Kent Academic Repository

Full text document (pdf)

Citation for published version

Barany, Zsofia L. and Siegel, Christian (2018) Job Polarization and Structural Change.   American Economic Journal: Macroeconomics, 10 (1).   pp. 57-89. ISSN 1945-7707.

DOI

Link to record in KAR

http://kar.kent.ac.uk/61271/

Document Version

Author's Accepted Manuscript

Job Polarization and Structural Change[*]

Zsófia L. Bárány[†]   Christian Siegel[‡]

20 March 2017

Abstract

We document that job polarization – contrary to the consensus – has started as early as the 1950s in the US: middle-wage workers have been losing both in terms of employment and average wage growth compared to low- and high-wage workers. Given that polarization is a long-run phenomenon and closely linked to the shift from manufacturing to services, we propose a structural change driven explanation, where we explicitly model the sectoral choice of workers. Our simple model does remarkably well not only in matching the evolution of sectoral employment, but also of relative wages over the past fifty years.

JEL codes: E24, J22, O41

Keywords: Job Polarization, Structural Change, Roy model

[†]Sciences Po. Email: zsofia.barany@sciencespo.fr
[‡]University of Kent. Email: c.siegel@kent.ac.uk

1   Introduction

The polarization of the labor market is a widely documented phenomenon in the US and several European countries since the 1980s.[1] This phenomenon, besides the relative growth of wages and employment of high-wage occupations, also entails the relative growth of wages and employment of low-wage occupations compared to middle-wage occupations. The leading explanation for polarization is the routinization hypothesis, which relies on the assumption that information and computer technologies (ICT) substitute for middle-skill and hence middle-wage (routine) occupations, whereas they complement the high-skilled and high-wage (abstract) occupations (Autor, Levy, and Murnane (2003), Autor, Katz, and Kearney (2006), Autor and Dorn (2013), Feng and Graetz (2014), Goos, Manning, and Salomons (2014), Michaels, Natraj, and Van Reenen (2014)).

The contribution of our paper is twofold. First, we document a set of facts which raise flags that routinization driven by ICT, although certainly playing a role from the 1980s onwards, might not be the only driving force behind this phenomenon. Second, based on these facts we propose a novel perspective on the polarization of the labor market, one based on structural change.

Our analysis of US data for the period 1950-2007 reveals some novel facts.[2] First, we document that polarization defined over occupational categories both in terms of employment and wages has been present in the US since the 1950s, which is long before ICT could have played a role. Second, we show that at least since the 1960s the same patterns for both employment and wages are discernible in terms of three broad sectors: low-skilled services, manufacturing and high-skilled services. Moreover, we confirm previous findings that a significant part of the observed occupational employment share changes are driven by sectoral employment shifts. Additionally we show that sectoral effects contribute significantly to occupational wage changes. Therefore understanding the sectoral labor market trends is important even for the occupational

---

[1]See for example Autor, Katz, and Kearney (2006) and Autor and Dorn (2013) for the US, Goos and Manning (2007) for the UK, Spitz-Oener (2006) and Dustmann, Ludsteck, and Schönberg (2009) for Germany, Goos, Manning, and Salomons (2009, 2014) for Europe.

[2]Analyzing the data until more recent years does not affect our findings; we chose 2007 as the final year to exclude the potential impact of the financial crisis.

trends.

Based on these facts, we propose a structural change driven explanation for these sectoral labor market trends. We introduce a Roy-type selection mechanism (Roy (1951)) into a multi-sector growth model, where each sector values a specific skill. Individuals, who are heterogeneous along a range of skills, optimally select which sector to work in. As long as the goods produced by the different sectors are complements, a change in relative productivities increases labor demand in the relatively slow growing sectors, and wages in these sectors have to increase in order to attract more workers.

In particular we assume that there are three types of consumption goods: low-end service, manufacturing and high-end service goods.[3] We break services into two as they comprise of many different subsets, e.g. dry cleaners vs. banking, which seem hardly to be perfect substitutes in consumption, as would be implied by having a single service consumption in households' preferences. In our model, we therefore treat low- and high-end services as being just as substitutable with each other as they are with manufacturing goods.

A change in relative productivities does not only affect relative supply, but through prices it also affects relative demand. Given that goods and the two types of services are complements, as relative labor productivity in manufacturing increases, labor has to reallocate from manufacturing to both service sectors. To attract more workers into both low- and high-skilled services, their wages have to improve relative to manufacturing. Since in the data we see that manufacturing jobs tend to be in the middle of the wage distribution, this mechanism leads to a pattern of polarization in terms of sectors, which is driven by the interaction of supply and demand for sectoral output.

We calibrate the model to quantitatively assess the contribution of structural change – driven by unbalanced technological progress – to the polarization of wages and employment. Taking measured labor productivity growth from the data and using existing estimates for the elasticity of substitution between sectors, we find that our model

---

[3]Buera and Kaboski (2012) also split services into low- and high-skilled: their selection is based on the fraction of college educated workers in the industry. Their main interest is linking the rising skill premium to the increasing share of services in value added, and they emphasize the home vs market production margin. Our focus is very different: sectoral wages.

predicts between 33 and 59 per cent of the relative average wage gain of high- and low-skilled services compared to manufacturing, and between 62 and 99 per cent of the change in employment shares. For this exercise, we quantify the adjustment of labor productivity growth needed to correct for selection effects in the calibrated model. Without these adjustments, productivity growth is understated in the expanding sectors, and is overstated in the shrinking sector; according to our model this would lead to an overstatement of annual productivity growth rate differentials by between 27 and 35 per cent.

This paper builds on and contributes to the literature both on polarization and on structural change. To our knowledge, these two phenomena until now have been studied separately.[4] However, according to our analysis of the data, polarization of the labor market and structural change are closely linked to each other, and according to our model, industrial shifts can lead to polarization.[5] Our theory highlights a particular connection between structural change and occupational structure. Structural change leads to the sector with the highest productivity growth to shrink in terms of employment, and to experience lower wage growth than the other sectors of the economy; the occupation which is used the most intensively in this sector also experiences employment and wage losses. In the period 1950 to 2007 around half of routine employment was in the manufacturing sector (and around 80 per cent of manufacturing hours were in routine jobs). Thus when manufacturing started to shrink after 1950, this led to a "hollowing out" pattern, as routine workers were in the middle both in terms of skills and wages.

The structural change literature has documented for several countries that as income increases employment shifts away from agriculture and from manufacturing towards services, and expenditure shares follow similar patterns (Kuznets (1957), Maddison (1980), Herrendorf, Rogerson, and Valentinyi (2014)). In particular the employment and expenditure share of manufacturing has been declining since the 1950/60s

---

[4]Acemoglu and Autor (2011) and Goos et al. (2014) look at the contribution of between-industry shifts to the polarization of occupational employment, but do not analyze the effect of structural change on the polarization of the labor market.

[5]While we focus in the main text on the link between industrial and occupational structure since 1950/1960, we show in Appendix A.10 using data over 1850–1940 that there is a close connection also over much longer time series, see Figure 16 and Table 13.

in the US, while those in services have been increasing. From an empirical perspective, we add to this literature by documenting that in the US the employment patterns are mimicked by the path of relative average wages. The economic mechanisms put forward in the literature for structural transformation combine specific features of preferences with some form of technological progress. Some papers assume non-homothetic preferences, such that changes in aggregate income – coming from technological progress – lead to a reallocation of employment across sectors (Kongsamut, Rebelo, and Xie (2001), Boppart (2014)). Other papers focus on differential total factor productivity (TFP) growth across sectors (Ngai and Pissarides (2007)) or on changes in the supply of an input used by different sectors with different intensities (Caselli and Coleman (2001), Acemoglu and Guerrieri (2008)).

We build on the model of Ngai and Pissarides (2007) closely, with one important modification: we explicitly model sectoral labor supply. As our goal is to study the joint evolution of employment and wages, we introduce heterogeneity in workers' skills, who endogenously sort into different sectors. In order to meet increasing labor demands in certain sectors – driven by structural change – the relative wages of those sectors have to increase. Since we model the sector of work choice, we can analyze the effects of structural change on relative sectoral wages, which is not common in models of structural change.[6] Another modification of Ngai and Pissarides (2007) is that we do not model capital, as our interest is in the heterogeneity of labor supply. The change in relative sectoral labor productivity can be driven by differential sectoral TFP changes or by capital accumulation and different sectoral capital intensities.[7] We stay agnostic about the origin of the differential labor productivity growth across sectors. As Goos et al. (2014) point out it is possible that part of this since the 1980s or 1990s is driven by different routine intensities and ICT. While the spread of ICT is not likely to be driving the differential sectoral productivity growth pre-1980, other forms of mechanization could have had a stronger impact in the manufacturing sector than in services over the longer time horizon that we study.

Ours is not the first paper to consider sectoral choice in a model of structural

---

[6]A notable exception is Caselli and Coleman (2001).

[7]For example if ICT is used more intensively in the manufacturing sector, and ICT becomes cheaper, then this would show up as an increase in the relative productivity of manufacturing workers.

change. The setup of Matsuyama (1991) is similar, where agents have different efficiencies across sectors, but focuses on the theoretical possibility of multiplicity of stationary steady states. Caselli and Coleman (2001) study the role of falling costs of education in the structural shift from agriculture to manufacturing, and they derive predictions about the relative wages in the farm and non-farm sector. Focusing on cross-country differences, Lagakos and Waugh (2013) show that self-selection can account for gaps in productivity and wages between agriculture and non-agriculture. Buera and Kaboski (2012) analyze the relation between the increasing value added share of the service sector and the increasing skill premium, without exploring their model's implications for sectoral employment or wages, whereas this is the focus of our paper.

The polarization literature typically focuses on employment and wage patterns after the 1980s or 1990s. We contribute to this literature by documenting that in the US the polarization of occupations in terms of wages and employment has started as early as the 1950s. As mentioned before, the leading explanation is routinization linked to ICT. While the spread of ICT is a convincing explanation for the polarization of labor markets after the mid-1980s, it does not provide an explanation for the patterns observed earlier.[8] But it is conceivable that similar factors, not linked to ICT, but to other forms of mechanization might have contributed to polarization earlier. Another explanation suggested in the literature are consumption spillovers. This argument suggests that as the income of high-earners increases, their demand for low-skilled service jobs increases as well, leading to a spillover to the lower end of the wage distribution (Manning (2004), Mazzolari and Ragusa (2013)). We do not incorporate such a mechanism in our model, as we strive for the most parsimonious setup featuring structural change, which does a good job in replicating the basic sectoral labor market facts since the 1960s.

The remainder of the paper is organized as follows: section 2 lays out our empirical findings, section 3 our theoretical model, section 4 the quantitative results, and section

---

[8]Another explanation is the increasing off-shorability of tasks (rather than finished goods), as first emphasized by Grossman and Rossi-Hansberg (2008). It has been argued that it is largely the middle-earning jobs that are off-shorable, but the evidence is mixed (Blinder (2009), Blinder and Krueger (2013), Acemoglu and Autor (2011)). Just as for the ICT routinization hypothesis, this mechanism could have explanatory power from the 1980s onwards.

5 concludes.

2 Polarization in the data

Using US Census data between 1950 and 2000 and the 2007 American Community Survey (ACS), we document the following three facts: 1) polarization in terms of occupations – contrary to the consensus – started as early as the 1950/60s, 2) wages and employment have been polarizing in terms of broadly defined industries as well, 3) a significant part of employment and wage polarization in terms of occupations is driven by industry level changes. The focus of our quantitative model is fact 2). We document fact 1) as most of the literature documents polarization in terms of occupations. The employment part of fact 3) has been documented in the literature using shift-share decompositions (Acemoglu and Autor (2011), Goos et al. (2014)), here we confirm it for our classification, data, and time horizon. We also conduct a similar decomposition for wages and show that a significant part of relative occupational wage changes are driven by industry effects. We report fact 3) to convince the reader whose main interest is in occupations, that for the full picture one needs to consider industries as well. In what follows we document each of these facts in detail.

2.1 Polarization in terms of occupations

In the empirical literature, polarization is mostly represented in terms of occupations. We document polarization in terms of two occupational classifications: we start from the finest balanced occupational codes possible, and then go to ten broad occupational categories.

Following the methodology used in Autor et al. (2006), Acemoglu and Autor (2011), and Autor and Dorn (2013), we plot the smoothed changes in log real wages and employment shares for occupational percentiles, where occupations are ranked according to their 1980 mean hourly wages.[9] The novelty in these graphs is that we show

---

[9]We split occupations into 100 groups, each representing 1 percent of employment in 1980. We smooth changes in log real hourly wages and employment shares with a locally weighted regression using a bandwidth of 0.8.
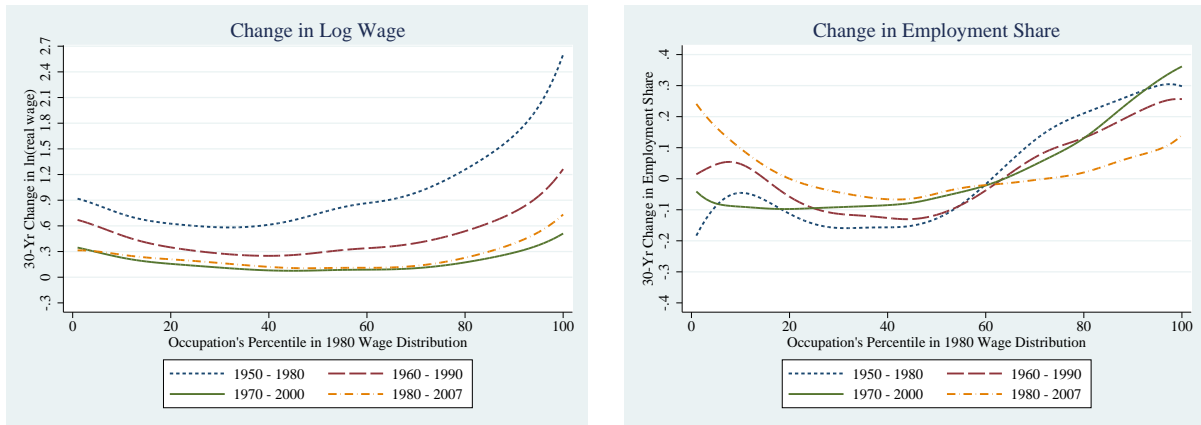
Figure 1: Smoothed changes in wages and employment

Notes: The data is taken from IPUMS US Census data for 1950, 1960, 1970, 1980, 1990, 2000 and the American Community Survey (ACS) for 2007. The sample excludes agricultural occupations/industries and observations with missing wage data; the details are given in Appendix A. Balanced occupation categories (183 of them) were defined by the authors based on Meyer and Osborne (2005), Dorn (2009) and Autor and Dorn (2013). The horizontal axis contains occupational skill percentiles based on their 1980 mean wages (see Appendix A for details). In the left panel the vertical axis shows for each occupational skill percentile the 30-year change in log hourly real wages, whereas in the right panel it shows the 30-year change in employment shares (calculated as hours supplied).

these patterns going back until 1950, rather than focusing only on the post-1980 period.[10] In both graphs, each of the four curves represent changes which occurred over a different 30-year period. The left panel in Figure 1 shows that there has been polarization in terms of real wages in all 30-year periods, since the real wage change is larger for low- and for high-ranked occupations than it is for middle-ranked occupations. The polarization of real wages is most pronounced in the first two 30-year intervals, but it is clearly discernible in the following ones as well from the slight U-shape of the smoothed changes. The right panel shows the smoothed employment share changes. The picture shows that employment did not move monotonically towards higher-wage occupations, instead it seems that middle-earning occupations lost the most in terms of employment. Thus employment polarization is present in the sense that the employment share in low- and high-wage occupations increased more (or decreased less) than in middle-wage occupations. Polarization in terms of employment is most pronounced in the last 30 years (1980-2007), but it seems to be present even in

---

[10]For comparability with the literature, we rank occupations based on their mean hourly wage in 1980. However, given that we look at a longer horizon than most of the literature, we also plot these changes against a different ranking of occupations, one based on the 1950 mean hourly real wages. The patterns look the same, see Figure 10 and the discussion in Appendix A.

the earlier periods.[11]

We focus on 30-year windows for two reasons. First, most of the literature documents polarization over periods longer than two decades. Second, we link polarization to structural change, which is a long-run phenomenon. One concern with showing 30-year windows, as in Figure 1, is that they stay silent on the exact timing of when polarization started. To address this, we show the decade-by-decade version of this in Figure 11 in Appendix A. This figure shows that these patterns do not necessarily hold for a decade-by-decade analysis, neither in the earlier nor in the later part of the sample. In some decades the top gains, whereas in others the bottom gains, but it is never the middle that grows the most in terms of employment shares.[12] However, between 1960 and 1970 there is clear evidence of polarization, therefore the early polarization patterns in the 30-year windows are not solely driven by changes after 1980.

A set of balanced occupational categories is needed to generate Figure 1. Meyer and Osborne (2005) develop a set of harmonized occupational codes for the 1950 to 2000 Census and the ACS data. Dorn (2009) aggregates Meyer and Osborne (2005) to achieve the finest possible balanced set of categories from 1980 onward. We base our categories on Meyer and Osborne (2005) and Dorn (2009) to similarly achieve the finest possible balanced set of occupations from 1950 onward. One concern with this approach is that despite the efforts of these authors, their harmonized and/or balanced categories are not truly comparable across Census years, and the reader might worry, that looking at a longer horizon, as we do, only exacerbates this problem. However, the biggest change in occupational classification occurred with the implementation of the Standard Occupational Classification (SOC) based occupation codes in the 2000 Census, when the hierarchical structure of occupational codes was drastically modified. While in previous Census years certain smaller occupation categories disappeared or entered, the main structure of occupations remained the same. [13] Therefore

---

[11]When ranking occupations in terms of the 1950, i.e. the initial, wage distribution, as is commonly done in this literature, employment polarization is much more noticeable since the 1960s, see Figure 10 in Appendix A.

[12]The fact that over 10-year windows the polarization patterns do not hold for a very fine classification of occupations is in line with the evidence in Acemoglu and Autor (2011), see their Figure 10. Nonetheless, with a coarser classification of occupations, one sees polarization decade-by-decade, see Figure 2.

[13]See the documentation of the Integrated Public Use Microdata Series (IPUMS) Census data.

looking at longer horizons does not worsen the comparability issue relative to existing literature which typically focuses on the post-1980 period.

Nonetheless, to minimize the comparability issues of fine occupational categories arising from the reclassifications across Census years, we aggregate up these fine categories to a coarser set of occupations. For these coarser categories there is less of a concern about the consistency over time. The bar charts in Figure 2 show for ten broad occupational groups percentage changes in total hours worked and in the mean log wages over ten year intervals. This categorization follows Acemoglu and Autor (2011) who show with this methodology the employment changes post 1980. We follow their ranking of occupations: the right three groups are the most educated and highest paid occupations, the four in the middle are middle-skilled, while the three on the left are the least educated and lowest paid occupations.
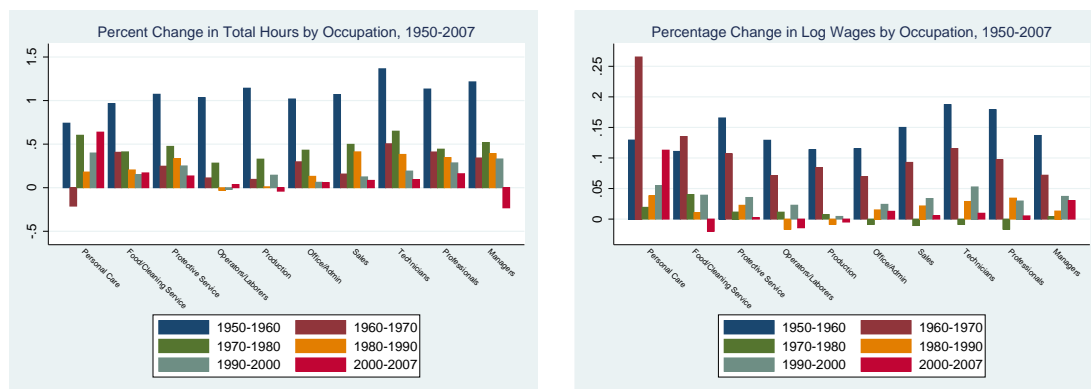
Figure 2: Polarization in broad occupational categories

Notes: These bar graphs show for ten broad occupational categories, as defined in Acemoglu and Autor (2011), the decade-by-decade percentage change in hours worked (left panel) and in the mean log wages (right panel).

We expand on Acemoglu and Autor (2011) in two ways: 1) by showing the trends in decennial employment growth from 1950 onwards (in the left panel), and 2) by also showing wage growth (right panel). While over 1950–1960, there is no clear pattern in the employment growth of these broad occupational categories, from 1960 onwards there is a U-shaped pattern. Total hours worked grew by more for occupations at the higher and at the lower end of the skill distribution than for those in the middle. A similar pattern is evident in the wage growth rates. Over each ten year period, wage growth was the lowest in middle-skilled occupations. We conclude that even when grouping occupations into broad categories following the methodology of Acemoglu

and Autor (2011), there is evidence of employment and wage polarization as early as 1960.[14]

2.2 Polarization in terms of sectors

Next we document the polarization of employment and wages in terms of three broad industries or sectors: low-skilled services, manufacturing, and high-skilled services.
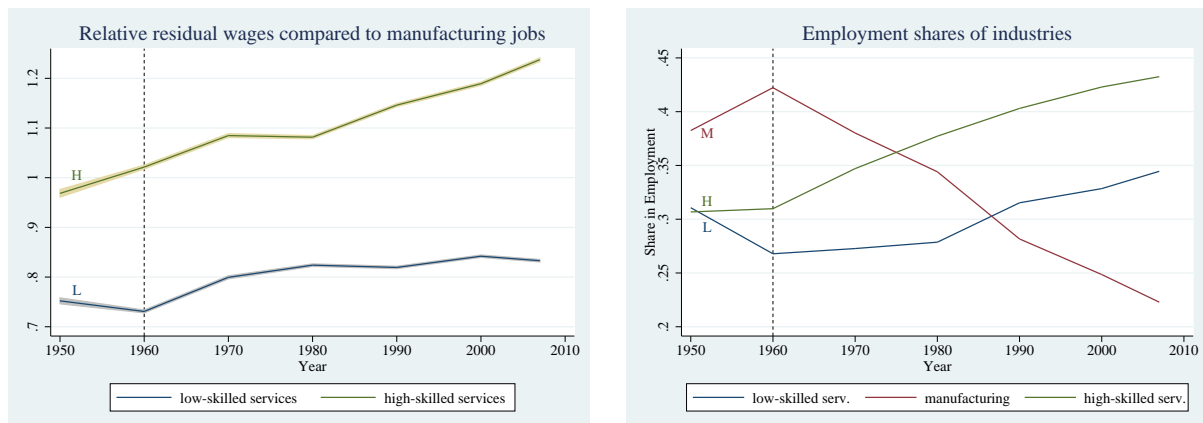
Figure 3: Polarization for broad industries

Notes: The data used is the same as in Figure 1. Each worker is classified into one of three sectors based on their industry code (for details of the industry classification see text and Appendix A.2). The left panel shows relative wages: the high-skilled service and the low-skilled service premium compared to manufacturing (and their 95% confidence intervals), implied by the regression of log wages on gender, race, a polynomial in potential experience, and sector dummies. The right panel shows employment shares, calculated in terms of hours worked. The dashed vertical line represents 1960, from when on manufacturing employment has been contracting.

Our classification for the manufacturing sector includes also mining and construction, as is common in the structural change literature (e.g. Herrendorf, Rogerson, and Valentinyi (2013)). As mentioned in the introduction, we split the remaining (service) industries into two categories based on consumption side considerations: within sectors the industries should be close substitutes, whereas across sectors they should be complements. The categorization is also guided by differences on the production side; in low-end services workers have much less education and earn much lower wages than in high-end services. For this reason we refer to them as low-skilled and high-

---

[14]In Figure 12 in Appendix A, we document polarization in terms of occupations in an even coarser classification. Following Acemoglu and Autor (2011) we classify occupation groups into three categories: manual, routine, and abstract. Again, we find that the middle earning group, the routine workers, lost both in terms of relative average wages and employment share to the benefit of manual and abstract workers.

skilled services.[15] As a result of the combined production and consumption side considerations we classify as low-end services the following industries: personal services, entertainment, low-skilled transport, low-skilled business and repair services, retail trade, and wholesale trade. High-end services comprise of professional and related services, finance, insurance and real estate, communications, high-skilled business services, utilities, high-skilled transport, and public administration.

Figure 3 documents the patterns of polarization both in terms of employment shares and wages for the above defined sectors between 1950/1960 and 2007. The right panel shows the path of employment shares: high-skilled services increase continuously, low-skilled services increase and manufacturing decreases from 1960 onwards.[16] In terms of wages, we plot the sector premium in high-skilled and low-skilled services compared to manufacturing, as well as their 95 percent confidence intervals. These sector premia are the exponents of the coefficients on sector dummies, which come from a regression of log wages where we control also for gender, race, and a polynomial in potential experience.[17] We plot these rather than the relative average wages, because in our quantitative exercise we do not aim to explain sectoral wage differentials that are potentially caused by age, gender or racial composition differences and the differential path of these across sectors.[18] As the graph shows, low-skilled service workers earn less, whereas high-skilled service workers earn more than manufacturing workers. Since the 1960s both low- and high-skilled service workers have been gaining in terms of wages compared to manufacturing workers.[19] To summarize, from 1960 onwards there is clear polarization in terms of these three sectors: the low- and high-skilled service workers gained in terms of employment and wages at the expense

---

[15]See Figure 3, as well as Figure 13 and Table 5 in Appendix A.

[16]Between 1950 and 1960 manufacturing employment increased, and low-skilled service employment fell.

[17]See Table 6 in Appendix A.3 for details of the regression.

[18]One might be concerned that the employment share changes are driven by changes in the age, gender, or race composition of the labor force. To assess this, we generate counterfactual industry employment shares by fixing the industry employment share of each age-gender-race cell at its 1960 level, and allowing the employment shares of the cells to change. This exercise confirms that to a large extent the employment share changes are not driven by the compositional changes of the labor force. See Figure 14 in Appendix A.5.

[19]These trends are very robust, they hold in the raw data (see the left panel in Figure 13 in Appendix A.3) as well as when constructing the sector premia using different specifications of the log wage regression (see Table 8 in Appendix A.4).

of the middle-earning, middle-skilled, manufacturing workers.

2.3 Polarization across occupations linked to industry shifts

To quantify the contribution of sectoral employment shifts to each occupation's employment share path, we conduct a standard shift-share decomposition.[20] The overall change in the employment share of occupation o between year 0 and t, $\Delta E_{ot} = E_{ot} - E_{o0}$, can be expressed as:

$$\Delta E_{ot} = \underbrace{\sum_i \lambda_{oi} \Delta E_{it}}_{\equiv \Delta E_{ot}^B} + \underbrace{\sum_i \Delta \lambda_{oit} E_i}_{\equiv \Delta E_{ot}^W}, \tag{1}$$

where $\lambda_{oit} = L_{oit}/L_{it}$ denotes the share of occupation o, industry i employment within industry i employment at time t, $E_{it} = L_{it}/L_t$ denotes the share of industry i employment within total employment at time t, Δ denotes the change between period 0 and t, and the variables without a time subscript denote the average of the variable between period 0 and period t. ΔE_o^B represents the change in the employment share of occupation o that is attributable to changes in industrial composition, i.e. structural transformation, while ΔE_o^W reflects changes driven by within sector forces.

Table 1 shows for three broad occupational categories the total employment share change and its decomposition between 1950 or 1960 and 2007 into a between-industry and a within-industry component. We use either 3 occupational and 3 sectoral categories, or similarly to Acemoglu and Autor (2011), 10 occupations and 11 industries to be sure that our results are not driven by the coarse categorization. This decomposition indicates that a significant part of the occupational employment share changes are driven by shifts in the industrial composition of the economy between 1950 and 2007. Between-industry shifts are the most important in manual occupations, and the least important in abstract occupations, where they still account for at least a quarter of the total change.

The magnitude of the role of between-industry shifts in our analysis is similar to

---

[20]An alternative way is to calculate how much occupational employment shares would have changed, if industry employment shares would have remained at their 1960 level. See Figure 15 in Appendix A.6.

Table 1: Decomposition of changes in occupational employment shares

| | Employment shares | | | |
| | 3 x 3 | | 10 x 11 | |
	1950-2007	1960-2007	1950-2007	1960-2007
Manual				
Total Δ	2.98	5.68	2.98	5.68
Between Δ	2.30	3.07	3.13	4.38
Within Δ	0.67	2.61	-0.15	1.30
Routine				
Total Δ	-19.79	-19.14	-19.79	-19.14
Between Δ	-5.66	-6.32	-9.73	-10.01
Within Δ	-14.13	-12.82	-10.06	-9.13
Abstract				
Total Δ	16.81	13.46	16.81	13.46
Between Δ	3.35	3.24	6.60	5.63
Within Δ	13.46	10.21	10.21	7.83

Notes: Same data as in Figure 1. For each occupational category, the first row presents the total change, the second the between-industry component, and the third the within-industry component over the period 1950 or 1960 to 2007. The first two columns use 3 occupations and 3 sectors, the last two use 10 occupations and 11 industries. The 10 occupations are the same as in Figure 2, while the 11 industries are: 1 personal services; entertainment and low-skilled business and service repairs, 2 low-skilled transport, 3 retail trade, 4 wholesale trade, 5 extractive industries, 6 construction, 7 manufacturing, 8 professional and related services and high-skilled business services, 9 finance, insurance, and real estate, 10 high-skilled transport and public utilities (incl. communications), 11 public administration.

that in Acemoglu and Autor (2011) and Goos et al. (2014).[21] Goos et al. (2014) argue that part of the between-industry shifts can be driven by routinization, which is a within-industry phenomenon. Since routinization has a bigger impact on industries where routine labor is used more intensively, employment might shift away from these industries. While this is a valid concern, routinization, linked to ICT, is not likely to be driving the faster productivity growth observed between the 1950s and 1980s. We also conduct an alternative shift-share decomposition where we use sectoral value added shares instead of sectoral employment shares. Even though the importance of the between industry component is somewhat smaller than in the standard shift-share

---

[21] Acemoglu and Autor (2011) use US Census data between 1960 and 2000, and the ACS 2008 data. Their focus is the declining importance of between-industry shifts from 1960-1980 to 1980-2007. We find some support for this in our decade-by-decade analysis shown in Table 9 in Appendix A.6. The relatively smaller contribution of between-industry shifts in later periods might be due to routinization kicking in after the 1980s, thus providing an extra force for within-industry reallocation of labor. Nonetheless, we find that even in the 1980-2007 period, between-industry shifts explain a significant part of occupational employment share changes. Goos et al. (2014) use data for 16 European countries between 1993 and 2010, and attribute a roughly equal role to between and within industry shifts in all occupations.

decomposition, it is nonetheless a substantial share of the overall change.[22]

We also decompose relative occupational wage changes into an occupation driven component and an industry driven component. We define the relative average wage of occupation o as the ratio of the occupation's average wage relative to the average wage in routine occupations:[23]

$$rw_{ot} \equiv \sum_i \frac{L_{iot}}{L_{ot}} \frac{w_{it}}{w_{rt}} \frac{w_{iot}}{w_{it}} = \sum_i \chi_{iot} rw_{it} p_{iot},$$

where $\chi_{iot} = L_{iot}/L_{ot}$ is employment in industry i and occupation o in period t relative to employment in occupation o in period t, $rw_{it} = w_{it}/w_{rt}$ is the ratio of the average wage in industry i in period t relative to the average wage of routine occupations in period t, and $p_{iot} = w_{iot}/w_{it}$ is the premium of occupation o in industry i in period t. In this three-way decomposition, the *industry effect* is itself composed of two parts: the first part captures changes coming from workers within an occupation moving across industries with different wages (L_{iot}/L_{it}). The second part captures that each industry's relative average wage path (w_{it}/w_{rt}) influences the overall relative average wage of the occupation. The *occupation effect* is driven by the change in the occupational premium within each industry (w_{iot}/w_{it}).

Table 2 shows the change in manual and abstract wages relative to routine wages, and their decomposition into an industry and an occupation component, between 1950 or 1960 and 2007. This Table confirms that both manual and abstract occupations have been gaining in terms of wages relative to routine jobs. The decomposition shows that between half and two thirds of the gain in relative manual wages is driven by industry effects, by the reallocation of manual labor to higher-paying industries, or by faster wage growth in industries where more manual workers are employed. For the increase in relative abstract wages the results are even more striking: all of the gain is driven by industry effects.[24]

---

[22]See Appendix A.7 for details.

[23]It is clear that the wage has to be a relative wage, otherwise the decomposition picks up the general upward trend in wages, and assigns it to the component which includes a wage change. Since we are interested in the path of manual and abstract wages relative to routine wages, it is natural to normalize by the average wage in routine occupations.

[24]In general there are three ways of conducting a three-way decomposition depending on which element one separates first. The other decompositions give virtually the same results, see Appendix

Table 2: Decomposition of changes in relative occupational wages

	Relative wages			
	3 x 3		**10 x 11**	
	1950-2007	1960-2007	1950-2007	1960-2007
Manual/Routine				
Total Δ	0.289	0.310	0.289	0.310
Industry Δ	0.180	0.148	0.225	0.218
Occupation Δ	0.108	0.162	0.064	0.093
Abstract/Routine				
Total Δ	0.327	0.240	0.327	0.240
Industry Δ	0.310	0.254	0.376	0.317
Occupation Δ	0.016	-0.014	-0.050	-0.077

Notes: Same data as in Figure 1. For each occupational category, the first row presents the total change, the second the industry component, and the third the occupation component over the period 1950 or 1960 to 2007. The first two columns use 3 occupations and 3 sectors, columns three and four 10 occupations and 11 industries.

To summarize, first we document that polarization defined over occupational categories both in terms of employment and wages has been present in the US since the 1950s. Second we show that the same patterns are discernible in terms of three broad sectors: low-skilled services, manufacturing and high-skilled services. Finally, we show that over the last six decades a significant amount of the employment share changes and relative wage changes in occupations are driven by the (employment) shifts across industries.

In the remainder of the paper we present a simple model of sorting and structural change to jointly explain the sectoral shifts in employment and the changes in average sectoral wages. We then calibrate the model to quantitatively assess how much of the polarization of sectoral employment and wages it can explain over the last fifty years, when feeding in sectoral labor productivity from the data.

3   Model

In order to illustrate the mechanism that is driving the polarization of wages and employment, we present a parsimonious static model, and analyze its behavior as productivity levels increase across sectors. The key novel feature of our model is that

---

A.8. The decomposition of the decade-by-decade wage changes shows broadly similar patterns, see Table 12 in Appendix A.9.

we assume that each sector values different skills in its production process. Relaxing the assumption of the homogeneity of labor allows us to derive predictions, not only about the employment and expenditure shares, but also about the relative average wages across sectors over time.

We assume that the economy is populated by heterogeneous agents, who all make individually optimal decisions about their sector of work. Every individual chooses their sector of work to maximize wages, in a Roy-model type setup. We assume that individuals are ex ante heterogeneous in their efficiency units of labor in low-skilled services, manufacturing and high-skilled services, and thus endogenously sort into the sector where the return to their labor is the highest.

Furthermore these individuals are organized into a stand-in household, which maximizes its utility subject to its budget constraint.[25] Households derive utility from consuming high- and low-skilled services and manufacturing goods.

The economy is in a decentralized equilibrium at all times: individuals make sectoral choices to maximize their wages, the stand-in household collects all wages and maximizes its utility by optimally allocating this income between low-skilled services, manufacturing goods and high-skilled services. Production is perfectly competitive, wages and prices are such that all markets clear. We analyze the qualitative and quantitative role of technological progress in explaining the observed sectoral wage and employment dynamics since the 1960s.

3.1 Sectors and production

There are three sectors in the model: high-skilled services (H), manufacturing (M), and low-skilled services (L). All goods and services are produced under perfect competition, and each sector uses only labor as an input into production.
The technology to produce in each sector is:

$$Y_j = A_j N_j \quad \text{for} \quad j \in \{L, M, H\}, \tag{2}$$

---

[25]We make the assumption of a stand-in household purely for expositional purposes. Given that the preferences we use are homothetic (see section 3.2.2), the resulting sectoral demands are equal to the aggregation of individual demands.

where A_j is productivity and N_j is the total amount of efficiency units of labor (efficiency labor for short) hired in sector J for production. Sector J firms are price takers, therefore the equilibrium wage per efficiency unit of labor (unit wage for short) in this sector has to satisfy:

$$\omega_j = \frac{\partial p_j Y_j}{\partial N_j} = p_j A_j \quad \text{for} \quad j \in \{L, M, H\}. \tag{3}$$

Note that the wage of a worker with a efficiency units of sector J labor when working in sector J is $a\omega_j$.

3.2 Labor supply and demand for goods

The stand-in household consists of a measure one continuum of different types of members. Each member chooses which one of the three market sectors to supply his one unit of raw labor in. The household collects the wages of all its members and decides how much low-skilled services, manufacturing goods and high-skilled services to buy on the market.

3.2.1 Sector of work

We assume that every member of the household works full time in one of the three market sectors. Since every member can work in any of the three sectors, and each member's utility is increasing in his own wages (as well as in all other members' wages), it is optimal for each worker to choose the sector which provides him with the highest wages.

Individuals are heterogeneous in their endowment of efficiency units of labor, $\mathbf{a} \in \mathbf{R}_+^3$, which is drawn from a time invariant distribution $f(\mathbf{a})$. This is an innate ability distribution, and as such is prior to any form of human capital that a worker might accumulate, for example by acquiring education. Even though we do not model this explicitly, one could think of our model as having a reduced-form educational choice in the following sense. If a worker given their ability selects a sector, say H, to work in, they will on the way get education, or other qualifications, as necessary to enter

18

that sector.[26]

The endowment, **a**, determines each individual's productivity in sector L, M and H. We assume that each dimension of ability corresponds to one sector, such that $a_l \equiv \mathbf{a}(1)$ denotes the individual's efficiency units of labor in low-skilled services, $a_m \equiv \mathbf{a}(2)$ in manufacturing, while $a_h \equiv \mathbf{a}(3)$ denotes his efficiency units in high-skilled services. Therefore the wage of an individual with $\mathbf{a} = (a_l, a_m, a_h)$ efficiency units of labor if working in sector J is $a_j \omega_j$. Given wage rates $\omega_l, \omega_m, \omega_h$ per efficiency unit of labor the optimal decision of any agent can be characterized as follows.

Result 1. *Given unit wage rates ω_l, ω_m and ω_h, the optimal sector choice of individuals can be characterized by two relative unit wages:*

$$\frac{\omega_l}{\omega_m} \quad and \quad \frac{\omega_h}{\omega_m}.$$

It is optimal for an individual with (a_l, a_m, a_h) efficiency units of labor to work in

- *sector L if and only if*

$$a_l \geq \frac{1}{\frac{\omega_l}{\omega_m}} a_m \quad and \quad a_l \geq \frac{\frac{\omega_h}{\omega_m}}{\frac{\omega_l}{\omega_m}} a_h. \tag{4}$$

- *sector M if and only if*

$$a_m \geq \frac{\omega_l}{\omega_m} a_l \quad and \quad a_m \geq \frac{\omega_h}{\omega_m} a_h. \tag{5}$$

- *sector H if and only if*

$$a_h \geq \frac{\frac{\omega_l}{\omega_m}}{\frac{\omega_h}{\omega_m}} a_l \quad and \quad a_h \geq \frac{1}{\frac{\omega_h}{\omega_m}} a_m \tag{6}$$

Figure 4 shows this endogenous sorting behavior. The left panel shows that for a given efficiency in high-skilled services, a_h^0, individuals who have relatively low efficiency in both manufacturing and low-skilled services sort into H (the green horizontally striped area), those with relatively higher efficiency in low-skilled services sort into L (the blue dotted area), while those with relatively higher efficiency in manufacturing sort into M (the red vertically striped area). The middle and right panel show

---

[26]Given that high-skilled services is the most intensive sector in college educated workers, the increase in college education that the US has witnessed over the last decades is in line with our model, as it generates an increase in the demand for high-skilled services, providing a link from structural change to educational attainment (see for example also Buera, Kaboski, and Rogerson (2015)).
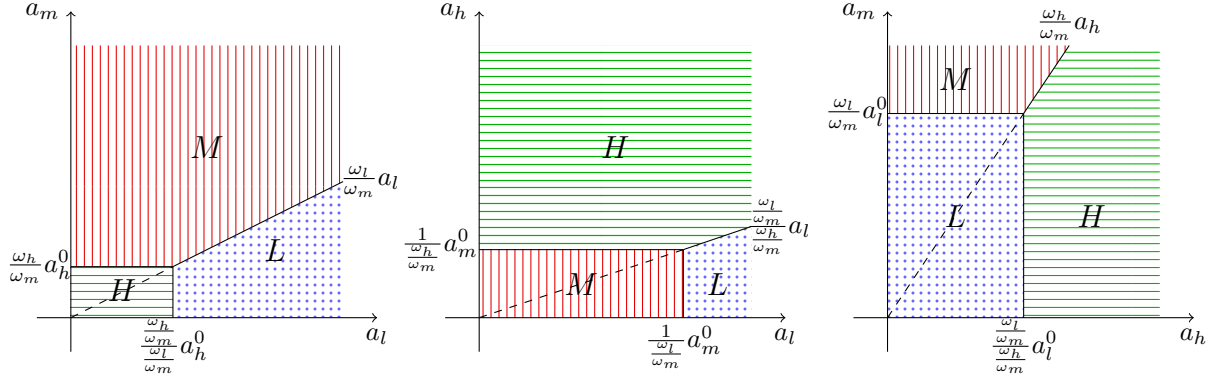
Figure 4: Optimal sector of work

This figure depicts the optimal sector of work choices as a function of ω_l/ω_m and ω_h/ω_m from three perspectives. The left panel shows for a given a_h^0 value the optimal sector choice for all $\{a_l, a_m\}$ pairs. The middle panel shows for a given a_m^0 value the optimal sector choice for all $\{a_l, a_h\}$ pairs. The right panel shows for a given a_l^0 value the optimal sector choice for all $\{a_h, a_m\}$ pairs. The blue dotted areas are where L is the optimal sector, the red vertically striped areas show where M is optimal, and the green horizontally striped areas show where H is optimal.

the same for a given manufacturing and low-skilled service efficiency. The comparative statics of optimal sorting with respect to changes in the relative unit wages is straightforward.

The optimal sector of work choices of individuals determine the effective labor supplies in the three markets:

$$N_l\left(\frac{\omega_l}{\omega_m}, \frac{\omega_h}{\omega_m}\right) = \int_0^\infty \int_0^{\frac{\omega_l}{\omega_m}a_l} \int_0^{\frac{\omega_l}{\omega_m}a_l}{\frac{\omega_h}{\omega_m}} a_l f(a_l, a_m, a_h) da_h da_m da_l, \tag{7}$$

$$N_m\left(\frac{\omega_l}{\omega_m}, \frac{\omega_h}{\omega_m}\right) = \int_0^\infty \int_0^{\frac{1}{\frac{\omega_l}{\omega_m}}a_m} \int_0^{\frac{1}{\frac{\omega_h}{\omega_m}}a_m} a_m f(a_l, a_m, a_h) da_h da_l da_m, \tag{8}$$

$$N_h\left(\frac{\omega_l}{\omega_m}, \frac{\omega_h}{\omega_m}\right) = \int_0^\infty \int_0^{\frac{\frac{\omega_h}{\omega_m}}{\frac{\omega_l}{\omega_m}}a_h} \int_0^{\frac{\omega_h}{\omega_m}a_h} a_h f(a_l, a_m, a_h) da_m da_l da_h. \tag{9}$$

The effective labor supply in sector J is the total amount of sector J efficiency units, a_j, supplied to that sector. This is not measurable in the data, but we observe hours worked, which corresponds to the raw labor supply, or employment share in the model. These employment shares are the mass of individuals who supply their la-

bor in the given sectors:

$$L_l\left(\frac{\omega_l}{\omega_m},\frac{\omega_h}{\omega_m}\right)=\int_0^{\infty}\int_0^{\frac{\omega_l}{\omega_m}a_l}\int_0^{\frac{\omega_l}{\omega_m}a_l}\frac{\omega_h}{\omega_m}}f(a_l,a_m,a_h)da_h da_m da_l,\tag{10}$$

$$L_m\left(\frac{\omega_l}{\omega_m},\frac{\omega_h}{\omega_m}\right)=\int_0^{\infty}\int_0^{\frac{1}{\frac{\omega_l}{\omega_m}}a_m}\int_0^{\frac{1}{\frac{\omega_h}{\omega_m}}a_m}f(a_l,a_m,a_h)da_h da_l da_m,\tag{11}$$

$$L_h\left(\frac{\omega_l}{\omega_m},\frac{\omega_h}{\omega_m}\right)=\int_0^{\infty}\int_0^{\frac{\frac{\omega_h}{\omega_m}}{\frac{\omega_l}{\omega_m}}a_h}\int_0^{\frac{\omega_h}{\omega_m}a_h}f(a_l,a_m,a_h)da_m da_l da_h.\tag{12}$$

In a similar vein, ω_l, ω_m and ω_h are the sectoral unit wages, which are in general also not observed in the data, but sectoral average wages are. These are simply the total earnings in a sector divided by the mass of people working in the sector, or equivalently the unit wage times the average worker efficiency ($\overline{a}_j = N_j/L_j$) in the given sector:

$$\overline{w}_j=\frac{\omega_j N_j}{L_j}\equiv\omega_j\overline{a}_j\quad\text{for}\quad j\in\{L,M,H\}.\tag{13}$$

3.2.2 Demand for consumption goods and services

Household members derive utility from low-skilled services, manufacturing goods and high-skilled services. The household allocates total income earned by household members to maximize the following utility:

$$\max_{C_l,C_m,C_h} u\left(\left[\theta_l C_l^{\frac{\varepsilon-1}{\varepsilon}}+\theta_m C_m^{\frac{\varepsilon-1}{\varepsilon}}+\theta_h C_h^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}\right)$$

$$\text{s.t.}\quad p_l C_l+p_m C_m+p_h C_h\leq\omega_l N_l+\omega_m N_m+\omega_h N_h$$

where u is any monotone increasing function, $\omega_l L_l+\omega_m N_m+\omega_h N_h$ are the total wages of household members, p_l, p_m, and p_h are the prices of the low-skilled services, the manufacturing goods, and the high-skilled services.

The household's optimal consumption bundle has to satisfy:

$$\frac{C_l}{C_m} = \left(\frac{p_l}{p_m}\frac{\theta_m}{\theta_l}\right)^{-\varepsilon}, \tag{14}$$

$$\frac{C_h}{C_m} = \left(\frac{p_h}{p_m}\frac{\theta_m}{\theta_h}\right)^{-\varepsilon}. \tag{15}$$

3.3 Competitive equilibrium and structural change

A competitive equilibrium is given by relative unit wage rates $\left\{\frac{\omega_l}{\omega_m}, \frac{\omega_h}{\omega_m}\right\}$, prices $\{p_l, p_m, p_h\}$, and consumption demands $\{C_l, C_m, C_h\}$, given productivities $\{A_l, A_m, A_h\}$, where individuals, households and firms make optimal decisions, and all markets clear.

Using goods market clearing in all sectors ($Y_j = C_j$ for $j \in \{L, M, H\}$), where the supply is given by (2), and the market clearing unit wage rates, (3), in the household's optimality conditions, (14) and (15), we obtain the following:

$$\frac{A_l}{A_m}\frac{N_l}{N_m} = \Big(\underbrace{\frac{\omega_l}{\omega_m}\frac{A_m}{A_l}}_{=p_l/p_m}\frac{\theta_m}{\theta_l}\Big)^{-\varepsilon},$$

$$\frac{A_h}{A_m}\frac{N_h}{N_m} = \Big(\underbrace{\frac{\omega_h}{\omega_m}\frac{A_m}{A_h}}_{=p_h/p_m}\frac{\theta_m}{\theta_h}\Big)^{-\varepsilon}.$$

The left hand side is the relative supply, while the right hand side is the relative demand for low- and respectively high-skilled services compared to manufacturing. A change in the relative productivity affects both the relative supply and the relative demand.

An increase in relative manufacturing productivity compared to low-skilled service productivity (A_m/A_l) has two direct effects: (i) it reduces the relative supply of low-skilled services, (Y_l/Y_m), and (ii) through an increase in the relative price of low-skilled services (p_l/p_m), it lowers the relative demand for low-skilled services. If low-skilled services and manufacturing goods are complements, $\varepsilon < 1$, the effect through relative prices is the weaker one, and relative supply falls by more than relative demand. To restore equilibrium, the relative supply of low-skilled services has to increase and/or its relative demand has to fall compared to manufacturing.

In order for the relative supply to increase, the efficiency units of labor hired in low-skilled services have to increase relative to manufacturing, which requires a rise in the relative unit wage, ω_l/ω_m. At the same time, a rise in the relative unit wage also increases the relative price of low-skilled services, thus lowering the relative demand. Similarly, an increase in A_m/A_h, through its effect on relative supply and relative demand requires a rise in ω_h/ω_m.

Using the optimal sorting of individuals, (7), (8) and (9), we obtain the following expressions, which allow us to formally analyze the comparative static properties of the equilibrium:

$$\frac{N_l\left(\frac{\omega_l}{\omega_m}, \frac{\omega_h}{\omega_m}\right)}{N_m\left(\frac{\omega_l}{\omega_m}, \frac{\omega_h}{\omega_m}\right)} \left(\frac{\omega_l}{\omega_m}\right)^{\varepsilon} = \left(\frac{A_m}{A_l}\right)^{1-\varepsilon} \left(\frac{\theta_m}{\theta_l}\right)^{-\varepsilon}, \tag{16}$$

$$\frac{N_h\left(\frac{\omega_l}{\omega_m}, \frac{\omega_h}{\omega_m}\right)}{N_m\left(\frac{\omega_l}{\omega_m}, \frac{\omega_h}{\omega_m}\right)} \left(\frac{\omega_h}{\omega_m}\right)^{\varepsilon} = \left(\frac{A_m}{A_h}\right)^{1-\varepsilon} \left(\frac{\theta_m}{\theta_h}\right)^{-\varepsilon}. \tag{17}$$

These two equations implicitly define the relative unit wages, $\frac{\omega_l}{\omega_m}$ and $\frac{\omega_h}{\omega_m}$, and in turn these fully characterize the equilibrium of the economy.

Proposition 1. *When manufacturing goods and the two types of services are complements ($\varepsilon < 1$), then faster productivity growth in manufacturing than in the two types of services ($dA_m/A_m > dA_h/A_h = dA_l/A_l$) leads to a change in the optimal sorting of individuals across sectors. In particular ω_l/ω_m and ω_h/ω_m unambiguously increase, while ω_l/ω_h can rise or fall. This results in an unambiguous increase in efficiency labor in L and in H, and a reduction in efficiency and raw labor in M.*

Proof. Total differentiation of (16) and (17). See Appendix B for details. □

Proposition 1 confirms the results of Ngai and Pissarides (2007) in terms of efficiency labor, rather than raw labor or employment shares: when sectoral outputs are complements in consumption, effective labor inputs need to increase in the sectors which become relatively less productive. As manufacturing productivity grows the fastest, efficiency labor has to move out of manufacturing into both low- and high-skilled services. Since individuals optimally sort into the sector with the highest re-

turn for them, this implies that the equilibrium relative unit wages have to adjust. Proposition 1 states what these adjustments entail. The adjustment to the new equilibrium requires sector M to be squeezed from both sides, ω_l/ω_m and ω_h/ω_m have to increase. This is very intuitive: sector M has to shrink, while sector L and H have to expand, which requires sector M unit wages to fall both relative to sector H and sector L unit wages. This implies that in sector M not only efficiency units of labor, but also the raw employment (share) falls, while the employment share of overall services expands. It is worth to note that these results hold for any underlying distribution of efficiency units of labor, $f(a_l, a_m, a_h)$. However, in general it is ambiguous whether the boundary between L and H shifts up or down (ω_l/ω_h increases or decreases), and thus also whether the employment share of both L and H increase. Since in general the boundary between sector L and H changes, one of the sectors looses workers to the other sector, which might imply that the employment share of one of the service sectors falls. The effective employment of both service sectors unambiguously increases, as their gain from sector M always outweighs their potential loss to the other service sector.

In terms of relative average wages, in general it is not possible to sign the changes predicted by the model. The reason is self-selection; the workers leaving manufacturing are the ones that have a relatively low efficiency. As a consequence, the average efficiency in manufacturing increases when its employment share decreases. This tends to increase the average wage in manufacturing compared to the other sectors, offsetting to some extent the direct effect of the falling relative manufacturing unit wage. Without further assumptions, it is conceivable that the indirect effect through the average efficiency might overturn the direct effect of changing unit wages. To see this, consider the average low(high)-skilled service wage relative to the average manufacturing wage (from (13)):

$$\frac{\overline{w}_j}{\overline{w}_m} = \frac{\omega_j \frac{N_j}{L_j}}{\omega_m \frac{N_m}{L_m}} = \frac{\omega_j}{\omega_m} \frac{\overline{a}_j}{\overline{a}_m} \quad \text{for} \quad j \in \{L, H\}.$$

From Proposition 1 we know that ω_j/ω_m increases. Due to the changing nature of self-selection average efficiency in M increases, while in J it decreases, thus in general the

change in average low(high)-skilled service wages relative to average manufacturing wages is ambiguous. In general, the overall direction of change of the relative average wages depends on the exact form of the underlying distribution. We explore the change in relative average wages in more detail in the quantitative analysis, and in all our simulations the relative average wages (for both L and H to M) move in the same direction as the relative unit wages.

Since the structural change literature focuses on employment shares and value added shares, we also investigate our model's implications for relative value added. We can show that relative sectoral value added shares increase in the sectors with lower productivity growth if the sectoral outputs are complements in consumption.

Proposition 2. *When manufacturing goods and the two types of services are complements ($\varepsilon < 1$), then faster productivity growth in manufacturing than in the two types of services ($dA_m/A_m > dA_h/A_h = dA_l/A_l$), increases the relative value added in both high- and low-skilled services compared to manufacturing:*

$$d\frac{p_h Y_h}{p_m Y_m} > 0 \quad and \quad d\frac{p_l Y_l}{p_m Y_m} > 0.$$

These results can be understood by considering the following. In this model, sectoral value added is equal to the sectoral wage bill: $p_i Y_i = p_i A_i N_i = \omega_i N_i$. Proposition 1 tells us for $j \in \{L, H\}$ that ω_j/ω_m increases, that N_j increases, and N_m falls. Both relative unit wages and effective labor changes increase the value added output of sector J relative to sector M.

The sectoral value added can be further expressed as $p_i Y_i = \omega_i N_i = \overline{w}_i L_i$, since the sectoral wage bill can be expressed as either sectoral unit wage times sectoral efficiency labor, or as sectoral average wage times sectoral raw labor. Using this latter expression we can show that

$$\frac{p_i Y_i}{p_j Y_j} = \frac{\overline{w}_i}{\overline{w}_j} \frac{L_i}{L_j}.$$

According to our model relative sectoral value added has to equal the product of relative sectoral average wages and relative sectoral employment shares. This result holds

even if we include capital in the model, unless one assumes either imperfect capital mobility across sectors, or different sectoral capital intensities. Since in the data the relative sectoral value added does not equal the product of relative sectoral average wages and employment shares, in our calibration we target relative average wages and sectoral employment shares, as it is the evolution of these two measures that is the focus of our paper.

4   Quantitative results

In this section we quantitatively assess the contribution of structural transformation to the polarization of employment and wages across sectors. To do this we consider the evolution of the competitive equilibrium in terms of employment shares and relative average sectoral wages as productivity increases in manufacturing and in both low- and high-skilled services. We calibrate our parameters to match key moments in 1960, and then feed in the exogenous process for labor productivity to generate predictions for the evolution of employment and wages. We choose 1960 as the starting point for the quantitative evaluation of the model, because as documented in section 2.2 the contraction of manufacturing employment is apparent in our data from 1960 onwards. We first describe the data targets and the calibration strategy, and then discuss the quantitative importance of our mechanism.

4.1   Calibration

Four of the key moments are calculated from the 1960 Census data. These are the relative average sectoral wages, $\frac{w_l}{w_m}$ and $\frac{w_h}{w_m}$, and the sectoral employment shares, L_l, L_m and L_h, which sum to one. Employment shares are calculated as share of hours worked, and relative average wages are the sector premia, both as in section 2.2. The fifth moment is the dispersion of the non-transitory component of log non-agricultural wages, as in Lagakos and Waugh (2013), which we estimate from Panel Study of Income Dynamics (PSID) data from the years 1968–1975.[27]

---

[27] We rely on PSID panel data, as it is not possible to calculate this parameter in cross-sectional Census data. We use data for the first 8 years which are available in the PSID, as we want to be as close to 1960

All parameters are time-invariant, and the only exogenous change over time is labor productivity growth. The following parameters need to be calibrated: the parameters of the utility function $\theta_l, \theta_m, \theta_h, \varepsilon$, the distribution of labor efficiencies, $f(a_l, a_m, a_h)$, and the initial sectoral labor productivities, $A_l(0), A_m(0), A_h(0)$.

We proceed in two steps. First, we calibrate the underlying distribution of labor efficiencies under the assumption that the 1960 employment shares are met. Second, given the distribution of labor efficiencies we calibrate the utility function and initial productivity parameters to match the 1960 employment shares.

The main idea behind the first step of the calibration is the following. For any given distribution, $f(a_l, a_m, a_h)$, there is a unique pair of relative unit wages, $\{\omega_l/\omega_m, \omega_h/\omega_m\}$, which results in the employment shares observed in the data in 1960. These relative unit wages in turn imply all the wages in the economy, including sectoral relative average wages and overall wage dispersion. We calibrate the parameters of the distribution to guarantee that if the observed employment shares are matched, then so are the relative average wages in 1960 and the dispersion of log wages.

For the baseline calibration we assume that labor efficiencies are drawn from a trivariate lognormal distribution. This assumption is not innocuous, but there are several reasons that support this choice. First, the lognormal distribution is particularly important in the selection literature (e.g. Roy (1951), Heckman and Sedlacek (1985), Borjas (1987), Mulligan and Rubinstein (2008)). Second, empirical wage distributions are skewed to the right and typically resemble a lognormal distribution. Assuming that the underlying distribution of abilities is trivariate lognormal does not guarantee that after selection the resulting one dimensional wage distribution is also going to be lognormal. Nonetheless, under our parametrization the model generates a wage distribution that resembles those we see in reality, in particular it has a long right tail. Third, the trivariate lognormal distribution allows for a very flexible correlation structure, the quantitative effect of which we explore.

Without loss of generality we normalize the mean of a_l, a_m and a_h to be unity.[28] Given these assumptions, the six parameters of the variance-covariance matrix are left

---

as possible, and to have a large enough sample size. See Appendix C.1 for the details of this estimation.

[28]For the trivariate lognormal distribution numerical simulations show that the path of employment shares and of relative average wages is independent from the level of these means.

to be calibrated. Let σ_j denote the standard deviation of efficiency in sector J, and ρ_{jk} the correlation between sector J and sector K efficiency. In our baseline calibration we set all pairwise correlations between workers' sectoral efficiencies to 0.3. For this set of correlations – over which we conduct extensive robustness checks in section 4.3 – we calibrate the three variance parameters of the distribution to match the relative average wages in 1960 and the dispersion of log wages.[29]

The parameters of the utility function and the initial labor productivities are left to be calibrated. Previous literature has found a very low elasticity of substitution between goods and services when output is measured in consumption value added terms. Ngai and Pissarides (2008) find that plausible estimates are in the range $(0, 0.3)$, while Herrendorf et al. (2013) find a value of $\varepsilon = 0.002$. While neither these papers, nor others in the literature, have estimated the elasticity for our sector classification, we use $\varepsilon = 0.002$ in our baseline calibration, and conduct sensitivity analysis with respect to this parameter. Of the remaining six parameters, $\theta_l, \theta_m, \theta_h$ and $A_l(0), A_m(0), A_h(0)$ only two ratios matter for the equilibrium of this economy, as can be seen from equations (16) and (17). We calibrate $\tau_l \equiv \left(\frac{A_m(0)}{A_l(0)} \right)^{1-\varepsilon} \left(\frac{\theta_m}{\theta_l} \right)^{-\varepsilon}$ and $\tau_h \equiv \left(\frac{A_m(0)}{A_h(0)} \right)^{1-\varepsilon} \left(\frac{\theta_m}{\theta_h} \right)^{-\varepsilon}$ to match the unit wages in 1960 found in the first step of the calibration. The calibrated parameters are summarized in Table 3.

It is well known that if individuals self-select based on their endowments of efficiency units and one cannot observe these efficiency units, then the measurement of changes in average wages or in labor productivity will be biased.[30] In our model, expanding sectors soak up, while contracting sectors shed relatively less efficient workers. This implies that the average efficiency in manufacturing increases, while the average efficiency in low- and high-skilled services falls over time, which – if left uncorrected – leads to an overestimation of productivity growth in manufacturing relative to both types of services.[31] To understand the potential magnitude of this bias, we cor-

---

[29]We use sectoral log wage dispersion only to calibrate the variance-covariance matrix of the underlying distribution of abilities. As our interest is in the evolution of employment shares and relative average wages, we leave the analysis of the implications of structural change for overall and sectoral wage dispersion for future work.

[30]Carneiro and Lee (2011) estimated the bias in the measurement of the skill premium, while Young (2014) pointed this out in the context of measuring productivity growth differentials across sectors.

[31]This is true only when manufacturing employment is shrinking, and low- and high-skilled service employment is expanding.

Table 3: Calibrated Parameters

	Description	Value
ρ_{lm}	correlation between L and M sector efficiency	0.3
ρ_{lh}	correlation between L and H sector efficiency	0.3
ρ_{mh}	correlation between M and H sector efficiency	0.3
σ_l^2	variance of L sector efficiency	0.12
σ_m^2	variance of M sector efficiency	0.30
σ_h^2	variance of H sector efficiency	0.35
ε	CES b/w L, M and H in consumption	0.002
τ_l	relative weight on L	0.54
τ_h	relative weight on H	0.84

The correlation values are set at 0.3, and the elasticity of substitution, ε, is taken from the literature. The remaining parameters are calibrated to match the sectoral employment shares, the relative average wages, and the overall dispersion of wages.

rect for this selection effect, and report both measured and model implied productivity growth rates.

Similarly to Ngai and Petrongolo (2014) we measure labor productivity growth by dividing the growth rate of industry level quantity indices from the Bureau of Economic Analysis (BEA) with the growth rate of industry level hours worked data from the Census/ACS data, and aggregating it up to our three sectors. Our model output can be written as $Y_j = A_j N_j = B_j L_j$ for each sector J, where L_j is the raw employment share (hours worked) and B_j the measured labor productivity, while A_j is the true productivity in the model and N_j the efficiency units of labor. To account for the difference between N_j and L_j – the selection effect – we compute the path of the A_js in the following way. Given the calibration of all of the fixed parameters of the model, we find the path of A_js such that the model implied growth in B_js is equal to the labor productivity growth measured in the data.

Table 4 shows in the first three columns the model implied, and in the last three columns the measured average annual labor productivity growth in low-skilled services, manufacturing and high-skilled services for each decade between 1960 and 2007, as well as for the entire period. According to our calculations the growth of labor productivity in manufacturing was higher than in either low- or high-skilled services in each of the decades considered. Under our calibration, the bias in the relative produc-

Table 4: Annual average labor productivity growth

| | Adjusted for selection | | | Based on raw labor | | |
	A_l	A_m	A_h	B_l	B_m	B_h
1960-1970	1.0077	1.0248	1.0132	1.0068	1.0279	1.0124
1970-1980	1.0095	1.0140	1.0116	1.0093	1.0148	1.0116
1980-1990	1.0051	1.0235	1.0108	1.0042	1.0270	1.0101
1990-2000	1.0265	1.0271	1.0120	1.0270	1.0286	1.0083
2000-2007	1.0072	1.0282	1.0162	1.0062	1.0320	1.0163
1960-2007	**1.0113**	**1.0231**	**1.0125**	**1.0109**	**1.0258**	**1.0115**

Labor productivity growth rates in the last three columns are calculated by dividing industry level quantity indices from the BEA with industry level employment growth data taken from the Census/ACS, and aggregating up to our three sectors. The first three columns give the path of model productivity growth, which for our baseline calibration result in the same measured labor productivity growth as in the data.

tivity differentials is small, but not negligible.[32] It is worth to note that both productivity growth and the relative productivity growth across sectors varied significantly decade by decade. For this reason, we evaluate the quantitative performance of the model in two ways: (i) by plugging in the average annual growth rates for the entire period; and (ii) by plugging in the decennial growth rates.

4.2 Wage and employment dynamics

To understand the strength of the mechanisms that we highlight, we simulate the competitive equilibrium of the economy at different productivity levels. We fix the preference and the sectoral efficiency distribution parameters at their calibrated values, and feed in the labor productivity growth shown in Table 4. Our ultimate interest is the endogenous path of employment shares and relative average wages.

Figure 5 plots the dynamics for our baseline calibration using average annual growth rates over the period 1960-2007 adjusted for selection (bottom left numbers in Table 4). The top left panel shows (in logs) the path of both the measured productivity (in the data and the model) in dotted lines, and the model productivity in solid lines for all three sectors. Since productivity growth is highest in the manufacturing sector, but

---

[32]Young (2014) finds that the implied bias might potentially be so large as to overturn the conventional wisdom of faster productivity growth in manufacturing. However, with our calibration this is not the case.
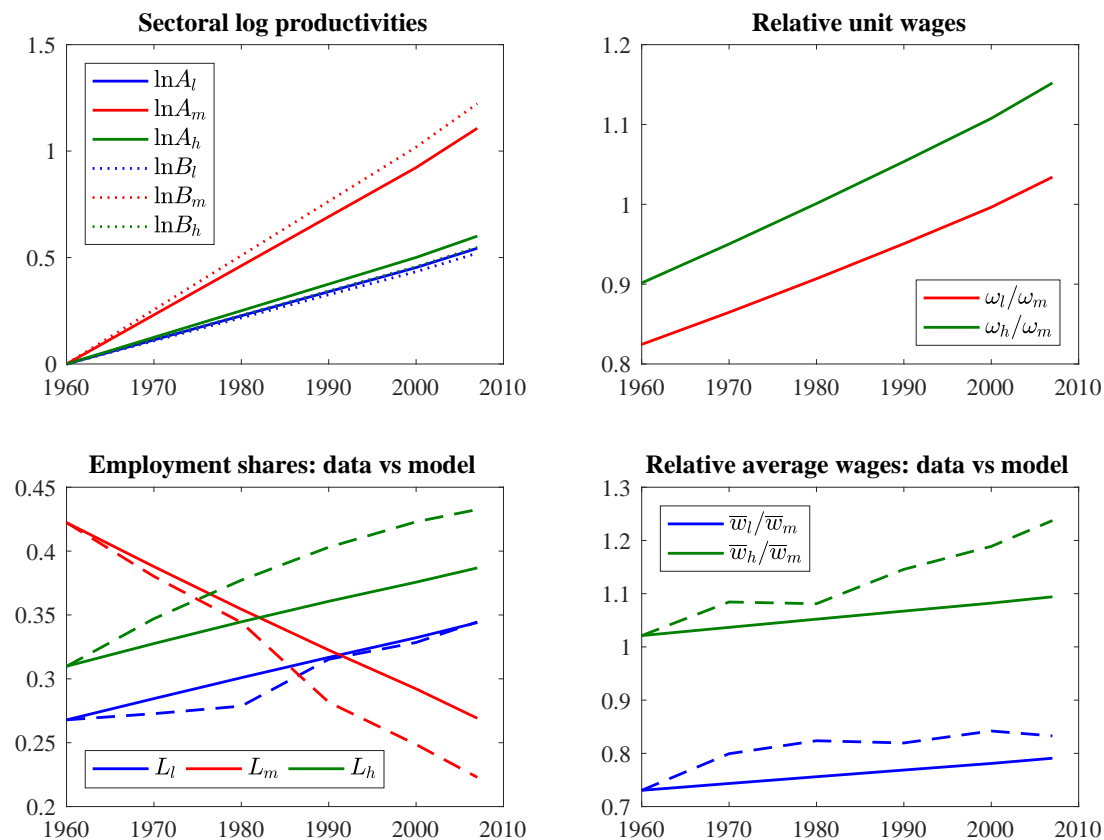
Figure 5: Transition of the benchmark model

The top left panel shows the observed average change in log labor productivity (dotted line) and model productivity (solid line). The top right panel shows the endogenous response of the relative average unit wages. The bottom left panel shows the predicted employment shares (solid line) and their actual path (dashed line), while the bottom right panel shows the model predicted (solid line) and actual path (dashed line) of the relative average sectoral wages.

manufactured goods and both types of services are complements in consumption, the increased demand for the output of all sectors in equilibrium is met through a reallocation of labor towards low- and high-skilled services, as we showed in Proposition 1.[33] The increased demand for labor in low- and high-skilled services puts an upward pressure on the unit wages in these sectors relative to the unit wage in manufacturing, which we plot in the top right panel. The bottom two panels show our model's predictions (solid lines) contrasted with the data (dashed lines) for our measures of interest. Not surprisingly, the model matches the 1960 employment shares (bottom left panel) and the 1960 relative average wages (bottom right panel) very well, as we targeted

---

[33]In Proposition 1 we assumed that productivity growth in the two types of services was equal. In the quantitative evaluation of the model we relax this assumption and take productivity growth from the data. Nonetheless, the results derived in the proposition hold in all our simulations.

these measures. But the model also does well in predicting the paths of employment shares and relative average wages after 1960. Our baseline model predicts at least 60 per cent of the change in the employment share of each sector.[34]

In our model the relative average wage changes are driven by changes in the relative unit wages and changes in the relative average sectoral labor efficiencies. As discussed in section 3.3, these two effects, in general, go in opposite directions, however the direct effect of the unit wages typically dominates the indirect effect that it has on average sectoral efficiencies. This is the case in our baseline calibration as well, and our model overall predicts about 33 per cent of the growth in the relative low-skilled service sector wages, and 59 percent of the growth in the relative average high-skilled service sector wages compared to manufacturing.[35]

The path of employment shares and relative average wages generated by the model are very smooth compared to the data. This is not surprising, as we assumed a constant annual growth rate of sectoral labor productivity between 1960 and 2007. However, Table 4 reveals that the growth rates have varied substantially over time. Figure 6 shows the simulated model contrasted with the data when feeding in the model productivity growth rates calculated for each period. The main difference in terms of productivity growth rates is that the growth rate in low-skilled services is very low in the initial decades, and it is very high between 1990 and 2000 when it is almost the same as in manufacturing. Another important thing to note is that the growth rate in high-skilled services is also quite high between 2000 and 2007. These changing productivity differences imply that initially high-skilled service employment expands more slowly, while low-skilled services expand more rapidly, and this pattern is reversed in 1990. While the model quantitatively does worse in the initial decades, the overall predicted changes are the same.

As discussed in section 3.3, while our model qualitatively matches the *changes* in expenditure shares, a model without capital intensity differences and with perfect cap-

---

[34]In the data the percentage point change in the employment share of high-skilled services / manufacturing / low-skilled services are 12 / -20 / 8, in our model these numbers are 8 / -15 / 8 respectively.

[35]In the data the relative average wage of L and H compared to M workers increased by 14 and respectively 21 per cent. In the simulation this is 8 per cent for L to M (25 per cent improvement in relative unit wages and 13 per cent decline in relative average efficiency), and 7 per cent for H to M (28 per cent increase in relative unit wages and a more than 15 per cent drop in relative average efficiency).
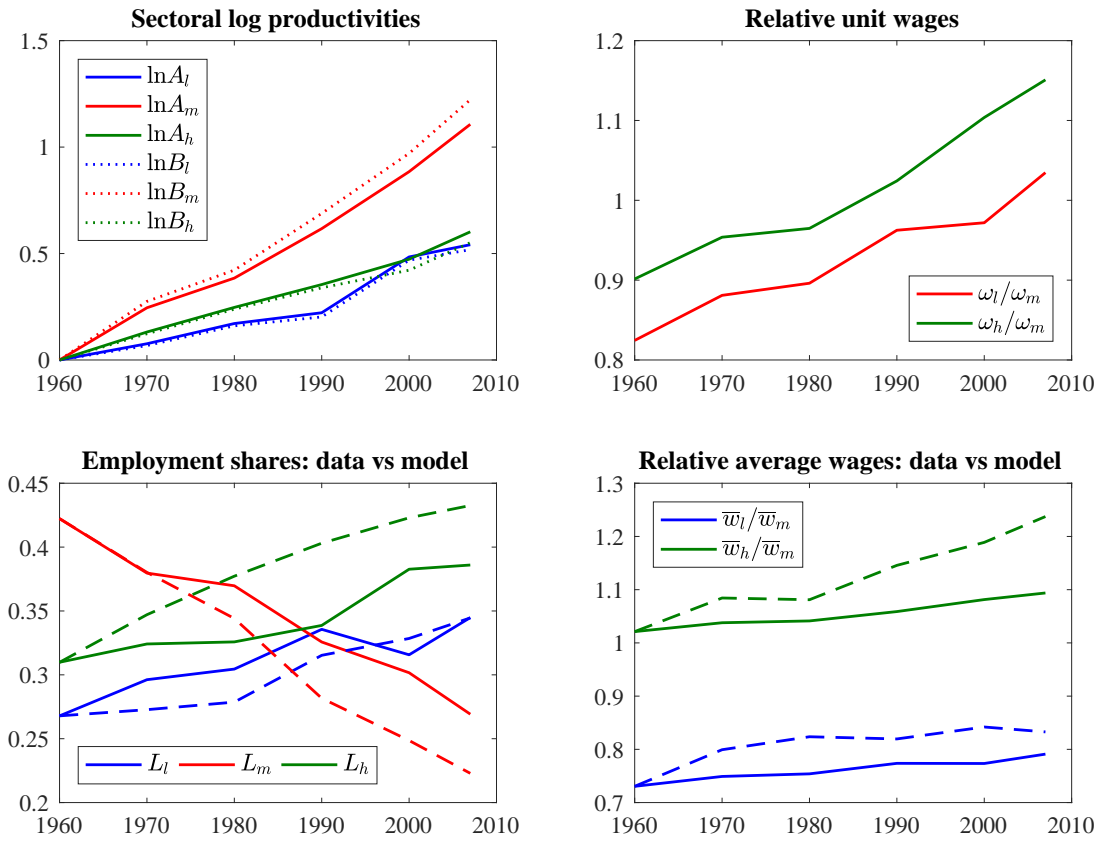
Figure 6: Transition of the model with decennial growth

The panels show the path of the same variables as in Figure 5, except that we feed in the change in labor productivity decade-by-decade, rather than the average for the entire period.

ital mobility across sectors cannot match the *level* of sectoral relative average wages, employment and expenditure shares jointly. Figure 7 shows the relative value added in low- and high-skilled services compared to the value added in manufacturing in the model (solid line) and in the data (dashed line) for the time-invariant productivity growth rates. The overall increase in high-skilled services relative to manufacturing is replicated quite well, while for low-skilled services the model over predicts the increase.[36]

---

[36]For H to M these numbers are 125% in the data, and 112% in the model, and for L to M 38% in the data and 120% in the model.
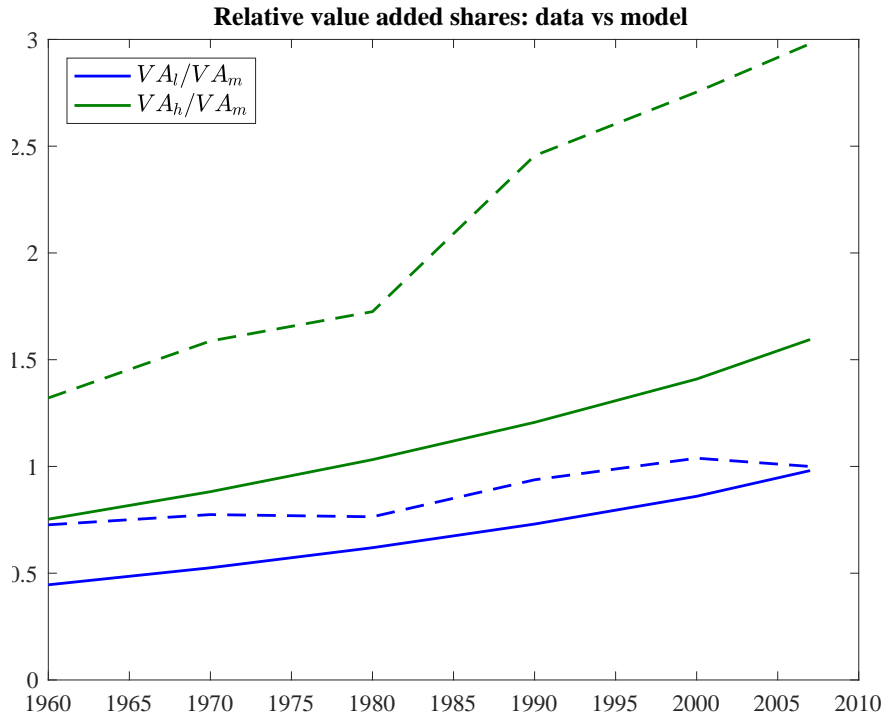
Figure 7: Transition of relative low- and high-skilled service value added

The graph shows the value added in low- and high-skilled services relative to manufacturing as predicted by the model (solid line) and in the data (dashed line).

4.3 Robustness checks

In our baseline calibration we assume that the correlation between any two of the underlying sectoral efficiency draws is 0.3, that the underlying distribution of abilities is trivariate lognormal, and that the elasticity of substitution in consumption is 0.002. Before we analyze the sensitivity of our results to the these assumptions, we first show that given our calibration strategy the model implied path of employment shares is independent of the assumed correlations and the assumed underlying distribution of abilities.

We calibrate the diagonal elements of the variance-covariance matrix of the distribution, and a combination of the utility function parameters and of the initial model productivities to match – among other targets – the employment shares in 1960. Thus, by construction, in the initial period the model implied employment shares are the same across *all* calibrations. Moreover, the model-implied path of employment shares is independent of the underlying ability distribution. This is due to the way we correct for selection effects in the measurement of sectoral productivity growth. In particular,

the A_js are such that – conditional on all fixed parameters – the model implied raw labor productivity growth (B_j) is equal to that measured in the data. This implies that the model implied growth in the B_js is also the same across *all* calibrations. This implies that as long as ε is the same across calibrations, the model implied path of employment shares is also the same, as we show below.

The model implied labor productivity growth can directly be matched to the model implied employment shares. To see how employment is allocated across sectors, consider the social planner's problem:

$$\max_{N_{lt}, N_{mt}, N_{ht}} u\left(\left[\theta_l C_{lt}^{\frac{\varepsilon-1}{\varepsilon}} + \theta_m C_{mt}^{\frac{\varepsilon-1}{\varepsilon}} + \theta_h C_{ht}^{\frac{\varepsilon-1}{\varepsilon}}\right]^{\frac{\varepsilon}{\varepsilon-1}}\right),$$

$$\text{s.t.} \quad C_{jt} = A_{jt}N_{jt} = B_{jt}L_{jt} \quad \text{for any } j = L, M, H.$$

The social planner's problem can be formulated either as choosing the efficiency units of labor assigned to each sector given the A_js or as choosing the employment shares given the B_js.[37] The first order conditions in terms of employment shares can be written as:

$$\frac{L_{jt}}{L_{kt}} = \left(\frac{\theta_k}{\theta_j}\right)^{-\varepsilon} \left(\frac{B_{kt}}{B_{jt}}\right)^{1-\varepsilon} = \left(\frac{\theta_k}{\theta_j}\right)^{-\varepsilon} \left(\frac{B_{k0}}{B_{j0}}\right)^{1-\varepsilon} \left(\frac{\frac{B_{kt}}{B_{k0}}}{\frac{B_{jt}}{B_{j0}}}\right)^{1-\varepsilon}$$

$$= \frac{L_{j0}}{L_{k0}} \left(\frac{\frac{B_{kt}}{B_{k0}}}{\frac{B_{jt}}{B_{j0}}}\right)^{1-\varepsilon} \quad \text{for any } t \text{ and } j, k \in \{L, M, H\}.$$

The last equality makes use of the fact that for all distributions τ_l and τ_h are calibrated such that the initial L_l, L_m and L_h are matched. This equation shows that given the growth rate of the measured labor productivity, as long as ε is the same across calibrations, the path of employment shares do not depend on the underlying distribution of sectoral efficiencies. Therefore we only report the path of relative average wages when conducting the sensitivity analysis with respect to the sectoral efficiency distribution.

Figure 8 summarizes the sensitivity of relative average wage paths to the assumed correlation structure, along with Table 14 in Appendix C.2. While average wages relative to manufacturing increase in both the high- and the low-skilled service sector

---

[37]Note that the path of A_js depends on the underlying distribution of abilities.
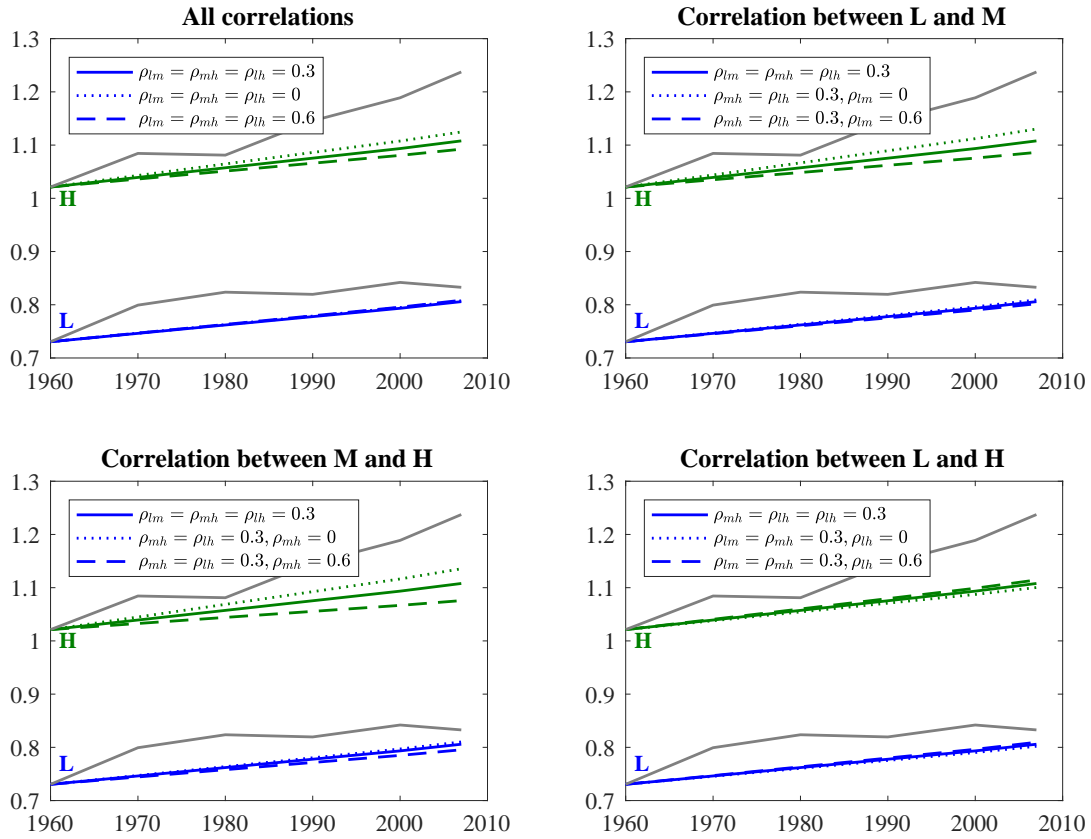
Figure 8: Transition under different correlation structures

The graph shows the path of relative average wages for different correlation structures. The gray solid line shows the data, the green lines show the average wage in high-skilled services relative to manufacturing, and the blue lines show the relative average wage of low-skilled services. The solid colored line shows our baseline calibration, whereas the dotted line represent the path for a lower correlation (0), and the dashed line for a higher correlation (0.6).

for all calibrations, the magnitude of these changes varies quite a bit. In Figure 8 the top left panel shows the wages for our baseline calibration (solid lines), which features a correlation of 0.3 between the underlying ability draws of any two sectors, along with one where this correlation is 0 (dotted lines) and 0.6 (dashed lines). This graph shows that while relative average high-skilled service wages increase a bit less when the correlation is higher, the path of relative wages in low-skilled services is hardly affected. Looking at the corresponding lines (top, middle bold and last) in Table 14, we can see that the higher correlation we assume, the calibration requires a less dispersed underlying efficiency distribution. With a less dispersed and more correlated distribution, the model predicts a smaller adjustment in the relative average sectoral wages, as the model gets closer to the case of homogeneous labor (i.e. similar to Ngai

and Pissarides (2007)). In the limiting case, where each individual is endowed with the same amount of efficiency units of labor in all three sectors, there is no selection. In order to have positive labor supply to all three sectors, unit wages have to be equalized. As unit wages are equalized, individuals are indifferent between working in any of the three sectors. Labor supply across sectors is random, average sectoral wages are equalized, implying relative average wages are independent of the level of technology in the economy. Thus it is not surprising that our model predicts smaller changes in relative average wages when we assume higher correlations across all three pairs of sectoral abilities.

In the other three panels of Figure 8 we show the relative average wage paths when only modifying one of the baseline correlations to 0 (dotted) and 0.6 (dashed): in the top right panel it is the correlation between sector L and M ability, in the bottom left it is between M and H and in the bottom right it is between L and H. These graphs show that in general the relative high-skilled service wages are more sensitive to the correlation structure, but the differences in the relative average wage paths are not large.

In our baseline calibration we assume that the underlying distribution of abilities is trivariate lognormal. At the beginning of this section we showed that the path of employment shares is independent of the assumed distribution of underlying abilities. However, the quantitative – and potentially even the qualitative – predictions regarding relative average wages are likely to be affected by not only the correlation structure, but also the functional form of the distribution. We recalibrate the model assuming a normal distribution, truncated at zero for all three abilities, for the same set of correlations as in Figure 8. As Table 15 in Appendix C.2 shows, our model predicts much larger changes in relative average wages when assuming that the underlying efficiencies are drawn from a truncated normal distribution. The model overpredicts the change in the relative low-skilled service wage by 35 to 79 per cent, while for relative high-skilled services wages it predicts between 88 and 112 per cent of the change.

In our baseline calibration we use $\varepsilon = 0.002$ for the elasticity of substitution between goods and services (measured in value-added terms) as estimated by Herrendorf et al. (2013), which is at the lower end of estimates reported by Ngai and Pis-

37
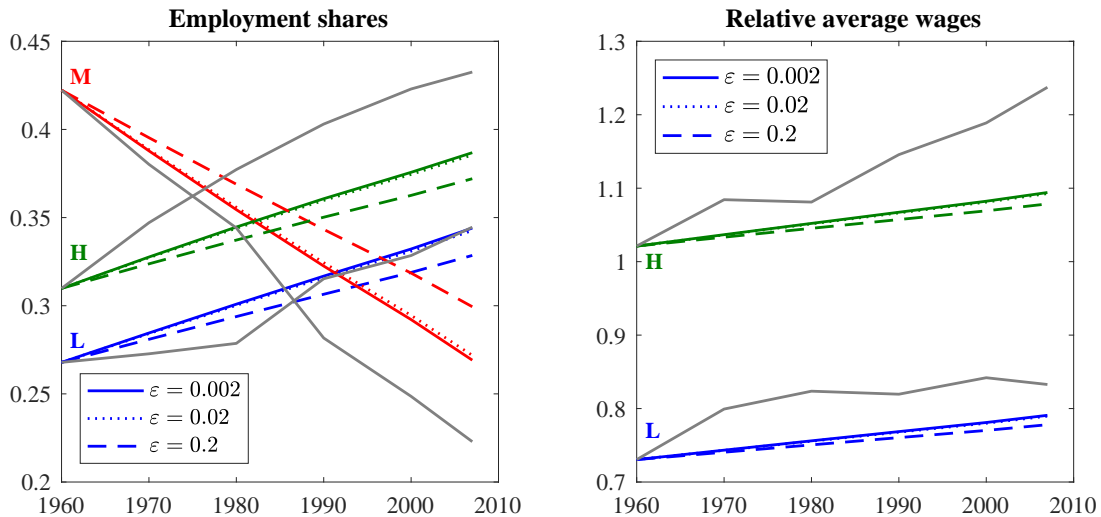
Figure 9: Transition under different elasticities of substitution

The graph shows the path of employment shares (left panel) and relative average wages (right panel) for different elasticities of substitution, against the data. The gray solid line shows the data. The solid colored lines show our baseline calibration ($\varepsilon = 0.002$), whereas the dotted lines represent the path for $\varepsilon = 0.02$, and the dashed lines for $\varepsilon = 0.2$.

sarides (2008). To see whether our results are robust to higher, yet plausible, values of this parameter, we explore how our results change when using $\varepsilon = 0.02$ or $\varepsilon = 0.2$, naturally recalibrating the other parameters to match our five targets. Figure 9 shows that qualitatively the transition paths look exactly the same. A higher elasticity of substitution implies that the effective employment in low- and high-skilled services have to increase less, and the effective employment in manufacturing has to fall less in order to meet equilibrium demands. This in turn implies less adjustment in employment shares and in relative average wages. Increasing the value of the elasticity of substitution takes the model's predictions further away from the time paths observed in the data. As can be seen in Figure 9 the transition paths look virtually the same for $\varepsilon = 0.002$ and $\varepsilon = 0.02$, while for $\varepsilon = 0.2$ the model predicts less adjustment, but it does reasonably well. This latter version of the model predicts 46 per cent of the increase in L and 27 per cent of the increase in H sector average wages relative to M. In terms of employment share changes, the model predicts at least half of the observed changes between 1960 and 2007.

4.4 Implications for occupational employment polarization

Even though our model is about sectoral labor market allocations, we can use it to calculate the fraction of occupational employment share changes that it predicts. Fixing the within-sector occupational employment shares at the 1960-2007 period's average, we can compute how much the sectoral labor reallocation – as predicted by our model – would explain of the (between-sector) occupational employment share changes identified in the shift-share decomposition of section 2.3. Our model predicts 82 per cent of the manual, 72 per cent of the routine, and 63 per cent of the abstract occupation's between-industry employment share changes. This is equivalent to 44 per cent of the total manual, 24 per cent of the total routine, and 15 per cent of the total abstract employment share change. However, given that our model abstracts from occupational differences within sectors, and thus by construction cannot explain any of the within-industry changes in occupational employment, we think that the model explains a quantitatively important fraction of total occupational employment share changes.

5 Conclusions

The literature on polarization of employment and wages has typically focused on occupations. We present a set of new empirical facts that suggest that in addition to reallocations between occupations within industries, also shifts between industries contribute to the polarization of labor markets. Moreover, we show that in terms of broadly defined industries, polarization was present as early as 1950-1960 and directly linked to the decline of manufacturing employment. Based on this evidence we propose a novel explanation, one based on structural change. A methodological contribution of our paper is that we develop a multi-sector model with heterogeneous labor in Roy-style fashion, the most parsimonious setup that yet allows heterogeneity in wages. An insight from our model is that unbalanced technological progress does not only lead to structural change, the reallocation of employment across sectors, but also affects sectoral average wages. We find that higher productivity growth in manufacturing than in low- and high-skilled services increases employment and wages in both

the low-skilled and the high-skilled service sector, thus leading to the polarization of the labor market. This simple model does remarkably well in predicting the sectoral wage and employment patterns of the last 50 years.

References

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

Daron Acemoglu and Veronica Guerrieri. Capital Deepening and Nonbalanced Economic Growth. *Journal of Political Economy*, 116(3):pp. 467–498, 2008.

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

David H. Autor, Frank Levy, and Richard J. Murnane. The Skill Content of Recent Technological Change: An Empirical Exploration. *The Quarterly Journal of Economics*, 118(4):pp. 1279–1333, 2003.

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

Alan S. Blinder. How Many US Jobs Might Be Offshorable? *World Economics*, 10(2): 41–78, 2009.

Alan S. Blinder and Alan B. Krueger. Alternative Measures of Offshorability: A Survey Approach. *Journal of Labor Economics*, 31(S1):S97 – S128, 2013.

Timo Boppart. Structural Change and the Kaldor Facts in a Growth Model with Relative Price Effects and Non-Gorman Preferences. *Econometrica*, forthcoming, 2014.

George J. Borjas. Self-selection and the earnings of immigrants. *The American Economic Review*, 77(4):531–553, 1987.

Francisco J. Buera and Joseph P. Kaboski. The Rise of the Service Economy. *American Economic Review*, 102(6):2540–69, September 2012.

Francisco J. Buera, Joseph P. Kaboski, and Richard Rogerson. Skill biased structural change. Working Paper 21165, National Bureau of Economic Research, May 2015.

Pedro Carneiro and Sokbae Lee. Trends in Quality-Adjusted Skill Premia in the United States, 1960-2000. *American Economic Review*, 101(6):2309–49, 2011.

Francesco Caselli and Wilbur John Coleman, II. The U.S. Structural Transformation and Regional Convergence: A Reinterpretation. *Journal of Political Economy*, 109(3): pp. 584–616, 2001.

David Dorn. *Essays on Inequality, Spatial Interaction, and the Demand for Skills*. PhD thesis, University of St. Gallen, September 2009. Dissertation no. 3613.

Christian Dustmann, Johannes Ludsteck, and Uta Schönberg. Revisiting the German Wage Structure. *The Quarterly Journal of Economics*, 124(2):843–881, 2009.

Andy Feng and Georg Graetz. Rise of the Machines: The Effects of Labor-Saving Innovations on Jobs and Wages. Working paper, 2014.

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

Maarten Goos, Alan Manning, and Anna Salomons. Job Polarization in Europe. *The American Economic Review Papers and Proceedings*, 99(2):pp. 58–63, May 2009.

Maarten Goos, Alan Manning, and Anna Salomons. Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, 104(8):2509–26, 2014.

Gene M. Grossman and Esteban Rossi-Hansberg. Trading Tasks: A Simple Theory of Offshoring. *American Economic Review*, 98(5):1978–97, 2008.

James Heckman and Guilherme Sedlacek. Heterogeneity, aggregation, and market wage functions: An empirical model of self-selection in the labor market. *Journal of Political Economy*, 93(6):1077–1125, 1985.

Berthold Herrendorf, Richard Rogerson, and Ákos Valentinyi. Two Perspectives on Preferences and Structural Transformation. *American Economic Review*, 103(7):2752–89, 2013.

Berthold Herrendorf, Richard Rogerson, and Ákos Valentinyi. Growth and Structural Transformation. In *Handbook of Economic Growth*, volume 2, chapter 6, pages 855–941. Elsevier, December 2014.

Piyabha Kongsamut, Sergio Rebelo, and Danyang Xie. Beyond Balanced Growth. *The Review of Economic Studies*, 68(4):pp. 869–882, 2001.

Simon Kuznets. Quantitative Aspects of the Economic Growth of Nations: II. Industrial Distribution of National Product and Labor Force. *Economic Development and Cultural Change*, 5(4):pp. 1–111, 1957.

David Lagakos and Michael E. Waugh. Selection, Agriculture, and Cross-Country Productivity Differences. *American Economic Review*, 103(2):948–80, 2013.

Angus Maddison. Economic Growth and Structural Change in the Advanced Countries. In I. Leveson and J. W. Wheeler, editors, *Western Economies in Transition: Structural Change and Adjustment Policies in Industrial Countries*, pages 41–65. London: Croom Helm, 1980.

Alan Manning. We Can Work It Out: The Impact of Technological Change on the Demand for Low-Skill Workers. *Scottish Journal of Political Economy*, 51(5):581–608, November 2004.

Kiminori Matsuyama. Increasing Returns, Industrialization, and Indeterminacy of Equilibrium. *The Quarterly Journal of Economics*, 106(2):pp. 617–650, 1991.

Francesca Mazzolari and Giuseppe Ragusa. Spillovers from High-skill Consumption to Low-skill Labor Markets. *The Review of Economics and Statistics*, 95(1):74 – 86, March 2013.

Peter B. Meyer and Anastasiya M. Osborne. Proposed Category System for 1960-2000 Census Occupations. Working Papers 383, U.S. Bureau of Labor Statistics, September 2005.

Guy Michaels, Ashwini Natraj, and John Van Reenen. Has ICT Polarized Skill Demand? Evidence from Eleven Countries over 25 years. *The Review of Economics and Statistics*, 96(1):60 – 77, March 2014.

Casey B. Mulligan and Yona Rubinstein. Selection, investment, and women's relative wages over time. *The Quarterly Journal of Economics*, 123(3):1061–1110, 2008.

L. Rachel Ngai and Barbara Petrongolo. Gender Gaps and the Rise of the Service Economy. Discussion Paper 9970, Centre for Economic Policy Research, May 2014.

L. Rachel Ngai and Christopher A. Pissarides. Structural Change in a Multisector Model of Growth. *The American Economic Review*, 97(1):pp. 429–443, 2007.

L. Rachel Ngai and Christopher A. Pissarides. Trends in Hours and Economic Growth. *Review of Economic Dynamics*, 11(2):239 – 256, 2008.

Andrew D. Roy. Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146, 1951.

Steven Ruggles, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. Integrated Public Use Microdata Series. Technical report, Minnesota Population Center, Minneapolis, MN, 2010. Version 5.0.

Alexandra Spitz-Oener. Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. *Journal of Labor Economics*, 24(2):pp. 235–270, 2006.

Alwyn Young. Structural Transformation, the Mismeasurement of Productivity Growth, and the Cost Disease of Services. *American Economic Review*, 104(11):3635–67, 2014.

Appendix

A Data appendix

We use data from the US Census of 1950, 1960, 1970, 1980, 1990, 2000 and the American Community Survey (ACS) of 2007, which we access from IPUMS-USA, provided by Ruggles, Alexander, Genadek, Goeken, Schroeder, and Sobek (2010). Following Acemoglu and Autor (2011) and Autor and Dorn (2013) we restrict the sample to individuals who were in the labor force and of age 16 to 64 in the year preceding the survey. We drop residents of institutional group quarters and unpaid family workers. We also drop respondents with missing earnings or hours worked data and those who work in agricultural occupations/industries or in the military. Our employment measure is the product of weeks worked times usual number of hours per week.[38] We compute hourly wages as earnings divided by the product of usual hours and weeks worked.

To construct the 30-year change graphs of Figure 1 and 10, and the 10-year change graphs of Figure 11 we follow the methodology used in Autor et al. (2006)), Acemoglu and Autor (2011), and Autor and Dorn (2013), which requires a balanced panel of occupations. Dorn (2009) and Autor and Dorn (2013) provide a balanced panel of occupational classifications ('*occ1990dd*') over 1980-2008, which we use to construct a balanced panel over 1950-2007 by aggregating occupational codes as needed. This leaves us with 183 balanced occupational codes. Figures 1, 10, and 11 plot the smoothed changes in average log hourly wages and total hours worked at each percentile of the occupational skill distribution. These skill percentiles are constructed by ranking the balanced occupations according to their 1950 (Figure 10 and top row of Figure 11) or 1980 mean hourly wages (Figure 1 and bottom row of Figure 11), and then splitting them into 100 groups, each making up 1 percentile of 1950 or 1980 employment.

Figure 10 shows the change in log real hourly wages and employment, similarly as Figure 1, with the difference that the ranking of occupations is based on their 1950 log

---

[38]Since in 1950 the Census did not include usual hours worked, we use hours worked last week instead. In 1960 and 1970 the Census asked only for an interval of hours and weeks worked last year; we use the midpoint of the interval given.
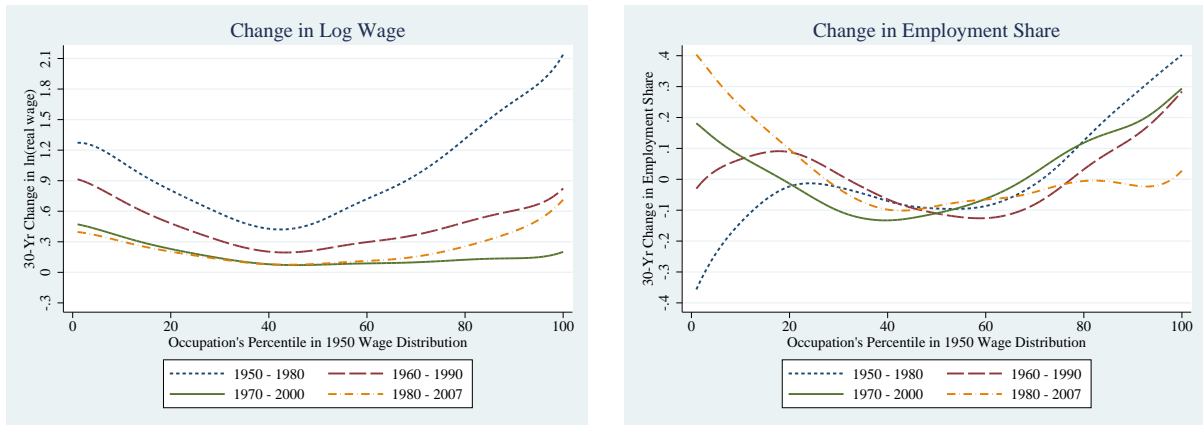
Figure 10: Smoothed wage and employment polarization 1950 ranking

Notes: Data and left and right panel same as in Figure 1, except occupations are ranked based on their 1950 mean wages.

real hourly wage. The graph reinforces the message of Figure 1. The left panel shows that wages have been polarizing from 1950 onwards, with the polarization most pronounced in the earlier periods. The right panel shows that the polarization of employment is present in all 30-year periods starting from 1960, with the most pronounced polarization between 1970-2000.

Figure 11 shows the wage and employment change of occupations for 10-year periods, with occupations ranked based on the 1950 wages (top row) and the 1980 wages (bottom row). These graphs show that polarization does not happen on a decade-by-decade basis. In some decades the top gains, while in others the bottom, but it is never the middle-wage occupations that gain the most in terms of wages or employment.

In the text we document polarization in terms of occupations for 183 and 10 occupation categories (in Figure 1 and Figure 2 respectively), here we show it for an even coarser classification. As in Acemoglu and Autor (2011) we classify occupation groups into the following categories: manual, routine, and abstract.[39] Figure 12 shows the patterns of polarization both in terms of wages and employment shares between 1950 and 2007 for these three broad categories. The right panel shows that the employment share of routine occupations has been falling, of abstract occupations has been increasing since the 1950s, while of manual occupations, following a slight com-

---

[39]Acemoglu and Autor (2011) have a similar graph of the path of employment shares of four occupation categories (abstract, routine cognitive, routine non-cognitive, manual) between 1960 and 2007. Here we show for 3 categories, starting from 1950, and more importantly, we also show the path of relative occupational wages.
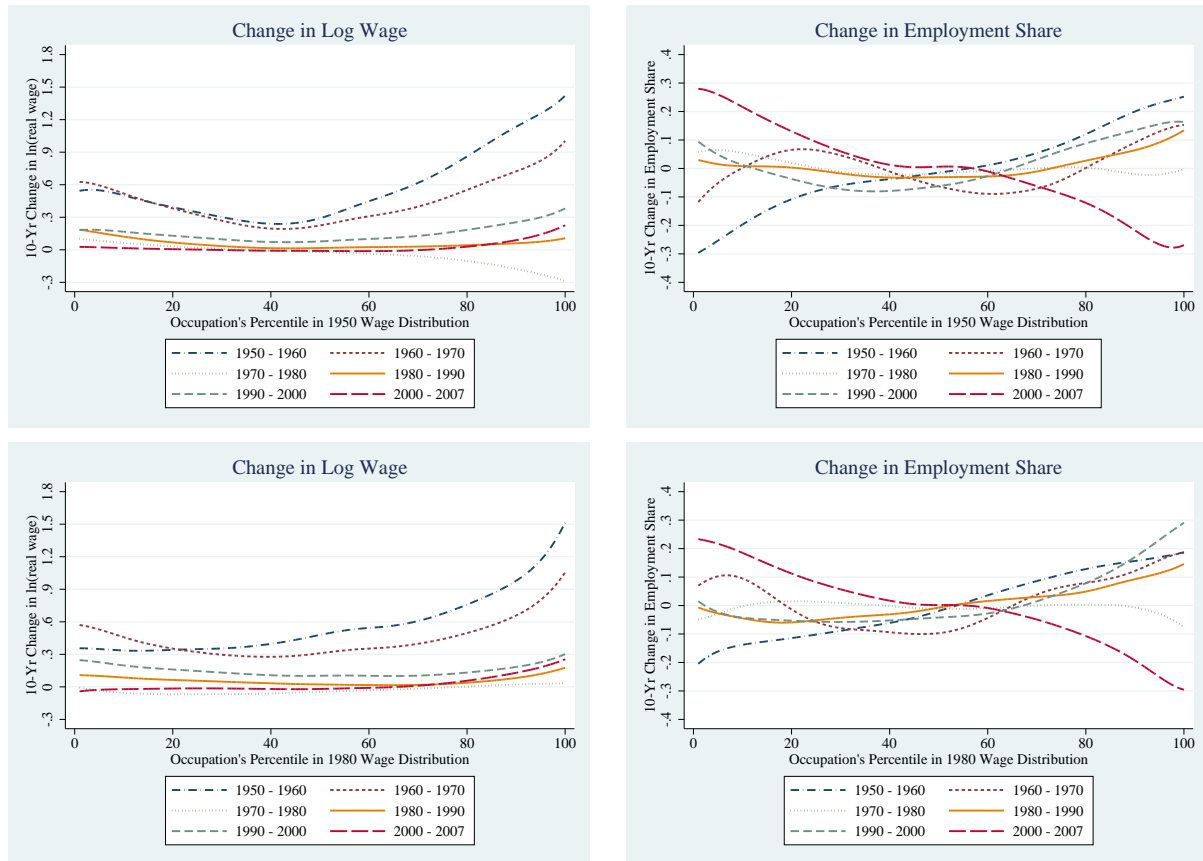
Figure 11: Smoothed wage and employment polarization, 10-year change

Notes: Data and left and right panel as in Figure 1. All panels show 10-year changes rather than 30-year changes. Occupations are ranked based on their 1950 mean wages in the top two panels, and based on their 1980 mean wages in the bottom two panels.

pression until 1960, has been steadily increasing. The left panel shows the path of the relative average manual and abstract wage compared to the routine wage. It is worth to note that, as expected, manual workers on average earn less than routine workers, while abstract workers earn more. However, over time, the advantage of routine jobs over manual jobs has been falling, and the advantage of abstract jobs over routine jobs has been rising. Thus, the middle earning group, the routine workers, lost both in terms of relative average wages and employment share to the benefit of manual and abstract workers. In other words, also in terms of these three broad occupations there is clear evidence for polarization.

A.1 Categorization of occupations

Following Acemoglu and Autor (2011) we classify occupations into three categories, which are used in Figure 12:

47
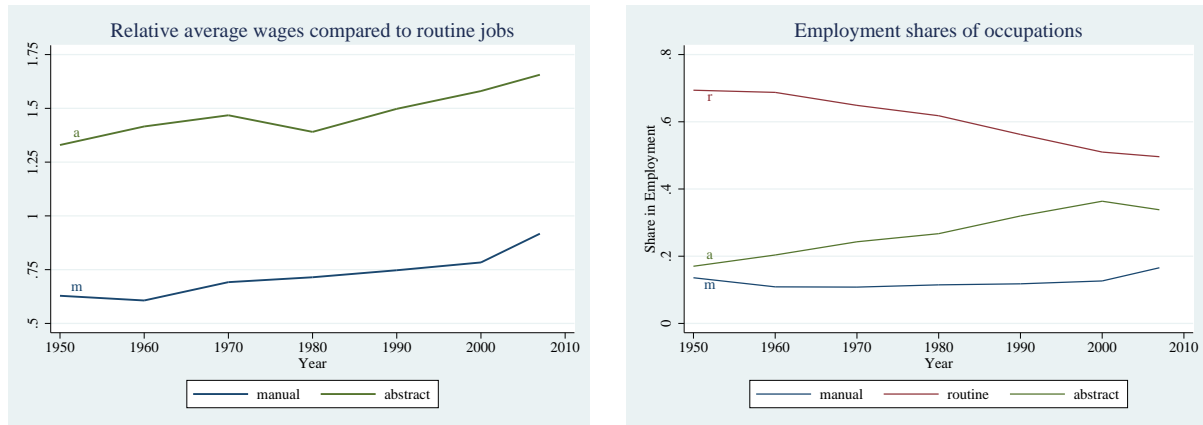
Figure 12: Polarization for broad occupations

Notes: Relative average wages and employment shares (in terms of hours) are calculated from the same data as in Figure 1. For details of the occupation classification see below.

- Manual (low-skilled non-routine): housekeeping, cleaning, protective service, food prep and service, building, grounds cleaning, maintenance, personal appearance, recreation and hospitality, child care workers, personal care, service, healthcare support;

- Routine: construction trades, extractive, machine operators, assemblers, inspectors, mechanics and repairers, precision production, transportation and material moving occupations, sales, administrative support;

- Abstract (skilled non-routine): managers, management related, professional specialty, technicians and related support.

A.2 Categorization of industries

Based on our theory we classify the industries into three sectors, which are used in Figure 3:

- Low-skilled services: personal services, entertainment, low-skilled transport (bus service and urban transit, taxicab service, trucking service, warehousing and storage, services incidental to transportation), low-skilled business and repair services (automotive rental and leasing, automobile parking and carwashes, automotive repair and related services, electrical repair shops, miscellaneous repair services), retail trade, wholesale trade;

- Manufacturing: mining, construction, manufacturing;

- High-skilled services: professional and related services, finance, insurance and real estate, communications, high-skilled business services (advertising, services to dwellings

48

and other buildings, personnel supply services, computer and data processing services, detective and protective services, business services not elsewhere classified), communications, utilities, high-skilled transport (railroads, U.S. Postal Service, water transportation, air transportation), public administration.

Table 5 summarizes the descriptive statistics for sectoral employment.

Table 5: Descriptive statistics by industry

	low-skilled services		manufacturing		high-skilled services	
	1960	2007	1960	2007	1960	2007
Highschool Dropout	55.15%	10.45%	56.29%	13.34%	30.68%	2.60%
Highschool Graduate	29.43%	36.77%	27.19%	39.98%	30.82%	19.19%
Some College	11.09%	33.50%	9.86%	26.79%	16.94%	30.84%
College Degree	3.82%	14.88%	5.63%	14.42%	14.26%	27.86%
Postgraduate	0.51%	4.40%	1.03%	5.48%	7.29%	19.50%
Avg Yrs of Education	10.26	13.08	10.21	12.86	12.21	14.69
Female Share	33.13%	48.00%	18.66%	21.85%	38.09%	54.35%
Foreign-Born Share	6.23%	18.05%	6.63%	20.02%	5.04%	12.88%

A.3   Occupation and sector premia

Figures 3 and 12 as well as our quantitative exercise focuses on relative average residual wages. We obtain these by regressing log hourly wages on sector dummies and on a set of controls, comprising of a polynomial in potential experience (defined as age - years of schooling - 6), dummies for gender, race, and born abroad.

Table 6: Regression of log hourly wages: sector effects

Year	1950	1960	1970	1980	1990	2000	2007
low-sk. serv.	-0.28***	-0.31***	-0.22***	-0.19***	-0.20***	-0.17***	-0.18***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
high-sk. serv.	-0.03***	0.02***	0.08***	0.08***	0.14***	0.17***	0.21***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Controls	yes	yes	yes	yes	yes	yes	yes
Observations	113635	459564	579290	958318	1094458	1235282	1308885
R^2	0.21	0.25	0.21	0.21	0.21	0.18	0.19

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Regression of log hourly wages: occupation effects

Year	1950	1960	1970	1980	1990	2000	2007
manual	-0.41***	-0.42***	-0.33***	-0.28***	-0.24***	-0.19***	-0.10***
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
abstract	0.17***	0.27***	0.32***	0.31***	0.39***	0.44***	0.50***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Controls	yes	yes	yes	yes	yes	yes	yes
Observations	113635	459564	579290	958318	1094458	1235282	1308885
R^2	0.23	0.28	0.25	0.26	0.27	0.26	0.26

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Tables 6 and 7 show the regression results. Since we omit the dummy for manufacturing, the implied relative wage of a low-skilled (high-skilled) service worker is given by the exponential of the estimated coefficient on the low-skilled (high-skilled) service sector dummy. The regression specification to compute residual occupational wages is analogue, with the sector dummies replaced by occupation dummies; we omit the dummy for routine occupations, such that relative wages compared to routine occupations are given by the exponential of the occupation dummies.
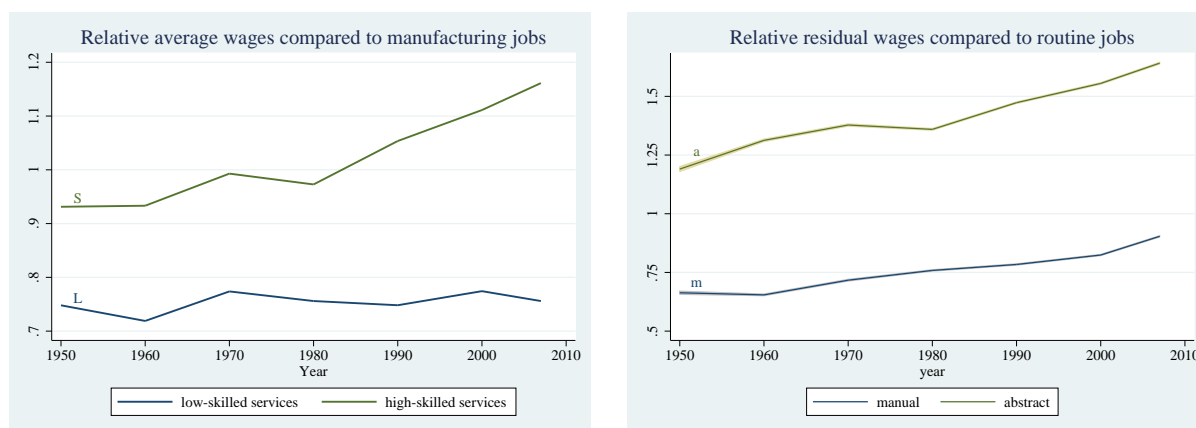
Figure 13: Wage polarization for sectors and occupations

Notes: Same data and classification as in Figure 3 and 12. The left panel shows the relative average wages of high-skilled and low-skilled service workers compared to manufacturing workers. The right panel shows the occupation premium for abstract and manual workers compared to routine workers, and their 95% confidence intervals, as estimated in Table 7.

In the text we show the coefficients on the sectoral dummies from the wage regressions, and in Figure 12 the relative average occupational wages. In Figure 13 we show the reverse: the sectoral relative average wages compared to manufacturing, and the coefficients on occupational dummies from a wage regression. The patterns are un-

changed.

A.4   Alternative wage specifications

In the main text we document the patterns of average low-skilled and high-skilled service wages relative to manufacturing by constructing the sector effects from a regression that controls for a set of observables, in order to remove effects stemming from changes in the composition of the workforce. In particular, we include a fourth-order polynomial in potential experience and dummies for gender, race, and foreign-born as covariates in the (log) wage regression. In Table 8 we show how the predicted sectoral relative wages change when adding further covariates to the regression and when restricting the sample to only men.

The first three rows show the baseline specification's prediction for sectoral relative wages in 1960 and 2007 as well as their percentage change over this period. These are the numbers against which we evaluate our quantitative model. In the rows below alternative sets of further controls are included in the regression. While the quantitative predictions naturally change, the patterns remain, showing an increase of low- and high-skilled service wages relative to manufacturing.

Table 8: Predicted sectoral relative wages in alternative wage regressions

Sample	additional controls	year	L to M	H to M
all	none	1960	0.731	1.021
		2007	0.833	1.238
		1960-2007	13.97%	21.16%
all	interaction of sectoral dummies and experience	1960	0.780	1.108
		2007	0.887	1.350
		1960-2007	13.77%	21.84%
all	dummies for three occupational categories	1960	0.771	0.998
		2007	0.857	1.100
		1960-2007	11.12%	10.20%
all	dummies for ten occupational categories	1960	0.743	0.937
		2007	0.807	1.032
		1960-2007	8.63%	10.11%
all	dummy for college degree	1960	0.734	0.960
		2007	0.832	1.057
		1960-2007	13.34%	10.09%
men	none	1960	0.7675	0.979
		2007	0.850	1.274
		1960-2007	10.75%	30.19%
men	interaction of sectoral dummies and experience	1960	0.788	1.032
		2007	0.880	1.357
		1960-2007	11.69%	31.52%
men	dummies for three occupational categories	1960	0.776	0.953
		2007	0.872	1.114
		1960-2007	12.41%	16.82%
men	dummies for ten occupational categories	1960	0.748	0.919
		2007	0.836	1.062
		1960-2007	11.78%	15.57%
men	dummy for college degree	1960	0.774	0.921
		2007	0.844	1.059
		1960-2007	9.05%	14.95%

Notes: The first 3 rows show the baseline specification's prediction for sectoral relative wages in 1960 and 2007 as well as their percentage change over this period. The subsequent blocks show the predictions when including (alternatively) additional covariates: interaction terms of sectoral dummies and experience, dummies for three occupational categories (manual, routine, abstract), dummies for ten occupational categories (as used in Figure 2), college dummy. The final set of rows show these predictions when restricting the sample to only men.

A.5 The role of gender and age composition changes

Figure 14 demonstrates that the sectoral employment share changes are not driven by changes in the age, gender, race composition of the labor force. The counterfactual industry employment shares are generated by fixing the sectoral employment share of each age-gender-race cell at its 1960 level, and allowing the employment shares of the cells to change. While it can be seen that the counterfactual employment shares (the dashed lines) qualitatively move in the same direction as the actual employment shares (the solid lines), in terms of magnitude the counterfactual employment shares move much less. This implies that the changing composition of the labor force is not the main driving force of the evolution of sectoral employment.
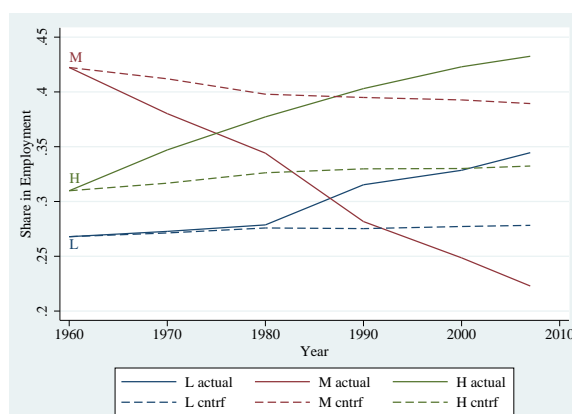
Figure 14: Counterfactual exercise: only changes in the gender-age composition of the labor force

Notes: Employment shares (in terms of hours) are calculated from the same data as in Figure 3. The actual data is shown as solid lines, while the dashed line show how the employment shares of industries would have evolved if only the relative size of gender-age cells in the labor force had changed over time.

A.6 The role of industry shifts in occupational employment shares

In Table 1 of the main text we showed a shift-share decomposition for the changes in occupational employment between 1950 and 2007, and alternatively between 1960 and 2007. In Table 9 we show this decomposition of employment share changes into a between-industry and a within-industry component for each decade. While we find a declining contribution of between-industry shifts since 1980, which might be due to routinization then taking off, again we find that a sizable part of the occupational employment share changes is due to shifts between industries.

Table 9: Decomposition of the changes in occupational employment shares by decade

	1950–60	1960–70	1970–80	1980–90	1990–00	2000–07
3 occupations, 3 sectors						
Manual						
Total Δ	-2.71	-0.07	0.67	0.31	0.85	3.93
Between Δ	-0.94	0.55	0.47	0.95	0.47	0.44
Within Δ	-1.76	-0.63	0.21	-0.65	0.39	3.48
Routine						
Total Δ	-0.65	-3.86	-3.09	-5.57	-5.24	-1.39
Between Δ	0.94	-1.41	-1.22	-1.58	-0.98	-0.70
Within Δ	-1.59	-2.45	-1.86	-3.99	-4.26	-0.69
Abstract						
Total Δ	3.35	3.93	2.41	5.27	4.39	-2.54
Between Δ	0.00	0.85	0.76	0.63	0.51	0.26
Within Δ	3.35	3.08	1.66	4.63	3.87	-2.80
10 occupations, 11 industries						
Manual						
Total Δ	-2.71	-0.07	0.67	0.31	0.85	3.93
Between Δ	-1.51	0.71	0.73	1.16	0.78	0.67
Within Δ	-1.19	-0.78	-0.06	-0.85	0.07	3.26
Routine						
Total Δ	-0.64	-3.86	-3.09	-5.57	-5.24	-1.39
Between Δ	0.85	-2.39	-1.96	-2.21	-1.80	-1.02
Within Δ	-1.49	-1.47	-1.12	-3.36	-3.44	-0.36
Abstract						
Total Δ	3.35	3.93	2.41	5.27	4.39	-2.54
Between Δ	0.67	1.69	1.23	1.05	1.02	0.36
Within Δ	2.69	2.25	1.18	4.21	3.37	-2.90

Notes: Same data as in Figure 1. For each occupational category, the first row presents the change in the share of employment (in terms of hours worked), the second the between-industry component, and the third the within-industry component for the time interval given at the top. The top panel uses 3 occupations and 3 sectors, the bottom panel 10 occupations and 11 industries.

As an alternative way to asses the importance of the employment reallocations between industries for the shifts in the broad occupation categories, we conduct the following counterfactual exercise: we fix the industry shares in employment (in terms of hours worked) at their 1960 levels and let the within-industry share of occupations follow their actual path, and compute how the occupational shares would have evolved in the absence of between-industry shifts. Figure 15 shows the resulting time series (dashed) and the actual data (solid). This exercise shows that if there had been only within-industry shifts, qualitatively the employment of the occupation categories would have evolved as in the actual data, but that quantitatively they cannot explain all of the changes. We therefore conclude that also between-industry shifts account for the polarization of occupational employment.
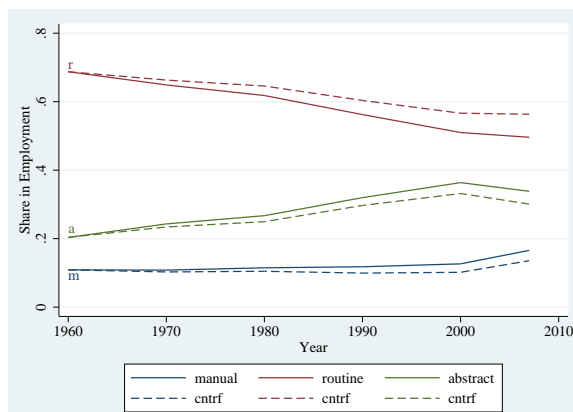
Figure 15: Counterfactual exercise: only-within industry shift of occupations

Notes: Employment shares (in terms of hours) are calculated from the same data as in Figure 13. The actual data is shown as solid lines, while the dashed line show how the occupational employment shares would have evolved in the absence of reallocations across industries.

A.7 Alternative shift-share decomposition

We also conduct an alternative shift-share decomposition, where we use industry level value added shares instead of employment shares. We construct hybrid occupational employment shares as

$$\tilde{E}_{ot} = \sum_i VA_{it}\lambda_{oit},$$

where VA_{it} is the share of industry i in total value added in period t, and λ_{oit} is the share of occupation o, industry i employment within industry i employment in period t, as defined earlier. In general $\tilde{E}_{ot} \neq E_{ot}$, where E_{ot} is simply the share of an occupa-

tion o in total employment, and which is given by $E_{ot} = L_{ot}/L_t = \sum_i E_{it}\lambda_{oit}$, where E_{it} is the share of a industry i in total employment, as before.

Given these hybrid occupational employment shares, we can decompose their change into a part that is driven by within industry occupational employment share changes, and a part that is driven by between industry shifts of value added.[40]

$$\Delta \tilde{E}_{ot} = \underbrace{\sum_i \lambda_{oi}\Delta VA_{it}}_{\equiv \Delta \tilde{E}_{ot}^B} + \underbrace{\sum_i VA_i \Delta\lambda_{oit}}_{\equiv \Delta \tilde{E}_{ot}^W}.$$

Table 10: Decomposition of changes in hybrid occupational employment shares

	Constructed employment shares 1960–2007	
	3 x 3	**10 x 11**
Manual		
Total Δ	3.65	3.35
Between Δ	1.32	0.98
Within Δ	2.36	2.37
Routine		
Total Δ	-19.53	-18.50
Between Δ	-5.29	-6.46
Within Δ	-14.24	-12.50
Abstract		
Total Δ	16.09	15.36
Between Δ	4.18	5.69
Within Δ	11.90	9.68

Notes: Same occupational employment share data as in Figure 1. The value added industry data come from the BEA. For each occupational category, the first row presents the total change, the second the between-industry component, and the third the within-industry component over the period 1960-2007. The first column uses 3 occupations and 3 sectors, column two uses 10 occupations and 11 industries for the decomposition.

Table 10 shows the changes in our hybrid occupational employment shares and their decomposition between 1960 and 2007, into a between-industry and a within-industry component. The value added data comes from the Bureau of Economic Analysis (BEA). Due to the lack of value added data for finer industry categories before the

---

[40]The change driven by shifts between sectors is calculated as the weighted sum of the change in sector i's value added share, ΔVA_{it}, where the weights are the average employment share of occupation o within sector i, $\lambda_{oi} = (\lambda_{oit} + \lambda_{oi0})/2$. The change driven by shifts within sectors is calculated as the weighted sum of the change in occupation o's share within sector i employment, $\Delta\lambda_{oit}$, where the weights are the average value added share of sector i, $VA_i = (VA_{it} + VA_{i0})/2$.

1960s, we decompose changes between 1960 and 2007. Table 10 suggests that between one fourth and one third of occupational employment changes are driven by between industry phenomena, regardless of whether we decompose 3 occupations in 3 sectors, or 10 occupations in 11 sectors. The importance of the between-industry component seems to be somewhat smaller than in the standard shift-share decomposition shown in Table 1, but it is nonetheless quite a substantial share of the overall change.

A.8 Three-way decomposition of relative wage changes

There are three ways of conducting a three-way decomposition:

$$\Delta rw_{ot} = \sum_i \underbrace{\underbrace{\frac{p_{iot}rw_{it} + p_{io0}rw_{i0}}{2}}_{=rw_{io}} \Delta\chi_{iot} + \sum_i \chi_{io}p_{io}\Delta rw_{it}}_{\text{industry effect}} + \underbrace{\sum_i \chi_{io}rw_i\Delta p_{iot}}_{\text{occupation effect}} \tag{18}$$

$$= \underbrace{\sum_i p_{io}rw_i\Delta\chi_{iot} + \sum_i \frac{\chi_{iot}p_{iot} + \chi_{io0}p_{io0}}{2}\Delta rw_{it}}_{\text{industry effect}} + \underbrace{\sum_i \chi_{io}rw_i\Delta p_{iot}}_{\text{occupation effect}} \tag{19}$$

$$= \underbrace{\sum_i p_{io}rw_i\Delta\chi_{iot} + \sum_i \chi_{io}p_{io}\Delta rw_{it}}_{\text{industry effect}} + \underbrace{\sum_i \frac{\chi_{iot}rw_{it} + \chi_{io0}rw_{i0}}{2}\Delta p_{iot}}_{\text{occupation effect}}, \tag{20}$$

where Δ denotes the change between period 0 and t, and the variables without a time subscript denote the average of the variable between period 0 and period t.

The first row is the decomposition we showed in the main body of the paper. The second row gives exactly the same results in terms of the breakdown between industry and occupation effects. The third row gives virtually the same results as summarized in Table 11.

A.9 Decomposition of relative wage changes by decade

In Table 2 of the main text we showed a decomposition of changes in relative occupational wages between 1950 and 2007, and alternatively between 1960 and 2007. In Table 12 we show this decomposition of relative wages changes into an industry and an occupation component for each decade.

Table 11: Alternative decomposition of changes in relative occupational wages

| | Relative wages | | | |
| | 3 x 3 | | 10 x 11 | |
	1950–2007	1960–2007	1950–2007	1960–2007
Manual/Routine				
Total Δ	0.289	0.310	0.289	0.310
Industry Δ	0.181	0.148	0.222	0.216
Occupation Δ	0.107	0.162	0.067	0.094
Abstract/Routine				
Total Δ	0.327	0.240	0.327	0.240
Industry Δ	0.310	0.254	0.381	0.323
Occupation Δ	0.016	-0.013	-0.054	-0.082

Notes: Same data as in Figure 1. For each occupational category, the first row presents the total change, the second the industry component, and the third the occupation component over the period 1950–2007 and over 1960–2007, based on the decomposition equation (20). The first two columns use 3 occupations and 3 sectors, columns three and four 10 occupations and 11 industries.

Table 12: Decomposition of the changes in relative average wages by decade

	1950–60	1960–70	1970–80	1980–90	1990–00	2000–07
	3 occupations, 3 sectors					
Manual/Routine						
Total Δ	-0.022	0.085	0.023	0.033	0.036	0.134
Industry Δ	0.024	0.061	-0.014	0.034	0.032	0.043
Occupation Δ	-0.046	0.024	0.037	-0.001	0.005	0.091
Abstract/Routine						
Total Δ	0.086	0.052	-0.077	0.107	0.083	0.076
Industry Δ	0.061	0.052	-0.017	0.105	0.075	0.046
Occupation Δ	0.025	-0.000	-0.060	0.002	0.008	0.029
	10 occupations, 11 industries					
Manual/Routine						
Total Δ	-0.022	0.085	0.023	0.033	0.036	0.134
Industry Δ	-0.006	0.076	-0.022	0.042	0.045	0.065
Occupation Δ	-0.016	0.009	0.045	-0.010	-0.008	0.068
Abstract/Routine						
Total Δ	0.086	0.052	-0.077	0.107	0.083	0.076
Industry Δ	0.065	0.067	-0.024	0.124	0.087	0.047
Occupation Δ	0.021	-0.015	-0.054	-0.017	-0.004	0.029

Notes: Same data as in Figure 1. For each occupational category, the first row presents the total change, the second the industry component, and the third the occupation component for the time interval given at the top, based on the decomposition equation (18). The top panel uses 3 occupations and 3 sectors, the bottom panel 10 occupations and 11 industries.

A.10 Historical data

Given that our model suggests that structural transformation leads to the employment compression of occupations most intensively used in the shrinking sector of the economy, we look at pre-1950 data to see whether this prediction also holds over longer horizons. There are some caveats to note. First, hours worked and wage data are not available, so we can only look at employment patterns in terms of persons employed, and we cannot analyze wage patterns. Given the lack of wage data, it is also hard to verify whether these labor market patterns resemble polarization or have different implications. Second, in the period 1850–1900 the Census used the 1880 occupational classification system, where workers' occupations were to some extent inferred from their industry.[41] This means that by construction there is a significant overlap between industry and occupation classifications prior to 1900. With these caveats in mind, we analyze the patterns of employment between 1850 and 1940. Since in the 1850s a large fraction of the workforce was employed in agriculture, we do not drop agricultural workers, but instead add extra categories for them, both as an occupation and as a sector.
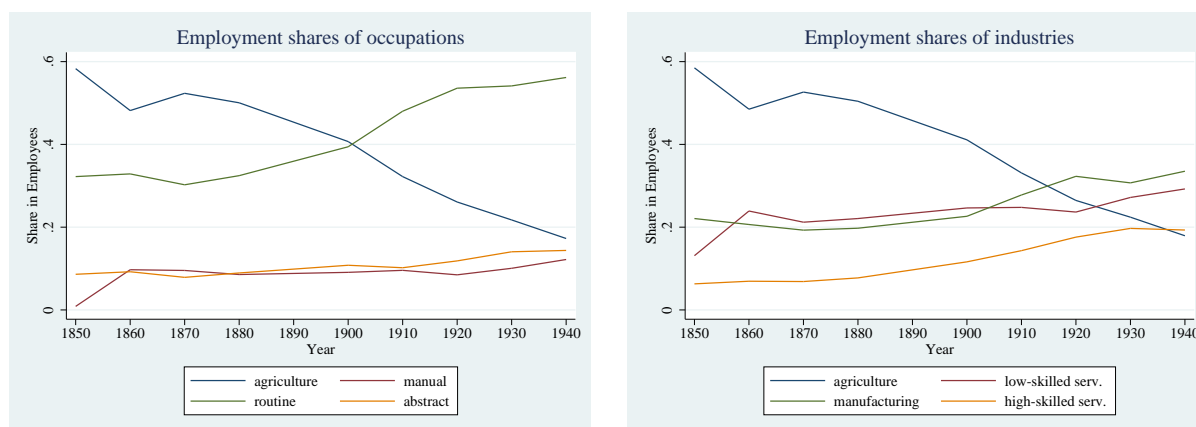
Figure 16: Employment patterns 1850–1940

Notes: The graphs is based on Census data between 1850 and 1940. Each worker is classified into one of four occupations based on their occupation code (*occ1950*) and one of four sectors based on their industry code (*ind1950*). Both graphs show employment shares in terms of number of people. The left panel shows employment shares in terms of occupations, while the right panel shows them in terms of sectors.

The employment share patterns are shown in Figure 16. This figure shows that the

---

[41] The IPUMS documentation writes: *"In 1850–1900, occupations are classified according to the 1880 system. The 1880 occupational classification was oriented more to work settings and economic sectors – what is now termed "industry" – than to workers' specific technical functions."*

defining trend in terms of sectors in the period 1850–1940 was the declining employment share of agriculture, and a slow increase in the other three sectors (low-skilled services, manufacturing and high-skilled services). In terms of occupations we see a parallel compression of agricultural occupations, a quite pronounced increase in routine workers, and a slow increase in manual and abstract workers. Thus even prior to 1950 we see quite a close connection between sectoral and occupational employment share trends.

Next we conduct a shift-share decomposition of occupational employment shares (as in section 2.3). This decomposition, summarized in Table 13, confirms what Figure 16 already suggests, that sectoral and occupational employment patterns are quite closely connected. The decomposition shows that almost all of the decline in agricultural occupations is driven by employment moving away from the agricultural sector; that abstract and routine employment are expanding due to the movement of labor into sectors where these occupations are used more intensively; and that manual employment is also partly expanding due to sectoral labor reallocation.

The historical data confirms that the structural transformation of the economy has a significant impact on occupational employment patterns even prior to the 1950s. In particular it seems that in the period 1850–1940 as the agricultural sector was shrinking, while manufacturing and low-and high-skilled services were increasing, the employment share in agricultural occupations fell, while the employment share in routine and abstract occupations increased, largely driven by the sectoral reallocation labor.

Table 13: Shift-share decomposition of occupational employment share changes

| | 4x4 | | | 12x14 | | |
	1850–1900	1900–1940	1850–1940	1850–1900	1900-1940	1850-1940
Agricultural						
Total Δ	-17.62	-23.43	-41.04	-38.57	-23.57	-68.61
Between Δ	-17.28	-22.61	-39.72	-21.49	-23.36	-53.23
Within Δ	-0.34	-0.81	-1.32	-17.08	-0.20	-15.37
Manual						
Total Δ	8.23	3.10	11.33	16.00	1.93	15.34
Between Δ	2.48	2.42	4.13	6.25	1.21	6.03
Within Δ	5.75	0.68	7.20	9.75	0.72	9.32
Routine						
Total Δ	7.21	16.74	23.95	4.49	20.96	26.30
Between Δ	9.06	15.78	24.99	13.81	19.31	31.67
Within Δ	-1.86	0.97	-1.03	-9.32	1.65	-5.37
Abstract						
Total Δ	2.18	3.58	5.76	2.82	4.35	7.15
Between Δ	5.73	4.42	10.61	6.63	7.67	15.28
Within Δ	-3.55	-0.84	-4.85	-3.81	-3.32	-8.13

Notes: Same data as in Figure 16. For each occupational category, the first row presents the total change, the second the between industry component, and the third the within industry component over the period 1850–1900, 1900–1940 and 1850–1940, based on the decomposition equation (1). The first three columns use 4 occupations and 4 sectors (as in Figure 16), the last three 12 occupations and 14 industries (same categories as in Table 1 with the following additional occupations:'farmers and farm managers', 'farm laborers', and industries: 'agriculture', 'forestry' and 'fishing').

B   Model appendix

Proof of Proposition 1. To simplify notation denote the relative unit wages by $\widehat{a}_m \equiv \frac{\omega_l}{\omega_m}$ and $\widehat{a}_h \equiv \frac{\omega_l}{\omega_h}$.

Starting from:

$$\frac{N_l\left(\widehat{a}_m, \frac{\widehat{a}_m}{\widehat{a}_h}\right)}{N_m\left(\widehat{a}_m, \frac{\widehat{a}_m}{\widehat{a}_h}\right)}\widehat{a}_m^\varepsilon = \left(\frac{A_m}{A_l}\right)^{1-\varepsilon}\left(\frac{\theta_m}{\theta_l}\right)^{-\varepsilon},$$

$$\frac{N_h\left(\widehat{a}_m, \frac{\widehat{a}_m}{\widehat{a}_h}\right)}{N_m\left(\widehat{a}_m, \frac{\widehat{a}_m}{\widehat{a}_h}\right)}\left(\frac{\widehat{a}_m}{\widehat{a}_h}\right)^\varepsilon = \left(\frac{A_m}{A_h}\right)^{1-\varepsilon}\left(\frac{\theta_m}{\theta_h}\right)^{-\varepsilon}.$$

A change in productivities triggers changes in the equilibrium cutoffs, $\widehat{a}_m$ and $\widehat{a}_h$, in such a way that the above conditions remain satisfied. Total differentiation then implies:

$$\varepsilon\frac{d\widehat{a}_m}{\widehat{a}_m} + \frac{dN_l}{N_l} - \frac{dN_m}{N_m} = (1-\varepsilon)\frac{d\frac{A_m}{A_l}}{\frac{A_m}{A_l}}, \tag{21}$$

$$\varepsilon\left(\frac{d\widehat{a}_m}{\widehat{a}_m} - \frac{d\widehat{a}_h}{\widehat{a}_h}\right) + \frac{dN_h}{N_h} - \frac{dN_m}{N_m} = (1-\varepsilon)\frac{d\frac{A_m}{A_h}}{\frac{A_m}{A_h}}. \tag{22}$$

Applying the Leibniz rule to the expressions for $N_l\left(\widehat{a}_m, \frac{\widehat{a}_m}{\widehat{a}_h}\right)$, $N_m\left(\widehat{a}_m, \frac{\widehat{a}_m}{\widehat{a}_h}\right)$ and $N_h\left(\widehat{a}_m, \frac{\widehat{a}_m}{\widehat{a}_h}\right)$, we get the following expressions for the change in the effective and raw labor supplies as a function of the change in $\widehat{a}_m$ and in $\widehat{a}_h$ is

$$dN_l\left(\widehat{a}_m, \frac{\widehat{a}_m}{\widehat{a}_h}\right) = \frac{\partial N_l}{\partial\widehat{a}_m}d\widehat{a}_m + \frac{\partial N_l}{\partial\widehat{a}_h}d\widehat{a}_h = \underbrace{\int_0^\infty\int_0^{\widehat{a}_h a_l} a_l^2 f(a_l, \widehat{a}_m a_l, a_h)da_h da_l}_{\equiv C_1 > 0}\cdot\widehat{a}_m d\widehat{a}_m$$

$$+ \underbrace{\int_0^\infty\int_0^{\widehat{a}_m a_l} a_l^2 f(a_l, a_m, \widehat{a}_h a_l)da_m da_l}_{\equiv C_2 > 0}\cdot\widehat{a}_h d\widehat{a}_h, \tag{23}$$

$$dN_m\left(\widehat{a}_m,\frac{\widehat{a}_m}{\widehat{a}_h}\right)=-\underbrace{\int_0^\infty\int_0^{\frac{\widehat{a}_h}{\widehat{a}_m}a_m}a_m^2 f(\frac{a_m}{\widehat{a}_m},a_m,a_h)da_h da_m}_{=C_3>0}\cdot\frac{1}{\widehat{a}_m^2}d\widehat{a}_m$$

$$-\underbrace{\int_0^\infty\int_0^{\frac{a_m}{\widehat{a}_m}}a_m^2 f(a_l,a_m,\frac{\widehat{a}_h}{\widehat{a}_m}a_m)da_l da_m}_{\equiv C_4>0}\cdot\frac{\widehat{a}_h}{\widehat{a}_m}\left(\frac{d\widehat{a}_m}{\widehat{a}_m}-\frac{d\widehat{a}_h}{\widehat{a}_h}\right),\qquad(24)$$

$$dN_h\left(\widehat{a}_m,\frac{\widehat{a}_m}{\widehat{a}_h}\right)=-\underbrace{\int_0^\infty\int_0^{\frac{\widehat{a}_m}{\widehat{a}_h}a_h}a_h^2 f(\frac{a_h}{\widehat{a}_h},a_m,a_h)da_m da_h}_{=C_5>0}\cdot\frac{1}{\widehat{a}_h^2}d\widehat{a}_h$$

$$+\underbrace{\int_0^\infty\int_0^{\frac{a_h}{\widehat{a}_h}}a_h^2 f(a_l,\frac{\widehat{a}_m}{\widehat{a}_h}a_h,a_h)da_l da_h}_{\equiv C_6>0}\cdot\frac{\widehat{a}_m}{\widehat{a}_h}\left(\frac{d\widehat{a}_m}{\widehat{a}_m}-\frac{d\widehat{a}_h}{\widehat{a}_h}\right).\qquad(25)$$

Similarly

$$dL_l\left(\widehat{a}_m,\frac{\widehat{a}_m}{\widehat{a}_h}\right)=\underbrace{\int_0^\infty\int_0^{\widehat{a}_h a_l}a_l f(a_l,\widehat{a}_m a_l,a_h)da_h da_l}_{\equiv\tilde{C}_1>0}\cdot\widehat{a}_m d\widehat{a}_m$$

$$+\underbrace{\int_0^\infty\int_0^{\widehat{a}_m a_l}a_l f(a_l,a_m,\widehat{a}_h a_l)da_m da_l}_{\equiv\tilde{C}_2>0}\cdot\widehat{a}_h d\widehat{a}_h,\qquad(26)$$

$$dL_m\left(\widehat{a}_m,\frac{\widehat{a}_m}{\widehat{a}_h}\right)=-\underbrace{\int_0^\infty\int_0^{\frac{\widehat{a}_h}{\widehat{a}_m}a_m}a_m f(\frac{a_m}{\widehat{a}_m},a_m,a_h)da_h da_m}_{\equiv\tilde{C}_3>0}\cdot\frac{1}{\widehat{a}_m^2}d\widehat{a}_m$$

$$-\underbrace{\int_0^\infty\int_0^{\frac{a_m}{\widehat{a}_m}}a_m f(a_l,a_m,\frac{\widehat{a}_h}{\widehat{a}_m}a_m)da_l da_m}_{\equiv\tilde{C}_4>0}\cdot\frac{\widehat{a}_h}{\widehat{a}_m}\left(\frac{d\widehat{a}_m}{\widehat{a}_m}-\frac{d\widehat{a}_h}{\widehat{a}_h}\right),\qquad(27)$$

$$dL_h\left(\widehat{a}_m,\frac{\widehat{a}_m}{\widehat{a}_h}\right)=-\underbrace{\int_0^\infty\int_0^{\frac{\widehat{a}_m}{\widehat{a}_h}a_h}a_h f(\frac{a_h}{\widehat{a}_h},a_m,a_h)da_m da_h}_{\equiv\tilde{C}_5>0}\cdot\frac{1}{\widehat{a}_h^2}d\widehat{a}_h$$

$$+\underbrace{\int_0^\infty\int_0^{\frac{a_h}{\widehat{a}_h}}a_h f(a_l,\frac{\widehat{a}_m}{\widehat{a}_h}a_h,a_h)da_l da_h}_{\equiv\tilde{C}_6>0}\cdot\frac{\widehat{a}_m}{\widehat{a}_h}\left(\frac{d\widehat{a}_m}{\widehat{a}_m}-\frac{d\widehat{a}_h}{\widehat{a}_h}\right).\qquad(28)$$

Plugging these into (21) and (22) and re-arranging we get:

$$\frac{d\widehat{a}_m}{\widehat{a}_m}\underbrace{\left[\varepsilon + \frac{C_1\widehat{a}_m^2}{N_l} + \frac{C_3\frac{1}{\widehat{a}_m}}{N_m} + \frac{C_4\frac{\widehat{a}_h}{\widehat{a}_m}}{N_m}\right]}_{\equiv B_1 > 0} + \frac{d\widehat{a}_h}{\widehat{a}_h}\underbrace{\left[\frac{C_2\widehat{a}_h^2}{N_l} - \frac{C_4\frac{\widehat{a}_h}{\widehat{a}_m}}{N_m}\right]}_{\equiv B_2} = \underbrace{(1-\varepsilon)\frac{d\frac{A_m}{A_l}}{\frac{A_m}{A_l}}}_{\equiv D_1},$$

$$\frac{d\widehat{a}_m}{\widehat{a}_m}\underbrace{\left[\varepsilon + \frac{C_6\frac{\widehat{a}_m}{\widehat{a}_h}}{N_h} + \frac{C_3\frac{1}{\widehat{a}_m}}{N_m} + \frac{C_4\frac{\widehat{a}_h}{\widehat{a}_m}}{N_m}\right]}_{\equiv B_3 > 0} - \frac{d\widehat{a}_h}{\widehat{a}_h}\underbrace{\left[\varepsilon + \frac{C_5\frac{1}{\widehat{a}_h}}{N_h} + \frac{C_6\frac{\widehat{a}_m}{\widehat{a}_h}}{N_h} + \frac{C_4\frac{\widehat{a}_h}{\widehat{a}_m}}{N_m}\right]}_{\equiv B_4 > 0} = \underbrace{(1-\varepsilon)\frac{d\frac{A_m}{A_h}}{\frac{A_m}{A_h}}}_{\equiv D_2}.$$

This leads to

$$\frac{d\widehat{a}_h}{\widehat{a}_h} = \frac{B_3 D_1 - B_1 D_2}{B_3 B_2 + B_1 B_4},$$
$$\frac{d\widehat{a}_m}{\widehat{a}_m} = \frac{D_2 B_2 + B_4 D_1}{B_3 B_2 + B_1 B_4},$$

where $B_3 B_2 + B_1 B_4 > 0$ can be easily verified by multiplying out the terms. Hence to determine the response in $\widehat{a}_m$ and in $\widehat{a}_h$, we only need to consider the sign of the numerator. If $D_1 = D_2 > 0$, i.e. the growth rate of A_l is equal to the growth rate of A_h, and lower than the growth rate of A_m, then the following expressions can be obtained:

$$\frac{d\widehat{a}_h}{\widehat{a}_h} = \frac{D}{B_3 B_2 + B_1 B_4}(B_3 - B_1) = \frac{D}{B_3 B_2 + B_1 B_4}\left(\frac{C_6\frac{\widehat{a}_m}{\widehat{a}_h}}{N_h} - \frac{C_1\widehat{a}_m^2}{N_l}\right), \tag{29}$$

$$\frac{d\widehat{a}_m}{\widehat{a}_m} = \frac{D}{B_3 B_2 + B_1 B_4}(B_2 + B_4) = \frac{D}{B_3 B_2 + B_1 B_4}\left(\varepsilon + \frac{C_5\frac{1}{\widehat{a}_h}}{N_h} + \frac{C_6\frac{\widehat{a}_m}{\widehat{a}_h}}{N_h} + \frac{C_2\widehat{a}_h^2}{N_l}\right) > 0. \tag{30}$$

As this shows, $\frac{d\widehat{a}_m}{\widehat{a}_m} > 0$. The sign of $\frac{d\widehat{a}_h}{\widehat{a}_h}$ is ambiguous in general, but it is straightforward that $\frac{d\widehat{a}_m}{\widehat{a}_m} - \frac{d\widehat{a}_h}{\widehat{a}_h} > 0$:

$$\left(\frac{d\widehat{a}_m}{\widehat{a}_m} - \frac{d\widehat{a}_h}{\widehat{a}_h}\right) = \frac{D}{B_3 B_2 + B_1 B_4}\left(\varepsilon + \frac{C_1\widehat{a}_m^2}{N_l} + \frac{C_2\widehat{a}_h^2}{N_l} + \frac{C_5\frac{1}{\widehat{a}_h}}{N_h}\right) > 0.$$

To summarize the changes in relative unit wages, $\omega_l/\omega_m = \widehat{a}_m$ and $\omega_h/\omega_m = \widehat{a}_m/\widehat{a}_h$ increases, while $\omega_l/\omega_h = \widehat{a}_h$ can increase or decrease.

These together with (24) and (27) imply that N_m and L_m always decrease. These

changes are:

$$dN_m\left(\widehat{a}_m, \frac{\widehat{a}_m}{\widehat{a}_h}\right) = -\left(C_3 \frac{1}{\widehat{a}_m} \frac{d\widehat{a}_m}{\widehat{a}_m} + C_4 \frac{\widehat{a}_h}{\widehat{a}_m}\left(\frac{d\widehat{a}_m}{\widehat{a}_m} - \frac{d\widehat{a}_h}{\widehat{a}_h}\right)\right) < 0,$$

$$dL_m\left(\widehat{a}_m, \frac{\widehat{a}_m}{\widehat{a}_h}\right) = -\left(\tilde{C}_3 \frac{1}{\widehat{a}_m} \frac{d\widehat{a}_m}{\widehat{a}_m} + \tilde{C}_4 \frac{\widehat{a}_h}{\widehat{a}_m}\left(\frac{d\widehat{a}_m}{\widehat{a}_m} - \frac{d\widehat{a}_h}{\widehat{a}_h}\right)\right) < 0.$$

By plugging in (29) and (30) into (23) we can show that effective employment in sector L increases:

$$dN_l\left(\widehat{a}_m, \frac{\widehat{a}_m}{\widehat{a}_h}\right) = C_1 \widehat{a}_m^2 \frac{d\widehat{a}_m}{\widehat{a}_m} + C_2 \widehat{a}_h^2 \frac{d\widehat{a}_h}{\widehat{a}_h}$$

$$= \frac{D}{B_3 B_2 + B_1 B_4}\left[C_1 \widehat{a}_m^2\left(\varepsilon + \frac{C_5 \frac{1}{\widehat{a}_h}}{N_h} + \frac{C_6 \frac{\widehat{a}_m}{\widehat{a}_h}}{N_h}\right) + C_2 \widehat{a}_h^2 \frac{C_6 \frac{\widehat{a}_m}{\widehat{a}_h}}{N_h}\right] > 0.$$

By plugging in (29) and (30) into (25) we can show that effective employment in sector H increases:

$$dN_h\left(\widehat{a}_m, \frac{\widehat{a}_m}{\widehat{a}_h}\right) = \frac{D}{B_3 B_2 + B_1 B_4}\left[C_5 \frac{1}{\widehat{a}_h} \frac{C_1 \widehat{a}_m^2}{N_l} + C_6 \frac{\widehat{a}_m}{\widehat{a}_h}\left(\varepsilon + \frac{C_6 \frac{\widehat{a}_m}{\widehat{a}_h}}{N_h} + \frac{C_2 \widehat{a}_h^2}{N_l}\right)\right] > 0.$$

$\square$

C   Quantitative results appendix

C.1   Non-transitory wage dispersion in the PSID 1968–1975

To calibrate the distribution of labor efficiencies in the quantitative model, we target as a fifth moment (besides relative average sectoral wages and employment shares in the 1960 Census data) the variance of the non-transitory component of log wages, which we estimate similarly to Lagakos and Waugh (2013). To compute this statistic, we require panel data. We therefore use data from the PSID family index from its launch in 1968 to 1975. To ensure that we are tracking individuals correctly over time, we restrict the sample to households in which neither the household head nor the spouse changes over this period. We restrict the sample further to individuals who are between 16 and 65 years of age, employed outside of agriculture, report hourly wages

that are not below the federal minimum wage, and whose wages are observed at least twice over 1968–1975.

Like Lagakos and Waugh (2013), we want to extract the non-transitory components in log wages to construct the calibration target for our model. We run a regression of log hourly wages on individual fixed effects and year fixed effects, and then compute the variance of the individual fixed effects. This gives a value of 0.187, which we use as our target for the dispersion of the non-transitory component in log wages.

C.2 Robustness

In section 4.3 of the main text we summarized how our result change when assuming a different underlying distribution of sectoral efficiencies and when varying the elasticity of substitution between goods and services (measured in value-added terms). Here we show in Table 14 and 15 the predictions of the model for various calibrations, assuming a (trivariate) log-normal distribution or a truncated normal distribution respectively.

Table 14: Robustness checks: different correlations vs the data

parameters								rel. avg. wage Δ	
ρ_{lm}	ρ_{mh}	ρ_{lh}	σ_l^2	σ_m^2	σ_h^2	τ_l	τ_h	L to M	H to M
0.0	0	0.0	0.13	0.36	0.45	0.56	0.86	10.53	10.13
0.0	0	0.3	0.13	0.36	0.40	0.53	0.83	10.76	10.67
0.0	0	0.6	0.13	0.37	0.34	0.50	0.79	10.90	11.19
0.0	0.3	0.0	0.14	0.32	0.39	0.58	0.84	9.79	7.75
0.0	0.3	0.3	0.13	0.33	0.34	0.56	0.81	10.31	8.49
0.0	0.3	0.6	0.14	0.34	0.29	0.53	0.77	10.81	9.13
0.0	0.6	0.0	0.14	0.30	0.34	0.61	0.82	8.88	5.34
0.0	0.6	0.3	0.14	0.31	0.29	0.58	0.78	9.76	6.37
0.0	0.6	0.6	0.14	0.32	0.24	0.55	0.74	10.67	6.96
0.3	0	0.0	0.12	0.32	0.45	0.54	0.89	8.53	8.68
0.3	0	0.3	0.12	0.32	0.41	0.51	0.85	8.74	9.28
0.3	0	0.6	0.12	0.33	0.35	0.49	0.81	8.68	9.79
0.3	0.3	0.0	0.13	0.29	0.40	0.57	0.87	7.71	6.24
0.3	**0.3**	**0.3**	**0.12**	**0.30**	**0.35**	**0.54**	**0.84**	**8.25**	**7.14**
0.3	0.3	0.6	0.13	0.31	0.30	0.51	0.80	8.59	7.91
0.3	0.6	0.0	0.13	0.27	0.35	0.60	0.85	6.54	3.54
0.3	0.6	0.3	0.13	0.28	0.31	0.57	0.81	7.56	4.94
0.3	0.6	0.6	0.13	0.29	0.27	0.54	0.77	8.30	6.00
0.6	0	0.0	0.13	0.28	0.46	0.52	0.92	6.41	6.91
0.6	0	0.3	0.12	0.29	0.41	0.50	0.88	6.57	7.52
0.6	0	0.6	0.12	0.29	0.35	0.47	0.84	6.25	7.85
0.6	0.3	0.0	0.13	0.25	0.41	0.55	0.90	5.41	4.31
0.6	0.3	0.3	0.12	0.26	0.36	0.52	0.87	6.11	5.42
0.6	0.3	0.6	0.12	0.27	0.32	0.49	0.83	6.33	6.21
0.6	0.6	0.0	0.14	0.23	0.36	0.58	0.89	3.47	0.78
0.6	0.6	0.3	0.13	0.24	0.32	0.55	0.85	5.18	2.96
0.6	0.6	0.6	0.13	0.25	0.28	0.52	0.81	5.99	4.36
Data								**14.00**	**21.17**

Notes: This table shows the calibration of the lognormal distribution as described in section 4.1 for all possible combinations of correlation structures where each correlation is from the $\{0, 0.3, 0.6\}$ set. The bold row in the middle shows our baseline calibration. The first three columns show the assumed correlations, the next five the calibrated parameters, and the final two show the implied relative average wage change of the low- and high-skilled service sector compared to manufacturing. The last row contains the change in these same measures between 1960 and 2007 in the data.

Table 15: Robustness checks: truncated normal distribution

parameters								rel. avg. wage Δ	
ρ_{lm}	ρ_{mh}	ρ_{lh}	σ_l^2	σ_m^2	σ_h^2	τ_l	τ_h	L to M	H to M
0.0	0.0	0.0	0.28	0.80	2.19	0.57	0.89	24.39	23.76
0.0	0.3	0.3	0.28	0.78	1.43	0.56	0.84	25.04	23.05
0.3	0.0	0.3	0.22	0.65	1.72	0.52	0.88	21.55	21.89
0.3	0.3	0.0	0.25	0.63	1.85	0.56	0.90	21.14	20.36
0.3	0.3	0.3	0.28	0.80	2.19	0.55	0.89	23.69	22.73
0.3	0.3	0.6	0.26	0.67	1.10	0.51	0.84	22.95	21.94
0.3	0.6	0.3	0.25	0.66	1.20	0.57	0.85	21.91	19.45
0.6	0.3	0.3	0.22	0.51	1.55	0.52	0.91	18.91	18.83
0.6	0.6	0.6	0.25	0.55	0.95	0.52	0.85	20.26	18.58
Data								**14.00**	**21.17**

Notes: This table shows the calibration of the truncated normal distribution as described in section 4.1 for nine correlation structures where each correlation is from the $\{0, 0.3, 0.6\}$ set. The first three columns show the assumed correlations, the next five the calibrated parameters, and the final two show the implied relative average wage change of the low- and high-skilled service sector compared to manufacturing. The last row contains the change in these same measures between 1960 and 2007 in the data.