Revisiting U.S. Wage Inequality at the Bottom 50%

Oren Danieli*
Tel Aviv University

July 19, 2022

Abstract

I propose a model of a skill-replacing routine-biased-technological-change (SR-RBTC). In this model, technology substitutes the usage of skill in routine tasks, in contrast to standard RBTC models which assume technology replaces the workers themselves. The SR-RBTC model explains three key trends that are inconsistent with standard RBTC models: why specifically middle-wages declined even though routine workers are dispersed across the entire bottom half of the wage distribution, why middle-wages stopped declining while the technological change continued, and why there is no substantial decline in the average wage of routine workers. I derive two new testable predictions from the model: a decrease in return to skill, and a decrease in skill level in routine occupations. I use an interactive-fixed-effect model to confirm both predictions. Since SR-RBTC violates the ignorability assumption required by standard decomposition methods, I introduce "skewness decomposition" to show that SR-RBTC is the main driver of bottom-half inequality trends.

*Email: orendanieli@tauex.tau.ac.il. I gratefully acknowledge financial support from the W.E. Upjohn Institute's Early Career Research Award program. I thank Lawrence Katz, David Autor, Roland Fryer, Thomas Lemieux, Ro'ee Levy, Jonathan Roth, Itay Saporta-Eksten, and Eran Yashiv for very helpful comments and suggestions. Seminar participants at Harvard, Tel-Aviv University, Hebrew University, Princeton University, the department of environmental economics at HUJI and Haifa University provided useful feedback. Elad Guttman provided outstanding research assistance.

1 Introduction

In recent decades, inequality at the bottom half of the U.S. wage distribution has been fluctuating. Figure 1 plots the evolution of the 90/50 and 50/10 wage ratios over time. At the upper half of the distribution, inequality has been steadily rising for more than three decades. At the bottom half, inequality was equally rising in the early 1980s due to skill-biased technological change (Katz and Murphy, 1992). Yet in the late 1980s, trends within the upper and lower halves of the distribution started to diverge, as bottom-half inequality started to decline. This generated a relative decrease in middle-wages that is often referred to as "wage polarization." Inequality at the bottom half then stopped declining and even started to moderately increase around the year 2000. The leading explanation for these trends is routine-biased-technological-change (RBTC): a decline in demand for routine tasks (Autor et al., 2006).

But while there is strong evidence that supports the existence of RBTC, this theory cannot explain some of the key stylized facts. Prominent models for RBTC (e.g., Acemoglu and Autor, 2011) argue that routine tasks, which can be more easily automated, require middle-skilled workers. As a result, middle-skilled workers should see the largest decline in their wages when automation expands. Yet, these models cannot explain why the decline in middle wages stopped around the year 2000, while evidence suggests that RBTC continued long afterward (Autor, 2014). It is also hard to support the assumption that RBTC reduces demand mostly for middle-skilled workers when routine workers are actually dispersed almost equally across the entire bottom half of the wage distribution (Autor and Dorn, 2013). Moreover, it is not clear why average wage declines in routine occupations are relatively modest, compared to the substantial decline in employment. Different decomposition methods found that the decrease in wages in routine occupations is too small to account for the aforementioned wage trends (Autor et al., 2005; Firpo et al., 2013). These unresolved puzzles leave room for consideration of other explanations for the fluctuations in bottom-half inequality (Hunt and Nunn, 2019), such as fluctuations in the real minimum wage (Piketty, 2014), decline in unions (Firpo et al., 2013), business cycles (Foote and Ryan, 2015), demand growth in the service sector (Autor and Dorn, 2013), or the low unemployment rate during the 1990s.

In this paper, I show theoretically and empirically that a small, yet important, modification to the RBTC model can solve all these puzzles. Instead of assuming that new technology is replacing workers, I assume that it replaces their skill. For instance, calculators replace the need for arithmetic skills and allow workers with low arithmetic abilities to perform calculation tasks equally well. As a result, the return to skill drops in routine occupations. This reduces wages for high-skill routine workers, but could actually increase wages for low-skill routine workers. Overall, the average decrease in wages for routine workers would be modest, but wage gaps within routine workers would drop considerably. At first, wage gaps at the bottom half of the distribution, where most routine workers are concentrated, decline. But gradually, the relative decline in their wages incentivizes middle-skilled workers to leave (or never join) routine occupations. At that point, while SR-RBTC continues, it does not affect middle-skilled workers since they do not perform routine tasks anymore. Instead, it mostly reduces wages of low-skilled workers, which makes inequality at the bottom half rise again.

I test this model in the data in two distinct ways. First, I confirm the above model predictions using an interactive fixed-effect model (IFEM). Using this model I find that the return to skill has sharply declined in routine occupations, making them the occupations with the lowest return to skill. I also show that with time, the composition of workers in routine occupations becomes less skilled until eventually, they employ the lowest-skilled workers on average. Second, I introduce a "skewness decomposition", which can quantify the wage impact of SR-RBTC. Using this decomposition, I find that almost the entire trend of wage polarization (93%) is attributed to occupations, and is mostly driven by the decrease in inequality in low-paying routine occupations, consistent with the prediction of the SR-RBTC model. Taken together, my results suggest that SR-RBTC is not only consistent with the data, but is also the main driver of bottom-half inequality trends in recent decades.

I start by outlining the theoretical framework of the paper. I construct a model where workers can be characterized by a one-dimensional continuous skill. Workers are employed in one of three occupations that vary in their return to skill (similar to Jung and Mercenier, 2014; Cortes, 2016). In equilibrium, workers are allocated to occupations based on comparative advantage. The lowest-skilled workers sort into the manual occupation, middle-skilled to the routine occupation, and the highest-skilled workers into the abstract occupation.

The model deviates from the majority of the previous literature by assuming that the new technology in the routine occupation substitutes the skill of workers in this occupation. This skill-replacing technology reduces the return to skill in the routine occupation, which generates larger wage decreases for higher-skilled routine workers. This differs from previous models (i.e., Acemoglu and Autor, 2011; Cortes, 2016) which assume a skill-neutral technological change, where the wage effects are identical to all routine workers, regardless of their skill levels. It also differs from skill-enhancing models (Jung and Mercenier, 2014) that make the opposite assumption that technology increases the return to skill for routine workers. It is conceptually similar to Downey (2021) who argues new technology benefits low-skill workers.

There are many examples of skill-replacing technology in routine occupations. Cashiers today do not need any arithmetic skills, as all calculations are automated; administrators typically do not need to memorize any procedures or customer details, as most of them are computerized; production workers rarely use their physical strength anymore as machines can perform many physical tasks. Yet, counterexamples exist as well. Therefore, the empirical part of this paper tests how well the SR-RBTC model fits the data.

SR-RBTC predicts a non-monotonic relationship between technological advancement and inequality at the bottom half of the distribution, consistent with the trends in recent decades. A decrease in return to skill in the routine occupation would generate the largest decrease in wages for the highest earning routine workers. Empirically, the highest earning routine workers used to be concentrated in the middle of the wage distribution.[1] A decrease in their wages would generate wage polarization. This creates an incentive for middle-wage workers to switch occupations, which generates a large decrease in employment in the routine occupation. The average wage in the routine occupation does not necessarily decrease as low-skill routine workers may benefit from the change.

At some point, the return to skill in the routine occupation would fall below its level in the manual occupation. When such a reversal of comparative advantage occurs, workers would completely reallocate and only the lowest-skilled workers would choose to work in the routine occupation. After this reversal, SR-RBTC would reduce wages for the remaining routine workers who would be at the bottom of the wage distribution. SR-RBTC would not affect middle-wage workers directly as they would no longer work in the routine occupation. Therefore, inequality at the bottom of the distribution would rise.

To directly test the model, I derive two new predictions that can distinguish a skill-replacing RBTC from a skill-neutral or skill-enhancing RBTC. First, the model predicts a decrease in return to skill in the routine occupation. Second, it predicts a gradual decline in the skill level of workers in the routine occupation. These trends should continue throughout the entire period of RBTC, starting in the late 1980s. At some point, the return to skill in the routine occupation should fall below the return in the manual occupation, leading to a reversal of comparative advantage. Following such reversal, the average skill level should be lower in the routine occupation, compared to the manual occupation.

To test these predictions, I estimate an interactive fixed-effects model (IFEM). IFEM is

[1] This is shown in Appendix Figure A1, which plots the average routine intensity index of workers' occupation by their wage percentile in 1990. Routine workers are concentrated at the bottom three quintiles of the wage distribution, and so the highest earning routine workers are around the median. The routine intensity index is defined in Appendix B. This result is similar to that of Autor and Dorn (2013, Figure 4).

a more general version of the standard fixed-effects model. It regresses log wages on a set of independent variables, including worker fixed effects that capture unobserved skill. The only difference from a standard fixed-effects model is that the worker fixed effects are interacted with the year and occupation. This interaction allows the return to the unobserved skill to vary over time and between occupations, as the SR-RBTC model predicts. Since the unobserved skill is estimated with noise, I instrument for the unobserved skill with years of schooling to prevent an attenuation bias. For this exercise, I use data from the Panel Study of Income Dynamics (PSID) between 1980–2011.

The results from the IFEM generate new empirical facts that are consistent with both model predictions. I find a sharp decrease in the return to skill in routine occupations starting in the late 1980s, exactly when inequality at the bottom half of the distribution started to decline. The return to skill in routine occupations continued to decrease for more than two decades. I also find that the average skill level in routine occupations, as measured using the IFEM, steadily fell during this period. As a result, routine workers became more concentrated at the bottom income quintile, instead of working in middle-wage jobs. Previous work investigating the compositional change in routine occupations has focused mainly on the employment decline in routine occupations (Goos and Manning, 2007; Goos et al., 2009, 2014), and the flow of workers in and out of routine occupations (Cortes, 2016; Cortes et al., 2020). However, there has been little discussion on the impact of these employment trends on the average skill level in routine occupations.

Estimates from the IFEM are consistent with a reversal of comparative advantage at the bottom of the distribution. I find that around 1987, before inequality started to decline at the bottom of the distribution, the return to skill was slightly higher in routine occupations compared to manual occupations. During this time, the average skill level of routine workers was substantially higher than the skill level of manual workers. With time, the return to skill in routine occupations fell far below its value in manual occupations. The average skill level of routine workers also declined and, by 2009, it fell below the average skill level in manual occupations. As a result, routine workers and especially those in administrative or operator occupations have the lowest level of skill across all occupational categories.[2]

In the final part of the paper, I use a decomposition exercise to show that SR-RBTC is substantial enough to account for the overall wage trend. While the IFEM results support

[2]By 2011 manual workers still earned less than routine workers on average, despite having a higher skill level. One reason for this is that routine workers are more experienced. Additionally, some of the lowest-skilled workers still worked in manual occupations and earned lower wages than routine workers with a similar level of skill. This is consistent with the model's prediction that after a large SR-RBTC the lowest-skilled workers earn a higher income if they work in the routine occupation.

the SR-RBTC hypothesis, they do not necessarily reject potential concurrent mechanisms for the fluctuations in inequality at the bottom half of the distribution. A common way to quantify the contribution of various mechanisms to the overall wage trends is by using a decomposition method (Juhn et al., 1993; DiNardo et al., 1996; Katz and Autor, 1999). Yet, previous attempts to decompose wage polarization have found that technological changes cannot explain the overall wage trend (Autor et al., 2005; Firpo et al., 2013). One important exception is a contemporaneous paper by Acemoglu and Restrepo (2021), who show technological change can explain the majority of the rise of inequality between skill groups. I show that SR-RBTC can explain both the increase and the decrease in inequality, for the entire bottom-half of the distribution, and not just across skill groups.

Most commonly used decomposition methods cannot quantify the impact of SR-RBTC since it violates the "ignorability assumption." Decomposition methods that simulate full counterfactual distributions (Juhn et al., 1993; DiNardo et al., 1996), as well as RIF-regressions (Firpo et al., 2009), all operate under this assumption of ignorability. In this context, the ignorability assumption implies that the distribution of the unobserved skill and its return are the same across occupations. But, the two key effects of SR-RBTC are exactly its impact on the composition of skill and the unequal return to skill across occupations.

Instead, I introduce a skewness decomposition to quantify the effect of SR-RBTC on the overall wage trends.[3] In analogy to inequality that can be measured with the second moment of the log wage distribution, wage polarization can be measured with the third moment of that distribution, namely, skewness. When inequality increases at the top and decreases at the bottom, the log wage distribution becomes more positively skewed and so this moment should increase. As expected, the skewness increases exactly when wage polarization occurs. The main advantage of using skewness to measure wage polarization is that, similar to variance, it can be decomposed into independent components (Mincer, 1974). Unlike in many commonly used decomposition methods, in skewness decomposition these components do not depend on any arbitrary choice of baseline year and are not affected by path dependence (Fortin et al., 2011).

Skewness decomposition breaks the trend of wage polarization into three components for each choice of categorical groups (e.g. occupations, industries, education levels, etc.). Similar to variance decomposition, skewness decomposition has a between-group and a within-group components. The between-group component captures any wage trend that affects all workers in the same group similarly (e.g., an overall decline in wages in routine occupations). The within-group component captures any unexplained trend that

[3]This is the first implementation of skewness decomposition on economics data.

is orthogonal to that grouping.

My focus will be on the third component, which captures violations of the ignorability assumption. This component measures the covariance between the mean and variance in each group. It is higher when higher-earning groups have larger inequality. Hence, if wage gaps decrease in lower-paying routine occupations, while they increase in higher-paying abstract occupations, as the model predicts, this will be captured by this covariance component. Therefore, skewness decomposition is particularly useful to quantify the impact of SR-RBTC. [4]

I find that 93% of the overall wage polarization is driven by occupational trends. I decompose the increase in skewness by occupations using data from the Current Population Survey Outgoing Rotation Groups (CPS-ORG). This data has a large sample size and measures the price of labor most precisely (Lemieux, 2006). I focus on the period of wage polarization during the 1990s, as afterward, inequality at the bottom of the distribution is relatively stable. Ninety three precent of the rise in skewness is driven by either the covariance component or the between-occupation component, and only 7% by the within-occupation component. For comparison, trying to decompose the increase in skewness by other categorical groups such as industry or education can explain only small portions of wage polarization.

The decomposition results indicate that RBTC is skill-replacing. Skill-neutral RBTC models (e.g., Acemoglu and Autor, 2011) assume a decrease in the price of routine tasks. Under this assumption wage changes are driven mainly by a decrease in the premium for routine occupations, and not by any distributional changes within routine occupations. Therefore, skewness decomposition should show a large increase in the between-group component. However, I find that almost the entire increase (78%) is in the covariance component. Moreover, I find that the covariance component increases during the period of wage polarization because of a decrease in inequality in routine occupations, exactly as predicted by the SR-RBTC model. This decrease in inequality in routine occupations was also documented by Lemieux (2007), and causally identified by Gaggl and Wright (2017). Using skewness decomposition I find that the decrease in inequality in routine occupations is actually the main driver of wage polarization, far beyond the decrease in the occupation premiums.

I conclude this paper by discussing alternative explanations for these wage trends and showing they are less consistent with my findings. I also briefly discuss similarities and discrepancies with bottom half inequality in other developed countries and their potential

[4] Since this decomposition could be useful in various other contexts in economics, I developed an R package for public use that can be downloaded from CRAN.

explanation.

2 Model

2.1 Occupational Sorting by Skill

I outline a model that highlights the difference in return to skill in each occupation, building on earlier work by Jung and Mercenier (2014) and Cortes (2016). Assume that workers have a one-dimensional skill, θ_i with some density function $f(\theta)$. This assumption is more general than the assumption of a discrete number of skill levels (Katz and Murphy, 1992; Autor et al., 2006; Acemoglu and Autor, 2011), but less general than assuming multidimensional skills (Roy, 1951).[5]

Occupations are characterized by their return to skill. To simplify, I will assume three occupations: manual, routine, and abstract. In each occupation $j \in \{M, R, A\}$, workers produce an intermediate good with a production function $\varphi_j(\theta_i)$. Assume that in baseline

$$\forall\theta : \frac{\partial \log \varphi_M(\theta)}{\partial\theta} < \frac{\partial \log \varphi_R(\theta)}{\partial\theta} < \frac{\partial \log \varphi_A(\theta)}{\partial\theta} \tag{1}$$

so that the manual occupation has the lowest return to skill, and the abstract occupation has the highest. Assume also that for every θ and every occupation $\frac{\partial \log \varphi_j(\theta)}{\partial\theta} > c$ for some constant $c > 0$, implying that the return to skill is strictly positive at any level.

Under the assumption of perfect competition, wages are set at the marginal productivity. Let p_j be the price of the intermediate good in occupation j. Therefore, if worker i is working in occupation j, she will earn

$$w_j(\theta_i) = p_j\varphi_j(\theta_i)$$

Workers sort into occupations based on comparative advantage. Condition 1 guarantees the existence of two thresholds θ_0, θ_1 such that any worker with $\theta_i < \theta_0$ will choose to work in the manual occupation, any worker with $\theta_0 < \theta_i < \theta_1$ will choose the routine occupation, and any worker with $\theta_i > \theta_1$ will choose the abstract occupation (Jung and Mercenier, 2014). Workers with a skill level that exactly equals the threshold will be

[5] In Section 5 I provide supportive evidence for this restriction by estimating a model with multidimensional skills and show that the skills are strongly correlated.

indifferent; hence the following two equations will hold in equilibrium:

$$\begin{aligned} p_M \varphi_M(\theta_0) &= p_R \varphi_R(\theta_0) \\ p_R \varphi_R(\theta_1) &= p_A \varphi_A(\theta_1) \end{aligned}$$

Figure A2 shows this graphically, by plotting the equilibrium log wages by skill level θ_i.

2.2 Routine-Biased Technological Change

I focus on technological change that improves productivity in the routine occupation. For simplicity, I assume that the technological change affects only φ_R directly, as this change is sufficient for explaining the inequality trends at the bottom of the wage distribution. Hence, φ_M, φ_A are left unchanged. However, wages in the manual and abstract occupations will be affected as well in a general equilibrium.

Specifically, I assume the following functional form:

$$\varphi_R(\theta_i; \tau) = \left(\theta_i^{\frac{\sigma-1}{\sigma}} + \tau^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

where τ is the level of technology and $\sigma > 0$ is the elasticity of substitution between technology and skill.

RBTC would then be modeled as an increase in the technology level τ over time. An increase in τ can be thought of as an improvement in the quality of computers or robots. Thus, RBTC enables every routine worker to produce more since $\frac{\partial \varphi_R(\theta,\tau)}{\partial \tau} > 0$. At the same time, it also enables the production of the same quantity with fewer workers.

While RBTC makes all routine workers more productive, some workers may experience larger productivity gains than others. This could affect income gaps within workers in the routine occupation. Whether income gaps increase, decrease, or remain constant depends on the value of the elasticity of substitution between technology and skill, σ.

Theorem 1. *Let $\theta_a, \theta_b \in (\theta_0, \theta_1)$ be the skill levels of two workers in the routine occupation where $\theta_a < \theta_b$. Let w_a, w_b denote their corresponding equilibrium wages. The effect of an improvement in technology τ on the wage ratio $\frac{w_b}{w_a}$ depends on σ such that*

$$sign\left(\frac{\partial \frac{w_b}{w_a}}{\partial \tau}\right) = sign(1-\sigma)$$

All proofs are given in Appendix A.

If σ equals 1, RBTC is skill-neutral as in Cortes (2016). The effect on log wages will be the same for all workers in the routine occupation. If $\sigma < 1$, as hypothesized by Jung and Mercenier (2014), the new technology increases gaps between skill levels. If $\sigma > 1$, technology is a substitute for skill, and the return to skill declines. For instance, cashiers no longer need to use any arithmetic, librarians no longer depend on their memory to locate books, and production workers rely on their physical strength much less than they used to. I will use Theorem 1 to show that $\sigma > 1$ is most consistent with the data.

The key difference between this model and previous models of RBTC is the focus on the substitutability between technology and skill, instead of the substitutability between technology and the routine workers themselves. Previous models have focused on the elasticity of substitution between technology and worker supply in routine occupations, requiring it to be larger than 1 to generate RBTC (Autor et al., 2003; Jaimovich et al., 2021). These papers assume routine workers have a homogeneous skill level (e.g., Acemoglu and Autor, 2011), and so RBTC is essentially skill neutral ($\sigma = 1$). I will show that replacing the assumption that technology substitutes labor, with the assumption that technology substitutes the usage of skill ($\sigma > 1$) yields predictions that are more consistent with the data.

2.3 General Equilibrium

I assume that the three intermediate goods are used jointly to produce a final good. I use M, R, A to denote the total amount produced from each intermediate good, which equals

$$\begin{aligned} M &= \int_{\theta_{min}}^{\theta_0} \varphi_M(\theta) \, \mathrm{d}\theta \\ R &= \int_{\theta_0}^{\theta_1} \varphi_R(\theta) \, \mathrm{d}\theta \\ A &= \int_{\theta_1}^{\theta_{max}} \varphi_A(\theta) \, \mathrm{d}\theta \end{aligned}$$

The final good is the output of a CES function, i.e.,

$$Y = (M^\rho + R^\rho + A^\rho)^{\frac{1}{\rho}}$$

where ρ<0. Hence the three intermediate goods are complementary, as found by Jaimovich et al. (2021).

While RBTC increases the production of routine goods R, routine workers do not necessarily benefit. This depends on whether there is a sufficient demand increase for those additional routine goods. The price of one unit of the routine good p_R would decrease due to the rise in quantity. Because of the complementarities ($\rho < 0$), the increased productivity in the routine occupation, makes the manual and abstract workers more productive as well which raises the prices of the goods they produce. Overall the share of the total output that is spent on routine workers $\frac{p_R R}{Y}$ declines.[6] This is summarized in the following theorem.

Theorem 2. *RBTC (i.e., an increase in τ) generates:*

1. *An increase in the production of the routine good ($\frac{dR}{d\tau} > 0$).*
2. *A decrease in the absolute price of the routine good ($\frac{dp_R}{d\tau} < 0$) and the relative price compared to abstract/manual good ($\frac{dp_R/p_j}{d\tau} < 0$ for $j \in \{M, A\}$).*
3. *A decrease in the share of the total income that is spent on routine goods ($\frac{d\frac{p_R R}{Y}}{d\tau} < 0$).*

These predictions coincide with predictions of various other models for RBTC. This is partly because the above predictions do not depend on any particular value for σ. Therefore, they can also be generated by a technological change that is skill-neutral as assumed in most earlier models. To distinguish between this and other models of RBTC, in the next two sections I will derive unique predictions for the case of a skill-replacing technology.

2.4 Skill-Replacing RBTC: First Stage

I now examine in more detail the case of SR-RBTC, where technology and skill are substitutes ($\sigma > 1$). In contrast to other models of RBTC, in this model, the impact of technology on bottom-half inequality is non-monotonic. I start with the first stage, where the increase in τ is still relatively small, such that the comparative advantage at Condition 1 still holds. A small increase in τ generates wage polarization and additional predictions that can be tested against the data.

Theorem 3. *Assume a skill-replacing technology ($\sigma > 1$). RBTC (i.e., an increase in τ) would generate the following:*

1. *A decrease in wage gaps between routine workers who do not switch occupations.*
2. *The highest skill routine workers would leave the routine occupation ($\frac{\partial \theta_1}{\partial \tau} < 0$).*
3. *The wage for the highest-skilled routine worker (θ_1) would decrease relative to all other workers.*

[6] This prediction was empirically shown by Eden and Gaggl (2018).

Figure 2a illustrates the results of Theorem 3. Since technology is skill-replacing, the return to skill in the routine occupation becomes flatter. This generates lower gaps between workers who stay in the routine occupation. A relative drop in middle-wages occurs since the most significant wage drop is for the highest-earning routine workers, which (empirically) are concentrated in the middle of the overall distribution of skill. As the return to skill declines, some of the highest-skilled routine workers will have their comparative advantage in the abstract occupation, and so θ_1 will drop.

The effect on θ_0 could go either way. If ρ approaches $-\infty$ (Leonteiff) θ_0 will increase, while if ρ is closer to 0 (Cobb–Douglas) θ_0 will decrease. Empirically, it seems that during the 1990s employment in manual jobs did increase, but not as fast as in abstract occupations (Acemoglu and Autor, 2011). In the specific case when θ is distributed uniformly, this could only occur if θ_0 increased, but by a smaller level compared to the decline in θ_1. The following theorem derives additional empirical conditions for this particular case. Appendix A proves a more general version of this theorem for any continuous distribution of θ, and the same conclusions hold.[7]

Theorem 4. *Assume a skill replacing technology ($\sigma > 1$), $\theta \sim U\left[\underline{\theta}, \bar{\theta}\right]$, and $0 < \frac{d\theta_0}{d\tau} < \left|\frac{d\theta_1}{d\tau}\right|$. In the routine occupation, RBTC would generate a decrease in: (i) employment, (ii) within-occupation inequality, and (iii) mean skill level (see Appendix A for formal definitions). Inequality within the abstract and manual occupation will rise and so overall wage trend would be U-shaped (wage polarization), such that for every $\theta_a < \theta_b \leq \theta_1$*

$$\frac{d\frac{w_b}{w_a}}{d\tau} \leq 0$$

and for every $\theta_1 \leq \theta_a < \theta_b$

$$\frac{d\frac{w_b}{w_a}}{d\tau} \geq 0$$

The U-shaped wage trend can be seen in the difference between the red line and the black line in Figure 2a. The productivity increase for routine workers is offset by the drop in prices. Therefore wages in the routine occupation fall relative to the other two occupations. Moreover, among routine workers, the relative drop in wages is most significant for the highest-skilled workers. The abstract occupation expands and now includes some additional less-skilled workers, which increases its within-occupation inequality. Taken

[7] For a general continuous distribution over $\theta \sim F$, the condition $0 < \frac{d\theta_0}{d\tau} < \left|\frac{d\theta_1}{d\tau}\right|$ is replaced by a more general requirement that the increase in employment in the abstract occupation ($\left|\frac{dF(\theta_1)}{d\tau}\right|$) is sufficiently larger relative to the increase in employment in the manual occupation ($\frac{dF(\theta_1)}{d\tau}$). When this conditions holds, the same conclusions are derived.

together these trends generate a U-shaped pattern where wages increase the most at the tails, and decrease the most around the middle of the skill distribution at the new value of θ_1.

In addition to the impact on wages, SR-RBTC also has an effect on employment in each occupation. Since there is not enough demand for all the new routine goods workers could potentially produce, some of them leave and employment in the routine occupation falls ("job polarization"). This decline in employment is driven more by the higher-skilled routine workers. As a result, routine workers become less skilled on average.

Along with employment, inequality within the routine occupation also declines for two separate reasons. First, it declines directly due to the decrease in the productivity gap. Second, it declines indirectly due to the compositional changes that make the remaining workers in the routine occupation more similar in their skill level.

2.5 Skill-Replacing RBTC: Second Stage

Wage polarization stops when middle-skilled workers' comparative advantage is no longer in the routine occupation. This occurs when the first inequality in Condition 1 no longer holds for all θ_i. At that point, higher-skilled workers in the routine occupation continue to leave. However, some of the employment decline is offset by a positive flow into the routine occupation from the bottom of the skill distribution.

Eventually, the comparative advantage flips, and bottom-half inequality starts to increase. At this point, the routine occupation is filled by the lowest-skilled workers. Any further increase in τ will still reduce wage gaps within routine workers. However, due to the decline in p_R, wages relatively drop for the lowest-paid workers as shown in Figure 2b. Hence at the bottom of the wage distribution, inequality could in fact rise.[8]

Yet, while wage trends change, the decline in employment trend still continues as workers continue to leave the routine occupation. Since wages decline in the routine occupation, more workers would prefer to leave and join the manual occupation. These predictions are summarized in the following theorem.

Theorem 5. *Assume a skill-replacing technology ($\sigma > 1$). There exists a $\tilde{\tau}$ such that for any $\tau \geq \tilde{\tau}$ and for any θ*

$$\frac{\partial \log \varphi_R(\theta;\tau)}{\partial \theta} < \frac{\partial \log \varphi_M(\theta)}{\partial \theta} \tag{2}$$

[8]Given that inequality declines within routine workers while it increases between manual and routine workers the overall impact of SR-RBTC on inequality at this stage depends on the specific parameters as well as the exact index used to measure bottom-half inequality.

When $\tau \geq \tilde{\tau}$ routine workers earn the lowest wages. Any additional SR-RBTC ($\tau \uparrow$) decreases employment in the routine occupation ($\frac{d\theta_0}{d\tau}$<0), as well as wage gaps within routine workers who do not switch occupations.

The key reason why the impact of SR-RBTC on the wage distribution changes over time is the change in the composition of routine workers. At first, when routine workers are middle-skilled, the main negative effect is concentrated around the median of the distribution. Later, when the routine occupation becomes a low-skilled job, the negative impact of SR-RBTC is concentrated at the bottom of the distribution.

This model of SR-RBTC is consistent with recent trends in bottom-half inequality, which were previously documented but could not be explained in terms of a skill-neutral technological change. Specifically, a skill-neutral technological change cannot explain why inequality at the bottom of the distribution rose again after its initial decline. It also cannot explain why the relative wage decrease was concentrated in the middle of the distribution when most routine workers are concentrated below the median. The first stage of SR-RBTC corresponds to trends in the late 1980s and 1990s and the second stage to trends in the 2000s and onwards.

The model generates two new predictions that can be tested against the data. First, it predicts a decline in the return to skill in the routine occupation. This generates a decrease in wage gaps for routine workers who do not switch occupations. Second, it predicts a decline in the average skill level of routine workers. Both trends should be sufficiently large so that at some point a reversal of comparative advantage occurs whereby the return to skill and the average skill level become higher in the manual than in the routine occupation.

In the following sections of the paper, I test these empirical predictions and show that they fit well with the data.

3 Methodology

This paper uses two separate empirical techniques. To test the predictions of the SR-RBTC model, I use an interactive fixed-effects model (IFEM). Then, to quantify the share of the overall wage trend that can be attributed to SR-RBTC, I use skewness decomposition.

3.1 Interactive Fixed-Effects Model

To test whether RBTC is skill-replacing, skill-enhancing, or skill-neutral, I estimate the return to skill directly, using an interactive fixed-effects model. This model is similar to a fixed-effects model that accounts for an unobserved skill, with the important distinction that the return to the unobserved skill may vary. Specifically, I will estimate the following equation for a worker i in occupation j in year t:

$$\log w_{ijt} = \beta_{jt} X_{it} + \lambda_{jt} + \alpha_{jt}\theta_i + \varepsilon_{ijt} \tag{3}$$

where λ_{jt} are occupation-year fixed effects and X_{it} is an additional control for experience squared.[9] The only difference from a fixed-effects model is that the individual fixed effects θ_i are interacted with a coefficient α_{jt}. This α parameter will be the focus of the analysis, as it measures the return to skill in each occupation in each year. The model predicts that $\alpha_{routine,t}$ will decrease as the skill-replacing technology advances. I use either three occupational categories (abstract, routine, manual) or nine (defined by the first digit of the occupational code).

The IFEM tests the SR-RBTC model by estimating wage gaps among workers who stay in the same occupation. The SR-RBTC model predicts that inequality within stayers will decline if and only if the technology is skill-replacing (Theorem 1). It is critical to focus on workers who do not switch occupations, since total inequality within occupations may also be driven by a composition effect, hence may decline under a skill-neutral or skill-enhancing RBTC as well.[10] This is especially concerning since we know that there is a substantial employment decline in routine occupations, driven prominently by the highest- and lowest-earning workers in those occupations (Cortes, 2016; Böhm et al., 2019). The change in the log wage gap between two workers with unobserved skill $\theta_b > \theta_a$ who have the same level of experience, and who stay in occupation j between period t' and t'' is on average $\left(\alpha_{jt''} - \alpha_{jt'}\right)(\theta_b - \theta_a)$. Hence the change in $\alpha_{j,t}$ determines the inequality trend among stayers.

Like a standard fixed-effects model, the interactive fixed-effects model also suffers from a nuisance parameter problem. Each θ_i can be estimated only from observations of a specific worker. Hence any estimation of it would be highly noisy. While the values of the θ_i parameters are not the focus of the analysis, this could still bias the estimates for α_{jt}. To solve this, I derive moments that do not depend on θ_i, as with a standard fixed-effects model. Define ν_{ijt} as

[9] I do not control for education level and experience as they are collinear with θ_i and λ_{jt}.

[10] Estimating wage gaps only among workers who do not switch occupations is not trivial. First because of regression to mean. Second, because workers who do not switch occupations are a selected sample. Finally, because it is hard to distinguish between a systematic change in wage gaps to a change in wage volatility. IFEM addresses all these problems.

$$\nu_{ijt} = \frac{1}{\alpha_{jt}}\left(y_{ijt} - \beta_{jt}X_{ijt} - \lambda_{jt}\right) = \theta_i + \frac{\varepsilon_{ijt}}{\alpha_{jt}}$$

For a given i, we can get a noisy estimate of θ_i as a function of y_i, X_i and the parameters α, β, λ by taking any weighted average over ν_{ijt},

$$\widehat{\theta}\left(y_i, X_i, \alpha, \beta, \lambda\right) = \sum_t \omega_{it}\nu_{ijt} = \theta_i + \widetilde{\varepsilon}_i \tag{4}$$

where $\sum_t \omega_{it} = 1$ for every i and $\widetilde{\varepsilon}_i = \sum_t \frac{\omega_{it}}{\alpha_{jt}}\varepsilon_{ijt}$. We can then use $\widehat{\theta}_i$ to calculate a linear combination of ε_{ijt} that does not depend on θ_i

$$\epsilon_{ijt}\left(y_i, X_i, \alpha, \beta, \lambda\right) = y_{ijt} - \beta_{jt}X_{it} - \lambda_{jt} - \alpha_{jt}\widehat{\theta}_i = \varepsilon_{ijt} - \alpha_{jt}\widetilde{\varepsilon}_i$$

I set $\omega_{ijt} = \frac{\alpha_{jt}^2}{\sum_{j't'}\alpha_{j't'}^2}$, which minimizes the mean squared error $\overline{\epsilon_{ijt}^2}$.

I estimate this model using the method of moments. The basic assumption of the model is that for every occupation j and every year t,

$$E\left[\varepsilon_{ijt}|j,t\right] = E\left[X_{it}\varepsilon_{ijt}|j,t\right] = E\left[\theta_i\varepsilon_{ijt}|j,t\right] = 0$$

Hence, as with a fixed-effects model, I can use moments of the form

$$E\left[\epsilon_{ijt}\left(y_i, X_i, \alpha, \beta, \lambda\right)|j,t\right] = 0$$
$$E\left[X_{ijt}\epsilon_{ijt}\left(y_i, X_i, \alpha, \beta, \lambda\right)|j,t\right] = 0$$

This generates $(p+1) \times J \times T - 1$ moments, where p is the number of X variables, J the number of occupations, and T the number of time periods.[11]

However, the total number of parameters is $(p+2) \times J \times T - 2$. To see why, note that the model has three types of parameters: β parameters ($p \times J \times T$), λ parameters ($J \times T$), and α parameters ($J \times T$). There are two degrees of freedom since θ_i can be identified only up to a linear transformation. Therefore, I pin $\alpha_{Abstract,1980} = 1$ and $\lambda_{Abstract,1980} = 0$. Taken together, there are $(p+2) \times J \times T - 2$ parameters.

Altogether, $J \times T - 1$ additional moments are missing, and so an additional instrument is necessary. This is not surprising, as we added $J \times T - 1$ parameters to the standard fixed-effects model by allowing the return to skill to vary. For this instrument to be informative it needs to be correlated with workers' skill θ_i.[12] However, we cannot simply

[11] The moments of the form $E\left[\epsilon_{ijt}\left(y_i, X_i, \alpha, \beta, \lambda\right)|j,t\right] = 0$ are linearly dependent, similar to a standard fixed-effects model.

[12] For instance Holtz-Eakin et al. (1988) use lagged variables as instruments. An alternative approach is to

use $\widehat{\theta}_i$ as it is correlated with ε_{ijt}.

A natural choice for such an instrument is years of schooling. Denote by S_i the years of schooling for individual i. I focus on a subsample of observations where workers have finished their schooling period, and so $S_{ijt} = S_i$. This adds $J \times T - 1$ moments of the form[13]

$$E\left[S_i \epsilon_{ijt}\left(y_i, X_i, \alpha, \beta, \lambda\right)|j,t\right] = 0$$

The identification assumption is $E\left[S_i \varepsilon_{ijt}\right] = 0$. This assumption implies that conditional on X_{ijt}, years of schooling affects wages only through its impact on overall skill $E\left[\theta_i|S_i\right]$.[14] Put differently, it implies that the return to education is identical to the return to any other skill that is not captured by years of schooling. This is already implied by the model that assumes only a one-dimensional skill, and hence no additional assumptions are needed.[15]

It is possible to estimate a more general model where skill can vary by occupation. Hence θ_{ij} is constant throughout time t, but varies across occupations j. This model would capture cases where different skills are used in different occupations. As a robustness test, I estimate this model as well, using the same instrument. The results in Appendix C indicate that θ_{ij} are very correlated across j, hence the one-dimensional skill assumption approximates the data quite well.

3.2 Decomposing Wage Polarization

Even if all the predictions of the SR-RBTC model were corroborated by the data, it would not disqualify other mechanisms that are potentially occurring simultaneously. For instance, institutional changes such as an increase in the real minimum wage (Piketty, 2014), or a decrease in the unionization rate (Firpo et al., 2013) can also coincide with SR-RBTC and potentially explain a significant portion of the trends as well.

Quantifying the importance of various potential explanations is often done using a

assume that ε_{ijt} has a constant variance that does not change over time, as in Ahn et al. (2001). This assumption allows them to use changes to the variance of ν_{ijt} in order to estimate α. However, this assumption is violated if wage volatility is not constant over time.

[13]Since S_i is constant for a given i, these moments are linearly dependent, and so their total number is $J \times T - 1$.

[14]The assumption still holds if S_i is only a signal for the skill level θ_i (Spence, 1973).

[15]This approach is similar to 2SLS where $\widehat{\theta}_i$ is the endogenous variable. The estimator $\widehat{\theta}_i$ is only a proxy for θ_i and is therefore endogenous. The instruments are S_i interacted with occupation and year, such that the link between education and skill can change over time (Carneiro and Lee, 2011). The moments used for 2SLS are a linear combination of the moments I use in the IFEM, and so would be pinned at zero as well, as the model is exactly identified. However, the approaches are not identical since $\widehat{\theta}_i$ is a function of the estimated parameters α, β, λ.

decomposition. Several decomposition methods have been developed in recent decades exactly for this task. These methods were proven especially useful in the study of the rise in income inequality in the 1980s. By showing that a large portion of the rise in inequality is driven by the rise in the return to education, they provided some of the most important evidence for skill-biased technological change.

Previous decomposition attempts have found that RBTC can account only for a small portion of the wage trends at the bottom of the distribution (Autor et al., 2005; Firpo et al., 2013). Earlier models of (skill-neutral) RBTC hypothesized that the recent wage trends were driven by changes in occupation premiums. Such changes are expected to be captured by the price component of various decomposition methods (e.g., Juhn et al., 1993; DiNardo et al., 1996; Firpo et al., 2009). Yet, the price component was not large enough to explain the main wage trends during this period, leaving room for other potential drivers.

Moreover, the decomposition methods discussed above are unable to quantify the impact of SR-RBTC, as it violates the ignorability assumption that underlies them. Commonly used decomposition methods assume that the distribution of wages conditional on observables does not change when the distribution of observables changes.[16] In this context, the critical observables are occupations or occupation characteristics. This assumption is innocuous if wages decline uniformly for all routine workers, as the skill-neutral RBTC model predicts. However, in an SR-RBTC the distribution of wages within occupations changes considerably, violating this assumption. This is because both the distribution of skill and the return to skill are changing within occupations. This generates a change in wage structure that includes an interaction of an observed characteristic (occupations) and an unobserved characteristic (skill). Most decomposition methods cannot accommodate such interactions without making strong assumptions such as ignorability, which are violated in this model (Fortin et al., 2011).

This problem is also relevant for a recentered influence function (RIF) regression. This method was used by Firpo et al. (2013) to show that at least some share of wage polarization can be attributed to RBTC. Firpo et al. (2013) were the first to document that inequality trends within occupations are asymmetric, and inequality drops in routine occupations, as predicted by the SR-RBTC model. They also suggested a model where the return to skill varies by occupation. However, the RIF regression they used cannot fully account for the impact of these trends on the overall wage polarization, due to the ignorability assumption. Specifically, RIF regression is valid when either skill and occupations

[16] Formally, ignorability assumes that the conditional distributions of wages $F_{w|X}(w|X=x)$ does not vary over time. This assumption implies invariance to conditional distributions, where $F_{w|X}(w|X=x)$ does not change when the marginal distribution of X (F_X) changes.

are independent, or when the distribution of unobserved skill is held fixed within occupations (Firpo et al., 2009). These assumptions are violated in the model, as well as in the data, as I will show.

To address this problem, I use a different decomposition, based on the skewness of the log wage distribution.

3.2.1 Skewness Decomposition

Wage polarization can be measured with skewness.[17] Skewness is the third standardized moment and is defined as

$$S(Y) = E\left[\left(\frac{Y - E[Y]}{\sigma}\right)^3\right] \tag{5}$$

It provides a measure for the asymmetry of the distribution relative to the mean. Appendix Figure A3 demonstrates the link between skewness and wage polarization by plotting the derivative of the empirical influence function at each quantile for a standard normal distribution. Intuitively, the figure shows the effect that a small increase in log wages has on the skewness, for each quantile of the distribution, when log wages are normally distributed. In particular, it shows that skewness increases exactly when wages at the edges increase relative to the middle. This pattern aligns quite well with the observed trends in wages by quantile that were shown by Autor et al. (2006) and I replicate in Section 7.1.

The main advantage of using skewness is that it has a simple decomposition. Letting Y be the standardized log wages, X be the category we wish to decompose by, and μ_3 be the third centralized moment ($\mu_3(Z) = E\left[Z^3 - E[Z]^3\right]$), we can write

$$S(Y) = \mu_3(Y) = \underbrace{E[\mu_3(Y|X)]}_{Within} + \underbrace{\mu_3(E[Y|X])}_{Between} + \underbrace{3COV(E[Y|X], V[Y|X])}_{Correlation} \tag{6}$$

This decomposition was theoretically discussed by Mincer (1974), but was never used on economics data. It is the third-moment equivalent of the variance decomposition formula.[18]

[17] In other contexts, polarization is typically measured with the fourth moment of the distribution (kurtosis). However, the term wage polarization refers to the polarization of the change in wages, where wages increase mostly at the top and at the bottom. The log wage distribution itself is not becoming more polarized or bipolar, and therefore the kurtosis will not necessarily change.

[18] The variance of log wages, which is commonly used to measure wage inequality, can be decomposed into two independent components for the variance within and between categories $V(\log w) = \underbrace{E[V(\log w|X)]}_{Within} + \underbrace{V(E[\log w|X])}_{Between}$.

The first and second components are quite standard. The first component $E[\mu_3(Y|X)]$ can be thought of as a "within" component. It captures the remaining skewness within each category. This component increases when the division into categories is orthogonal to the increase in skewness, and therefore can be thought of as a residual component. The second component, $\mu_3(E[Y|X])$ captures skewness between groups, which is the skewness due to differences between group averages. This component increases if the increase in wage polarization is due to a similar change in wages for all workers in a group, compared to other groups. This includes changes in occupation premiums, return to education, etc.

The third component captures the correlation between the levels and inequality in each categorical group. Formally, this component measures the covariance between the conditional mean and variance for each value of X. When highly paid groups also have larger inequality, inequality will be higher at the top than at the bottom of the overall distribution, making the distribution more positively skewed.

With this covariance component, we can capture trends that violate ignorability. The covariance component allows us to have interactions between unobserved characteristics (e.g., ability), and observed characteristics (e.g., occupation). Hence it can quantify changes to the wage structure that cannot be detected by other methods.

Thus, the covariance component allows measuring the wage impact of SR-RBTC. According to SR-RBTC, inequality increases in the abstract occupation, since skill gaps increase when lower-skilled workers join this occupation. By contrast, inequality decreases in the routine occupation (Theorem 4). The effect on inequality in the manual occupation could go either way, yet since manual occupations are only a small portion of all occupations, their overall impact would be small. Taken together, we expect a wage trend that is exactly captured by this component: inequality is rising in the higher-paying more abstract occupations and declining in the lower-paying more routine occupations. Indeed, the covariance component will turn out to be responsible for most of the increase in skewness during the period of wage polarization.

Skewness decomposition also provides another test of whether RBTC is skill-replacing or skill-neutral. In SR-RBTC, most of the increase in skewness is due to the increase in the covariance component. By contrast, in a skill-neutral RBTC, most of the increase in skewness should be due to the between component. This is because according to a skill-neutral RBTC, recent wage trends are driven mostly by the decrease in price of routine goods p_R. Such a decrease in price has an identical effect on all workers in a given occupation. This is exactly the case when we expect a large effect on the between component.[19]

[19]The covariance component can also increase in a skill-neutral RBTC due to some particular composi-

Skewness decomposition has several important properties that make its results more robust. It allows quantifying wage polarization using a single index. Similar to variance decomposition, skewness decomposition breaks skewness into independent components. This means that there is no problem of path dependence nor any need to arbitrarily define a baseline year, which are common problems in other popular decomposition methods (Fortin et al., 2011).

While in this paper I use skewness decomposition to study wage polarization, it could also be applied to any distribution where the third moment is of interest. There are various cases in economics where we know that the distribution is very skewed, and the level of skewness has important implications. Some examples are the distribution of the return to patents, firm productivity, the distribution of capital ownership, and raw wages (without logs). Any variation in these distributions over time or across places can be analyzed with skewness decomposition. To simplify and encourage the usage of skewness decomposition by more researchers I provide an R package that implements it.[20]

4 Data

This paper combines three sources of data. To estimate the interactive fixed-effects model panel data is required. I use the Panel Study of Income Dynamics (PSID) between 1980–2011. This data was chosen due to its long panel. I measure income using hourly wage (annual income divided by hours worked) as this best captures the real price of labor, which is the focus of the model.[21] I use the full core sample (SRC) without weights for every individual whose wage is available.[22]

Whenever a panel structure is not needed, including the skewness decomposition exercise, I use a larger dataset from the Current Population Survey Outgoing Rotation Group (CPS-ORG). The CPS-ORG provides the most accurate representative sample of hourly wages (Lemieux, 2006). I use the same sample definition as given in Acemoglu and Autor (2011). Observations with missing wages are dropped. The main results hold when using imputations instead. Sampling weights are used in all analyses.

tional changes. This is why I also estimate the IFEM, which shows that inequality within routine occupations declines regardless of any compositional changes.

[20]The package offers both skewness and variance decomposition. It can be easily applied to any data, for any choice of an outcome variable Y and a categorical variable X. It also provides an analytical calculation of the standard errors.

[21]Inequality trends in the PSID are similar to the ones in the CPS.

[22]I include workers who appear only once in the data because, even though they do not affect the estimation of the IFEM parameters, they still affect the estimation of the average skill level in each occupation.

One important limitation of hourly wage data is its relatively high level of measurement errors. This problem is particularly severe at both tails of the distribution. Misreporting of working hours could lead to extremely high or extremely low values of hourly wages. Therefore, I drop the top and bottom 5% of the positive wages throughout the paper. The level of 5% minimizes the drop of data, without generating substantial fluctuations between consecutive years in the skewness estimator. It is also similar to the data cut choices made by earlier papers in this literature (Katz and Murphy, 1992; Autor et al., 2008). Smaller cuts also yield similar but noisier results, particularly for the skewness estimates.[23]

Most of the skewness decomposition analysis is focused on the years between 1992–2002. This is due to a significant revision of the occupational classification system that took place before and after this period, which makes comparisons to other years less precise. As I will show, most of the increase in polarization occurred during this time period. For robustness, I also implement an analysis over a longer period using the occupational crosswalk constructed by Autor and Dorn (2013), and show the main results hold.

I maintain a consistent definition for routine occupations, similar to earlier papers in the literature. I first translate all versions of occupational coding into a uniform coding, using the Autor and Dorn (2013) crosswalk. I then define all administrative, operator, and production occupations as routine, based on their 1-digit category. All managerial, professional, and technician occupations are classified as abstract. Sales, services, and agricultural occupations are classified as manual. This is a similar classification to that used in previous studies (e.g., Acemoglu and Autor, 2011) with one important exception: I do not classify sales occupations as routine occupations.[24] The analysis by 1-digit occupational category is fully consistent with previous literature.

Finally, I use data from the Occupational Information Network (O*NET) to measure the routine-level of each occupation more accurately, in cases where a continuous index can be accommodated. This dataset contains 400 scales to describe various aspects of each occupation, based on a worker survey. I use the same index for the routine content of occupations as Acemoglu and Autor (2011). This index summarizes six questions that proxy the routine level of the job.

More details on the data are provided in the data appendix (B).

[23] Cornfeld and Danieli (2015) analyze skewness in the Israeli labor market, using the entire distribution since measurement errors in the Israeli data are not as severe. They reach very similar conclusions.

[24] While information on task components suggests that sales is a routine occupation, I do not see the same wage and employment patterns in sales as in other routine occupations. Be that as it may, including sales in the routine occupations will not strongly affect the results as it is a small share of workers compared to the other routine occupations. However, it will make the results weaker. One potential explanation for this is that sales occupations are not as easily automated as one might infer from their O*NET description.

5 The Decline in Return to Skill in Routine Occupations

The main prediction of the SR-RBTC model is that the return to skill declines in the routine occupation. Based on Theorem 1, such a decrease in the skill gap is only consistent with a skill-replacing RBTC ($\sigma > 1$). In this section, I provide empirical evidence for this prediction. I first show reduced-form evidence that the education premium has declined in routine occupations. I then use an interactive fixed-effects model to get a direct estimate of the return to skill and its trends in all occupational categories.

5.1 The Education Premium in Routine Occupations

In cross-sectional data, the education premium reflects not only the return to education, but also potential differences in unobserved ability across education groups (Card, 1999). Therefore, any changes in the education premium over time could reflect both changes in the return to skill, as well as changes in the skill composition of workers across education levels. In order to focus on the changes in the return to skill, I measure the changes in the education premium in a panel setting, controlling for differences in ability.

I estimate the education premium using a (standard) fixed-effects model. I define education level by years of schooling. Specifically, I use the following model to estimate the education premium for the subsample of routine workers in the PSID

$$\log w_{it} = \gamma_t S_i + \psi_t + \theta_i + \rho_t X_{it} + \varepsilon_{ijt} \tag{7}$$

where S_i measures years of schooling, ψ_t are year fixed effects, θ_i are worker fixed effects, and X_{it} is an additional control for experience squared. I focus on the coefficient γ_t, which captures how much the wage gap between more- and less-educated routine workers changes over time.

This specification focuses on changes in the return to skill, holding skill composition fixed. This is done by focusing on wage changes over time. Assuming that the skill differences between workers is constant over time, the fixed effects (θ_i) guarantee that any trend in γ_t is not driven by compositional changes, which are controlled for. For instance, if $\gamma_t > \gamma_{t+1}$, it implies that for a given set of workers, the gap between more and less educated workers is decreasing over time.

I find that the wage gap between more and less educated workers has decreased since the late 1980s. Appendix Figure A4 plots the estimated coefficient γ_t relative to its value in 1980.[25] While the results are noisy due to the small sample size of the PSID, they still indicate that since the late 1980s the education premium has declined in routine occupations. This is consistent with the timing when bottom-half inequality starts to decrease. In order to account for skill differences within education levels, compare the trends in routine occupations to abstract and manual occupations, and improve precision, I estimate the interactive fixed-effects model.

5.2 IFEM Results

The IFEM estimation results are consistent with the predictions of the SR-RBTC model. I estimate Equation 3 as described in Section 3.1. I start by analyzing the estimation results in detail for specific years, before analyzing the full sample period.

In 1987, approximately the last year before bottom-half inequality starts to decline, the estimation results are consistent with the initial pre-SR-RBTC model predictions. Panel A of Figure A5 plots the expected log wage of workers in 1987 as a function of their skill θ_i in the three different occupational categories. This figure highly resembles the theoretical prediction in Figure 2a. Return to skill, α_{jt}, which is the slopes in the graph is highest in the abstract category, lower in the routine category, and lowest in the manual category. As a result, the lowest-skilled workers can earn their highest wage in manual occupations, the highest-skilled workers can earn the highest wage in abstract occupations and middle-skilled workers earn the highest wage in routine occupations. Moreover, the indifference point between routine and abstract occupations is very close to zero, which is the average skill level in the economy. Hence, if workers sort optimally into occupations, the highest-skilled routine workers have an average skill level.

Following the model prediction, by 2011, there is a reversal of comparative advantage. Panel B of Appendix Figure A5 plots the expected log wage as a function of a worker's skill for the three occupational categories in 2011. At this point, the slope is lowest in routine occupations, which reflects their lowest return to skill ($\alpha_{routine,2011}$). The lowest-skilled workers are expected to earn the highest wages in the routine occupation.[26]

Figure 3 extends the analysis over more years and shows that the return to skill has steadily declined in routine occupations since the late 1980s. This is approximately when wage polarization and the decline in routine employment start. The figure plots $\alpha_{j,t}$ in

[25] Since this specification is only focused on changes, it cannot estimate the absolute return to education.

[26] Unlike in the model, the IFEM estimation results predict that wages in abstract occupations are on average higher than in manual occupations, for every skill level. Workers still choose to sort into manual occupations for reasons beyond the scope of the model, such as transition costs, search frictions, compensating differentials, multidimensional skills, etc.

log units for the three broad occupational categories. Since there is a degree of freedom in this estimation, I pin $\log \alpha_{abstract,1980}$ to 0. The figure shows that the return to skill in routine occupations has dropped substantially. The value of $\log \alpha$ decreased by more than 0.7, which correspond to a 50% cut between its peak value in 1987 and 2011. This means that conditional on age, the average return to skill was reduced by more than half. Hence, skill gaps were substantially compressed in routine occupations among stayers.

The other two occupational categories did not see a similar sharp decline. For manual occupations, $\log \alpha_{manual,t}$ remains very stable at around -0.2. In the late 1980s, the return to skill in manual occupations was below that in routine occupations, as assumed in the model (Condition 1). Yet because the return to skill declined in routine occupations while remaining relatively stable in manual occupations, their ranking reversed in the 1990s. This matches the prediction of Theorem 5. Abstract occupations also see some decline in return to skill, mostly after 1994, which supports recent evidence on a reversal in demand for cognitive skill (Beaudry et al., 2016). While interesting in itself, this decline is significantly smaller compared to the decline in routine occupations and is not large enough to change the ranking of occupational categories based on their return to skill.[27]

I find similar results when repeating the analysis separately by gender. I estimate Equation 3 separately for females and males. The results are presented in Appendix Figure A6. For both genders, I find a similar decrease in the return to skill in routine occupations.

The same pattern of results emerges when using nine occupational categories based on 1-digit occupational coding. I estimate Equation 3 allowing the return to skill (α_{jt}) to vary by 1-digit occupation category and year. Figure 4 plots the coefficient for α_{jt} in log units for each 1-digit occupation category in three years: 1985, 1997, 2011.[28] In 1985, before wage polarization starts, the return to skill is in accordance with the assumption of the model: the return to skill is largest in the abstract occupations (managers, professionals, and technicians), lowest in the service occupations, and in between for routine occupations (administrative, operator and production). One noticeable exception is sales occupations, which seem to have a return to skill in the range of the abstract occupations despite often being classified as routine.

[27] Figure 3 documents a decline only in return to skill in abstract occupations ($\frac{\partial \log \varphi_A(\theta)}{\partial \theta}$), and not a general decline in the occupation premium (p_A). Wages in abstract occupations are still higher relative to other occupations, as seen in panel B of Figure A5 for 2011. Inequality within abstract occupations is still rising, potentially due to lower-skilled workers joining these occupations. See Section 7.2 for further discussion.

[28] In this plot I omit agricultural workers, who comprise only a small share of the labor force. Yet these workers experience similar trends to other manual workers in the service sector, as seen in Appendix Figure A7.

The return to skill then drops only in the routine occupations. All four routine occupations, including sales, see a decline in return to skill between 1985–1997. At the same time, the other four occupations (managers, professionals, technicians, and services) experience an increase in their return to skill. Later, between 1997–2011, there is a decline in the return to skill in all occupations, but it is sharper in the routine ones, and especially in the administrative category. By 2011, the occupations with the lowest return to skill are the four routine occupations.

Appendix Figure A7 shows the trends in more detail. It plots the value of $\alpha_{j.t}$ for each 1-digit occupational category by year. Administrative workers and operators and to some extent also production workers, the three occupations classified as routine, experience a significant drop in return to skill. By contrast, in service and agricultural occupations, the return to skill has not shown any decline. During the 2000s service occupations had a return to skill around –0.1 in log units, substantially higher than the routine occupations during this time. This could potentially explain why workers from routine occupations decided to switch service occupations (Autor and Dorn, 2013), which I discuss in more depth in the next section.

In accordance with the SR-RBTC model, the IFEM assumes a one-dimensional skill. I test the implication of this assumption in Appendix C. I estimate the model allowing for multidimensional skills where θ_{ij} can vary by occupation. I find that skills are highly correlated across occupations, hence the one-dimensional skill assumption is a reasonable approximation for the purposes of this research.

Overall, these results fit well with the predictions of the skill-replacing RBTC model. They rule out a skill-neutral or skill-enhancing technological change. Theorem 1 shows that a decrease in the skill premium is consistent with a skill-replacing technology ($\sigma > 1$). The interactive fixed-effects model shows that (conditional on cohort) wages of routine workers who do not switch to a different occupational category become more similar and less affected by skill. This process started in the late 1980s and continued until 2011. This fits the prediction of Theorems 3 and 5 that wage gaps between stayers will decrease in both stages of RBTC.

The drop in the return to skill in routine occupations is most harmful to the highest-skilled routine workers. Workers with the largest values of θ_i are most affected when $\alpha_{j,t}$ drops. This is why the model predicts that these workers will leave routine occupations (Theorem 3), and that overall routine occupations will become less skilled (Theorem 4). In the next section, I provide evidence for such employment trends.

6 The Reversal of Comparative Advantage

This section presents evidence that the employment decline in routine occupations was predominantly driven by higher-skilled routine workers. As a result, routine occupations' average skill level fell below that of manual workers. This explains why inequality at the bottom of the distribution stopped declining and started rising back, even though SR-RBTC continued.

I estimate each worker's skill using the interactive fixed-effects model. For each worker i, I estimate $\widehat{\theta_i}$ using Equation 4. Since $\widehat{\theta_i}$ is estimated separately for every worker, it relies only on a small number of observations, making it a very noisy estimate. To solve this problem, I analyze the average value of $\widehat{\theta_i}$ for large groups of workers.

Specifically, I examine the average value of $\widehat{\theta_i}$ for each occupational category in a given year. I do this both by dividing occupations into the three main categories (abstract, routine, manual) as well as by the 1-digit classification. I normalize $\widehat{\theta_i}$ to have a mean of zero for each cohort based on the year in which the worker entered the labor market. Therefore this is an analysis of relative skill within cohorts.[29]

I find a substantial and steady decline in the average skill level of routine workers. Figure 5 plots the average skill level by occupational category and year. At the beginning of the sample period, in the early 1980s, routine workers were middle-skilled. Their average $\widehat{\theta_i}$ was very close to zero, which is the population average. Over the next three decades, the skill composition of routine workers steadily declined, reaching -0.2 at the end of the period. Since the return to skill in routine occupations in 2011 ($\alpha_{routine,2011}$) was about 0.4 (-0.9 in log units), it follows that if routine workers had been as skilled as they were at the beginning of the period they would have earned an additional 8 log points (about 8%) of their 2011 wages.

By the end of the sample period, routine occupations employed the lowest-skilled workers. While the average skill level of routine workers declined, the average skill level of manual and abstract workers remained fairly stable. As a result, in 2009, the average skill level of routine workers fell for the first time below that of manual workers.[30]

I find similar trends when repeating the analysis separately by gender. Appendix Figure A8 presents the average skill level by occupational category and year when estimating

[29] The estimated skill $\widehat{\theta_i}$ is higher for older cohorts due to the positive return to experience.

[30] Since the analysis is done within cohort, it does not imply that wages in routine occupations are lower than in manual occupations. Older workers who earn higher wages due to their experience are more likely to work in routine occupations. Moreover, as shown in Panel B of Appendix Figure A5, low-skilled workers who still sort into manual occupations earn even lower wages than similar-skilled routine workers. This is because the return to skill is higher in manual occupations.

Equation 3 separately for the subsample of females and males. For both genders, I find a decline in the average skill level of routine workers. However, for females, average skill levels for routine workers are still higher than manual by 2011.

I find very similar trends using the 1-digit classification of occupations. Figure 6 plots the average skill level in 1985, 1997, and 2011. At the beginning of the period, routine occupations were middle-skilled, where all four routine occupational categories (including sales) had a skill level between -0.1 to 0.1. In the following periods, administrators, operators, and production workers became significantly less skilled.

Other occupations, like services, actually saw a rise in the average skill level of their workers. In 2011, service workers had higher skills than administrative workers and operators. This fits well with the prediction of the model that since the lowest-skilled workers were now employed in routine occupations, manual occupations such as services would see an increase in the skill level of their workers. Appendix Figure A9 plots the results for all years.

I also find that while in 1990 many middle-wage workers worked in routine occupations, by 2010 this is no longer the case. Figure 7 plots the average routine intensity in each occupation for each wage percentile. It adds two additional years, 2000 and 2010, to the original plot of 1990 I discussed earlier (Figure A1). Between 1990 and 2000, routine intensity fell mostly for wages above the 40th percentile. In the following decade between 2000–2010, routine employment fell mostly between the 20th and 40th wage percentiles, perhaps because not many workers in routine occupations were left in higher percentiles.

There are at least two potential explanations for this decline in the number of middle-wage routine workers. First, the decline could be driven by middle-skilled workers leaving, or never joining routine occupations. Workers in the middle and the upper half of the wage distribution may have switched to occupations with a lower routine intensity. This corresponds to a decline in θ_1 in the SR-RBTC model, as predicted by Theorems 3 and 5. Second, the decline could be driven by lower wages in routine occupations. This corresponds to a decline in p_R, as predicted by Theorem 2 and empirically shown by Cortes (2016). Either way, routine workers are now concentrated in much lower percentiles of the wage distribution than they were in the past. Therefore, any further RBTC is not expected to generate a decline in middle wages.

These findings fit very well with the model's predictions. Since wages decline mostly for the highest-skilled routine workers, they are the first to leave these occupations, as predicted by Theorem 3. As a result, the overall skill level decreases, as predicted by Theorem 4. At some point, the average skill level of routine workers falls below that of manual workers, as predicted by Theorem 5.

This compositional change explains why inequality at the bottom of the distribution stopped decreasing and started increasing. Once middle-wage workers were no longer employed in routine occupations, they were no longer affected by SR-RBTC as before. Since routine workers were now the lowest-skilled workers, any further SR-RBTC was working mostly against the lowest earning workers. This generated an increase in inequality at the bottom half of the wage distribution.

7 Quantifying the Overall Impact of SR-RBTC

So far I have shown that the main predictions of the model are consistent with the data. This section shows that SR-RBTC is also substantial enough to account for almost the entire trend of wage polarization. Using skewness decomposition, I show that wage polarization is driven almost entirely by occupational trends. Moreover, the effect is not driven by the drop in the premium in routine occupations as predicted by a skill-neutral RBTC. Instead, I find that the decline in inequality in low-paying routine occupations is the main driver of wage polarization, consistent with the SR-RBTC model.

7.1 Evidence From Skewness Decomposition

I start by showing that skewness is indeed a good measure for wage polarization. Figure 8 shows the trend in skewness between 1979–2012. The rise in skewness aligns very well with the timing of wage polarization as depicted in Figure 1. Skewness increased between the late 1980s and the early 2000s, exactly when the 90/50 gap was rising and the 50/10 gap was falling.

The rise in skewness is driven by trends in all parts of the wage distribution. An increase in skewness occurs when the distribution becomes more tilted toward the left-hand side. This corresponds to an increase in the gap between middle and high wages and a decrease in the gap between middle and low wages. Appendix Figure A10 presents a bin scatter of the change in wages between 1992–2002, for 20 quantiles. This generates a U-Shape that was previously shown by Autor et al. (2006, 2008). The U-Shape received qualitatively resembles the EIF derivative plotted in Figure A3. This suggests that skewness rose in this period because of the rise in wages both at the top and at the bottom of the distribution, making it a good fit to measure wage polarization.

I decompose the rise in the skewness of the distribution into three components, namely, within, between, and covariance components, as described in Equation 6, for different choices of categorical groups (X). I first focus on the period between 1992-2002 since data on other years uses different occupational coding (see Section 4). As Figure 8 shows this time period includes a big portion of the overall increase in skewness.

Decomposing by occupations can explain almost the entire rise in skewness. Figure 9 presents the decomposition by 3-digit occupational coding. The figure depicts the change in each of the aforementioned components since 1992, as well as in the sum of the three, which equals the total change in skewness.[31] The first interesting conclusion from this figure is the importance of occupation in explaining the trends in skewness. The within component, which captures the part that is unrelated to occupational trends, explains only a very small share of the overall increase. That small share might also be the result of classification errors.

Most of the increase in skewness is due to the covariance component. This is indicated by the blue area in Figure 9. Table 1 displays the values of each component and its overall contribution to the rise in skewness. The within component accounts for only 7% of the rise in skewness. Hence 93% of the rise in skewness is related to occupational trends. The between component, which is driven by the trends in mean wages in each occupation, accounts for 15% of the overall trend. The majority (79%) of the increase is driven by the rising correlation between the mean and the variance of log wages in occupations. In other words, the rise in skewness is due to the growing correlation between wage levels and inequality levels in each occupation. As I discussed in Section 3, this type of correlation is not captured by other decomposition methods, which is why earlier work potentially underestimated the contribution of occupational trends.

These results align better with the hypothesis that RBTC is generating wage polarization than with institution-related hypotheses. The theory of RBTC argues that its effect is driven predominantly through occupations. Therefore, the fact that wage polarization, as measured with skewness, is driven by occupations greatly supports this hypothesis. By contrast, institutional changes do not operate directly through occupations.

Moreover, the results are most consistent with a skill-replacing RBTC. Earlier models of a skill-neutral RBTC (Autor et al., 2006; Acemoglu and Autor, 2011) argue that there is a drop in the price of routine tasks, making wages fall equally for all workers in routine-intense occupations. Such a trend would have been captured by the between component as it generates the same effect for all workers in the same occupation. While this component can be seen to rise, it generates only 15% of the overall rise in skewness. Instead, the substantial rise in the covariance component suggests that the effect is mostly driven by the asymmetric trends within occupations. Such trends could include a decrease in the

[31] Appendix Figure A11 performs the same exercise using imputed wages for observations where wages are not reported and reaches very similar results.

return to skill in routine occupations, as described in the SR-RBTC model.

The results are not driven by any other worker characteristic in the data. Since occupations are correlated with workers' skill levels or industries, it is important to verify that occupations are not just proxying for some other worker characteristics. In Appendix Figures A12 and A13 I show the same decomposition results by industry, education, and experience. Clearly, in those cases the within component is much larger, suggesting that a great portion of the trend in skewness is unrelated to these categories.

Moreover, most of the increase in the between and covariance components in those cases is due to their correlation with occupations. Appendix D discusses how to decompose by more than one category using a linear model. I then use this method to decompose jointly by occupation and industry or education. The results in Appendix D show that the increase in skewness is driven almost entirely by occupation and not other observables.

Looking at a longer time period yields similar results. In Appendix Figure A14, I run the same decomposition between 1988, when skewness starts to rise, and 2012. Within this time period the occupational coding changes. While this analysis uses an occupational crosswalk (Autor and Dorn, 2013), changes in the baseline coding still might be affecting the results, making changes between 1991–1992 and 2002–2003 potentially biased. With that caveat in mind, we still see a similar pattern in the earlier period between 1988–1992. Most of the increase in skewness is driven by the covariance component. In the period after 2002, when wage polarization stops, skewness is stable, as are the three different components.

7.2 The Decline in Inequality Within Routine Occupations

The increase in correlation between wage levels and inequality that is driving wage polarization could be driven by different explanations. The increase could be because of trends in the wage levels, wage inequality, or perhaps the composition of workers in each occupation. The following section presents evidence that the main driver is the decrease in inequality in low-paying routine occupations, as predicted by Theorem 4.

During the 1990s, inequality trends within occupations were strongly correlated with the wage levels in those occupations. High-paying occupations saw an increase in inequality, while low-paying occupations saw a decrease. Figure A15 shows this by plotting the change in the variance of log wages from the beginning of the studied period (1992/3) to its end (2001/2) as a function of mean log wages. Changes in inequality, measured with the variance of log wages, are correlated with the occupation wage levels. This fact was previously documented by Lemieux (2007).

In fact, the trends in within-occupation inequality can explain the full rise in the covariance component. The covariance component is calculated by using

$$COV\left(E\left[Y|X\right],V\left[Y|X\right]\right)=\sum_{x}\Pr\left(X=x\right)E\left[Y|X=x\right]V\left(Y|X=x\right)$$

where in this case Y is log wages and X is 3-digit occupations. Most of the increase stems from changes in the variance of log wages in different occupations, $V\left(Y|X=x\right)$. To show this, I fix the share of workers and the expected log wage in each occupation to their averages throughout the period. Thus, I allow only the variance to vary between years. Formally, I calculate the following counterfactual partial-equilibrium covariance for t between 1992 and 2002:

$$\widetilde{COV}\left(\overline{E\left[Y_t|X_t\right]},V\left[Y_t|X_t\right]\right)=\sum_{x}\overline{\Pr\left(X=x\right)}\;\overline{E\left[Y|X=x\right]}V\left(Y_t|X_t=x\right) \tag{8}$$

where $\overline{\Pr\left(X=x\right)}$ and $\overline{E\left[Y|X=x\right]}$ are simple averages of the share of workers and mean log wages over all years between 1992–2002.

I find that the asymmetric trends in within-occupation inequality can explain the entire increase in covariance. Figure 10 compares the real value of the covariance to its counterfactual value from Equation 8. The counterfactual trend closely follows the real trend. Therefore, if the share of workers and the mean log wage in each occupation were held fixed, we would still get the same increase in the covariance, and hence the same increase in skewness and wage polarization. Letting the share of workers or the expected log wage vary while holding other factors fixed does not yield any similar results. This exercise demonstrates that the increase in covariance, and hence the increase in wage polarization, is mostly the result of the asymmetric changes in within-occupation inequality, as measured with the variance of log wages.

Wage polarization occurs mainly during the 1990s because that is the only period in which inequality in lower-paying occupations dropped. Figure A16 shows the trend in occupational inequality by decade. For each decade, it presents a bin scatter plot of the average change in within-occupation inequality, measured with the variance of log wages. Occupations are binned into deciles by their mean log wage. The increase in inequality in high-paying occupations is a long-standing trend.[32] However, the 1990s are unique

[32] In Section 5 I showed some decline in the return to skill in abstract occupations. Yet inequality is still rising in high-paying occupations which are mainly abstract. This is potentially driven by a compositional change. As employment in abstract occupations grows, they employ lower-skilled people as well, which

for their decrease in inequality in low-paying occupations. This is why inequality drops at the bottom of the wage distribution only in the 1990s, generating wage polarization instead of the overarching increase in wage inequality across all parts of the wage distribution, as in other decades.

The drop in inequality in low-paying occupations is driven mostly by routine occupations. Figure A17 presents a bin scatter plot of the changes in within occupation variance for routine and non-routine occupations. I divide occupations into 10 bins separately for routine and non-routine occupations, based on their initial wage decile in 1992. I then plot the mean change in the variance of log wages between 1992–2002. While there is some drop in inequality in low-paying occupations that are non-routine, the trend is substantially stronger for routine occupations. This is also in accordance with the findings in Firpo et al. (2013) who show using O*NET data that routine occupations tend to have a stronger decrease in variance.

Overall, these findings fit well with the predictions of the model. Most of the wage polarization is related to occupational trends, which supports the explanation of RBTC. The trend is driven mostly by the asymmetric trends in within-occupational inequality. Inequality is decreasing in low-paying, mostly routine occupations, while it is increasing in high-paying occupations. This fits well with the predictions of the SR-RBTC model in Theorem 4, and explains why we see a U-shaped wage trend during the 1990s.

8 Discussion

This paper makes three distinct contributions. First, I provide a theoretical explanation for the fluctuations in inequality at the bottom half of the wage distribution since the late 1980s. I introduce the notion of SR-RBTC, where new technology does not replace the workers themselves, but rather the skills they use in their work. The reduction in the return to skill reduces inequality levels within routine occupations. This makes overall inequality decline at the bottom half of the wage distribution where most routine workers are concentrated. SR-RBTC changes the skill composition of routine workers, pushing middle-wage workers out of routine occupations. When routine occupations employ mostly low-wage workers, SR-RBTC lowers wages at the very bottom of the distribution, and inequality at the lower half of the distribution increases.

Second, using IFEM I document two new empirical facts that are consistent with this model. I find that the return to skill declined in routine occupations significantly more

generates larger wage gaps within these occupations.

than it did in other occupations. This is consistent with a skill-replacing technology. I also show evidence that the decline in employment in routine occupations is driven mostly by the higher-earning routine workers, who suffer the most from the decline in return to skill. While in the early 1980s the skill level of routine workers was similar to that of the general population, by 2009 routine occupations employed the lowest-skilled workers in the labor force.

Third, using skewness decomposition I show that 93% of wage polarization is driven by occupational trends. These trends cannot be captured by most decomposition methods that rely on the assumption of ignorability, which is violated in an SR-RBTC model. I find that the 1990s are unique for their decrease in inequality within routine occupations. Even though most routine workers were earning below median wages, most of the relative decline in wages was around the middle of the wage distribution, because of the wage compression in routine occupations.

Skewness decomposition could be a useful empirical tool for studying any trend that affects the skewness of a distribution. In economics, we study many skewed distributions (wealth distribution, returns to patents, etc.) or models that generate an increase in skewness (e.g., economics of superstars model, as in Rosen, 1981; König, 2019).

Other explanations do not fit these empirical patterns as well as SR-RBTC. Several alternative explanations for the decline in bottom-half inequality in the 1990s focus on institutional changes, such as an increase in the real minimum wage (Piketty, 2014) or a decline in unionization (Lemieux, 2007). Other explanations focus on high growth and low unemployment rates as potential drivers for the increase in lower wages. Finally, it is possible that trade shocks have lead to some of the changes in the wage distribution (Autor et al., 2013). However, none of these explanations is expected to work through occupations in particular, more than through education levels or industries. For instance, while some occupations are more unionized than others, industries are likely better proxies for unionization status. Moreover, while these mechanisms could generate a decrease in inequality within lower-paying occupations, as part of the overall decrease in lower-half inequality, it is unclear why their impact would be mostly on low-paying routine occupations and not, for example, service jobs.

Among theories that focus on occupational-related trends, SR-RBTC best fits the empirical findings. Generally, theories related to technology seem to fit the data better as most of the trends are related particularly to routine occupations, which can be automated more easily (Autor et al., 2006; Goos et al., 2014). Skill-neutral or skill-enhancing RBTC models are inconsistent with the clear decline in return to skill in routine occupations. They also do not predict the decrease in skill level in routine occupations. Another theory that could potentially explain the bottom-half inequality trends is a positive demand shock for service occupations (Autor and Dorn, 2013). SR-RBTC also predicts an increase in demand for manual occupations due to the complementarities between occupations. Therefore several predictions of the SR-RBTC model overlap with the predictions in Autor and Dorn (2013). But one important distinction has to do with who is expected to leave routine occupations. A demand shock in service occupations should attract more workers from the bottom of the skill distribution as they have a comparative advantage in manual jobs. However, most employment decline in routine occupations is driven by the highest-skilled routine workers, making it more consistent with SR-RBTC.

Institutional explanations are potentially more relevant for bottom half inequality in other countries. Other developed countries have more dominant labor market institutions compared to the U.S. (Blau and Kahn, 2002). Such institutions tend to have a larger impact on bottom-half compared to upper-half inequality. For instance, Broecke et al. (2016) find that minimum wage levels are substantially more associated with bottom half inequality than upper half inequality. As a result, technological changes could have less effect on bottom half inequality in other developed countries. This could explain why similar inequality patterns are not seen in other countries, despite having similar patterns of job polarization (Naticchioni et al., 2014; Goos et al., 2014). Interestingly, in Israel, which is one of the few countries to experience similar fluctuations in bottom half inequality, similar patterns are detected (Cornfeld and Danieli, 2015).

While this paper does not provide causal identification, my findings are consistent with those of previous papers that have studied the causal effect of RBTC on firm wage distribution. Gaggl and Wright (2017) exploit a natural experiment where exposure to technology varies by firms. They find that the new technology generates wage compression within routine workers in a given firm. In this paper, I show that this wage compression is actually the main driver of wage polarization and not a side-effect.

One direction for future research would be to explore why technology in routine occupations is skill-replacing. This could be because skill-replacing technologies are easier to develop. Alternatively, it is possible that employers invest more in technology that requires more skill/training to save higher wages costs (Feng and Graetz, 2020).

References

Acemoglu, Daron, and David Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In . Vol. 4, Part B of *Handbook of Labor Economics*, , ed. Orley Ashenfelter and David Card, Chapter 12, 1043–1171. Elsevier.

Acemoglu, Daron, and Pascual Restrepo. 2021. "Tasks, automation, and the rise in US wage inequality." National Bureau of Economic Research.

Ahn, Seung Chan, Young Hoon Lee, and Peter Schmidt. 2001. "GMM Estimation of Linear Panel Data Models with Time-Varying Individual Effects." *Journal of Econometrics*, 101(2): 219–255.

Autor, David. 2014. "Polanyi's Paradox and the Shape of Employment Growth." National Bureau of Economic Research Working Paper 20485.

Autor, David, and David Dorn. 2013. "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market." *American Economic Review*, 103(5): 1553–1597.

Autor, David, David Dorn, and Gordon H Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, 103(6): 2121–2168.

Autor, David, Frank Levy, and Richard Murnane. 2003. "The skill content of recent technological change: An empirical exploration." *The Quarterly Journal of Economics*, 118(4): 1279–1333.

Autor, David, Lawrence F. Katz, and Melissa S. Kearney. 2006. "The Polarization of the US Labor Market." *American Economic Review*, 96(2): 189–194.

Autor, David, Lawrence F. Katz, and Melissa S. Kearney. 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics*, 90(2): 300–323.

Autor, David, Lawrence Katz, and Melissa S Kearney. 2005. "Rising Wage Inequality: The Role of Composition and Prices." National Bureau of Economic Research Working Paper 11628.

Beaudry, Paul, David A Green, and Benjamin M Sand. 2016. "The great reversal in the demand for skill and cognitive tasks." *Journal of Labor Economics*, 34(S1): S199–S247.

Blau, Francine D, and Lawerence M Kahn. 2002. *At home and abroad: US labor market performance in international perspective.* Russell Sage Foundation.

Böhm, Michael J, Hans-Martin von Gaudecker, and Felix Schran. 2019. "Occupation, growth, skill prices, and wage inequality."

Broecke, Stijn, Glenda Quintini, and Marieke Vandeweyer. 2016. "Wage Inequality and Cognitive Skills: Re-Opening the Debate." National Bureau of Economic Research.

Card, David. 1999. "The Causal Effect of Education on Earnings." *Handbook of labor economics*.

Carneiro, Pedro, and Sokbae Lee. 2011. "Trends in quality-adjusted skill premia in the United States, 1960-2000." *American Economic Review*, 101(6): 2309–49.

Cornfeld, Ofer, and Oren Danieli. 2015. "The Origins of Income Inequality in Israel-Trends and Policy." *Israel Economic Review*, 12(2): 51–95.

Cortes, Guido Matias. 2016. "Where Have the Middle-Wage Workers Gone? A Study of Polarization Using Panel Data." *Journal of Labor Economics*, 34(1): 63–105.

Cortes, Guido Matias, Nir Jaimovich, Christopher J. Nekarda, and Henry E. Siu. 2020. "The Dynamics of Disappearing Routine Jobs: A Flows Approach." *Labour Economics*, 65: 101823.

DiNardo, John, Nicole M Fortin, and Thomas Lemieux. 1996. “Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach.” *Econometrica*, 1001–1044.

Downey, Mitch. 2021. “Partial automation and the technology-enabled deskilling of routine jobs.” *Labour Economics*, 69: 101973.

Eden, Maya, and Paul Gaggl. 2018. “On the Welfare Implications of Automation.” *Review of Economic Dynamics*, 29: 15–43.

Feng, Andy, and Georg Graetz. 2020. “Training requirements, automation, and job polarisation.” *The Economic Journal*, 130(631): 2249–2271.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2009. “Unconditional Quantile Regressions.” *Econometrica*, 77(3): 953–973.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2013. “Occupational Tasks and Changes in the Wage Structure.” *University of British Columbia Working Paper*.

Foote, Christopher L., and Richard W. Ryan. 2015. “Labor-Market Polarization Over the Business Cycle.” *NBER Macroeconomics Annual*.

Fortin, Nicole M, Thomas Lemieux, and Sergio Firpo. 2011. “Decomposition Methods in Economics.” In *Handbook of Labor Economics*. Vol. 4, , ed. Orley Ashenfelter and David Card, 1–102. Elsevier.

Gaggl, Paul, and Greg C. Wright. 2017. “A Short-Run View of What Computers Do: Evidence From a UK Tax Incentive.” *American Economic Journal: Applied Economics*, 9(3): 262–94.

Goos, Maarten, Alan Manning, and Anna Salomons. 2009. “Job Polarization in Europe.” *American Economic Review*, 99(2): 58–63.

Goos, Maarten, Alan Manning, and Anna Salomons. 2014. “Explaining Job Polarization: Routine-Biased Technological Change and Offshoring.” *American Economic Review*, 104(8): 2509–2526.

Goos, Maarten, and Alan Manning. 2007. “Lousy and Lovely Jobs: The Rising Polarization of Work in Britain.” *Review of Economics and Statistics*, 89(1): 118–133.

Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen. 1988. “Estimating Vector Autoregressions with Panel Data.” *Econometrica*, 56(6): 1371–1395.

Hunt, Jennifer, and Ryan Nunn. 2019. “Is Employment Polarization Informative about Wage Inequality and Is Employment Really Polarizing?” National Bureau of Economic Research Working Paper 26064.

Jaeger, David A. 1997. “Reconciling the old and new census bureau education questions: Recommendations for researchers.” *Journal of Business & Economic Statistics*, 15(3): 300–309.

Jaimovich, Nir, Itay Saporta-Eksten, Henry Siu, and Yaniv Yedid-Levi. 2021. “The macroeconomics of automation: Data, theory, and policy analysis.” *Journal of Monetary Economics*, 122: 1–16.

Juhn, Chinhui, Kevin M Murphy, and Brooks Pierce. 1993. "Wage Inequality and the Rise in Returns to Skill." *Journal of Political Economy*, 101(3): 410–442.

Jung, Jaewon, and Jean Mercenier. 2014. "Routinization-Biased Technical Change and Globalization: Understanding Labor Market Polarization." *Economic Inquiry*, 52(4): 1446–1465.

Katz, Lawrence F, and David Autor. 1999. "Changes in the Wage Structure and Earnings Inequality." In *Handbook of Labor Economics*. Vol. 3, , ed. Orley Ashenfelter and David Card, 1463–1555.

Katz, Lawrence F., and Kevin M Murphy. 1992. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." *Quarterly Journal of Economics*, 107(1): 35–78.

König, Felix. 2019. "Technical Change and Superstar Effects: Evidence From the Roll-Out of Television." *Centre for Economic Performance Working Paper*.

Lemieux, Thomas. 2006. "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *American Economic Review*, 96(3): 461–498.

Lemieux, Thomas. 2007. "The Changing Nature of Wage Inequality." *Journal of Population Economics*, 21(1): 21–48.

Lemieux, Thomas. 2010. "What Do We Really Know about Changes in Wage Inequality?" In *Labor in the New Economy*. , ed. Katharine G. Abraham, James R. Spletzer and Michael Harper, 17–59. National Bureau of Economic Research.

Mincer, Jacob. 1974. *Schooling, Experience, and Earnings. Human Behavior & Social Institutions No. 2.* Columbia University Press.

Naticchioni, Paolo, Giuseppe Ragusa, and Riccardo Massari. 2014. "Unconditional and conditional wage polarization in Europe."

Piketty, Thomas. 2014. *Capital in the 21st Century.* Harvard University Press.

Rosen, Sherwin. 1981. "The Economics of Superstars." *American Economic Review*, 71(5): 845–858.

Roy, Andrew D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers*, 3(2): 135–146.

Spence, M. 1973. "Job Market Signaling." *Quarterly Journal of Economics.*

Figures and Tables

Tables

Year	Total Skewness	Within	Between	3COV
1992	.0946	.0334	-.04171	0.1030
$N = 130,502$		(.0037)	(.0032)	(.0053)
2002	.1898	.0397	-.0276	0.1777
$N = 108,145$		(.0045)	(.0036)	(.0062)
Δ 1992-2002	.0951	.0063	.0141	0.0747
		(.0058)	(.0048)	(.0081)
	100%	6.6%	14.9%	78.5%

Table 1: Skewness Decomposition by 3-Digit Occupation

Skewness decomposition based on Equation 6. The three components sum to the overall skewness (Equation 6). Wages at the top and bottom 5% were dropped (see Section 4). Standard errors are calculated analytically (using the delta method).
Source: CPS Outgoing Rotation Groups

Figures

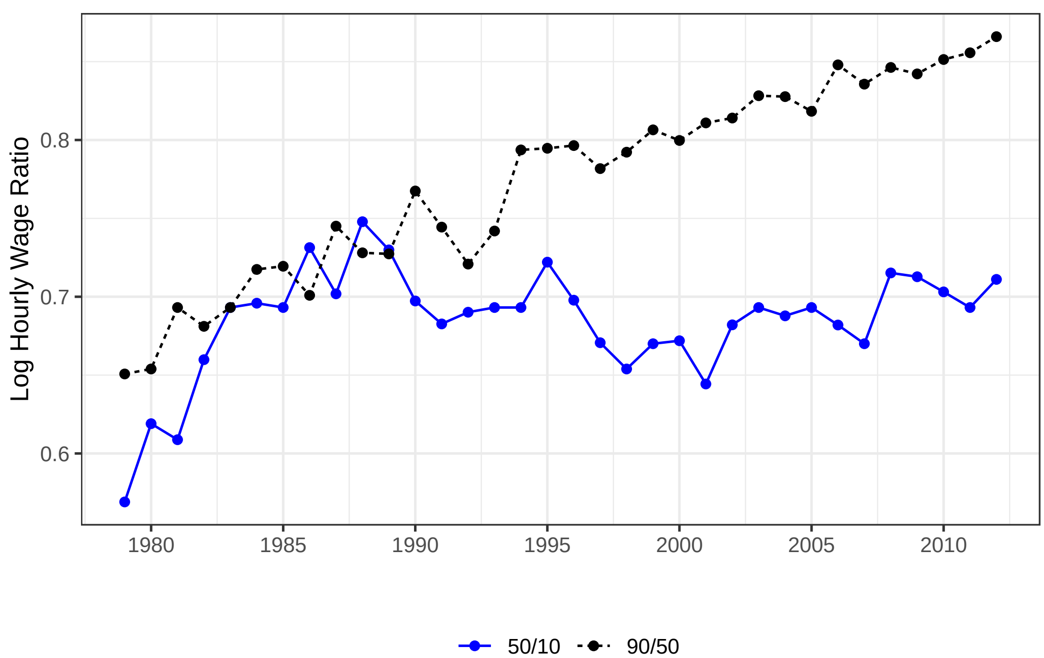
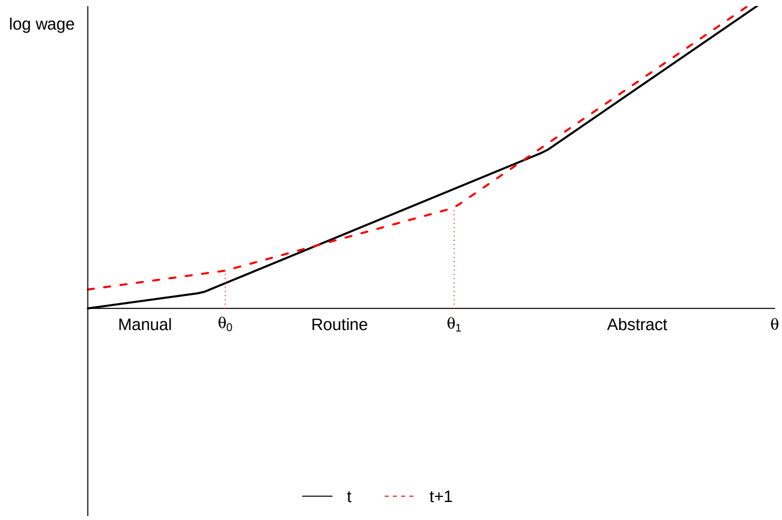
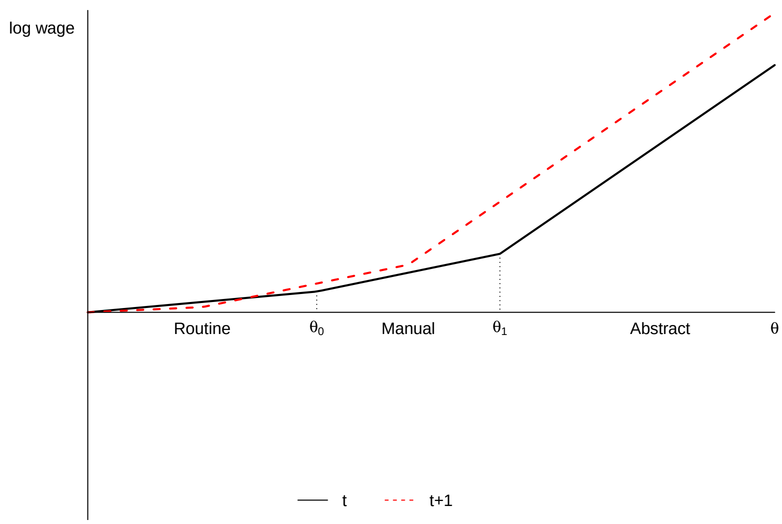


Figure 1: 90/50 and 50/10 Log Hourly Wage Ratio

Quantiles are calculated for all workers with positive earnings at the hours level, using sample weights multiplied by hours worked.
Source: CPS Outgoing Rotation Groups



(a) Change in Log Wage for Small SR-RBTC



(b) Change in Log Wage for Large SR-RBTC

Figure 2: Illustrated Log Wages by Skill in Equilibrium

These figures illustrate the equilibrium sorting of workers into occupations and their log wages as a function of their skill θ. The dashed red line represents equilibrium log wages in a later time period when technology has further advanced (increase in τ). Panel A represents a small technological change, that reduces the slope of log wages as a function of θ only in the routine occupation. Panel B describes the equilibrium after a large technological advancement and a reversal of comparative advantage such that the slope in the routine occupation is lower than the slope in the manual occupation (Condition 2 replaces Condition 1).

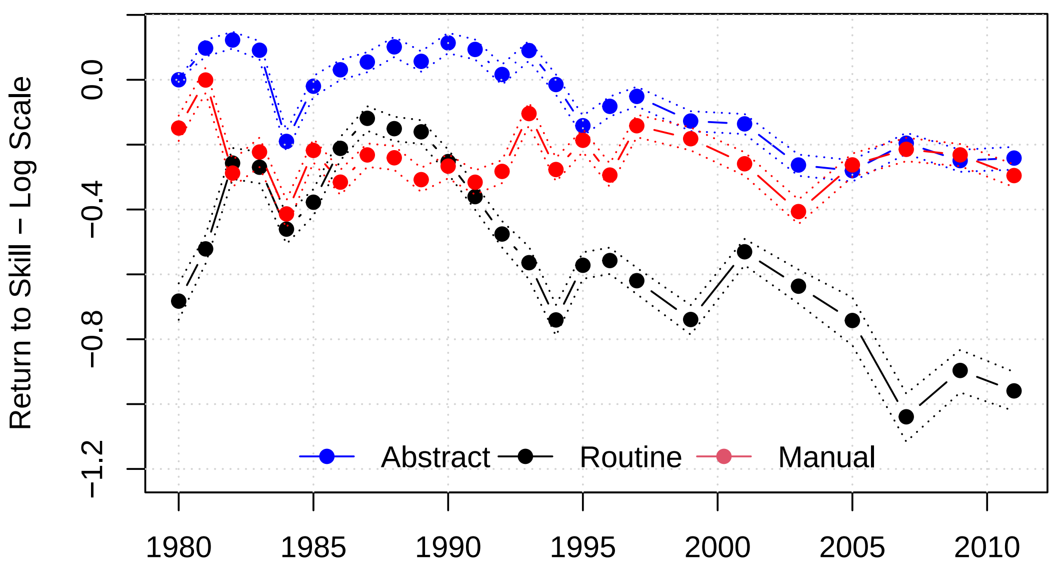


Figure 3: Return to Skill (a_{jt}) by Occupational Category

This figure presents the return to skill (α_{jt}) in log units for the three occupational categories. Return to skill is calculated using an interactive fixed-effects model (Equation 3). The log return to skill in the abstract occupation in 1980 is fixed to zero, hence all other values are relative to that year and occupational category. Routine workers are defined as workers in administrative, production, or operator occupations, classified by the first occupational coding digit. Abstract workers include managers, technicians, and professionals. Manual includes service, sales and agricultural occupations. Autor and Dorn (2013) occupational crosswalk is used for a consistent definition of occupations over time. Dashed lines represent 95% confidence intervals.
Source: PSID

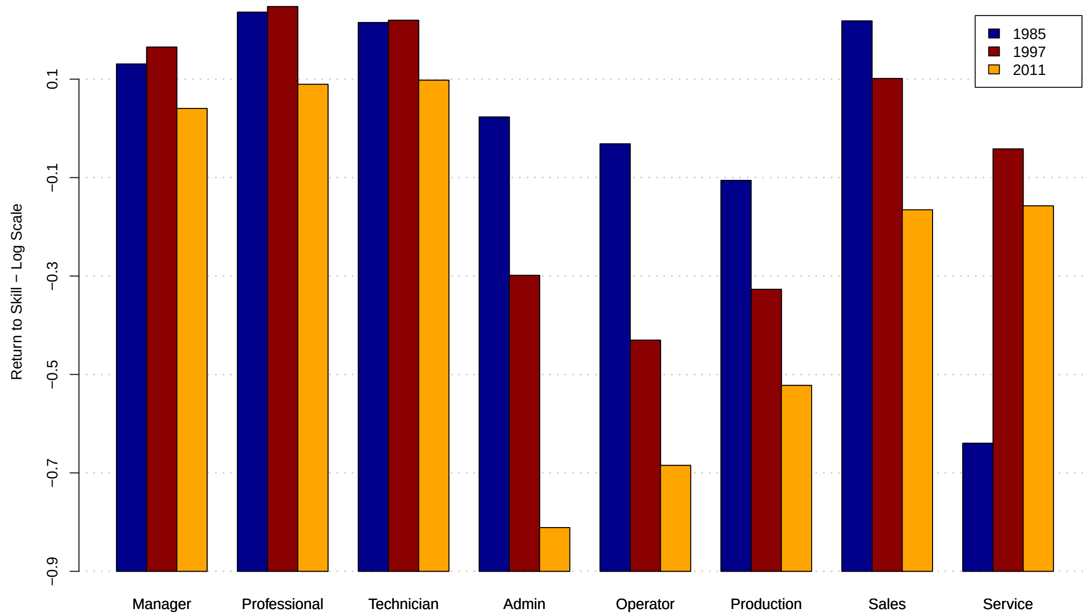


Figure 4: Return to Skill (α_{jt}) by 1-Digit Occupational Category

This figure presents the return to skill (α_{jt}) in log units for eight 1-digit occupational categories. The log return to skill in administrator occupations in 1980 is fixed to zero, hence all other values are relative to that year and occupational category. Results for all years are available in Appendix Figure A7. Return to skill are calculated in an interactive fixed-effects model (α_{jt}, using Equation 3). α_{jt} varies by 1-digit occupation and year. Autor and Dorn (2013) occupational crosswalk is used for a consistent definition of occupations over time.

Source: PSID

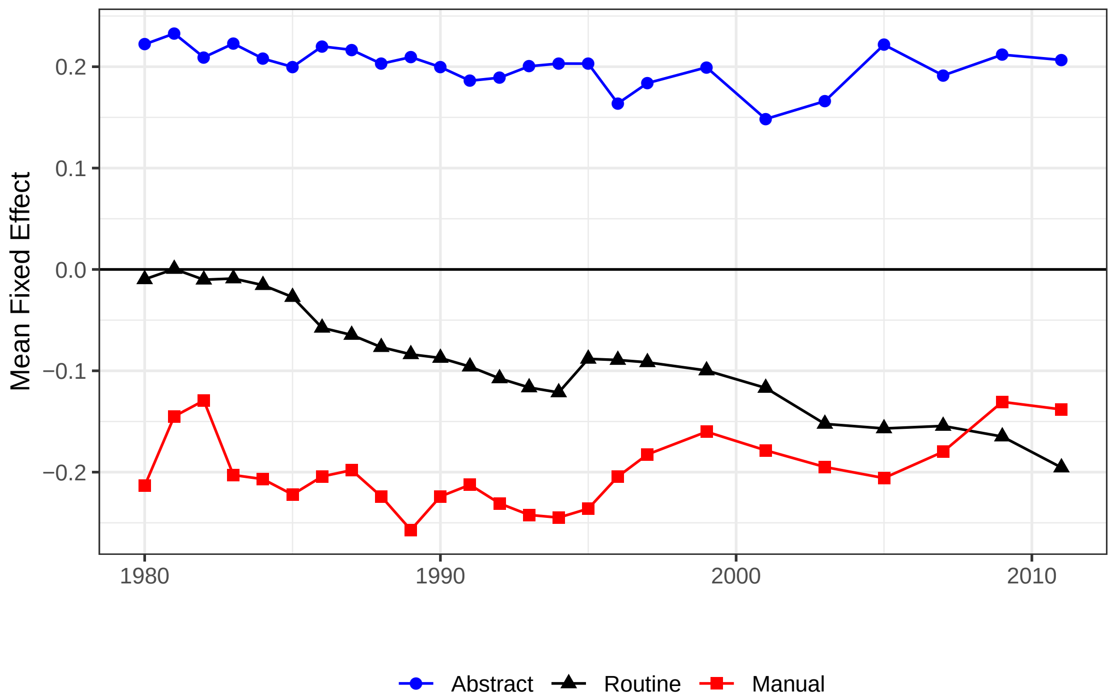


Figure 5: Mean Skill Level ($\widehat{\theta}_i$) by Occupational Category

Mean level of $\widehat{\theta}_i$ by occupational category and year. $\widehat{\theta}_i$ is calculated using Equation 4, and demeaned at the cohort level, where cohorts are defined based on year of entry into the labor market. Routine workers are defined as workers in administrative, production, or operator occupations, classified by the first occupational coding digit. Abstract workers include managers, technicians, and professionals. Manual workers include service, sales, and agricultural occupations. Autor and Dorn (2013) occupational crosswalk is used for a consistent definition of occupations over time.
Source: PSID

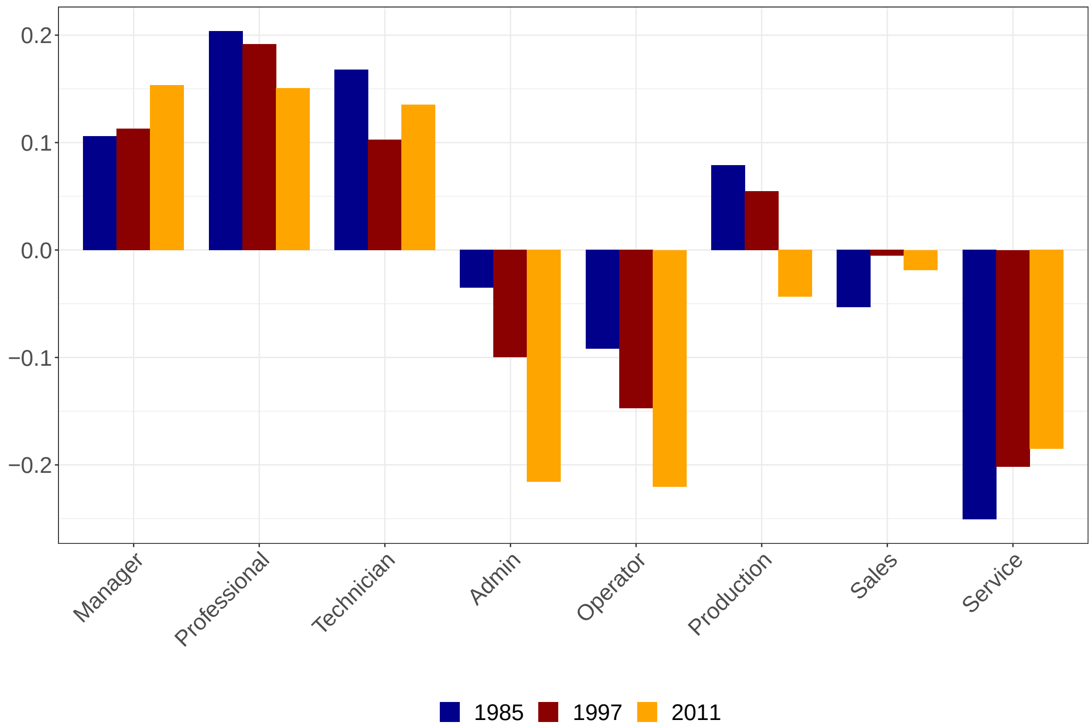


Figure 6: Mean Skill level ($\widehat{\theta}_i$) by1-Digit Occupational Category

Mean level of $\widehat{\theta}_i$ by occupational category and year. $\widehat{\theta}_i$ is calculated using Equation 4 and demeaned at the cohort level, where cohorts are defined based on year of entering the labor market. The Autor and Dorn (2013) occupational crosswalk is used for a consistent definition of occupation over time.
Source: PSID

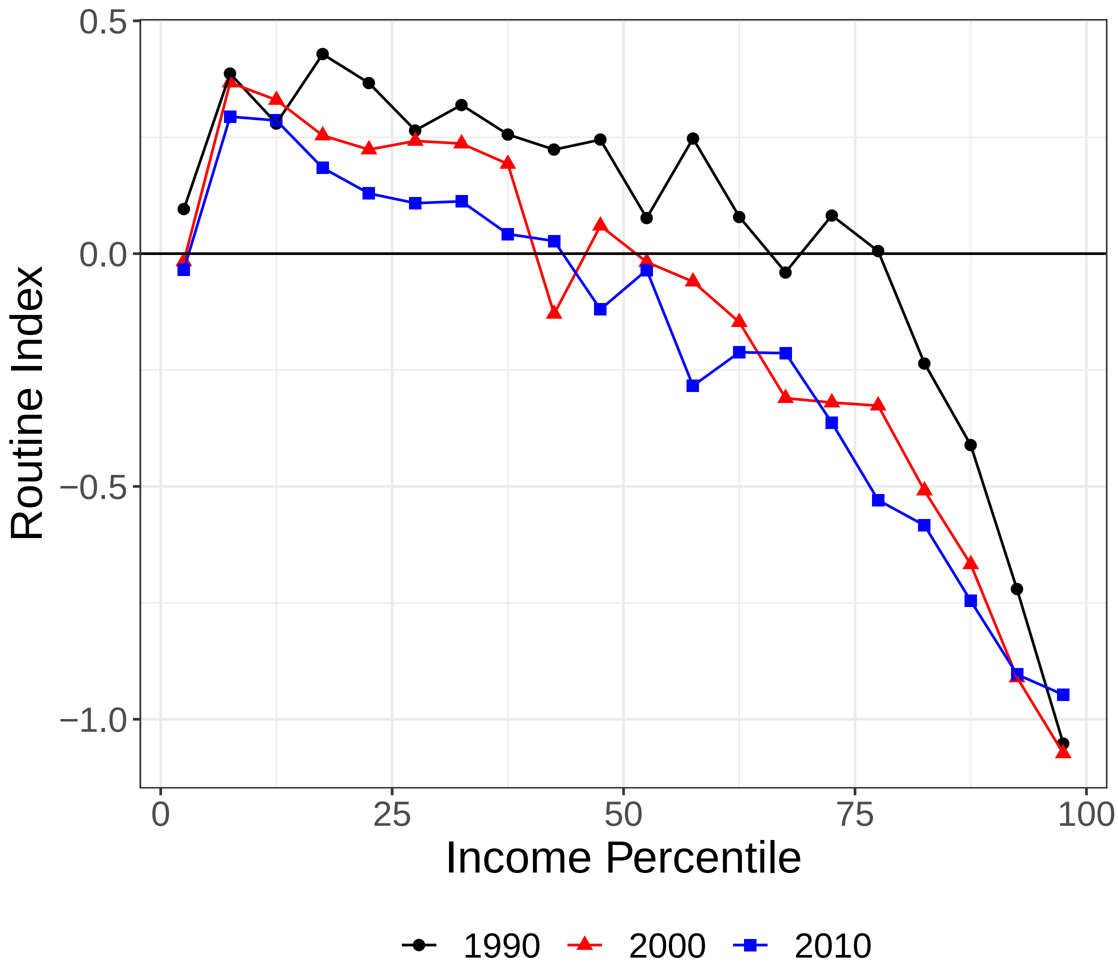


Figure 7: Routine Intensity of Occupation by Wage Percentile

This figure plots the average routine intensity by wage bins for 20 equal-sized bins. Bins are based on workers' hourly wages. The routine intensity is calculated at the occupational level as in Acemoglu and Autor (2011). It is the average of routine manual and routine cognitive indices, both standardized, such that the population average is 0. More details are provided in Appendix B. I use the occupation classification in Autor and Dorn (2013) for consistency across decades. Sample weights are used.
Source: CPS Outgoing Rotation Groups and O*NET

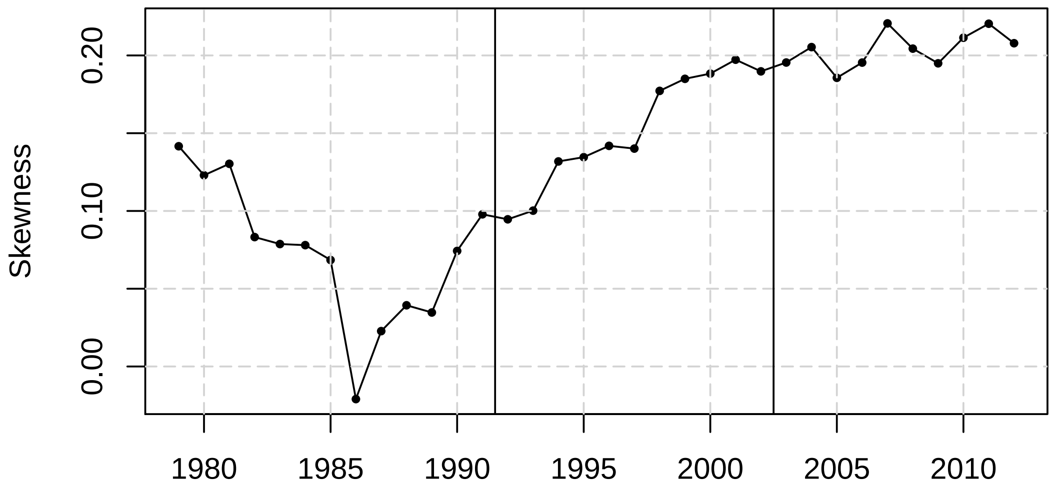


Figure 8: Skewness of Log Hourly Wage

Skewness (Equation 5) of the log wage distribution by year. Sample weights are used. Vertical lines represent changes in occupational coding. Wages at the top and bottom 5% were dropped (see Section 4).
Source: CPS Outgoing Rotation Groups

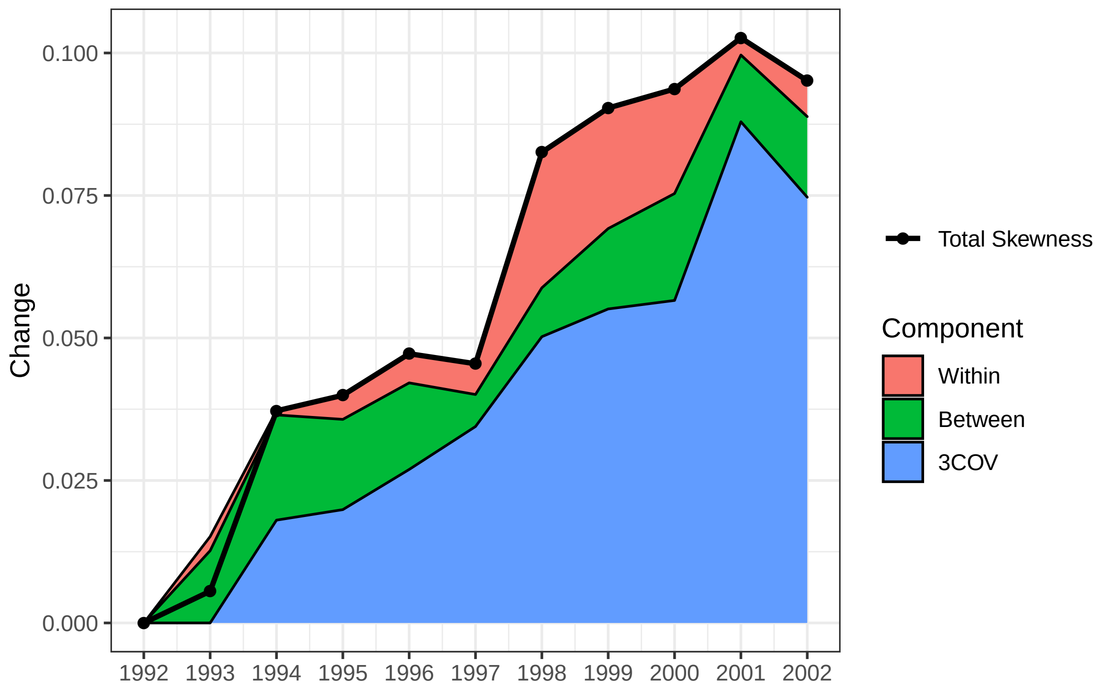


Figure 9: Skewness Decomposition by 3-Digit Occupation

Skewness decomposition based on Equation 6. Changes in each component (within, between, covariance) are plotted relative to the baseline year (1992). The three components sum to the overall skewness (Equation 6). Wages at the top and bottom 5% were dropped (see Section 4).

Source: CPS Outgoing Rotation Groups

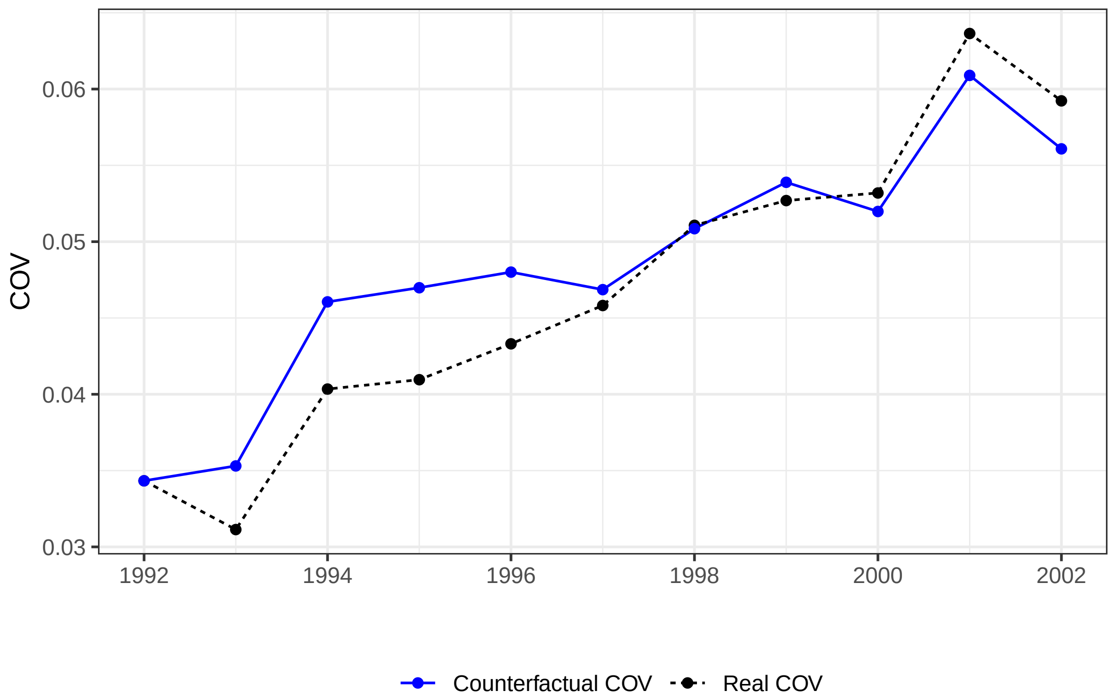


Figure 10: Covariance of Expectation and Variance of Log Wages by Occupation

This figure plots the covariance of mean log wage and variance of log wage by occupation, $COV\left(E\left[\log w|occ\right], V\left(\log w|occ\right)\right)$ (black line). The counterfactual covariance (in blue) is calculated by fixing $E\left[\log w|occ\right]$, and the share of workers in each occupation to their average throughout the period, allowing only the variance within each occupation to change (Equation 8.) Wages at the top and bottom 5% were dropped (see Section 4).
Source: CPS Outgoing Rotation Groups

For Online Publication

A Proofs

Theorem 1 Let $\theta_a, \theta_b \in (\theta_0, \theta_1)$ be the skill levels of two workers in the routine occupation where $\theta_a < \theta_b$. Let w_a, w_b denote their corresponding equilibrium wages. The effect of an improvement in technology τ on the wage ratio $\frac{w_b}{w_a}$ depends on σ such that

$$sign\left(\frac{\partial \frac{w_b}{w_a}}{\partial \tau}\right) = sign\,(1-\sigma)$$

Proof. The wage ratio is

$$\frac{w_b}{w_a} = \frac{\varphi_R(\theta_b;\tau)}{\varphi_R(\theta_a;\tau)} = \left(\frac{\theta_b^{\frac{\sigma-1}{\sigma}} + \tau^{\frac{\sigma-1}{\sigma}}}{\theta_a^{\frac{\sigma-1}{\sigma}} + \tau^{\frac{\sigma-1}{\sigma}}}\right)^{\frac{\sigma}{\sigma-1}}$$

$$\frac{\partial \frac{w_b}{w_a}}{\partial \tau} = \tau^{-\frac{1}{\sigma}}\left(\theta_b^{\frac{\sigma-1}{\sigma}} + \tau^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{\sigma-1}}\left(\theta_a^{\frac{\sigma-1}{\sigma}} + \tau^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1-2\sigma}{\sigma-1}}\left(\theta_a^{\frac{\sigma-1}{\sigma}} - \theta_b^{\frac{\sigma-1}{\sigma}}\right)$$

Since $\tau^{-\frac{1}{\sigma}}\left(\theta_b^{\frac{\sigma-1}{\sigma}} + \tau^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{\sigma-1}}\left(\theta_a^{\frac{\sigma-1}{\sigma}} + \tau^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1-2\sigma}{\sigma-1}} > 0$

$$sign\left(\frac{\partial \frac{w_b}{w_a}}{\partial \tau}\right) = sign\left(\theta_a^{\frac{\sigma-1}{\sigma}} - \theta_b^{\frac{\sigma-1}{\sigma}}\right) = sign(-\frac{\sigma-1}{\sigma}) = sign\,(1-\sigma)$$

□

Theorem 2 RBTC (i.e., an increase in τ) generates:

1. An increase production of the routine good ($\frac{dR}{d\tau} > 0$).
2. A decrease in the absolute and relative price of the routine good ($\frac{dp_R}{d\tau} < 0$) and the relative price compared to abstract/manual good ($\frac{dp_R/p_j}{d\tau} < 0$ for $j \in \{M, A\}$).
3. A decrease in the share of the total income that is spent on routine goods ($\frac{d\frac{p_R R}{Y}}{d\tau} < 0$).

Proof. The proof follows the order of the claims in the theorem:

1. Under the same allocation of workers, when τ increases R increases while M, A are not changed. Hence, Y must be larger in GE (otherwise, Y is not maximized). If

R does not increase then M, A must increase in order for Y to increase. Assume without lost of generality that A increases. From the FOC of the CES function we have

$$\frac{p_R}{p_A} = \left(\frac{R}{A}\right)^{\rho-1}$$

Hence, if A increases and R decreases, p_R/p_A increases. Therefore, for the previous equilibrium level of θ_1

$$p_R\varphi_R(\theta_1) > p_A\varphi_A(\theta_1)$$

which implies that θ_1 increases until equality is reached. Therefore, A must decrease, in contradiction to the assumption.

2. From the FOC we have

$$p_R = \left(\frac{R}{Y}\right)^{\rho-1}$$

Hence, it is sufficient to show R increases more than both M, A such that $\frac{R}{Y}$ increases (and p_R declines). Assume without loss of generality that R/A decreases. By a similar argument as before, if R/A decreases, θ_1 increases and hence A decreases (contradiction). Hence, R/A must increase and hence so must $\frac{R}{Y}$.

Relative prices must also decrease since $\frac{p_R}{p_A} = \left(\frac{R}{A}\right)^{\rho-1}$ and $\rho - 1 < 0$.

3. The share is

$$\frac{p_R R}{Y} = \frac{R^\rho}{Y^\rho} = \frac{R^\rho}{M^\rho + R^\rho + A^\rho} = \frac{1}{\left(\frac{M}{R}\right)^\rho + \left(\frac{A}{R}\right)^\rho + 1}$$

since R increases relative to both M, A, $\left(\frac{M}{R}\right)^\rho$ and $\left(\frac{A}{R}\right)^\rho$ increase and therefore $\frac{p_R R}{Y}$ decreases.

$\square$

Theorem 3 Assume a skill-replacing technology ($\sigma > 1$). RBTC (i.e., an increase in τ) would generate the following:

1. A decrease in wage gaps between routine workers who do not switch occupations.
2. The highest skill routine workers would leave the routine occupation ($\frac{\partial\theta_1}{\partial\tau} < 0$).
3. Wages for the highest skill routine worker (θ_1) would decrease relative to all other workers.

Proof.

1. By theorem 1.

2. By the following lemma (where $\varphi_\tau^R(\theta_0,\tau) = \frac{\partial \varphi^R(\theta_0,\tau)}{\partial \tau}$)

Lemma 6. $\frac{\varphi_\tau^R(\theta_0,\tau)}{\varphi^R(\theta_0,\tau)} > \frac{R_\tau}{R} = \frac{\int_{\theta_0}^{\theta_1} \varphi_\tau^R(\theta_i,\tau)}{\int_{\theta_0}^{\theta_1} \varphi^R(\theta_i,\tau)} > \frac{\varphi_\tau^R(\theta_1,\tau)}{\varphi^R(\theta_1,\tau)}$

Proof. We use

$$\frac{\partial^2 \log \varphi}{\partial \tau \partial \theta} = \frac{1-\sigma}{\sigma} \frac{1}{\left(\theta_i^{\frac{\sigma-1}{\sigma}} + \tau^{\frac{\sigma-1}{\sigma}}\right)^2 (\tau\theta)^{\frac{1}{\sigma}}} < 0$$

Hence, for any positive $b > a$, we have

$$\frac{\varphi_\tau(a,\tau)}{\varphi(a,\tau)} > \frac{\varphi_\tau(b,\tau)}{\varphi(b,\tau)}$$

$$\frac{\varphi_\tau(a,\tau)}{\varphi_\tau(b,\tau)} > \frac{\varphi(a,\tau)}{\varphi(b,\tau)}$$

Defining $b = \theta_1$ and taking the integral over a between $[\theta_0, \theta_1]$ yields $\frac{R_\tau}{R} = \frac{\int_{\theta_0}^{\theta_1} \varphi_\tau^R(\theta_i,\tau)}{\int_{\theta_0}^{\theta_1} \varphi^R(\theta_i,\tau)} > \frac{\varphi_\tau^R(\theta_1,\tau)}{\varphi^R(\theta_1,\tau)}$. Similarly defining $a = \theta_0$, taking the inverse of the above inequality, and integrating over $b \in [\theta_0, \theta_1]$ yields $\frac{\varphi_\tau^R(\theta_0,\tau)}{\varphi^R(\theta_0,\tau)} > \frac{R_\tau}{R}$.

$\square$

Using this lemma we can show that θ_1 decreases. To do so, we use the equations

$$p_R \varphi_R(\theta_1) = p_A \varphi_A(\theta_1)$$

$$p_R \varphi_R(\theta_0) = p_M \varphi_M(\theta_0)$$

and the CES structure

$$\log \varphi_A(\theta_1) - \log \varphi_R(\theta_1,\tau) = \log p_R - \log p_A = (\rho-1)\log R - (\rho-1)\log A$$

and similarly for M. Define two functions of θ_0, θ_1, τ such that

$$f_M = \log \varphi_M(\theta_0) - \log \varphi_R(\theta_0,\tau) + (\rho-1)\, m(\theta_0) - (\rho-1)\, r(\theta_0,\theta_1,\tau)$$

$$f_A = \log \varphi_A(\theta_1) - \log \varphi_R(\theta_1,\tau) + (\rho-1)\, a(\theta_1) - (\rho-1)\, r(\theta_0,\theta_1,\tau)$$

where $m(\theta_0) = \log M(\theta_0)$, $r(\theta_0,\theta_1) = \log R(\theta_0,\theta_1)$, and $a(\theta_1) = \log A(\theta_1)$. In equilibrium $f_M, f_A = 0$ as the FOCs hold. Using the implicit theorem function we can derive $\frac{d\theta_1}{d\tau}$. Taking the derivative by τ we have

$$\frac{\partial f_M}{\partial \tau} = -\frac{\varphi_\tau^R(\theta_0,\tau)}{\varphi^R(\theta_0,\tau)} - (\rho-1)\frac{R_\tau}{R}$$
$$\frac{\partial f_A}{\partial \tau} = -\frac{\varphi_\tau^R(\theta_1,\tau)}{\varphi^R(\theta_1,\tau)} - (\rho-1)\frac{R_\tau}{R} > 0$$

where the last inequality is from the previous lemma and $\rho < 0$. Taking the derivative by θ_0, θ_1 and using Condition 1

$$\frac{\partial f_M}{\partial \theta_0} = \frac{\varphi_M'}{\varphi_M} - \frac{\varphi_R'}{\varphi_R} + (\rho-1)\, m'(\theta_0) - (\rho-1)\frac{R_{\theta_0}}{R} < 0$$

$$\frac{\partial f_A}{\partial \theta_1} = \frac{\varphi_A'}{\varphi_A} - \frac{\varphi_R'}{\varphi_R} + (\rho-1)\, a'(\theta_1) - (\rho-1)\frac{R_{\theta_1}}{R} > 0$$
$$\frac{\partial f_A}{\partial \theta_0} = -(\rho-1)\frac{R_{\theta_0}}{R} < 0$$
$$\frac{\partial f_M}{\partial \theta_1} = -(\rho-1)\frac{R_{\theta_1}}{R} > 0$$

From the implicit function theorem we have

$$\nabla\theta(\tau) = -\begin{pmatrix} \frac{\partial f_M}{\partial \theta_0} & \frac{\partial f_M}{\partial \theta_1} \\ \frac{\partial f_A}{\partial \theta_0} & \frac{\partial f_A}{\partial \theta_1} \end{pmatrix}^{-1} \left(\frac{\partial f_M}{\partial \tau}, \frac{\partial f_A}{\partial \tau}\right)$$

or

$$\nabla\theta(\tau) = -\frac{1}{det}\begin{pmatrix} \frac{\partial f_A}{\partial \theta_1} & -\frac{\partial f_M}{\partial \theta_1} \\ -\frac{\partial f_A}{\partial \theta_0} & \frac{\partial f_M}{\partial \theta_0} \end{pmatrix} \left(\frac{\partial f_M}{\partial \tau}, \frac{\partial f_A}{\partial \tau}\right)$$

The determinant is negative. Hence

$$\frac{\partial \theta_1}{\partial \tau} = c * \left(-\frac{\partial f_A}{\partial \theta_0}\frac{\partial f_M}{\partial \tau} + \frac{\partial f_M}{\partial \theta_0}\frac{\partial f_A}{\partial \tau}\right)$$

$$\frac{\partial \theta_0}{\partial \tau} = c * \left(\frac{\partial f_A}{\partial \theta_1}\frac{\partial f_M}{\partial \tau} - \frac{\partial f_M}{\partial \theta_1}\frac{\partial f_A}{\partial \tau}\right)$$

when $c > 0$.

Plugging in the values that were previously calculated and using Lemma 6 yields

$$\frac{\partial\theta_1}{\partial\tau}=\left(\frac{\varphi_M'}{\varphi_M}-\frac{\varphi_R'}{\varphi_R}+(\rho-1)\,m'(\theta_0)\right)\frac{\partial f_A}{\partial\tau}-(\rho-1)\,\frac{R_{\theta_0}}{R}\left(\frac{\partial f_A}{\partial\tau}-\frac{\partial f_M}{\partial\tau}\right)=$$

$$=\left(\frac{\varphi_M'}{\varphi_M}-\frac{\varphi_R'}{\varphi_R}+(\rho-1)\,m'(\theta_0)\right)\frac{\partial f_A}{\partial\tau}-(\rho-1)\,\frac{R_{\theta_0}}{R}\left(-\frac{\varphi_\tau^R(\theta_1,\tau)}{\varphi^R(\theta_1,\tau)}+\frac{\varphi_\tau^R(\theta_0,\tau)}{\varphi^R(\theta_0,\tau)}\right)<0$$

For θ_0 we have

$$\frac{\partial\theta_0}{\partial\tau}=\left(\frac{\varphi_A'}{\varphi_A}-\frac{\varphi_R'}{\varphi_R}+(\rho-1)\,a'(\theta_1)\right)\frac{\partial f_M}{\partial\tau}-(\rho-1)\,\frac{R_{\theta_1}}{R}\left(-\frac{\varphi_\tau^R(\theta_0,\tau)}{\varphi^R(\theta_0,\tau)}+\frac{\varphi_\tau^R(\theta_1,\tau)}{\varphi^R(\theta_1,\tau)}\right)$$

however, the sign of this expression could go both ways.

3. Among routine workers, $\sigma>1$ implies that the largest decline in wages is for the highest-skilled workers. Their wages also fall compared to those of abstract workers since abstract workers see a change only in p_A and at θ_1:

$$\frac{\partial\log p_R}{\partial\tau}+\frac{\partial\log\varphi_R}{\partial\tau}(\theta_1)-\frac{\partial\log p_A}{\partial\tau}+0<0 \tag{9}$$

as otherwise θ_1 would not go down.

Finally, for manual workers, if $\frac{d\theta_0}{d\tau}\geq 0$ (employment in the manual occupation is weakly increasing) then at θ_0 we have

$$\frac{\partial\log p_R}{\partial\tau}+\frac{\partial\log\varphi_R}{\partial\tau}(\theta_0)-\frac{\partial\log p_M}{\partial\tau}+0<0$$

and since $\frac{\partial\log\varphi_R}{\partial\tau}(\theta_1)<\frac{\partial\log\varphi_R}{\partial\tau}(\theta_0)$ we get

$$\frac{\partial\log p_R}{\partial\tau}+\frac{\partial\log\varphi_R}{\partial\tau}(\theta_1)<\frac{\partial\log p_M}{\partial\tau} \tag{10}$$

and so wages at θ_1 fall relative to all manual jobs.

If $\frac{d\theta_0}{d\tau}<0$ then $\frac{dM}{d\tau}<0$, and since $\frac{dA}{d\tau}>0$ it must be that $\frac{d\left(\frac{M}{A}\right)^{\rho-1}}{d\tau}>0$. Using $\frac{p_M}{p_A}=\left(\frac{M}{A}\right)^{\rho-1}$ we get that wages in the manual occupation increase faster than in abstract occupation and using 9 we get that 10 also holds.

$\square$

Theorem 4 Assume a skill replacing technology ($\sigma > 1$), and

$$\frac{\left|\frac{dF(\theta_1)}{d\tau}\right|}{\frac{dF(\theta_0)}{d\tau}} > \frac{E\left[\theta|\theta_0 < \theta < \theta_1\right] - \theta_0}{\theta_1 - E\left[\theta|\theta_0 < \theta < \theta_1\right]} \tag{11}$$

In the routine occupation, RBTC would generate a decrease in:

1. Employment ($\theta_1 - \theta_0$)
2. Within-occupation inequality ($V\left[w_i|\theta_i \in [\theta_0, \theta_1]\right]$)
3. Mean skill level ($E\left[\theta_i|\theta_i \in [\theta_0, \theta_1]\right]$) .

Inequality within the abstract ($V\left[w_i|\theta_i > \theta_1\right]$) and manual occupation ($V\left[w_i|\theta_i < \theta_0\right]$) will rise and so overall wage trend would be U-shaped (wage polarization), such that for every $\theta_a < \theta_b \leq \theta_1$

$$\frac{d\frac{w_b}{w_a}}{d\tau} \leq 0$$

and for every $\theta_1 \leq \theta_a < \theta_b$

$$\frac{d\frac{w_b}{w_a}}{d\tau} \geq 0$$

Note that this is a more general version of the theorem in the main text. Specifically, when $\theta \sim U$, condition 11 is equivalent to $0 < \frac{d\theta_0}{d\tau} < \frac{d\theta_1}{d\tau}$.

Proof. Employment in the routine occupation decreases since θ_1 decreases and θ_0 increases.

Within the routine occupation, inequality would decrease since the skill distribution is now more equal ($V\left[\theta_i|\theta_i \in [\theta_0, \theta_1]\right]$ decreases), and conditional on skill wage gaps are smaller (Theorem 1). Same argument will also apply for other inequality measures.

Mean skill is $E\left[\theta_i|\theta_i \in [\theta_0, \theta_1]\right] = \frac{\int_{\theta_0}^{\theta_1} \theta f(\theta) d\theta}{\int_{\theta_0}^{\theta_1} f(\theta) d\theta}$. Taking the derivative yields

$$\frac{\left[\frac{d\theta_1}{d\tau}\theta_1 f\left(\theta_1\right) - \frac{d\theta_0}{d\tau}\theta_0 f\left(\theta_0\right)\right] \int_{\theta_0}^{\theta_1} f\left(\theta\right) d\theta - \int_{\theta_0}^{\theta_1} \theta f\left(\theta\right) d\theta \left(\frac{d\theta_1}{d\tau} f\left(\theta_1\right) - \frac{d\theta_0}{d\tau} f\left(\theta_0\right)\right)}{\left(\int_{\theta_0}^{\theta_1} f\left(\theta\right) d\theta\right)^2}$$

which is always positive when condition 11 holds.

Within the abstract and the manual occupation, inequality increases as there is a larger variation in skill.

The derivative $\frac{d \log w}{d\tau}(\theta)$ is U-shaped in θ with a minimum at θ_1. This is because for manual workers, the wage effect is $\frac{d \log p_M}{d\tau}$. For routine workers, it is $\frac{d \log p_R}{d\tau} + \frac{d \log \varphi_R}{d\tau}$. And for abstract workers, it is $\frac{d \log p_A}{d\tau}$. The wage increase in the abstract occupation is larger than that in the routine occupation close to θ_1 because from the decrease in θ_1 if follows that

$$\frac{d \log p_R}{d\tau} + \frac{d \log \varphi_R(\theta_1)}{d\tau} < \frac{d \log p_A}{d\tau}$$

Among routine workers, wages increase relatively for lower-skilled since $\sigma > 1$. And manual workers see a larger increase from the wage change in θ_0 since

$$\frac{d \log p_R}{d\tau} + \frac{d \log \varphi_R(\theta_0)}{d\tau} < \frac{d \log p_M}{d\tau}$$

as otherwise θ_0 would not increase.

Within the manual and abstract occupations, wage impact is the same. □

Theorem 5 Assume a skill-replacing technology ($\sigma > 1$). There exists a $\widetilde{\tau}$ such that for any $\tau \geq \widetilde{\tau}$ and for any θ

$$\frac{\partial \log \varphi_R(\theta; \tau)}{\partial \theta} < \frac{\partial \log \varphi_M(\theta)}{\partial \theta}$$

When $\tau \geq \widetilde{\tau}$ routine workers earn the lowest wages. Any additional SR-RBTC ($\tau \uparrow$) decreases employment in the routine occupation ($\frac{d\theta_0}{d\tau}$<0), as well as wage gaps within routine workers who do not switch occupations.

Proof. Let θ_{min} be the lowest value of θ. Therefore, for every θ_i,

$$\frac{\partial \log \varphi_R(\theta_i; \tau)}{\partial \theta} \leq \frac{\partial \log \varphi_R(\theta_{min}; \tau)}{\partial \theta}$$

By assumption $\exists c$ s.t. $\forall \theta_i$ $\frac{\partial \log \varphi_j(\theta)}{\partial \theta} > c > 0$. And by the definition of φ_R

$$\lim_{\tau \to \infty} \frac{\partial \log \varphi_R(\theta; \tau)}{\partial \theta} = \lim_{\tau \to \infty} \frac{1}{\left(\theta_i^{\frac{\sigma-1}{\sigma}} \tau^{\frac{1}{\sigma}} + \tau\right)} = 0$$

Therefore, there exists $\widetilde{\tau}$ such that for every $\tau \geq \widetilde{\tau}$ and every θ_i,

$$\frac{\partial \log \varphi_R(\theta_{min}; \tau)}{\partial \theta} < c$$

and using $\frac{\partial \log \varphi_j(\theta)}{\partial \theta} > c$ and $\frac{\partial \log \varphi_R(\theta_i;\tau)}{\partial \theta} \leq \frac{\partial \log \varphi_R(\theta_{min};\tau)}{\partial \theta}$ we get

$$\frac{\partial \log \varphi_R(\theta_i;\tau)}{\partial \theta} < \frac{\partial \log \varphi_M(\theta_i)}{\partial \theta}$$

For the second part of the theorem (for $\tau \geq \widetilde{\tau}\ \frac{d\theta_0}{d\tau} < 0$) the equilibrium condition for θ_0 is

$$\log \varphi_M(\theta_0) - \log \varphi_R(\theta_0, \tau) = \log p_R - \log p_M = (\rho - 1)\log R - (\rho - 1)\log M$$

Defining

$$f(\theta_0, \tau) = \log \varphi_M(\theta_0) - \log \varphi_R(\theta_0, \tau) - (\rho - 1)\log R + (\rho - 1)\log M$$

and using the implicit function theorem we get

$$\frac{\partial f}{\partial \tau} = -\frac{\varphi_\tau^R(\theta_0, \tau)}{\varphi^R(\theta_0, \tau)} - (\rho - 1)\frac{R_\tau}{R} > 0$$

where the last inequality follows from Lemma 6 (for which now the upper bound of the routine occupation is θ_0).

Taking the derivative by θ_0, we have

$$\frac{\partial f}{\partial \theta_0} = \frac{\varphi_M'}{\varphi_M} - \frac{\varphi_R'}{\varphi_R} + (\rho - 1)\, m'(\theta_0) - (\rho - 1)\frac{R_{\theta_0}}{R} > 0$$

where $m'(\theta_0) < 0$ as a result of the following lemma:

Lemma 7. *When $\tau \geq \widetilde{\tau}$ $sign\left(\frac{dM}{d\tau}\right) = sign\left(\frac{dA}{d\tau}\right)$*

Proof. When $\tau \geq \widetilde{\tau}$, θ_1 separates between manual and abstract workers. Using the equilibrium condition we have

$$\log \varphi_M(\theta_1) - \log \varphi_A(\theta_1) = \log p_A - \log p_M = (\rho - 1)\log A - (\rho - 1)\log M$$

If $\frac{dM}{d\tau} > 0 \Rightarrow \frac{dA}{d\tau} > 0$; else $\frac{d\frac{p_A}{p_M}}{d\tau} > 0 \Rightarrow \frac{d\log \varphi_M(\theta_1) - \log \varphi_A(\theta_1)}{d\tau} > 0 \Rightarrow \frac{d\theta_1}{d\tau} < 0$, where the last part is because

$$\frac{d\log \varphi_M(\theta_1) - \log \varphi_A(\theta_1)}{d\tau} = \frac{\partial \log \varphi_M(\theta_1) - \log \varphi_A(\theta_1)}{\partial \theta_1}\frac{\partial \theta_1}{\partial \tau}$$

□

By this lemma, if θ_0 increases and M increases, A increases as well, implying that θ_1 decreases, which is a contradiction (M cannot increase if θ_0 increases and θ_1 decreases).

Taken together

$$\frac{d\theta_0}{d\tau} = -\frac{\frac{\partial f}{\partial \tau}}{\frac{\partial f}{\partial \theta_0}} < 0$$

□

B Data Appendix

B.1 CPS-ORG

The CPS-ORG provides the most accurate representative sample of hourly wages (Lemieux, 2006).

I use the same sample definitions as Acemoglu and Autor (2011), who kindly made their cleaned data files available online. See Acemoglu and Autor (2011) data appendix for exact definitions of the sample and sample sizes. Observations with missing wages are dropped. The main results hold when using imputations instead. Sampling weights are used in all CPS-ORG analyses. Education categories are equivalent to those employed by Autor et al. (2003) based on the consistent classification system proposed by Jaeger (1997).

One important limitation of the CPS data is its relatively high level of measurement errors. This problem is particularly severe at both tails of the distribution. Misreporting of working hours could lead to extremely high or extremely low values of hourly wages. Moreover, the CPS applies top coding to prevent the identification of individuals with extremely high income. Therefore, I drop the top and bottom 5% of the positive wages. Similar methods have been applied in previous work that used this data (Katz and Murphy, 1992; Autor et al., 2006, 2008; Acemoglu and Autor, 2011).

Since my analysis focuses on hourly wages, I multiply the CPS weights by the number of hours worked to obtain the real price of an hour of labor, as explained in Lemieux (2010). This procedure is also consistent with the literature.

B.2 PSID

I use information from both the individual survey and the family survey. The oversampling of low-income households and immigrants samples are not used as it was added only in the 1990s. Therefore, I do not use sampling weights (similar to Cortes, 2016). However, as the analysis focuses on hourly wages, I weight observations by the number of weekly hours to obtain the real price of an hour of labor, similar to the CPS exercise.

I drop observations with hourly wages at the top or the bottom 5%, as with the CPS. Observations in which wage is imputed are also omitted.

The education variable is defined based on a survey question that inquiries about the highest grade/years of schooling that the respondent has completed. A value of 16 indicates a college graduate. This variable is capped at 17.

Occupations are coded using the Autor and Dorn (2013) occupational crosswalk for census coding.

B.3 O*NET

Routine Index: Acemoglu and Autor (2011) construct two indices: routine-manual and routine-cognitive. These indices are based on occupational averages of survey responses in the O*NET database. I take the average of both (standardized) indices. The routine manual index includes questions on:

- Pace determined by speed of equipment.
- Controlling machines and processes.
- Spend time making repetitive motions.

The routine cognitive index includes questions on:

- Importance of repeating the same tasks.
- Importance of being exact or accurate.
- Structured v. unstructured work (reverse).

I thank the authors for sharing their data with me.

C Multidimensional Skill

I estimate Equation 3 allowing θ_i to vary by j for the three broad occupational categories (abstract, routine, manual). This allows for a different skill to be used in each occupational category, as in a Roy model.

I find that the correlations between θ_{ij} for a given value of i are between 0.69–0.83 as shown in Table A1. The correlations are estimated for workers that chose to switch occupations (51% of the sample). Since θ_{ij} is measured with a high level of noise, the results are downward biased. This high level of correlation suggests that the one-dimensional skill restriction is a reasonable approximation for this exercise. For the rest of the paper, I assume that θ_i is fixed.

D Decomposing by More Than One Category

Similar to variance decomposition, skewness decomposition can also be easily extended to accommodate linear models. Assume the following simple linear model when Y is standardized:

$$Y = \sum_i X_i$$

Using simple algebra we get

$$\mu_3(Y) = \sum_i \mu_3(X_i) + \sum_i \sum_{j \neq i} COV\left(X_i^2, X_j\right) + \sum_i \sum_{j \neq i} \sum_{k \neq i,j} E\left[X_i X_j X_k\right] \tag{12}$$

Therefore, we can decompose the skewness of Y into a linear combination of: (i) the skewness of the linear components, (ii) the covariance of the linear components second and first moments, and (iii) the triple multiplication of all three distinguished components. Though this decomposition includes a large number of different terms, many of them are constant at zero.

For example, writing Y as the sum of its conditional expectation in X and a residual ε

$$Y = E\left[Y|X\right] + \varepsilon$$

and using Equation 12 yields Equation 6, by the law of iterated expectations over X.

Linear skewness decomposition can be applied to decompose any linear model. This is useful for comparing occupations directly to other categories, such as industry and education. I show this for the equation

$$\ln w_i = occ_i + ind_i + \varepsilon_i \tag{13}$$

where occ_i and ind_i are occupation and industry dummies. I then decompose the increase in skewness using Equation 12. Figure A18 presents the results. Most of the increase in skewness is generated by the increase in the correlation between the occupation premium occ_i and the residual variance ε_i^2. The equivalent component for industries (in green) is negligible. All other components, such as the skewness between occupations or industries, the correlation between the occupation premium and the industry premium variance and others are aggregated and plotted in red. Altogether, they comprise only a small share of the increase.

To do the same exercise for occupations with observable skills I estimate a Mincer equation with occupational dummies

$$\ln w_i = occ_i + \beta X_i + \varepsilon_i \tag{14}$$

where X_i includes years of schooling, experience and experience squared.

In this estimation, the occupation premiums are conditional on the workers' observed skills. Therefore, I decompose βX_i into mean occupation skill level and within-occupation skill difference

$$E\left[\beta X_i | occ_i\right] + \left(\beta X_i - E\left[\beta X_i | occ_i\right]\right)$$

such that the first component captures the average skill level in an occupation and the second component captures the skill part that is orthogonal to the occupation.

I implement a linear skewness decomposition into four components using the following equation

$$\ln w_i = occ_i + E\left[\beta X_i | occ_i\right] + \left(\beta X_i - E\left[\beta X_i | occ_i\right]\right) + \varepsilon_i$$

I find that the two main components are the correlation of ε^2 with both occ_i and $E\left[\beta X_i | occ_i\right]$. This means that the correlation of the inequality of the unobservables (the variance of ε) with occupational wage levels is due to both occupation premium (occ_i) and the mean skill level at the occupation ($E\left[\beta X_i | occ_i\right]$). Hence, inequality is large in occupations that pay more and have higher-skilled workers, consistent with the SR-RBTC model. Categories that are unrelated to occupations are still negligible.

E Appendix Figures

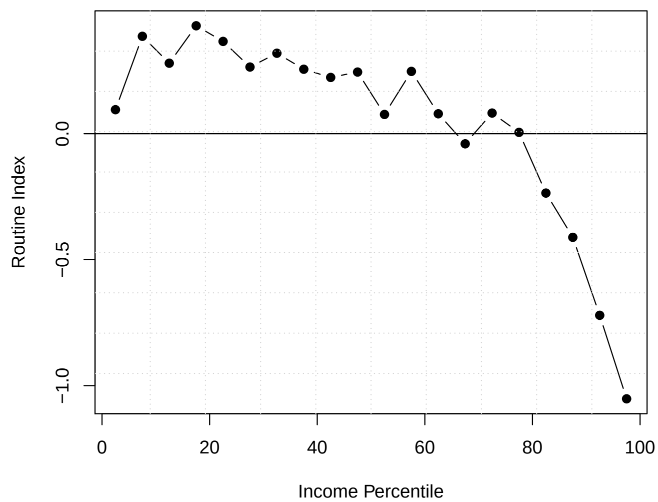


Figure A1: Routine Intensity of Occupation by Wage Percentile – 1990

This figure plots the average routine intensity index by wage bins for 20 equal size bins. Bins are based on workers' hourly wages in 1990. The routine index is calculated at the occupational level as in Acemoglu and Autor (2011). It is the average of routine manual and routine cognitive indices, both standardized, such that the population average is 0. More details in Appendix B. I use the occupation classification in Autor and Dorn (2013) for consistency across decades. Sample weights are used.
Source: CPS Outgoing Rotation Groups and O*NET

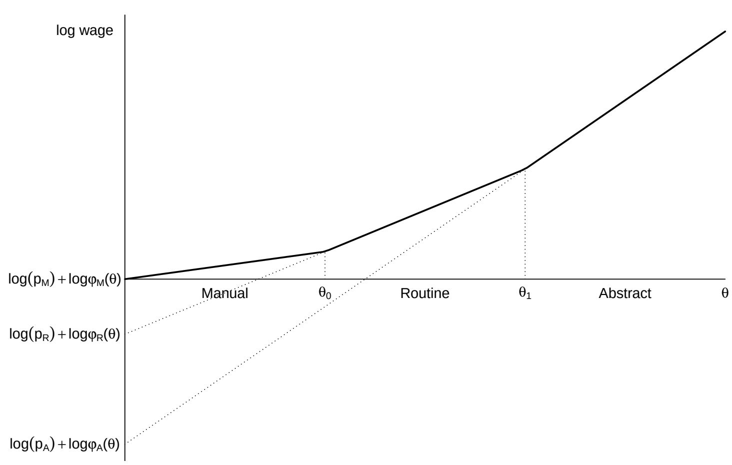


Figure A2: Equilibrium Log Wage by Skill

Sorting into occupations in a Jung and Mercenier (2014) model. The bold line represents equilibrium log wages as a function of θ. Dashed lines are off-equilibrium wages in other (suboptimal) occupations. θ_0, θ_1 are the threshold skill levels in which workers are indifferent between two occupations.

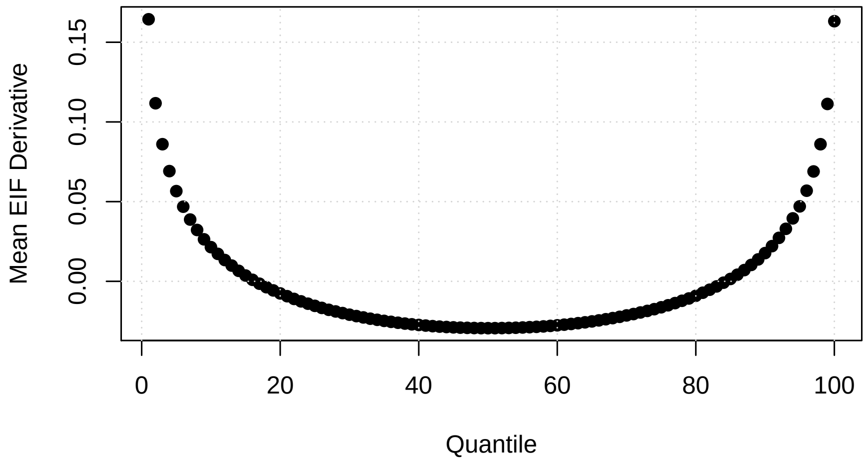


Figure A3: Derivative of EIF on Skewness for Standard Normal Distribution
The empirical influence function is a function from the value of a given observation x_i to some statistic $T_n(x)$ (in this case, the empirical skewness), taking the other observations x_{-i} as given. I calculate this for a sample of $n = 100$. I sample 1,000 samples of 100 observations from a standardized Normal distribution, and calculate numerically the derivative at the kth-order statistic at the sample point. The figure shows the mean over the 1,000 samples of this derivative.

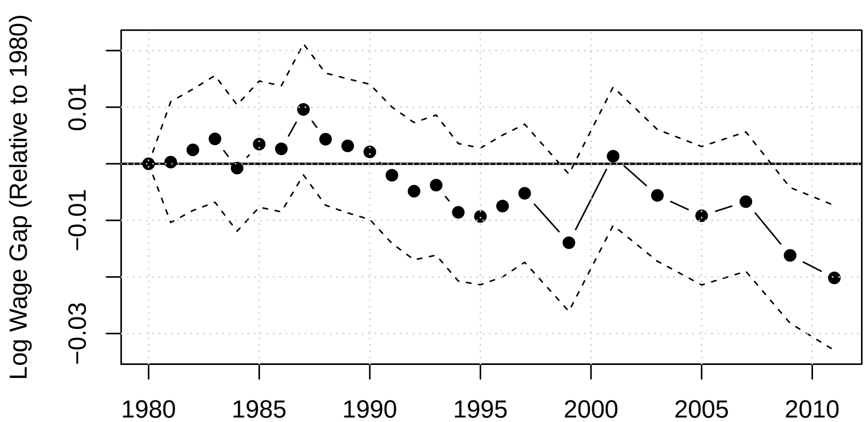
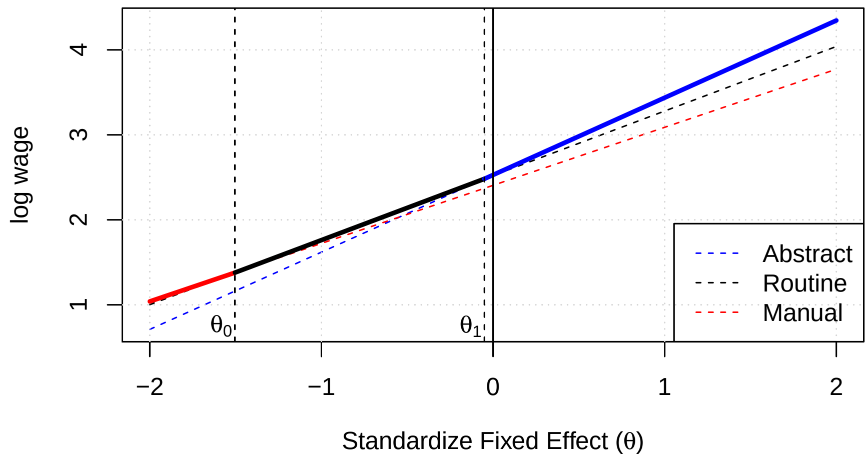
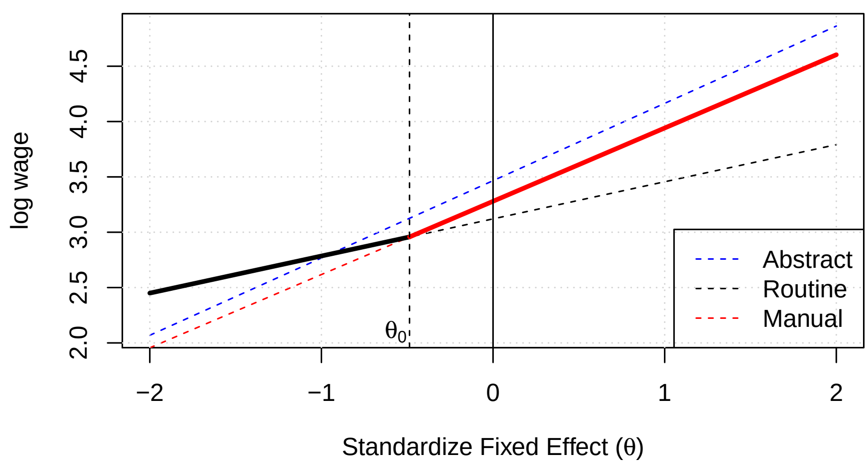


Figure A4: Education Premium in Routine Occupations
This figure plots the change in wage gaps between workers with different years of schooling relative to 1980. Education premium is estimated with coefficient γ_t in Equation 7. Data includes all routine workers in the PSID. See Section 4 for a definition of routine occupations.
Source: PSID



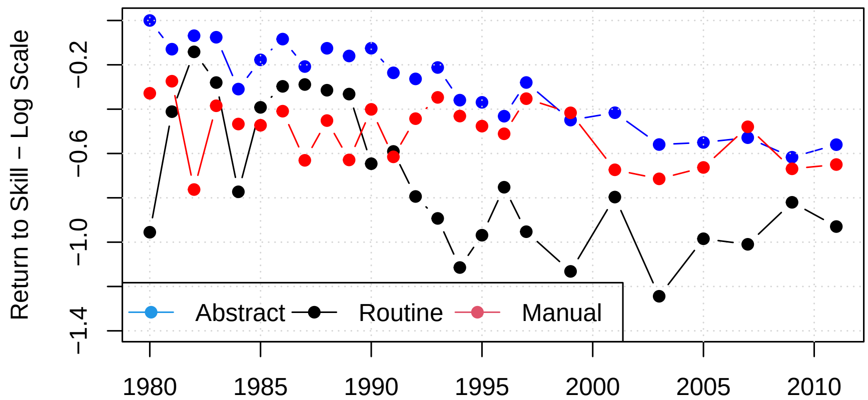
(a) 1987



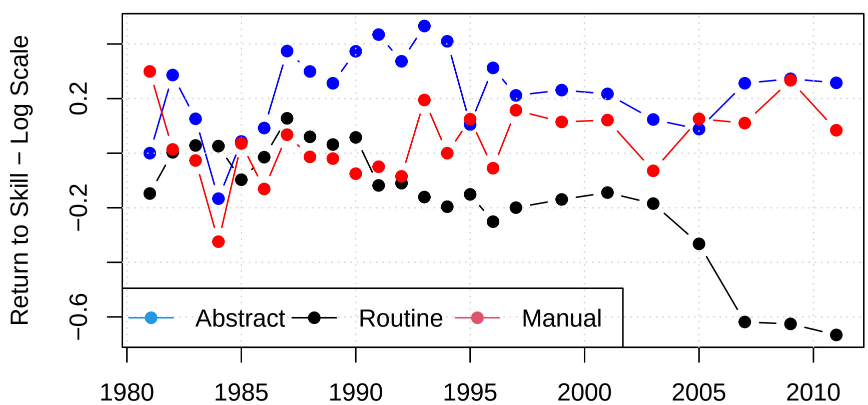
(b) 2011

Figure A5: IFEM Expected Log Wage by Occupational Category

This figure plots the expected log wage as a function of the standardized worker fixed effect θ and occupational category for the years 1987 and 2011 using the estimation of the IFEM (Equation 3). The slope in each occupation is determined by the parameter $\alpha_{j,t}$, where a higher slope implies a larger return to skill. Worker fixed effects are standardized to have a mean of zero in every cohort and a standard deviation of 1 overall. In Panel A the bold line represents the highest expected wage for each skill level. The dashed vertical lines mark the indifference points between two occupational categories that correspond to θ_0 and θ_1 in the model. Panel B focuses on the routine and manual categories. The bold line represents the highest expected wage between these two categories. θ_0 represents the indifference point between them. The abstract category is also drawn for reference (in blue). In contrast to the model prediction, the wage in the abstract category dominates wages in the manual category (see discussion in footnote 26).
Source: PSID



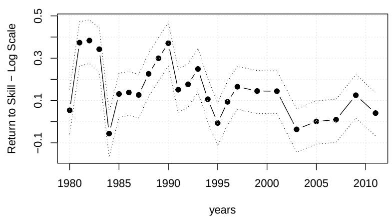
(a) Females



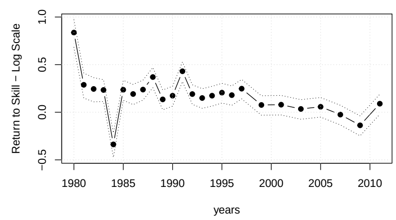
(b) Males

Figure A6: Return to Skill (a_{jt}) by Occupational Category

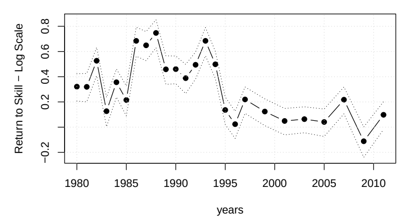
This figure presents the return to skill (α_{jt}) in log units for the three occupational categories separately by gender. Return to skill is calculated using an interactive fixed-effects model (Equation 3). The log return to skill in the abstract occupation in 1980 is fixed to zero, hence all other values are relative to that year and occupational category. Routine workers are defined as workers in administrative, production, or operator occupations, classified by the first occupational coding digit. Abstract workers include managers, technicians, and professionals. Manual includes service, sales and agricultural occupations. Autor and Dorn (2013) occupational crosswalk is used for a consistent definition of occupations over time.
Source: PSID



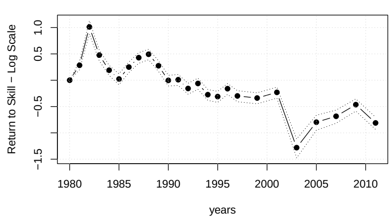
(a) Managers



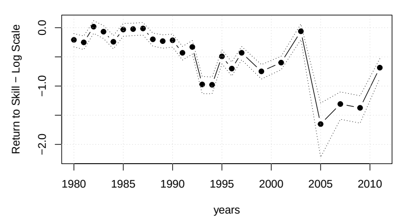
(b) Professional



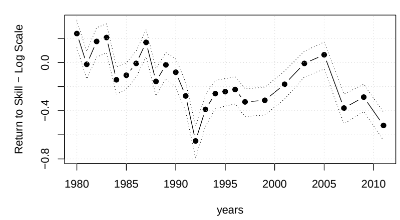
(c) Technicians



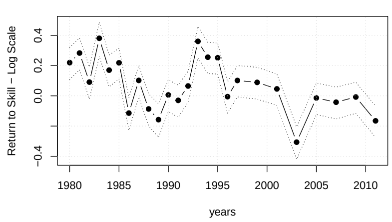
(d) Administrators



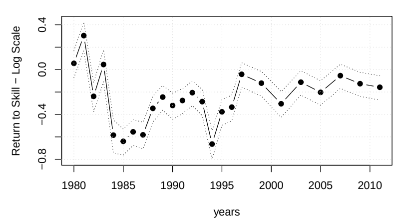
(e) Operators



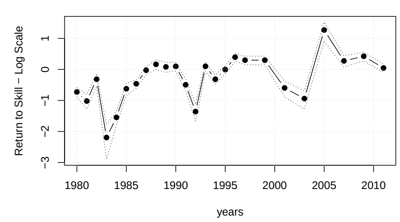
(f) Production



(g) Sales



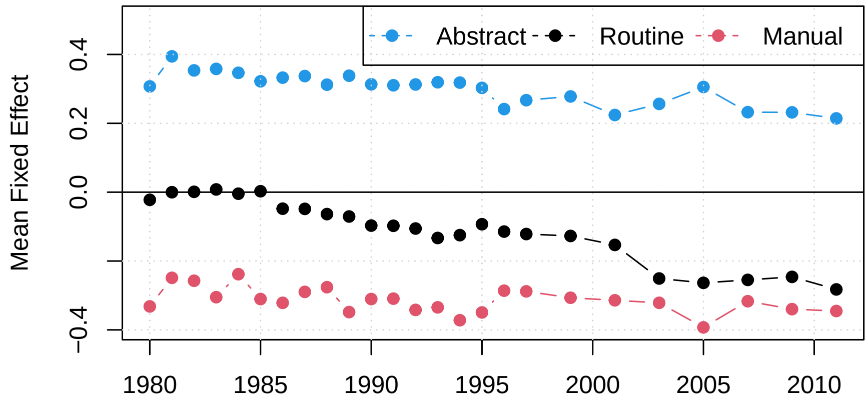
(h) Services



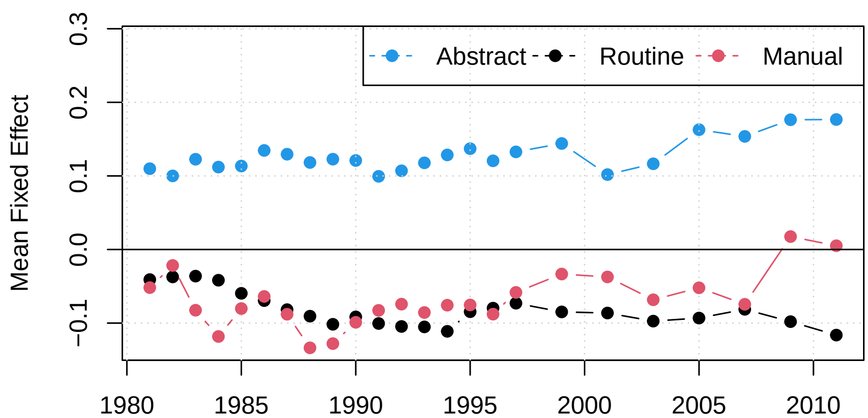
(i) Agriculture

Figure A7: Return to Skill (α_{jt}) by1-Digit Occupation

Returns to skill are calculated in an interactive fixed-effects model (α_{jt}, using Equation 3). α_{jt} varies by 1-digit occupation and year. The Autor and Dorn (2013) occupational crosswalk is used for a consistent definition of occupations over time.
Source: PSID



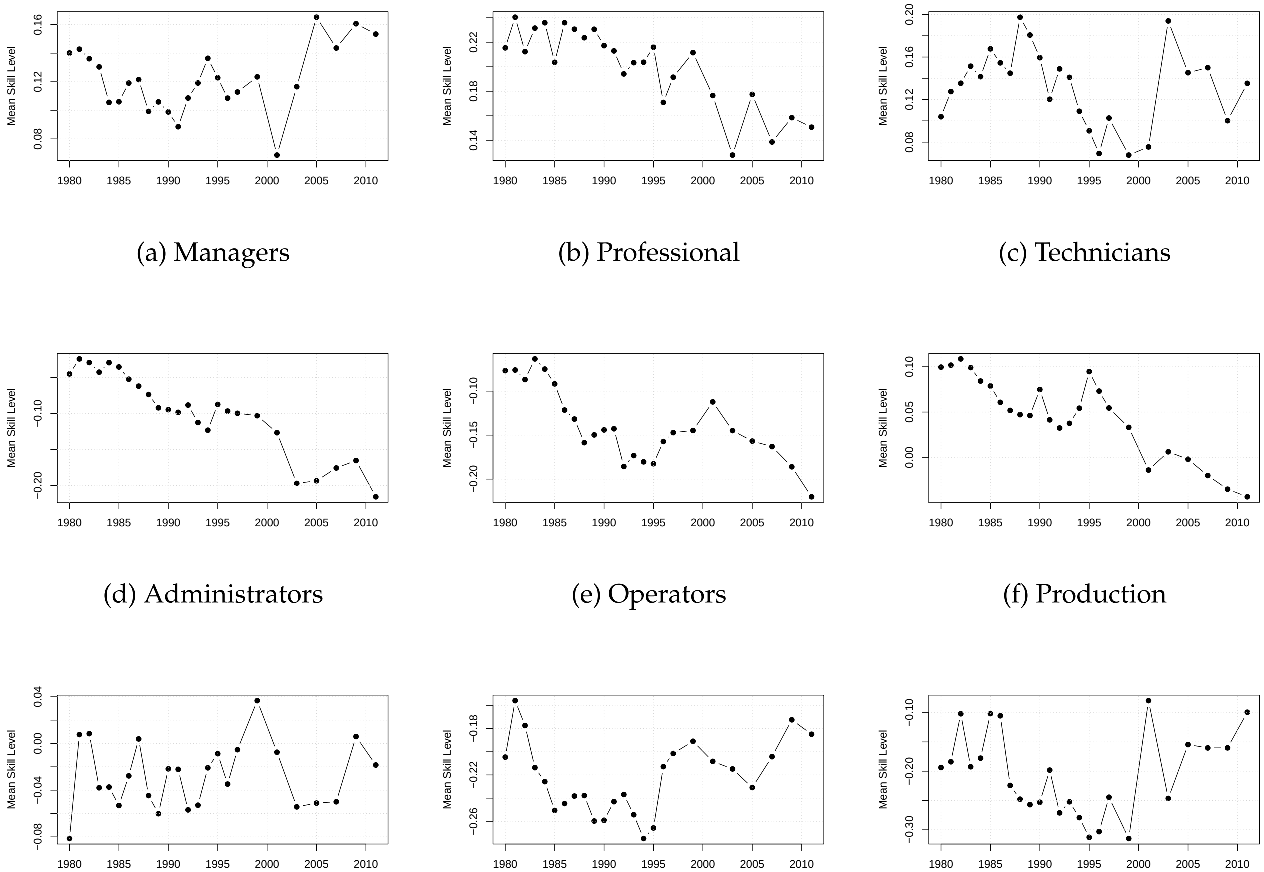
(a) Females



(b) Males

Figure A8: Mean Skill Level ($\widehat{\theta}_i$) by Occupational Category

Mean level of $\widehat{\theta}_i$ by gender, occupational category and year. $\widehat{\theta}_i$ is calculated using Equation 4 estimated separately by gender, and demeaned at the cohort level, where cohorts are defined based on year of entry into the labor market. Routine workers are defined as workers in administrative, production, or operator occupations, classified by the first occupational coding digit. Abstract workers include managers, technicians, and professionals. Manual workers include service, sales, and agricultural occupations. Autor and Dorn (2013) occupational crosswalk is used for a consistent definition of occupations over time.
Source: PSID



(a) Managers (b) Professional (c) Technicians

(d) Administrators (e) Operators (f) Production

(g) Sales (h) Services (i) Agriculture

Figure A9: Mean Skill Level ($\widehat{\theta}_i$) by1-Digit Occupation

Mean level of $\widehat{\theta}_i$ by occupational category and year. $\widehat{\theta}_i$ is calculated using Equation 4, and demeaned at the cohort level, where cohorts are defined based on year of entering the labor market. The Autor and Dorn (2013) occupational crosswalk is used for a consistent definition of occupations over time.
Source: PSID

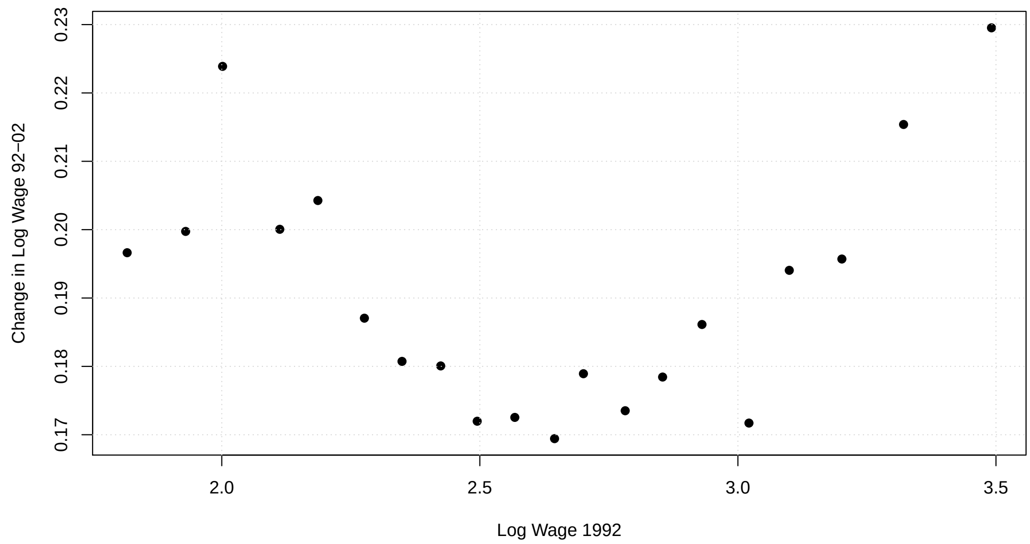


Figure A10: Bin Scatter Plot - Change in Log Wages 1992–2002

Change in log wages in each of 20 equal-sized quantiles. Quantiles are calculated separately for both 1992 and 2002. The x-axis shows the value of the mean log wage in each quantile. The y-axis plots the difference in mean log wages in each of the 20 quantiles between 1992–2002. Sample weights are used.
Source: CPS Outgoing Rotation Groups

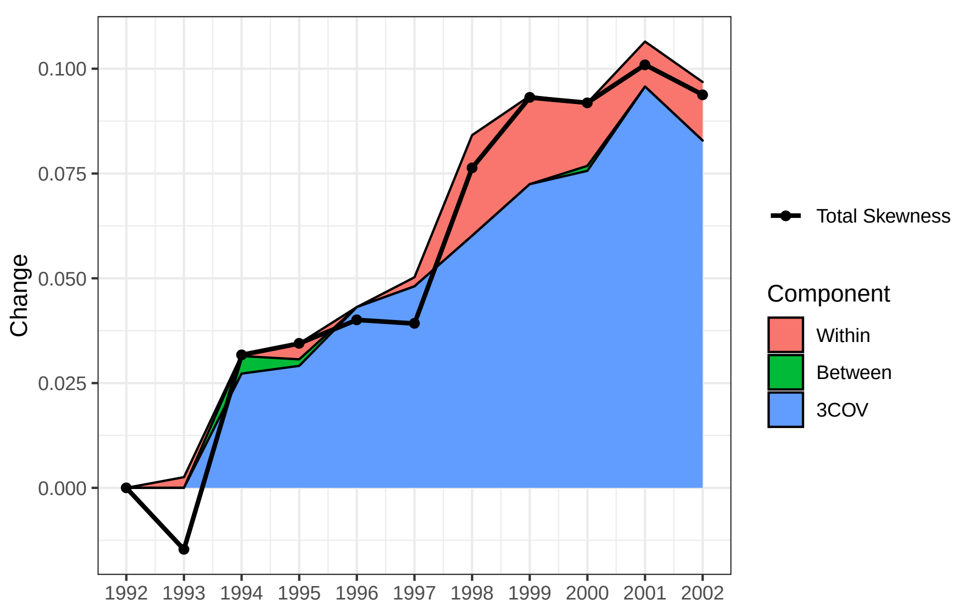


Figure A11: Skewness Decomposition by 3-Digit Occupation with Imputed Wages

Skewness decomposition based on Equation 6. This figure replicates the results in Figure 9 including imputed wages. Changes in each component are plotted relative to the baseline year (1992). Wages at the top and bottom 5% were dropped (see Section 4).
Source: CPS Outgoing Rotation Groups

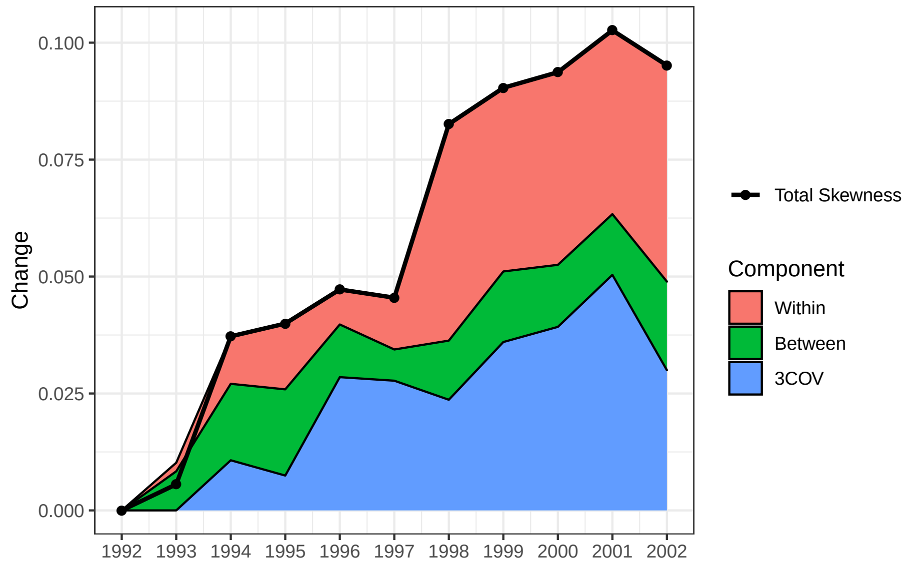


Figure A12: Skewness Decomposition by 3-Digit Industry

Skewness decomposition (Equation 9) by 3-digit industry categories. Changes in each component are plotted relative to the baseline year (1992). Wages at the top and bottom 5% were dropped (see Section 4).
Source: CPS Outgoing Rotation Groups

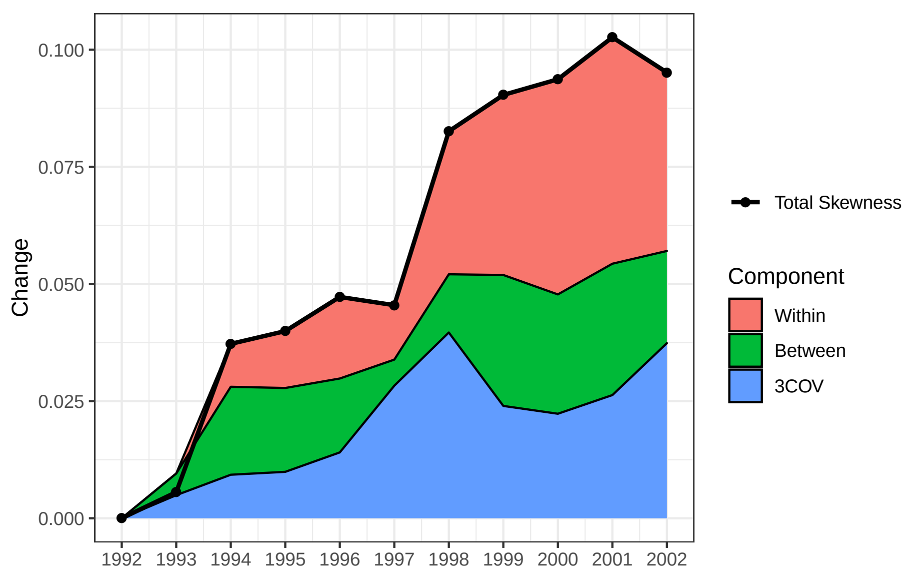


Figure A13: Skewness Decomposition by Education and Experience

Skewness decomposition (Equation 9) by the interaction of years of schooling and years of experience. Changes in each component are plotted relative to the baseline year (1992). Wages at the top and bottom 5% were dropped (see Section 4).
Source: CPS Outgoing Rotation Groups

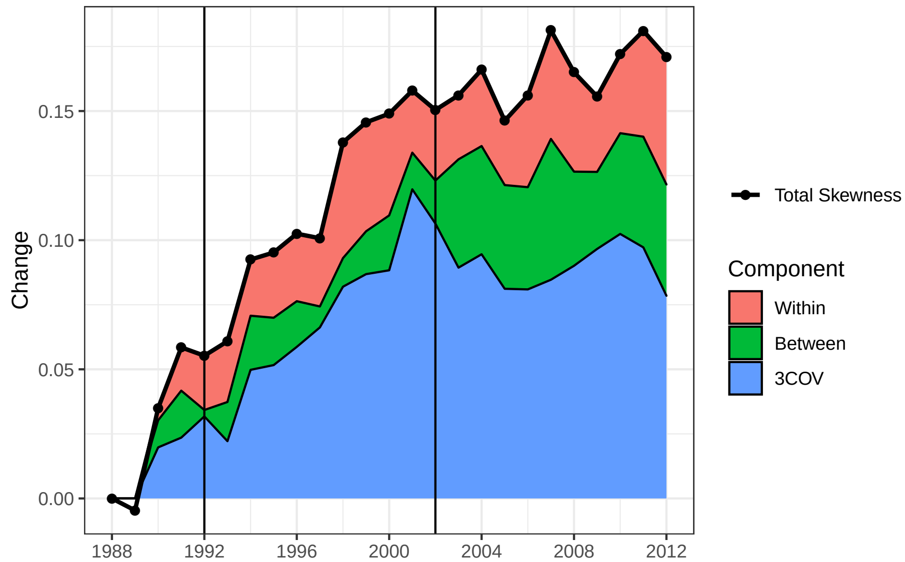


Figure A14: Skewness Decomposition by Occupation 1988–2012

This figure repeats the exercise of Figure 9 for a longer time period. Decomposition is based on Equation 13. Vertical lines represent changes in 3-digit occupational coding. Occupational coding is based on the Autor and Dorn (2013) crosswalk. Changes are reported relative to the baseline year (1988), which is approximately the beginning of the rise in skewness. Wages at the top and bottom 5% were dropped (see Section 4).
Source: CPS Outgoing Rotation Groups

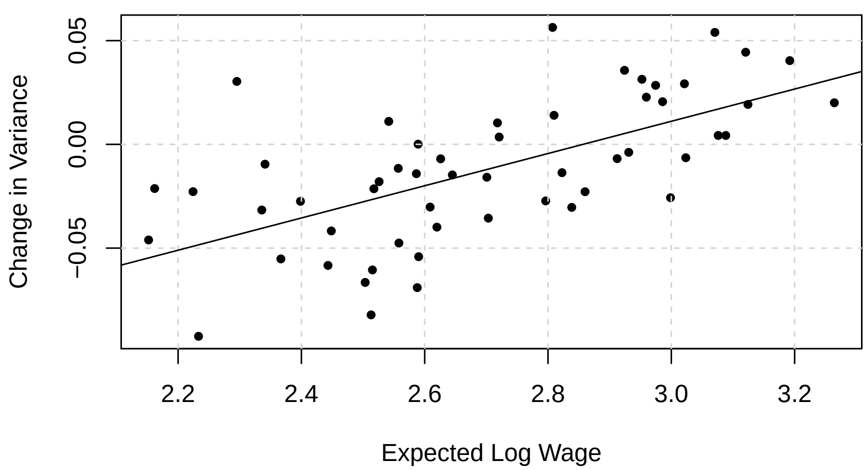


Figure A15: Changes in Occupational Inequality 1992/3–2001/2 by Mean Log Wages

This figure plots all occupations with at least 0.5% of the total working hours (top 47 out of 501 occupations that include 53% of the total working hours). The expected log wage (X-axis) is the average log wage in an occupation during the entire period (1992–2002). Change in Variance (Y-axis) is the difference between the average of the first and last two years (I pool two years together to reduce errors due to small sample size). The line is the best linear fit to the points. Wages at the top and bottom 5% were dropped (see Section 4).
Source: CPS Outgoing Rotation Groups

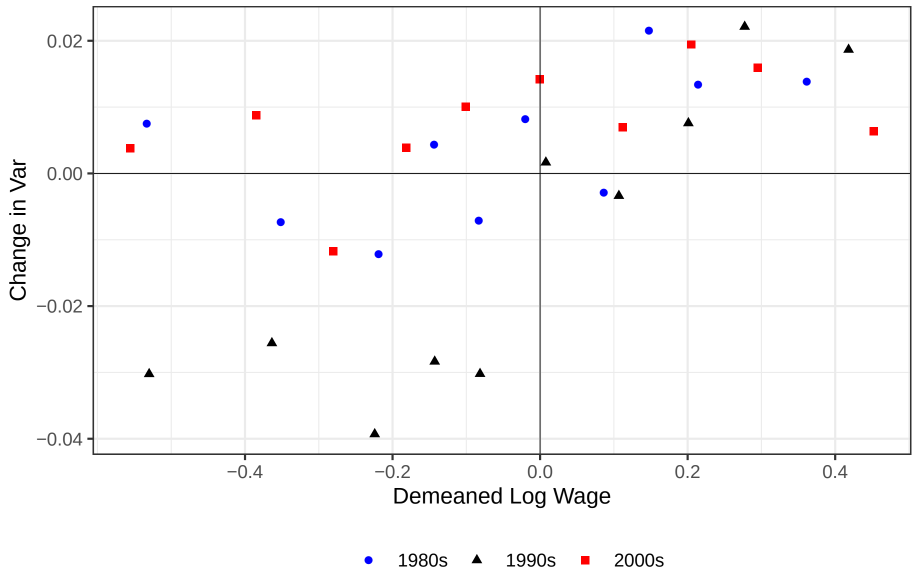


Figure A16: Binned Changes in Occupational Inequality by Mean Log Wages

This figure plots the change in occupational variance as a function of occupation mean for three periods. The periods were chosen such that occupational coding is held fixed: 1979–1991, 1992–2002, 2003–2012. In each period I divide occupations into 10 equal size bins of occupations based on occupation mean log wage, weighted by occupation size. Each point calculates the mean log wage in the baseline year, and the change in within occupation variance in this period. The distribution of log wages is demeaned separately in each period.
Source: CPS Outgoing Rotation Groups

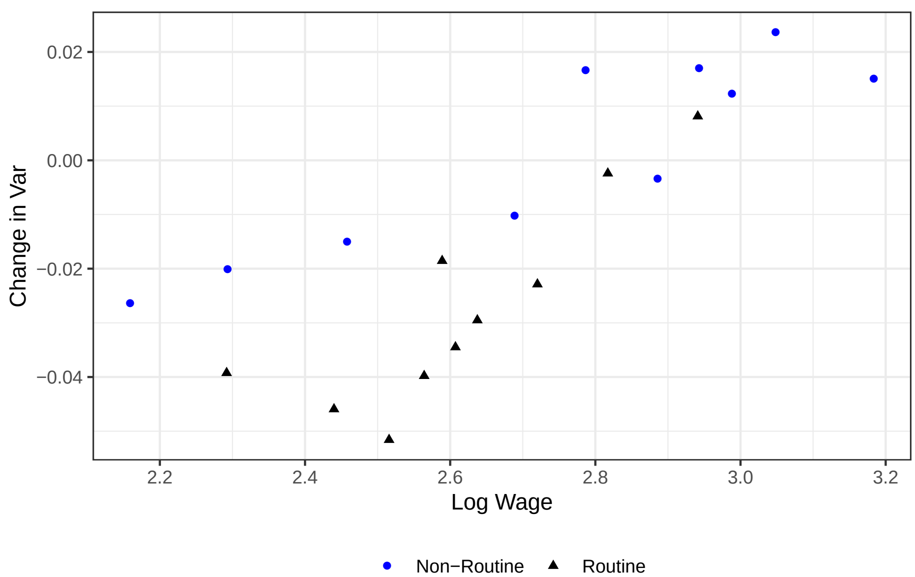


Figure A17: Binned Changes in Occupational Inequality by Mean Log Wages 1992-2002

This figure plots the change in occupational variance as a function of occupation mean separately for routine and non-routine 3-digit occupations. Occupations are binned based on occupation mean log wage separately for routine and non-routine occupations using 10 equal-sized bins (deciles) of occupations, weighted by occupation size. Each point displays the mean log wage in the baseline year 1992, and change in the variance between 1992–2002 when occupational coding is fixed. Routine occupations include all occupations classified as administrators, operators, and production workers based on 1-digit occupation coding.
Source: CPS Outgoing Rotation Groups

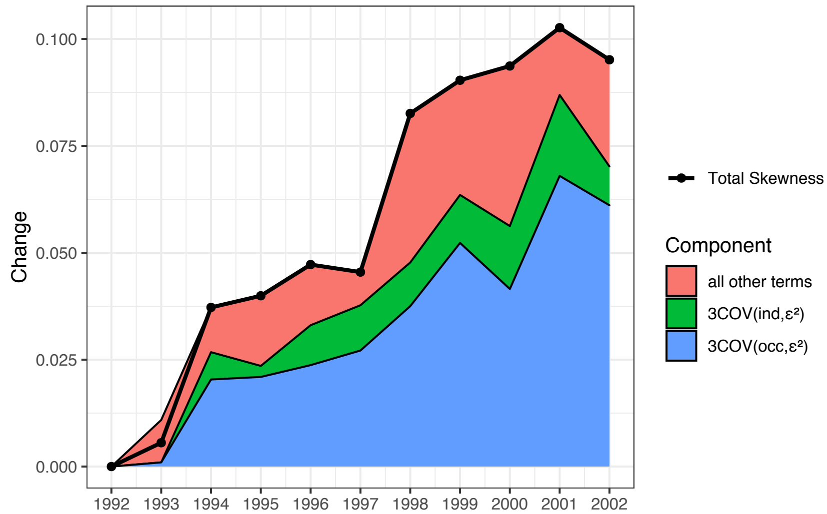


Figure A18: Skewness Decomposition by Occupation and Industry

Linear skewness decomposition (See Appendix D) by occupation and industry (Equation 13). $COV\left(occ, \varepsilon^2\right)$ (in blue) and $COV\left(ind, \varepsilon_i^2\right)$ (in green) are the covariance of occupation and industry premiums with the unexplained variance and are plotted separately. All other terms are aggregated (in red). Changes in each component are plotted relative to the baseline year (1992). Wages at the top and bottom 5% were dropped (see Section 4).
Source: CPS Outgoing Rotation Groups

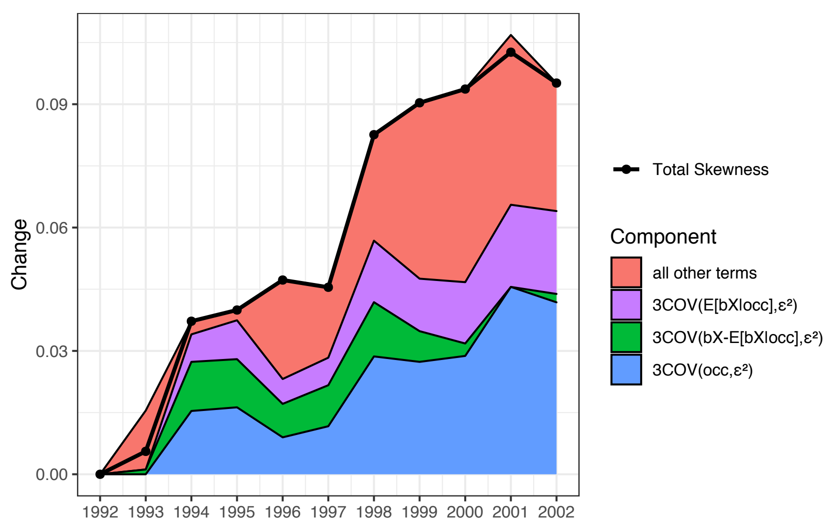


Figure A19: Skewness Decomposition by Occupation, School, Experience

Linear skewness decomposition (see Appendix D) by occupations, years of schooling and years of experience, where education and experience premiums are separated into the occupation mean and a residual (Equation 14). The covariances with the unexplained variance are plotted separately while all other terms are aggregated. occ is the occupation premium for each 3-digit occupation (conditional on education and experience), $E\left[\beta X|_i occ\right]$ is the mean of education and experience premiums in each 3-digit occupation, and $\beta X_i - E\left[\beta X_i|occ_i\right]$ are the demeaned premiums for education and experience. Changes in each component are plotted relative to the baseline year (1992). Wages at the top and bottom 5% were dropped (see Section 4).
Source: CPS Outgoing Rotation Groups

F Appendix Tables

	Abstract	Routine	Manual
Abstract	1		
Routine	.74	1	
Manual	.83	.69	1

Table A1: Correlation of Occupational Skills

Pearson correlation coefficient between $\widehat{\theta_{ij_0}}, \widehat{\theta_{ij_1}}$ for pairs of occupational categories. The values of $\widehat{\theta_{ij}}$ are estimated using Equation 3, allowing θ_i to vary by the three occupational categories. Routine workers are defined as workers in administrative, production or operator occupations, classified by the first occupational coding digit. Abstract workers are defined as workers in managerial, professional and technician occupations. Manual workers are defined as workers in service, sales, and agriculture. Each correlation is calculated using all workers who ever worked in both categories.
Source: PSID