Complex-Skill Biased Technical Change and Labor Market Polarization*

Colin Caines†
University of British Columbia

Florian Hoffmann‡
University of British Columbia

Gueorgui Kambourov§
University of Toronto

Very preliminary and incomplete.

Abstract

We document that, regardless of their position in the 1980 skill distribution (as measured by the average occupational wage in 1980), *complex* occupations that require the ability to abstract, analyze, make connections, or form decisions have experienced higher wage and employment growth over the 1980-2005 period than *simple* occupations that do not require such skills. We use the German BiBB dataset in order to classify occupations into complex and simple and map that classification into the US occupational codes. The occupations in the top third of the 1980 skill distribution are classified as *managerial* and *college* and are treated as an extreme case of complex occupations. Our main finding is that, controlling for the percentile in the 1980 skill distribution, complex occupations have experienced 11.2 percentage points higher wage growth than simple occupations over the period. Complex occupations have also experienced a higher employment growth, conditional on the percentile of the 1980 skill distribution, than simple occupations. Once we control for the occupation's type − simple, complex, managerial, and college − we find that the routine task intensity index for an occupation has no effect on the observed wage and employment growth. Therefore, we conclude that a complex-skill biased technical change − an increase in the demand for complex skills relative to the demand for simple skills − is a plausible explanation for the observed wage and employment growth patterns observed in the data over the 1980-2005 period. This finding has significant implications for government policy since apprenticeship programs, in addition to college education, can assist workers in accumulating complex skills and facilitate their transition into the complex occupations in the economy.

JEL Classification: .

Keywords: .

*This version: July 14, 2015.

†Department of Economics, University of British Columbia, 1873 East Mall, Vancouver, BC, V6T 1Z1, Canada. E-mail: colin.cw.caines@gmail.com.

‡Department of Economics, University of British Columbia, 1873 East Mall, Vancouver, BC, V6T 1Z1, Canada. E-mail: Florian.Hoffmann@ubc.ca.

§University of Toronto, Department of Economics, 150 St. George St., Toronto, Ontario M5S 3G7, Canada. E-mail: g.kambourov@utoronto.ca.

1 Introduction

A large literature has studied the effect of the computerization and routinization of the production process on the observed changes in the wage and employment distribution over the 1980-2005 time period. For example, Autor and Dorn (2013) (AD) have recently provided evidence that the computerization and routinization of occupations in the middle of the 1980 skill distribution (as measured by the average occupational wage in 1980) has lead to a relatively slower wage and employment growth in those occupations over the period. In contrast, occupations at the low end of the skill distribution in 1980, the so called service occupations, did relatively well over the 1980-2005 period in terms of wage and employment growth due to the fact the production process in these occupations is not easily computerizable. Finally, the occupations at the high end of the skill distribution in 1980 − college and professional occupations − have experienced relatively high wage and employment growth over the period. However, the average experience of an occupation in the bottom and the middle of the 1980 skill distribution hides significant variation which is informative of the underlying forces affecting the dynamics of wage and employment occupational growth. If we focus on the occupations in the middle of the 1980 skill distribution − which on average experienced a slow wage and employment growth due to the performance of such occupations as truck drivers − some occupations did actually quite well: air pilots experienced significant wage and employment growth. The same is true at the lower end of the 1980 skill distribution: cooks did quite well in terms of wage and employment growth, but cleaners did not.

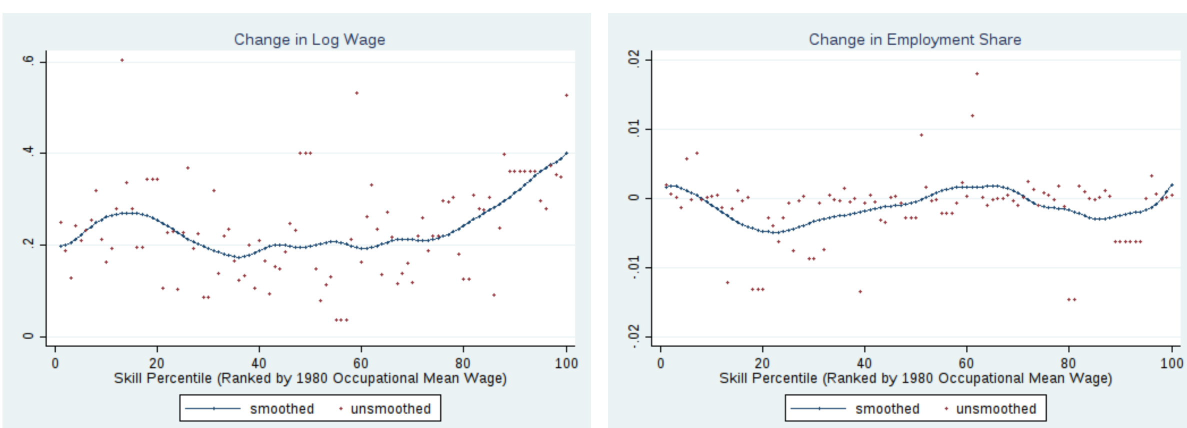
In this paper we show that, regardless of their position in the 1980 skill distribution, *complex* occupations that require the ability to abstract, analyze, make connections, or form decisions have experienced higher wage and employment growth over the 1980-2005 period than *simple* occupations that do not require such skills. We use the German BiBB dataset in order to classify occupations into complex and simple and map that classification into the US occupational codes. The occupations in the top third of the 1980 skill distribution are classified as *managerial* and *college* and are treated as an extreme case of complex occupations. Our main finding is that, controlling for the percentile in the 1980 skill distribution, complex occupations have experienced 11.2 percentage points higher wage growth than simple occupations over the period. Complex occupations have also experienced a higher employment growth, conditional on the percentile of the 1980 skill distribution, than simple occupations. Once we control for the occupation's type − simple, complex, managerial, and college − we find that the routine task intensity index for an occupation has no effect on the observed wage and employment growth. Therefore, we conclude that a complex-skill biased technical change − an increase in the demand for complex skills relative to the demand for simple skills − is a plausible explanation for the observed wage and employment growth patterns observed in the data over the 1980-2005 period. This finding has significant implications for government policy since apprenticeship programs, in addition to college education, can assist workers in accumulating complex skills and facilitate their transition into the complex occupations in the economy.

The paper is structured as follows. We begin with the recent analysis in Autor and Dorn (2013) on wage and employment polarization and its relationship to the growth in low-skill services. We highlight significant deviations from the average non-parametric relationship between AD's measure of occupational skill content in 1980 and occupation-level changes in wages and employment. These deviations are frequent in the middle of the occupational skill distribution, that is, among jobs which are thought to have disappeared. Since the level of aggregation is the 3-digit occupation, we argue that such deviations should not be interpreted as statistical outliers, but rather as deviations with economic content that contain valuable information about the sources of the changing wage structure. Once one takes into account these deviations we argue that the popular notion of disappearing middle-class jobs masks important exceptions. We then use an alternative way of measuring the skill content of occupations and show that it has considerable explanatory power to explain the deviations from AD's non-parametric trend line. Our categorization does not rely on the concept of computerization, but more explicitly on the type of skills that are likely to be important for being productive in a particular occupations. We therefore argue that our approach moves the study of the changes in the occupational wage and employment structure closer to the literature on the returns to educational interventions.

2 Motivation

In this section we conduct a descriptive analysis of the changes in the occupational wage and employment structure that have taken place over the last few decades.

Figure 1: Wage and Employment Growth by Skill Percentile.



We start with reproducing AD's graphical description of polarization of the wage and employment distribution. To this end we have collected IPUMS Census and ACS data for which we imposed identical sample restrictions like those in Autor and Dorn (2013).[1] Changes in log-wages and employment on the 3-digit occupational level are shown in the two panels of Figure 1. The skill content of an occupation is measured by its average wage in 1980, and each 3-digit occupation is sorted into a percentile of the occupation-level wage distribution in 1980. As with Autor and Dorn (2013) we find that there are a number of low-skill occupations that fared quite well in terms of wage growth between 1980 and 2005. Occupations between the 10th and 25th percentile of this distribution experienced a wage growth similar to occupations at approximately the 80th percentile. In contrast, occupations in the middle of the skill distribution had the lowest wage growth, on average. Similar evidence of wage polarization has been documented extensively not only in US data, but also in data for various industrialized countries, such as the UK, Germany, and Australia. Corresponding with the distribution of wage changes in the first panel, the second panel shows that employment growth was above average exactly for the occupation groups that experienced above-average wage growth.

A popular theory of these changes, commonly associated with the analysis in Autor et al. (2003), is that medium-skilled occupations have been computerized because they are intensive in routine tasks, leading to their decline in terms of employment shares or relative wages. An implication is that occupations that cannot be computerized easily have performed relatively well. AD highlight the role of the rise in low-skill services for explaining the large wage growth of occupations that had a low standing in the 1980 skill distribution. Consistent with this hypothesis, we find that low-skill service occupations had a wage growth of 21 percentage points, compared with the 11 percentage point growth observed amongst Production, Craft and Repair occupations and 9 percentage point growth seen by Operators, Fabricators and Laborers over the same time period. Table 1 shows that among five broadly defined occupation groups, only Professional and Managerial occupations and the Service occupations experienced positive employment growth.

The computerization hypothesis, which is about the shape of technological changes, is evidently a powerful explanation for this polarization. However, it is difficult to draw any particular policy conclusions from this, partly because of the limited role of skills in such explanations. An educational intervention that increases skills from a very low to a median level would be predicted to have no positive impact because median-skilled occupations are exactly those with the smallest average wage and job growth. But the occupations contained in the middle of the skill distribution are very heterogeneous. For example, it includes truck drivers and auto body repairers, both of which are likely to require different sets of skills. As a result one may wonder if both of them had low wage and employment growth. It is therefore also useful to look at the "raw" data, rather than the "smoothed" one by the kernels in Figure 1. This figure uncovers substantial variation of wage growth around the kernel, and it is particularly large in the lower and the middle thirds of the

[1] For this reason, we do not provide a detailed description. For the descriptive analysis we have downloaded their do-files and run them on our sample.

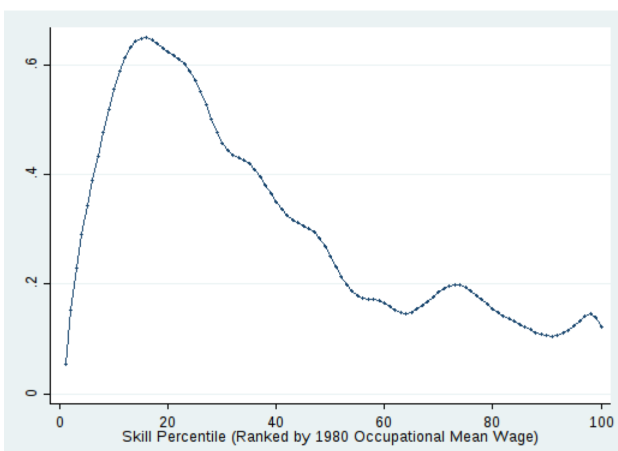
Table 1: Change in Employment Share from 1980 to 2005 by Occupation Group

Occupation Group	100×change in log employment share
Managerial and Professional Specialty Occupations	0.0814
Technical, Sales, and Administrative Support Occupations	8.28×10^{-4}
Service Occupations	0.0901
Precision Production, Craft and Repair Occupations	-0.0429
Operators, Fabricators, and Laborers	-0.104

skill distribution. Importantly, in both parts of the distribution there are many occupations with a wage growth above 20 percentage points. Wage polarization therefore seems to be generated by a larger fraction of occupations experiencing low wage growth − but not by the disappearance of occupations with high wage growth − in the middle third.

The starting point of our descriptive analysis is the view that deviations from the non-parametric relationship between average wages in 1980 and subsequent wage and employment growth contain valuable information about the underlying economic forces shaping their trends. Given the level of aggregation in these figures this seems to be a more reasonable interpretation than viewing them as outliers.

Figure 2: Share of Routine-Intensive Occupations by Skill Percentile



The particular question we are asking is whether there is a systematic difference in the skill content between occupations that fared well and those that did not, *conditional on AD's skill measure*. This is very different from AD's question of why occupations in the middle of the skill distribution did poorly relative to those at the edges, answered by the computerization- and re-routinization-hypothesis. In fact, computerization is not likely to be a satisfying answer to our question, as can be concluded from looking at our Figure 2. This figure shows the share of routine intensive occupations for each skill percentile and replicates Figure 4 in AD. As a first pass, the inverse of this figure tracks the wage growth kernels quite closely indeed, consistent with the computerization hypothesis. But whereas there seems to be a strong correlation between wage growth and routine task content at the edges of their distributions, several issues are apparent. First, while the kernel of routine task-intensity decreases monotonically over the entire middle third of the skill distribution, there does not seem to be a corresponding increase of the share of occupations with large wage growth. This can also be seen in the kernel in Figure 1, which is flat throughout the region. Second, there is significant variation in wage growth throughout the bottom two thirds of the skill distribution. Hence, there are low and high wage growth occupations, no matter the routine task intensity. Third, the task-content of occupations itself demonstrates significant variation over the wage distribution. Many of the occupations in the middle of the skill distribution for example are not ascribed as being routine according to the Autor and Dorn RTI measure. For instance, truck driving is, in terms of employment shares, one of the largest among this group of occupations, and it has a wage growth of close to zero. Yet, under Autor and Dorn's measure, it is *not* a routine occupation.

We conclude that to understand the distribution of wage and employment growth throughout the *entire* skill distribution one needs to go beyond the focus on service occupations and routine task intensity. Our discussion above points out two particularly promising avenues for enriching the analysis. First, the kernel smoothing of wage growth data and the task intensity measure masks important data variation. In particular, while they visionalize a striking polarization of the wage and employment growth distributions it is not true that traditionally middle-class jobs are uniformly loosing ground. We therefore investigate the systematic differences between growing and declining occupations, conditional on their average wage in 1980.

Second, defining the skill content of an occupation by its average wage in 1980 is problematic as well. For routine occupations are often assigned a fairly high skill measure, and yet we do not find it natural to think of them as medium-skill occupations. Conversely, many occupations that are in the middle of the skill distribution and that are defined as routine seem to involve non-trivial skills that may gain from educational interventions. Oftentimes, these are also the occupations that fared quite well in terms of wage growth. For example, auto body repairers are a routine occupation that is at the 41st percentile of the skill distribution. It experienced a wage growth of 19 percent, not much lower than low-skill service occupations. Even more extreme are precision makers, repairers, and smiths, an occupation at the 36th percentile of the skill distribution with a wage growth of 30 percent, and legal assistants and paralegals, just one percentile above, with a wage growth of 50 percent.

3 Complex vs. Simple Occupations

3.1 Overview

To investigate which occupations have experienced wage and employment gains, no matter their average wage in 1980, we start with a deeper investigation of the type of skills that are being employed. To this end we abandon the focus on computerization because we do not view it as an economic primitive, but rather as a reflection of the underlying skills that are required as an input. This is consistent with insights from the literature on computational complexity and computer algorithms, which emphasizes that computers serve to improve on processes that humans can perform as well. Such improvements are possible only to some extent because computers lack the cognitive abilities of humans. Hence, computers do not compete with humans, but rather with some of their skills. These skills are likely to be of the routine type, such as copying and pasting, sorting or packing. We view these tasks as those for which workers are unlikely to become much more productive if they are subjected to some educational intervention, which turns out to be a useful hypothesis when testing the quality of our categorization of occupations below.

We also hypothesize that low skill intensity is a necessary, but not sufficient condition, for computerization. That is, there are both low skill occupations that can be computerized and low skill occupations that cannot. For example, stock and inventory clerks perform tasks that are likely to be increasingly computerized, and we view them as low-skill occupations. Their wage growth was -4 percent. Likewise, motion picture projectionists is an occupation that has nearly died out because it can be computerized. Indeed, it has a wage growth of -32 percent. However, most transporting occupations are not computerized (yet), but we view them as low skill occupations, in contrast to AD. Consistent with our view that skill content, rather than computerization, is the fundamental characteristic explaining the rise or decline of certain occupations, transportation occupations have generally had very low wage growth.

3.2 Defining Occupation Groups via the German BiBB Data

The General Approach to Classifying Occupations. A central deviation of our paper from AD is how we define low- and medium-skill occupations. We follow the methodology in Caines et al. (2015), and rather than using occupation-level average wages in 1980, we sort 3-digit-occupations into islands based on the types of skills that are employed and the intensity with which they are used. Roughly speaking, the low-skill island consists of occupations that do not require the ability to abstract, analyze, making connections, or form decisions. These include occupations with a high routine task intensity, implying that they can be computerized, but also occupations like truck-driving or mail carriers which have not been computerized so far and therefore fall under the non-routine category in AD. The remaining islands include occupations with increasing intensity of the abilities to perform well on more abstract or complex tasks. An interesting example which helps clarifying the difference in our approach to the one in AD is the two occupations "truck drivers" and "airplane pilots." Both occupations have a very similar routine intensity in AD and are viewed as "non-routine." In contrast, according to our view truck driving is a low-skill occupation because it does not require any high-level skills for abstraction and cognitive processing, but airplane pilots are. Interestingly, wage growth for truck drivers was 3.8 percentage points and for airplane pilots was 17.2 percentage points, consistent with our view.

To operationalize this idea, we take the unusual step of relying on German data rather than US data on occupation-specific task usage. The particular data set we are using is called BiBB and has been employed for example by Spitz-Oehner (2006) and Gathmann and Schonberg (2010). There are two major advantages to this approach. First, the BiBB is a worker-level survey that asks employees about the extent to which certain tasks, which we will describe in a moment, are important on their job. Because the data also report the 3-digit occupation a worker is employed in, we can generate occupation-level task usage data from these worker-level data. This is different from the US data, where task usage is not reported by workers and are already aggregated to the occupation level. Second, we can perform a quality check of our categorization by using information about the share of workers in a particular island who hold a formal apprenticeship degree. As highlighted above, we view AD's approach as particularly suitable to study the edges of the distribution of their skill variable. However, it is less successful in explaining the high wage growth in a substantial number of occupations in the middle of this distribution which are viewed as "disappearing." The German apprenticeship system is the largest of its kind in the world and is traditionally viewed as fostering skills in occupations associated with the manufacturing sector − exactly those that are in the "hollowed out" section of Figure 1. We expect that occupations attracting a large share of apprentices are likely to be those lying above the line in Figure 1.

The disadvantage of using the BiBB data is of course that task usage in an occupation may not be the same in Germany and the US and that we may not be able to match occupations in the German and the US data. Both of these issues can be addressed to some extent, as we will do below.

The BiBB Data and the Classification Algorithm. To quantify the skill requirements defining an occupation we rely on the German BIBB data set, a survey of employees on "qualification and working conditions in Germany." The BiBB is a repeated cross-sectional data set, with samples drawn representatively from the working population, including self-employed individuals, in 1979, 1986, 1992, 1999, and 2006. Each of the waves have approximately 35,000 observations on the worker level. We only keep male workers in Western Germany who are born after 1935, who are between 25 and 60 years of age, and are not self-employed. The variable of our interest reports task usage on the job, constructed from surveying workers about the main tasks performed on the job among a list of approx. 20 tasks. Examples of tasks are "equip and operate machines," "repair, renovate, reconstruct," "serve, accommodate," "calculate, keep books," or "employ, manage, organize, coordinate." Unfortunately, task categories are not consistent across waves and actually become coarser in more recent years. Given that it is possible to construct a set of comparable task categories for the first three waves only, we do not use the BiBB data for 1999 and 2006.

Table 2: Aggregated and Detailed Task Groups, BiBB Data, 1979, 1985, 1992

Aggregated	Detailed, time-consistent
Routine Manual	equip and operate machines
	cultivate, excavate
	pack, ship, transport, sort, archive, drive vehicles
	cleaning, laundry, garbage disposal
Non-Routine Manual	manufacture, produce, finish, process, cook, bake
	repair, renovate, reconstruct
	construct, install
	serve, accomodate
	secure, guard
	nurse or treat others
Routine Cognitive	clerical
Non-Routine Cognitive	research, evaluate, measure, test, design, plan, sketch
	calculate, bookkeeping
	programming, IT
	execute laws, interpret rules
	employ, manage personnel, organize, coordinate
Interactive	sell, buy, advertise
	teach or train others
	publish, present, entertain others

Since the BiBB contains the 3-digit occupational code for the job a surveyed individual is currently employed in, we can construct measures of task usage on the occupational level. Our first step adopts the approach in Gathmann and Schonberg (2010) and aggregates the detailed task information to 5 large groups of tasks. These are: manual routine, manual non-routine, cognitive routine, cognitive non-routine, and interactive. Table 2 provides a list of detailed and aggregated tasks. For each of the 3-digit occupations we then calculate the fraction of individuals who report these tasks to be a main part of their job. This readily identifies four main groups of occupations defined by their task inputs: Occupations that predominantly use routine tasks, such as land workers, plastic processors, or packers; occupations that predominantly use non-routine tasks, such as plumbers, motor vehicle repairers, office specialists, or nurses; occupations that are upgrades of other occupations, such as advanced technicians instead of simple technicians, foremen, and management on various levels; and finally occupations that can only be accessed with a college- or university degree because of institutional requirements. For that reason, we define the four islands as "simple," "complex," "advanced/managerial," and "college" and assign every 3-digit occupation to one of these islands.

We think of simple occupations as those that require little specific skills and do not require the ability to abstract, analyze a problem and coming up with a solution, or making decisions. Educational interventions are therefore not very likely to increase worker productivity significantly. This conjecture is supported by the evidence on the relative share of workers with an apprenticeship degree in these occupations, provided below. Tasks in this category, whether manual or cognitive, include those which can be computerized, such as operating machines or performing clerical duties. They also include other simple tasks, like cleaning, which may be harder to computerize. This is a crucial difference to AD. In contrast, non-routine tasks, whether manual or cognitive, require some ability to abstract and process complex information. Workers performing such tasks may gain from educational interventions. Again, we provide evidence for this conjecture below.

In a next step, we sort each of our 3-digit occupations into one of the 4 islands. While advanced- and college-occupations are defined by exogenous characteristics − updates from another occupation or occupational requirements, respectively − the assignment into simple and complex occupations is less clear-cut. We therefore use the following heuristic: First, we sort an occupation into island 1 (2) if more than two-thirds of workers employed in this occupation report the (non-)routine task as the main part of their job. This does not match all occupations to some island. Second, we assign occupations for which at least 3 tasks have a high reporting fraction into the complex occupation, based on the idea that such occupations require the combination of various tasks and are therefore sufficiently complex. Third, occupations that are observed to be mechanized and automatized over time, as reflected by an increase of the routine task over time, are assigned to the simple island. This heuristic assigns all 3-digit occupations to one of our four islands. A major advantage of this algorithm is that it does not require using information on the interactive task, which we find difficult to categorize.

In a final step we match the island assignment on the 3-digit occupational level to the US Census data. This step is not entirely straightforward since the occupational codes for the German and the US data, though quite similar, are not harmonized. We match in both directions on key words. That is, we first look of a 3-digit occupation in the US data that matches closest to a particular 3-digit occupation in the German data, based on the wording of the occupational code. For the unmatched occupations we then to the same in the opposite direction. In the end, there are a number of German 3-digit codes that are matched to multiple 3-digit US codes and vice versa. We will conduct a number of simple tests for the quality of this match below

Testing the Quality of our Categorization: Apprentices and Islands in the German SIAB Data. As with any categorization of occupations into aggregate groups, there are a number of judgement calls to be made that cannot be justified by hard evidence. In our case, the fundamental idea that simple tasks do not require skills to solve complex problems and are therefore unlikely to require educational interventions is vague. An implication of this hypothesis is that workers are not likely to voluntarily sort into costly training programs for low-skill occupations since the returns are likely to be very small. At the same time, employers are less likely to demand specific qualifications for occupations that are simple because they do not increase the productivity of their workers. Conversely, workers in complex occupations may see significant gains from specific training and therefore should be more likely to enroll in apprenticeship programs. We will test this prediction by calculating the shares of workers with an apprenticeship program in each of our islands. To see the logic behind this test, we first offer a brief description of the apprenticeship program in Germany.

Germany's apprenticeship program offers occupation-specific training to high-school graduates together with government-sponsored general education and is currently the largest training program of its kind in the world. An interesting feature of this program is that on-the-job training and its content is highly regulated, with firms requiring certification to be able to hire trainees, and explicitly designed to develop occupation-specific human capital. Indeed, one of the requirement for an employer to be allowed to train a worker is that she guarantees that upon completion of a degree, an apprentice is able to perform the tasks required to be successful in the training occupation. This is evaluated via theoretical and practical tests.

Individuals can apply for apprenticeships after completion of a secondary degree. Programs are designed to provide occupational skills and, depending on the training occupation, take two to three years to completion. They are offered in over 500 occupations, ranging from carpenter, mason, cook or industrial-, electrical- or car-mechanic to nurse, lab technician or financial accountant. Besides training on the job, apprentices are required to visit a government-sponsored school of general education ("Berufsschule") that teaches skills such as mathematics, languages, social sciences, and accounting. Approximately sixty percent of an apprenticeship program takes place on the job and the rest in school. Firms that are interested in hiring an apprentice need to acquire a certification from industry-specific employer associations first. Once certified, employers searching for apprentices post vacancies, commit to providing appropriate training for a particular occupation, and pay an occupation-specific training wage that is negotiated between unions and employer-associations. Standards for on-the-job training that need to be followed by firms are set by employer associations in coordination with the Federal Employment Agency. Individuals with a secondary degree apply to these vacancies and, once accepted, are subject to a probation period. The apprenticeship degree is by far and large the most common educational degree in Germany, with over two-thirds of the German workforce holding one. In contrast, only slightly more than ten percent have a university-or technical college degree.

To calculate the share of apprentices in each island we rely on a second data set, called the SIAB. We use its confidential version, which is a 2%-extract from German administrative social security records for the years 1975 to 2008.[2] The SIAB is representative of the population of workers who are subject to compulsory social insurance contributions or who collect unemployment benefits. This amounts to approximately 80% of the German workforce, excluding self-employed and civil servants. The 3-digit occupation codes in the SIAB and the BiBB are identical since they are collected by the same institution. We can therefore match the island designation for each 3-digit occupation to the SIAB data. Though we do not use these data extensively in this study, it nevertheless is quite useful. First, it is much larger than the BiBB; second, it provides a well-defined education measure which is unlikely to be plagued by measurement error; and third, it includes information on earnings.

Earnings information is valuable because it allows us to document that enrolling in an apprenticeship degree has opportunity costs. Conditional on occupation, apprentices take a significant wage cut relative to workers who enter the labor market directly after completion of high school. This is due to the fact that wages for apprentices are set by collective bargaining and need to be low enough to provide employers the incentive to create vacancies for apprenticeship spots. It is likely that the raw earnings difference we document here is biased downwards relative to the true opportunity costs since it is likely to be the more able individuals who choose to enrol and turn out to be accepted into an apprenticeship program.

Regarding the share of apprentices in each occupation we find that it is much higher in island 2 than in island 1. We view these results as strong evidence in favor of our occupational classification. In particular, the more complex the tasks are, the higher the incentive to acquire more education. Conversely, if an occupation does not require the ability and skills to solve complex tasks, at least to some extent, then there is very little reason for workers to incur the opportunity costs of a program that is designed to increase skills. In equilibrium, employers will not demand formal degrees for these types of occupations because it does not make their workers more productive.

Crosswalks: Islands, Education and Task Inputs in the US Census and DOT Data. We test the quality of our crosswalk from the German to the US occupation codes on the 3-digit

[2]These data are collected by the "Institut fuer Arbeits-und Berufsforschung" (IAB) (Institute for Employment Research) at the German Federal Employment Agency.

Table 3: Average Educational Attainment

	< Grade 12		Grade 12		College	
	1980	2005	1980	2005	1980	2005
Island 1	0.327	0.138	0.457	0.504	0.216	0.358
Island 2	0.206	0.090	0.427	0.394	0.367	0.516
Island 3	0.115	0.035	0.316	0.256	0.569	0.709
Island 4	0.022	0.007	0.083	0.083	0.895	0.910
AD RTI - Nonintensive	0.225	0.088	0.357	0.346	0.418	0.565
AD RTI - Intensive	0.192	0.081	0.464	0.412	0.344	0.507

level using a similar logic as in the last section. In particular, if the importance of solving complex tasks increases in the number of the island, we should expect average education to be the highest in island 4 and the lowest in island 1. Table 3 shows average educational attainment by island designation. In both the 1980 Census and 2005 ACS the share of individuals with some college education is higher within the complex island than in the simple island. Similarly the share of high-school dropouts and individuals with only a grade 12 education is higher in island 1 for both samples. It is interesting to compare this with the average educational attainment observed in AD's classification of routine and non-routine occupations. AD classify an occupation as being routine-intensive if the measured value of its routine task intensity index, which the authors construct from DOT data, is in the top third amongst all occupations. As can be seen in Table 3, consistent with our island designation, the share of college- educated workers is higher in the routine-intensive occupations. However, it is noteworthy that for both the 1980 and 2005 samples the share of college dropouts is also lower amongst the routine-intensive occupations.

Another useful exercise is to correlate our task intensity measures with AD's DOT data. AD follow Autor, Katz, and Kearney (2006) (AKK) in collapsing the DOT task intensity measures of Autor et al. (2003) (ALM) to three dimensions: abstract, manual, and routine. AKK intend for these to roughly correspond to high, medium, and low-skill work, respectively. Table 4 shows the correlation matrix between our task measures derived from the BIBB and the AD DOT measures, weighted by 1980 occupation shares. We argue that AD's division of tasks into abstract, manual, and routine makes it difficult to identify the role that solving complex tasks plays in an occupation. AD's measure of abstract task input is positively correlated with our measures of non-routine cognitive and interactive tasks, and their measure of routine task input negatively correlated with our measures of interactive task. The correlation coefficients between AD's measure of manual task input and our measures of routine manual and non-routine manual task inputs are 0.40 and

0.30 respectively, however we consider these types of tasks to be distinctly different in the types of skills that they require and we hypothesize that wage growth should be different between them. Similarly, the correlation coefficient between AD's routine task measure and our measures of non-routine manual and routine cognitive measures are 0.21 and 0.29 respectively.

Table 4: Correlation between AD and BIBB Task Measures

	Manual		Cognitive		Interactive
	Routine	Non-routine	Routine	Non-routine	
Abstract (AD)	-0.543	-0.291	-0.001	0.559	0.367
Routine (AD)	0.006	0.208	0.286	0.036	-0.474
Manual (AD)	0.400	0.304	-0.362	-0.348	-0.218

The differing method of assigning task usage implies that a number of occupations designed as routine by AD are complex under our island designation and vice-versa. In total, 61 occupations assigned as being routine-intensive by AD are located in our complex island and 35 occupations in our simple island are not designated as being routine-intensive by AD. The average wage growth from 1980 to 2005 is more than twice as large in the former group than in the latter (26.5 percentage points vs. 12.4 percentage points). Among the occupations that AD list as routine-intensive that we do not classify as simple are a number of financial service occupations that are located relatively high in the 1980 wage distribution and that we do not intuitively think of as being routine or simple occupations, including Accountants, Financial Managers, and Real Estate Occupations.

4 Main Results

A major motivation for our reclassification of low- and medium skill occupations was Figure 1, which shows that there is a substantial fraction of occupations above the average relationship between average occupational wages in 1980 and wage growth in the "hollowed-out" part of its distribution. Likewise, there are quite a few occupations with low wage growth at the left end of this relationship. Both of these findings are important deviations from the polarization hypothesis. A natural first step is thus to ask if occupations in island 2 are more likely to be above than below the kernel than occupations in island 1. This should be the case under our hypothesis that the ability to solve complex tasks has been rewarded, no matter the average occupational wages in 1980. Table 5 provides supporting evidence. Only one quarter of occupations in island 1 had a wage growth that was higher than predicted by the non-parametric relationship of wage growth as a function of 1980 average wages. In contrast, almost 50 percent of occupations in island 2 are above this line.

Table 5: Island Statistics

Island	Porportion Above Kernel	Change in Employment Share
1	25.00%	-5.04%
2	48.62%	1.29%
3	66.67%	1.72%
4	55.00%	2.04%

One problem with such aggregate statistics is that it masks heterogeneity across the different regions of the kernel. Given the local nature of the kernel estimate, this can lead to a distorted picture of the fraction of high wage growth occupations in each island. To see this in the extreme, first consider island 4, which is a high-skill occupation when using AD's measure. Yet, the table shows that only 55 percent of occupations in this island have wage growth that is above the prediction line. The reason is that nearly all of these types of occupations are located at the right side of Figure 1, meaning that all of them experienced high wage growth. Furthermore, almost no occupations from other islands are located in this area of AD's occupational skill measure. Local averaging thus takes place only over occupations in island 4, and the share above and below the kernel will be approximately fifty percent by construction.

4.1 1980-2005 wage growth

We therefore estimate the differences in wage growth between islands conditional on occupational skill as measured by the 1980 average wages using parametric regressions. The data used in the regression are on the occupation level, which produces 312 observations. Results are shown in Table 6, where the coefficients on the three island dummies are relative to island 1 (our regressions include an intercept). We start with a a specification that controls linearly for AD's skill measure and that uses the occupation specific employment shares in 1980 as weights. As expected, island 1 contains occupations that have, on average, the lowest wage growth, no matter where they started in 1980. The average difference in wage growth increases monotonically in the number of the island. Wage growth for occupations in island 2 was 11.2 percentage points higher than in island 1. The corresponding estimates for island 3 and island 4 are 15.3 and 18.8 percentage points. All estimates are highly significant.

In the next few columns we increase the order of the polynomial in the skill measure to account for the polarization in wage growth which cannot be captured by a linear specification. We continue finding statistically significant positive wage growth for islands 2 to 4 relative to island 1. The estimated effect for island 2 is particularly robust in magnitude. Somewhat surprising, the point

Table 6: Wage Regression

Dependent Variable: Change in Log Hourly Wages 1980-2005					
Independent Variable	(i)	(ii)	(iii)	(iv)	(v)
Island 2	0.112***	0.0991***	0.107***	0.104***	
	(5.52)	(4.91)	(5.30)	(5.17)	
Island 3	0.153***	0.0857**	0.0782**	0.0854**	
	(4.28)	(2.17)	(2.00)	(2.17)	
Island 4	0.188***	0.138***	0.134***	0.138***	
	(5.06)	(3.54)	(3.48)	(3.57)	
Routine-Intensive				0.0293	0.0363*
				(1.53)	(1.85)
Order of Skill Poly.	1	2	3	3	3
$N = 312$					

$^{*}\, p < 0.1$; $^{**}\, p < 0.05$; $^{***}\, p < 0.01$

estimates for islands 3 and 4 become substantially smaller and, for the former, only marginally significant (10 percent level). Importantly, we find that it is impossible to reject the hypothesis that wage growth in islands 2 and 3 were identical. These findings are not surprising, and should be interpreted with caution. The reason is that the higher the order of the polynomial, the more local our estimates are, and we have the same problems as discussed above. In particular, there are very few cases for which an island 1 and island 4 occupation have a similar skill index, and these cases may be outliers or misallocations of a 3 digit occupation into a particular island. In contrast, there are many cases in which islands 1 and 2 are ranked closely in terms of AD's skill measure, which explains that the point estimate for island 2 is rarely affected by the order of the polynomial in occupational skill. At the same time, it may not be very surprising that wage growth in islands 2 and 3 are similar because we defined the latter to be an update of the former. Moving to island 3 may therefore come with a one-time wage increase, but not a permanently higher wage growth.

One concern with this specification is that we may identify the positive and highly robust coefficients from occupations at the left end of Figure 1, that is low-skill occupations by AD's measure that saw particularly large wage gains. We therefore reestimate our regression on various subsamples. We start splitting the sample by terciles of the 1980 skill distribution. The results for the linear specification are shown in Table 7. Again we find that occupations in island 2 experiences higher wage growth than occupations in island 1. These differences are significant on the 1 percent level for the bottom and the middle tercile of the 1980 occupational wage distribution. Interestingly, the effect is particularly pronounced in the middle part, that is the "hollowed-out", section. Hence, the occupations above the kernel in middle of Figure 1 are indeed occupations we view as utilizing complex tasks, as expected.[3]

Table 7: Wage Regression

Dependent Variable: Change in Log Hourly Wages 1980-2005			
Independent Variable	Bottom Skill Tercile	Middle Skill Tercile	Top Skill Tercile
Island 2	0.0870***	0.129***	0.0727
	(3.30)	(3.70)	(1.21)
Island 3	-0.0659	0.164	0.0561
	(-0.09)	(0.90)	(0.87)
Island 4	0.145	0.0456	0.133**
	(0.39)	(0.35)	(2.08)
Order of Skill Poly.	1	1	1
N	101	95	116

$^{*}\, p < 0.1$; $^{**}\, p < 0.05$; $^{***}\, p < 0.01$

The coefficient estimates on occupations in the other two islands are insignificant for the two lower terciles, and the estimate for island 4 is marginally significant for the highest tercile. This is not surprising since there are very few occupations for these two islands in the lower two terciles and almost only such islands in the upper tercile. As a consequence, there is not sufficient statistical power to estimate these effects precisely.

For further robustness we also estimate the models on different aggregate groups of occupations. These results are shown for a linear function of the skill index in Table A-1 in the appendix. We find significant estimates on the island 2 dummy for technical, sales and administrative support occupations, for service occupations, and for operators, fabricators and laborers. The estimates for a sample of managerial and professional specialty occupations and for precision craft and repair occupations are insignificant and point in the wrong direction. These two groups have very little overlap with other islands. Specifically, the first group is comprised almost exclusively of occupations in island 4, and the second group contains almost by definition almost only occupations from

[3]These results are highly robust to inclusion of higher-order polynomials.

island 2. There are therefore no valid comparison groups to estimate the differences in wage growth across islands within these broadly defined occupation groups. For similar reasons, we are not able to estimate precisely and effects for islands 3 and 4.

To increase precision, we aggregate these groups further. The next column shows results for a sample merging the high-skill occupations, that is managerial, professional and technical occupations. Without including a control for the AD skill index we get significant estimates for all three dummy variables. However, once we control for the skill index only the island 2 effect remains significant. This is likely driven by the sales and administrative support occupations which have occupations in island 1 and island 2. It is clear from these results that such types of occupations have experienced larger wage gains if they required some complex tasks.

Next, we consider a group of "middle-class" occupations, that is the type of occupations that are thought to have lost ground and are positioned in the middle, hollowed-out section of Figure 1. These are occupations related to production, repairing, operating and fabricating. As hypothesized, we find a highly significant estimate for occupations in island 1 (1 percent), albeit somewhat smaller than the corresponding point estimates in other occupation groups. The coefficient on island 3 is positive and twice as large. Furthermore, since there are a number of these occupations in island 3, we have enough precision to attain significance on the 10 percent level. These results are one of the main findings from our paper. It is on average true that traditionally middle-class jobs have lost ground over the last three decades. However, occupations in this part of the skill distribution that rely on complex tasks have still kept up with other type of jobs in terms of wage growth.

Another interesting conclusion one can draw from this table is that the focus on service occupations is indeed too narrow to understand the wage growth distribution. The results from column 3 show that the distinction between jobs that require complex tasks and those that only require simple tasks is particularly important for this group of occupations. In particular, whether service occupations require simple or more complex tasks has large explanatory power for differences in wage growth. In fact, there was a large variation of wage growth within low-skill service occupations. It was the low-skill service occupations that rely partially on complex tasks which experienced the largest wage gains. Whether an occupation can be computerized or not can therefore not be the whole explanation. Skills being employed play a crucial role as well.

It is worth discussing a number of key implications from this empirical analysis. First, in the aggregate wage growth is monotonically increasing in the number of the island. However, in our disaggregated wage-growth regressions we were not able to estimate any significant effects for the third and fourth island once we control for the AD skill index. The reason is that such occupations are concentrated at the high end of the skill distribution where island 1 and island 2 are essentially absent. It is therefore not possible to compare wage growth between all islands conditional on the skill index. We do not think of this as a large problem, for the following reason. On the one hand, we view AD's theoretical and empirical framework to be particularly adequate for occupations on the right end of Figure 1, which are predominantly occupations in islands 3 and 4. On the other hand, their framework has substantial limitations for the middle part and, as we have found, also for the lower end of Figure 1. For it is here where their routinization index cannot account for the large variation of wage growth among occupations in similar skill percentiles. As we have shown, the distinction between complex and simple tasks is a major factor as well, and this applies to low-skill service occupations as well as manufacturing-related occupations.

Second, it is true that among island 1 occupations it is the low-skill service occupations, such as cleaning, that fared the best. This however can potentially be explained by an interaction between changing demand for household services and labor market frictions, such as compensating wage differentials or discrimination. It therefore seems that once one conditions on the skill content of an occupation in the sense that it requires solving complex tasks or not, the concept of computerization becomes obsolete.

4.2 1980-2005 employment growth

To complete our descriptive analysis, we also estimate our regressions using employment rather than wage growth as a dependent variable. Table 5 shows that employment in island 1 has declined substantially, while it has increased in all other islands. Furthermore, the relationship between island number and employment growth is monotonic. It is natural to ask whether this trend holds within skill percentiles or whether reallocation across islands occurs across the skill distribution. The results from a regression of employment shares (as measure by share of total hours worked) on island dummies and a skill polynomial are shown in Table 8. Generally we find that at the 10 percent significance level island 2 and 3 occupations grew in employment share relative to island 1 occupations, conditional on the skill index. The effect is not effect by controlling for AD's indicator of routine-intensive occupations. Hence, within a skill percentile as defined by AD the share of employment in island 2 grew relative to island 1 concurrent with the wage changes documented earlier.

5 Conclusion

In this paper we document the role played by complex skills – those that govern the ability of abstract, analyze, make connections, or form decisions – in explaining polarization in the labor market over the past 30 years. Previous literature has focused on the role of technical change, most often via the computerization and routinization of occupations, in dampening wage and employment growth in the middle of the skill distribution. We argue that such an approach abstracts from substantial variation observed in the middle of the wage distribution and show that, throughout the 1980 wage distribution, occupations that require the performance of complex skills experienced significantly higher wage and employment growth from 1980 to 2005. Therefore, a relative increase

Table 8: Employment Regression

Dependent Variable: Change in Employment Share 1980-2005					
Independent Variable	(i)	(ii)	(iii)	(iv)	(v)
Island 2	0.000694*	0.000697*	0.000634*	0.000668*	
	(1.87)	(1.88)	(1.69)	(1.78)	
Island 3	0.00126*	0.00125*	0.00116*	0.00119*	
	(1.91)	(1.89)	(1.74)	(1.77)	
Island 4	0.00101*	0.000956	0.00100	0.000911	
	(1.73)	(1.57)	(1.65)	(1.48)	
Routine-Intensive				-0.000166	-0.000323
				(-0.48)	(-0.95)
Order of Skill Poly.	1	2	3	2	2
$N = 321$					

$^{*}\, p < 0.1$; $^{**}\, p < 0.05$; $^{***}\, p < 0.01$

in demand for complex skills, in other words a complex-skill based technical change, is a plausible explanation for wage and employment growth patterns observed over the period.

References

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

Autor, D. H., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal Economics 118*(4), 1279–1333.

Caines, C., F. Hoffmann, and G. Kambourov (2015). Training programs, skills, and human capital: A life-cycle approach. Mimeo.

Gathmann, C. and U. Schonberg (2010). How general is human capital? A task-based approach. *Journal of Labor Economics 28*(1), 1–49.

APPENDIX

A Additional Tables

Table A-1: Wage Regression

Dependent Variable: Change in Log Hourly Wages 1980-2005

Indep. Variable	Occ. Group A	Occ. Group B	Occ. Group C	Occ. Group E	Occ. Group F	Occ. Group A & B	Occ. Group E & F	Occ. Group A, B, E, F
Island 2	-0.0628	0.0973**	0.134***	-0.0121	0.0465**	0.0941***	0.0472***	0.111***
	(-0.63)	(2.63)	(4.65)	(-0.28)	(2.61)	(3.04)	(3.26)	(4.87)
Island 3	-0.174*	-0.0507	0	0.0381	-0.0147	-0.00831	0.112***	0.155***
	(-1.74)	(-0.48)	(.)	(0.68)	(-0.04)	(-0.19)	(2.99)	(3.98)
Island 4	-0.163	0.0275	0.157*	0	0	0.0107	0	0.190***
	(-1.64)	(0.10)	(1.76)	(.)	(.)	(0.24)	(.)	(4.69)
Order of Skill Poly.	1	1	1	1	1	1	1	1
N	88	67	33	65	59	155	124	279

$^{*}\ p < 0.1$; $^{**}\ p < 0.05$; $^{***}\ p < 0.01$

Occupational Groups	
A	Managerial & Professional Speciality Occupations
B	Technical, Sales and Administrative Support Occupations
C	Service Occupations
E	Precision Production, Craft and Repair Occupations
F	Operators, Fabricators and Laborers