980s, as shown in Panel B of Figure 2.

Intuitively, between-group inequality is a weighted average of the (squared) deviation of each group's wage from the aggregate mean wage, with the weight of each group being equal to its relative size, as shown in Equation (3). The mean wage is itself a weighted average of the skilled and unskilled wages. When the share of skilled workers is below 50%, the aggregate mean wage will be closer to the unskilled wage than to the skilled wage, and hence the group that has the smaller deviation from the aggregate mean wage (the unskilled group) has a larger weight. Everything else equal, if the share of skilled workers increases, this increases inequality for two reasons: First, the weight on the group that has a larger deviation from the aggregate mean wage (the skilled group) increases. Second, the increase in the share of skilled workers increases the aggregate mean wage, thus increasing the deviation between the unskilled wage and the aggregate mean wage. On the other hand, if the share of skilled workers is above 50%, the aggregate mean wage will be closer to the skilled wage, and further increases in the share of skilled workers will further increase the weight on the group that has the smaller deviation from the aggregate mean wage, while also shrinking the magnitude of that gap. This intuition applies more generally to any measure of wage inequality that is based on group deviations from the aggregate mean wage, as we discuss in further detail below.

The implication of Equations (6) through (8) is that, in an environment where the share of skilled workers, the unskilled wage level, and the skill premium are all simultaneously changing, between-group inequality does not have to change in the same direction as the skill premium. Crucially, when the share of skilled workers is above 50%, increases in the skill premium and increases in the skill share will have opposing effects on between-group inequality, and the net change will depend on which force dominates.[5,6]

To further illustrate this result, Panel C of Figure 2 plots the observed evolution of between-group wage inequality along with three counterfactuals: one in which the skill premium is allowed to evolve over time as observed in the data but the unskilled wage level and the share of skilled workers are kept as in 1980; one where the unskilled wage level is allowed to evolve over time as observed in the data but the skill premium and the share of skilled workers are kept as in 1980; and one where the share of skilled workers is allowed to evolve over time as observed in the data but the unskilled wage level and the skill premium are kept as in 1980.

In the 1980s, given that the skill share was below 50%, the rise in the skill share and the rise in the skill premium both contributed to an increase in between-group inequality.

[5] These opposing forces are distinct from the race between technology and the supply of skilled workers highlighted by the SBTC model, as we discuss below.

[6] The fact that increases in educational attainment can have inequality-increasing effects has been made in a different context by Bourguignon et al. (2005) and is discussed by Firpo & Portella (2019). Their argument, however, is conceptually distinct from ours, as it relates to the convexity of the returns to education. In our simpler setting with two skill levels, this potential (additional) inequality-inducing effect of rising educational attainment is not present.

Since the 1990s, however, the two forces have worked in opposite directions: the increases in the skill premium have pushed between-group inequality up (as shown in the red dashed line), while the increases in the skill share have pushed between-group inequality down (as shown in the orange dashed line). The falling unskilled wage level also put downward pressure on between-group inequality (as shown in the green dashed line) up until 2010. Importantly, between 2010 and 2019, the downward pressure from the increasing skill share dominated the upward pressure from the increasing wage premium and from the increasing unskilled wage, leading to the overall decline in between-group inequality.

The key conclusion from this analysis is that, in an environment where the skill premium, the unskilled wage level, and the share of skilled workers are all simultaneously changing, there is no direct link between changes in the skill premium and changes in between-group wage inequality.[7] If the share of skilled workers is above 50% and rising, an increase in the skill premium may be observed alongside a *decrease* in between-group inequality, as we in fact observe over the most recent decade in the U.S. in Figure 1. While changes in between-group inequality moved closely in line with changes in the skill premium in the 1980s and 1990s, this has become less so in more recent decades.

3.4 Link to SBTC Model

The SBTC model argues that the skill premium is determined as the outcome of a "race" between technological change – which increases the relative demand for skilled workers – and rising educational attainment – which increases the relative supply of this group. All else equal, technology is argued to push the skill premium (and inequality) up, while increases in education are argued to push the skill premium (and inequality) down.

More formally, if we use z_s to denote a skill-biased technology shock, the model predicts that $\partial\omega/\partial z_s > 0$ and, all else equal, given the assumption that the number of workers of each skill type is exogenously given and that all workers have a perfectly inelastic labor supply, the shock has no effect on the skill share h.[8] In terms of between-group inequality, based on the definition in Equation (5) and the result in Equation (6), this implies:

[7] This result is of course more general in the sense that it would apply in any context that considers wage differentials between different groups of workers (defined by any characteristic; not necessarily skills): a growing (shrinking) wage gap between the groups does not necessarily imply more (less) between-group inequality, if the size of the groups is also changing.

[8] For simplicity, we consider here a version of the SBTC model in which shocks to technology and shocks to the supply of skills only affect relative wages, without affecting unskilled wages in levels, but we note below how our results are strengthened if we also allow unskilled wage levels to be affected by these shocks.

$$\frac{\partial V^B}{\partial z_s} = \frac{\partial V^B}{\partial \omega} \frac{\partial \omega}{\partial z_s} > 0 \tag{9}$$

Hence, a skill-biased technological change shock (which does not impact the skill share of the economy) will unambiguously lead to an increase in the skill premium and an increase in between-group inequality.[9]

On the other hand, an increase in education (holding the skill bias of technology constant), increases the skill share h and endogenously leads to a reduction in the skill premium ω in the model. Denoting the education shock as z_e, this implies that $\partial h/\partial z_e > 0$ and $\partial \omega/\partial z_e < 0$. If we consider the implication for between-group inequality, we have that:

$$\frac{\partial V^B}{\partial z_e} = \frac{\partial V^B}{\partial h} \frac{\partial h}{\partial z_e} + \frac{\partial V^B}{\partial \omega} \frac{\partial \omega}{\partial z_e} \tag{10}$$

Equation (10) shows that the shock has two effects on between-group inequality. The first is the direct effect due to the change in the skill share; the second is the indirect effect due to the change in the skill premium. From Equation (6), we know that $\partial V^B/\partial \omega > 0$, so given that $\partial \omega/\partial z_e < 0$, the second term in Equation (10) is unambiguously negative. This is the channel traditionally highlighted in SBTC models, whereby an increase in the skill supply will put downward pressure on the skill premium and will therefore push inequality downwards. Equation (10), however, shows that there is an additional direct channel (the first term in the equation) through which the change in the supply of skills impacts between-group inequality. We know from Equation (8) that the sign of $\partial V^B/\partial h$ depends on the value of h, so the overall sign of $\partial V^B/\partial z_e$ in Equation (10) depends on the value of h and the magnitude of the different terms. In particular, if $h > 0.5$, then $\partial V^B/\partial h < 0$ and the whole term in Equation (10) is negative. However, if $h < 0.5$, then $\partial V^B/\partial h > 0$ and the sign of $\partial V^B/\partial z_e$ will depend on whether the direct or the indirect effect dominates. In other words, in response to a shock that exogenously increases the supply of skilled workers, between-group inequality may go up or down, even if the skill premium unambiguously falls.[10]

[9] This result is exacerbated if the unskilled wage level is allowed to change. Technological progress is assumed to have a positive impact on the wages of all workers (though disproportionately so for skilled workers). This means that $\partial \overline{w}_L/\partial z_s > 0$, which implies that, as a result of the skill-biased technology shock, between-group inequality would increase not only because of the increase in the skill premium, but also because of the increase in the unskilled wage level.

[10] This ambiguity is compounded further if we allow the unskilled wage level to change. Specifically, the model would predict that $\partial \overline{w}_L/\partial z_e > 0$, due to the reduction in the supply of unskilled workers. This would put positive pressure on $\partial V^B/\partial z_e$, which could further offset the negative pressure on between-

More generally, within the context of the SBTC model, if there are shocks that affect both the skill premium and the skill share of the economy, the sign of the change in the between-group variance is ambiguous, even if the sign of the change in the skill premium can be determined. Such shocks can include not only changes in education, but also shocks such as technology or trade that induce changes in labor supply (e.g. if labor supply is not perfectly inelastic or if technology or trade shocks affect individuals' educational choices). The ambiguity also applies in cases where there are independent but simultaneous shocks that affect the relative demand and the relative supply of skills, such that both the skill share and the skill premium are simultaneously changing.

3.5 Results with Other Inequality Indices and Skill Groupings

The results that we have discussed are not limited to our choice of the variance of log wages as a measure of inequality. Similar trends are observed when we compute the between-group component of two other standard measures of inequality: the Gini and the Theil indices.

The between-group component of the Gini coefficient can be obtained by estimating the area under the Lorenz curve through a linear interpolation technique, using the average log wage and the population share of each of the two skill groups. It can be written as:

$$Gini_{bet} = 1 - \left[h^2\left(\frac{\overline{w}_H}{\overline{w}}\right) + (1-h)^2\left(\frac{\overline{w}_L}{\overline{w}}\right) + 2h(1-h)\left(\frac{\overline{w}_L}{\overline{w}}\right)\right] \tag{11}$$

The between-group component of the Theil index is given by:

$$Theil_{bet} = h\left(\frac{\overline{w}_H}{\overline{w}}\right)\ln\left[\frac{\overline{w}_H}{\overline{w}}\right] + (1-h)\left(\frac{\overline{w}_L}{\overline{w}}\right)\ln\left[\frac{\overline{w}_L}{\overline{w}}\right] \tag{12}$$

The evolution of these two indices is shown in Figure 3. Consistent with what we observe for the between-group variance of log wages, these inequality indices grow from 1980 to 2010, but fall thereafter, in spite of the continued increase of the skill premium in the most recent decade. This demonstrates that the lack of a direct empirical link between the skill premium and between-group inequality (in an environment where group sizes are changing) is a more general result, and not specific to the wage variance.

It is also interesting to consider the evolution of the skill premium and between-group inequality separately for men and women. This is shown in the top panel of Figure 4. While the estimated skill premium is higher for women than for men (see the top left

group inequality arising from the fall in the skill premium.

panel), overall between-group inequality is higher for men than for women (as shown in the top right panel). For men, we observe the same pattern as in the aggregate in terms of the opposing signs for the change in the skill premium and the change in the between-group variance of log wages between 2010 and 2019. For women, both measures continue to increase in the 2010s, albeit at a smaller pace.

The discrepancy between the empirical evolution of the skill premium and overall between-group inequality can arise when the majority of the workforce is classified as skilled, and when this group continues to grow. While classifying workers as skilled or unskilled based on whether they have at least some college education has been the norm in the literature, it may be the case that it has now become more relevant to consider other (stricter) definitions of what constitutes a skilled worker. Panel B of Figure 4 shows results based on two alternative definitions of skills: one where only those with at least four years of college education are considered skilled, and one where workers are considered skilled if they are employed in a high-skill high-wage non-routine cognitive occupation.[11]

As the bottom left panel of Figure 4 shows, the skill premium based on these alternative measures of skills has been steadily rising since the 1980s. As the bottom right panel shows, between-group inequality based on these categorizations has also steadily risen. This is because, although the share of skilled workers according to these definitions has been rising over time (i.e. an increasing share of the workforce has at least four years of college education, and an increasing share of the workforce is employed in non-routine cognitive occupations), this share remains below 50% in both cases. Hence, changes in group sizes have compounded the changes in the skill premium in driving the rise in between-group inequality.

Overall, these results suggest that analyses based on these alternative definitions of skills may provide better guidance about the evolution of overall inequality in recent decades. It is worth noting, however, that the share of skilled workers based on these alternative definitions is steadily approaching 50%, so the empirical disconnection between the evolution of the skill premium and the evolution of between-group inequality that we observe based on the traditional definition of skills may also arise in future years even when using these stricter measures.

[11]This classification is based on the polarization literature, which groups occupations based on the tasks that workers perform; see Autor et al. (2003); Acemoglu & Autor (2011). We use the mapping of occupation codes from Cortes et al. (2020).

3.6 Patterns at the Local Labor Market Level

As a final illustration of the potential empirical divergence between the skill premium and the between-group component of inequality, we use data at the Commuting Zone (CZ) level. To do so, we make use of the local labor market crosswalk files from Autor & Dorn (2013) and Autor et al. (2019), which provide a probabilistic matching of the smallest Census geographic units to the CZ level. CZs are defined as clusters of U.S. counties that are characterized by strong within-cluster and weak between-cluster commuting ties, with the total number of these locations being 741.

Figure 5 plots the cross-sectional relationship between the skill premium and the between-group variance of log wages at the CZ level for both 1980 and 2019. Each circle corresponds to a particular CZ, with the size of the circle being proportional to the CZ's total employment in that particular year. Although the figure shows that there is indeed a strong positive correlation between the skill premium and between-group inequality at the CZ level, one can identify many location pairs where one location has a higher skill premium, but a lower level of between-group inequality. For example, in 2019 the skill premium in New York was 1.099, while in Washington DC it was 1.106. In spite of having a higher skill premium, Washington DC's between-group wage variance was lower than New York's (0.080 vs 0.083).

Figure 6 plots the relationship between within-CZ changes in the skill premium and within-CZ changes in between-group log wage variances in each of the four decades in our data.[12] The top right and bottom left quadrants in each of the four panels contain the locations where the skill premium and between-group inequality moved in the same direction (i.e. they either both increased or both decreased) during the particular decade. Meanwhile, locations in the top left and in the bottom right quadrants experienced oppositely-signed changes in the two measures. The graph indicates the fraction of CZs in each quadrant, as well as the fraction of total employment comprised by these locations (in brackets).

As the figure shows, in the earlier decades, almost all CZs experienced changes in the skill premium that were of the same sign as the change in their between-group inequality. The link, however, became weaker in the most recent decade, with the proportion of CZs with oppositely signed changes increasing from 2% in the 1980s to 8% in the 2010s. Notably, many of these locations with oppositely signed changes are large in terms of their total employment. In fact, around one-fifth of the total workforce in 2019 is found

[12]The size of the circle for each CZ is proportional to its total employment at the end of the corresponding decade.

in locations where the skill premium went up, but between-group inequality fell between 2010 and 2019.

The weakening over time in the relationship between changes in the skill premium and changes in between-group inequality is driven by the growth in the share of locations where skilled workers represent more than half of the local workforce, along with the fact that the skill premium tended to rise less during the most recent decade (making it more likely that the upward pressure on between-group inequality coming from the rise in the skill premium could be offset by the downward pressure coming from the rise in the share of skilled workers). Overall, these results further confirm that inferring inequality implications from changes in the skill premium has become particularly problematic over recent years.

4 Conclusions

A vast literature aimed at understanding the nature and causes of wage inequality has focused on the skill premium as a key object of interest. While we do not dispute that the skill premium may be of interest *per se*, we show in this paper that, when the share of skilled workers exceeds 50% and continues to grow (as is currently the case nationally in the U.S. as well as in many local labor markets), there is no unambiguous empirical link between changes in the skill premium and changes in the between-group component of standard wage inequality measures (such as the wage variance, the Gini index or the Theil index). Increases in the skill premium may in fact occur alongside *decreases* in between-group wage inequality – as we indeed observe in the data. Hence, in an environment in which the share of skilled workers is changing, understanding the evolution of the skill premium is not sufficient in order to understand broader inequality trends in the economy, even if one is solely interested in understanding the between-group component of inequality. This is especially true when changes in the skill premium are relatively small. In such cases, changes in the share of skilled workers may be the dominant factor driving changes in between-group inequality. Researchers and policymakers should therefore be weary about focusing (solely) on the skill premium when trying to understand why wage inequality in a particular economy is rising.

Our results also show that policies to expand access to higher education will have important consequences for inequality which go beyond the channel that the skill-biased technological change (SBTC) literature has focused on, i.e. the idea that an increase in the supply of skills decreases inequality because it puts downward pressure on the skill premium. In particular, in societies where the share of skilled workers is below 50%, increases in educational attainment have the direct effect of increasing measured between-group inequality. Although expanding education is clearly desirable for many reasons beyond the impact on inequality, it is important for policymakers to keep in mind that an educational expansion in an economy where the share of skilled workers is below 50% may not translate into a reduction in aggregate measures of between-group inequality: If the direct group size effect dominates the indirect effect via the skill premium, between-group inequality will increase. On the other hand, in societies where the share of skilled workers is above 50%, an increase in educational attainment reduces between-group inequality not only by putting downward pressure on the skill premium, but also through its direct effect due to the change in group sizes. Indeed, rising educational attainment has led to a reduction in (between-group) inequality in the U.S. in the 2010s, but not because of the reason emphasized by the SBTC literature (given that the skill premium has actually continued to rise during this time period), but rather because of the novel channel that we emphasize in this paper which is related to the group size effect.

Finally, our results suggest that it has become increasingly relevant to focus on more stringent definitions of skill that go beyond the traditional distinction based on whether individuals have at least some college education. Moreover, the rise in inequality *among* individuals with some college education implies that college attendance does not necessarily guarantee a high income stream. While going to college remains an importance insurance mechanism against low incomes, it is less so now than in the past. The rising empirical importance of within-group inequality also highlights the need to move beyond standard models that focus on between-group changes towards models that allow us to better understand why heterogeneity in wages within groups is rising.

References

Acemoglu, D. (2002). Technical change, inequality, and the labor market. *Journal of Economic Literature*, 40(1), 7–72.

Acemoglu, D. & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of Labor Economics*, 4, 1043–1171.

Acemoglu, D. & Restrepo, P. (2021). Tasks, automation, and the rise in us wage inequality. *NBER Working Paper No. 28920*.

Autor, D., Dorn, D., & Hanson, G. (2019). When work disappears: Manufacturing decline and the falling marriage market value of young men. *American Economic Review: Insights*, 1(2), 161–78.

Autor, D. H. & Dorn, D. (2013). The growth of low skill service jobs and the polarization of the U.S. labor market. *American Economic Review*, 103(5), 1553–1597.

Autor, D. H., Katz, L. F., & Kearney, M. S. (2008). Trends in US wage inequality: Revising the revisionists. *The Review of Economics and Statistics*, 90(2), 300–323.

Autor, D. H., Katz, L. F., & Krueger, A. B. (1998). Computing inequality: Have computers changed the labor market? *The Quarterly Journal of Economics*, 113(4), 1169–1214.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, 118(4), 1279–1333.

Bourguignon, F., Ferreira, F. H., & Lustig, N. (2005). *The microeconomics of income distribution dynamics in East Asia and Latin America*. The World Bank and Oxford University Press.

Card, D. & Lemieux, T. (2001). Can falling supply explain the rising return to college for younger men? A cohort-based analysis. *The Quarterly Journal of Economics*, 116(2), 705–746.

Carneiro, P. & Lee, S. (2011). Trends in quality-adjusted skill premia in the United States, 1960–2000. *The American Economic Review*, 101(6), 2309–2349.

Cortes, G. M., Jaimovich, N., Nekarda, C. J., & Siu, H. E. (2020). The dynamics of disappearing routine jobs: A flows approach. *Labour Economics*, 65, 101823.

Elbers, C., Lanjouw, P., Mistiaen, J. A., & Özler, B. (2008). Reinterpreting between-group inequality. *The Journal of Economic Inequality*, 6(3), 231–245.

Firpo, S. & Portella, A. (2019). Decline in wage inequality in Brazil: A survey. *The World Bank Policy Research Working Paper 9096*.

Goldin, C. & Katz, L. F. (2008). Long-run changes in the wage structure: Narrowing, widening, polarizing. *Brookings Papers on Economic Activity*, 2007(2), 135–165.

Goos, M. & Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in Britain. *The Review of Economics and Statistics*, 89(1), 118–133.

Johnson, G. E. (1997). Changes in earnings inequality: The role of demand shifts. *Journal of Economic Perspectives*, 11(2), 41–54.

Juhn, C., Murphy, K. M., & Pierce, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy*, 101(3), 410–442.

Katz, L. F. & Murphy, K. M. (1992). Changes in relative wages, 1963–1987: Supply and demand factors. *The Quarterly Journal of Economics*, 107(1), 35.

Mishel, L., Shierholz, H., & Schmitt, J. (2013). Don't blame the robots: Assessing the job polarization explanation of growing wage inequality. *EPI-CEPR Working Paper*.

Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J., & Sobek, M. (2020). Integrated Public Use Microdata Series (IPUMS) USA: Version 10.0 [dataset].

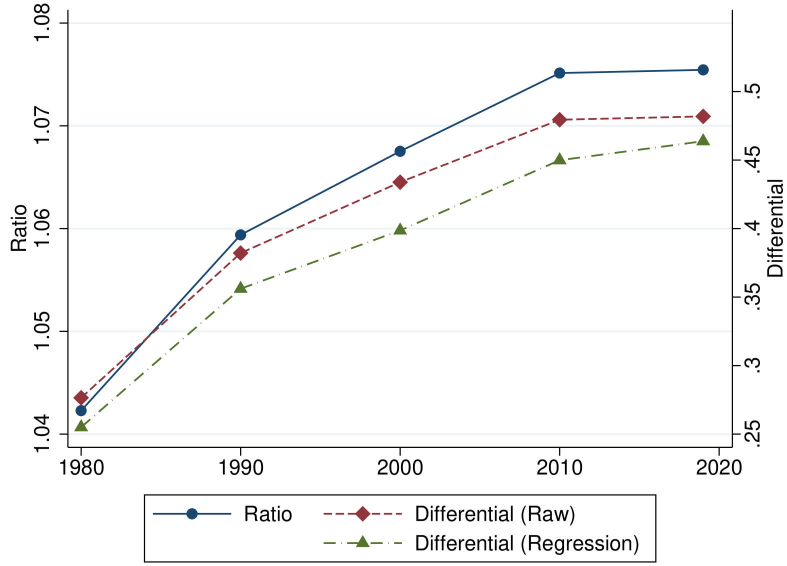
Steelman, A. & Weinberg, J. A. (2005). What's driving wage inequality? *FRB Richmond Economic Quarterly*, 91(3), 1–17.

Tinbergen, J. (1974). Substitution of graduate by other labor. *Kyklos*, 27(2), 217–226.

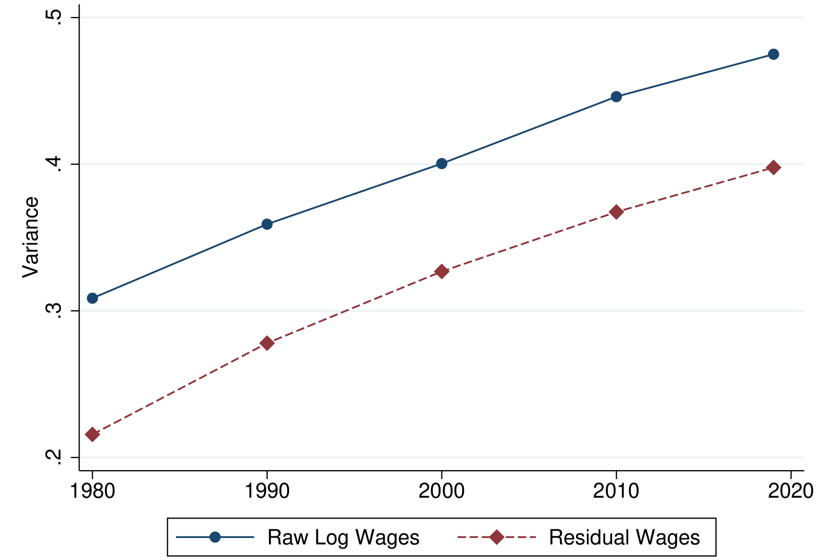
Tinbergen, J. (1975). *Income Differences: Recent Research*. North-Holland Publishing Company.

Figure 1: Evolution of the Skill Premium and the Variance of Log Wages

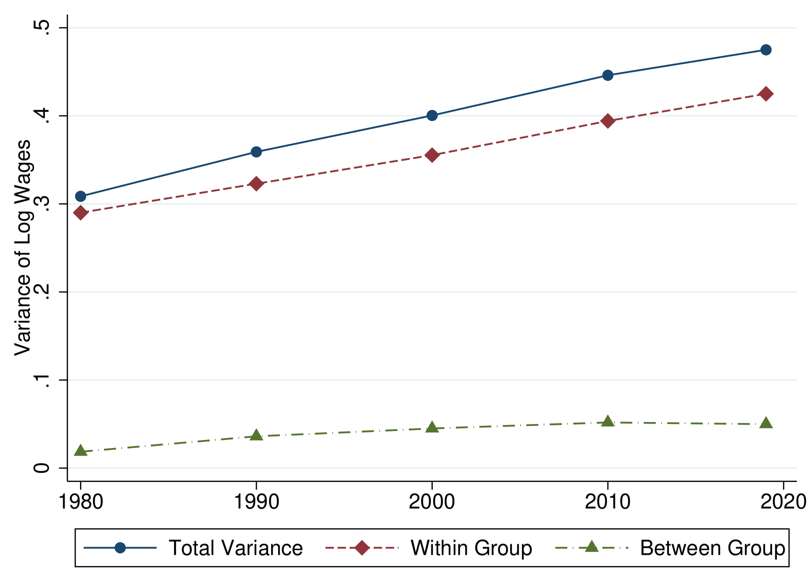
Panel A: Skill Premium



Panel B: Wage Variance

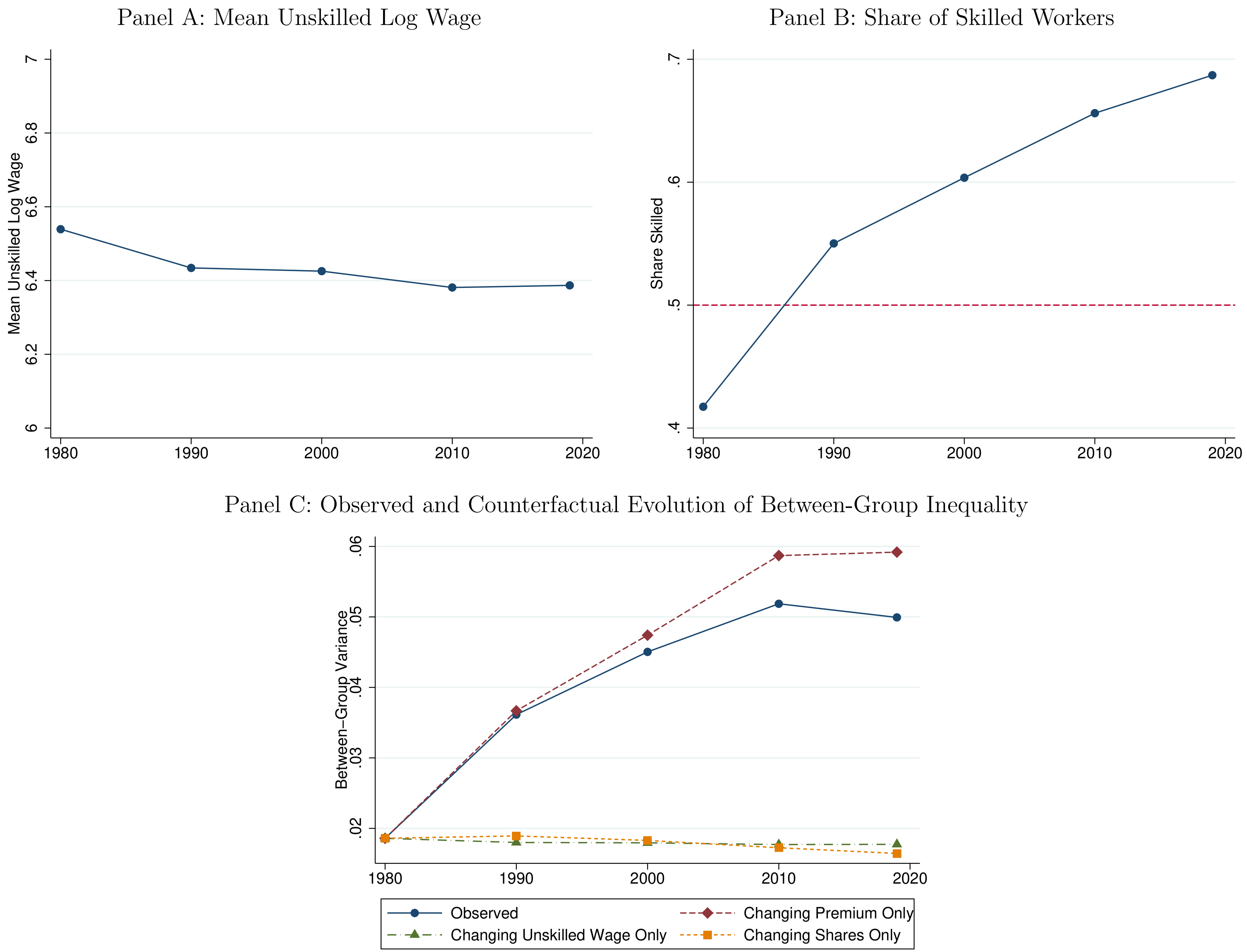


Panel C: Decomposition of Wage Variance



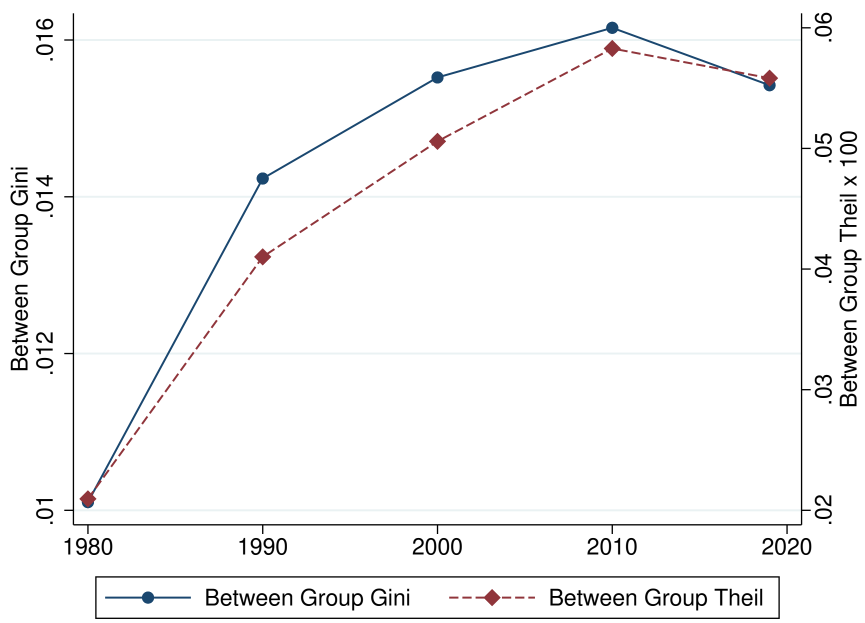
Note: The figures are based on Census and American Community Survey data for full-time full-year workers aged 16-64. Skilled workers are those with at least some college education. Panel A plots the ratio and raw differential between average log real weekly wages of skilled and unskilled workers, and a regression-based estimate that controls for observable characteristics. Panel B plots the variance of log wages and of residual wages (controlling for observable characteristics). Panel C decomposes the variance into between- and within-group components, based on the two skill groups. See text for details.

Figure 2: Mean Unskilled Log Wage, Skilled Worker Share, and Counterfactual Between-Group Variances



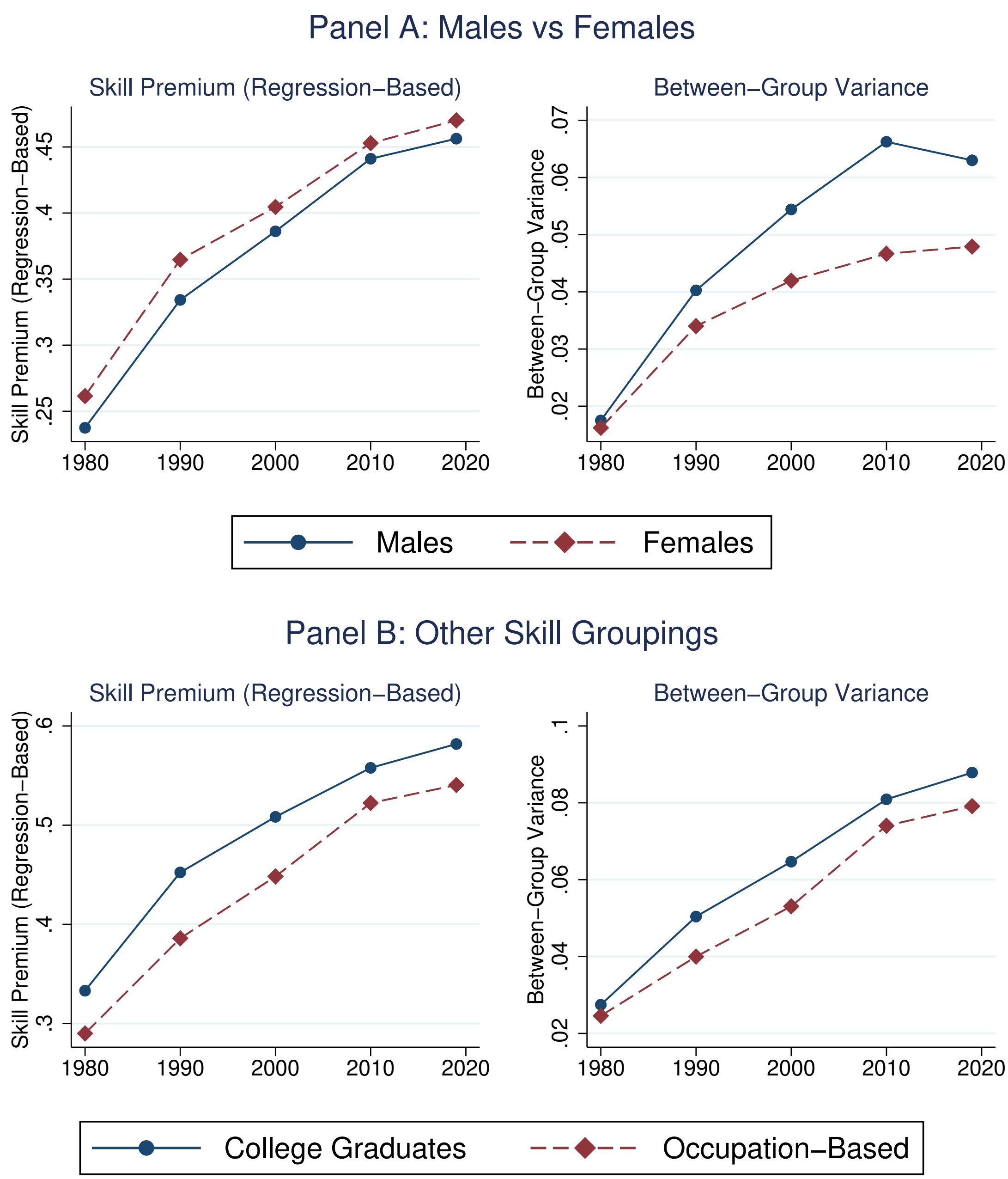
Note: The figure is based on Census data for 1980, 1990 and 2000, and American Community Survey data for 2008-2010 and 2017-2019 for full-time full-year workers aged 16-64. Skilled workers are those with at least some college education. Panel A shows the mean log wage of unskilled workers. Panel B shows the share of skilled workers. The solid line in Panel C shows the observed between-group variance of log wages. The dashed lines display counterfactual between-group variances where only one of the three factors (either the skill premium, the unskilled log wage level, or the share of skilled workers) is allowed to change over time as observed in the data, while the other two factors are held constant at their 1980 levels.

Figure 3: Evolution of the Between-Group Components of the Gini and the Theil Indices



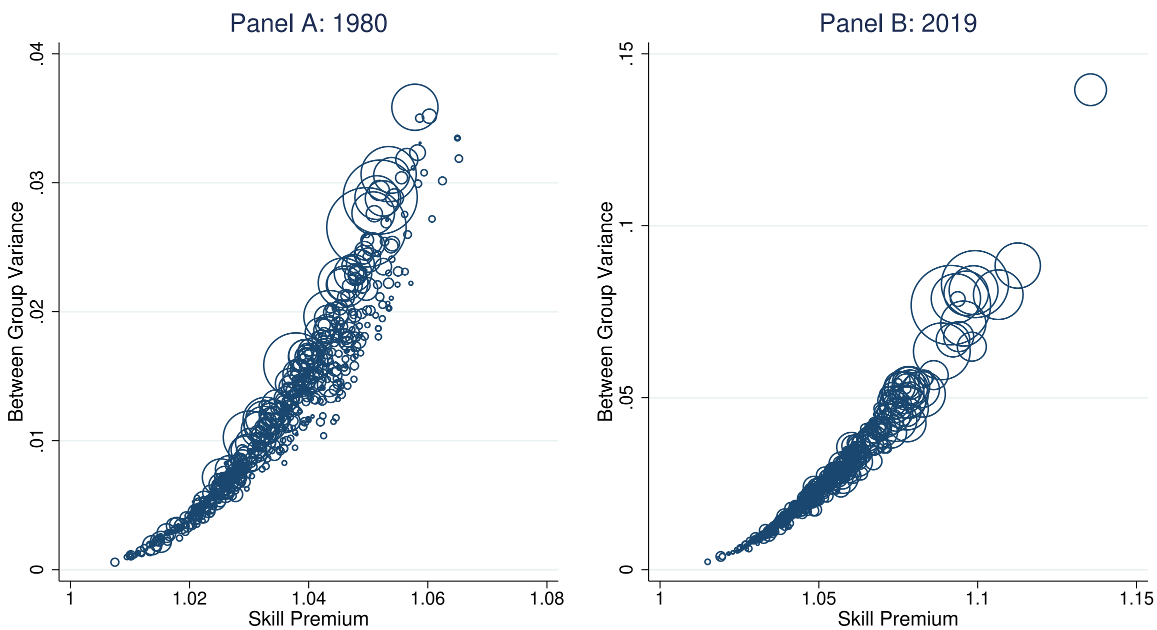
Note: The figure is based on Census data for 1980, 1990 and 2000, and American Community Survey data for 2008-2010 and 2017-2019 for full-time full-year workers aged 16-64. The figure displays the between-group component of the Gini and Theil indices for the distribution of log wages; see Equations (11) and (12). Groups are based on skill levels, with skilled workers being those with at least some college education.

Figure 4: Skill Premium and Between-Group Inequality with Other Worker Groupings



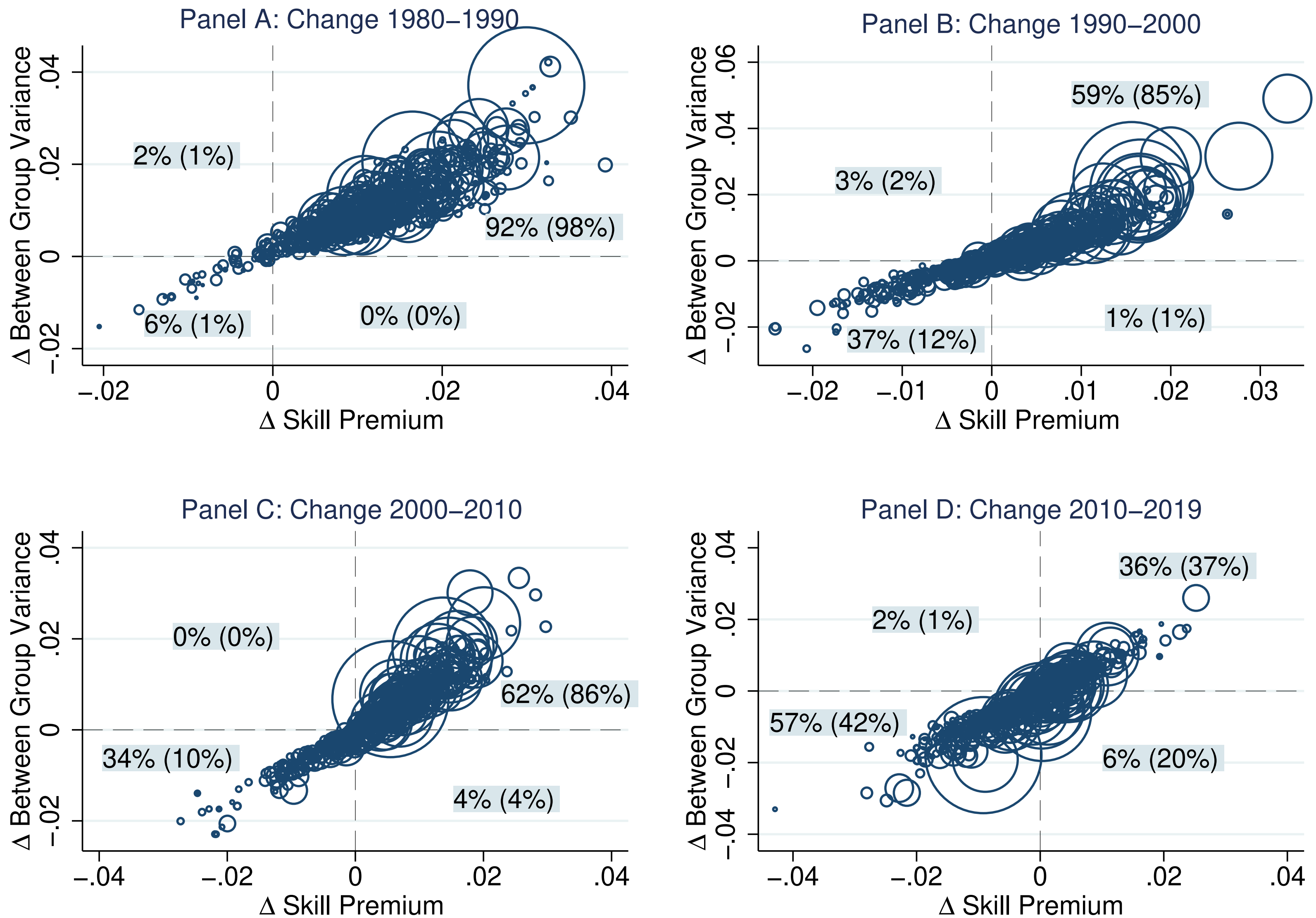
Note: The figure is based on Census data for 1980, 1990 and 2000, and American Community Survey data for 2008-2010 and 2017-2019 for full-time full-year workers aged 16-64. In Panel A, skilled workers are those with at least some college education. Panel B considers two alternative definitions of skilled workers. For the patterns in the solid blue line, skilled workers are those with at least four years of college education. For the patterns in the dashed red line, skilled workers are those employed in high-skill non-routine cognitive occupations, following the classification of Cortes et al. (2020).

Figure 5: Cross-Sectional Relationship between the Skill Premium and Between-Group Log Wage Variance at the Commuting Zone Level



Note: The figure is based on Census data for 1980, 1990 and 2000, and American Community Survey data for 2008-2010 and 2017-2019 for full-time full-year workers aged 16-64. Each circle represents a commuting zone, with the size of the circle being proportional to the commuting zone's total employment in the corresponding year. The skill premium is computed as the ratio of the average log wage of skilled workers relative to the average log wage of unskilled workers. Skilled workers are those with at least some college education.

Figure 6: Relationship between Changes in the Skill Premium and Changes in Between-Group Log Wage Variance at the Commuting Zone Level



Note: The figure is based on Census data for 1980, 1990 and 2000, and American Community Survey data for 2008-2010 and 2017-2019 for full-time full-year workers aged 16-64. Each circle represents a commuting zone, with the size of the circle being proportional to the commuting zone's total employment at the end of the corresponding decade. The skill premium is computed as the ratio of the average log wage of skilled workers relative to the average log wage of unskilled workers. Skilled workers are those with at least some college education. The figure indicates the share of commuting zones in each quadrant, as well as the share of total employment comprised by those commuting zones (in brackets).