Polarization: A Supply-Side Mechanism*

Benedikt Dengler
Federal Ministry of Economic Affairs and Energy, Germany†

Vinzenz Ziesemer
Erasmus University Rotterdam

September 2020

Abstract

This paper provides empirical evidence that most jobs require both manual and cognitive skills. The latter are more important at the top of the earnings distribution, and are acquired through education. The paper embeds these findings into a tractable framework, connecting two disparate strands of the literature on earnings inequality: As in public economics, tax incentives to human capital accumulation drive the pre-tax earnings distribution. As in skill-biased technological change, relative prices of the skills are determined in general equilibrium. Analyzing the model, it finds that incentive changes in taxation like those that occurred in the second half of the 20$^{\text{th}}$ century can lead to polarization of the labor market. This supply- or policy-driven explanation is complementary to the demand- or technology-driven explanations found in the existing literature.

*Author contacts: Dengler: benedikt.dengler@eui.eu, Ziesemer: vinzenz.ziesemer@eui.eu.

We would like to thank Árpád Ábrahám, Enghin Atalay, Simon Evenett, Piero Gottardi, Bas Jacobs, Dirk Krueger and Coen Teulings for comments and discussions.

†Disclaimer to be added.

1 Introduction

1.1 Motivation

It is a well know argument from the public economics literature that high levels of average taxes, and especially tax progressivity (average tax rates that rise in income), play an important role in shaping optimal human capital investments.[1] In particular, high average and marginal tax rates reduce human capital investment, and therefore pre-tax wages. This is typically considered more relevant the more skilled workers are (Guvenen, Kuruscu, and Ozkan, 2014). As a consequence, progressive tax schedules compress the distribution of earnings.[2] Since the 1970s, average income tax levels, as well as their progressivity (the extent to which average tax rates rise in income) have fallen dramatically in the United States. Estimated average tax schedules for the years 1983 and 2003 are depicted in Figure 1.

While these results are suggestive, other theories of the earnings distribution focus on the relative prices for different skills. For example, the theory of Skill-Biased Technological Change (SBTC)[3] assumes that an increased supply of high-skilled labor increases the relative price of low-skilled labor.

This paper combines both perspectives to tackle the primary challenge to both types of models: the 'polarization' phenomenon. Polarization refers to the observation that, starting in the 1980s, jobs in the middle of the earnings distribution have seen less growth in wages and employment than those at the top or bottom. This holds true both in the US (Autor and Dorn, 2013) and across advanced economies (Goos, Manning, and Salomons, 2014). This development coincided with growth in overall earnings inequality, i.e. a growing difference between the top and the bottom. Figure 2 displays these phenomena for the United States.[4]

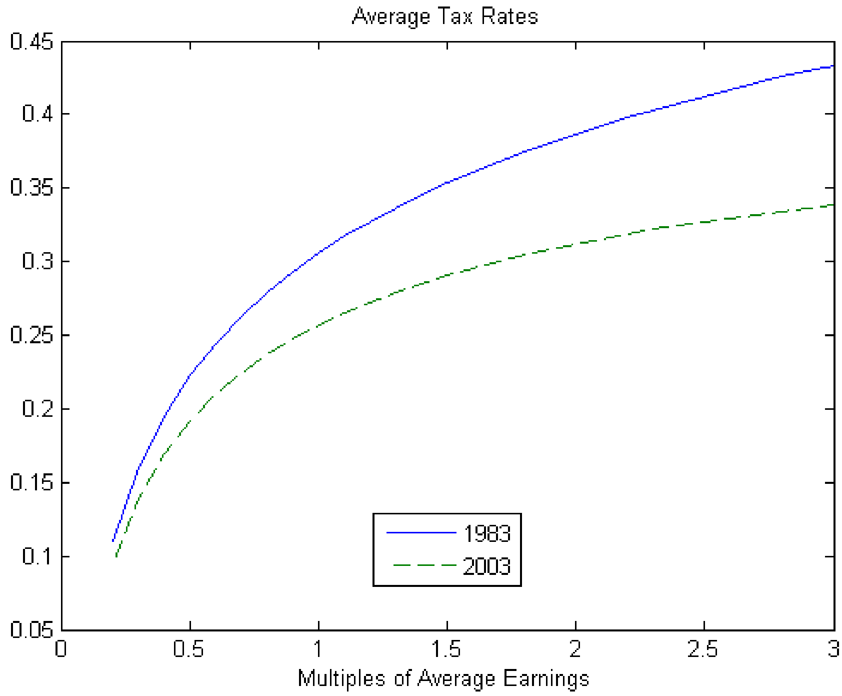
[1] See for example Bovenberg and Jacobs (2005) for a static setting, and Stantcheva (2017) for a dynamic extension. In this line of research, human capital is considered one-dimensional, and general equilibrium effects on wages play no role. Recently, a literature has developed that considers the original Mirrlees problem when many types interact in general equilibrium. A recent contribution is by Sachs, Tsyvinski, and Werquin (2016). The formation of types (or human capital) has so far been taken as exogenous. The same applies to previous work by Teulings (2005), which provided a framework for tracing out general equilibrium effects across many types.

[2] Guvenen, Kuruscu, and Ozkan argue that this mechanism is key to understanding differences in earnings inequality across countries and over time, showing that changes to tax schedules explain up to two-thirds of the change in the US wage premium between 1973 and 2003.

[3] Katz and Murphy (1992); Berman, Bound, and Machin (1998)

[4] Importantly, all of this refers to *relative* wage growth. In *absolute* (real US Dollar) terms, wage growth is positive everywhere, but relatively flat along the lower half of the distribution and steeply rising in the upper half.

Figure 1: Estimated average tax rates in the United States (from Guvenen, Kuruscu, and Ozkan, 2014).

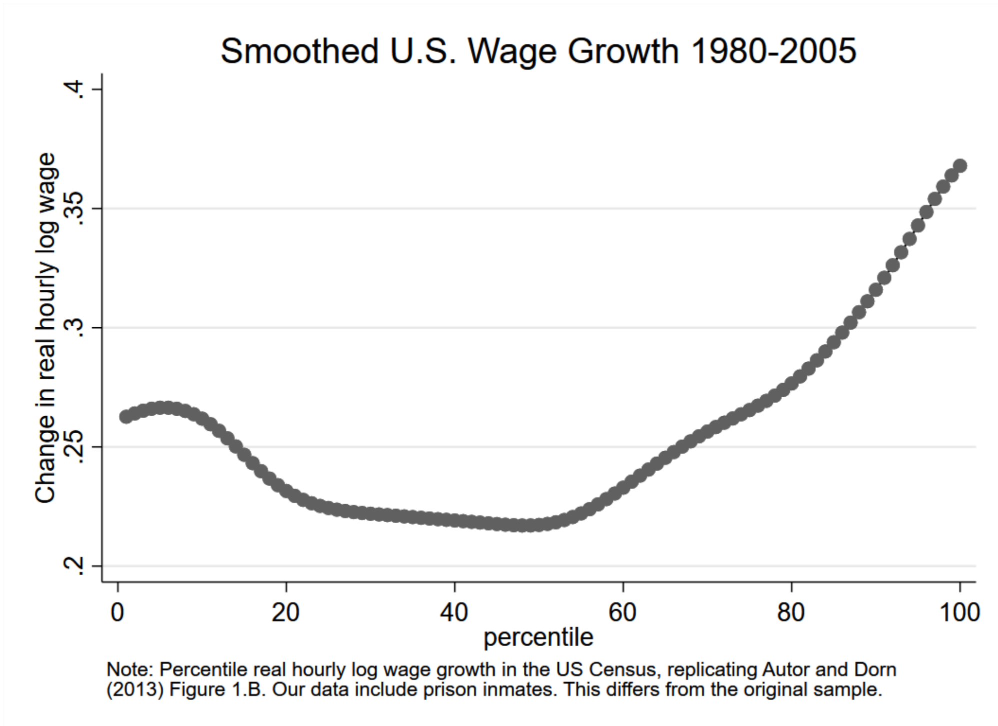


A number of explanations have been put forward to explain what is different about jobs in the middle of the income distribution, such as offshorability, competition from China, or declines in unionization rates in the manufacturing sector, but consensus has formed around the view that these jobs have a higher degree of 'routineness', and are therefore more susceptible to automation (by machines, robots, and computers). In short, polarization of the labor market is seen as demand-driven, and attributed to exogenous technological forces. See Autor et al. (2010) for a review of this literature.

This paper takes an entirely different and complementary approach. It extends the analysis of incentive changes as in the public economics literature to a setting where all workers simultaneously need two skills: manual and cognitive skills. Cognitive skills are more important at the top of the earnings distribution, and more learnable.[5] A reduction in tax progressivity increases incentives to acquire cognitive skills at the top. This extends the supply of cognitive skills (a quantity effect), raising the relative price of manual skills (a price effect). The former effect mostly benefits the top of the distribution, where incentives matter most, while the latter benefits the bottom of the distribution, where labor relies heavily on manual skills.

[5]A similar setup has been used by Guvenen and Kuruscu (2010) and Guvenen and Kuruscu (2012) to study skill-biased technological change. However, in both papers the general equilibrium wage effect is deliberately shut off by choosing a linear production technology. This precludes interaction effects between relative skill quantities and prices of the type we study in this paper.

Figure 2: Wage Inequality Growth and Polarization in the United States



The middle stays behind in relative terms.

We begin our analysis with a set of empirical results. Combining data on occupational skills from the Dictionary of Occupational Titles (DOT) with Census data, this paper establishes that there seem to be *two* relevant dimensions of job skills: cognitive and manual skills. Most jobs require a combination of both. We find the importance of these skills to be heterogeneous over the distribution of earnings: manual skills play a relatively larger role at the bottom of the distribution, cognitive skills play a larger role at the top. We will also argue that cognitive skills depend heavily on schooling decisions. We discuss the details of this analysis in Section 2.

These empirical findings have important implications for the impact of tax incentives: First, tax incentives are more relevant for those at the top of the income distribution. This is because cognitive skills are subject to individual investment – and therefore incentives – to a much larger extent than manual skills, and because cognitive skills dominate at the top of the income distribution. Second, in general equilibrium, a change in the relative amount of cognitive skills may affect the relative prices of the two skill types and therefore individual pre-tax earnings - a channel that is absent in models of one-dimensional human capital, but common in the literature on SBTC.

Motivated by these empirical findings, Section 3 sets out a life-cycle model in continuous time, in which earnings are derived from cognitive and manual skills, cognitive skills are subject to endogenous investment decisions, and relative wages are determined in general equilibrium. Importantly, in our model of skills it is not different education levels that map into different skill types as is standard in the SBTC literature. Instead, each individual uses some amount of both skills. This also sets it apart from existing models of polarization, which generate polarization by introducing more than two types that each specialize in one skill (see Acemoglu and Autor, 2011).[6] Several choices keep the model tractable, allowing for an analytical solution of individual choices despite the life-cycle nature of the model.

Section 4 exploits the extraordinary tractability of our model to study the effects of tax progressivity in our setting theoretically. We emphasize two implications: inequality and polarization. Just like in the uniform human capital model, a decline in tax progressivity impacts the top of the income distribution more than the bottom, thereby increasing income inequality in absolute terms. Additionally, having two types of skills, polarization arises. This is because lower tax progressivity increases the relative supply of cognitive skills more than that of manual skills, thereby increasing the latter's relative price. If this effect is strong enough, it can even increase the wages of those at the bottom relatively more than of those at middling levels of the distribution. In short, the tails of the distribution potentially show a relatively stronger response to changes in tax progressivity, with a reduction in progressivity causing polarization of pre-tax earnings. We discuss the underlying mechanism in further detail in Section 4.

The earnings distribution is likely subject to a multitude of economic forces, and no single mechanism will be able to fully account for the changes that took place during the period in which polarization arose. In this paper, we attempt to study our supply-side mechanism in isolation. That limits the extent to which we can observe its empirical implications in data. Nevertheless, we include a qualitative comparison of the model's macro-economic implications to data, both over countries and across time, in Section 5. Using OECD data on income distributions across countries and over time, we construct measures of income inequality and measures of tax progressivity across countries. Patterns that are easily accounted for by our multi-dimensional model are pervasive. We also discuss implications for changes over time, and the limitations present in verifying these.

Section 6 takes an enriched version of the model that can be solved numerically. Parameters are now tied down so that the baseline version of the enriched model matches relevant moments of the US economy. The model is then used to study the quantitative impact of a

[6] In Lindenlaub (2017) individuals also derive earnings from several skills. She analyzes a matching model with multi-dimensional skills and shows how different rates of technological change between different skills can lead to polarization. Unlike our model, she takes all skills as exogenously given.

typical decline in tax progressivity.[7] In order to do so, we use the decline of US tax progressivity since the 1980s that we discussed above as our experiment. We calibrate the model economy to match moments from the US economy in the early 2000s. We then compare the steady-state earnings distribution of this economy with the counter-factual tax progressivity of 1983 to the one in 2003 and calculate the rate of change in earnings.

The main goal of this exercise is to gauge the general quantitative "bite" of the human capital investment channel on changes in the earnings distribution, rather than wanting to account for the empirical change in earnings growth over the same period. This would require at the very least taking into account the transition period as well as cohort composition effects, both of which our model is silent on. More generally, we are looking at our mechanism in isolation, whereas in reality several channels are likely to have played a role in the rise of polarization. The results from the experiment indicate that the model captures growth in overall wage inequality reasonably well. Most of the change comes out of the upper half of the income distribution, in line with the empirical evidence. The results further indicate that the polarization effect exists, and is quantitatively sizable but smaller than what we observe empirically. In conclusion, our mechanism has impact under quantitatively relevant variations in policy. We also argue why our estimate of the mechanism's quantitative implications might be seen as a lower bound.

Our results contribute to two separate literatures. First, they explain why the lower half of the income distribution responds little to changes in tax progressivity. Existing papers focusing on uniform human capital, such as Guvenen et al. (2014), lack explanatory power in this region of the distribution. By adding the general equilibrium relative price effect, our model complements the direct incentive effect studied in their paper with the general equilibrium price effect. The latter works primarily in the lower half of the distribution and helps to limit the increase in total inequality.

Second, existing theories of polarization are primarily labor demand driven. Autor and Dorn (2013) introduce a third 'routine' skill category and explain polarization through increasing automation of 'routine'-intensive tasks, reducing the demand for jobs located in the middle of the wage distribution. By adding general equilibrium relative price effects to the traditional skilled-unskilled dichotomy of the endogenous human capital literature, we are able to generate qualitatively similar changes in the earnings distribution without resorting

[7] In this paper, we focus on tax progressivity. As Guvenen, Kuruscu, and Ozkan demonstrate, allowing for flexible labor supply makes tax levels a disincentive in the accumulation of human capital as well.In the context of a cross-country comparison, Guvenen, Kuruscu, and Ozkan find that differences in tax progressivity are a more important determinant of differences in inequality than are differences average tax rates, explaining our focus on the former.

to a third type of skill. Given the complexity of the earnings distribution, there are likely many underlying factors at work. Consequently, we see our channel as complementary to the skill-demand based explanations put forward in the existing literature.

Throughout, we emphasize that taxation is just one particular type of disincentive to human capital formation. In principle, there are many other distortions driving a wedge between public and private returns to education that differ across parts of the population. Two of the major trends in the 2^{nd} half of the 20^{th} century have been declines in gender and race based discrimination, both in education and the labor market. Hsieh et al. (2016) attribute about 25% of total economic growth in the US between 1960 and 2010 to changes in discrimination against women and minorities. It also seems reasonable to think that these trends correspond to a reduction in wedge progressivity, since discrimination is likely more salient towards the top of the distribution. Thus, for the remainder of this paper, one may want to think of 'wedges' more generally whenever we discuss taxes. Our quantitative results regard taxation only, so that investigating the role of discrimination for polarization is our main suggestion for further research. Section 7 concludes and provides further such directions for future research.

2 Manual and Cognitive Skills

We use data on the skills required in a number of occupations, from the Dictionary of Occupational Titles. We analyze the structure of these data using a statistical technique (Principal Component Analysis) that allows us to reduce the dimensionality of the data and subsequently interpret them. We find that skills are best summed up by two dimensions: cognitive and manual skills, both of which are important. In order to map the skill content into the wage distribution, we link the DOT data to the Census. This also enables us to investigate how skill measures relate to other observables, such as education. It turns out that the cognitive component of skills is strongly correlated to measures of education, while the manual component is not. Below, we discuss our data sources in more detail. Empirical methodology and results are presented thereafter.

2.1 Data

2.1.1 DOT

We use the 'Current Population Survey (CPS), April 1971, Augmented With DOT Characteristics and Dictionary of Occupational Titles (DOT)', obtained from the ICPSR. This version of the CPS was augmented with data on occupation characteristics from the 4th edition of the Dictionary of Occupational Titles (DOT). The 4th edition of the DOT is unique, in the sense that it is the final edition of the so-called 'Analyst Database': Over decades, starting in the mid-1930s, the United States Employment Service led an effort to systematically document the skills required to perform a range of occupations. This was done by sending trained occupational analysts to job sites, where they would complete standardized questionnaires on occupation content. While the database was revised since, the focus after the 4th edition of the DOT shifted to the generation of O*NET data, which are based on surveys of employees and employers, and therefore much less suitable for comparison across occupations. Following much of the literature, we therefore choose to use the 4th edition DOT. The main advantage of using the augmented CPS database is that it provides us with numbers of workers per occupation in the original DOT occupation classification.

The 4th edition DOT provides information on 46 variables of skills needed for and characteristics of 3886 DOT occupations (some examples: Marine Architect, Die-Designer Apprentice, Weather Observer, Hypnotherapist). In the nationally representative CPS database that we use, we also have a proportion of the working population for each of these occupations. The 46 variables consist of the analyst's answers to a wide variety of questions per occupation:

1. To what extent does the job relate to data, people, things? (3 questions)
2. What educational development is required (reasoning, mathematical, language, vocational)? (4 questions)
3. To what extent are aptitudes like intelligence important, or finger dexterity? (11 questions)
4. What temperaments relate to some occupation? (10 questions)
5. What are the physical demands of the job? (6 questions)
6. What physical environment does the job take place in? (7 questions)
7. To what interests does the job relate? (5 questions)

The survey includes clear and detailed instructions on how to answer these questions, making the answers comparable across occupations. Because the questions in the database vary in type and topic, and their number is large, researchers typically make ex-ante decisions on which variables to use. For example, while some questions clearly relate to skills, others clearly do not (such as interest and environment variables). We try to keep any pre-selection to a minimum, and include the first three categories of questions on the DOT when we perform Principal Component Analysis. These categories, comprising 18 questions, have in common that they must all be answered on a numeric scale that suggests some form of cardinal interpretation. (This is generally not true for the other categories: they are of the ‘yes or no’ type.) Hence, we treat them cardinally not to lose information. They also all clearly relate to skills, rather than the environment or personal characteristics of the typical person performing the job. We provide more detail on the 18 questions with our empirical results.

2.1.2 Census

We obtain a crosswalk between DOT occupation codes and Census 1990 occupation codes from the Analyst Resource Center (amongst others associated with the US Department of Labor). Whenever several DOT codes map into one census code, we take the average of component scores as the component score for that Census occupation. This procedure results in 452 occupations.

We use US Census data from IPUMS for all non-skill data (wages, hours worked, employment shares over Census occupations, education, and so forth), where we take a random sample of 50 thousand observations for each of the census years we use. Non-farm hourly wage rates are constructed by combining wage income and non-farm business income, following the example of the Census itself, and correcting for the number of weeks worked and the number of hours worked in a typical week. We reflate all wages to 2012 levels using the ‘CPI total items for the United States’ from the Federal Reserve Bank of St. Louis. We remove all occupations which are not present throughout our sample, as well as all farm occupations. The final number of occupations for which we have data in our sample is 308.

Appendix A details the construction of population percentiles of income (hourly wages) from these data.

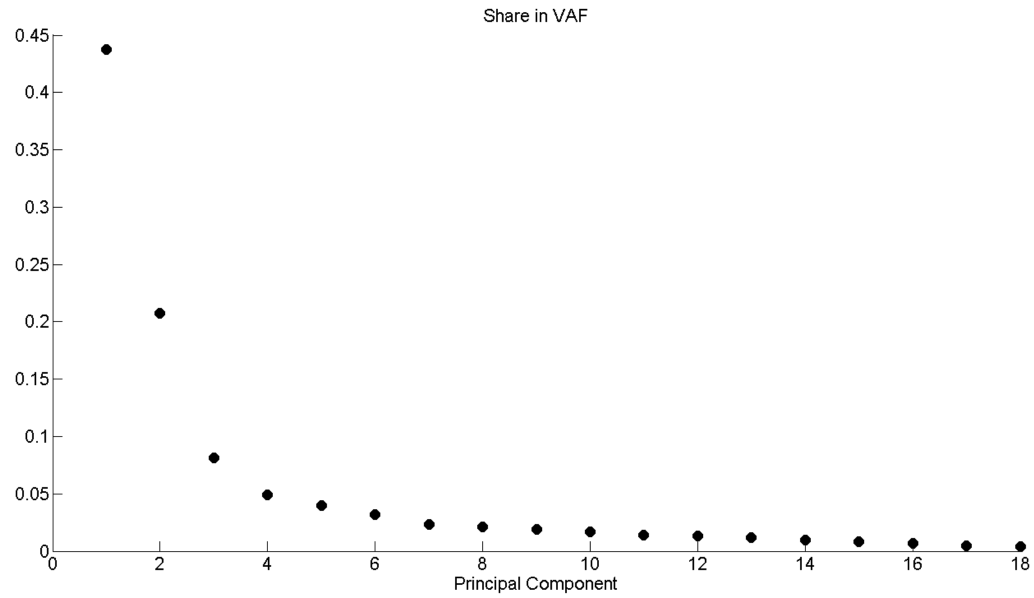
2.2 Empirics

2.2.1 PCA

The leftmost column of Table 1 shows the labels of the 18 DOT questions we include in our analysis. This set is still large, so that we want to reduce it for more tractable empirical analysis. We think of these skills as *ex-ante* equally important indicators of underlying core skills, and want to find out what these underlying skills look like. One method that allows doing so is Principal Component Analysis (PCA).

PCA is a relatively standard technique for dimension reduction, that creates new variables by linear combinations of existing ones. Its objective is to maximize the variance of the new variable, which is called a principal component. Each subsequent component’s vector of weights to the variables is assumed orthogonal to the previous ones'. (To make this problem well-defined, variables are first standardized to have mean zero and standard deviation one, and the total weight given to each of them is restricted to be no more than one.) The optimality condition for this problem is a simple eigenvector-eigenvalue decomposition, which yields as many components (eigenvectors) as there are variables, with all components orthogonal to each other. The corresponding eigenvalues relate directly to the variance accounted for by each component. One can simply think of the components as new dimensions: the dimensions are rotated such that the first dimension explains as much variance as possible, thereafter the second, and so on. Thus, the components are identified up to sign and scaling. Variance accounted for by each component are displayed in Figure 3. Clearly, the first two components dominate the others in explanatory power: they jointly explain more than 60% of the variance in the data, while no other component explains more than 10%. The third component and further component do not seem to pick up a fundamentally different aspect of skills, but rather seem to modulate the first two. Full PCA results are included in Appendix B.

Figure 3: PCA Scree Plot



What do the components look like? Table 1 shows the correlation between the first two components with the DOT skill measures over occupations (weighted by their share in employment). Those questions that are negatively correlated with the first component are highlighted. A brief study of the category groups with positive (negative) correlations with the first (second) component unambiguously leads to the conclusion that the first component relates to measures of cognitive ability, while the second component relates to physical skills.

Table 1: Component Correlations

Variable	Component 1: 'Cognitive'	Component 2: 'Manual'
Data	0.49	-0.31
People	0.45	-0.27
Things	-0.34	0.12
GED Reasoning	0.47	-0.36
GED Mathematical	0.46	-0.39
GED Language	0.48	-0.38
Specific Vocational Prep.	0.36	-0.22
Intelligence	0.51	-0.44
Verbal	0.51	-0.49
Numerical	0.43	-0.62
Spatial	0.05	0.41
Form Perception	0.05	-0.09
Clerical Perception	0.53	-0.51
Motor Coordination	-0.44	-0.07
Finger Dexterity	-0.63	-0.14
Manual Dexterity	-0.57	0.21
Eye-Hand-Foot Coord.	-0.14	0.44
Color Discrimination	-0.28	0.46

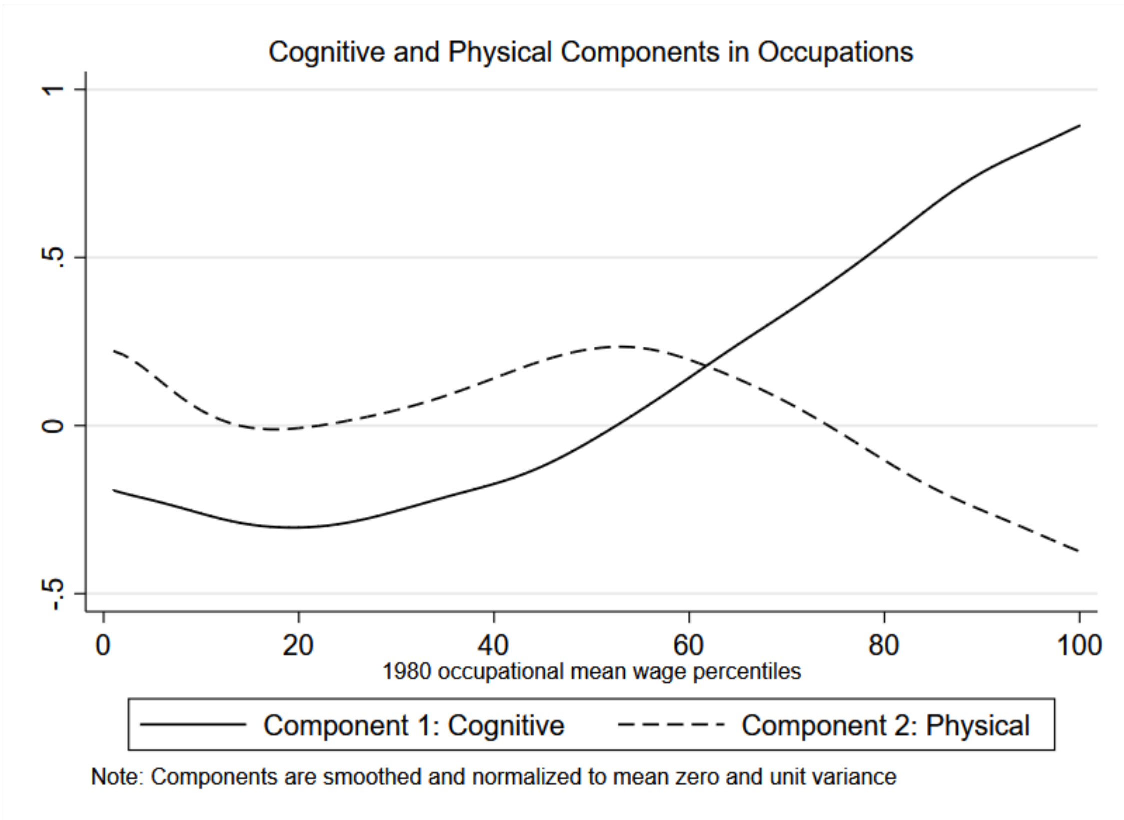
The orthogonality assumption inherent to the method has the natural economic interpretation that these are truly different underlying skills: being good at one does not mechanically imply being good at another. At the same time, there can certainly still be correlation in abilities in the population of observed occupations. In fact, the correlation between observed occupation scores on the first two components is -0.25: those with more cognitive ability tend to be less able physically, and vice versa.

The main drawback of our type of analysis is that we look for underlying skill categories in the data *per se*, i.e. not in relation to the wages or schooling decisions we expect them to explain. Our results in first instance only aim to have explanatory power with regards to the variables observed in the DOT. Several arguments speak for our approach nevertheless. First, the clear advantage of this approach is that our measures are in some sense still direct measures of skills, even if they are compounded and rely on analysts. Any explanatory power they have in our further analysis is not due to how we have produced them. Second, and related, the questions included in the DOT were included for a reason: because they were believed to be relevant measures of occupational skill. Last but not least, there is a related literature in which data on skills are directly related to wages. The most important reference in this regard is Yamaguchi (2012), who estimates a structural model of wage development in relation to unobserved skills using the same data on occupational skills as we do. He also finds two underlying skill factors to be of major importance, which he refers to as cognitive and motor tasks.

2.2.2 Covariates

We investigate how the results of the principal component analysis described above relate to the wage distribution. Figure 4 plots the first two components over population wage percentiles. As one would expect, the cognitive component is of minor importance in the lower end of the wage distribution, and starts to increase in importance somewhere below the median. In contrast, the physical component is relatively flat for the lower half of the distribution. Above the median it continuously declines with increasing skill level.

Figure 4: Components over the Distribution



Taken together, the PCA results imply that the multi-dimensionality of human capital or skills captured in the DOT can be summarized in two main factors, which we call cognitive and manual. As expected, the physical skill is relatively more important in the lower half of the wage distribution, while the cognitive becomes increasingly important for the higher skilled occupations. We view this as an interesting result, as the skilled-versus-unskilled dichotomy has a long tradition in the analysis of human capital.

Empirically, the skilled-versus-unskilled distinction has often been proxied for by years of education or by comparing college educated workers to those without college education. Figure 5 compares the PCA cognitive component to these traditional skill measures. It plots the cognitive component alongside the average years of education and the share of college graduates across wage percentiles in the population. All three measures behave very similarly in a qualitative sense. They are systematically lower for the lower half of the distribution and exhibit a break around the median. In the right half, all three measures are quickly increasing. Table 2 presents correlations between these measures of education and the two components, confirming the results from Figure 5: Education strongly correlates with cognitive skills, but not with manual skills.

Figure 5: Components and Education

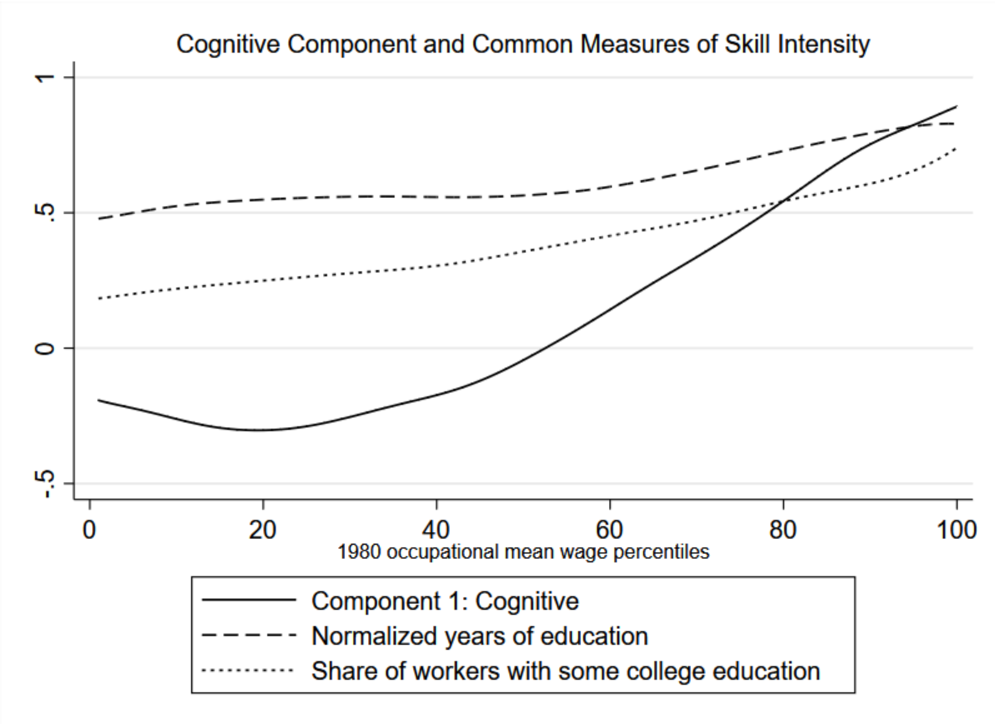


Table 2: Correlation between skill components and education measures

Correlations	'Cognitive'	'Manual'
Years of Education	0.97	-0.83
College Share	0.98	-0.70

To sum up, the empirical results suggest that both manual and cognitive skills are important, manual skills predominate in occupations at the bottom of the income distribution while cognitive skills predominate in the upper half. Furthermore, cognitive skills are highly correlated with traditional measures of schooling, which suggests that the cognitive skill is formed through education.

These observations have implications for the design of our theoretical framework. In particular, they suggest a model of skills where both manual and cognitive skills are important. This is much in line with the literature on Skill-Biased Technological Change, however there is one crucial difference: In the SBTC literature, it is commonly assumed that agents exclusively supply one type of labor, skilled or unskilled. Empirically, the most common strategy for mapping workers to skill types is by applying a cut-off for years of education above (below) which workers are categorized as skilled (unskilled). In contrast, the PCA results presented above suggest that both skills are continuous, i.e. workers in each occupation supply a *bundle* of cognitive and manual skills, rather than just one of the two. This precludes the mapping from years of schooling to skill categories. We will discuss the implications for our theoretical framework in the next section.

3 Model

In the remainder of the paper, we will explore both theoretically and quantitatively how changes to tax policy distort skill accumulation incentives, and thereby impact the shape of the earnings distribution.

The below begins by setting the environment. Agents are heterogeneous by two types of skills: manual and cognitive. Manual skills are considered innate. Wages accruing to both skills are set in general equilibrium, so that they depend on the relative supply of both skills. The accumulation of cognitive skills is modeled using a time-in-school setup. Several of our modeling choices serve to keep the framework tractable, thereby allowing us to simplify the individual life-cycle problem to a much more tractable static one, in spite of the complications presented by multi-dimensional skills and a non-linear tax function.

3.1 Environment

Human Capital Accumulation We assume that the manual skill $\mathcal{H}_m$ is innate, i.e. it cannot be accumulated and depends only on manual ability α_m. In contrast, the cognitive skill $\mathcal{H}_s$ is subject to human capital accumulation or schooling.

To have manual skills not respond at all should be seen as a simplifying assumption. Key to our mechanism is that those at the bottom of the distribution respond less to incentive changes. The previous section has attempted to demonstrate that cognitive skills are strongly positively related to formal measures of schooling and vocational training, whereas manual skills are not.

This does leave the possibility that learning on-the-job is relatively more important for manual skills. This seems unlikely however, because occupations at the bottom of the income distribution show much less wage growth over the life cycle than occupations at the top. Thus, those in occupations at the top learn more (marketable) skills over the life cycle. The final remaining question is whether this could be beause the individuals in those occupations are simply better at learning. Yamaguchi (2012) uses combined data on occupations, skills, and life-cycles of earnings to disentangle sources of wage growth by education level and skill. He finds what we assume: that wage growth over the life cycle is mostly due to growth in cognitive skills, rather than manual ones.

$\mathcal{H}_s$ is accumulated by spending time in school at the beginning of an agent's life. Agents are active for one period of time, $t \in [0,1]$ and begin their active period in school. They can leave at any time $x \in [0,1]$ to begin working. Individuals are endowed with cognitive and manual abilities $\alpha = (\alpha_s,\, \alpha_m)$. Ability α is the source of heterogeneity in the population. It is continuously distributed with pdf $f(\alpha)$ over a finite and positive support $[\underline{\alpha}, \bar{\alpha}]^2$. As we will see shortly, the time-in-school setup we choose here helps keep the individual problem tractable.

Investment in human capital takes both time and money, but we model the time component only. At some level of abstraction, one can simply think of one in terms of the other, since time that was to be spent on schooling can be converted into money at the wage rate. However, conversion in the opposite direction does not work. In addition, money can be borrowed, observed, and therefore also made deductible from taxable income. All of this does not apply to time, making the mechanism we describe more fundamental than what can simply be resolved by education policies.

The efficiency of time in school in the model depends on individual's ability α, the amount of human capital already accumulated and the time spent in school, according to a schooling function $s(\mathcal{H}_{s,t}, \alpha_s, t)$. Cognitive human capital $\mathcal{H}_s(\alpha_s, x) > 0$, is assumed to be a continuous and twice differentiable function. Finally, we assume that $\frac{\partial \mathcal{H}_s(\alpha_s,x)}{\partial x} > 0$: accumulated human capital is a strictly positive function of time spent in school.

After quitting school, human capital stays constant for the remaining active time of the agents. Human capital accumulation thus follows a differential equation

$$\frac{\partial \mathcal{H}_{s,t}}{\partial t} = \begin{cases} s(\mathcal{H}_{s,t}, \alpha_s, t), & t \leq x \\ 0, & t > x \end{cases} \tag{1}$$

and the amount of cognitive human capital while working is the same as the level of human capital at time x, $\mathcal{H}_s(\alpha, x)$.

Multi-Dimensional earnings A continuum of agents of mass one derives (pre-tax) earnings from two skills, manual m and cognitive s, quantities of which are measured by $\mathcal{H}_m$ and $\mathcal{H}_s$, respectively. Individuals are assumed to supply both of their skills to the market simultaneously. Labor supply is assumed inelastic, and set to one. This implies that each individual supplies $\mathcal{H}_m(\alpha_m)$ units of manual labor and $\mathcal{H}_s(\alpha_s, x)$ units of cognitive labor when working. Instantaneous gross earnings y of an individual can now be described as

$$y = w_m \mathcal{H}_m(\alpha_m) + w_s \mathcal{H}_s(\alpha_s, x). \tag{2}$$

Here, w_m and w_s are wage rates for efficiency units of manual and cognitive skills, taken as given by the agents.

On the production side of the economy, final output is produced by an aggregate production function taking the total amounts of manual and cognitive skills in the economy as inputs,

$$Y = F(M, S).$$

Here, M and S are aggregate amounts of manual and cognitive skills in the economy and given by

$$S = \int_{\underline{\alpha}}^{\overline{\alpha}} \mathcal{H}_s(\alpha, x) f(\alpha)\, d\alpha$$

and

$$M = \int_{\underline{\alpha}}^{\overline{\alpha}} \mathcal{H}_m(\alpha) f(\alpha)\, d\alpha = 1,$$

in light of the above normalization of manual skills. We assume competitive input markets, thus wages w_m and w_s are given by their respective marginal products.

3.2 Individual Problem

Consumption prices are taken as the *numeraire*. Markets are complete, there are no sources of uncertainty, and a single asset completes the market: agents can save and borrow asset a without limit (except for repayment at $t = 1$) at the discount rate r. We do not model the capital stock of the economy in general equilibrium, which is consistent with an economy that is 'small' and open to foreign goods and capital, or with an aggregate production technology that is linear in capital.

Taxes are assumed to be paid instantaneously over the rate of income y. This is a crucial model ingredient: in reality, tax schedules are applied yearly, which is a short frequency compared to the length of the life cycle. In our continuous time model, we capture this by applying the tax schedule to the wage rate at any instance. All of this keeps the problem tractable.

An individual's problem then looks as follows:

$$\max_{\substack{x\in[0,1],\\ \{c_t\}_{t\in[0,1]}}} \quad \int_0^1 e^{-rt}\frac{c^{1-\sigma}}{1-\sigma}dt \tag{3}$$

$$\begin{aligned}
&\text{subject to} \quad \forall t: \\
&\frac{\partial a_t}{\partial t} = -c_t + a_t r \quad \text{if} \quad t \leq x, \\
&\frac{\partial a_t}{\partial t} = y_t(1-\tau(y_t)) - c_t + a_t r \quad \text{if} \quad t > x, \\
&a_0 = 0, \quad a_1 \geq 0, \\
&c_t \geq 0.
\end{aligned}$$

Agents decide on the duration of their education and on the life-cycle profile of consumption and savings. The government levies taxes on earnings in order to meet wasteful government spending target G. We assume that the average earnings tax rate $\tau(\cdot)$ is governed by two parameters, responsible for the level ϕ and the degree of tax progressivity θ, $\tau(\cdot) = \tau(\cdot\,;\phi,\theta)$.

A specific example of such an earnings tax function is $\tau(y) = 1 - \frac{1}{\phi}(y)^{-\theta}$. This is the tax function we will use in our quantitative exercise in Section 6. The function has a property that will turn out to be convenient. Let $T(y) = y\tau(y)$ denote total taxes to be paid, so that $T'(y)$ is the marginal tax rate. A tax function is considered progressive if marginal taxes are higher than average taxes everywhere. For this function, we have that $\frac{1-T'(y)}{1-\tau(y)} = 1-\theta$, so that taxes are progressive if $\theta > 0$, and more so (by this measure) as θ rises.

In typical models of taxation over the life cycle one needs marginal tax rates that increase with income for tax schedules to have an effect on human capital investment (as long as there is no labor-leisure choice; cf. Heckman, 1976). In our setup this will be less straightforward, due to the separation of the life-cycle into exclusive learning and working periods. In the special case of the above tax function, the influence of the tax schedule on the human capital investment decision in our problem is captured entirely by θ, which in turn can be seen as a function of both marginal and average tax rates.

3.3 Simplification of the Individual Problem

Life-cycle problem (3) above has a much simpler equivalent. First, since individuals can only decide between going to school or working full-time, the amount of human capital after school is fully determined by the time spent in school, x, and cognitive ability, α_s: $\mathcal{H}_s(\alpha_s, x)$. Second, because markets are complete, agents smooth consumption and the choice of the optimal x is unconstrained. As there are no other choices in the model, the agent now simply maximizes lifetime after-tax income with respect to time in school.

Taxes are applied at every instant, but income is zero at first and constant thereafter. Thus, we can just think of the tax rate being applied to the income level after school, and levied over the part of the life-cycle not spent in school.

These steps deliver a much simpler version of the individual problem:

$$\begin{aligned} \max_{x\geq 0} \quad & (1-x)y(1-\tau(y;\phi,\theta)) \\ & \text{subject to:} \\ & y = w_m\mathcal{H}_m(\alpha_m) + w_s\mathcal{H}_s(\alpha_s, x). \end{aligned} \tag{4}$$

For the functional forms and parameterization chosen in the quantitative section below, the solution to x is always unique. Until then, it is assumed to be.

In the following, we will be interested in how changes in the tax progressivity θ shape the income distribution in this environment. As discussed, the optimal schooling decision will be directly governed by the degree of tax progressivity. From now on, we will therefore directly work with $h(\alpha_s, \theta) \equiv \mathcal{H}_s(\alpha_s, x^*)$ instead of human capital $\mathcal{H}_s(\alpha_s, x)$, where x^* is the argmax of individual income maximization problem (4).

4 Tax Policy, Growing Inequality, and Polarization

We now formalize the notions of earnings inequality and earnings polarization within the context our framework, and then provide conditions for either of them to arise. As will become clear, earnings polarization is a special case of earnings inequality growth, with additional restrictions on the relative movements of earnings within the lower tail of the distribution.

4.1 Policy Changes and the Earnings Distribution

We study the effect of changes in tax policy, in particular changes to the tax progressivity θ on the shape of the earnings distribution. We do so in a rather abstract sense, using the simplified framework we have derived above. For tractability, we further assume that there is no heterogeneity in the manual skill, so that we can normalize its level to one: $\mathcal{H}_m = 1$. This implies that the aggregate amount of manual skill, M is also equal to one. We now also surpress the subscript of α_s to avoid cumbersume notation, so that α now simply refers to learning ability for cognitive skills. When we discuss the effects of policy changes in this model, we will essentially be comparing steady states.

We call into memory the effect taxes have on individual choices at given prices. Investing in education enables individuals to achieve higher earnings in a shorter time span. Recall that in our framework taxes are not applied to life-time income, but instead levied on instantaneous earnings in order to resemble real-world income taxation. For a given life-time income, positive tax progressivity punishes higher per-period earnings relative to an earnings profile that spreads out lower earnings over a larger fraction of the life-cycle. Through this channel, tax progressivity reduces the optimal amount of time spent in school, x^* and cognitive human capital, $\mathcal{H}_s(\alpha, x^*)$.

Our general object of interest is given by

$$\frac{\frac{\partial y}{\partial \theta}}{y}.$$

We are interested in effects of an abstract policy, represented by a change in θ: a lower θ represents less progressive taxes.[8] We then want to see what this reform does to earnings in percentage change terms, since this is the theoretical equivalent of Figure 2. In particular, we will then be interested in how these policy effects differ across different parts of the income distribution.

It turns out, the object above can be easily decomposed into separate parts as follows:

$$\begin{aligned}
\frac{\left(\frac{\partial y}{\partial \theta}\right)}{y} &= \frac{\left(\frac{\partial\left(\frac{y}{w_s}\right)}{\partial \theta}\right)}{\left(\frac{y}{w_s}\right)} + \frac{\left(\frac{\partial w_s}{\partial \theta}\right)}{w_s} \\
&= \frac{\left(\frac{\partial\left(\frac{w_m}{w_s}\right)}{\partial \theta}\right) + \frac{\partial h}{\partial \theta}}{\left(\frac{w_m}{w_s}\right) + h} + \frac{\left(\frac{\partial w_s}{\partial \theta}\right)}{w_s} \\
&\equiv \underbrace{\frac{w'_\theta}{(w+h)}}_{\text{price effect}} + \underbrace{\frac{h'_\theta}{(w+h)}}_{\text{quantity effect}} + \underbrace{\frac{{w_s}'_\theta}{w_s}}_{\text{level effect}}
\end{aligned} \tag{5}$$

The last line just restates the previous one using the definitition $w \equiv \frac{w_m}{w_s}$. Percentage changes in income have been separated into three terms. The first two terms describe the potential trade-off policy created in a multi-dimensional model: on the one hand, policies can increase (or decrease) incentives to acquire human capital, which we call a *quantity effect*, but when they do so for all individuals this increases (decreases) the overall supply of learnable skills in the economy, which can decrease (increase) their relative price - a *price effect*. Their relative importance and strength depends on an individual's schooling responsiveness to policy. This responsiveness will in principle depend on the level of the ability parameter α, generating potentially non-linear effects of policy changes on income changes. The last term above affects all individuals equally in percentage terms. It arises because wage effects are described in skill premium terms, but a policy reform can also impact the overall productivity level in an economy - hence the name *level effect*.

[8] More formally, θ remains undefined in this section: it is simply a parameter that has influence on the formation of cognitive skills, and therefore wages. Its influence is defined by the assumptions below.

Polarization in this environment arises if relative income changes in response to a policy change are stronger in the tails of the income distribution than in the center of the distribution. Since in the model, income is entirely determined by ability, this is equivalent to comparing income responses for different ability levels. Formally, inequality growth and polarization in response to tax policy changes can be defined in terms of relative income changes as follows.

Definition 1. *Inequality Change and Polarization. Inequality change exists in response to a policy change in θ if for $\underline{\alpha}$ and $\overline{\alpha}$ the following inequality holds:*

$$\left.\frac{\left(\frac{\partial y}{\partial \theta}\right)}{y}\right|_{\alpha=\overline{\alpha}} < \left.\frac{\left(\frac{\partial y}{\partial \theta}\right)}{y}\right|_{\alpha=\underline{\alpha}} < 0. \tag{6}$$

Polarization exists if in addition to equation (6), for some $\hat{\alpha} \in (\underline{\alpha}, \overline{\alpha})$ the following holds as well:

$$\left.\frac{\left(\frac{\partial y}{\partial \theta}\right)}{y}\right|_{\alpha=\underline{\alpha}} < \left.\frac{\left(\frac{\partial y}{\partial \theta}\right)}{y}\right|_{\alpha=\hat{\alpha}} < 0. \tag{7}$$

Definition 1 restates income inequality growth and polarization in response to a decline in tax progressivity θ in concise terms. First, the inequality aspect, i.e. high income individuals pulling away even further from the rest of the population, requires a stronger relative income response to a policy change for high ability individuals than low ability individuals. Second, the non-monotonicity in the lower tail of the income distribution distinguishes polarization from general trends in overall inequality: low-income individuals are able to partially catch-up to medium income individuals, while overall inequality still increases.

4.2 Conditions for Inequality Growth and Polarization

Our reformulation of the life-cycle problem (3) as income maximization problem (4) allows us to establish sufficient conditions for polarization to arise in our framework. The intuition for the result that follows further below is contained in a reformulation of decomposition (5) in terms of tax policy elasticities:

$$\varepsilon_\theta^y = \varepsilon_\theta^w \frac{w}{w+h(\alpha)} + \varepsilon_\theta^h(\alpha)\frac{h(\alpha)}{w+h(\alpha)} + \varepsilon_\theta^{LE}. \tag{8}$$

Equation (8) rewrites the total relative earnings elasticity in terms of a weighted sum of elasticities of the price and the quantity effect. (The level effect is also there, but it is uninteresting because it does not vary by income levels.) Polarization can arise because the elasticity of the price effect is the same for all abilities (because it only depends on prices), while the elasticity of the quantity effect (under asumptions) grows in α. In addition, the weights also change in ability, since for higher ability the share of income generated from cognitive human capital increases. Under light assumptions, we can have non-monotone changes in relative income across abilities. In particular, comparing levels of α starting with the lowest, if $\varepsilon_\theta^h(\alpha)$ is small in absolute magnitude for small and medium α, then the first term will dominate the lower half of the distribution. It will shrink as income goes up. Only once $\varepsilon_\theta^h(\alpha)$ is large enough in absolute magnitude will the absolute magnitude of the overall elasticity begin to grow in earnings.

In the following, we will make this reasoning more precise by first laying out the assumptions sufficient for polarization to arise, and then go through the precise mechanism. The main purpose of this exercise is to formalize the intuition just laid out.

Definition 2. *Define $\hat{\alpha} \in (\underline{\alpha}, \overline{\alpha})$ as any α that generates a local extremum in the tax elasticity of income,*

$$\left.\frac{\partial \varepsilon_\theta^y(\alpha)}{\partial \alpha}\right|_{\alpha=\hat{\alpha}} = 0.$$

Inequality change and polarization in Definition 1 were defined for an arbitrary interior $\hat{\alpha}$. Definition 2 restricts, as we will see below, the interior $\hat{\alpha}$ to the ability level that under the below assumptions minimizes (in absolute terms) the income elasticity $\varepsilon_\theta^y(\alpha)$.

Assumption 1. *Shape of the human capital elasticity. Human capital elasticity $\varepsilon_\theta^h(\alpha)$ behaves relative to the relative wage elasticity ε_θ^w as follows for different ability levels:*

- *The human capital quantity elasticity is increasing and convex in ability level α in absolute terms, $\frac{\partial \varepsilon_\theta^h(\alpha)}{\partial \alpha} < 0$ and $\frac{\partial^2 \varepsilon_\theta^h(\alpha)}{\partial \alpha^2} < 0$, and at the lower bound approximately zero: $\frac{\partial \varepsilon_\theta^h(\underline{\alpha})}{\partial \alpha} \approx 0$.*
- *For abilities $\alpha \leq \hat{\alpha}$, the human capital elasticity is lower in absolute terms than the elasticity of relative prices, $\varepsilon_\theta^h(\alpha) - \varepsilon_\theta^w > 0$.*
- *For high ability individuals, the human capital elasticity is higher in absolute terms than the elasticity of relative prices, $\varepsilon_\theta^h(\overline{\alpha}) - \varepsilon_\theta^w < 0$.*

Assumption 1 states that for low ability individuals, the relative price elasticity is stronger than the quantity elasticity. The quantity elasticity is increasingly growing in ability and for abilities high enough, it becomes larger than the relative price elasticity in absolute terms.

Assumption 2. *Human capital function. Human capital accumulation is strictly convex and positive in α, $h'_\alpha, h''_{\alpha\alpha} > 0$. In addition, for all $\alpha \in [\underline{\alpha}, \overline{\alpha}]$, the following restriction on the* *shape of the optimal human capital quantities holds:*

$$\frac{1}{w+h} < \frac{2(h'_\alpha)^2}{h''_{\alpha\alpha}}.$$

Recall from above that $h(\alpha,\theta)$ is the cognitive human capital resulting from the optimal schooling decision of the agent, $h(\alpha,\theta) \equiv \mathcal{H}_s(\alpha,x^*)$. Therefore, Assumption 2 effectively imposes restrictions shape of the schooling technology. In particular, it is required that the human capital is convex in ability, but cannot increase too quickly: $h''_{\alpha\alpha}$ has to be sufficiently small.

Result 1. *Given Assumption 1, inequality changes as defined in Definition 1 occur in response to a change in tax policy θ. If in addition Assumption 2 also holds, polarization as defined in Definition 1 occurs as well.*

Proof: For the inequality change part, we will show that income elasticity $\varepsilon^y_\theta(\alpha)$ is larger in absolute terms for $\overline{\alpha}$ than for $\underline{\alpha}$. For the polarization part, we will show that under Assumption 2, there is a unique $\hat{\alpha}$ as defined in Definition 2 and this $\hat{\alpha}$ is the argmax of the maximum of $\varepsilon^y_\theta(\alpha)$ (minimum in absolute terms).

Inequality change: To show that $0 > \varepsilon^y_\theta(\underline{\alpha}) > \varepsilon^y_\theta(\overline{\alpha})$, we will first show that $\varepsilon^y_\theta(\underline{\alpha}) - \varepsilon^{LE}_\theta > \varepsilon^w_\theta$ and second that $\varepsilon^y_\theta(\overline{\alpha}) - \varepsilon^{LE}_\theta < \varepsilon^w_\theta$. To show the former, consider

$$\begin{aligned}
&\varepsilon^y_\theta(\underline{\alpha}) - \varepsilon^{LE}_\theta > \varepsilon^w_\theta \\
\Leftrightarrow\, &\varepsilon^w_\theta \frac{w}{w+h(\alpha)} + \varepsilon^h_\theta(\alpha)\frac{h(\alpha)}{w+h(\alpha)} > \varepsilon^w_\theta \\
\Leftrightarrow\, &\varepsilon^w_\theta \frac{w}{w+h(\alpha)} + \varepsilon^h_\theta(\alpha)\frac{h(\alpha)}{w+h(\alpha)} > \varepsilon^w_\theta \frac{w}{w+h(\alpha)} + \varepsilon^w_\theta \frac{h(\alpha)}{w+h(\alpha)} \\
\Leftrightarrow\, &\varepsilon^h_\theta(\underline{\alpha}) > \varepsilon^w_\theta,
\end{aligned}$$

where the last inequality holds by Assumption 1. For the high ability case $\varepsilon^y_\theta(\overline{\alpha}) - \varepsilon^{LE}_\theta < \varepsilon^w_\theta$, a similar argument holds. Together, this implies that $\varepsilon^y_\theta(\underline{\alpha}) - \varepsilon^{LE}_\theta > \varepsilon^w_\theta > \varepsilon^y_\theta(\overline{\alpha}) - \varepsilon^{LE}_\theta$. $\varepsilon^y_\theta(\underline{\alpha}) > \varepsilon^y_\theta(\overline{\alpha})$ is implied by the last inequalities, establishing inequality change as defined in equation (1.6) from Definition 1.

Polarization: To show that the unique global maximum of $\varepsilon^y_\theta(\alpha)$ is at $\hat{\alpha}$, we show first that $\hat{\alpha}$ is the only extremum, and second that the first derivative is strictly larger (smaller) than zero for all α smaller (larger) than $\hat{\alpha}$. The first derivative of $\varepsilon^y_\theta(\alpha)$ with respect to α is given by

$$\frac{\partial \varepsilon^y_\theta(\alpha)}{\partial \alpha} = \left(\varepsilon^h_\theta(\alpha) - \varepsilon^w_\theta\right)\frac{wh'_\alpha}{(w+h)^2} + \frac{\partial \varepsilon^h_\theta(\alpha)}{\partial \alpha}\frac{h}{w+h}. \tag{9}$$

Therefore, $\frac{\partial \varepsilon_\theta^y(\alpha)}{\partial \alpha} > (<)\, 0$ boils down to

$$\left(\varepsilon_\theta^h(\alpha) - \varepsilon_\theta^w\right)\frac{wh'_\alpha}{(w+h)^2} > (<) \; - \frac{\partial \varepsilon_\theta^h(\alpha)}{\partial \alpha}\frac{h}{w+h}.$$

For the left-hand side, Assumption 1 implies that $\left(\varepsilon_\theta^h(\alpha) - \varepsilon_\theta^w\right)$ is strictly declining in α, positive for $\underline{\alpha}$ and negative for $\overline{\alpha}$. Define $\tilde{\alpha}$ as the α such that $\varepsilon_\theta^h(\tilde{\alpha}) - \varepsilon_\theta^w = 0$. Note that by Assumption 1 $\overline{\alpha} > \tilde{\alpha} > \hat{\alpha}$ holds. Assumption 2 implies that $\frac{wh'_\alpha}{(w+h)^2}$ is strictly decreasing in α and strictly positive. Together this implies that the left-hand side is strictly declining in α for $\alpha < \tilde{\alpha}$, positive for $\alpha < \tilde{\alpha}$ and negative for $\alpha > \tilde{\alpha}$. For the right-hand side, Assumption 1 implies that $-\frac{\partial \varepsilon_\theta^h(\alpha)}{\partial \alpha}$ is ≈ 0 for $\underline{\alpha}$ and strictly increasing and positive for all $\alpha > \underline{\alpha}$.

Taken together, this implies that $\frac{\partial \varepsilon_\theta^y(\underline{\alpha})}{\partial \alpha} > 0$ and $\frac{\partial \varepsilon_\theta^y(\overline{\alpha})}{\partial \alpha} < 0$. Furthermore, since the left-hand side is strictly declining while positive and the right-hand side strictly increasing and positive, there exists exactly one α for which $\frac{\partial \varepsilon_\theta^y(\alpha)}{\partial \alpha} = 0$. This proves the existence of a unique $\hat{\alpha}$ as defined in Definition 2. Since for all $\alpha < \hat{\alpha}$, we have that $\frac{\partial \varepsilon_\theta^y(\alpha)}{\partial \alpha} > 0$ and for all $\alpha > \hat{\alpha}$, we have that $\frac{\partial \varepsilon_\theta^y(\alpha)}{\partial \alpha} < 0$, $\hat{\alpha}$ is a global maximum. Since $\hat{\alpha} \in (\underline{\alpha}, \overline{\alpha})$, this establishes polarization as defined in equation (1.7). $\square$

The aim of this section has been to detail conditions on the price and quantity effects for earnings polarization to arise in our framework with two types of skills and general equilibrium price effects. We show that depending on the shape of the elasticities, polarization can arise in our framework of two-dimensional skills and general equilibrium skill price effects. Rather than examining the relevance of our two key assumptions here, we later calibrate an enriched version of our model in which these results indeed arise.

In the following, we will first present some reduced-form cross-country evidence for our mechanism in the next section, and then try to quantify the economic importance of this supply-side channel in Section 6.

5 Models versus Data

We now compare the implications of our model to what we know about the data. We begin with an analysis of cross-country data, followed by a discussion of the literature on relevant time trends.

5.1 Across Countries

Tax systems differ in progressivity across countries (Figure 6). Our model makes clear predictions on the role of progressivity in income inequality: more progressive tax systems produce less inequality as measured by relative earnings in the income distribution. This is driven by changes in the upper half of the income distribution, while inequality in the bottom half varies little with tax progressivity. We now investigate these predictions in cross-country data, for which we need measures of tax progressivity and relative earnings inequality.[9]

Coen-Pirani (2017) sets forth a method to obtain measures of tax progressivity from OECD data, which works as follows: If we assume that both gross and net earnings are log-normally distributed and that taxes follow the functional form that was introduced as an example above, measured Gini coefficients of gross and net earnings can be used to back out an estimate of θ. A panel data set of Gini coefficients is available from the OECD Income Distribution Database. We use data on the working age population (ages 18 to 65), using the income definition that the OECD followed until 2011 for better availability and comparability of data. Data are available for about 30 OECD member countries, covering a period from the mid 1970s to 2015. Coverage is thin for earlier years, but improves towards the end of the sample. Because the panel is rather unbalanced, we average the resulting measures of tax progressivity for the years 2010–2015, and use this as a cross-section of country-level tax progressivity.

Also available from the OECD is an unbalanced panel of relative earnings inequality measures across countries and over time. The underlying population are full-time employees of either gender. These include the earnings ratio of the 90th percentile cut-off to the 50th percentile cut-off, or *90-50 ratio*, and the same for the 50th and 10th percentile, the *50-10 ratio*. While these two measures describe relative inequality above and below the median, the resulting *90-10 ratio* measures inequality. We choose to use these measures because their movement has a close correspondence to what we consider in our theoretical exposition: if relative (percentage) changes are the same across the distribution, then these measures will remain unchanged with tax progressivity. We again average over the years 2010–2015. The overlap between the two datasets consists of 32 countries.

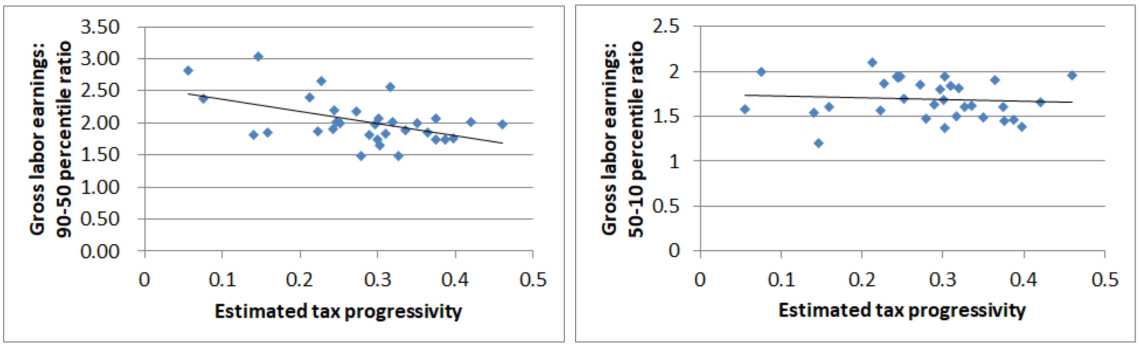
Next, we estimate the (linear) impact of tax progressivity on earnings inequality at different points in the distribution. Results of OLS regressions of the latter on the former are displayed in Table 3. Figure 6 presents the results graphically. Tax progressivity is generally associated with a reduced relative earnings inequality. For the 90-10 ratio, the slope is statistically significant at the 5% level. For the 90-50 ratio the slope is statistically significant at the 1% level. For the 50-10 ratio, the slope is not statistically significant, even at the 10% level. While tax progressivity has quite some explanatory power in the upper half of the distribution, as measured by the R^2, this is not true for the lower half of the distribution.

[9]We think of tax systems as exogenous to the remainder of the economy. If tax progressivity is in some way a response to higher earnings inequality, this would counteract our mechanism and make it harder to find a correspondence between theory and data.

Table 3: Regression results

	Inequality measure		
	90-10 ratio	90-50 ratio	50-10 ratio
Progressivity θ	-1.17	-1.92	-0.19
	(0.54)	(0.59)	(0.41)
Constant	Yes	Yes	Yes
Observations	32	32	32
R^2	0.14	0.26	0.01

Figure 6: Tax progressivity and Inequality across Countries



All these results align very well with our model prediction, even as we are looking at simple linear relations. In a more extensive quantitative exercise, Guvenen, Kuruscu, and Ozkan (2014) analyze the responses of a one-dimensional model of human capital to changes in tax progressivity. Their model has several added features, such as flexible labor supply, and a more flexible functional form for average tax rates. While they show that their model does well in accounting for 90-10 ratios, it is less succesful in disentangling 50-10 ratios. Our analysis suggests that this is due to the multi-dimensional nature of skills, which is most relevant to the bottom half of the distribution.

5.2 Over Time

Tax progressivity in the United States has declined dramatically since the 1970s (Figure 1). That same observation applies to many other countries (Guvenen, Kuruscu, and Ozkan, 2014). What implications would this have had for other observables, in particular changes to the shape of the earnings distribution? The literature reviewed above suggests that the earnings distribution is shaped by a number of different forces. That fact significantly limits the extent to which we can verify the direct impact of tax changes on inequality through our mechanism empirically. Nevertheless, we attempt to provide a qualitative discussion.

Educational decisions are decisions for the long run. Agents' expectations of future policies are therefore key to the empirical mechanism we describe, and observed transitions may be slow. In any case, one would expect that younger cohorts react more strongly to incentives than older ones, so that empirically it should be the younger cohorts that cause polarization. This is indeed what the empirical literature finds. Cortes, Jaimovich, and Siu (2017) document that the fall in what they call 'routine' occupations in the middle of the distribution can be largely attributed to two groups: young and prime-aged men with low levels of education where it comes to 'routine manual' occupations, and young and prime-aged women with intermediate levels of education where it comes to 'routine cognitive' occupations. In terms of age structure, this lines up well with the implications of our mechanism. While our model does not speak to gender *per se*, the gender differences these authors highlight underline our main suggestion for further research: changes in labor market discrimination may be important. We will come back to this in more detail in the concluding Section 7.

The observed changes in tax incentives predict rising inequality, in gross wages and even more so net of taxes, and polarization. While these phenomena can also be observed in the data, underlying quantities and prices of human capital cannot. Did wage inequality grow due to greater differences in human capital or due to rising prices for the highly skilled? Our mechanism would suggest the former. On the other hand, the theory of Skill-Biased Technological Change takes educational attainment at face value as a measure of human capital. Because educational attainment has slowed down, it concludes that prices for the labor of the highly skilled must have risen.

Separating human capital quantities from their price is a central empirical challenge in the labor literature, and existing evidence is scarce. One approach is to identify an age in the life-cycle at which human capital is unlikely to change much, and attribute wage changes at that age to changes in the price of human capital. This is the approach followed by Bowlus and Robinson (2012). These authors do not find large changes in prices at all, attributing changes in the wages of different educational groups to changes in human capital. This would be more in line with our mechanism than for example SBTC. To cause polarization, our general equilibrium effect would require a growing relative price of manual versus cognitive skills. Price estimates by such skill types are unfortunately unavailable.

Similar caveats apply to direct measures of human capital (such as schooling attainment), measured skill premia, and returns to schooling (before and after tax). While our model makes predictions for each of these, it is not clear what the relevant empirical counterpart is. A number of possible comparisons are further complicated by the fact that our model is not a growth model, so that it cannot account for longer-run trends in these data.

Finally we return to our initial comment: our mechanism is unlikely to have been the only relevant change during the period. Other explanations focus on secular technological developments that have shaped the wage distribution through changes in labor demand. These explanations are complementary to ours as long as relative prices of skills move in the same direction as in our model. That holds for the literature that describes how middling 'routine' occupations are more prone to automation. The same applies to papers that explain the growth of service occupations at the bottom of the distribution through changes in demand. SBTC fits our model less well, since it starts with the assumption that it is prices of human capital that have caused inequality to grow. Future research will hopefully shed further light on this debate.

Having discussed evidence for some general predictions of our framework, we will now use it to investigate the quantitative relevance of our mechanism.

6 An Enriched Model

In this section, we extend our model to include heterogeneity in manual (non-learnable) skills and choose some functional forms. We then parameterize our model to reproduce several key stylized facts of the US economy, and use it to evaluate counterfactual policies.

6.1 Model Description

From here on out, the individual problem is the same as in (3) above. A continuum of agents, whose total mass equals one, live for $t \in [0, 1]$, first goes to school until $t = x$ and then works. When in school ($x \leq t$), individuals build learnable human capital according to the following law of motion:

$$\frac{\partial h_{s,t}}{\partial t} = \beta t^{\beta-1} \alpha_s h_{s,t}. \tag{10}$$

Thereafter, $\frac{\partial h_{s,t}}{\partial t} = 0$. This function resembles more conventional human capital functions such as the one due to Ben-Porath (1967), but the time-in-school structure keeps the model computationally simple. Time in school is more productive for the more able and educated, but diminishes over time. Both α_s and the existing stock $h_{s,t}$ linearly enter the production of additional human capital, with the benefit that the differential equation has an analytical solution: $h(\alpha_s, x) = h_{s,0} e^{(\alpha_s x^\beta)}$.

$h_{s,0}$ is assumed linear in α_s, so that the two are perfectly correlated. This simplifies the problem significantly at little cost. Non-learnable human capital is given by $h_{m,t} = h_{m,0} = \alpha_m$. Both skills are assumed to be independently drawn from normal distributions, resulting in a tuple (α_m, α_s) for each individual. When working ($x > t$), individuals derive income from both types of human capital:

$$y_t = w_m h_{m,t} + w_s h_{s,t}. \tag{11}$$

As mentioned above, we use the tax function $\tau(y) = 1 - \frac{1}{\phi}(y)^{-\theta}$. Using the simplification of (3) to the static problem of (4), we obtain a simple solution for x:

$$\frac{1}{1-x} = \frac{(\frac{\partial y}{\partial x})}{y}\frac{1-T'(y)}{1-\tau(y)} = \frac{w_s(\frac{\partial h(\alpha_s,x)}{\partial x})}{y}(1-\theta). \tag{12}$$

Several modeling choices are jointly responsible for allowing this level of tractability. As a result, we can now numerically analyze a life-cycle model that features two dimensions of skill, endogenous formation of human capital, and general equilibrium effects.

Before moving on to parameterize our model, we now complete the definition of the equilibrium. We consider overlapping generations such that the population distribution is always in steady state. Let the distribution of type tuples $(\alpha_m, \alpha_s) \in \mathcal{A}$ be denoted by λ. Define human capital aggregates as follows (where $I_{[\cdot]}$ is an indicator function):

$$H_m = \int_0^1 \int_{\mathcal{A}} h_{m,t} I_{[t>x]}\, d\lambda\, dt, \tag{13}$$

$$H_s = \int_0^1 \int_{\mathcal{A}} h_{s,t} I_{[t>x]}\, d\lambda\, dt. \tag{14}$$

Aggregate production takes place using the following production function:

$$Y = F(H_m, H_s) = A\Big[\gamma H_m^{\rho} + (1-\gamma)H_s^{\rho}\Big]^{\frac{1}{\rho}}. \tag{15}$$

The elasticity of substitution between the two inputs is given by $\frac{1}{1-\rho}$, and γ is a share parameter. We normalize output so that $A = 1$.

A government sets taxes τ_c and $\tau_n(\cdot)$. Its budget is balanced by expenditures G that are assumed not to influence any of the above:

$$\int_0^1 \int_{\mathcal{A}} c_t\tau_c + y_t\tau_n(y_t)I_{[t>x]}\, d\lambda\, dt = G. \tag{16}$$

Definition 3. *A stationary equilibrium of the model is defined as:*
Wages w_m, w_s,
allocations H_m, H_s,
government spending G,
decision rules for x, $\{c_t\}_{t\in[0,1]}$ $\quad\forall\quad (\alpha_m, \alpha_s) \in \mathcal{A}$
such that given the parameters of the model the following holds:

- *individual decision rules solve problem (3)*

- *goods markets clear:*

$$Y = \int_0^1 \int_{\mathcal{A}} c_t\, d\lambda\, dt \tag{17}$$

- *labor markets clear (equations (13) and (14))*

- *wages equal marginal products (of equation (15))*

- *and the government budget constraint is balanced (equation (16)).*

6.2 Parameterization

Equilibria of the economy are found numerically. Parameters are set to match moments of the data in the early 2000s. In doing so, the following parameterizations of initial abilities and human capital stocks is used. Let $\tilde{\alpha}_s$ denote a standard normal distribution, winsorized at three standard deviations.

$$\alpha_s = \mu_s + \sigma_s \tilde{\alpha}_s, \tag{18}$$

$$h_{s,0} = 1 + (\tilde{\alpha}_s - \underline{\tilde{\alpha}}_s)\psi_s. \tag{19}$$

$\underline{\tilde{\alpha}}_s$ is the lowest level of $\tilde{\alpha}_s$. The lowest level of $h_{s,0}$ is normalized to 1, while average learning ability, the spread in learning ability, and the spread in initial learnable human capital is controlled by parameters. Likewise,

$$h_{m,0} = 1 + (\alpha_m - \underline{\alpha}_m)\psi_m, \tag{20}$$

where α_m is standard normal and ψ_m controls the spread of initial non-learnable human capital.

Table 4: Parameters and moments

Parameter	Value	Moment	Model	Data
σ	2.857	Elasticity of intertemporal substitution	0.350	0.350
ψ_m	0.140	Earnings variance at start of working life versus overall	0.524	0.500
ψ_s	28.229	Gini coefficient of gross earnings	0.347	0.440
μ_s	0.947	Average share of working age spent in school	0.031	0.034
σ_s	0.225	Variance of share in school	0.002	0.002
β	0.856	Share with zero education after age 18	0.473	0.456
ρ	0.286	Elasticity of substitution in production	1.400	1.400
γ	0.520	Non-learnable share of output	0.248	0.250

Table 4 reports data moments. Some of our model parameters are straightforwardly informed by moments of the data, while for others much less clear-cut measures are available. We use the midpoint of the range of elasticities of intertemporal substitution reported in Havranek (2013) to set the same in the model (σ), but that parameter does not influence any of the results we report. The spread of both initial human capitals is important for overall earnings variation, and their relative size helps determine the extent to which that variation is present at age 0. Thus, we target the Gini coefficient of gross earnings as reported by the OECD for the year 2000. We also target a ratio of earnings variance at age 0 versus earnings variance overall of 1/2. While we do not have a precise estimate for this number from the data, research using the life-cycle of earnings Huggett, Ventura, and Yaron (2011) suggests about two thirds of earnings are pinned down after tertiary education. Finally, to determine the average and spread of ability, we target the share of a potential 48 years of working life from age 18 that is spent in school (i.e. college and beyond), the variance of these shares, and the share of pupils who do not spend any time in college. We calculate the data moments from the 2000 Census sample described in the above, where all education beyond 12th grade is counted as taking place during the adult life cycle.

Finally, the parameters in the production function are key to size general equilibrium responses. Unfortunately, no reduced form results on general equilibrium effects between skills as we describe them are available. Instead, we rely on evidence on general equilibrium effects between college educated and non-college educated labor. Here, a large body of evidence suggests an elasticity of substitution of about 1.4 (see for example Katz and Murphy (1992) and Ciccone and Peri (2005)). Because these two groups would both use either type of human capital, we take the view that this is a very conservative estimate of the two elasticity of substitution that is relevant to our model. To tie down the share parameter of the production technology, we target the share of non-learnable human capital in output. Again, no direct evidence is available, so that we tentatively set this target to 25%.

We estimate the tax function used in the above from tax rates at different levels of average US earnings for 2003, and then do the same for 1983, following Guvenen, Kuruscu, and Ozkan (2014) (we use the same data as those authors). This results in an estimate $\theta = 0.119$ for 2003, which is used for parameterizing the model, and an estimate of $\theta = 0.188$ for 1983, which we use in our counter-factual analysis below. The corresponding figures for ϕ are 1.343 and 1.446, respectively. Because we are interested in the incentive effects of precisely these tax schedules, G is set to clear the government's budget constraint each time and is purposefully not held constant.

In our numerical procedure, we also include a linear consumption tax to ensure that we match average effective tax rates on labor well. Consumption taxes are set to 7.5%, following the 2003 Figure reported in McDaniel (2007).

Table 4 also demonstrates the model's ability to match the data. Overall, model moments are close to data moments, although the model does struggle to create sufficient earnings heterogeneity to match the economy's inequality levels. This would most likely be addressed by incorporating more sources of heterogeneity into the framework (such as income risk, labor supply, and accumulation of skills on the job). Because that would imply losing the model's tractability, we do not pursue this route.

6.3 Results

To analyze the results of tax progressivity, we compare the steady state earnings distribution of the 1983 estimate of θ to the steady state distribution with the 2003 estimate. We think of this as a counter-factual reform in which tax progressivity was reduced. The procedure yields a reform that is realistic, both in shape and magnitude. We would not want to argue that our results are empirical in the sense that they have bearing on the change in the period. (For that to be the case, one would want to consider other factors, as well as the transition from one steady state to another.) Rather, we are looking for a counter-factual experiment that gives us a feeling for the effect sizes in our model.

We then turn to measures of inequality. Indeed, reducing the progressivity parameter has increased the 90-10 ratio about one-for-one, which is what we also find in our cross-country analysis. This increase can be almost entirely attributed to the upper half of the distribution i.e. the 90-50 ratio. Again, this is entirely in line with our cross-country findings. These results give us confidence that the model adequately captures the reaction of the earnings distribution to tax progressivity.

Figure 7a shows the results graphically (labeled 'baseline'). It is apparent that some polarization occurs, but little: the bottom wages grow a few tenths of percent more than those with the lowest wage growth. The top grows by almost 7% more than the lowest point.[10] To bear out polarization given the large increase in inequality in the top half, we show the same graph but restricted to the lower half of percentiles in Figure 7b.

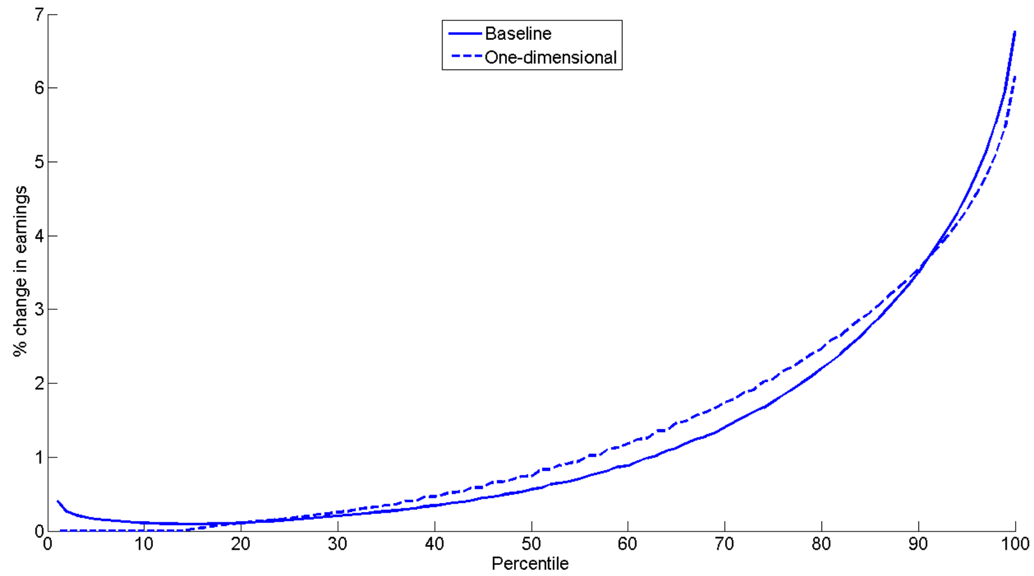
Figure 8 decomposes the baseline experiment into the three components of equation (5). The quantity effect, which is driven by incentives to accumulate human capital, dominates. But the price effect, representing the general equilibrium effect that makes manual labor relatively more valuable, is strong enough at the bottom to create a non-linearity. The level effect, as expected, is the same for all.

One might consider the effect sizes we present conservative. The elasticity of substitution between the two skill types may be smaller in practice, leading to larger price effects: the elasticity has been measured in the previous literature using data on college versus non-college educated labor. However, that categorization is a noisy measure of the underlying skills that our theory predicts is relevant. This would lead to an overestimation of the elasticity in a

[10]For those interested, we report that this is 6% and 45%, respectively, of the equivalent empirical change in the period. As already noted, we do not want to encourage such empirical interpretations too much.

Figure 7: Relative earnings change under counter-factual reform

(a) Full distribution



(b) Lower half of the distribution

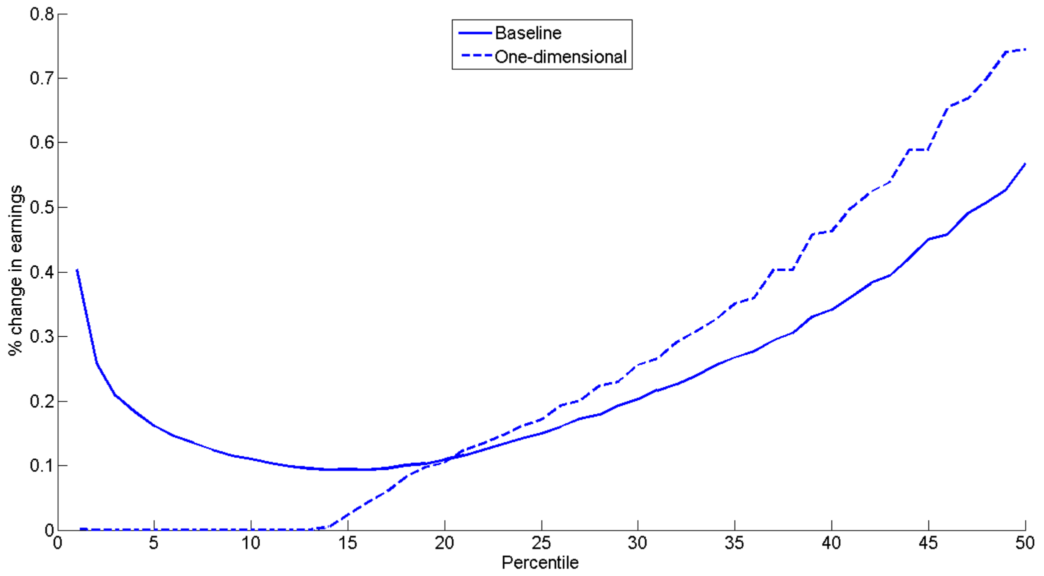
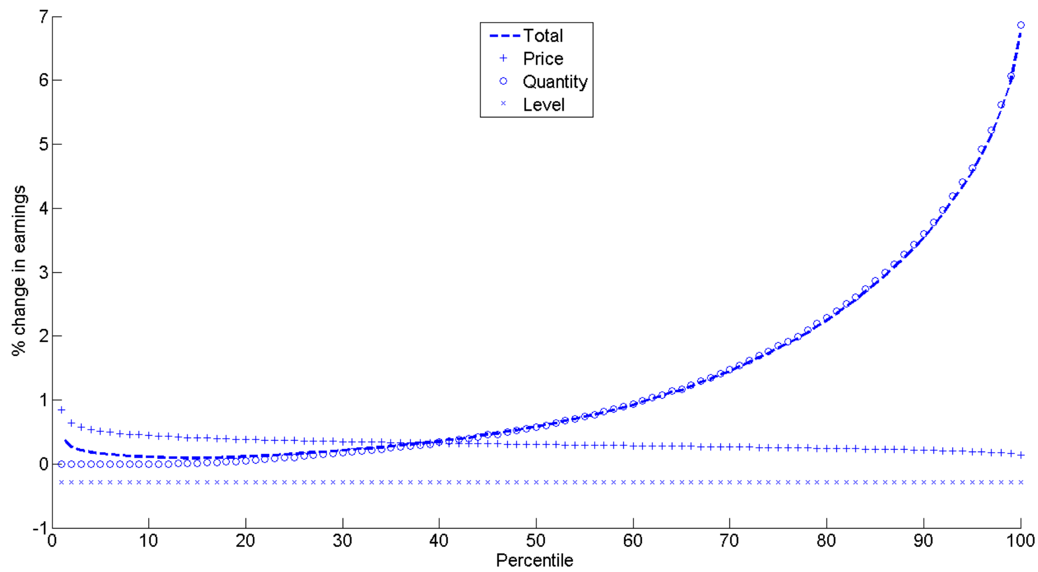


Figure 8: Relative earnings change decomposed



typical regression methodology (e.g. in that of Katz and Murphy (1992)) due to attenuation bias, reducing the price effect (which goes to zero as the elasticity goes to infinity). Second, we have not included leisure, which works as an amplifying mechanism (cf. Guvenen, Kuruscu, and Ozkan (2014)). Third, our view of human capital is a very limited one, because we only focus on time in school. The same incentives would however also affect learning on the job, making the overall impact much larger. In addition, in this paper we focus on the part of the labor wedge originating exclusively from income taxation. There exist other sources for the labor wedge, in particular discrimination. Since this is outside the current model, we will postpone a detailed discussion to the concluding Section 7.

Finally, we know little about the empirical counterpart of γ, the share of output the model attributes to manual skills. The next subsection contains a formal sensitivity analysis. As will become clear, the sensitivity is relatively small. This is reassuring, as it implies that our results are relatively robust to changes in γ.

6.4 Sensitivity Analysis

The main moment of which we are uncertain is the one informing γ, the share of output that is contributed by non-learnable skills. At the same time, this parameter is obviously crucial in assessing the importance of our mechanism: in the absence of non-learnable skills, the model collapses to a uniform human capital model. To make this clear, we re-calibrate the model setting the moment for γ to zero, which results in $\gamma = 0$ (and slight changes to some of the other parameters). Figures 7a and 7b also show the results in this case (labeled 'one-dimensional'). While the result is similar for overall inequality, polarization has disappeared. The effect on inequality within the bottom half of the population is now much more straight-forward.[11]

We provide a more formal analysis of the sensitivity of γ in the remainder of this Section. Our parameters can be interpreted as estimates of an indirect inference procedure: They are the result of minimizing the distance between the data moments described in Table 4, the vector of which we will now call $\hat{s}$, and the model moments that we will call $s(\theta)$ (where θ is the vector of parameters). Defining $\hat{g} = \hat{s} - s(\theta)$, we then used θ to minimize $\hat{g}' I \hat{g}$ (where I is the identity matrix that we use as weights) and reported the argmin $\hat{\theta}$ of our problem in Table 4.

Andrews, Gentzkow, and Shapiro (2017) establish a methodology for measuring the sensitivity of parameter estimates to estimation moments. They suggest reporting an estimate of the matrix $\Lambda = -(G'WG)^{-1}G'W$, where G is the Jacobian of the probability limit of $\hat{g}$ at the true parameter values θ_0, and W is the weighing matrix (the identity matrix in our case). The advantage of their method is that it is computationally simple to find a point estimate of G, and therefore Λ: because our objective vector $\hat{g}$ is additive and only $s(\theta)$ depends on the parameters, we can simply calculate the numeric Jacobian matrix S of our model moments $s(\theta)$ at the estimated parameter value $\hat{\theta}$. In short, we have that our *sensitivity* estimate is given by $\Lambda = S^{-1}$.

How should these sensitivity estimates be interpreted? Entry λ_{ij} of Λ tells us, roughly, how large the local impact of a change in data moment j is on parameter i. It can be used to calculate the asymptotic bias in our estimates associated with an alternative hypothesis on the data moments, as long as the alternative is sufficiently close to the data moments we report. More straightforwardly, we can use it to discuss the sensitivity of our estimates to the data moments. That is a particularly appealing feature in light of the uncertainty around some of the data moments that we report above. Because a unit change in the data moments is not always easy to interpret, we instead opt to report results relevant to a 1% change of each data moment. This is achieved by multiplying λ_{ij} by a percent of data moment j. The results are in Table 5.

Two parameters, σ and ρ, are only sensitive to the one moment on which they depend by a closed-form relation (the latter's sensitivity measure is zero in the table due to round-off). ψ_s takes on larger values, and so is generally more responsive in level terms. ψ_m and ψ_s react most heavily to moments that describe the distribution of earnings and schooling. As

[11] The 'one-dimensional' graph in Figure 7b appears to display a kink that is not actually there: investment in education is always non-zero due to an Inada condition in human capital formation. The visual effect arises because levels of the human capital distribution have been compressed to a percentile scale.

Table 5: Sensitivity Analysis

Moment nr.	1	2	3	4	5	6	7	8
σ	-0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ψ_m	0.00	0.00	0.00	-0.09	0.66	-0.00	0.00	0.00
ψ_s	0.00	-49.89	27.28	-253.95	-4373.20	-4.12	0.14	-11.16
μ_s	0.00	0.05	-0.06	0.31	4.49	0.01	-0.00	-0.00
σ_s	0.00	-0.00	0.00	-0.03	-0.20	-0.00	0.00	0.00
β	0.00	0.02	-0.01	0.06	1.60	0.00	-0.00	-0.00
ρ	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
γ	0.00	-0.00	0.00	0.04	-0.58	0.00	0.00	0.01

Note: Model moments are 1 – Elasticity of inter-temporal substitution, 2 – Earnings variance at start of working life versus overall, 3 – Gini coefficient of gross earnings, 4 – Average share of working age spent in school, 5 – Variance of share in school, 6 – Share with zero education after age 18, 7 – Elasticity of substitution in production, 8 – Manual human capital share of output. See also Table 4.

we would hope, parameters describing learning ability and the formation of human capital indeed react most strongly to those moments that describe the distribution of schooling. The parameter γ reacts strongly to the 5th moment, the variance of schooling, which clearly plays an important role in the determination of the model's parameters.

As discussed above, we have very little information about the 'non-learnable share of output', which is the eighth and last data moment in Table 5 above. It turns out that this moment does play some role in the determination of γ, but not a large one. That there is some sensitivity is quite in line with our expectation, given the analysis included above where we set $\gamma = 0$. The fact that the sensitivity is not extremely large is reassuring, since it implies that our results would change little if our target of γ was somewhat off.

7 Conclusion

This paper has provided evidence on the importance of different skills over the earnings distribution. These results suggest that occupations require multiple skills, and that both incentives to accumulate cognitive skills and the relative price of cognitive (versus manual) skills are of importance to the earnings distribution. The paper has set up a tractable model that allows for the joint analysis of both elements. In doing so, it has provided an alternative mechanism through which labor market polarization may arise.

In the paper we focus exclusively on the part of the labor wedge originating from taxation. An important additional source of the labor wedge originates from discrimination. Over the second half of the 20$^{\text{th}}$ century (labor market) discrimination against women and non-white groups arguably decreased a lot. There is growing evidence that the decline in discrimination has been quantitatively important for US macroeconomic outcomes. Dwyer (2013) provides evidence that polarization in employment has been driven to a substantial part by women increasingly entering the labor market, primarily in the tails of the distribution. Related evidence by Cortes et al. (2017) has been discussed in the above. Hsieh et al. (2016) estimate that about 25% of US output per capita growth between 1960 and 2010 can be attributed to an improved allocation of talent due declines in discrimination in the labor market and in access to education. Decreasing the price of education for a substantial share of the working population would have a similar effect as the decline in tax progressivity, by increasing the relative payoff of spending time in school. Similarly, if declines in discrimination take place in the form of 'breaking the glass ceiling', they might over-proportionally improve labor market outcomes for high-earning women, again resembling declines in tax progressivity. Potentially, these results therefore imply that the decline of progressivity of the effective labor wedge has been a lot larger than the decline in the explicit tax wedge. In this case, our results may understate the importance of the supply-side polarization channel.

Future research may lead in a number of directions. First, fundamental questions on our model of the labor market remain of interest. For example, credibly exogenous variation in skill levels might illuminate the prices paid for different levels (or bundles) of skills. Second, further research into the distributional effects of reduced discrimination against minority groups in the labor market seems warranted. Finally, while the emphasis in this paper has been on positive implications, one might ask what optimal tax and education policies look like in a model like ours. In the presence of general equilibrium effects, tax disincentives to the formation of human capital are more harmful than is traditionally assumed, likely warranting less progressive tax schedules.

A Population Percentiles

Population percentiles are obtained as follows. For each occupation in our sample we obtain mean hourly log wages $\overline{w}^{occ}$ and the share of the population employed in the respective occupation x^{occ}. We sort occupations by their mean log hourly wage in 1980. We construct percentile employment shares x^{perc} and average wages $\overline{w}^{perc}$ by mapping the occupation population shares into population percentiles. In particular, we assign to each percentile, the share of each occupation falling into the respective percentile using the 1980 population share per occupation. Doing this, we obtain the following conversion matrix C of dimensions $\#occ \times 100$, which by definition maps the vector of occupation population share vector x^{occ} into population percentile vector x^{perc}:

$$C'_{1980} x^{occ}_{1980} = x^{perc}_{1980} \equiv \mathbf{1}.$$

By construction, the population percentiles obtained in this way are 1 in 1980. Percentile mean wages in 1980 are obtained by multiplying occupational mean wages with the conversion matrix:

$$\overline{w}^{perc}_{1980} = C'_{1980} \overline{w}^{occ}_{1980}.$$

To obtain the change in employment shares between 1980 and 2010, we first calculate how occupational employment changed in terms of 1980 percentiles and then compute the rate of change. In particular, we take the conversion matrix from occupations into percentiles in 1980 and multiply it with the occupation employment share vector in 2010 as follows:

$$\Delta_{2010-1980} x^{perc} = C'_{1980} (x^{occ}_{2010} - x^{occ}_{1980}).$$

This calculation converts 2010 occupational population shares into the 1980 percentile bins. If for an occupation the employment share increased (decreased) relative to 1980, this will result in the respective population percentile to increase (decrease) as well.[12] For the calculation of changes of wages we similarly multiply the 1980 conversion matrix with 2010 occupational mean log hourly wages and obtain the growth rate by taking the difference of the 2010 and 1980 percentile wages.

B PCA Results

The table below displays the full PCA results. Each column represents the correlation between a component and the original variables. The table begins with the component that explains the largest share of variance, then the second largest, and so forth.

[12] Note that by using this strategy, we are restricted to the analysis of occupations which are present in both 1980 and 2010. Thus, we remain silent on the effects of vanishing and newly appearing occupations on the aggregate wage and employment distribution. The procedure follows the approach taken by Autor and Dorn (2013).

Data	0.49	-0.31	0.25	0.51	0.49	0.12	0.13	0.31	0.16	0.42	0.60	0.62	0.49	-0.31	0.77	-0.06	0.81	0.70
People	0.45	-0.27	0.34	0.39	0.56	0.47	0.15	0.05	0.19	0.20	0.57	0.68	0.70	0.32	0.70	0.33	0.46	0.36
Things	-0.34	0.12	-0.49	-0.15	-0.38	-0.42	-0.15	0.30	0.56	-0.19	-0.01	-0.45	-0.55	0.11	-0.29	0.09	0.47	0.45
GED Reasoning	0.47	-0.36	0.22	0.50	0.62	0.27	0.09	0.33	0.16	0.38	0.64	0.68	0.52	0.01	0.66	-0.19	0.86	0.77
GED Mathematical	0.46	-0.39	0.22	0.49	0.45	0.04	0.06	0.43	-0.02	0.55	0.56	0.62	0.45	-0.01	0.64	-0.35	0.83	0.72
GED Language	0.48	-0.38	0.27	0.54	0.61	0.34	0.04	0.43	0.14	0.32	0.65	0.69	0.65	0.04	0.67	-0.22	0.75	0.70
Specific Vocational Preparation	0.36	-0.22	0.04	0.46	0.36	0.03	0.17	0.24	0.29	0.31	0.46	0.61	0.31	-0.20	0.58	-0.13	0.88	0.75
Intelligence	0.51	-0.44	0.26	0.54	0.60	0.25	0.05	0.36	-0.02	0.39	0.72	0.69	0.52	-0.09	0.54	0.02	0.80	0.75
Verbal	0.51	-0.49	0.37	0.61	0.55	0.34	0.09	0.19	0.00	0.37	0.76	0.68	0.67	-0.15	0.60	-0.01	0.72	0.69
Numerical	0.43	-0.62	0.31	0.38	0.59	0.02	-0.06	0.38	0.11	0.59	0.58	0.53	0.36	-0.02	0.53	-0.22	0.71	0.72
Spatial	0.05	0.41	-0.59	0.07	0.13	-0.63	0.21	0.12	0.19	0.12	-0.15	0.02	-0.19	0.16	0.14	-0.03	0.50	0.56
Form Perception	0.05	-0.09	-0.32	0.49	-0.04	-0.09	-0.06	0.34	0.42	0.30	0.04	-0.11	-0.00	-0.03	-0.04	0.21	0.63	0.76
Clerical Perception	0.53	-0.51	0.60	0.42	0.38	0.45	-0.45	0.34	0.07	0.41	0.73	0.40	0.41	-0.09	0.41	-0.04	0.47	0.56
Motor Coordination	-0.44	-0.07	-0.47	-0.37	-0.55	0.01	-0.35	0.05	0.25	-0.14	0.08	-0.51	-0.24	0.08	-0.42	0.28	0.27	0.49
Finger Dexterity	-0.63	-0.14	-0.31	-0.00	-0.37	-0.05	-0.40	0.15	0.39	-0.03	0.04	-0.48	-0.15	0.08	-0.36	0.22	0.40	0.57
Manual Dexterity	-0.57	0.21	-0.54	-0.38	-0.45	-0.48	-0.13	-0.19	0.22	-0.47	-0.57	-0.70	-0.36	0.24	-0.53	0.11	0.15	0.15
Eye-Hand-Foot Coordination	-0.14	0.44	-0.51	-0.43	-0.30	-0.26	0.48	-0.20	-0.22	-0.69	-0.38	-0.17	-0.61	0.46	0.09	0.20	-0.16	0.01
Color Discrimination	-0.28	0.46	-0.05	-0.25	-0.09	0.16	0.51	0.60	0.39	0.00	-0.24	-0.27	-0.14	-0.00	-0.02	0.28	0.20	0.43

References

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

Andrews, I., M. Gentzkow, and J. M. Shapiro (2017). Measuring the sensitivity of parameter estimates to estimation moments. *The Quarterly Journal of Economics 132*(4), 1553–1592.

Autor, D. et al. (2010). The polarization of job opportunities in the us labor market: Implications for employment and earnings. *Center for American Progress and The Hamilton Project*.

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

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *The Journal of Political Economy 75*(4), 352–365.

Berman, E., J. Bound, and S. Machin (1998). Implications of skill-biased technological change: international evidence. *The quarterly journal of economics 113*(4), 1245–1279.

Bovenberg, A. L. and B. Jacobs (2005). Redistribution and education subsidies are Siamese twins. *Journal of Public Economics 89*(11), 2005–2035.

Bowlus, A. J. and C. Robinson (2012). Human capital prices, productivity, and growth. *The American Economic Review 102*(7), 3483–3515.

Ciccone, A. and G. Peri (2005). Long-run substitutability between more and less educated workers: evidence from us states, 1950–1990. *The Review of Economics and Statistics 87*(4), 652–663.

Coen-Pirani, D. (2017). Geographic mobility and redistribution: A macro-economic analysis.

Cortes, G. M., N. Jaimovich, and H. E. Siu (2017). Disappearing routine jobs: Who, how, and why? *Journal of Monetary Economics 91*, 69–87.

Dwyer, R. E. (2013). The care economy? gender, economic restructuring, and job polarization in the us labor market. *American Sociological Review 78*(3), 390–416.

Goos, M., A. Manning, and A. Salomons (2014). Explaining job polarization: Routine-biased technological change and offshoring. *The American Economic Review 104*(8), 2509–2526.

Guvenen, F. and B. Kuruscu (2010). A quantitative analysis of the evolution of the us wage distribution, 1970–2000. *NBER Macroeconomics Annual 24*(1), 227–276.

Guvenen, F. and B. Kuruscu (2012). Understanding the evolution of the us wage distribution: A theoretical analysis. *Journal of the European Economic Association 10*(3), 482–517.

Guvenen, F., B. Kuruscu, and S. Ozkan (2014). Taxation of human capital and wage inequality: A cross-country analysis. *The Review of Economic Studies 81*(2), 818–850.

Havranek, T. (2013). Publication Bias in Measuring Intertemporal Substitution. *Working Papers IES 2013*(15).

Heckman, J. J. (1976). A life-cycle model of earnings, learning, and consumption. *Journal of political economy 84*(4, Part 2), S9–S44.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2016). The allocation of talent and us economic growth.

Huggett, M., G. Ventura, and A. Yaron (2011). Sources of lifetime inequality. *The American Economic Review 101*(7), 2923–2954.

Katz, L. F. and K. M. Murphy (1992). Changes in relative wages, 1963-1987: Supply and demand factors. *The Quarterly Journal of Economics*, 35–78.

Lindenlaub, I. (2017). Sorting multidimensional types: Theory and application. *The Review of Economic Studies 84*(2), 718–789.

McDaniel, C. (2007). Average tax rates on consumption, investment, labor and capital in the OECD 1950-2003.

Sachs, D., A. Tsyvinski, and N. Werquin (2016). Nonlinear tax incidence and optimal taxation in general equilibrium.

Stantcheva, S. (2017). Optimal taxation and human capital policies over the life cycle. *Journal of Political Economy 125*(6), 1931–199.

Teulings, C. N. (2005). Comparative advantage, relative wages, and the accumulation of human capital. *Journal of Political Economy 113*(2), 425–461.

Yamaguchi, S. (2012). Tasks and heterogeneous human capital. *Journal of Labor Economics 30*(1), 1–53.