

Decomposing Wage Polarization

Oren Danieli*

Harvard University

April 23, 2019

Preliminary and Incomplete

Abstract

Several papers have argued that the relative decline in middle-wages in the U.S. during the 1990s (“Wage Polarization”) is the result routine-biased-technological-change (RBTC). However, some questions remain open: Why did middle wages decline when most of the routine workers are paid below median wages? Why did wage polarization stop around the year 2000? And why weren’t standard decomposition methods able to show a direct link between RBTC and wage polarization. This paper uses novel empirical methods to address these questions by showing that RBTC caused a decrease in returns to skill at routine-heavy occupations. I use a decomposition method based on the third moment (skewness), which quantifies the contributions of different factors to the overall increase in wage polarization. I find that the drop in inequality at routine occupations is the main driver of wage polarization. Using panel data I show this drop is driven by a drop in return to skills in those occupations. This is mostly affecting higher-skilled workers in routine occupations, which were earning close to median wages. It was not captured with other decomposition methods as returns to skill are asymmetric, and also increase in some occupations. Finally, I find that consistent with a drop in returns to skill, the employment declines in routine-heavy occupations (“job polarization”) is only occurring for higher-skilled workers. Once higher-skill workers leave routine occupations, any further decrease in demand for routine occupations is only affecting low-wage workers, explaining why wage polarization has stopped.

*Email: odanieli@fas.harvard.edu. I am grateful to Lawrence Katz, David Autor, Roland Fryer, Thomas Lemieux and Eran Yashiv for very helpful comments and suggestions.

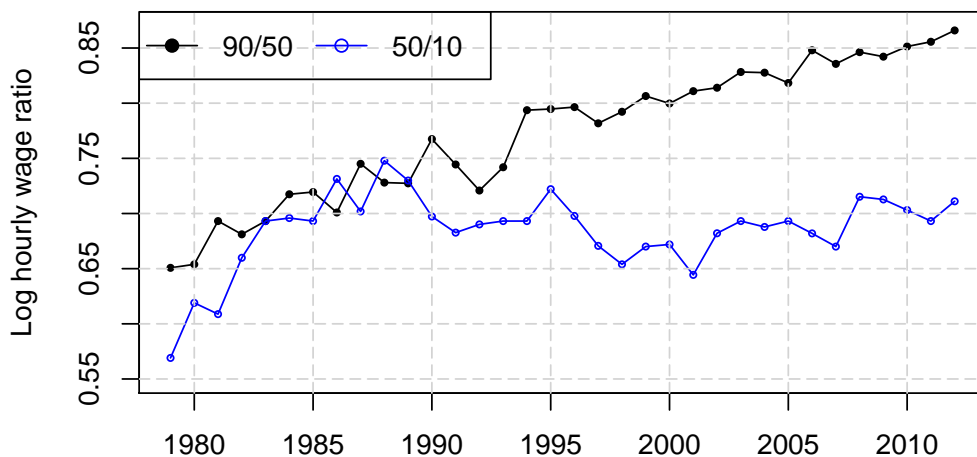


Figure 1: All Population 90/50 and 50/10 Log Hourly Wage Ratio

Quantiles are calculated at the hours level, using sample weights multiplied by hours worked.

Source: CPS Outgoing Rotation Groups

1 Introduction

During the decade of the 1990s, the U.S. labor market has experienced wage polarization - a substantial relative decline in middle wages. Wages around the median declined compared to both wages at the top and at the bottom. Figure 1 show this trend was different from the broad increase in inequality at all parts of the distribution that occurred both before, in the early 1980s, and after, starting from the early 2000s. The leading explanation for this trend is a Routine-Biased-Technological-Change (RBTC): a decline in demand for routine tasks that used to require middle-skill workers, but now can be automated (Autor et al., 2006). This hypothesis is supported by the extensive decline in employment at routine-heavy occupations in most developed countries (Goos et al., 2014).

But while there is a variety of strong evidence that support the mere existence of RBTC, very little is known about whether the magnitude and the shape of its effect on wages can explain large portions of the decline in middle-wages. It is also harder to explain why RBTC is mostly affecting middle-wage workers, where many routine workers are actually concentrated at the bottom of the distribution. Another puzzle is why the decline in middle wages stopped in mid-2000s when evidence suggest RBTC continues long after that. These unresolved questions left room to consider other explanations to wage po-

larization such as an increase in 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.

One reason why we know so little about the relative contribution of each explanation to the general trend is that we do not have the right tool to decompose wage polarization. While there is quite a diverse toolkit to decompose wage inequality, wage polarization possess new challenges that need to be addressed. The key challenge is that, by definition, wage polarization has asymmetric trends as the wage gap increases at the upper tail of the wage distribution, but decreases at the lower tail. Most existing decomposition methods can isolate the effects of prices and compositional changes in the overall trends in wages. But if prices changed in opposite directions at different parts of the distribution, this will not be captured.

This paper will show that RBTC does in fact account for almost the entire trend of wage polarization. I will do so using a decomposition method which I call “Skewness Decomposition”. Using this method I estimate that 79% of wage polarization are due to asymmetric trends in occupations: a rise in inequality at high paying occupations and a decrease at low paying occupations. Both of these trends are generating a relative drop in middle wages. But while rising inequality at high paying occupations has been steady for several decades, the period of the 1990s is unique for its decline in inequality at low-paying occupation. Using panel-data I show that this is driven by a wage compression at routine occupations.

Using a simple model, I show how RBTC can explain these findings, and address several puzzles about the effect of RBTC on wages. The key distinction from previous models is that the new technology is not equally substitutional to all workers in routine tasks. Instead, new technology is substitutional to the usage of skills in those tasks. This generates a decrease in return to skills, only in those heavily-routine occupations. Even though most routine workers are earning below median wages (Autor and Dorn, 2013), such a decrease in return to skill is affecting mostly middle-wage workers, because that’s the typical wage level for the higher-skilled workers in routine occupations. As a result, higher-skilled workers leave these occupations, gradually making routine occupations composed of more low-skilled workers, without many middle-wage workers. From that point, any further RBTC is generating a decrease in demand for low-skill workers, and inequality starts increasing again at the lower tail of the distribution.

I start by outlining the theoretical framework of the paper. The key assumptions is that workers are characterized by a one dimensional skill, and occupations vary in their returns to that skill. This means that workers with a higher skill level will have a com-

parative advantage in occupations with a higher return to skill. Occupations with mainly manual tasks have the lowest return to skill, and occupations with mainly abstract tasks have the highest return. Occupations with mostly routine tasks provide comparative advantage to workers from the middle of the skill distribution. In the empirical section, I test these assumptions and show that they work reasonably well in the data.

The key parameter in the model is the elasticity of substitution between the skill and the new technology. The model is general enough to allow for a technological change that is skill neutral, as in many previous models (i.e., [Acemoglu and Autor, 2011](#); [Cortes, 2016](#)). It also allows for a technological change that is increasing returns to skill as in [Jung and Mercenier \(2014\)](#), or decreasing as causal evidence suggests ([Gaggl and Wright, 2017](#)). The model will allow us to derive different predictions for each one of these cases, which we will then be able to test in the data.

In order to quantify the overall effect of RBTC on the wage distribution, we need to use some decomposition method. Decomposition methods are the standard way to analyze changes in distributions across time, or between groups ([Fortin et al., 2011](#)). For instance, using variance decomposition, it is very straightforward to see that most of the increase in income inequality during the 80s was through increasing gaps between education groups ([Katz and Autor, 1999](#)). To study the effect of RBTC, we would like to decompose by occupations, as it is the key component of the underlying theory behind it.

However, there are various reasons why existing decomposition methods would not work. Since wage polarization could both increase or decrease inequality, methods that are designed for inequality analysis would be irrelevant. Instead we need a new statistic that measures polarization, that could be decomposed. More general decomposition methods, that simulate full counterfactual distributions ([Juhn et al., 1993](#); [DiNardo et al., 1996](#)) or that could decompose any statistic ([Firpo et al., 2009](#)) also pose a different set of problems. Generally, they are harder to interpret due to reasons such as path-dependence or an arbitrary choice of baseline year ([Fortin et al., 2011](#)).

But more specifically, these methods are not suitable to capture asymmetric trends. Broadly speaking, most decomposition methods are decomposing any change in wages to change in composition, prices and a residual. In the context of occupations, they can capture any wage trend that is driven by transitions between occupations, or change in mean wages at each occupation. But as the model predicts, and as the results suggest, a large share of the trend in wages is through asymmetric effect on inequality within occupation. This will be missed by those methods that will classify it in the residual component.

Skewness decomposition can solve all of those problems. In analogy to inequality that

can be measured with the second moment of the distribution of log wages, polarization can be measured with the third moment of that distribution - the skewness. As expected, the skewness is rising exactly when wage polarization is increasing. The advantage of using skewness is that, similar to variance, it can be decomposed into independent components, that do not depend on any arbitrary choice of year or order of components.

Skewness decomposition breaks the trend in wage polarization into three components, for each choice of groups. The researcher can choose a group such as education level, occupations, industry etc. and decompose the increase in polarization by that group. Similar to variance decomposition, skewness decomposition has a between-group and a within-group components. The between-group captures any trend that is driven by changes in prices between the groups. The within component captures any unexplained trend that is orthogonal to that grouping. But there's an important third component that captures the correlation between the mean and variance at each group. This component will be higher when higher paying groups have larger inequality. This allows skewness decomposition to capture some types of asymmetric wage changes, that are missed in other methods.

Applying skewness decomposition to data on the wage distribution in the U.S. gives a direct evidence that wage polarization is driven by occupational trends. So far, the main suggestive evidence that linked wage polarization to occupations was employment and wage decline in routine-heavy occupations during the same period of time ([Acemoglu and Autor, 2011](#); [Cortes, 2016](#)). Those routine occupations tend to pay low to middle wages, suggesting that this drop in demand could be linked to wage polarization. My results quantify the share of the rise in skewness that can be attributed to occupations and show it can explain 93% of the overall increase. Comparing these results to the result when decomposing by other categories, such as industries or education, shows clearly that the trend is driven by occupations. I use data from the CPS outgoing rotation group since it measures the price of labor most precisely ([Lemieux, 2006](#)).

But in contrast to the prediction of earlier models for RBTC, the effect is driven by asymmetric inequality trends within occupations. Earlier models ([Acemoglu and Autor, 2011](#)) assume that all routine workers have the same skill level, and that RBTC simply changes the price of routine tasks. This setting predicts that wage changes would be made through the decline in premium for routine occupations, and doesn't allow for any distributional changes within routine occupations. If that was the case, skewness decomposition should have shown that the increase in skewness is in the "between" component. However, I find that almost the entire increase is in the correlation component. Wage polarization happened because inequality increased at high paying occupations, but de-

creased at low-paying occupations. This trend was documented by Lemieux (2007), and using skewness decomposition I find it is actually the main driver of wage polarization, far beyond the changes in occupation premiums.

Wage polarization is driven by the drop in inequality in heavily routine occupations. While inequality at high-paying occupations is steadily increasing for decades, the drop in inequality at low paying occupations is unique to the 1990s. This is the reason why wage gaps are falling at the bottom of the distribution, generating wage polarization during that period. Most of this drop is in low-paying routine occupations, while other low paying occupations like services don't experience any such trend.

There are two distinctive explanations for this trend, and I use panel data to decide between them. One reasonable explanation is that the drop in demand for workers in routine tasks made the highest and lowest paid workers to leave, thus making the distribution more equal. I show the model generates this pattern when RBTC is skill neutral. Cortes (2016) shows empirical evidence that generally workers from the edges of the wage distribution are more likely to leave, so potentially this trend exacerbated during the period of wage polarization. An alternative explanation is that returns to skills have declined, generating a wage compression even without any transitions. To distinguish the two explanations I use panel data from the PSID.

I find that a decrease in returns to skill at routine occupations is the key driver of wage compression in routine occupations, and hence of wage polarization. I estimate an interactive fixed effect model, that allows the return to skill to vary across time between occupations. I find evidence for a decrease in return to skill in routine occupations. This decrease in routine occupations makes wages to drop mostly for the highest earning routine workers, who are concentrated around the median of the overall wage distribution. As a result, higher-skilled workers have the highest incentive to leave routine occupations. Indeed, I find that almost all of the employment decline in routine occupations ("job polarization") is driven by workers with above average skills. The share of below average workers in routine heavy occupations remains steady.

These results explain how RBTC can first generate a decline in middle wages, which then stops and turns into a decline in lower wages. At first, there are several high skill workers in routine occupations. Because wages in routine occupations are relatively lower, those workers' wages are close to the median wage of the entire labor market. The drop in demand is strongest for those workers, who experience the largest drop in their wages and eventually leave. This generates both wage and job polarization. Gradually, the composition of routine occupations becomes more affluent with low skilled workers. Any further decline in demand for routine tasks would then be a drop in demand for

lower-wage workers and generate a decline in lower wages.

Other explanations cannot explain these findings. Institutional explanations like an increase in minimum wage or decline in unionization, as well as macroeconomic explanation like low unemployment do not fit these empirical findings. In general, these explanations are not particularly related to occupations, more than they are to industries, education levels or other partitions of the data. Moreover, they shouldn't affect workers in routine occupations any differently from workers in services or in more abstract-heavy occupations. Increase in demand for service occupations seems to be more of a result of polarization and not the main driver of it. If it were, we would expect most of employment decline to be driven by lower-skilled workers, which is the opposite of what I find.

The rest of the paper is organized as follows: Section 2 outlines a theoretical framework for RBTC. In Section 3 I will discuss skewness decomposition in detail and its advantages over existing methods. Section 4 describes the data sets used throughout the paper. Section 5 will present the results of skewness decomposition and show wage polarization is driven by occupational trends, supporting the RBTC hypothesis. Section 6 will present the evidence for a decrease in returns to skill at routine occupations, using panel data. Section 7 will conclude by reexamining all the evidence and show how they fit the model predictions.

2 Model

In this section I'll present the theoretical framework for RBTC that will be used throughout the paper. I use a model that highlights the different return to skill in each occupation, in the spirit of [Jung and Mercenier \(2014\)](#) and [Cortes \(2016\)](#). The model allows to introduce new technology that could be substitutional, complimentary or neutral to the workers' skill. In each of those cases technology will affect the distribution of wages within occupation, between occupations and the composition of workers in each occupation, but in different ways. I'll discuss the differences in these predictions and how they can be observed in the data.

2.1 Simple Model of Occupational Sorting

Assume workers have a one dimensional skill. We will mark this skill level by θ_i . This assumption is more general than models that assume only a discrete number of skills ([Katz and Murphy, 1992](#); [Autor et al., 2006](#); [Acemoglu and Autor, 2011](#)), but less general

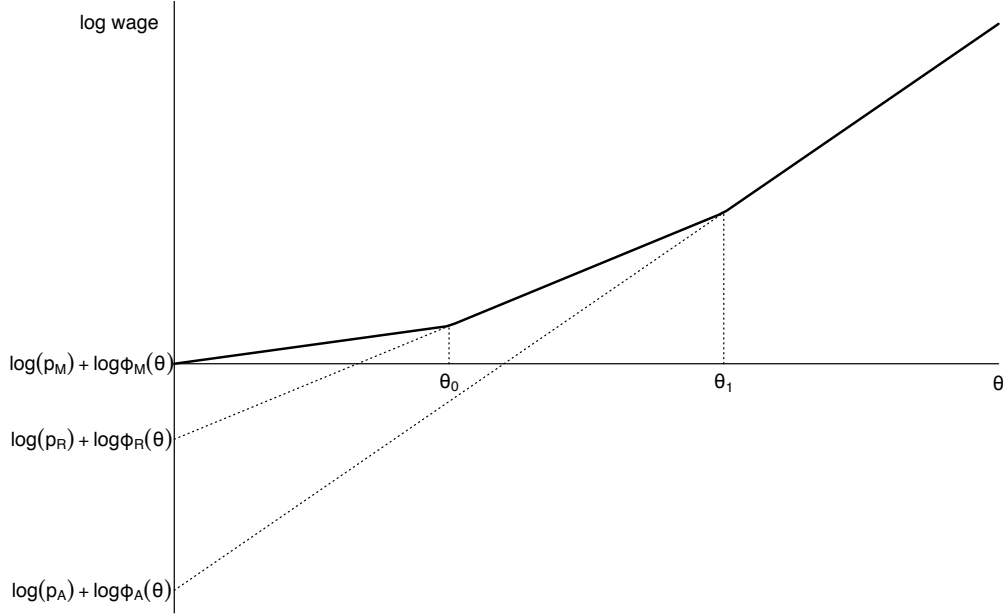


Figure 2: Equilibrium log Wage by Skill

than models that allow for multi-dimensional skills (Roy, 1951). I will test this assumption empirically in Section 6, and show that adding more dimensions of skills doesn't improve precision substantially.

Occupations will be 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

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

so the manual occupation has the lowest usage of skill, and abstract has the highest.

Perfect competition sets wages at their marginal productivity. I assume that identical firms are competing for the same workers. 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 will sort into occupations based on comparative advantage. Condition 1

guarantees that there would be two thresholds θ_0, θ_1 such that any worker with $\theta_i < \theta_0$ will choose to work in manual, any worker with $\theta_0 < \theta_i < \theta_1$ will choose routine, and $\theta_i > \theta_1$ will choose abstract (Jung and Mercenier, 2014). Workers with skill level that exactly equals the threshold will be indifferent, hence the following two equations will hold in equilibrium

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

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

Figure 2 presents this graphically, by plotting the equilibrium log-wages by skill level θ_i .

2.2 Technological Change

I assume that the three intermediate goods, together with capital are producing a final good. Mark the total amount produced from each intermediate good by M, R, A with

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

and assume total capital K is fixed.¹ The final good is the output of a CRS function $Y = F(M, R, A, K)$.

I focus on a technological change that is shifting the production of routine goods from labor to capital. I assume that technological change is affecting only φ_R directly, and φ_M, φ_A are left unchanged for simplicity. However, wages in the manual and abstract occupations will be affected as well in general equilibrium. Specifically I assume the following functional form

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

where τ is the level of technology. σ is the elasticity of substitution between technology and skills.

Capital is complimentary to routine work, so if routine workers produce more, a larger share goes to capital. To simplify, I assume a specific functional form to F where

$$Y = M^{\alpha_1} (R^\rho + K^\rho)^{\frac{\alpha_2}{\rho}} A^{1-\alpha_1-\alpha_2}$$

and $\rho < 0$. The Cobb-Douglas structure implies that a constant share of Y goes to manual

¹The results could be generalized for a case where capital has increasing marginal costs.

workers, and abstract workers.² A constant share of $\alpha_2 Y$ goes jointly to routine workers and capital owners.

RBTC would then be modeled as an increase in technology level τ . This allows for every worker to produce more, since $\frac{\partial \varphi_R(\theta, \tau)}{\partial \tau} > 0$. Therefore, RBTC will generate an increase in R , which would increase the share going to capital from total output on the expense of total share to routine workers, since they are complements. An increase in τ can be thought of as an improvement in quality of computers or robots. While this makes each worker more productive, it also allows to produce the same quantity with fewer workers, making labor prices drop.

To understand the effect of RBTC on the labor market, it's key to know whether the elasticity of substitution between technology and skills (σ) is greater, equal, or less than 1. The effect of an increase in τ on different levels of θ depends on

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

which has the same sign as $1 - \sigma$. If σ is 1, RBTC is skill neutral as in [Cortes \(2016\)](#). The effect on log wages will be the same for all workers in routine occupations. If $\sigma < 1$, as hypothesized by [Jung and Mercenier \(2014\)](#), the new technology is increasing gaps between skill levels. At some point, routine occupations could actually have a comparative advantage for the highest skill workers. If $\sigma > 1$, technology is substitutional to skills, and returns to skill will decline.

In any case, simple decomposition methods will not capture the entire effect of RBTC. If $\sigma \neq 1$, RBTC is not only changing the premium for routine occupations, it also changes wage gaps within routine occupations. Even if $\sigma = 1$, the distribution of wages within occupations will change. The lowest (highest) skill workers will leave to manual (abstract) occupation. This will decrease inequality at routine occupations, and increase it in manual and abstract occupations. Any such changes are not captured by a decomposition method that is focusing on prices and compositions.

²A more general CES function with complementarities between occupations (as empirical evidence suggest) could be used instead, and the results will only be amplified. In case of complementarities, RBTC that allows for higher R will cause larger share of total income to go to manual and abstract occupations. This will increase the decline in routine wages. For simplification we're using CD.

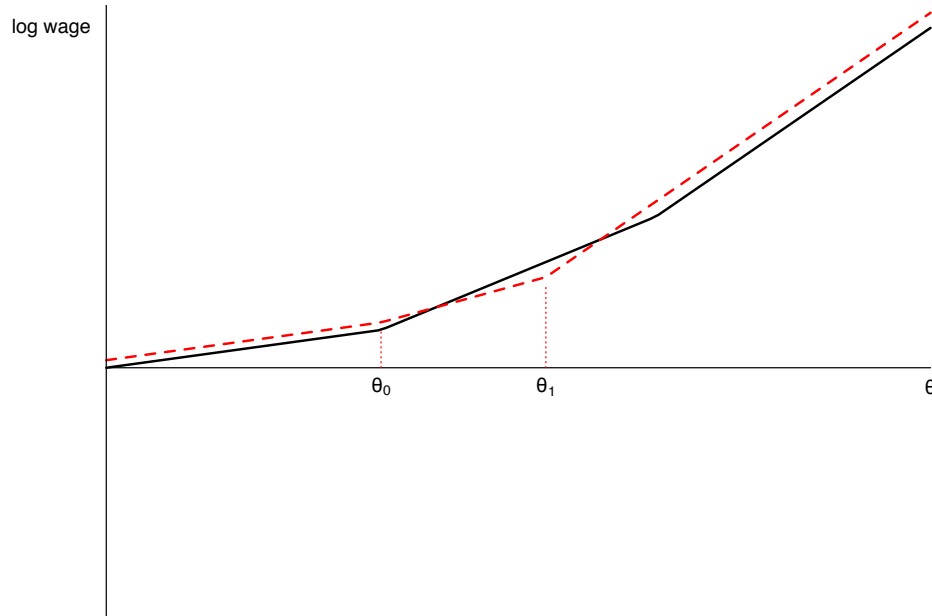


Figure 3: Change in log Wage for Small RBTC

2.3 Decrease in Return to Skill

I will focus in more details on the case where technology and skills are substitutional, as it will have the highest fit to my empirical results. I will describe the different stages of a gradual increase in τ , and how they can be measured empirically.

At first stage, we would see wage polarization. Wage polarization will be generated by a decrease in inequality in routine occupations. Figure 3 illustrates this case where the return to skill in routine occupations becomes flatter. This generates lower gaps between workers who stay at routine occupations. The most significant wage drop is for workers at the top wage levels of the routine occupation, which are approximately in the middle of the overall distribution of skills. As a result, higher skill routine workers will now have their comparative advantage in the abstract occupations, so θ_1 will drop. This will generate a drop in employment at routine occupations (“job polarization”). The effect on θ_0 is unclear. Overall, we expect job polarization to be driven mostly by higher skilled workers. I will argue that this behavior closely fits the empirical findings on wage and employment patterns in the late 1980s and 1990s.

Wage polarization stops when low skilled workers start having a comparative advan-

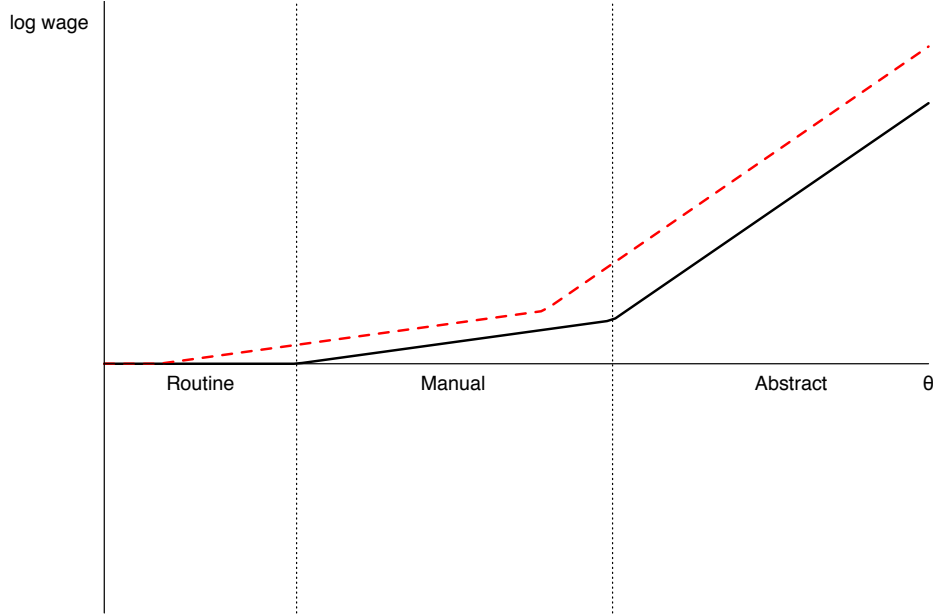


Figure 4: Changes in log Wages for Large τ

tage in routine occupations. This will occur when the first inequality in Equation 1 no longer holds for all θ_i . At that point, higher skill routine workers at routine occupations will continue to leave. But some of the employment drop would be offset by joining into routine occupations from the bottom of the skill distribution.

Finally, comparative advantage will be flipped, and inequality will start growing at the bottom. This will occur when

$$\forall \theta : \frac{\partial \log \varphi_R(\theta)}{\partial \theta} < \frac{\partial \log \varphi_M(\theta)}{\partial \theta} \quad (2)$$

At this point, routine occupations employ the lowest skilled workers. Any further increase in τ will make wages relatively drop for the lowest paid workers as shown on Figure 4. A growing share of routine goods will be produced by capital. Hence, job polarization will continue as more workers will leave routine occupations. More routine goods will be produced, making workers in manual and abstract occupations more productive. This will increase their wages, and so increase inequality at all parts of the distribution. I will argue that this behavior fits empirical findings from mid 2000s onwards.

3 Skewness Decomposition

The main empirical tool I use in this paper to quantify RBTC in the data is skewness decomposition. Before diving into its details, I'll briefly review alternative decomposition methods and why they are unable to capture the effects of RBTC based on the theoretical framework. I'll then talk about how skewness decomposition can address all these challenges, and how it will be able to quantify the contribution of the predicted wage changes from the model.

3.1 Challenges to Standard Decomposition Methods

The first challenge with decomposing wage polarization is that it's unclear which statistic we should decompose. When studying inequality, various indices such as variance of log wages, or Gini, could be used to measure inequality levels. These indices are useful to see how inequality varies across different time periods or countries. They can also be naturally decomposed, allowing researchers to better understand why inequality varies. Variance of log wages for instance, can be easily decomposed into a between and a within component

$$V(\log w) = \underbrace{E[V(\log w|X)]}_{\text{Within}} + \underbrace{V(E[\log w|X])}_{\text{Between}}$$

This very simple decomposition allowed us to learn that a large share of the increase in inequality during the 1980s was due to an increase in inequality between education groups.³ Ideally, we would want to be able to perform a similar exercise for wage polarization.

However, there's no clear single index to measure wage polarization. Since inequality is rising at the top but declining at the bottom, we cannot use the same indices to measure wage polarization. So far, we have used both the 90/50 and the 50/10 wage ratio to describe wage polarization. We could potentially use some combination of those two measures. However quantiles generally cannot be decomposed as elegantly, forcing us to use more general decomposition methods.

More general methods have their own drawbacks. Methods such as [Juhn et al. \(1993\)](#); [DiNardo et al. \(1996\)](#) are constructing counterfactual distributions in partial equilibrium, holding some components fixed. Since the full distribution is constructed, every statistic can then be calculated, making those methods very general. However, the interpretation of these methods is harder, as some arbitrary choices could potentially affect the results

³See [Yitzhaki and Schechtman \(2013\)](#) for Gini Decomposition.

quite substantially. The order of components, which is usually completely arbitrary, has an effect on the contribution of each component.⁴ The choice of a baseline year is also arbitrary, and could affect the results as well (Lemieux, 2010). Moreover, those methods cannot accommodate any changing in the coding of the category that is used, making it hard to study long periods of time.

Most importantly, most methods don't have the right component to capture the effect of RBTC. Most decomposition methods are decomposing the changes in the wage distribution into changes in composition, prices and an unexplained part. So for example, decomposing by occupations, will allow to study the effect of occupational transitions, and trends in mean wage at each occupation. As we discussed in Section 2.2, RBTC generates important trends within occupations. Inequality within manual, routine and abstract occupations could change substantially as well. Moreover, inequality could change in opposite directions in each occupation. This is why previous papers have not been able to show RBTC is generating most of wage polarization.

The closest approach to this paper is perhaps a re-centered influence function (RIF) regression. This method was used to study wage polarization in a paper by Firpo et al. (2013). This method doesn't suffer from path-dependence, but still requires an arbitrary choice of baseline year. The main drawback of this method is that it doesn't capture trends within occupations. Firpo et al. (2013) document that inequality trends within occupations are asymmetric, and inequality drops at routine occupations. But the RIF-regression cannot quantify the effect of these trends, as they are not reflected in occupation premiums (prices), nor in occupation compositions.

Skewness decomposition will be able to address all these challenges.

3.2 Skewness Decomposition

Wage polarization can be measured with skewness. Skewness is the third standardized moment and is defined by

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

It provides a measure for the asymmetry of the distribution relative to the mean. Figure 5 shows graphically the link to wage polarization by plotting the derivative of the em-

⁴These methods are building counterfactual distributions allowing an additional component (such as composition, prices or residual) to vary each time. The effect of that component is the change in the statistic at each stage. But reordering the components will change the results (the effect of prices holding composition fixed or allowing it to change is different). See Fortin et al. (2011) for an extensive discussion on this.

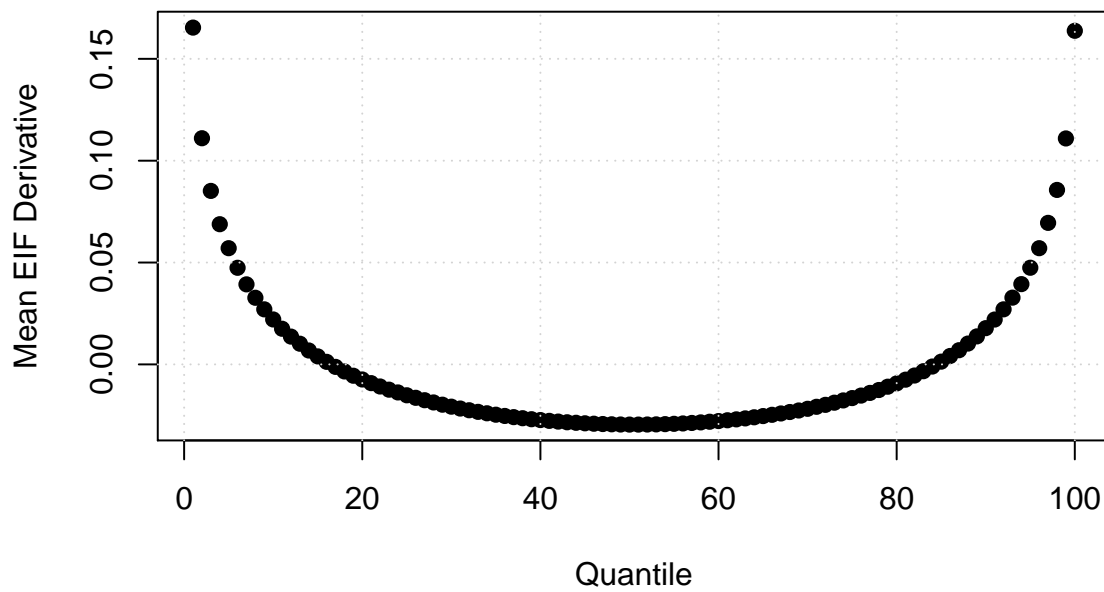


Figure 5: Derivative of EIF on Skewness for Standard Normal Distribution

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 k th order statistic at the sample point. The figure shows the mean over the 1,000 samples of this derivative.

pirical influence function at each quantile for a standard normal distribution. Intuitively, it shows the effect on skewness of a small increase in log wages for each quantile of the distribution, in case the log wage distribution was normally distributed. This shows that skewness increases exactly when wages at the edges increase compare to the middle. This pattern aligns quite well with the real trends in wages by quantile that I will show in Section 5.1.

The main advantage of using skewness is that it has a simple decomposition. Writing Y as standardized log wages, X some category we wish to decompose by, and μ_3 as the third centralized moment we can write

$$S(Y) = \mu_3(Y) = \underbrace{E[\mu_3(Y|X)]}_{\text{Within}} + \underbrace{\mu_3(E[Y|X])}_{\text{Between}} + \underbrace{3\text{COV}(E[Y|X], V[Y|X])}_{\text{Correlation}} \quad (4)$$

This decomposition was previously introduced by [Mincer \(1974\)](#), and it is the third moment analogous for the variance decomposition formula. Similar to variance decomposition, skewness decomposition break skewness into independent components as well. This means that there is no problem of path-dependence or any need to arbitrary define a baseline year.

The first and second component are quite standard. $E[\mu_3(Y|X)]$ can be thought of as a “within” component. It captures the remaining skewness within each category. It should be high when our division into categories is orthogonal to the increase in skewness, and therefore can be thought of as a residual component. $\mu_3(E[Y|X])$ captures skewness between groups and captures skewness due to differences between group averages. This component will be high if the increase in wage polarization is because of changes in prices for all workers in a group. For instance, changes in occupation premiums, return to education etc.

The third component captures the correlation between the levels and inequality at each group. Formally, it is 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, making the distribution more positively skewed. This component will allow to capture patterns like we predict from the theoretical framework, where higher paid occupations also have higher inequality. This pattern will be missed if we use a method that only decompose to occupational composition and prices. Therefore, it will be critical in order to test if wage polarization is indeed related to occupations and RBTC. This component will turn out to be capturing most of the increase in skewness during the period of wage polarization.

While in this paper I will use skewness decomposition to study wage polarization, it could 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 or raw wages (without logs). Any variation in these distribution across time or places can be analyzed with skewness decomposition. Similar decompositions exist for higher moments as well.⁵

4 Data

This paper uses two sources of data to get both, a large sample, and a panel structure.

The main analysis is done using the CPS outgoing rotation group. Since the theories I examine are related to the real price of labor, they are best captured using hourly wages.⁶ The CPS outgoing rotation group provides the most accurate representative sample of hourly wages (Lemieux, 2006). I use the years 1979-2012, that capture the early increase in inequality, wage polarization, and the return to increase in inequality. I use the same definition of the sample as (Acemoglu and Autor, 2011).⁷ Missing wages are dropped, which doesn't seem to affect the results as I'll show for the main results.

One important limitation of this data is its relatively high level of measurement errors. This problem is particularly severe at the edges of the distribution. Misreporting of working hours could lead to extremely high or extremely low values of hourly wages. As Figure 5 shows, skewness is especially sensitive to very high and very low incomes, making measurement error a large problem for this exercise. To deal with this, I drop the top and bottom 5% of the positive wages for the skewness analysis. The level of 5% was chosen in order to take the minimal cut, without substantial fluctuations between consecutive years in the skewness estimator. However, smaller cuts would also yield similar results, only noisier.⁸

In most of the analysis I focus on the years between 1992-2002. The reason is that a significant change in occupational coding has taken place before and after this period so

⁵A general way to find the decomposition is to write $Y = E[Y|X] + \varepsilon$ and then use Newton's Binomial formula on $E[(E[Y|X] + \varepsilon)^n]$. See also Appendix A.

⁶I multiply the CPS weights in the number of hours worked to focus on the real price of an hour of labor as explained in Lemieux (2010), which is also consistent with the literature.

⁷See Acemoglu and Autor (2011) for exact definition of sample sizes. I thank the authors for publicly sharing their cleaned data online.

⁸Cornfeld and Danieli (2015) analyze skewness in the Israeli labor market, using the entire distribution since measurement errors are not as severe, and reach very similar conclusions.

it is difficult to make informative comparisons to other years.⁹ As I will show, most of the increase in polarization occurred during this time period. I will separately implement several tests on the years before and after this period, and show that no similar trends occurred.

I maintain a consistent definition for routine occupations, similar to the previous literature. I first translate all occupational coding into the same coding, following [Autor and Dorn \(2013\)](#). I then define all administrative, operation and production jobs as routine, based on the 1-digit category. This is a similar classification to previous literature (for instance, [Acemoglu and Autor, 2011](#)) with one important exception - I do not classify sales occupations as routine occupations.¹⁰

In order to analyze transition and selection into occupations, I also use panel data. Specifically, I use the Panel of Income Survey Data in Section (PSID) for the same years. This data was chosen due to its long panel. Because it only includes a small sample of workers, I do not use it for the main analysis. I use the full core sample ("SRC"), without weights. The over-sample of low income household and the immigrant samples that were added in the 1990s are not used.

5 Results

5.1 Skewness Decomposition: Results

Using skewness decomposition, I will now show that wage polarization is indeed driven by occupational trends. This finding fits the theory that wage polarization is driven by RBTC as hypothesized by [Autor et al. \(2006\)](#). However, I'll show that the effect is more nuanced than previously thought. The effect is not driven by a drop in wages at middle-skill occupation, but more due to the drop in inequality within low-paying occupations.

I will start by showing that skewness does indeed capture wage polarization well. [Figure 6](#) 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](#). Starting from the late 1980s, until the early 2000s, when the 90/50 gap is rising and 50/10 gap is falling, we see an increase in skewness.

⁹Unlike other decomposition methods, skewness decomposition allows for the coding of X to change. However we cannot rule out that some of the trends in each component is driven by changes in coding.

¹⁰While information on task components suggests that sales is a routine occupation, I do not see the same wage nor employment patterns in these occupations. Including sales in 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 them weaker. This could be suggesting that sales occupations are not as easily automated as could be inferred from their O*NET description.

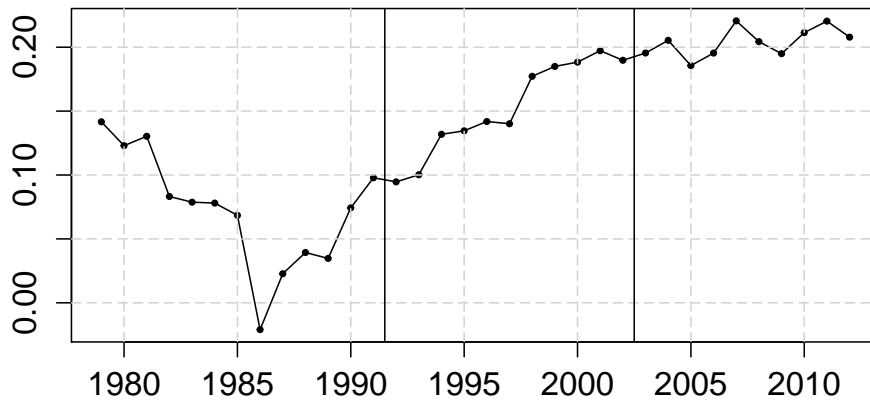


Figure 6: Skewness of Log Hourly Wage

Skewness (Equation 3) for a given year sample. Sample weights are used. The vertical lines are where changes in occupational coding took part. Wages at the top and bottom 5% were dropped (see Section 4).
 Source: CPS Outgoing Rotation Groups



Figure 7: Bin Scatter - Change in Log Wages 1992-2002

Change in log wages in each of 20-quantiles. 20 equal size quantiles are calculated for both 1992, 2002. The x -axis shows the value of mean log wage in each quantile. The y -axis plots the difference in mean log wages in each of the 20 quantiles from the end (2002) to start (1992) year. Sample weights are used.
 Source: CPS Outgoing Rotation Groups

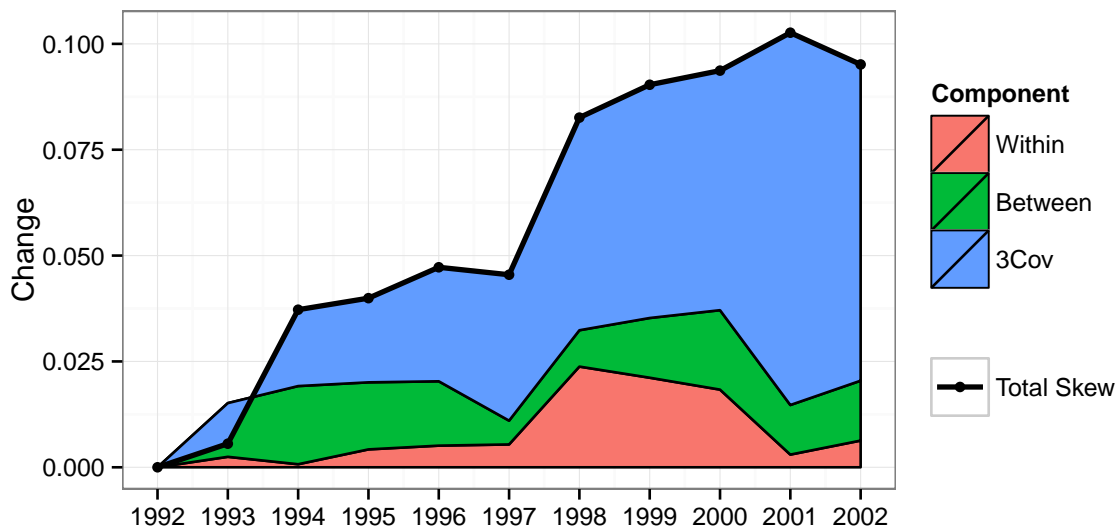


Figure 8: Skewness Decomposition by 3-Digit Occupation

Skewness decomposition based on Equation 4. Changes since base year (1992). The three components sum to the overall skewness (Equation 4). Wages at the top and bottom 5% were dropped (see Section 4). Source: CPS Outgoing Rotation Groups

The rise in skewness is driven by trends in all parts of the distribution. An increase in skewness occurs when the distribution becomes more tilted towards the left hand side. This means an increase in the gap between the middle and high wages and a decrease in the gap between the middle and low wages. Figure 7 plots 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 5. This implies that skewness has grown because of both the rise in wages at the top and at the bottom, making it a good fit to measure wage polarization.

I will decompose the rise in skewness of the distribution into three components, as described in Equation 4 for different choices of grouping (X). I will focus on the period between 1992-2002 since other years have different occupational coding (see Section 4). As Figure 6 shows this time period includes a big portion of the increase in skewness. We can decompose by any variable that is in the data, such as occupations, industries, education etc. I will look for grouping where the increase in skewness is captured by the between and then covariance component. Since any increase in the within component could be thought of as part of the increase that is unrelated to this choice of X variable.

Decomposing by occupations can explain almost the entire rise in skewness. Figure 8

presents the decomposition by 3-digit occupational coding.¹¹ The figure draws the change in each component since 1992, and the sum of the three, which will always equal the total. The first interesting conclusion from this figure is the importance of occupations in explaining the trends in skewness. The within component, which captures the part that is unrelated to occupations is very small, and could also be the result of classification errors. Therefore, .089 of the .095 total rise in skewness (93%), can be attributed to occupations.

Most of the increase in skewness is due to the covariance component. The between component, which is driven by the trends in mean wages in each occupation, can explain only 15% of the overall trend. The majority (79%) of the increase is through the rising correlation between the mean and the variance of log wages in occupations. In other words, the growing correlation between wage levels and inequality levels at each occupation. As I discussed in Section 3.1, this type of correlation is not captured by other decomposition methods, which is why earlier work heavily underestimated the contribution of occupational trends.

The results are not driven by any other worker characteristic I can observe in the data. Since occupations are correlated with workers skills or industries it is important to verify that occupations are not just proxying for some other worker characteristics. In Appendix Figures 23 and 24 I show the same decomposition results by industry, and education and experience. Clearly, in those cases the within component is much larger, suggesting that great portions of the trend in skewness is unrelated to these categories. Moreover, I can show that most of the increase in the between and covariance component in those cases is due to their correlation with occupations. Appendix A discusses how to decompose by more than one category using a linear model. Appendix Figures 25, 26 shows that by doing so we get that the increase is almost entirely through occupations.

These results strongly support the hypothesis that RBTC is generating wage polarization. The theory of RBTC argues that it affects differently workers performing different tasks. Occupations are the best proxy for tasks we have in most data sets. The fact that wage polarization, as measured with skewness, is driven by occupations greatly supports this explanation. Other possible explanations are not directly linked to occupations in any particular way.

However, these results also teach us new things about the way RBTC is affecting wages. Earlier models of RBTC such as Autor et al. (2006); Acemoglu and Autor (2011), would have predicted that the effect will be captured in the “between” component. These models argue that there is a drop in the price of routine tasks, making wages fall equally for all workers in routine-heavy occupation. This would have been captured in the be-

¹¹Appendix Figure 22 performs the same exercise using imputed wages and reaches very similar results.

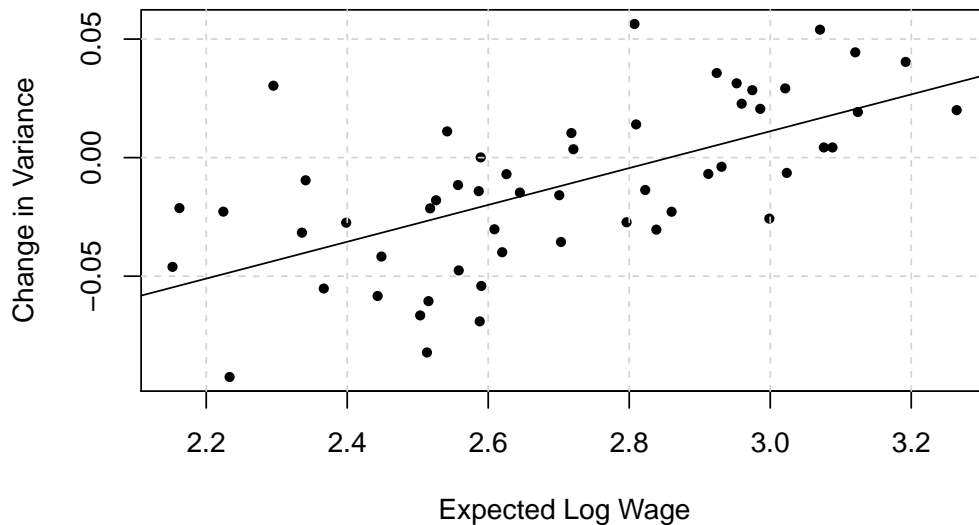


Figure 9: Changes in Variance by Expectation of Low Wages for Top Occupations 1992/3-2001/2

Wages at the top and bottom 5% were dropped (see Section 4). Includes 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 is the average of the entire period (1992-2002), and the variance 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.

Source: CPS Outgoing Rotation Groups

tween component. The rise in the covariance component suggest that the effect is actually driven by the asymmetric trends in occupations, which I will now discuss.

5.2 Asymmetric Trends in Occupational Variances

The increase in correlation between wage levels and inequality that is driving wage polarization, could be driven by different explanations. It's unclear whether this rise is because of trends in the wage levels, or wage inequality in occupations, or maybe occupational compositions. I will show that the main driver is the drop in inequality in at low-paying routine occupations. This is another support for the RBTC explanation for wage polarization.

During the 1990s the change in inequality within occupations was strongly correlated to the occupation wage level. High paying occupations saw an increase in inequality, and low paying occupations saw a decrease. Figure 9 shows this by plotting the change in variance of log wages from the beginning of the period (1992/3) to its end (2001/2) as a function of expected log wages. Changes in inequality, measured with the variance of log

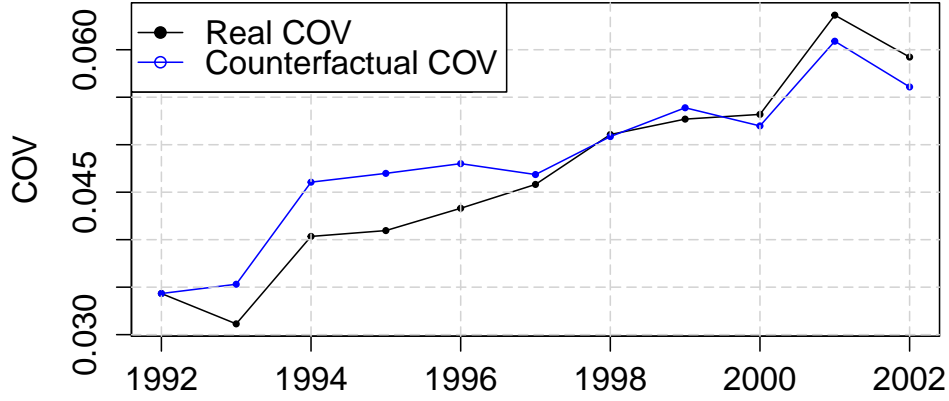


Figure 10: Covariance of Expectation and Variance of Log Wages at Different Occupations
 Covariance of mean log wage and variance of log wage by occupation: $COV(E[\log w|occ], V(\log w|occ))$. Wages at the top and bottom 5% were dropped (see Section 4). Counterfactual covariance is calculated by fixing $E[\log w|occ]$, and the share of workers in each occupation to their average throughout the period (Equation 5.)
 Source: CPS Outgoing Rotation Groups

wages, are clearly correlated with the initial wage levels. This fact was also documented before by Lemieux (2007).

In fact, the trends in with-occupation inequality can explain the full rise in the covariance component. I use counterfactual partial-equilibrium wage distribution to show that. The covariance is calculated by

$$COV(E[Y|X], V[Y|X]) = \sum_x \Pr(X = x) E[Y|X = x] V(Y|X = x)$$

Most of the increase stems from changes in the variance of log wages at different occupations ($V(Y|X = x)$). To show that, I will fix the share of workers and the expected log-wage in each occupation to their averages throughout the period. Thus, I will allow only the variance to vary between years. Formally, I will calculate a counterfactual partial-equilibrium covariance

$$COV(E[\widetilde{Y|X}], V[Y|X]) = \sum_x \Pr(X = x) \overline{E[Y|X = x]} V(Y_t|X_t = x) \quad (5)$$

for t between 1992 and 2002.

I find that the asymmetric trends in variance can explain the entire increase in covariance. Figure 10 compares the real value of the covariance to its counterfactual value from

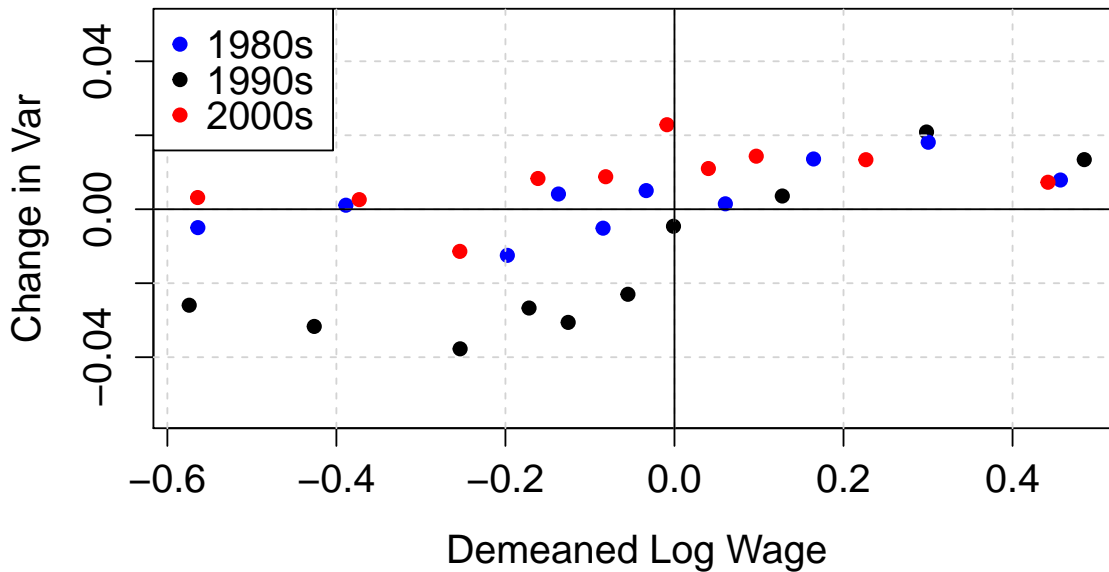


Figure 11: Change in $V[\ln w|occ]$ by $E[\ln w|occ]$ - Binned Scatter Plot

In each decade I bin occupations into 10 equal size bins of occupations, weighted by occupation size. Each point calculate the mean log wage at base year, and change in Variance from the start to the end of decade. In each decade, I de-mean the distribution of log wages, to have mean of zero. Each period includes the time period without changing to occupational coding. Hence, the three periods are 1979-1991, 1992-2002,2003-2012.

Source: CPS Outgoing Rotation Groups

equation 5. The counterfactual trend closely follows the real trend. This means that 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 wage polarization. Letting the share of workers, or the expected log wage to vary while other factors are fixed does not yield any similar results. From this we infer that indeed the increase in covariance, and thus the increase in wage polarization is mostly the result of the asymmetric changes in occupational variances.

Wage polarization in the 1990s is driven by the drop in inequality at lower paid occupations. Figure 11 shows the trend in occupational inequality by decade. For each decade, it plots a bin scatter of the changes in variance of log wages, by occupation mean log wage decile. The increase in inequality at high paying occupations is a long-standing trend. However, the 1990s are unique for their drop in inequality at low-paying occupations. This is why inequality is dropping at the bottom of the distribution only at the 1990s, generating wage polarization instead of an increase in wage inequality as in other decades.

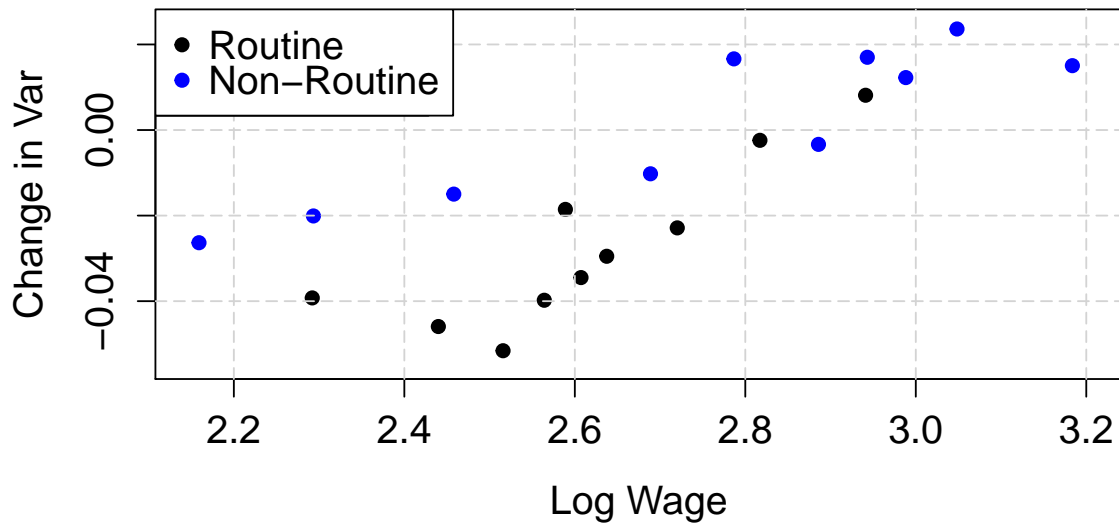


Figure 12: Change in $V[\ln w|occ]$ by $E[\ln w|occ]$ 1992-2002

I bin occupations separately for routine and non-routine occupations into 10 equal size bins (deciles) of occupations, weighted by occupation size. Each point calculate the mean log wage at base year, and change in Variance from the start to the end of decade 1992-2002.

Source: CPS Outgoing Rotation Groups

The drop in inequality at low paying occupations is driven mostly by routine heavy occupations. Figure 12 plots the changes in variance for routine and non-routine occupations. I bin occupations by their initial income decile in 1992 separately for routine and non-routine occupations and plot the mean change in the variance of log wages between 1992-2002. While there's some drop in inequality at low paying occupations that are non-routine, the trend is definitely stronger for routine occupations. This is in accordance with the findings in [Firpo et al. \(2013\)](#) who find using the O*NET data that routine occupations tend to have a stronger decrease in variance.

The most significant relative decline in wages, happened at the top of the income distribution in routine occupations. Figure 13 plots the (demeaned) wage trend in each wage decile, separately for workers in routine and non-routine occupations. Throughout the entire distribution, wages increased in non-routine occupations above the average increase in the labor market (that is pinned to 0). Wages relatively dropped in routine occupations, but not in an equal manner. The drop is much more significant at the top of the distribution. This generates a drop in inequality at routine occupations.

This decline could be driven by either a change in sorting into occupations, or real decrease in wages for higher skilled workers. In the next section I'll use panel data to

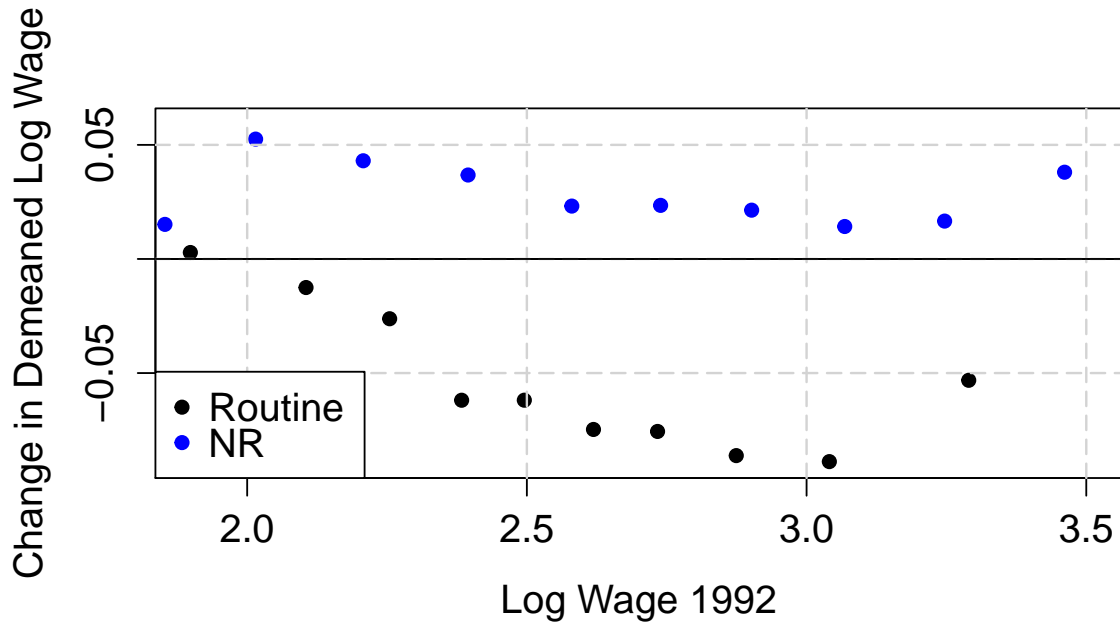


Figure 13: Change in Relative Wages 1992-2002, Routine and Non-Routine Occupations
 I bin workers separately for routine and non-routine occupations into 10 equal size bins (deciles) . Each point calculate the mean log wage at 1992, and change in log mean wage 1992-2002, for this decile.
 Source: CPS Outgoing Rotation Groups

decide between these two explanations.

6 Decrease in Return to Skill - Evidence from Panel Data

The decline in top wages in routine occupations could be driven by different explanations, which can only be distinguished using panel data. It is well established that during the years of wage polarization, there is also job polarization - a significant drop in employment at routine occupations. If job polarization is changing the composition of routine occupations to have less workers from the top and the bottom of the skill distribution (as argued by Cortes, 2016), we will get exactly this trend of wage compression in routine occupations. In the theoretical framework, this is the case where $\sigma = 1$. It is also possible that demand decline is more significant for higher-skill workers, generating wage compression even for workers who don't switch occupations. This is the case when new technology is substitutional to skill ($\sigma > 1$). Using the PSID data I will show that the evidence support the latter explanation.

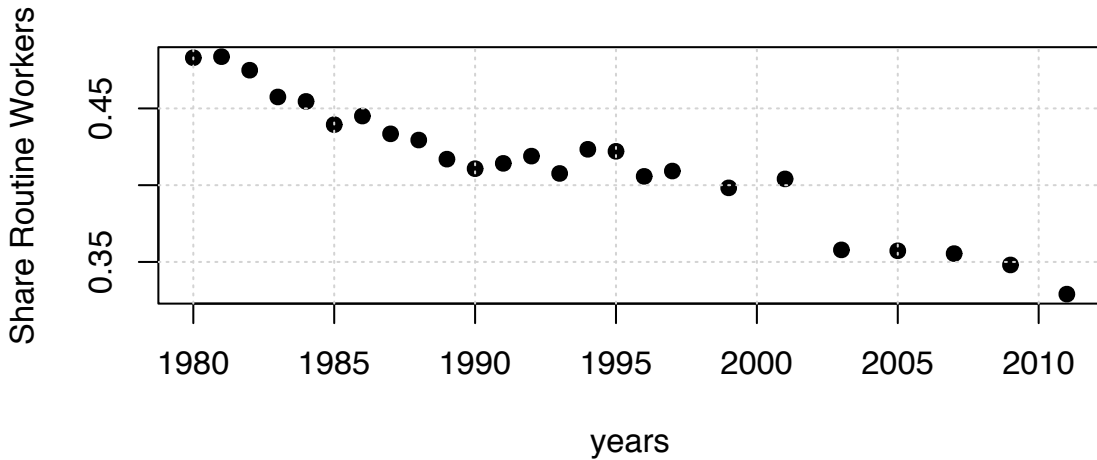


Figure 14: Share of Routine Workers - PSID

Share of routine workers from overall sample. Routine workers are defined as workers in administrative, production or operator occupations, classified by the first occupational coding digit.
Source: PSID

6.1 Decline in Routine Occupations

I'll start by showing directly that both employment and wage premiums dropped in routine occupations, repeating some of the analysis of [Cortes \(2016\)](#).

The decline in employment is shown in [Figure 14](#). The PSID uses the same occupational coding since 1980, allowing us to compare the share of workers in routine occupations consistently across various years. While in 1980, the share of routine workers in the PSID sample was close to half of all workers it dropped to about a third by 2011. The trend seems long and steady throughout the entire period, and not particularly stronger in any period of time.

I use a fixed effect model to study the trend in occupational wage premiums. I estimate the following standard fixed effect model

$$\log w_{ijt} = \beta X_{it} + \lambda_{jt} + \theta_i + \varepsilon_{ijt} \quad (6)$$

where X_{it} includes a quadratic in experience, λ_{jt} is occupation by year fixed effect for three occupations: manual, routine and abstract and θ_i is individual fixed effect.

I find the premium for routine occupations has decline steadily. [Figure 15](#) shows the wage premium of routine workers compared to manual and abstract occupations ($\lambda_R - \lambda_M, \lambda_R - \lambda_A$). The premium for routine occupations seems to decline compared to both

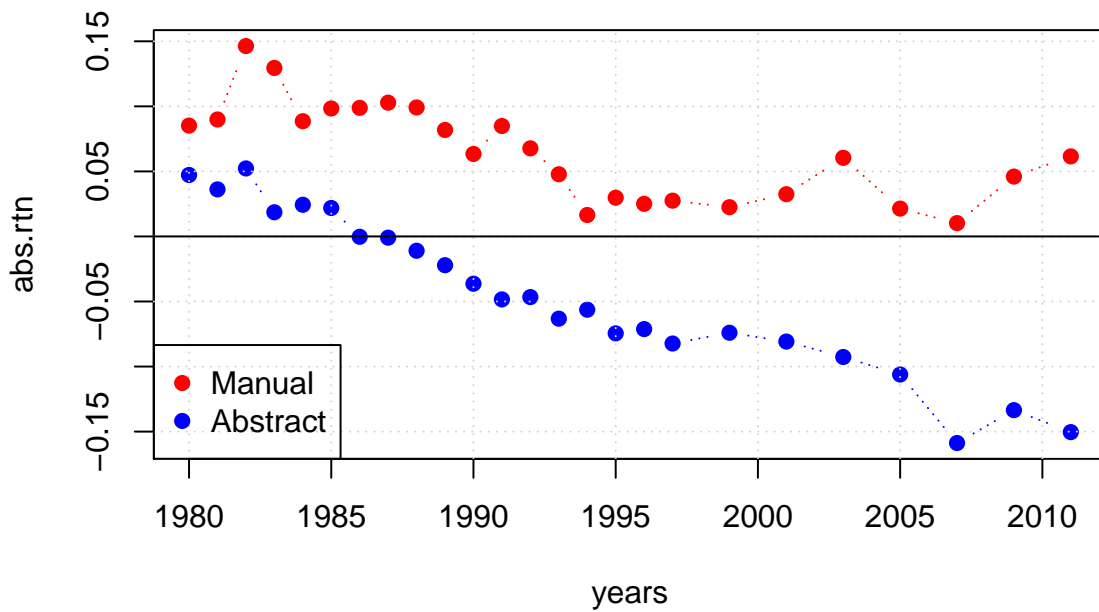


Figure 15: Wage Premium for Routine Occupation Compared to Abstract/Manual

Wage premiums are estimated using a fixed effect model (Equation 6). Routine workers are defined as workers in administrative, production or operator occupations, classified by the first occupational coding digit. Similarly abstract includes all managerial, professional and technician occupations. Manual includes service, sales and agriculture. Taking routine workers as the reference category, I plot minus one times the coefficient for manual and abstract by year.

Source: PSID

other alternatives.

By construction, this setting assumes occupational wage premiums are the same for all skill levels. To test that, we would need a more general model, that I will use in the next section.

6.2 Interactive Fixed Effect Model

To test if returns to skills are changing over time, we need to go beyond the standard fixed effect model. In a fixed effect model, the log wage ratio of two workers with a skill level of θ_0, θ_1 will always remain $\theta_0 - \theta_1$, as long as they are in the same occupation. Wage compression within occupation is therefore impossible in this kind of model. Therefore, we would need a more general setting.

I will estimate an interactive fixed effect model, where returns to skill could change. Specifically, I will estimate the following equation

$$\log w_{ijt} = \beta_{ijt}X_{it} + \lambda_{jt} + \alpha_{jt}\theta_i + \varepsilon_{ijt} \quad (7)$$

The only difference from Equation 6 is that the individual fixed effects θ_i are interacted with an occupation time-varying coefficient α_{jt} . This α parameter will be the focus of this analysis, as it measures the returns to skills in each occupation at each year. I will use either three categories of occupation, or 1-digit. I will estimate the model using a least square estimator (Bai, 2009).

Alternative ways to estimate this model yield very similar results. Least square method is consistent when the number of observations per individual is large enough. Since the estimates for $\hat{\theta}_i$ is estimated with noise, it could be correlated with ε_{ijt} . While this correlation is asymptotically zero, it's unclear whether the number of periods used is sufficiently large. Therefore, I also estimate the model using an alternative approach.

I instrument for θ_i with worker's years of education. Generally, other approaches must require an additional source of information, such as an instrument.¹² Here, instead of assuming $\hat{\theta}_i$ and ε_{ijt} are uncorrelated, I assume that years of education are uncorrelated with ε_{ijt} . This assumes that years of education is only affecting log wages through its effect on θ_i .¹³

¹²For instance Holtz-Eakin et al. (1988) use lagged variables as instruments.

¹³Put differently, this approach assumes that trends in returns to education are identical to trends in returns to skill that are not captured with years of schooling. This is implied by the model that assumes only a one-dimensional skill.

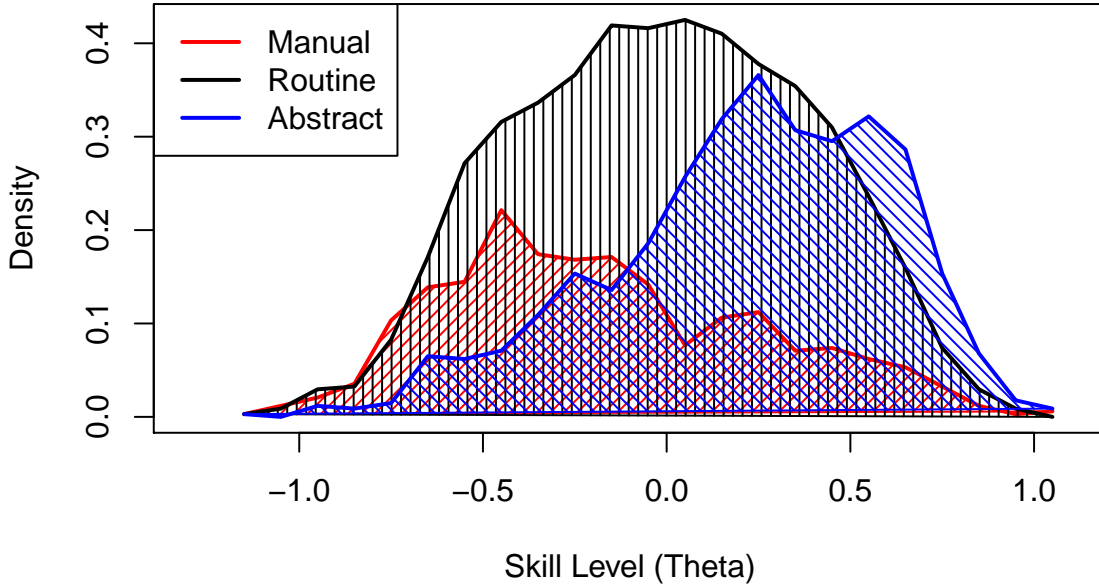


Figure 16: Estimated PDF of Skill (θ_i) by Occupation Category in 1981

Histogram for $\hat{\theta}_i$ from an interacted fixed effect model (Equation 7), by three occupation categories. Routine workers are defined as workers in administrative, production or operator occupations, classified by the first occupational coding digit. Similarly abstract includes all managerial, professional and technician occupations. Manual includes service, sales and agriculture. Source: PSID

6.3 Returns to Skill by Occupation

The estimation of Equation 7 supports the case of $\sigma > 1$ that is described in the model. That is, technological progress is substitutional to skills in routine occupations.

Supporting the assumption of the model, I find evidence that middle-skill workers have a comparative advantage in routine-intense occupations. Figure 16 plots the estimated density of θ_i for each one of the three occupational categories in 1981. At this point, routine occupations employ a large share of workers. On average, workers in manual occupations are the less skilled, workers in abstract occupations are the most skilled, and workers in routine occupations have an average level of skill. In the context of the model, since workers sort into occupations based on comparative advantage, this implies that returns to skill are highest at abstract occupation, then in routine, and smallest in manual. However, because manual occupations include a smaller share of all workers, the majority low skilled workers are still in routine occupations.

I test that using a one-dimensional skill is a reasonable assumption for this context. I do this by allowing θ_i to vary by j . This allows for a different skill to be used in each

	Abstract	Routine	Manual
Abstract	1		
Routine	.74	1	
Manual	.71	.66	1

Table 1: Correlation of Occupational Skills

Pearson correlation coefficient between $\widehat{\theta}_{ij_0}, \widehat{\theta}_{ij_1}$ for pairs of occupation categories. $\widehat{\theta}_{ij}$ are estimated using Equation 7, 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. Similarly abstract includes all managerial, professional and technician occupations. Manual includes service, sales and agriculture. Each correlation is calculated using all workers who worked at both categories.
Source: PSID

occupation, as in a Roy model. I find that the correlation between θ_{ij} for a given value of i are between .66-.74, as shown in Table 1. Since we can only estimate the correlation for workers that chose to switch occupations, they might be upward biased. But more than half (51%) of workers do switch occupations and so most of the sample is used. Moreover, θ_{ij} is measured with a high level of noise, which biases the results downward. The high level of correlation suggest that we're not losing much precision by allowing for only one skill.

The key finding from estimating the interactive fixed effect model is that there is a decline in returns to skill in routine occupations. The return to skill in routine occupation is measured with the parameter α_{Rt} in Equation 7. I plot its estimated value for each year in Figure 17. Since there is one degree of freedom in this equation, I pin $\alpha_{R,1980}$ to 1. At the beginning of the period, during the 1980s return to skill actually increase quite significantly. Then, exactly in the years of wage polarization and the decrease in inequality in routine occupations, there is a large drop in return to skill. This shows that during that time, wages compressed in routine occupations, even without any changes in their composition.

This pattern doesn't repeat itself in every occupation. I estimate Equation 7 allowing the return to skills (α_{jt}) to vary by 1-digit occupation and year. Figure 18 plots the coefficient for α_{jt} for each 1-digit occupation. Almost all occupations experienced an increase in return to skill during the 1980s. This fits the existing literature on the rise of inequality in the 1980, that argues that returns to education is sharply increasing in that period (Katz and Murphy, 1992).

Starting from the 1990s, the patterns are different by occupation. Administrative workers and Operators and to some extent also Production workers, the three occupations classified as routine, have a significant drop in return to skill. In contrast, in service

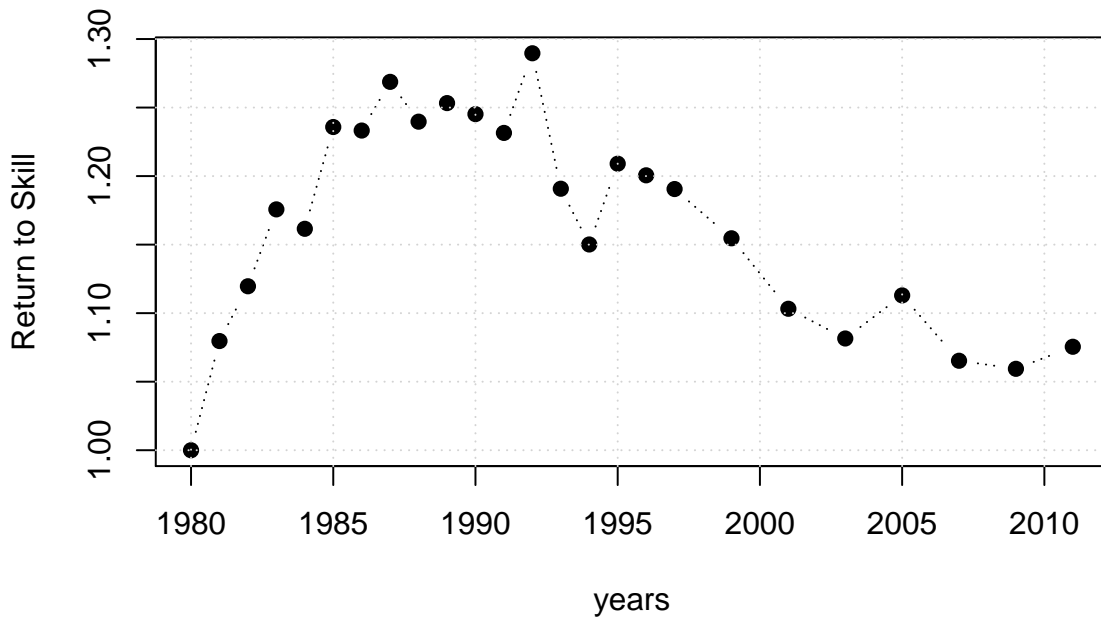
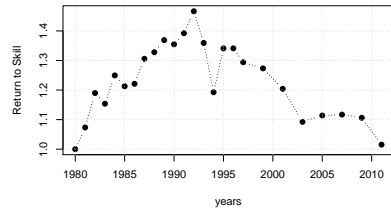
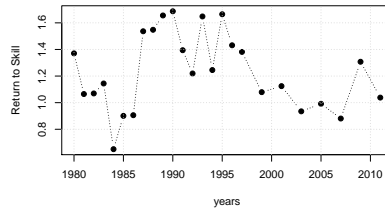


Figure 17: Returns to Skill (α_{Rt}) in Routine Occupations

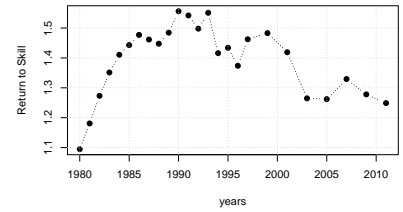
Return to skill are calculated in an interactive fixed effect model (α_{Rt} , using Equation 7). Routine workers are defined as workers in administrative, production or operator occupations, classified by the first occupational coding digit.
 Source: PSID



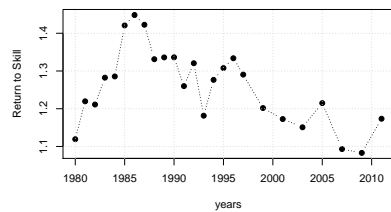
(a) Administrative



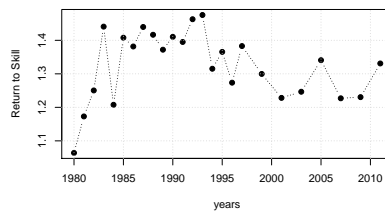
(b) Agriculture



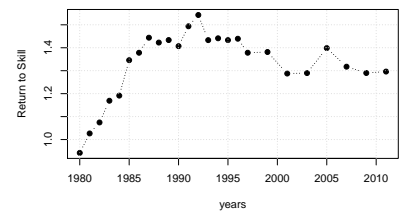
(c) Managerial



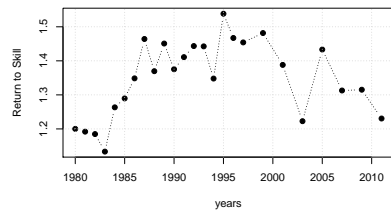
(d) Operators



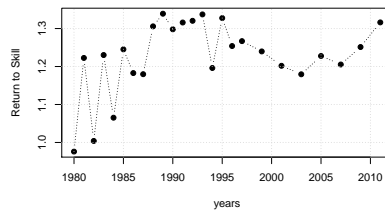
(e) Production



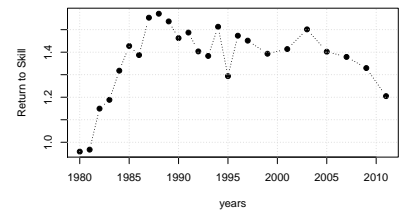
(f) Professional



(g) Sales



(h) Service



(i) Technicians

Figure 18: Return to Skill (α_{jt}) by 1-digit Occupation

Return to skill are calculated in an interactive fixed effect model (α_{jt} , using Equation 7). α_{jt} varies by 1-digit occupation and year. Source: PSID

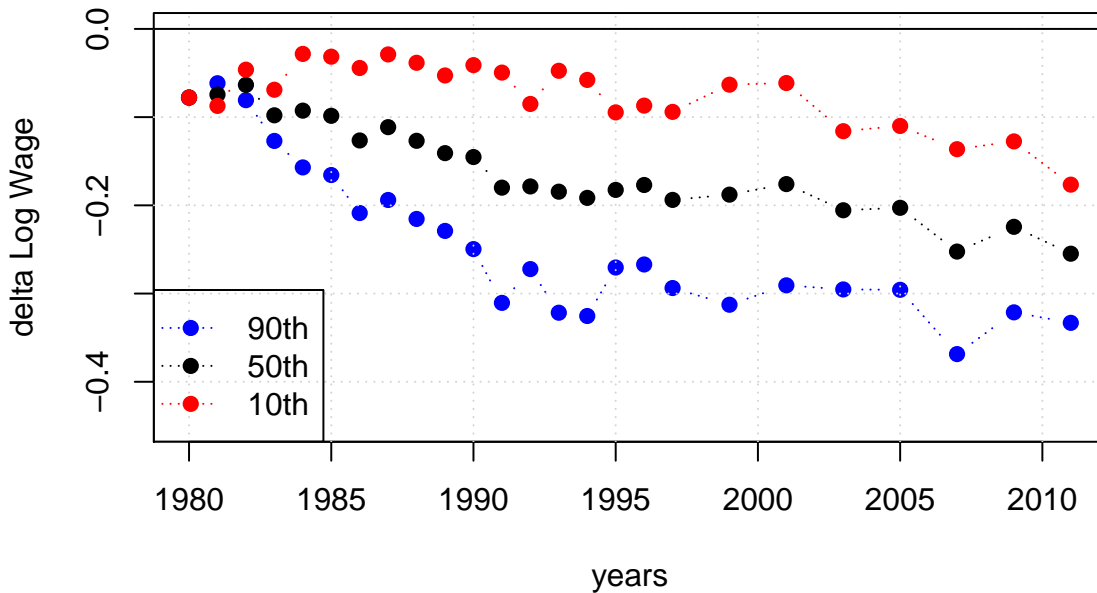


Figure 19: Occupation Premium for Routine vs Abstract by Skill Percentile

Difference in log predicted wage for workers in routine versus abstract occupation for three percentiles at the distribution of θ_i . θ_i are defined net of age and cohort. Routine workers are defined as workers in administrative, production or operator occupations, classified by the first occupational coding digit. Similarly abstract includes all managerial, professional and technician occupations. Source: PSID

occupations, return to skill have not shown any decline. This category includes most of the manual workers. During the 2000s it had a return value of 1.3, compared to 1.1-1.2 in administrative and operator occupations. This could potentially explain why workers from routine occupations decided to join service occupation (Autor and Dorn, 2013).

Abstract occupations vary in their trends in return to skills. For professional workers, it's hard to see any decline in the return to skill, which is among the highest in any occupation. But managerial and technicians do seem to have some decline. This decline seems to be most prominent in the 2000s, which fits the theory of Beaudry et al. (2016) for a reverse in the demand trend for skilled labor. This requires further investigation that is beyond the scope of this paper.

As a result, decrease in the wage premium for routine occupations is sharper for higher skilled workers. In Figure 15 I followed previous literature that showed the decrease in routine premium for the average routine worker. But this masks significant heterogeneity by skill level. To account for those I plot the change in wage premiums accounting for the different levels in α_{jt} , for skill levels at the 10th, 50th and 90th percentile.

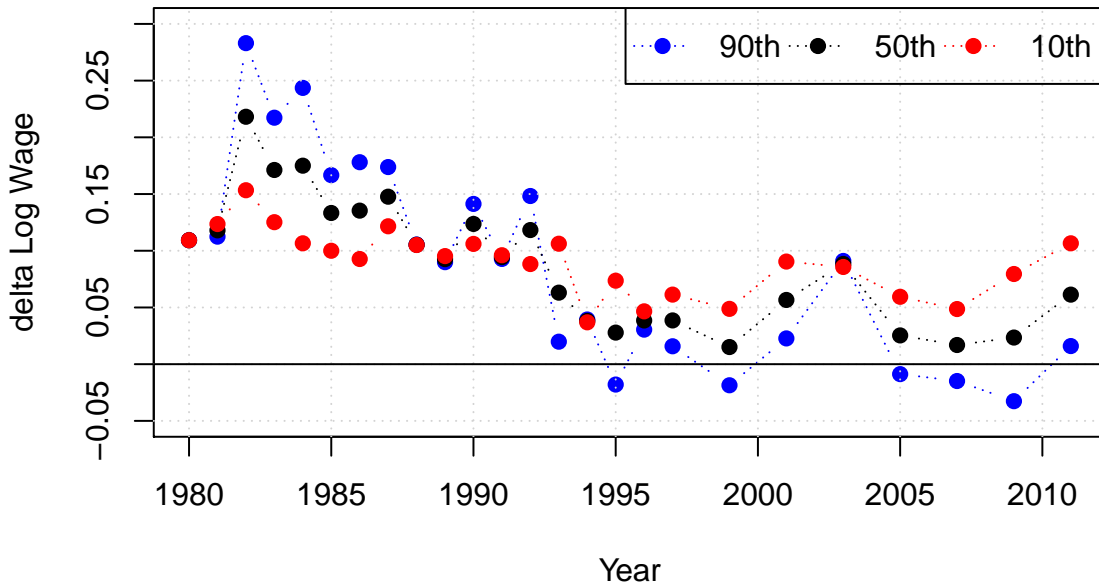


Figure 20: Occupation Premium for Routine vs Abstract by Skill Percentile

Difference in log predicted wage for workers in routine versus manual occupation for three percentiles at the distribution of θ_i . θ_i are defined net of age and cohort. Routine workers are defined as workers in administrative, production or operator occupations, classified by the first occupational coding digit. Similarly manual includes all service, sales and agriculture workers. Source: PSID

Figure 19 shows the results comparing routine to abstract occupations. There is a significant decline in all parts of the skill distribution. This can explain why we see a large growth in the share of workers in abstract occupations. However, the decline is sharper for higher skilled workers. This implies that comparative advantage for skilled workers in abstract occupations became more significant.

The comparison of routine to manual occupations is even more striking. Figure 20 plots the result. During the 1980s, the premium for moving to routine occupations from manual occupations was positively correlated with skills. This generated a comparative advantage of higher skilled workers in routine occupations, compared to manual occupations. But during the 1990s the direction of the correlation flips. In the 2000s, the premium is higher for lower skilled workers, and is in fact even negative in some years for the higher skill ones. This implies that comparative advantage flipped during this period, making manual occupations more suitable for higher-skilled workers compared to routine. Overall, the decline in premium is close to zero for low skilled workers. This suggest that we can expect most of the drop in employment to be from higher skilled

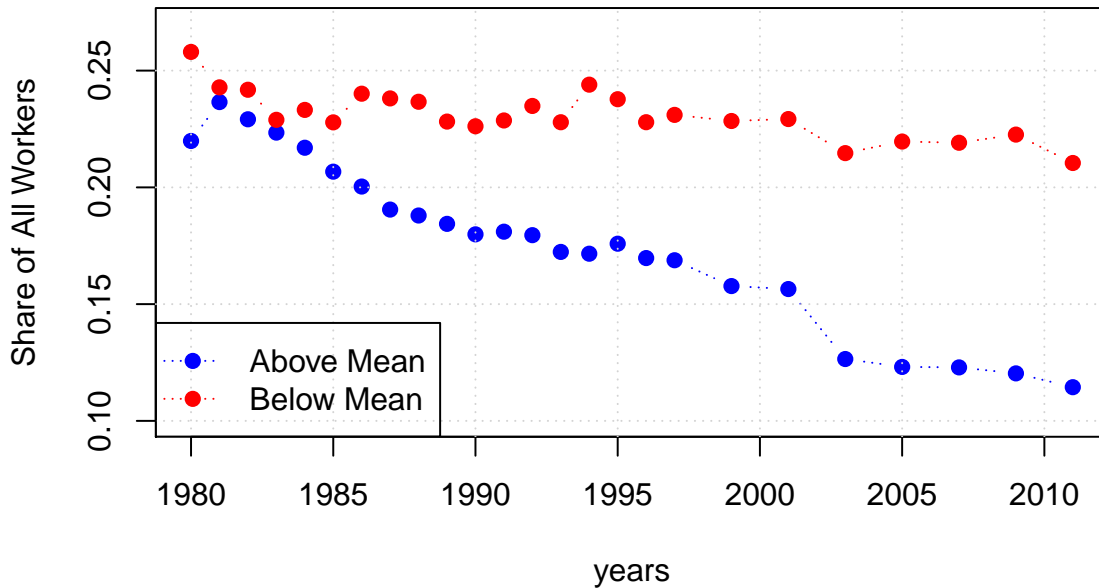


Figure 21: Share of Routine Workers by Skill

Share of workers in routine occupations binned by level of θ_i . $\hat{\theta}_i > 0$ is above mean, and otherwise below. θ_i are defined net of age and cohort. Routine workers are defined as workers in administrative, production or operator occupations, classified by the first occupational coding digit. Share is taken from all workers in sample. Source: PSID

workers.

Appendix Figures 27 repeats these results using years of education as instruments for skills. As discussed in the previous section, a least square approach would be inconsistent if the number of periods is not large enough. This is solved when instrumenting for θ_i with years of education. I find that the results are fairly similar, and the same conclusions hold.

As predicted by the model and by the wage trends I estimate, job polarization is driven almost entirely from higher skilled workers. Figure 21 plots the decline in the share of routine workers separately for workers above and below the mean skill level. While in the early 1980s there has been an equal share of above and below mean workers in routine occupations, this is far from the case later. There has been a steady decline in the share of above mean workers in routine occupations, cutting their share in more than one half during this period. At the same time, the share of workers below mean in routine occupations stays fairly stable around 23% throughout most of the period. As a result, routine occupations gradually became low skill occupations, and are no longer middle-

skill.

7 Discussion

This paper uses novel methods to address the main puzzles regarding the effect of RBTC on wages. The empirical findings closely align with the predictions of the model I outlined for such technological change. Together they explain why at first, in the 1990s wages relatively drop in the middle of the distribution, why later in the 2000s wages relatively drop at the bottom of the distribution, and why this was not captured by standard decomposition methods.

I show that the decrease in middle wages in the 1990s is the result of a technology that is replacing the usage of skills in routine occupations. I first use skewness decomposition to show that almost the entire trend of wage polarization is driven by occupational trends, which are impossible to capture with most other decomposition methods. I find that the 1990s are unique for their decrease in inequality within routine occupations. Even though most routine workers are earning below median wages, most of the relative decline in wages is around the middle, because of the wage compression in routine occupations.

Using an interactive fixed effect model I show that returns to skill declined in routine occupations. I use panel data to show that wages compressed in routine occupations even regardless of any employment trends. In the context of the model, this fits only the case where technology is substitutional to skills ($\sigma > 1$). Technology blurs the skill differences between workers, making all workers more equally productive. This generates the largest drop in demand for skilled workers in routine occupations. Supporting this notion, I find that job polarization is driven almost entirely by the decline in employment of higher-skill workers in routine occupations.

The model, and the empirical finding also provide an answer for why wage polarization stopped. Since the return to skills declined in routine occupations, middle skilled workers no longer have comparative advantage in them. The panel results show that gradually, routine occupations became low-skilled. As a result, starting approximately from early 2000s technological improvement in routine occupations do not affect middle wage workers anymore, since they no longer work there.

Other explanations do not fit these empirical patterns as well. Institutional changes, such as an increase in minimum wage could potentially generate a decline in middle wages by increasing lower wages. Decline in unions could also affect middle-wage workers more (Lemieux, 2007). High growth rates and low unemployment could also poten-

tially boost lower wages, and so generate a relative decline in middle wages. However, neither of these explanations is expected to work through occupations in particular, more than through education levels or industries. The effect on occupations also should not be different for routine occupations, than it is on manual occupations such as service jobs.

An increase in demand for service occupations seems to be more of an outcome than a cause to wage polarization. A demand increase in service occupations should have attracted more workers from the bottom of the skill-distribution. However, I find that job-polarization is mostly the result of employment drop for above average skill workers.

The decrease in return to skill also fits the findings in research that study 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 is generating wage compression within routine workers in a given firm. In this paper I quantify that this wage compression is actually the main driver of wage polarization.

Looking forward, if RBTC continues, its negative effect on low-income workers should only increase. As routine occupations have a growing share of low-skilled workers, any further shift towards capital in performing their main tasks will be more similar to a classic skilled-biased technological change. However, new technologies may start affecting the returns to skill in manual and abstract occupations. This will generate different wage trends beyond the scope of the theoretical framework I discussed here.

The results of this paper highlight the need for additional research on the micro-foundations of the decrease in return to skill in routine occupations. In the model, I used a general macro-level production function to explain how new technology replaces skill. This production function could be a result of a technology that is replacing the usage of human memory, or performs complicated calculations, making these skills less valuable. Hence, leaving workers with very basic simple tasks that don't depend on skill as they don't leave room for mistakes.

An alternative explanation is that the price of switching to the new technology is only justified when it saves the costs of the more skilled workers. In this case, the technology could in theory replace every worker. But its price will only be justified if workers are earning above its cost, which is why it's replacing mostly the high earners. Using matched employer-employee data to test if employers that are paying higher wages are adopting more technology could be an interesting exercise. However such data exist mostly in European countries that do not always experience the same wage trends ([Massari et al., 2013](#)).

References

- Acemoglu, D., Autor, D., 2011. Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings. Vol. 4, Part B of Handbook of Labor Economics. Elsevier, pp. 1043–1171.
URL <http://www.sciencedirect.com/science/article/pii/S0169721811024105>
- Autor, D., Dorn, D., 2013. The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review* 103 (5), 1553–1597.
URL <http://www.aeaweb.org/articles.php?doi=10.1257/aer.103.5.1553>
- Autor, D., Katz, L. F., Kearney, M. S., 2006. The Polarization of the US Labor Market. *The American Economic Review* 96 (2), 189–194.
- Autor, D., Katz, L. F., Kearney, M. S., Apr. 2008. Trends in U.S. Wage Inequality: Revising the Revisionists. *Review of Economics and Statistics* 90 (2), 300–323.
URL <http://dx.doi.org/10.1162/rest.90.2.300>
- Bai, J., 2009. Panel data models with interactive fixed effects. *Econometrica* 77 (4), 1229–1279.
- Beaudry, P., Green, D. A., Sand, B. M., 2016. The great reversal in the demand for skill and cognitive tasks. *Journal of Labor Economics* 34 (S1), S199–S247.
- Cornfeld, O., Danieli, O., 2015. The origins of income inequality in israel-trends and policy. *Israel Economic Review* 12 (2), 51–95.
- Cortes, G. M., 2016. Where have the middle-wage workers gone? a study of polarization using panel data. *Journal of Labor Economics* 34 (1), 63–105.
- DiNardo, J., Fortin, N. M., Lemieux, T., 1996. Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica: Journal of the Econometric Society*, 1001–1044.
- Firpo, S., Fortin, N. M., Lemieux, T., 2013. Occupational Tasks and Changes in the Wage Structure (February).
URL http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1778886

- Firpo, S., Fortin, N. M., Lemieux, T., *Econometrica*, S., May, N., Firpo, B. S., Fortin, N. M., Lemieux, T., 2009. Unconditional Quantile Regressions. *Econometrica* 77 (3), 953–973.
- Foote, C., Ryan, R., 2015. Labor-market polarization over the business cycle. NBER Macroeconomics Annual.
URL http://www.richmondfed.org/conferences_and_events/research/2012/pdf/labor_markets_conf_2012_foote_ryan.pdf
- Fortin, N., Lemieux, T., Firpo, S., 2011. Decomposition methods in economics. *Handbook of labor economics* 4, 1–102.
- Gaggl, P., Wright, G. C., 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, M., Manning, A., Salomons, A., Aug. 2014. Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review* 104 (8), 2509–2526.
URL <http://pubs.aeaweb.org/doi/abs/10.1257/aer.104.8.2509>
- Holtz-Eakin, D., Newey, W., Rosen, H. S., 1988. Estimating vector autoregressions with panel data. *Econometrica: Journal of the Econometric Society*, 1371–1395.
- Juhn, C., Murphy, K. M., Pierce, B., 1993. Wage inequality and the rise in returns to skill. *Journal of political Economy*, 410–442.
- Jung, J., Mercenier, J., 2014. Routinization-biased technical change and globalization: Understanding labor market polarization. *Economic Inquiry* 52 (4), 1446–1465.
- Katz, L., Autor, D., 1999. Changes in the wage structure and earnings inequality. *Handbook of labor economics* 3, 1463–1555.
- Katz, L., Murphy, K., 1992. Changes in relative wages, 1963-1987: Supply and demand factors. *Quarterly Journal of Economics* 107 (1), 35–78.
URL <http://www.nber.org/papers/w3927>
- Lemieux, T., Jun. 2006. Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill? *The American Economic Review* 96 (3), 461–498.
URL <http://www.jstor.org/stable/30034058>
- Lemieux, T., Sep. 2007. The changing nature of wage inequality. *Journal of Population Economics* 21 (1), 21–48.
URL <http://link.springer.com/10.1007/s00148-007-0169-0>

- Lemieux, T., 2010. What Do We Really Know About Changes in Wage Inequality? In: Labor in the New Economy. National Bureau of Economic Research, Inc, pp. 17–59.
URL <http://econpapers.repec.org/RePEc:nbr:nberch:10812>
- Massari, R., Naticchioni, P., Ragusa, G., et al., 2013. Unconditional and conditional wage polarization in europe. *Comunicazioni Spontanee*, 2.
- Mincer, J., 1974. Schooling, Experience, and Earnings. *Human Behavior & Social Institutions* No. 2.
- Piketty, T., 2014. Capital in the 21st Century.
URL <http://piketty.pse.ens.fr/files/Piketty2014Capital21cUS.pdf>
- Roy, A., 1951. Some thoughts on the distribution of earnings. *Oxford economic papers* 3 (2), 135–146.
URL <http://www.jstor.org/stable/2662082>
- Yitzhaki, S., Schechtman, E., 2013. The Gini Methodology - A Primer on a Statistical Methodology.
URL <http://link.springer.com/content/pdf/10.1007/b98971.pdf>

A Decomposing By More Categories

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

$$Y = \sum_i X_i$$

Using simple algebra

$$\mu_3(Y) = \sum_i \mu_3(X_i) + \sum_i \sum_{j \neq i} COV(X_i^2, X_j) + \sum_i \sum_{j \neq i} \sum_{k \neq i, j} E[X_i X_j X_k] \quad (8)$$

Therefore, we can decompose the skewness of Y into a linear combination of the skewness of its linear components, the covariance of their second and first moments, and the triple multiplication of all three distinguished components. I will call this decomposition using Equation 8 - linear skewness decomposition. Though this decomposition includes a lot of different terms, many of them has an expectation of zero.

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

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

and using Equation 8 yields Equation 4, using the law of iterated expectations over X .

Alternatively we can estimate any linear model, and use this formula. This is useful to compare occupations directly to other categories. I show this for the equation

$$\ln w_i = occ_i + ind_i + \varepsilon_i \quad (9)$$

where occ_i and ind_i are occupation and industry dummies. I then decompose the increase in skewness by Equation 8. Figure 25 presents the results. Most of the increase in skewness comes from the correlation of occ_i and ε_i^2 . That's the correlation of the part of occupation premium that is orthogonal to industry, and inequality within occupation and industry. The equivalent component for industries (in green) is negligible. All other components, such as the skewness between occupations or industries, the correlation of occupation premium and industry premium variance and others are aggregated and plotted in red. All together 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

$$\ln w_i = occ_i + \beta X_i + \varepsilon_i \quad (10)$$

Similar to the case with industries, the covariance for occupations and residuals is still capturing almost half of the increase. Note that this is the correlation with the occupation premium, for workers with the same level of skills.

However, as opposed to the case of industries, the correlation of skill group mean wage, and variance

$$COV(\beta X_i, \varepsilon_i^2)$$

also seem to matter. To further investigate that I decompose βX_i to mean occupation skill and within occupation skill difference

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

Decomposing by the four components

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

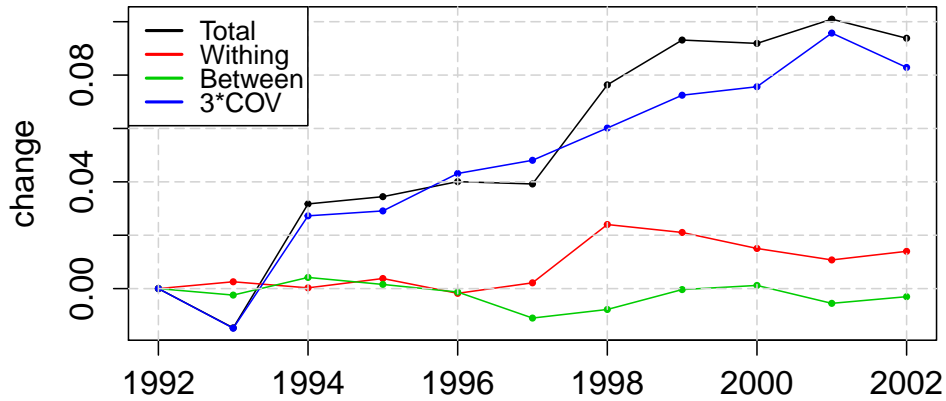


Figure 22: Skewness Decomposition by 3-Digit Occupation With Imputed Wages
 Changes since base year (1992). Wages at the top and bottom 5% were dropped (see Section 4). Includes wages that were imputed.
 Source: CPS Outgoing Rotation Groups

I find that all of the increase in $COV(\beta X_i, \varepsilon_i^2)$ comes from $COV(E[\beta X_i | occ_i], \varepsilon_i^2)$ as can be shown in Figure 26. At total, the main two components are the correlation of ε^2 with both occ_i and $E[\beta X_i | occ_i]$. This means that the correlation of variance with occupational wage levels is stemming from both, occupation premium (occ_i) and mean skill level at the occupation ($E[\beta X_i | occ_i]$). But similar to industries, categories that are unrelated to occupations are still negligible.

B Appendix Figures

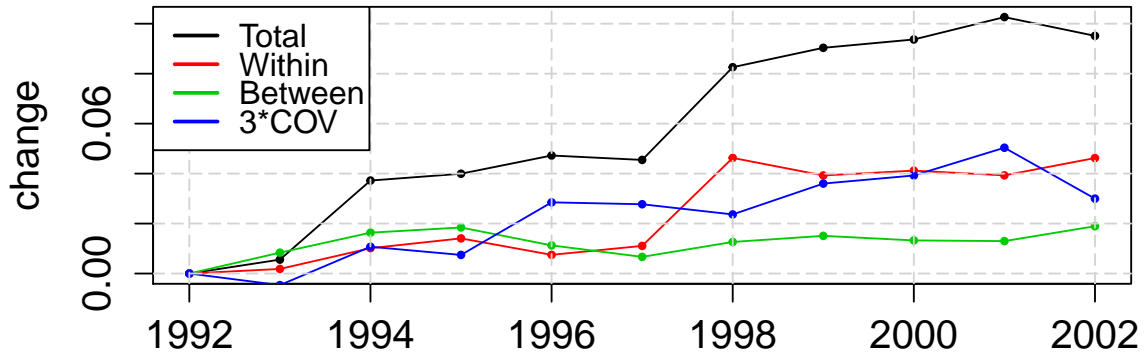


Figure 23: Skewness Decomposition by 3-Digit Industry

Changes since base year (1992). Wages at the top and bottom 5% were dropped (see Section 4).
 Source: CPS Outgoing Rotation Groups

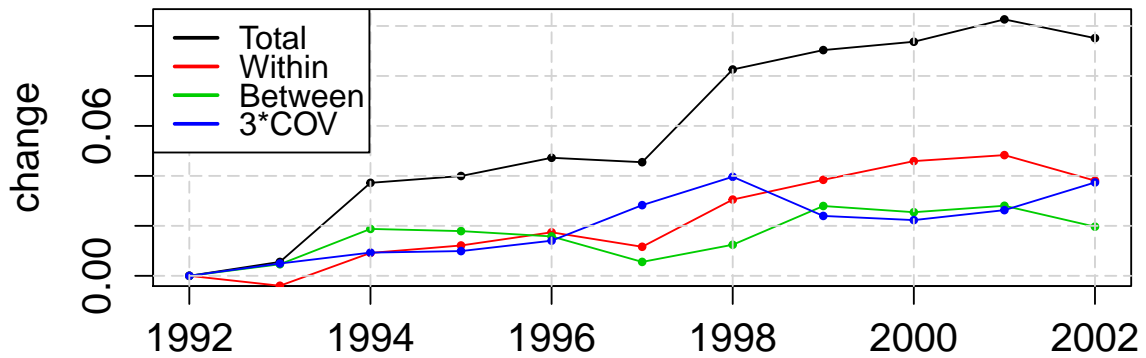


Figure 24: Skewness Decomposition by Education and Experience

Changes since base year (1992). Wages at the top and bottom 5% were dropped (see Section 4).
 Source: CPS Outgoing Rotation Groups

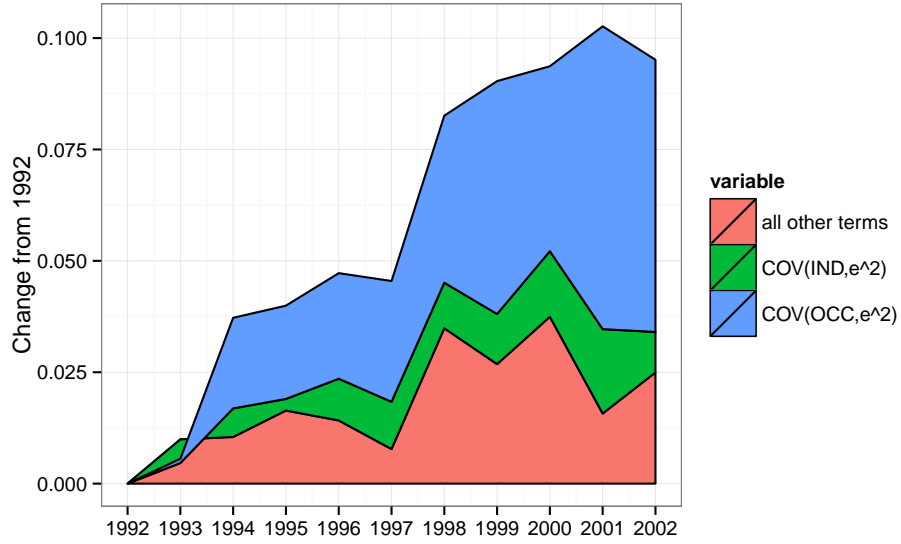


Figure 25: Skewness Decomposition by Occupation and Industry

Changes since base year (1992). Wages at the top and bottom 5% were dropped (see Section 4). Decomposing based on Equation 9. I plot only two components separately and the rest are aggregated (in red). $COV(occ, \varepsilon^2)$ and $COV(ind, \varepsilon_i^2)$ are the covariance of occupation and industry premium with the unexplained variance.

Source: CPS outgoing rotation groups.

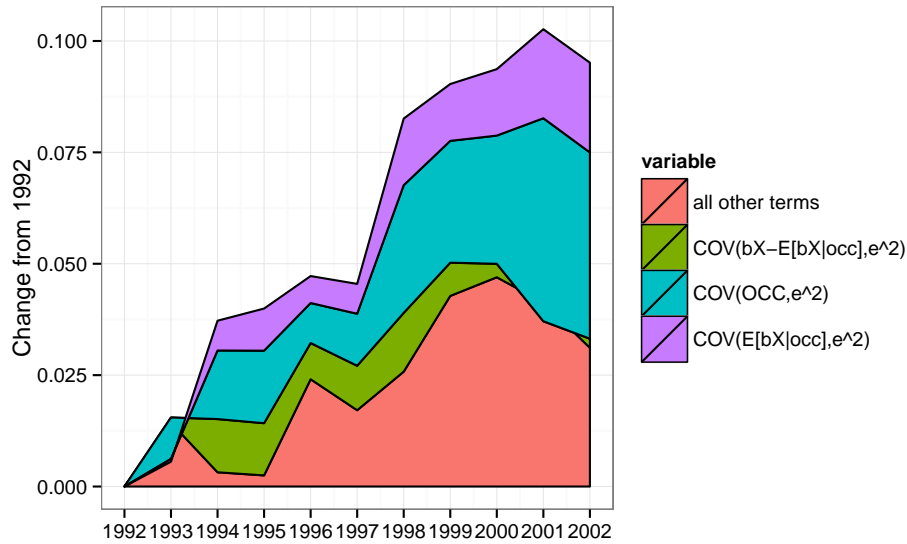
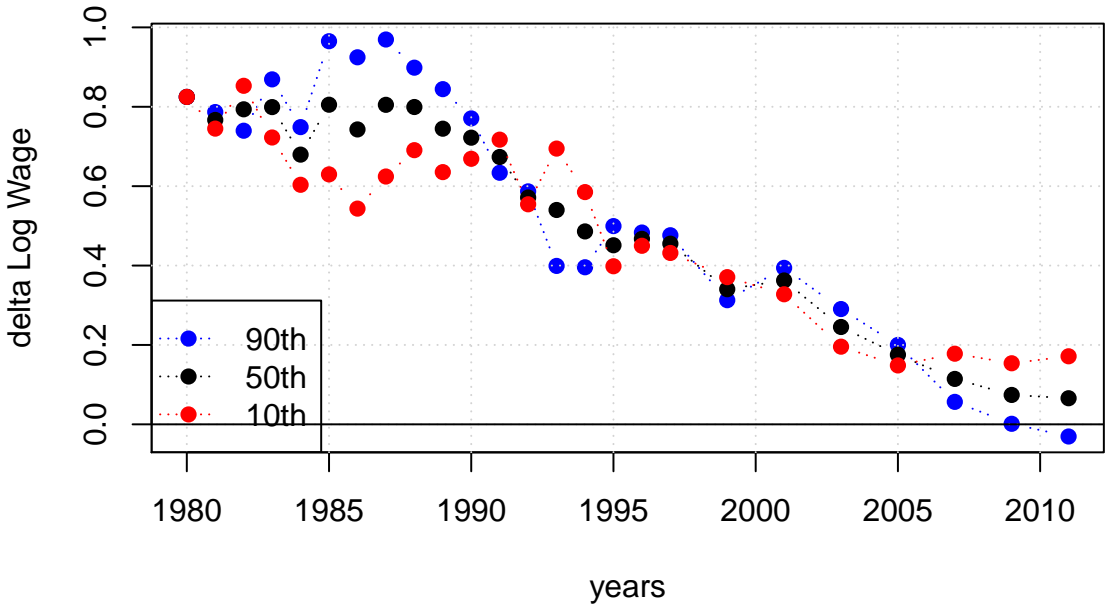


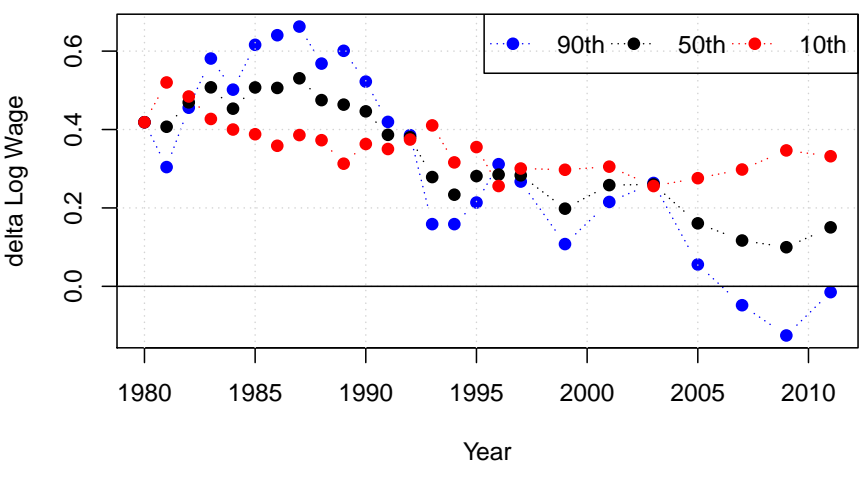
Figure 26: Skewness Decomposition by Occupation, School, Experience

Changes since base year (1992). Wages at the top and bottom 5% were dropped (see Section 4). Decomposing based on Equation 9. I plot only two components separately and the rest are aggregated (in red).

Source: CPS outgoing rotation groups.



(a) Routine/Abstract



(b) Routine/Manual

Figure 27: Occupation Premium by Skill Percentile with Instruments

Difference in log predicted wage for workers in routine versus abstract/manual occupation for three percentiles at the distribution of θ_i . Estimated with an interactive fixed effect model, using years of education as instruments. θ_i are defined net of age and cohort. Routine workers are defined as workers in administrative, production or operator occupations, classified by the first occupational coding digit. Similarly manual includes all service, sales and agriculture workers. Source: PSID