Wage Inequality and Occupation Polarization: A Dynamic Perspective

Tomás R. Martínez*

July, 2019

Abstract

In this paper, I argue that job polarization has long lasting effects in the U.S. wage structure. I suggest that, by changing the cross-cohort occupational structure, polarization can impact returns to experience and future wages. Firstly, I document that polarization has different impact across workers of different ages and education. Young workers disproportionally moved to low and high wage occupations in comparison to old workers, with significant differences between educational groups. Secondly, I document substantial heterogeneity in the level and growth of the returns to experience by occupation. Using an overlapping generations model with endogenous education and occupational choice, I show that if there exist complementarities between young and old labor, job polarization can affect the returns to experience. Quantitatively, I use the model to estimate the effect of technological and demographic changes in the U.S. wage structure, accounting for the transition path dynamics. During the transition, because of cohort imbalances and occupation switching costs, inequality is higher: college premium can be almost 10% higher than in the steady state and the relative wage of the median with respect to the top occupation is 12% worse. This culminates in a clear policy recommendation: the decrease of occupation switching costs, accelerating the transition and increasing wages of vulnerable groups.

JEL Codes: E24, I26, J24, J31.

Keywords: Job Polarization, Technological Change, Returns to Experience, Skill Premium, Occupational Choice.

*Department of Economics, Universidad Carlos III de Madrid, Calle Madrid 126, 28903 Getafe, Spain. Email: torodrig@eco.uc3m.es. I am grateful to my supervisor Andrés Erosa for his guidance and support. I thank Matthias Kredler, Jan Stuhler and participants of the 2017 Doctoral Workshop of Quantitative Dynamic Economics, the 2017 ENTER Jamboree and UC3M Student Workshop for the helpful comments and suggestions. I acknowledge financial support by grant ECO2014-56676-C2-1-P from the Ministerio de Economía y Competitividad. All errors are my own.

1 Introduction

In the past 40 years, the U.S. labor market has experienced a substantial decrease in the labor share of middle wage occupations. The disappearing of these jobs concurrent with the expansion of the ones with high and low wages, so called "job polarization", has been linked to the expansion of automation technology, as it provided a substitution for routine tasks. In a seminal work, Autor, Levy, and Murmane (2003) suggested that the introduction of computers, a cheap form of automation technology, replaced occupations with higher routine task in the workplace[1]. Hence, routine-intensive occupations, which are typically in the middle of the wage distribution, are displaced as the cost of automation technologies go down.

Concomitant to the phenomena of job polarization, other changes have happened to the U.S. wage and employment structure over this time. A well known fact is the increase in the education premium simultaneously with the supply of college educated workers (Katz and Murphy (1992) and Katz and Autor (1999)). Less known, but similarly important, are the changes of the experience-wage profiles for different education groups and cohorts (Elsby and Shapiro (2012) and Kambourov and Manovskii (2005)). Low educated workers have experienced a flattening in their wage profiles, while college graduates observed a steepening in theirs. Although the first fact has long been connected to skill-biased technological change[2] (SBTC), most explanations of the second fact have lied on the differences of labor supply of different cohorts (as in Jeong et al. (2015)).

In this paper, I argue that changes across cohorts are crucial to understand the inequality dynamics between occupations, education and age groups. Firstly, I show that occupation polarization has different impact across cohorts. Young workers disproportionally moved to low and high-wage occupations in comparison to old workers. In addition, I decompose the polarization for different educational groups. Despite the fact that the phenomena was widespread over all the distribution of education, I point out that low-educated workers have a higher likelihood to move to a lower wage occupation, whereas college educated workers tend to climb to a higher wage occupation. This antagonistic propensity, induced by the polarization, in the absence of other changes, increased the share of college educated workers in high returns to experience occupations, as well the share of non-college educated workers in the low returns to experience ones. Finally,

[1] Another strand of the literature have argued that trade may have played a role in polarization. See Goos et al. (2014) and Autor et al. (2015).

[2] Krusell et al. (2000), Acemoglu (2002).

I document that the returns to experience by occupation is almost entirely increasing at their wage level, meaning that low-wage occupations also tend to have low returns to experience. High-skill occupations have up to 25% higher returns to experience than low skill occupations. Moreover, returns to experience have been increasing in the top of the distribution but decreasing in the bottom, these differences can account to up to 10 p.p.

Thus, providing that there exists complementarities between the type of labor and experience supplied by alternate cohorts, the price of experience will endogenously respond to cross-cohort movements. In fact, since younger cohorts are more educated and are employed in distinct occupations than their older counterparts, the price of experience should increase as observed in the data. Furthermore, changes in returns to experience may potentially affect occupational choice today, and impact the observed cross-sectional occupational polarization.

In light of these facts, I propose a life cycle general equilibrium model with endogenous occupational and educational choice. Workers select an occupation and supply inelastically three type of tasks: abstract, routine and manual. Also, as they get old, they accumulate experience and are able to supply a differentiated type of labor, generating the usual concave life cycle profile of wages. Yet, occupational choice is sticky: only a fraction of workers can switch occupations later in life.

In my model, all the occupations are a composition of the three tasks but in different shares. When a worker chooses an occupation, she is effectively picking a bundle of tasks to supply. This is a departure from the polarization literature that sometimes treats tasks and occupations interchangeable. In the usual polarization models, workers of different ability sort themselves into supply exclusively one task. Biased technological change alters the sorting, pushing more (less) workers to supply abstract and manual (routine) tasks. In my case, routine-biased technical change (RBTC) also affects the price of the three tasks. However, instead of sorting through tasks, RBTC affects sorting through occupations, as it increases the wage of abstract and manual-intensive occupations and decreases the wage of the routine-intensive ones.

Moreover, the model captures the heterogeneity in returns to experience observed in the data. Similarly to Katz and Murphy (1992) and Jeong, Kim, and Manovskii (2015), I embed two type of experience-labor: young and old. Workers supply both types of labor, increasing the amount of old labor supplied and decreasing young labor as they age. The fundamental assumption is that the age supply schedule is task specific. For instance, if the supply schedule is steeper for abstract tasks, then abstract-intensive occupations have steeper wage profiles. This structure not only gives us substantial heterogeneity in the returns to experience by occupation, but also, allows these returns to adjust *endogenously* over the time, when distinct cohorts change their relative supply of tasks. Therefore, as old and young labor are not perfectly substitutes and occupation movement is costly (so older cohorts do not have the same occupational structure than the young ones), job polarization can potentially drives movements in the returns to experience.

As usual in the literature, the key mechanism is the complementarity between different types of education, tasks and experience. Extending the usual nested CES structure, the model is able to jointly explain the changes in the wage structure under the period studied. In addition to the two demand shifters (SBTC and RBTC), I allow various demographic supply shifters, such as the increase of college share, the increase of abstract and manual-intensive occupations (and decrease of routine-intensive occupations), the increase in occupation mobility and the increase in the cohort size (the baby boomers). A major implication of these mechanisms is that cohort effects disappear when demographics adjust, and old and young cohorts have the same education and occupation structure. To capture potential effects from the transition dynamics, the model is calibrated to replicate the 40 years' transition path of the U.S. labor market. Theoretically, I suggest that assuming that the economy is in the steady state produces an upward bias in the value of going to college.

Quantitatively, I show that the transition matters for inequality. During the transition, because old workers do not go back to college, the college premium is exceptionally higher. Similarly, because occupation switching costs are higher for old cohorts, a demand shift biased against routine occupations, shift disproportionally more young workers. In addition, during a transition, as not every worker can switch out from the dying routine occupations, the excess of supply in the these occupations decreases the relative wage even further. To account for this mechanism, I simulate the model to a future steady state, where all cohorts can choose freely their occupation and schooling. In this steady state, college premium is 10% lower than in 2010. Also, the median occupation wage is 22% lower than the 90th occupation and 23% higher than the 10th occupation, in comparison to 34% and 12% in 2010, providing evidence that once the technological transition vanishes, inequality between groups decrease substantially. However, returns to experience responds very little in the new steady state, as in 2010, part of the demographic transition is already completed. The policy recommendation is clear: occupation switching costs drag the transition, preventing workers to move to high paid occupations. Therefore, decreasing these costs is essential to accelerate the transition and increase the wage of vulnerable groups (low educated workers in routine-intensive occupations). The model also allows me to revisit an old question in labor economics: what is the role of demand in shaping inequality between groups in the U.S? To tackle this, I fix the demand shifters and simulate the counterfactual U.S. labor market of the last 40 years. The simulations tell that changes in the demand were the main drivers of inequality. Without these changes, the college premium would have been almost 10 p.p. lower in 2010 than in 1970. This would create a disincentive for workers to go to college, making their share to be 7 p.p. lower. Finally, in the counterfactual world, the difference between the 90th occupation and the median (in terms of average wages) shrinks 7 p.p, while the difference between the median and the 10th occupation stays the same.

I contribute to different branches of the literature. Broadly, I contribute to the literature on the the skill-biased technical change and on the labor market structure. Throughout the last two decades, a wide range of papers have connected technological change with the increase in inequality[3]. Katz and Murphy (1992), Katz and Autor (1999) and others have documented the simultaneously persistent increase in the college premium and in the supply of college educated workers. The accepted explanation for this phenomena is the combination of cheap capital with some form of capital skill complementarities as in Krusell et al. (2000) and Acemoglu (2002). Nevertheless, the early models on SBTC are mute about wage and job polarization, making the literature move to the "routinization" hypothesis initiated by Autor, Levy, and Murmane (2003). I contribute by providing a model that can generate polarization relying on SBTC, unifying both frameworks.

The shift toward the routine-biased technological change have produced an incipient but growing literature. Goos et al. (2009) provide evidence that polarization is a diffused phenomena in developed economies and documented it in Europe. Autor and Dorn (2013) have unified the employment and wage polarization to the growth of low skill services occupation. More recently, some papers aimed to decompose the polarization in different demographic groups. Autor and Dorn (2009) use the age composition of different occupations in different local labor markets to evaluate the employment response of workers of different groups. They found that while educated workers may reallocate up or down in the wage distribution (depending specially on their age), low educated workers are shifting uniformly toward the low payed occupations. Similarly, Cortes et al. (2017) perform an accounting exercise on the decline of routine intensive occupations. They

[3]See Acemoglu and Autor (2011) for a survey.

notice that young and low educated men and young and intermediate educated women lower their propensity to work in routine occupations while increasing their propensity to work in low wage occupations and not participate in the labor market. Regarding the non-routine cognitive occupations, they found that most of the increase in the labor share comes from composition rather propensity changes. This points for the increase in college educated workers, a group with high propensity to work in non-routine cognitive occupations, as one explanation to the right tail of the polarization. Finally, Cerina et al. (2017) decompose occupation polarization by gender and, using a model with endogenous market and production, attribute women polarization to the increase in women labor force participation.

Despite the crescent literature, very few papers have focused on the dynamic effects of occupation polarization[4]. My work aims to fill this gap by providing evidence that occupation polarization has a strong and important cohort component, and that can feedback on the contemporaneous occupational choice. This presents indication that polarization has not only increased inequality over a cross-section, but also potentially over the individual life-cycle.

The rest of this paper proceed as follows. In the next section I discuss the empirical evidence. In Section 3, I present the model. In Section, 4 I discuss the calibration and the fit in the data. Section 5 describes the model dynamics and the counterfactual experiment. Finally, in 7, I conclude.

2 Empirical Evidence

In this section, I present some empirical regularities about the evolution of the wage structure in the U.S. My analysis uses data from the March Current Population Survey (CPS) spanning the period from 1970 to 2012. My sample is restricted to male workers[5], who are not self employed and were full-time and full-year, defined as working 40 weeks or more per year and 35 or more hours per week. I also drop every worker younger than 17 years old and older than 60 years. The resulting sample has 1,067,202 observations over the 1970-2012 period, an average of around 25 thousand per year. To investigate

[4]A notable exceptions is Cortes (2016). He analyses the transition out of routine intensive occupation and its subsequent wage growth. An important result is the observed faster wage growth from the workers who switched out.

[5]Much of our analysis is using potential experience as a proxy of experience. It is a well known fact that this is a poor proxy for women as they tend to interrupt their labor force participation much more frequently than men.

the changes in the occupational structure, I build on the crosswalk of David Dorn[6] that produces a time-consistent occupation categories for U.S. Census and extended whenever it was necessary. For many of the occupation statistics I present, I pool the two previous years and two following years[7] to increase power. All calculations are weighted by their respective survey weight.

2.1 Job Polarization by Age

Fact 1. *Job Polarization is more pronounced in young cohorts. In particular, the difference between cohorts is larger in the lower tail of the distribution.*

A substantial body of the literature have shown that the U.S. labor market has become polarized: the labor share of abstract-tasks intensive occupations and manual-tasks intensive occupations have increased, while routine intensive occupations have disappeared. In this section, I argued that job polarization is higher for young workers. In Figure 1, I show that the "hollowing out" of the distribution is much more pronounced in younger than in older workers. In short, the change in the top and in the bottom of the occupation distribution for workers between 20 and 29 years old in 2010, in comparison with the same group in 1980, is larger than for workers of older age. In fact, for the low-wage occupations there is a clear monotone increasing relationship between the increase in employment and age. While workers of age between 20 and 29 years old have more than double employment in some low wage occupations, workers of between 50 and 59 years old had an increase of less than the 20 p.p.

The evidence of Figure 1 is consistent with the findings of Autor and Dorn (2009). They argue that routine intensive occupations - usually in the middle of the occupation wage distribution - are getting older relatively to non-routine occupations. Moreover, by exploiting cross-section changes in employment shares of local labor markets, they show that the decline in routine occupations is uniformally larger for young than old workers.

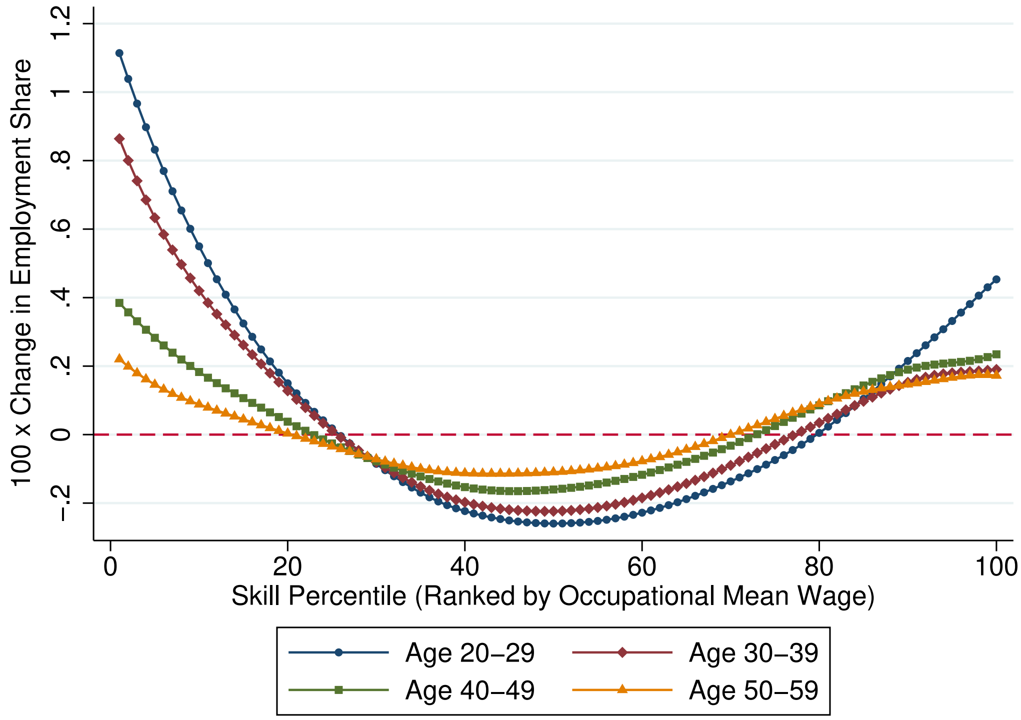
This fact is relatively intuitive. It is well known that young workers are, on average, more mobile than old workers[8]. A simple explanation is that the technological change that displaces routine occupations is a time specific shock that affects all cohorts. However, because occupation specific human capital and other moving costs, workers of different ages can have a very different response. Old workers with high occupation specific human

[6]See Autor and Dorn (2013).

[7]E.g. the occupation statistics for 1980 includes 1978 up to 1982

[8]For instance, see Kambourov and Manovskii (2008).

Figure 1: Smoothed Changes in Employments by Occupational Skill Percentile 1980 to 2010: By Age Groups



Notes: The Occupation Mean Wage is computed by pooling all age groups and years from 1978 to 1982. The change in employment share is computed by taking the difference of occupation shares of pooled years from 1978 to 1982 and the pooled years from 2008 to 2012. This is done by all age groups independently. The distribution is smoothed using a locally weighted regression with bandwidth 0.75.

capital have less incentives to start over and move to a different career. Moreover, because they are closer to retirement age, they will not reap as much benefits from the change as young workers.

2.2 Job Polarization by Education

Fact 2. *The increase in the share of high-wage occupations is accounted almost exclusively by: (i) The increase in the share of high-wage occupations in college educated workers, and (ii) the increase in college share.*

The increase in the number of college educated workers is a widely studied phenomena. In 1970, a bit less than 20% of all working men had a college degree, while, in 2010, the number gets to almost 35%. In this section, I add evidence on how college and non-college workers were affected by polarization and how the increase in the college workers interacts with this phenomena.

In Figure 2, I show the typical job polarization picture, as in Autor and Dorn (2013) and Acemoglu and Autor (2011), disaggregated by college and non-college workers. Panel A exhibits two opposite tendencies: albeit the disappearing of middle-skill occupations happened for both college and non-college workers, the share of non-college workers in lower skill occupations have increased much more relatively to college workers. Not only that, non-college workers are responsible for almost the entire share of low-skill occupations, effectively taking into account most of the polarization in the lower part of the distribution. In the upper part of the distribution, on the other hand, college educated workers have increased their share in high-skill occupations relatively to non-college workers. Moreover, if I disaggregate the non-college workers in high school dropouts, high school graduates and workers with some college, the tendency of low educated workers to move towards low-skill occupations and high educated workers to move to high-skill occupations becomes even more evident. Similar evidence on the heterogeneity in the movement of workers with different schooling level is discussed in Autor and Dorn (2009) and Cortes et al. (2017). My results confirm theirs at the 3-digits occupational level.

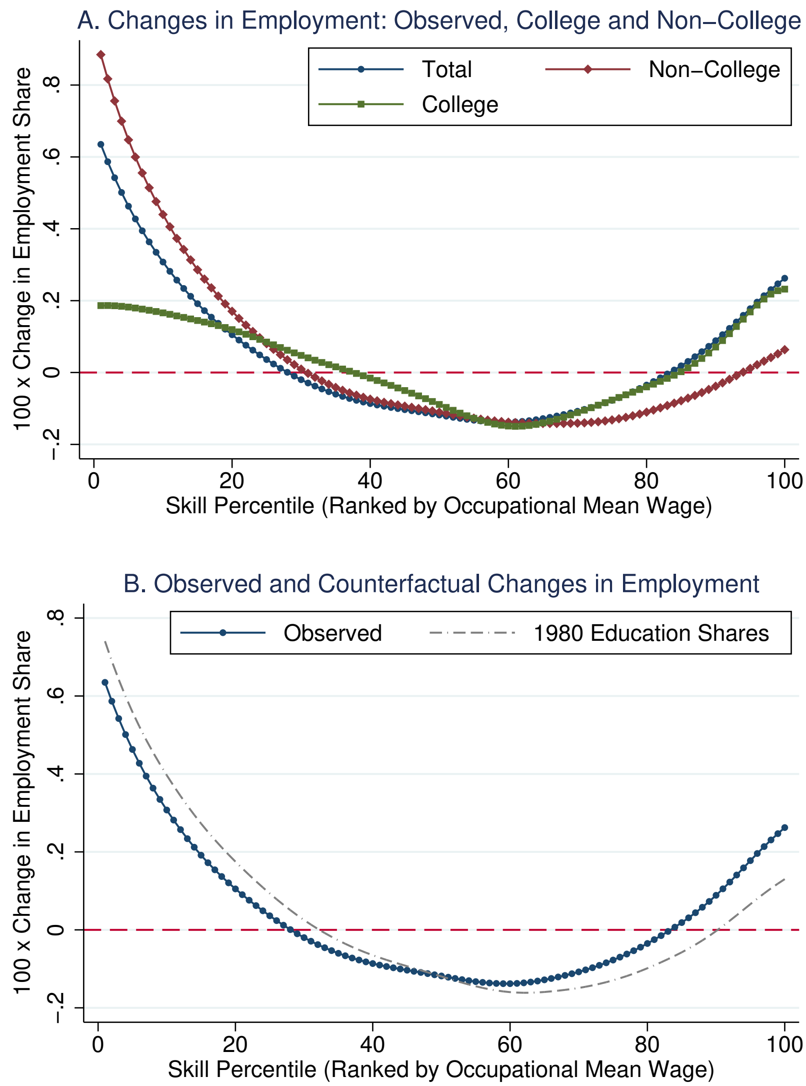
Panel A provides evidence that college and non-college have contributed to the polarization differently. This picture, however, is incomplete, since the U.S. population is much more educated in 2010 than in 1980. In panel B, using standard decomposition methods[9], I show how the polarization graph for men would look like if we maintain the shares of high school dropouts, high school graduates, workers with some college and college graduate unchanged at the 1980 level. The counterfactual distribution tells us that the increase of years of education accounted for some of the increase in the share of workers in high-skill occupations (up to a half in some parts of the distribution). The intuition for this result is clear, as some high-skill, high-wage occupations such as physicians, engineers, and others can only be performed with more advanced studies. This support the view that job polarization and the increased in the share of college workers are different sides of a related phenomena.

The increase in the college share has an obvious cohort component[10]. College decision

[9]I have used methods described in details in Fortin et al. (2011) and in Dinardo et al. (1996). Specifically, I re-weight our sample using the prediction of a logit regression for the probability that an individual belongs to the 1980-82 sample on education dummies.

[10]Figure A.2 presents the college share by labor market entry year.

Figure 2: Smoothed Changes in Employments by Occupational Skill Percentile 1980 to 2010: By Educational Groups and Constant 1980 Education Shares



Notes: panel A plots changes in employment share by 1980 occupation percentile wage rank for all the workers, college only and non-college only. Panel B plots changes in employment share fixing the share of high school dropouts, high school graduates, workers with some college and college graduate at the 1980 level. To calculate the shares and wages of 1980, I pool all years from 1978 from 1982. In 2010, I pool all years from 2008 from 2012. The distribution is smoothed using a locally weighted regression with bandwidth 0.75.

usually happens before labor market entry, and few workers pursuit college education later in life. Thus, in the case that college educated workers respond differently to changes in technology, demand shifters and polarization may present cohort effects. Furthermore, these effects would be present in a demographic transition, only ceasing to exist after the differences across generations disappear.

2.3 Returns to Experience by Occupation

Fact 3. *Returns to experience by occupation is increasing in their wage level. Furthermore, returns to experience of high-wage occupations have increased relatively to low-wage occupations.*

In this section, I document a new fact: the returns to experience by occupation is almost entirely increasing in their average wage level. This means that the average wage growth within an occupation is positively correlated with the rank of that occupation in the wage distribution. This fact has an important implication: providing that the average initial wage level is positive correlated with their average wage level and the individual only switches to close occupations (in terms of the position of the occupation in the wage distribution), the differences in returns to experience across occupations can account to some of the increase of the earning inequality over the life-cycle.

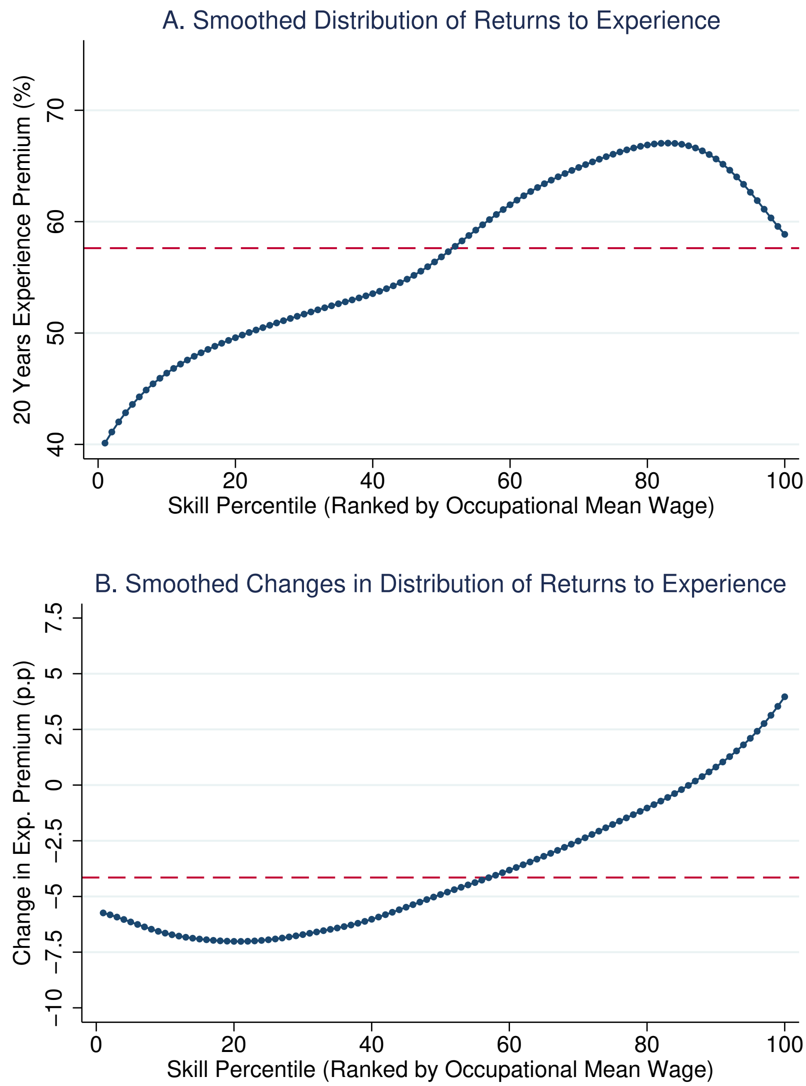
Figure 3 plots the distribution of returns to 20 years of potential experience by occupation at the 3-digit level and the change between the distribution of 1980 and 2010. In Panel A, I show that high-skill occupations have up to 25% higher returns to experience than low-skill occupations, and the distribution is almost entirely increasing, as only after the 95th percentile the returns to experience begins to decrease. Panel B tells that this underline distribution, over the years, has become even more steeper, as the returns to experience have increased more for high-wage occupations. In fact, returns to experience of low-skill occupations have decreased almost 7.5 p.p.

One could argue that during a transition the cross-section distribution of the returns to potential experience might not be the same as the ones truly experienced by a given cohort. In Figure A.4, I show that, using synthetic cohorts, the pattern remains robust. Low-skill occupations have low returns to experience while high skill ones have high returns. Similarly, Panel B tells that the returns to experience in a given cohort are increasing in the high skill occupations with respect to the low skill ones.[11].

It is well established that there exists non-negligible occupation specific human capital (Kambourov and Manovskii (2008) and Huckfeldt (2018)). But, at which extent the

[11]We also point that differences in the returns to experience within a cohort were made between the distribution of 1980 and 1990. Unfortunately we cannot observe the entire cohort of the individuals entering the labor market in 2000s, therefore underestimating the returns of experience for that cohort.

Figure 3: Smoothed Distribution of Returns to Experience and the Changes from 1980 to 2010 in Returns to Experience by Occupation



Notes: The distribution of returns to experience is constructed from a 20 years of experience prediction using occupation specific coefficients of a regression of log wage on a quadratic polynomial of potential experience, a full set of education dummies and a full set of year fixed effects. The distribution is smoothed using a locally weighted regression with bandwidth 0.75. Occupations are ranked by their mean wage in 1980. The change is calculated taking the difference between the distribution of 1970-82 and 2000-12.

returns to tenure differs among occupations? The results above provide evidence that the difference can be substantial. A natural first step is to ask why this is the case. One occupation is composed by a set of tasks regularly performed by the worker. I believe that complex-task intense occupations have steeper learning curves and therefore high returns to experience. Table 1 shows the correlation between the returns to experience and the task intensity indexes used in Acemoglu and Autor (2011). Not surprisingly, occupations intense in abstracts cognitive tasks tend to have higher returns to experience. Routine cognitive tasks appear to have close to zero correlation with returns to experience, while occupations intense in physical tasks, routine and non-routine, have low returns to tenure[12].

Table 1: Correlation between Returns to Experience and Task Indexes

Task Index	Correlation
Non-routine Cognitive: Analytical	0.296^{**}
Non-routine Cognitive: Interpersonal	0.177^{**}
Routine Cognitive	0.022
Routine Manual	-0.155^{**}
Non-routine Manual: Physical	-0.164^{**}

Notes: Returns to experience is constructed from a 20 years of experience prediction using occupation specific coefficients of a regression of log wage on a quadratic polynomial of potential experience, a full set of education dummies and a full set of year fixed effects. The task index are composite measures of different O*NET Work Activities Scale, see the Appendix for details. I drop occupations with negative returns and returns higher than 100%, totaling 321 occupations. $**p < 0.01$.

Moreover, as I have shown before, during the last 40 years, the difference between cohorts have accentuated, as young cohorts tend to be more educated and work on high skill or low skill occupations. In the case there exists complementarities between different age groups (such as in Katz and Murphy (1992), Card and Lemieux (2001) and Jeong et al. (2015)), changes in the relative supplies of cohorts can endogenously change returns to experience. In the case of an increase in the relative supply of young educated workers in manual and abstract intensive occupations, one should observe a increase in the price of experience exactly for these groups. In Panel B of Figure 3, I show that the indeed there is an increase in the price of experience in high skill occupations. In low skill occupations, however, there has been a decrease in the price of experience, which is inconsistent with an increase in the relative supply of young workers in manual occupation but since it interacts

[12] Adda et al. (2017) make a similar argument, pointing that abstract occupations have high wage losses, in comparison to routine occupations, after a work interruption of 1-3 years.

non-trivially with the increase non-college workers, one has to be careful to analyze which effects dominate.

Outside of the degree of complexity of the occupations and the relative cohort supply, other factors can generate differences in the returns to experience. Providing there exist non-transferable occupation specific human capital, higher occupation switching among low skill workers might undermine the returns to occupations in low skill occupations. A similar argument can be made even if the switching patterns are the same for high skill and low skill workers, but the degree of *human capital transferability* differs for some occupations. Another concern is the use of potential experience as a proxy of actual experience. If workers of low skill occupations on average spend more time unemployed, then this proxy will underestimate the returns of experience of low skill occupation. These explanations are out of the scope of this paper, but are an interesting avenue for future research.

3 Model

I study simultaneously the effect of the change in the college premium, the experience premium and occupation polarization in different cohorts. In particularly, I am interested on how the polarization, induced by technological change, affects the wage structure of workers of different skills and cohorts. For that I develop an overlapping generations model with endogenous occupation and education choice where occupations supply different tasks. The model builds on the class of models that study how demand shifts (such as SBTC and RBTC) affect the wage structure exploiting the complementarities of different types of labor. The classical example is the increasing of the college premium studied by Katz and Autor (1999) and Krusell et al. (2000). Nevertheless, the same methodology was used to study the change of returns to experience (Katz and Murphy (1992) and Jeong et al. (2015)) and the wage polarization (Autor and Dorn (2013) and Cortes (2016)). A limitation of some of these studies is that it abstracts from the worker side and focus on the structure of the production function - usually a nested CES - to derive log linear forms that can be easily estimable as a linear regression. Another limitation is the focus on the cross-section and the static decisions. I extended the supply side along two dimensions. First, I incorporate a Roy-type of occupation decision similar to Hsieh et al. (2018) and an educational choice. Second, I model the life-cycle with different cohort size. This gives the possibility to analyze not only demand shifts but also supply shifts such as the increasing of cohort size given by the incorporation of baby boomers in the labor market and changes in the educational and occupational structure of different generations.

3.1 Environment

The economy is populated by overlapping generations of workers living H periods. I refer t to calendar time and τ to the cohort of worker[13]. Every generation comprises a continuum of workers with size $n(\tau)$. In an initial pre-period, workers draw a college cost $\kappa \sim H_t(\kappa)$ from a time-variant distribution and decide to be a college type or not, by comparing their lifetime expected utility. After their education decision is made, they draw, from an education and time specific Frechet distribution, a set of occupation abilities $\epsilon \equiv (\epsilon_1, \epsilon_2, ..., \epsilon_M)$ and decide an occupation $j \in (1, 2, ..., M)$ to enter the labor market.

Every period a worker of age h, education type e and a vector of occupation abilities ϵ supplies fully her labor to occupation j. Occupation is sticky: by the end of the period the worker has to stay in the same occupation with probability $\beta_{te}(h)$, otherwise she is free to re-optimize and choose another occupation[14]. Utility is linear and there are no savings, thus the per period value of a worker is defined as:

$$W_t(h, j, \epsilon, e) = \tilde{w}_{te}(h, j)\epsilon_j + \beta_{te}(h)W_{t+1}(h+1, j, \epsilon, e) + (1 - \beta_{te}(h))\mathbb{E}_\epsilon[\max_{j'} \left\{W_{t+1}(h+1, j', \epsilon', e)\right\}] \qquad \text{for } h = 1, ..., H-1 \tag{1}$$

$$W_t(h, j, \epsilon, e) = \tilde{w}_{te}(h, j)\epsilon_j \qquad \text{for } h = H \tag{2}$$

where $\tilde{w}_{te}(h, j)\epsilon_j$ is the wage of a worker of age h, that chose education e, occupation j, with ability ϵ at time t and is a composition of his own ability and the occupation efficient wage. Conditional on the age and education, the occupation efficient wage is a composition on the market price of three tasks: Abstract, Routine and Manual[15]:

[13] Since, I am not modeling the entire life-cycle in the strict sense, cohort here means labor market cohort. That is, the year the worker enter the labor market.

[14] This encompass many sources of frictions, from occupation specific human capital to moving frictions.

[15] One interpretation is that occupations follow a hedonic pricing, where their wage reflect the tasks performed.

$$\tilde{w}_{te}(h,j) = a_j^A w_{te}^A(h) + a_j^R w_{te}^R(h) + a_j^M w_{te}^M(h) \tag{3}$$
$$a^A + a^R + a^M = 1$$
$$a^i \in [0,1] \quad \text{for } i \in (A,R,M)$$

where (a_j^A, a_j^R, a_j^M) represent the occupation-specific composition shares. The separation between Abstract, Routine and Manual tasks are not unusual in the polarization literature[16]. Where I depart from the literature is how the price of tasks is mapped to the final wage. Most of the models in the literature use tasks and occupation interchangeable. Their main idea is that workers sort themselves according some skill distribution and supply exclusively one type of tasks, e.g. high-skill workers supply only abstract tasks, whereas low-skill supply only manual tasks. In my model, all workers supply all three tasks (except in the extreme case where the composition share is 0). Nevertheless, some occupations supply more of one type of task than others. Thus, whenever the price of a task increases (or decreases) the overall wage paid by all the occupations increases (or decreases) by their own share of that task.

Returns to experience is modelled by following Katz and Murphy (1992) and Jeong et al. (2015). I assume two types of labor, young and old, which are supplied during the workers life-cycle. Younger workers supply relatively more young labor, and as they age, the relative amount of old labor they supply increases. The relative supply follows a task and education specific function. Hence the price of task i for a worker of age h is given by:

$$w_{te}^i(h) = \eta_{iey}(h) \cdot w_{tey}^i + \eta_{ieo}(h) \cdot w_{teo}^i \tag{4}$$

where w_{tey}^i and w_{teo}^i are education and time varying task prices for young and old labor, respectively.

[16] Autor and Dorn (2013), Acemoglu and Autor (2011), Cortes et al. (2017) and Cortes (2016) use a similar, or sometimes slightly more detailed, definition of tasks.

3.2 Educational Choice

A worker decides her education by comparing the lifetime expected utility of being a college type and paying the college tuition[17]. Let the value of a worker born in cohort $\tau = t - h + 1$ with education e be:

$$V(e, \tau) = \mathbb{E}_{\epsilon} \left[\max_{j \in (1, \ldots, M)} \left\{ W_{\tau}(1, j, \epsilon, e) \right\} \right]. \tag{5}$$

Then, the share of workers that decide to be a college type is defined by a simply cutoff rule:

$$\begin{aligned} e(\kappa, \tau) &= C \qquad \text{if} \quad V(C, \tau) - \kappa \geq V(NC, \tau), \\ e(\kappa, \tau) &= NC \qquad \text{otherwise.} \end{aligned} \tag{6}$$

3.3 Occupational Choice

Occupational decision draws from Hsieh et al. (2018). Once the workers decide their own education, they draw a vector of occupation specific ability ϵ from a multivariate Frechet distribution with shape θ and time-age-education-occupation varying scale π_{thej}:

$$F_{the}(\epsilon_1, \epsilon_2, \ldots, \epsilon_M) = \exp \left\{ - \sum_{j=1}^{M} \pi_{thej} \epsilon_j^{-\theta} \right\}. \tag{7}$$

The parameter θ controls the dispersion of ability, as higher parameter values are associated with smaller dispersion. The scale π_{thej} governs the size of ϵ, with higher values associated with a higher probability of drawing a high value of ϵ_j for some occupation j. I assume that the draws of each ϵ_j are uncorrelated.

After the worker has drawn the set of abilities, the problem boils down to choose the occupation that yields the maximum lifetime income. Given that with probability β occupation is sticky, the occupation decision depends not only on contemporaneous wages but also on future wages discounted by beta. This is true for every occupation

[17]The college cost κ includes not only tuition, but also any other non-pecuniary cost that the agent may incur when being a college type.

decision faced by the worker in her lifetime. Exploiting the properties of the extreme value distribution, I denote the share of workers that choose occupation j using the following proposition:

Proposition 1. *Without loss of generality consider an arbitrary time and education type. Let* $\epsilon_j \hat{w}(h,j) = \epsilon_j \sum_{k=0}^{H-h} \prod_{n=1}^{k} \beta(h+k-n)\tilde{w}(h+k,j)$ *be the discounted sequence of future wages in occupation* j*, where* $\tilde{w}(h,j)$ *is the wage paid in efficient units for a worker of age* h *and occupation* j *and* ϵ_j *the ability of the worker in occupation* j*.*

1. *The probability that a worker at age* $h-1$ *chooses occupation* j *is given by:*

$$p_j(h) = \frac{\pi_{hj}\hat{w}(h,j)^{\theta}}{\sum_s \pi_{hs}\hat{w}(h,s)^{\theta}} \tag{8}$$

2. *The share of occupation* j *at age* h *is given by:*

$$\begin{aligned} s_j(1) &= p_j(1) \\ s_j(h) &= \beta(h-1)s_j(h-1) + (1-\beta(h-1))p_j(h) && h = 2,..,H \\ s_j(h) &= p_j(1)\prod_{k=1}^{h-1}\beta(k) + \sum_{k=2}^{h}(1-\beta(k-1))p_j(k)\prod_{r=1}^{h-k}\beta(h-r) && h = 1,..,H \end{aligned} \tag{9}$$

3. *The average ability of workers of occupation* j *at age* h*:*

$$E(\epsilon|h,j) = \overline{E(\epsilon|1,j)}\prod_{k=1}^{h-1}\beta(k) + \sum_{k=2}^{h}(1-\beta(k-1))\overline{E(\epsilon|k,j)}\prod_{r=1}^{h-k}\beta(h-r) \quad h = 1,..,H \tag{10}$$

where $\overline{E(\epsilon|h,j)} = \left(\frac{\pi_{hj}}{p_j(h)}\right)^{1/\theta}\Gamma\left(1-\frac{1}{\theta}\right)$ *and* $\Gamma(.)$ *represents the Gamma function.*

Proposition 1 states that the probability the worker chooses an occupation j is a function of the discounted lifetime wage paid by that occupation and the scale of the distribution π_{hj}. Intuitively, it is easy to see that if occupation mobility is too small, the likelihood the worker is still working in the chosen occupation is higher and future wages have a higher weight in the contemporaneous occupation decision. Obviously, the higher is the scale factor π_{hj}, the higher is the probability the worker has high ability and therefore chooses occupation j.

The second part of the proposition defines the occupation employment shares of the economy conditional by age. The occupation share of age h is a function of the occupation structure in age $h-1$ and the fraction $(1-\beta)$ of workers who were allowed to re-optimize and move to a different occupation. Since, β does not depend on j and the law of large number applies, the share of workers who were free to choose a new occupation is again given by equation (8), making the occupation share of a given cohort a function of previous occupational decisions.

This equation indicates how sluggish the economy adjusts. Even if demand shifters are putting upward pressure on the wage of some occupations, because of occupation stickiness, only a fraction of old workers switch. The economy will gradually adjust once new cohorts enter the labor market and the old retire.

Lastly, I characterize the average ability of workers by age and occupation. The relationship between the occupation probabilities and the average ability is clear: if the probability of choosing a given occupation j is low (with respect to its scale), only workers with relatively high ability will choose it, increasing the average ability. Then, equation (10) summarizes the average ability in an occupation for a given given cohort. Equivalently to equation (9), the average ability will be the weighted average of the ability of workers that choose their occupation today and the average ability of previous choices.

Intuitively, Proposition 1 allows to solve the value function (1) for a large arbitrary number of occupations without dealing with the curse of dimensionality. The key point is that once I know the occupation wage of every age and education, it is relatively trivial to compute the worker's lifetime income and retrieve the occupation shares and the value function. The underline assumption behind this result is that ability and the occupation wage are multiplicative separable, and occupational choice only affects the continuation value through wages. If this was not the case (for instance, if β depended on j or ϵ), I would not be able to separate the occupation lifetime income and worker's ability.

3.4 Technology

Technology is standard. I follow the literature to include demand shifters in a series of nested CES production function. The economy produces a final good $Y(t)$, given by a representative firm that takes two intermediary inputs, services and goods:

$$Y(t) = \Lambda(t)S(t)^{\alpha}G(t)^{1-\alpha}, \tag{11}$$

where $\alpha \in (0,1)$ is the income share of services with respect to goods. Production of both goods and services are given by two perfectly competitive representative firm. The production technology of goods requires both abstract and routine tasks and is given by:

$$G(t) = \left(A(t)^{1-1/\gamma} + (\Lambda^R(t)R(t))^{1-1/\gamma}\right)^{(\gamma/(\gamma-1))} \tag{12}$$

with the elasticity of substitution between tasks $\gamma > 0$ and the routine-biased demand shifter $\Lambda^R(t)$. Denote $L_{i,C}$ the total labor supplied by college workers and $L_{i,NC}$ the total labor supplied by non-college workers to task $i \in \{A, R, M\}$. The intermediates input task $A(t)$ and $R(t)$ are produced given another CES aggregator:

$$A(t) = \left(\Lambda^S(t)L_{A,C}(t)^{1-1/\rho_a} + (1-\Lambda^S(t))L_{A,NC}(t)^{1-1/\rho_a}\right)^{(\rho_a/(\rho_a-1))} \tag{13}$$

$$R(t) = \left(\Lambda^S(t)L_{R,C}(t)^{1-1/\rho_r} + (1-\Lambda^S(t))L_{R,NC}(t)^{1-1/\rho_r}\right)^{(\rho_r/(\rho_r-1))}, \tag{14}$$

where $\rho_i > 0$ is the elasticity of substitution between college and non-college labor and $\Lambda^S(t) \in (0,1)$ the college-biased demand shifter. Low-skill services is produced using only the total units of the manual task $M(t)$:

$$S(t) = M(t) = L_{M,C}(t) + L_{M,NC}(t), \tag{15}$$

in which for simplicity, I assume that college and non-college labor are perfectly substitutes in the production of manual tasks. Finally, I aggregate total young labor L_{ieY} with total old labor L_{ieO}, for each task-education pair:

$$L_{ie}(t) = \left(L_{ieY}(t)^{1-1/\delta_{ie}} + \omega_{ie}L_{ieO}(t)^{1-1/\delta_{ie}}\right)^{(\delta_{ie}/(\delta_{ie}-1))}, \tag{16}$$

where $\delta_{ie} > 0$ is the elasticity of substitution between young and old labor and the relative share $\omega_{ie} > 0$.

3.5 Equilibrium

Let the aggregate supply of labor for every triple of task $i \in (A, R, M)$, education $e \in (C, NC)$ and labor experience $x \in (y, o)$ be:

$$L_{iex}(t) = \sum_{j}^{M} a_j^i \sum_{h=1}^{H} s_{tej}(h) e(t-h+1) n(t-h+1) \eta_{iex}(h) \mathbb{E}\left[\epsilon_j | t, e, h, j\right], \tag{17}$$

where the internal sum aggregates over cohort-varying supply shifters such as: occupation shares $s_{tej}(h)$, education shares $e(t-h+1)$, cohort size $n(t-h+1)$, young/old relative supplies $\eta_{iex}(h)$ and average ability $\mathbb{E}\left[\epsilon_j | t, e, h, j\right]$. The outside sum, using the task composition a_j^i, aggregates over all the occupations in the economy. In the appendix, I describe carefully the algorithm used to solve numerically the model. Formally, equilibrium is defined:

Definition 1. *An equilibrium is a set of allocations for individuals:* $\{\{\{e(\kappa, \tau), j\}_{t=0}^{\infty}\}_{\kappa \in H}\}_{\epsilon \in F}$, *allocations for firms:* $\{\{L_{i,C,Y}(t), L_{i,C,O}(t), L_{i,NC,Y}(t), L_{i,NC,O}(t)\}_{i \in (A,R,M)}\}_{t=0}^{\infty}$ *and prices:* $\{\{w_{t,C,Y}^i, w_{t,C,O}^i, w_{t,NC,Y}^i, w_{t,NC,O}^i\}_{i \in (A,R,M)}\}_{t=0}^{\infty}$ *such that:*

1. *Given prices, the allocations maximizes the lifetime utility of the workers.*

2. *Given prices, the allocations for the firms maximizes their profits.*

3. *Market clearing: Equation (17) clears the market for all t and i.*

4 Quantitative Analysis

4.1 Calibration

Most of the literature attribute the increasing of the college premium from the end of 1970s onwards. Equivalently, the job market begun to polarize at 1980. I calibrate my model from 1970 to 2010 to take into account for the entire transition of these 40 years. I decide that a cohort "lives" for 35 years ($H = 35$). The model is solved with $M = 263$. The number of occupations M reflects the 3-digit code from the CPS used in the empirical analysis. Finally, to increase statistical power and reduce computational burden, I solve the model in intervals of 5 years. To facilitate exposure, it is useful to separate the data inference in three groups: parameters externally calibrated without solving the model, parameters internally calibrate while solving the model and estimated parameters to match selected moments.

4.1.1 Parameters calibrated without solving the model

The first step is to choose a set of parameters to calibrate before solving the model. Firstly, I follow Hsieh et al. (2018) and use their θ. They use the shape of the extreme value distribution to match the dispersion of the residuals from a regression of log wages on occupation-age-education dummies. Secondly, the cohort size $n(\tau)$ is calibrated given the population growth weighted by the labor force participation of prime age men[18]. I normalize the generation entering the labor market in 1970 to 1.

To get the occupation composition shares (a_j^A, a_j^R, a_j^M) I draw heavily from the ONET (Occupation Information Network). The initial step is to define measures of Abstract, Routine and Manual tasks. I follow most of the polarization literature and use a composition of tasks to create the measures, in particular I use the same composition as Acemoglu and Autor (2011). Then, I standardize them assuming they follow a uniform distribution with support $U(0,1)$, such that the occupation with the lowest measure gets 0 and the highest gets a 1. Finally, I divide each task by the sum of all three, normalizing the sum of the compositions to one for all occupations.

To calibrate the occupation stickiness $\beta(h)_{et}$, I use information on the rising occupational mobility in the U.S. Using the PSID, Kambourov and Manovskii (2008) documents the increase of occupational mobility by age and educational group by 3-digit occupations conditional on a wide range of characteristics. I use the estimated parameters of their probit specifications to calibrate β by time, age and education.

To alleviate the computational cost in the estimation of the demand shifters, I decide to estimate the parameters regarding the supply of young and old labor and the elasticity of substitution of their CES outside of the model. This approach shares many similitude to Katz and Murphy (1992). First, I use the task composition share to get the contribution of each task to the final occupation wage. Then, I estimate $\eta_{iey}(h)$ and $\eta_{ieo}(h)$ by projecting worker's wages on age interactions with the average wage at no experience and the wage at 35 years of experience. The identifying assumption is that at no experience, workers only supply young labor, while at the oldest age they supply exclusively old labor. In all

[18]Technically, as I do not observe the two youngest cohort when they are prime age, I assume they have the same size as the cohort "born" in 2000. This does change the results.

the other age bins they supply a bundle of young and old labor. In summary:

$$w_{ietz} = \sum_{h=1}^{H} \eta_{iey}(h)\overline{w}_{iet}^{0}\mathbb{1}_{h=exp_z} + \sum_{h=1}^{H} \eta_{ieo}(h)\overline{w}_{iet}^{35}\mathbb{1}_{h=exp_z}. \tag{18}$$

Where w_{ietz} denotes the part of the wage (not in logs) given by task i, of worker z at time t[19]. The binary indicator $\mathbb{1}_{h=exp_z}$ takes the value 1 when worker z has experience h. Then, I recover the set of parameters $(\delta_{ie}, \omega_{ie})$ by exploiting the structure of the production function, an usual approach in the literature. Specifically, the CES implies that the ratio of the Marginal Product of Labor is of the form:

$$\frac{w_o}{w_y} = \omega \left(\frac{L_y}{L_o}\right)^{1/\delta}. \tag{19}$$

Taking logs from both sides and using the supply schedules η_{iey} and n_{ieo} to compute L_y and L_o, a linear regression yields the parameters δ and ω.

Table 2: Estimated Parameters - Experience and Task Production Function

Parameter	**Description**	**Value**
	College	
$\omega_{A,C}$	Relative share of young and old labor: abstract	1.7138
$\delta_{A,C}$	Elasticity of substitution between young and old labor: abstract	61.1053
$\omega_{R,C}$	Relative share of young and old labor: routine	1.6373
$\delta_{R,C}$	Elasticity of substitution between young and old labor: routine	16.8867
$\omega_{M,C}$	Relative share of young and old labor: manual	1.5719
$\delta_{M,C}$	Elasticity of substitution between young and old labor: manual	10.8269
	Non-college	
$\omega_{A,NC}$	Relative share of young and old labor: abstract	1.5313
$\delta_{A,NC}$	Elasticity of substitution between young and old labor: abstract	9.4371
$\omega_{R,NC}$	Relative share of young and old labor: routine	1.5980
$\delta_{R,NC}$	Elasticity of substitution between young and old labor: routine	13.8734
$\omega_{M,NC}$	Relative share of young and old labor: manual	1.6164
$\delta_{M,NC}$	Elasticity of substitution between young and old labor: manual	19.4850

Table 2 presents the estimated parameters for the experience-education-task produc-

[19] In practice, I estimate this equation for every pair i and e.

tion function. At first glance, I acknowledge a couple of details. The relative share of young and old labor ω identifies the fixed experience premium of every task and is, in general, higher for college educated than for non-college workers. Moreover, the elasticity of substitution is lower for non-college workers. This means that for non-college workers, experience and young labor are less substitutes than for college workers. This masks substantial heterogeneity within tasks. The production function of abstract tasks for college workers is almost linear, implying perfect substitution between young and old labor. Nevertheless, for non-college workers, abstract task has the lowest elasticity of substitution. The elasticities of substitution have important implications for the model. The highest is the elasticity of substitution, the lowest impact cohort effects has on returns to experience. The results tell that occupations with high share of non-college workers will respond more to changes that affect asymmetrically different cohorts.

4.1.2 Parameters calibrated solving the model

While the model is solved, I match exactly the education shares and the occupation structure of workers of every cohort. To match the college shares, I follow Heathcote et al. (2010). I assume that the distribution of college cost is log normal $\kappa \sim \log N(\mu_t, \sigma)$, with time-varying mean. The variance of the distribution is pin down by assuming that the mean is the same in both the initial and the ending steady state. Accordingly, the time varying mean matches exactly the share of college workers by cohort.

Then, I use π_{hj} to match the occupation employment share[20] conditional on the calibrated $\beta(h)$. To do so, I first compute the model lifetime income $\hat{w}(h, j)$ for all occupations of a given cohort, age and education. Then, I fix $s_j(h)$ to match the data exactly. For the youngest worker, the occupation employment share is equal to the probability of choosing the occupation, $s_j(1) = p_j(1)$. Thus, to recover π_{1j}, I use $\hat{w}(1, j)$ implied by the model and invert equation 8[21]. For the subsequent ages, I use the calibrated $\beta(h)$ and the employment share of the previous age group $s_j(h-1)$ to proceed recursively and identify $p_j(h)$. Then, again, I use $\hat{w}(h, j)$ and invert equation 8 to get the π_{hj}. Doing that for every age-education-time pair, fully identifies π_{thej}.

Table 3 summarizes the calibrated parameters and their respective targets. Notice that the μ, σ and π are used to capture any other unobserved factor outside supply and demand that might have changed the shares of college workers and occupations.

[20]Time and education subscript omitted.

[21]In practice, since I have to find all π_j simultaneously, I solve a linear system of M unknowns.

Table 3: Calibrated Parameters

Parameter	Description	Target / Source
a_j^i	Occupation composition shares	Task intensity (ONET)
$\beta(h)_{et}$	Prob. of re-choosing occupation	Occ. mobility (Kambourov and Manovskii (2008))
θ	Shape of ability distribution	Hsieh et al. (2018)
μ_t	Mean of college cost	College share by cohort (CPS)
σ	Std. deviation of college cost	College share by cohort (CPS)
π_{jet}	Mean of ability distribution	Occupation employment share (CPS)
$n(\tau)$	Cohort size	Pop. size and labor force participation (CPS)
$\eta_{iey}(h)$	Units of young labor supply	Estimated from CPS (see text)
$\eta_{ieo}(h)$	Units of old labor supply	Estimated from CPS (see text)

4.1.3 Parameters estimated

Established the calibration, I focus the attention on the remaining parameters. The following list is left to be estimated: $\chi = \left(\alpha, \gamma, \rho_a, \rho_r, \Lambda(t), \Lambda^R(t), \Lambda^S(t)\right)$, which consists of the elasticity of substitutions and shares between education types, tasks, and goods and services. To be parsimonious, I approximate the demand shifters $\Lambda^R(t)$ and $\Lambda^S(t)$ by a second degree polynomial, and $\Lambda(t)$ by a linear trend. Finally, I construct a measure between the distance of the moments generated by the model and their counterparts in the data, and fit the parameters to minimize this distance.

To estimate the model, I solve it assuming that 1970 is a steady state (e.g. agents believe prices will not change), and from 1975 until 2010 let the agents fully internalize in their contemporaneous decision changes in future prices. Once the model reaches 2010, I assume the economy is a new steady state, meaning that prices stay constant forever. After I solve the model, I simulate it and target the following moments from data: the average log wage of college and non-college workers; the correlation of the average wage by occupations with their abstract, routine and manual task content; and the wage polarization by percentile from 1970 to 2010;

Despite the parameters are chosen simultaneously to match the data, some are closely related to a particular target. An increase in the demand shift $\Lambda^R(t)$ over the time combined by gross complementarity between abstract and routine task ($\gamma \in (0,1)$), increases the relative demand of abstract to routine tasks. The share of services α, pin down the pass-through of the increase in the relative price of services (manual tasks) to goods, when there is a increase in G. The college-biased demand shifter $\Lambda^S(t)$ matches the changes in the college premium over time. The elasticity of substitution between college and non-college labor for abstract and routine task (ρ_a, ρ_r), jointly with $\Lambda^S(t)$, match the response of the college premium in both abstract and routine intensive occupations.

Table 4: Estimated Parameters

Parameter	Description	Value
α	Income share of services	0.7148
γ	Elasticity of substitution between abstract and routine tasks	0.4003
ρ_a	Elasticity of substitution between college and non-college in abstract tasks	1.5171
ρ_r	Elasticity of substitution between college and non-college in routine tasks	1.1759
Λ_0	Neutral technological change (degree 0 coefficient)	2.4506
Λ_1	Neutral technological change (degree 1 coefficient)	-0.04949
Λ_0^s	College biased technological change (degree 0 coefficient)	19.6069
Λ_1^s	College biased technological change (degree 1 coefficient)	2.9451
Λ_2^s	College biased technological change (degree 2 coefficient)	0.1768
Λ_0^r	Routine biased technological change (degree 0 coefficient)	11.9931
Λ_1^r	Routine biased technological change (degree 1 coefficient)	0.1088
Λ_2^r	Routine biased technological change (degree 2 coefficient)	0.0199

Notes: Polynomial of the college biased technological progress is input in a logistic function to bound the share between 0 and 1. Targets used in the estimation are average log wage of both college and non-college workers, the correlation between the average wage across occupations and their abstract, routine and manual task content, and the wage polarization by percentile from 1970 to 2010.

Table 4 summarizes the estimated parameters. I focus in a couple of key details. First, it is not surprising that both the college-biased technological change $\Lambda^S(t)$ and the routine-biased technological change $\Lambda^R(t)$ are increasing over time, given that the increase in college premium and the wage polarization were used as target. The neutral technological change $\Lambda(t)$, however, has decreased over the time period to compensate for the decline of around 6% of average real wage of the median occupation. Second, the elasticity of substitution γ is less than 1, implying gross complementarity between abstract and routine tasks. Finally, the elasticities of substitution between college and non-college are both larger than 1, being the elasticity of the abstract production function larger than the routine one. The reason is twofold: first, the usual complementarity in high and low-skilled workers is already nested by the abstract and routine tasks. Thus, within the production of every task, college and non-college are gross substitute. Second, given the production technology of the model, there exists, effectively, one college premium for abstract and another for routine tasks. Since the demand shift $\Lambda^s(t)$ is the same in both tasks, the difference in ρ captures the heterogeneous response of the two college premiums from changes in the relative supplies of college workers. The higher degree of substitution of abstract with respect to routine tasks, then, means that the college premium in abstract intensive occupations respond less to the increase in supply of college workers.

4.2 Model Fit

I now discuss how the model fits the calibration targets. In Table 5, I show how the model replicates the occupation polarization along thirty years. It compares the difference of occupation employment share of 2010 with 1980 both in the model and in the data. Although, by construction, the model should match almost exactly the occupation share of the data, this does not imply that it generates the hollow out of the changes in the occupation employment distribution. This can be the case if the occupational mean wage rank generated by the model is not exactly the same of the one observed in the data. Table 5 displays that the model does replicate the increase in low skill occupation employment, but it does not generate enough polarization in the higher end of the distribution, specially in later years. This mean that it does a good job in ranking the occupations in the lower portion of the distribution, but possible assign some occupations with high wage in the data in the middle portion of the distribution[22].

Table 5: Calibration Results

Target	Data	Model
Job Polarization 1980-2010 (p.10)	0.260	0.299
Job Polarization 1980-2010 (p.25)	-0.042	-0.061
Job Polarization 1980-2010 (p.50)	-0.201	-0.176
Job Polarization 1980-2010 (p.75)	0.022	0.108
Job Polarization 1980-2010 (p.90)	0.260	-0.043
Avg ln wage - College (avg. of all t)	3.340	3.287
Avg. ln wage - Non-College (avg. of all t)	2.862	2.821
Correlation Occ. Wage with Abstract	0.672	0.795
Correlation Occ. Wage with Routine	-0.512	-0.568
Correlation Occ. Wage with Manual	-0.621	-0.699

Regarding average log wage, Table 5 shows that the model matches well the average wage inequality between college and non-college. Also, Figure A.6 (in the appendix) displays that the model matches the increase in the college premium in the U.S. along

[22]See Figure A.7 for the complete distribution

these forty years. Given that I am using data on men only, it is not surprising that the average wage has been decreasing, specially for low educated workers.

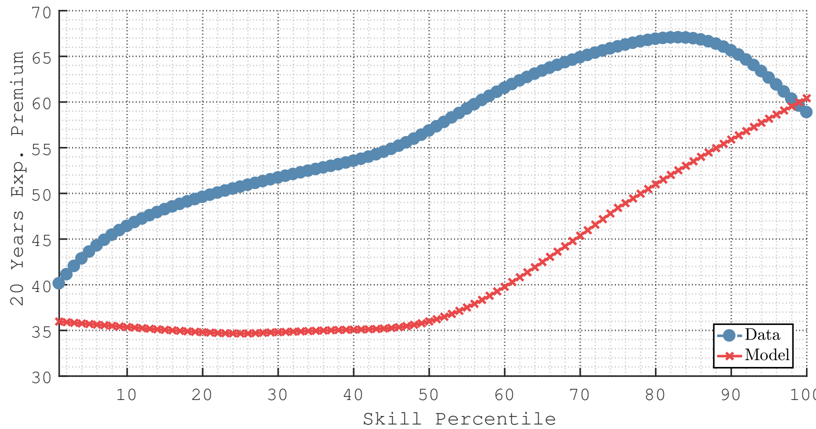
Finally, I present the correlation of the occupation mean wage with the abstract, routine and manual task content given by model and the data. The theory replicates the fact that abstract intensive occupations have higher wage, while routine and manual tasks are associated with lower average wage. It does, in fact, generate too much dispersion across occupations tasks, as abstract tasks have a higher positive correlation, and routine and manual tasks have lower negative correlation with occupation wage than in the data. Thus, in the case of the polarization picture, the model is likely not accounting for high-wage occupations with low content of abstract tasks.

Regarding the non-target moments, Figure 4 exhibits the distribution of returns to experience by occupation. It replicates qualitatively the graph observed in Figure 3. High skill occupations have up to 25 p.p. higher returns to experience than low skill occupations. In addition, there is a clear monotone increase in returns to experience around the 50th percentile. Although, the model predicts a relatively flat returns to experience for low wage occupations, the fact that it delivers a similar occupation distribution of returns to experience without directly targeting this moment is encouraging. The reason for the shape of the distribution of prices to experience in the model is given by the composition effects from tasks and education. Returns to experiences are specific to every education and task pair. Thus, even assuming that there were no differences in the supply of young and labor by occupation, because they are a composition of tasks, they have a cross-section distribution of returns to experience. In addition to that, some occupations have a higher share of college educated than others. Whenever an occupation is performed by a college worker instead of a non-college, it has, on average, higher returns to experience.

In Figure 5, I plot the evolution of the returns to experience in the model. The model predicts that, from 1980 to 2010, returns to experience should increase by around 6 p.p. However, this masks substantial heterogeneity by occupation, as, the price of experience has increased by up 10 p.p. for low skill, while high-skill occupations have increased by less than 5 p.p.

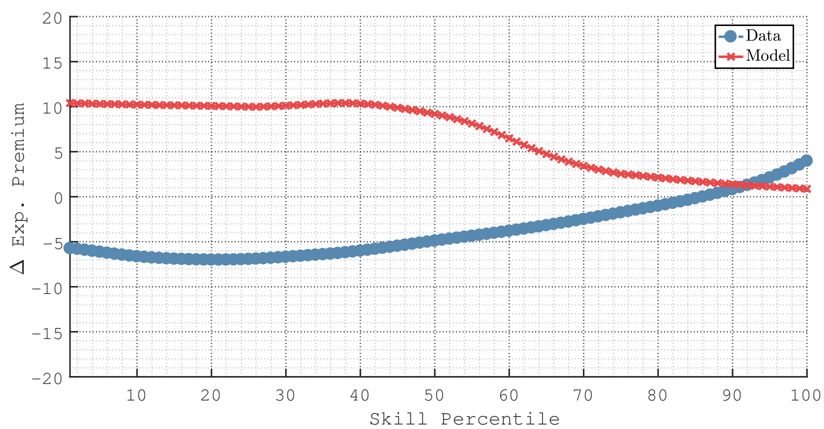
Given that this picture differs substantially than the data, one should ask what is driving the change in returns to experience in the model. As explained before, changes in returns to experience can be driven by changes in the composition of the occupation (e.g. share of college workers in the occupation) and by changes in the relative supply of young to old workers. In 2010, the relative supply of old to young labor was higher than

Figure 4: Returns to Experience by Occupation: Model



Notes: Returns to experience is calculated by taking the difference of the average log wage of workers of 20 to 25 years of experience and the average log wage of workers with 0 to 5 years of experience in the same occupation and averaging in all the model periods. The distribution is

Figure 5: Change in Returns to Experience by Occupation: Model



Notes: The change in the returns to experience is calculated by taking the difference of the average returns to experience by occupation in the model in 2010 and in 1980.

in 1980, that should give a decrease in the experience premium. However, because there is an overall increase of college educated workers in all occupations, and they have, on average, higher returns to experience, the observed change is positive.

5 Model Dynamics

In this section, I discuss the model static and dynamic response to the multiple demand and supply shifters present in the U.S. economy from 1970 to 2010.

5.1 Static Response

Demand shifters. In this section, for simplicity, I assume an unexpected and permanent increase in the level of Λ^R and Λ^S. As discussed before, given the calibration presented in table 4, an increase in both demand shifters put upward pressure in the wages of abstract intensive occupations and college educated workers, respectively. However, because of cohort specific frictions, different workers may, potentially, have different responses.

In the case of the college-biased technological change Λ^S, because the college decision happens only once in the beginning of the lifetime, the only supply response would come from new entrants in the labor market. Ceteris paribus, the increase in the relative supply of young college educated workers with respect to old workers, increases the price of experience for the college educated but decreases for the non-college educated. Furthermore, because the distribution of ability is education specific - college workers have higher probability to have larger ability draws in abstract intensive occupations - the increase in the college share shifts the occupational structure of young cohorts toward high skill occupations, and, as a consequence of general equilibrium effects, decreases the wage of these occupations *for all* cohorts.

The initial increase in the routine-biased shifter $\Lambda^R(t)$ has two effects on the relative wages of occupations: the direct increase in the relative price of abstract tasks with respect to routine tasks, and the indirect effect on the increase of the price of services with respect to goods coming from the change in the production of goods G. The direct effect accounts for the "right hand side" of the wage polarization, while the indirect effect accounts for "left hand side", the increase in the wage of low-skill manual-intensive services occupation. Similarly to the case of the college-biased technological change, the change in the $\Lambda^R(t)$ presents strong cohort effects. This is due to the fact that young workers are much more mobile than old workers, and thus, will have a stronger response from the change in wages. Again, the extra supply of young cohorts to abstract and manual tasks endogenously increases the price of experience for these tasks and decreases for the routine ones. The impact on the college premium is, a priori, unclear. Polarization increases wages in both extremes of the occupation distribution, the final effect would depend on the propensity of college and non-college to choose abstract, routine or manual intensive occupations[23].

Supply shifters. While demand shifters can be interpreted as time effects, supply shifters are characterized by being mainly cohort effects. In the model, the cohort size $n(\tau)$ is calibrated to replicate the population growth and labor force participation of American men. Assuming no changes in education or occupation employment share, a permanent increasing in the cohort size, at first, increases the price of experience for all groups. Later, after the smaller cohorts retire, the returns to experience recovery to their initial level. The increase in the price of experience, nevertheless, is different for every occupation and education. Because the elasticity of substitution for young and old labor is different for every education and task, even in the absence of any other change in demand and supply, an increase in cohort size has heterogeneous effects on the wages of every group in the model. Likewise, changes in the average college costs μ_t or in the average occupation ability π_{thej} can induce a similar response to the returns to experience[24].

5.2 Dynamic Response

In the model, I assume the effects of future wages are fully internalized by the agents, and they perfectly foresee changes in prices during the transition path. This implies that not only future changes in demand and supply shifters impact the contemporaneous decisions of the agents, but also their current decisions affect future prices.

As before, take an unexpected and permanent increase in the level of $\Lambda^S(t)$ which raises the share of young college educated workers and the contemporaneous college premium and returns to experience for college workers. Nonetheless, as new cohorts will be more educated, *future* college premium and returns to experience will be lower, decreasing the incentives to go to college today. In fact, because of the overlapping generation structure of the economy, even without further changes in the level of technology $\Lambda^S(t)$, future cohorts will face lower incentives to go to college[25]. In short, assuming that the economy is in a steady state can potentially overstate the value to go to college.

[23]In the presented calibration, the wage effect on the low-skill occupations is stronger than in high-skill occupations. One explanation is that both demand shifters are jointly calibrated and the college-biased shifter accounts for a large part of the wage effect on the high-skill occupations.

[24]Recall that both parameters are calibrated to match the change in college and occupation shares unexplained by wage changes.

[25]In the extreme case, the fact that all the cohorts cannot adjust instantaneously to shifts in demand may generate "echo effects".

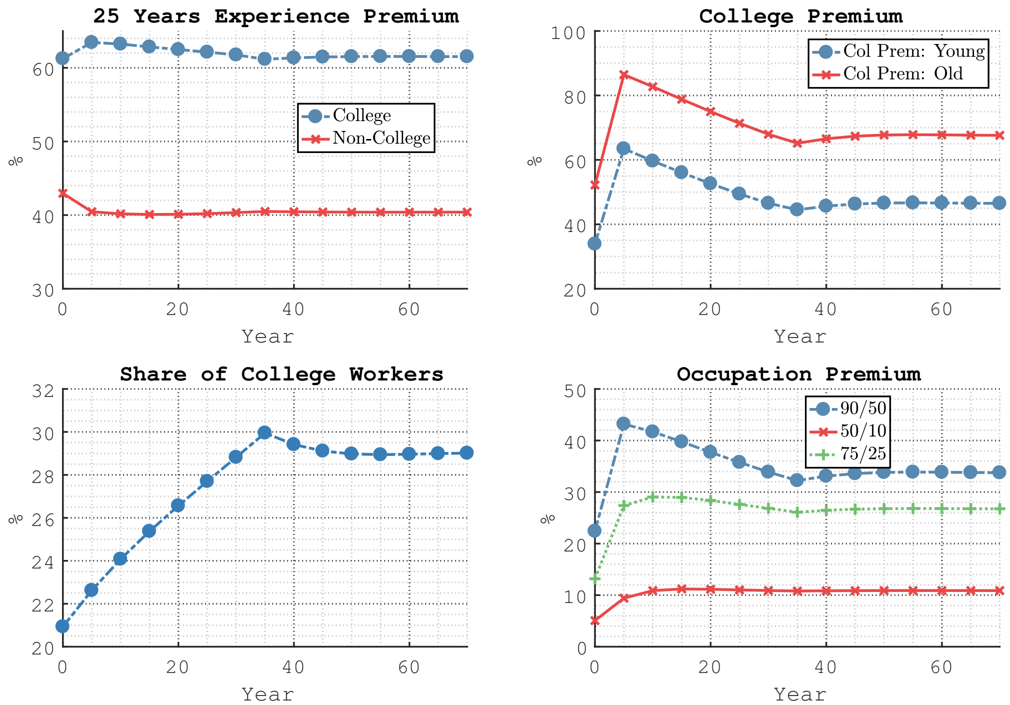
Although in the case of $\Lambda^R(t)$ the intuition is similar, there is an extra layer of complexity, the occupation mobility. In general, the adjustment of the occupation supply is faster than the education decision as, not only the entrants are allowed to optimize, but also a fraction of the older cohorts. As discussed before, just by the fact that younger workers are more mobile, they will have a strong response from contemporaneous wage changes. On top of that, because the continuation value of working in a given occupation is age dependent, the occupational choice of workers of different age will respond differently to changes in future wages. The first reason is the obvious case that an old worker is closer to retirement age, thus, she reaps very few benefits from an occupational change. On the other hand, young workers are so mobile that they give very low weight to future wages in their contemporaneous occupation decision. It turns out that the model predicts that men just entering prime age (around 15 years of experience), which are young enough to have many years in the labor market and old enough to not change occupation too often, are the ones that tend to respond more to future wages, giving 2.5 more weight to future wages than new labor market entrants.

Since occupational mobility has been increasing in the U.S. during the studied period, an obvious question is ask how this change is affecting the mechanisms behind the model. It turns out that, in the model, occupation switching costs drive mostly the dynamic effects of polarization. To gather intuition, suppose workers can move across occupations frictionless ($\beta = 0$). Any current change in occupation wages generate a similar response from all cohorts, prompting very small differences in labor supply across cohorts and almosto no change in the price to experience. In addition, future wage changes would have zero impact on the current occupational choice[26].

In figure 6 I show the response of an unexpected and permanent increase in the demand shift Λ^S. In the upper right corner, one observes the typical increase in the college premium followed by a demand shift biased to college workers. Given that new entrants are still making their educational decision, they are the only one who respond to the demand shift by having more college educated workers. The endogenous supply response from the new entrants, ceteris paribus, shift the relative supply of old versus young workers and that, because of complementarities between cohort shift the relative wage of old versus young workers, the experience premium. In the upper left corner, the graph plots exactly this movement. While for college workers the experience premium

[26]Mathematically, one could see that if $\beta = 0$, then $\epsilon_j \hat{w}(h, j) = \epsilon_j \tilde{w}(h, j)$, meaning that occupational choice (equation (8)) only depends on current wages.

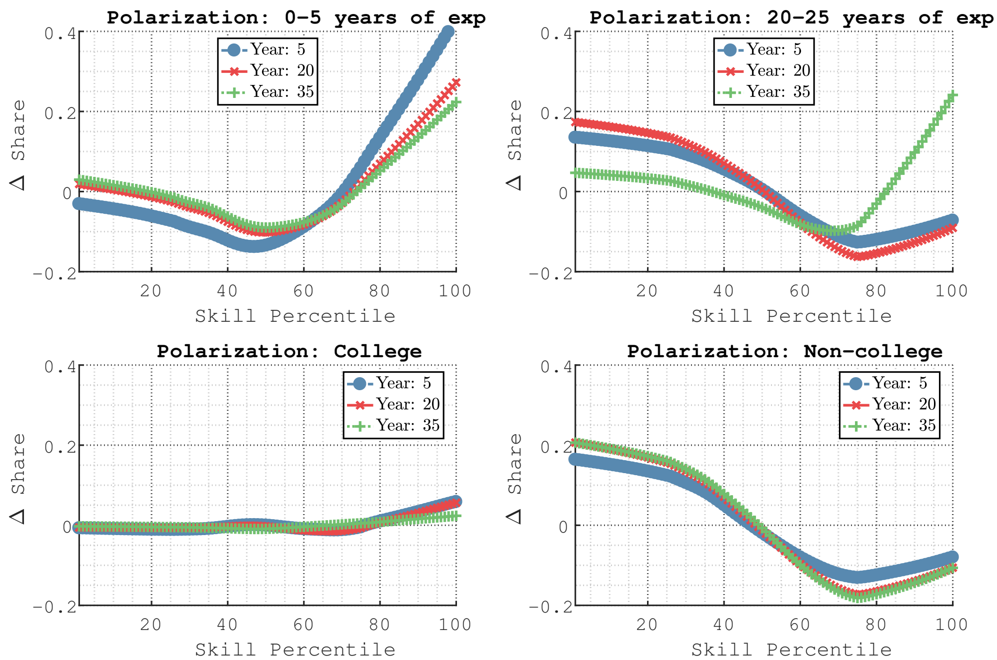
Figure 6: Unexpected and permanent increase in $\Lambda^S(t)$



goes up, for non-college workers there are more non-college and old relative to young pushing the experience premium down. As the economy transit out to new steady state, the increase in the supply of college educated workers push down the college premium - albeit still higher than the initial level - and shifts the experience premium closer to their initial level.

Following the effect in inequality, figure 7 displays the effect on polarization. The upper left corner shows that younger cohorts increase their propensity in working in high-wage occupations. That is because more educated cohorts have a higher propensity in working in a high-wage occupation. Given that there exist complementarities between the high-wage abstract occupations and the low-wage manual occupations, the biased demand shift also prompts a response of workers toward low-skill occupations. As the lower right corner shows, the response comes only from non-college educated workers.

Figure 7: Unexpected and permanent increase in $\Lambda^S(t)$



6 Counterfactuals

6.1 The Role of Demand

After taking into account the role of the dynamic response from the supply side, I revisited the role of demand in shaping inequality. Figure 8 plots the inequality dynamics in the calibrated baseline model. As expected, the usual inequality trends of the U.S. labor market are highlighted. The college premium rose together with the share of college workers, while the wage premium in high wage occupation increase relative to the median occupation.

In Figure 9, I fix $\Lambda^S(t)$, $\Lambda^R(t)$ and $\Lambda(t)$ to the 1970 level and simulate the model to calculate the counterfactual wage inequality. In this version, all the exogenous forces that increase the relative demand for certain groups of workers are shut down. The college premium - that increased by almost 10 p.p in the baseline - decrease by 8 p.p in the counterfactual simulation. The college share still increases due to exogenous supply forces, but the lack of demand forces damps the endogenous response of workers, decreasing the share of college workers by 7 p.p.

Finally, the absence of demand shifter also have a profound effect on occupation wage inequality. In the baseline, although the occupation premium increased in all groups, the

Figure 8: Inequality Dynamics in the Baseline Model

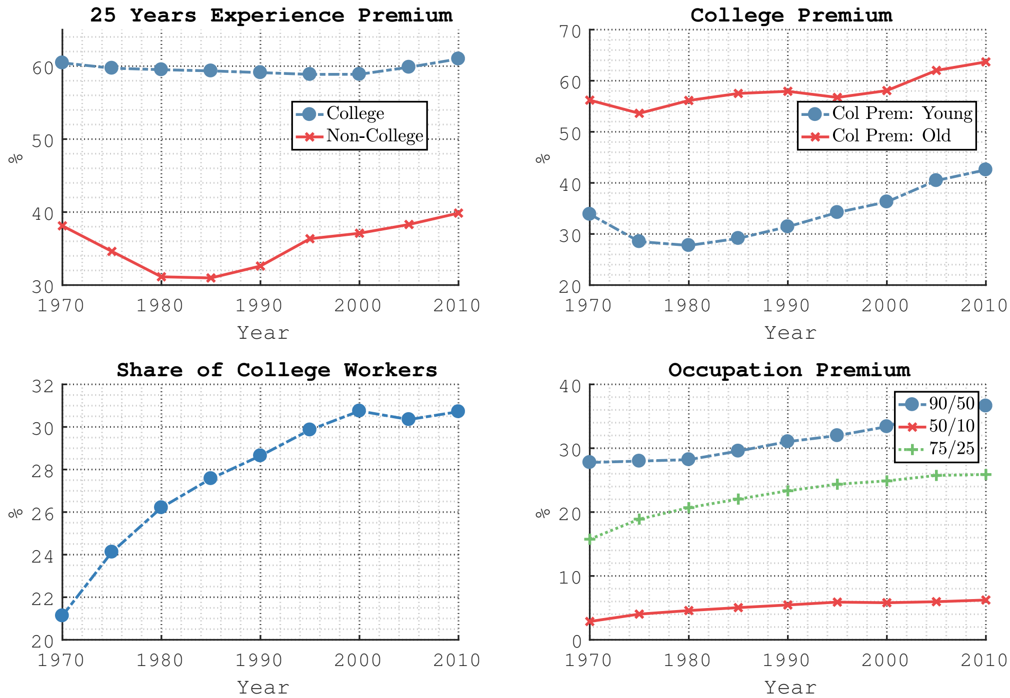
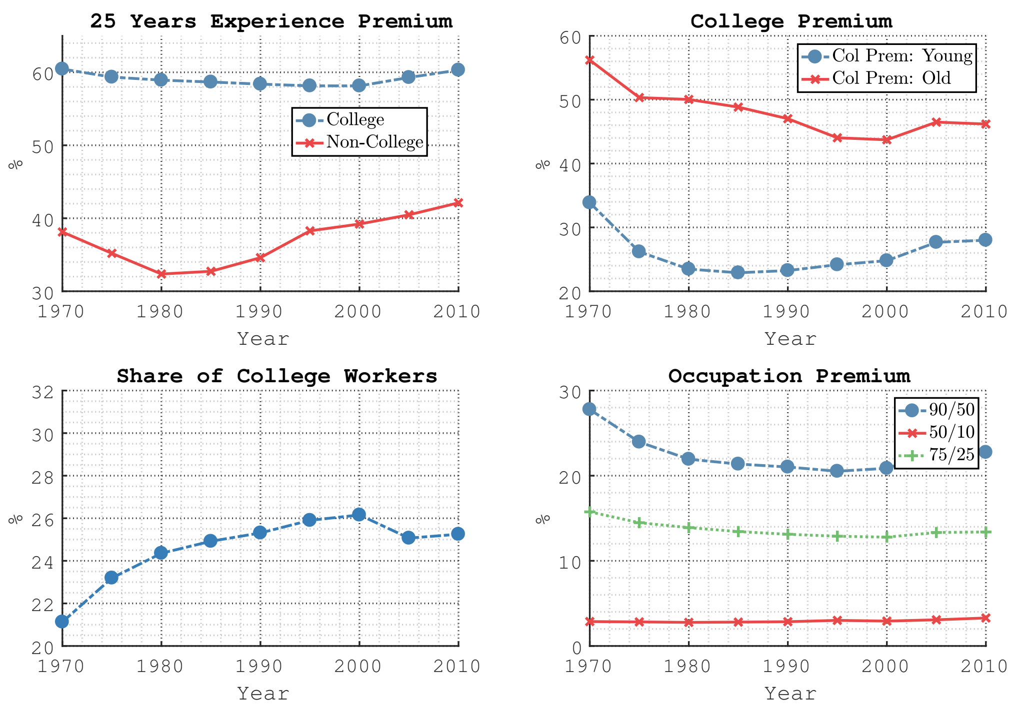


Figure 9: Inequality Dynamics without Demand Shifters



effect was stronger in the top than in the bottom. In the bottom right panel of Figure 9, it is noticeable that the increase in demand accounts for the increase in top income inequality. Without the demand effect, the occupation premium between the occupation in the 90th centile and the median decreases by 6 p.p. Nevertheless, the inequality between the median and the occupations in the 10th centile are roughly unchanged. This discrepancy is intimately related to the decrease in the college premium. Top occupations have a higher share of college educated workers and the decrease in the college premium accounts for part of the decreasing in top income inequality.

6.2 The Transition Effects

Through the lens of the model, the last 40 years of the U.S. labor market is characterized by being in a transition path. There has been big differences in terms of cohort size because of the entry of baby boomers in 1980s, every new cohort is more educated than the last one, occupational differences between young and old workers became more divergent every year and the estimates of demand shifters have changed substantially from 1970 to 2010. One interesting question the model is able to answer is, after all the transition effects phase out, what the U.S. wage structure looks like? To answer this, I simulate the model to a future steady state using the 2010 parameters, allowing all cohorts to optimize freely without any demographic friction.

Table 6: Inequality comparison, 2010 and future steady state

Statistic	**Simulation: 2010**	**Simulation: Steady State**
College Premium (%)	52.51	41.91
20 Years Ret. to Experience - College (%)	60.91	59.64
20 Years Ret. to Experience - Non-College (%)	39.80	40.37
Occupation Mean Wage 90/50 (%)	34.29	22.33
Occupation Mean Wage 50/10 (%)	12.30	23.26
Occupation Mean Wage 75/25 (%)	25.77	25.65

By comparing the future steady state with 2010, table 6 shows that, in many dimensions, inequality is higher during the transition path. In the last period the model is fit, college premium is around 52%, while in the future steady state it decreases to 41%, a decrease by more than 10 p.p. The intuition is clear, in 2010 the entrant cohort has a college share of almost 38%, whereas only 31% of prime age men have a college education. Once I allow all cohorts to choose their education, the number of college workers increase substantially, decreasing the college premium. On the other hand, the price of experience moves only slightly, decreasing around 1 p.p. for college workers and increasing by the same amount for non-college. The explanation for this small movement is that the main reason the price of experience changes is large differences in relative cohort size. The cohort size attained its peak in 1980, decreasing for every subsequent cohort ever since. In 2010, the cohort entering the labor market is, in reality, a bit smaller than the one that entered 1970. In the model, the majority of the baby boomers already retired in 2010 and demographic transition is already close to an end.

Finally, to connect to the job polarization literature, I show how occupation inequality changed. First, notice that, in 2010, occupations in the 90th percentile (ranked by occupation mean wage), earned almost 35% percent more than the median occupation. On the other hand, the median occupation earned only 12% more than the occupation in the 10th percentile. In the future steady state, nonetheless, the wage of the median occupation increases relative to the others, shrinking the difference to the 90th percentile to 22% and increasing the difference to the 10th percentile to 23%. This movement in the wage of middle-skill occupations is consistent with wage and occupation polarization. During the transition, routine-biased technological change puts downward pressure on middle-skill wages, but because of moving costs, there are still too many workers in the "dying" occupations. In the steady state, however, all the workers move out from the routine-intensive occupations, helping the wages of these occupations to rebound.

7 Conclusion

In this paper, I develop an overlapping generations model with endogenous occupational and educational choices to investigate the multiple changes in the U.S. wage structure over the last 40 years. The model is tractable and is built to account for multiple demand and supply shifters: the college-biased technological change, the routine-biased technological change, the changes in cohort size, and the change in educational and occupation share across cohorts. Empirically, I provide evidence that occupation polarization affected disproportionally workers of different cohorts and education. I argue that this is linked with the heterogeneity in the returns to experience by occupation in level and changes.

Based on these facts, I take my model to the data. Given the life-cycle structure of the workers, the model is well suited to evaluate the dynamics in the transition path. In the model, because college decision is made in the beginning of their life and occupation switching is costly, old workers cannot respond as promptly to demand shifters as the young ones. During the transition, technological change may introduce cohort effects that have long lasting effects in the wage structure: routine-intensive occupations will have too many uneducated old workers with respect to abstract and manual. This effects only vanish in the very long run. Comparing 2010 with a future steady state, where all cohorts are allowed to freely choose their occupation and education, I find that, during the transition, the college premium is 11 p.p. higher and the difference between the average wage of the occupation in the 90th percentile and the median is 12 p.p higher than the steady state.

Finally, the model is used to understand the contribution of changes in the demand in the inequality in the U.S. To answer this, I simulate the model from 1970 to 2010 and produce the statistics of a counterfactual U.S. labor market. The demand shifters were responsible for 19 p.p. of the increase in the college premium and for 7 p.p. of the increase in the share of college workers. Moreover, they account for part of the increase in top income inequality, as the premium paid between the top and median occupation was about 15 p.p. lower in the counterfactual world, while the premium between the median and the bottom was only 4 p.p lower.

Albeit the model is fully efficient, in the sense that workers are doing what is best, given the observed technology, the adjustment in the transition is slower when occupation mobility is lower. This has a clear policy implication. A decrease in the occupation switching costs, specially for old workers, can help individuals to move out from dying routine occupations. Policies that provide information, training and assistance in reallocating workers - through the lights of the model this represents a decrease of the probability of re-choosing the occupation - accelerate the transition, reduce inequality and may even increase aggregate production.

There are still several avenues for future research. First, although I allowed for the increase in occupational mobility in a simple reduced-form manner, reallocation can be a possible endogenous response from technological change. Carefully modelling occupation switching in a technological transition is not a trivial task, but it can certain enrich the analysis in the future. Second, the heterogeneous returns to experience by occupation comes from the differences in the old labor supply schedule by education and task. A more systematic investigation on the sources of the wage differential between young and old workers by occupation would definitely complement my framework. Third, a well known fact is the decline in the labor force participation of prime age men. Some papers have pointed that job polarization may be behind this fact (Cortes et al. (2017)), while others have connected this trend to changes in returns to experience (Elsby and Shapiro (2012)). Since my framework accounts for both facts, extended it to incorporate labor force participation can be worth exploring further. Fourth, I abstract completely from the transition in and out of unemployment. Obviously, adding human capital depreciation and/or search frictions can shed light on the sources of the slow adjustment during a transition. Finally, in my framework, the workers fully forecast and respond to changes in future prices. This behavior may be unconnected with reality and future research could address this issue.

References

Acemoglu, D. (2002). Technical Change, Inequality, and the Labor Market. *Journal of Economic Literature 40*(1), 7–72.

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

Adda, J., C. Dustmann, and K. Stevens (2017). The Career Costs of Children. *Journal of Political Economy 125*(2), 293–337.

Autor, D. and D. Dorn (2009). This job is "Getting Old": Measuring changes in job opportunities using occupational age structure. *American Economic Review: Papers & Proceedings 99*(2), 45–51.

Autor, D. H. and D. Dorn (2013). The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market. *American Economic Review 103*(5), 1553–97.

Autor, D. H., D. Dorn, and G. H. Hanson (2015). Untangling Trade and Technology: Evidence from Local Labour Markets. *Economic Journal 125*(584), 621–646.

Autor, D. H., F. Levy, and R. Murmane (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics 118*(4), 1279–1333.

Card, D. and T. Lemieux (2001). Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis. *The Quarterly Journal of Economics 116*(2), 705–746.

Cerina, F., A. Moro, and M. P. Rendall (2017, March). The Role of Gender in Employment Polarization. ECON - Working Papers 250, Department of Economics - University of Zurich.

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

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

Dinardo, J., N. M. Fortin, and T. Lemieux (1996). Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach. *Econometrica 64*(5), 1001–1044.

Elsby, M. W. L. and M. D. Shapiro (2012). Why does trend growth affect equilibrium employment? A new explanation of an old puzzle. *American Economic Review 102*(4), 1378–1413.

Fortin, N., T. Lemieux, and S. Firpo (2011). Decomposition Methods in Economics. In *Handbook of Labor Economics*, Volume 4, pp. 1–102. Elsevier Inc.

Goos, M., A. Manning, and A. Salomons (2009). Job polarization in europe. *American Economic Review: Papers & Proceedings 99*(2), 59–63.

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

Heathcote, J., K. Storesletten, and G. L. Violante (2010). The Macroeconomic Implications of Rising Wage Inequality in the United States. *Journal of Political Economy 118*(4), 681–722.

Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2018). The Allocation of Talent and U.S. Economic Growth.

Huckfeldt, C. (2018). Understanding the Scarring Effect of Recessions. 2018 Meeting Papers 1207, Society for Economic Dynamics.

Jeong, H., Y. Kim, and I. Manovskii (2015). The price of experience. *American Economic Review 105*(2), 784–815.

Kambourov, G. and I. Manovskii (2005). Accounting for the Changing Life-Cycle Profile of Earnings. *Society of Economic Dynamics Meeting Paper 231*.

Kambourov, G. and I. Manovskii (2008). Rising occupational and industry mobility in the United States: 1968-97. *International Economic Review 49*(1), 41–79.

Katz, L. and K. Murphy (1992). Changes in Relative Wages, 1963-1987: Supply and Demand Factors. *The Quarterly Journal of Economics 107*(1), 35–78.

Katz, L. F. and D. H. Autor (1999). Chapter 26 Changes in the wage structure and earnings inequality. In *Handbook of Labor Economics*, Volume 3, pp. 1463–1555. Elsevier Masson SAS.

Krusell, P., L. E. Ohanian, J. V. Rios-Rull, and G. L. Violante (2000). Capital-skill complementarity and inequality: a macroeconomic analysis. *Econometrica 68*(5), 1029–1054.

Appendix

A Additional Figure and Tables

Figure A.1: College Premium and College Share

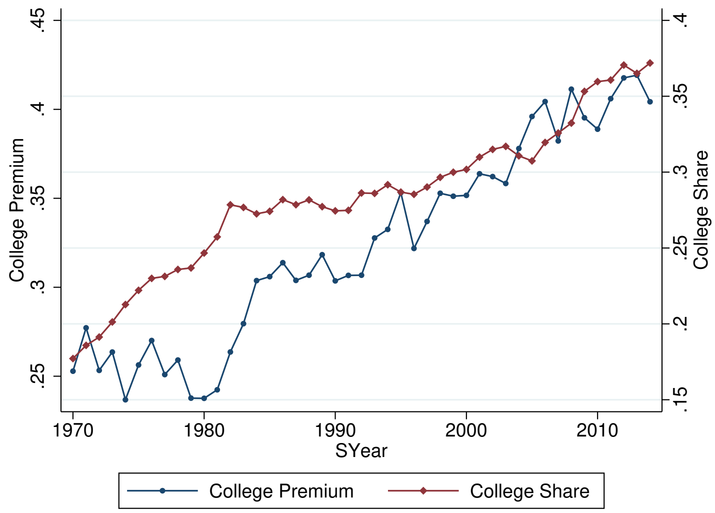
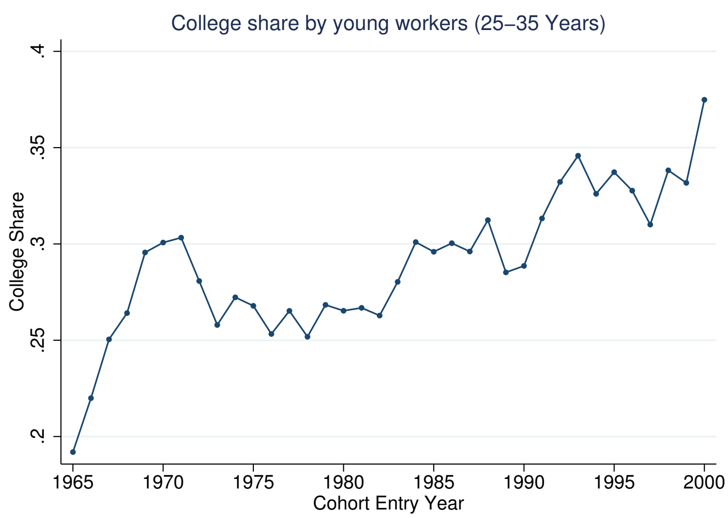


Figure A.2: College Share by Cohort



Notes: College share of workers between 25-35 by the year they have enter the labor market. CPS data.

Figure A.3: Population Share by Age Group

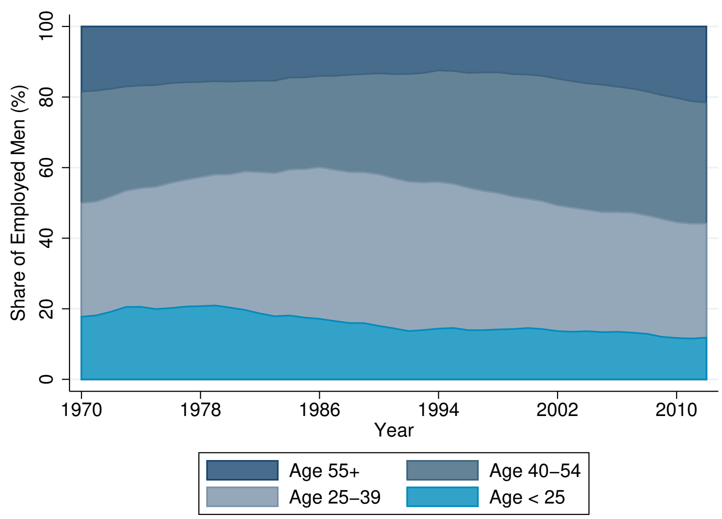
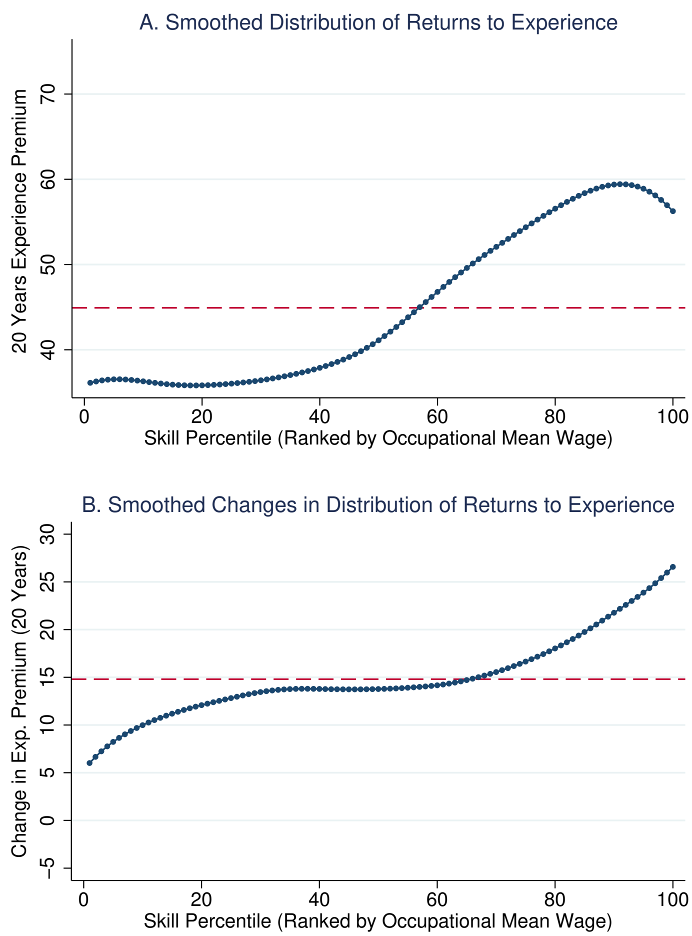


Figure A.4: Smoothed Distribution of Returns to Experience and the Changes in Returns to Experience by Occupation: By Cohort



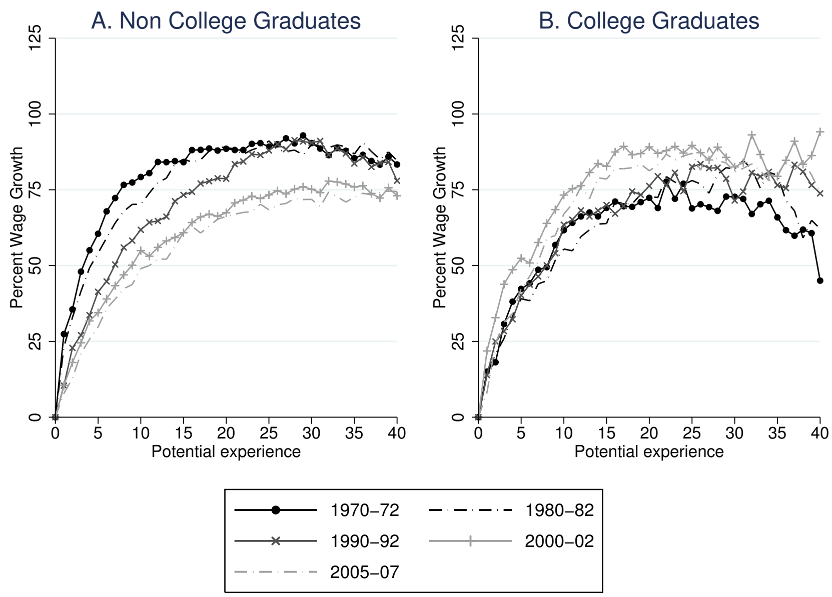


Figure A.5: Changes in the experience-wage profile for college and non-college graduates, 1970-2005.

Notes: Experience coefficients of a regression of log wages over a full set of experience dummies and year fixed effects. Non-college graduates include all individuals with less than 16 years of education, while college graduates include individuals with more or equal than 16 years of education. CPS data.

Figure A.6: Average wage fit: College and Non-college

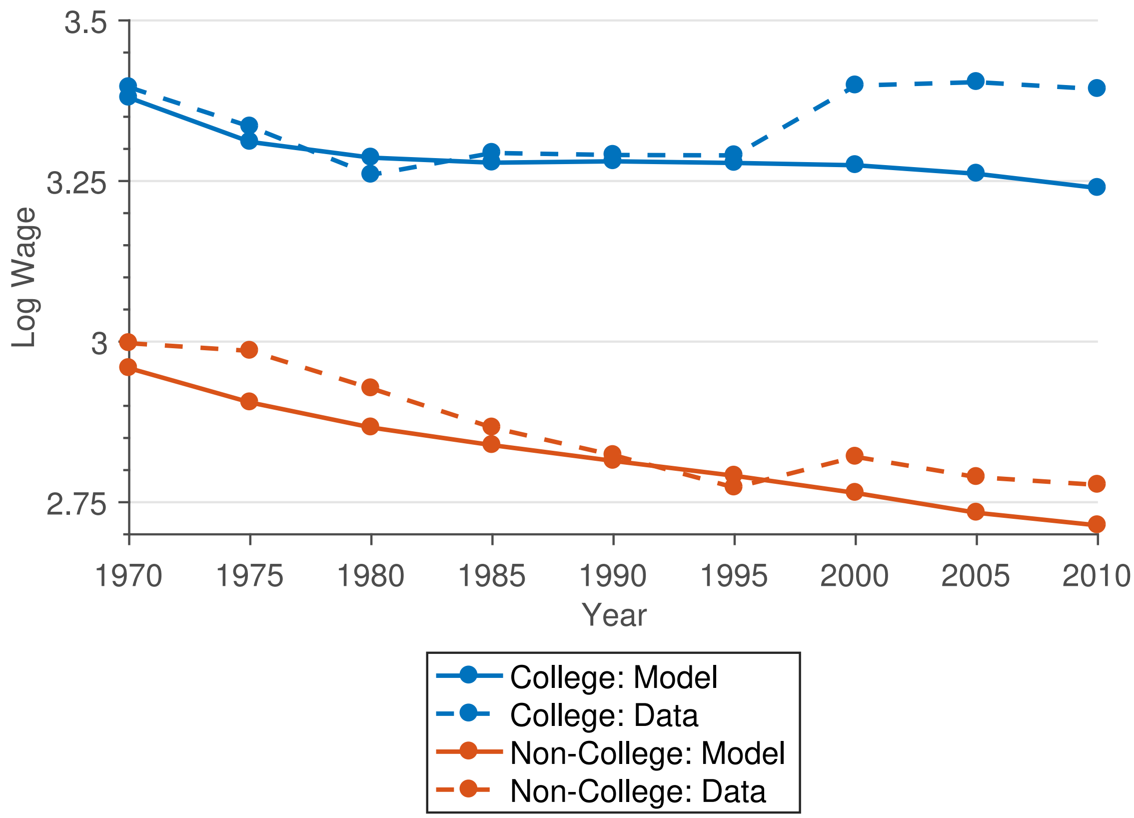
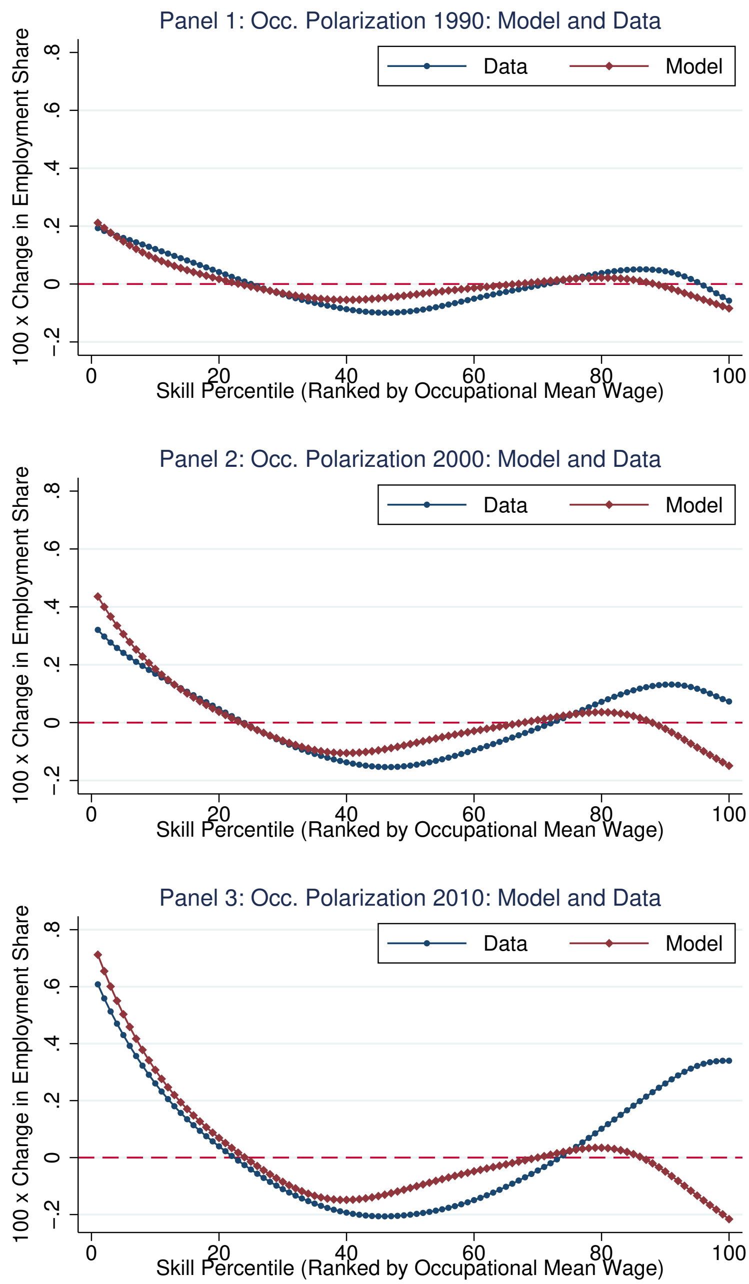


Figure A.7: Occupation Polarization: Model fit



Notes: The Occupation Mean Wage is computed by pooling all age groups and years from 1978 to 1982 in the CPS. Job polarization is computed by taking the difference in employment 2010 to 1980. The distribution is smoothed using a locally weighted regression with bandwidth 0.75.

B Theory Appendix

B.1 Proof of Proposition 1

Let the value of a worker of age h in occupation j with ability vector ϵ. Without loss of generality, consider any time t and education e, and, for simplicity, consider β to be the same across different ages (notice that under this assumption $\prod_{n=1}^{k} \beta(h+k-n) = \beta^k$):

$$W(h,j,\epsilon) = \tilde{w}(h,j)\epsilon_j + \beta W(h+1,j,\epsilon) + (1-\beta)\mathbb{E}_{\epsilon}[\max_{j'}\left\{W(h+1,j',\epsilon')\right\}] \quad \text{for } h = 1,...,H-1 \tag{20}$$

$$W(h,j,\epsilon) = \tilde{w}(h,j)\epsilon_j \qquad \text{for } h = H \tag{21}$$

Let $A(h+1) = (1-\beta)\mathbb{E}_{\epsilon}[\max_{j'}\left\{W(h+1,j',\epsilon')\right\}]$, then I can rewrite equation 20 as:

$$W(H-1,j,\epsilon) = \tilde{w}(H-1,j)\epsilon_j + \beta\tilde{w}(H,j) + A(H)$$

$$W(H-n,j,\epsilon) = \epsilon_j \sum_{k=0}^{n} \beta^k \tilde{w}(H-n+k,j) + \sum_{k=0}^{n-1} \beta^k A(H-1-n+k)$$

$$W(1,j,\epsilon) = \epsilon_j \sum_{k=0}^{H-1} \beta^k \tilde{w}(1+k,j) + \sum_{k=0}^{H-2} \beta^k A(2+k)$$

$$W(h,j,\epsilon) = \epsilon_j \underbrace{\sum_{k=0}^{H-h} \beta^k \tilde{w}(h+k,j)}_{\hat{w}(h,j)} + \underbrace{\sum_{k=0}^{H-h-1} \beta^k A(h+1+k)}_{\overline{A(h)}}$$

Proof. 1. Denote the value to be in occupation j by $w(\hat{h},j)\epsilon_j$ at age. Where ϵ_j is the ability to work in occupation which is draw from a multivariate Frechet distribution: $F(\epsilon_1,\epsilon_2,...,\epsilon_M) = \exp\left\{-\sum_{s=1}^{M} \pi_{hs}\epsilon_s^{-\theta}\right\}$. Without loss of generality, let us find the share of workers that choose occupation 1 ate age h. Define this share as p_1:

$$
\begin{aligned}
p_1(h) &= P(W(h,1,\epsilon) > W(h,s,\epsilon)) \qquad \forall s \neq 1 \\
p_1(h) &= P(\hat{w}(h,1)\epsilon_1 + \overline{A(h)} > \hat{w}_(h,s)\epsilon_s + \overline{A(h)}) \qquad \forall s \neq 1 \\
&= P(\hat{w}(h,1)\epsilon_1 > \hat{w}(h,s)\epsilon_s) \qquad \forall s \neq 1 \\
&= P\left[\epsilon_2 < \frac{\hat{w}(h,1)}{\hat{w}(h,2)}\epsilon_1, \epsilon_3 < \frac{\hat{w}(h,1)}{\hat{w}(h,3)}\epsilon_1, ..., \epsilon_M < \frac{\hat{w}(h,1)}{\hat{w}(h,M)}\epsilon_1\right] = P\left[\epsilon_s < \alpha_s \epsilon_1\right] \quad \forall \neq 1 \\
&= \int F_{h1}(\epsilon, \alpha_2\epsilon, ..., \alpha_M\epsilon)d\epsilon \\
&= \int \theta\pi_{h1}\epsilon^{-\theta-1}\exp\left\{-\bar{\alpha}\epsilon^{-\theta}\right\}d\epsilon \\
&= \frac{\pi_{h1}}{\bar{\alpha}}\int \bar{\alpha}\theta\epsilon^{-\theta-1}\exp\left\{-\bar{\alpha}\epsilon^{-\theta}\right\}d\epsilon = \frac{\pi_{h1}}{\bar{\alpha}}\int dF(\epsilon) = \frac{\pi_{h1}}{\bar{\alpha}}
\end{aligned}
$$

where we define $\alpha_s \equiv \frac{\hat{w}(h,1)}{\hat{w}(h,2)}$ and $\bar{\alpha} \equiv \sum_s \pi_s \alpha_s^{-\theta}$. Thus, we can write $p_1(h)$ as:

$$
\begin{aligned}
p_1(h) =& \frac{\pi_{h1}}{(\hat{w}(h,1)/\hat{w}(h,1))^\theta \pi_{h1} + (\hat{w}(h,2)/\hat{w}(h,1))^\theta \pi_{h2} + ... + (\hat{w}(h,M)/\hat{w}(h,1))^\theta \pi_{hM}} \\
p_1(h) =& \frac{\pi_{h1}\hat{w}(h,1)^\theta}{\sum_s \pi_{hs}\hat{w}(h,s)^\theta}
\end{aligned}
$$

and it is equivalent to write for any occupation share $p_j(h)$. □

Proof. 2. At age 1, since all workers are choosing their occupation, is trivial to see that the occupational share is given by: $s_j(1) = p_j(1)$. Then for all other ages, occupational share is given by the composition between the share β of workers who had to stay in the same occupation plus the rest who were allowed to reoptimize:

$$
s_j(h) = \beta s_j(h-1) + (1-\beta)p_j(h)
$$

So iterating forward:

$$
\begin{aligned}
s_j(2) &= \beta p_j(1) + (1-\beta)p_j(2) \\
s_j(3) &= \beta^2 p_j(1) + \beta(1-\beta)p_j(2) + (1-\beta)p_j(3) \\
&\quad ... \\
s_j(h) &= \beta^{h-1}p_j(1) + (1-\beta)\sum_{k=2}^{h}\beta^{h-k}p_j(k)
\end{aligned}
$$

□

Proof. 3. First, let us compute the average ability of new entrants in an occupation j at age h: $\overline{\mathbb{E}\left[\epsilon_j|h,j\right]} = \mathbb{E}[\epsilon^*|h]$. To calculate this I need to know the distribution of the chosen occupation. Denote

$$y^*(h) \equiv \max_j \left\{W(h,j,\epsilon))\right\} = \max_j \left\{\hat{w}(h,j)\epsilon_j + \overline{A(h)}\right\} = \max_j \left\{\hat{w}(h,j)\epsilon_j\right\} = \hat{w}(h)^*\epsilon^*$$

as the utility of the chosen occupation. Then:

$$\begin{aligned} P[y^*(h) < z] &= P[y(h)_j < z] \qquad \forall j \\ &= P\left[\epsilon_j < \frac{z}{\hat{w}(h,j)}\right] \qquad \forall j \\ &= F\left(\frac{z}{\hat{w}(h,1)}, \ldots, \frac{z}{\hat{w}(h,M)}\right) \\ &= \exp\left\{-\sum_s \pi_s \hat{w}(h,s)^\theta z^{-\theta}\right\}. \end{aligned}$$

Using the distribution of y^* it is straightforward to derive the distribution of the chosen occupation ϵ^*:

$$\begin{aligned} P[y^*(h) < z] &= P\left[\epsilon^* < \frac{z}{\hat{w}^*(h)}\right] \\ &= \exp\left\{\frac{-\sum_s \pi_{hs}\hat{w}(h,s)^\theta (z/\hat{w}^*(h))^{-\theta}}{\hat{w}^*(h)^\theta}\right\} \\ &= \exp\left\{\frac{-\sum_s \pi_{hs}\hat{w}(h,s)^\theta x^{-\theta}}{\hat{w}^*(h)^\theta}\right\} = P[\epsilon^* < x] \\ &= \exp\left\{-x^{-\theta}\pi_h^* \frac{\sum_s \pi_{hs}\hat{w}(h,s)^\theta}{\pi_h^* \hat{w}^*(h)^\theta}\right\} \\ &= \exp\left\{-\frac{\pi_h^*}{p^*(h)} x^{-\theta}\right\} = G(x|h). \end{aligned}$$

Where I use the previous definition of $p^*(h) \equiv \frac{\pi_h^* \hat{w}^*(h)^\theta}{\sum_s \pi_{hs}\hat{w}(h,s)^\theta}$. Now the I have the distribution of the chosen occupation, I can get any moment for all occupations j:

$$\begin{aligned} \overline{\mathbb{E}(\epsilon_j^k|h,j)} &= \int_0^\infty \epsilon^k dG(\epsilon|h) \\ &= \int_0^\infty \frac{\theta\pi_{hj}}{p_j(h)} \epsilon^{-\theta-1+k} \exp\left\{-(\pi_{hj}/p_j(h))\epsilon^{-\theta}\right\} d\epsilon. \end{aligned}$$

Using the definition of the Gamma function: $\Gamma(\alpha) \equiv \int_0^\infty x^{\alpha-1}e^{-x}dx$ and applying the change of variable $x = h(\epsilon) \equiv \frac{\pi_j}{p_j}\epsilon^{-\theta}$, and $f_x(x) = g_\epsilon\left(h^{-1}(x)\right)\left|\frac{\partial}{\partial x}h^{-1}(x)\right| = e^{-x}$ so:

$$\begin{aligned}\overline{\mathbb{E}(\epsilon_j^k|h,j)} &= \left(\frac{\pi_{hj}}{p_j(h)}\right)^{\frac{k}{\theta}} \int_0^\infty x^{-k/\theta}e^{-x}dx \\ &= \left(\frac{\pi_{hj}}{p_j(h)}\right)^{\frac{k}{\theta}} \Gamma\left(1-\frac{k}{\theta}\right)\end{aligned}$$

Finally, it is easy to see that I can do the same thing I did for the occupational shares for to compute the average ability of a given age $E(\epsilon|h,j)$. First, $\overline{E(\epsilon|h,1)} = E(\epsilon|h,1)$. Then, iterate in the following function:

$$E(\epsilon|h,j) = \beta E(\epsilon|h-1,j) + (1-\beta)\overline{E(\epsilon|h,j)}$$

I find the average ability of a given age h and occupation j:

$$E(\epsilon|h,j) = \beta^{h-1}\overline{E(\epsilon|h,1)} + (1-\beta)\sum_{k=2}^{h}\beta^{h-k}\overline{E(\epsilon|k,j)}$$

□

C Numerical Appendix

To solve the model, I first solve for the equilibrium in the initial steady state, then, I solve for the transition years assuming the last year is a steady state. The following outline summarize the numerical algorithm:

1. Guess a sequence of allocations $L_{iex}^{l}(t)$ for all t.
2. Compute occupation wages $\tilde{w}(h,j)$ and the discounted sequence of future wages $\hat{w}(h,j)$ for all h, j and t.
3. Using $\hat{w}(h,j)$ compute the law of motion of occupation shares $s_{tej}(h)$ and the average ability by occupation $E(\epsilon|.)$.
4. Using the shares and abilities recover the value function for college and non college $V(e,\tau)$ and the college share by cohort $e(\tau)$.
5. Aggregate the labor supply by occupation, age and time to get new allocations: $L_{iex}^{l+1}(t)$.

6. if $||L_{iex}^{l}(t) - L_{iex}^{l+1}(t)||$ is close enough stop. Otherwise update the guess L_{iex}^{l}.