Innovation on Tools and the Rise of Skill Premium *

Hyejin Park[†], Younghun Shim[‡]

October 9, 2021

Preliminary and Incomplete

Abstract

This paper measures innovation on tools used by different occupations and studies its impact on the increasing skill premium. First, we match the description of tools from Wikipedia with patent text data using textual analysis to measure the innovation on tools. Then, we study its relation with the labor market variables at the occupation level. We find 1) innovation on tools grew more in skill-intensive occupations. 2) it is positively associated with wage and employment growth across occupations. 3) it is positively correlated with the skill premium and skill intensity growth within each occupation. Motivated by this reduced-form evidence, we build a model where tool innovation increases the demand of occupations, potentially more for skilled workers. Parameters are estimated through the Generalized Method of Moments. We find that tool innovation accounts for 61% of the total demand factor that contributed to the skill premium increase in 1980-2015.

Keywords: Skill-biased technical change, Skill Premium, Innovation, Text Analysis of Patents

JEL codes: J24, J31, O33, O47

*We thank Ufuk Akcigit, Erik Hurst, Simon Mongey, Josh Morris-Levenson for their helpful comments. We also thank all the participants in AMT seminar. All errors are ours.

[†]University of Chicago; Email Address: hjpark0619@uchicago.edu

[‡]University of Chicago; Email Address: yhshim@uchicago.edu

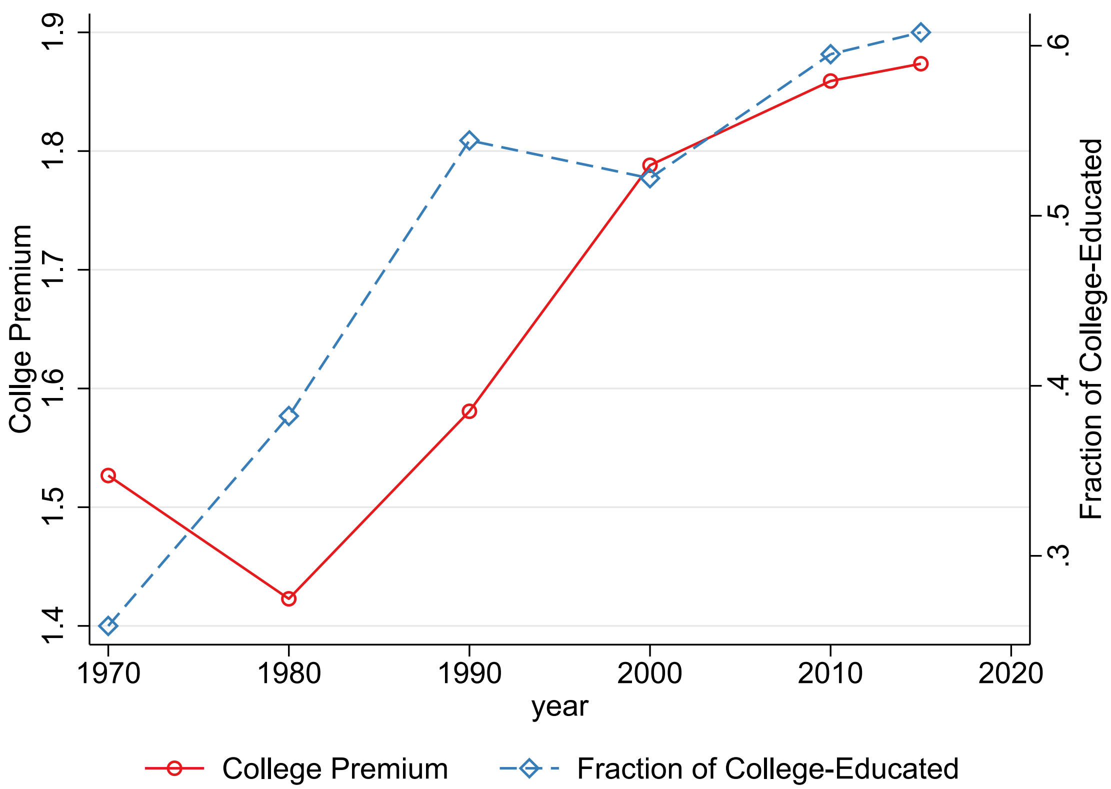
1 Introduction

Economic growth has been associated with uneven benefits to workers. Skill premium, defined by the mean wage difference between skilled and unskilled labor, has grown steadily over time. At the same time, the fraction of blue-collar workers has decreased while the shares of service and office workers have increased substantially since the 1980s. Such worker heterogeneities have been studied to evaluate the need for policy intervention such as worker retraining programs or income reallocation schemes.

Meanwhile, economists have considered technological changes in the demand side of the labor market as the driver of the uneven incidences of economic growth (See, for example, Hornstein et al. (2005)). This is because, in almost all cases, the skill group of workers that is favored by the wage growth also experienced employment growth. Figure 1 shows that there have been increases in skill premium as well as in the share of college-educated workers since 1980. Moreover, rapid improvements of capital goods and modernization of the production process have affected the labor market during the same period. Most views hypothesize that the complementarity of new technologies is heterogeneous across worker groups depending on their education level or occupations (e.g., Nelson and Phelps (1966), Krusell et al. (2000), Autor et al. (2003)). However, the lack of measures for technical change over the heterogeneous workers makes it difficult to find direct evidence of the hypothesis.

This paper constructs a direct measure of technical change on tools used by different occupations and studies its effect on the increase in the skill premium. We use the list of tools used in each occupation constructed in O*NET. These tools are expected to be complementary to labor, but the complementarity is confirmed by data. Then, we measure the technical change on each tool using patent data. Since patents are not naturally connected with tools or any other commodity codes, we use textual

Figure 1: College premium and share of college graduates



analysis to match patents to tools. To be specific, we match the most relevant tools for each patent by calculating the similarity between the abstract of patents and the Wikipedia page of the tools. Patents that are matched to each tool are used as the measure of innovation on the tool. The measure is then aggregated at the occupation level and connected to the labor market variables such as wage, employment, and skill premium.

Our tool innovation measure has two strengths. First, it directly measures technological developments while previous studies often rely on proxies to measure technological changes. Krusell et al. (2000) and Caunedo et al. (2021) use changes in the price of capital goods as a proxy for technical changes. While this approach was successful to explain the observed changes in skill premium, it is not clear what this indirect measure captures. The price of capital goods might be affected by various supply and demand factors other than technical changes. On the other hand, this paper uses patent data that are commonly used in innovation literature to measure technical changes (e.g., Akcigit et al. (2017)). Second, it covers universal tools with the entire occupations in the economy. Although many papers study important technology such as computers (Autor et al., 2003; Burstein et al., 2019) or robots (Acemoglu and Restrepo, 2020), this paper covers expansive technology, which allows cross-sectional variation across a broader set of occupations.

With the technology measure and the labor market variables at the occupation level, we first provide reduced-form results. We show that tool innovation grew more in skill-intensive occupations. Then, we find that tool innovation is positively associated with the change of wage, employment, skill premium, and skill intensity at the occupation level by running long-difference OLS. This suggests that tool innovation increases the demand for occupation and favors skilled workers. Since the unobserved demand factor of occupation might be correlated with the tool innovation, there is a potential endogeneity problem in the OLS. We constructed an instrumental variable by collecting "upstream" patents that are not in the same industry sector but induce knowledge spillover to a given patent class[1]. The results from IV are qualitatively the same as the OLS.

Next, to quantify the effect of tool innovation on skill premium in a general equilibrium setting, we build a simple aggregate production function with occupational inputs

[1]For example, we use patents in electricity as IV for patents in the IT sector. We assume that innovation in electricity is correlated with innovation in IT as there is knowledge spillover from electricity to IT, but it is not correlated with other demand shocks of IT workers. Knowledge spillover is estimated by the citation network of patents as in Cai and Li (2019)

and estimate parameters of the production function. A representative firm has an aggregate production function that aggregates tasks done by occupations, and skilled and unskilled workers are imperfect substitutes in each occupation. Motivated by the reduced form evidence, we assume that tool innovation increases labor productivity. We also allow the complementarity with the productivity of tools to be different for skilled and unskilled labor. Based on the model, we estimate structural parameters using the Generalized Method of Moments. We then run a counterfactual exercise to identify the quantitative importance of the tool innovation channel. We find that the tool innovation alone can explain 61% of the total demand factor that contributed to the skill premium increase. The remaining 39% is an unobserved demand factor that favors skilled workers.

1.1 Related Literature

We first contribute to the long debate on the rising skill premium. Skill premium, often represented by the wage premium of college-educated relative to non-college-educated, has steadily risen since the 1980s in most developed and developing countries. Many previous studies attribute the increase in the skill premium to the skill-biased technical change (Acemoglu, 2002b; Violante, 2008). Their arguments rely on the complementarity of capital equipment or computer technology to the skilled labor as in Greenwood and Yorukoglu (1997) and Krusell et al. (2000). Those studies often use time-series fluctuations on skill premium and technology investments to infer the source of increases in skill premium. We, on the other hand, investigate occupation-specific developments on tools represented by the number of patents and how these technological changes affect the wage, employment, and skill premium at the occupation level. This cross-sectional variation makes it possible to control secular trends in demand for skilled labor that do not necessarily result from changes in production equipment, such as changes of firm or trade structures.

This paper is one of the first papers that estimate the aggregate production function with occupational inputs. The structure is comparable to the task-based approaches which became increasingly popular after the 2000s. Since the seminal work by Autor et al. (2003), the unit of analysis for the impact of technical changes on the labor market has been a task, which is often categorized as routine, cognitive, abstract, or manual. Technical changes in computerization or robotization are regarded as increases in the capital that substitutes labor inputs in cognitive and manual tasks. These task-based approaches offer a powerful framework for the analysis of labor-substituting technologies both empirically and theoretically (Autor and Dorn, 2013; Acemoglu and Restrepo, 2018; Cortes et al., 2017). We contrarily focus on labor-augmenting technologies that are potentially be more skill-biased, and the unit of analysis is occupation-specific tasks. Occupation is a more informative unit of analysis in this case because of variations in tools used in each occupation. As long as some tools have more technical changes than others and those tools are used by only a subset of occupations, the differences in the wage or employment changes can be regressed on those innovations on tools even when both occupations have non-routine and abstract tasks.

This paper is related to the literature on the quantitative occupation-choice model. Lee and Wolpin (2006) are one of the earlier studies that incorporated the occupation-choice model to quantify the contribution of technology growth in the service sector, as opposed to the decrease in mobility costs, on the employment growth of the service sector. A recent study by Hsieh et al. (2019) analyzes the occupation choice of workers heterogeneous of race and gender to analyze the impact of reallocation of worker talent across different occupations over time. This paper also addresses the issue of workers allocation over different occupations. This paper emphasizes the technological origin of labor reallocation from the 1980s to the 2010s. Because of technological developments being heterogeneous across occupations, skilled workers in an occupation are reallocated from other occupations. Moreover, tool innovation is more complementary to skilled labor than to unskilled labor. As a result, with more innovation on tools over time, unskilled labor is substituted more with skilled labor within occupations. Cortes et al. (2017) also consider the technical change as a source of disappearing routine jobs, but they do not measure and use technological changes directly in their analysis. Burstein et al. (2019) focuses on computer usage at workplaces. They use occupation-level variations in fractions of workers using computers on the job to quantify the importance of skilled workers' comparative advantage in computer usage for analysis of rises in skill premium since the 1980s. This paper carefully constructs the measure of innovation on tools using patent data and uses the occupational differences in the number of patents matched to each occupation.

Lastly, this paper is related to a growing literature that applies textual analysis to patent data. Kelly et al. (2018) calculate text similarity between patents and identify breakthrough innovations which are distinct from the previous inventions but similar to the following inventions. Argente et al. (2020) match patents with product data from Nielson, by calculating text similarity between patents and product descriptions. Zhestkova (2020) matches patents with Wikipedia articles to cluster technology sectors by using similar textual analysis techniques. Bloom et al. (2021) match patent with an earnings call and job posting to study the diffusion of disruptive technology over time. Webb (2019) would be the most relevant papers with ours where he matches patent with the occupation's task description to measure the exposure to automation. He finds that the occupations with higher exposure to automation experienced decreasing wages and employment. While he studies technology that substitutes labor, we focus on the technology of tools that are complement labor.

The remainder of the paper is organized as follows. Section 2 explains the data used.

Section 3 shows reduced-form evidence. Section 4 presents the result from structural estimation. Section 5 runs counterfactual exercises. Section 6 concludes.

2 Data

2.1 Overview

Our goal is to measure innovation on tools and study its impact on occupation-level outputs such as employment and wage. We measure innovation from patent data. One big challenge is how to match patents to different tools. We calculate text similarity between tool and patent to overcome this challenge. By classifying the patents into different tools, we quantify the innovation on each tool. Then, we aggregate this measure of innovation into occupation level by using the tool - occupation data from O*NET. Lastly, we connect this with occupation level variables such as wage, employment, college premium from Census data. Figure 2 summarizes the whole procedure.

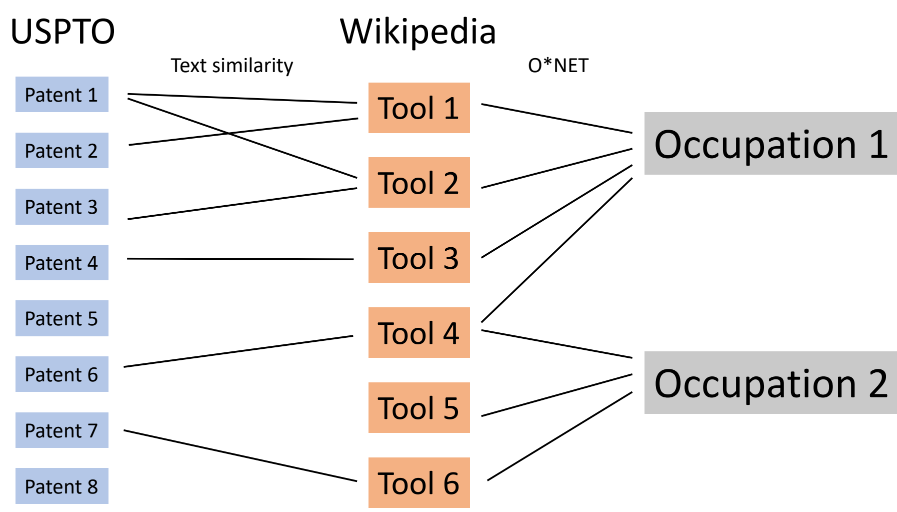
2.2 Innovation on Tools

Occupation - tool data Tool data is from "tools used" data in O*NET[2], where we can see a list of tools used by different occupations. O*NET collects data about tools such as machines or equipment that are used by workers and essential to perform their occupation roles (Dierdorff et al., 2006). For example, aerospace engineers use tools such as lasers, atomic force microscopes, and construction laborers use asphalt saws, electric drills. We have 775 occupations where each of them has 39 tools on average[3]. There are 4,180 unique tools in the data. Tools have their title and United Nations

[2]We use the version of 25.0 which is updated in August 2020.

[3]Its median is 29 and the standard deviation is 36.4.

Figure 2: Summary of the data construction process



Standard Products and Services Code (UNSPSC).

Tool description data To calculate text similarity between tools and patents, we need a sufficiently long description of them. We have an abstract and full article for the patent but we only have a short title for the tool. We decide to use the Wikipedia page of tools as the description of the tool. Wikipedia has broad coverage of products, and its articles usually include a technical description, which makes it easy to match with patents. We search the title of tools using Wikipedia API[4] and download the entire text of the corresponding article. Among 4,180 tools, we could find Wikipedia pages for 1,825 tools.

Patent data To measure the innovation, we use patent data from the United States

[4]'wikipediaapi' package in Python can be found in https://pypi.org/project/wikipedia/. We downloaded the data in 02/28/2021.

Patent and Trademark Office (USPTO)[5]. It has the entire patent data registered in the U.S. We use application year, technology classes, type of patents, title, and abstract of patents. We use application year instead of grant year for the analysis since the application year is closer to the actual innovation year. We restrict our samples to utility patents and exclude design patents to focus on technological improvement. In the end, we have 6.1 million utility patents in 1970-2015.

2.3 Matching patent with tool

Our main goal is to find corresponding tools for each patent. We use natural language processing to calculate the similarity between Wikipedia articles of tools and patent text following the literature such as Argente et al. (2020) and Zhestkova (2020). The similarity score is calculated by counting the proportion of overlapped words between two texts.

Before matching the two texts, we follow the common procedure in natural language processing literature to clean the texts. First, we remove "Stopwords". "Stopwords" are the most common words in English and do not have important meanings. For example, "is", "where", "have" are classified as "stopwords". We remove them to avoid matching two texts just because they share a lot of the function words but do not share meaningful words. Then, we lemmatize words to convert words into their standard form[6]. For example, we change "generating" or "generated" to "generate". Lemmatizing helps us to match words that have the same meaning but in different forms. We repeat the title of the patent and the Wikipedia page 3 times as they summarize important information. The description of the IPC class at 4 digit level is also included 3 times. As most of the patent abstracts are shorter than 5,000 characters,

[5]The bulk file is downloaded through https://patentsview.org
[6]We use the NLTK package in python. https://pypi.org/project/nltk/

we truncate both Wikipedia and patent texts by 5,000 characters in length.

Next, we calculate the pairwise similarity between patents and tools. We transform the cleaned texts into two-word combinations which are called "bigrams". For example, there are bigrams such as "combustion engine", "air-fuel" in the oxygen sensor patent. Following Bloom et al. (2021), we use bigrams instead of single words since they have more clear meaning than single words[7]. For instance, "ratio" appears in a lot of documents and does not have a clear meaning while "air-fuel ratio" has a more clear meaning.

Then, we vectorize each text and compute cosine similarity. This cosine similarity represents the share of overlapped bigrams between two texts. We also consider the fact that the importance of words would be smaller if they are used commonly. We use the term frequency-inverse document frequency (TF-IDF) to appropriately weigh words. ω_{ij} which is the weight of words i in document j is as below.

$$\begin{aligned} \omega_{ij} &= TF_{ij} \cdot IDF_i \\ TF_{ij} &= \frac{f_{ij}}{\sum_i f_{ij}} \\ IDF_i &= \log(\frac{J}{\sum_j \mathbb{1}\{i \in j\}}) \end{aligned} \tag{1}$$

J is the number of total documents. Therefore, IDF_{ij} is higher when the bigram appears frequently in the document but is lower when it appears in other documents as well. This transformation helps us to match two texts that have meaningful common words. The final similarity is between 0 to 1 by construction. If the score is 0, there are no common words and if the score is 1, the two texts are identical.

We assign patents to tools using the computed similarity. Some innovations might not

[7] Our results are qualitatively similar when we use single words instead of bigrams.

be relevant to any of the tools in the data and some innovations might be relevant to many tools. Therefore, we allow multiple matching or non-matching depending on the similarity score. We keep at most 5 tools for each patent and keep the matching if the similarity score is higher than 0.025[8]. As a result, 27% of patents are matched with at least one tool. Table 1 shows the summary statistics of patents for each tool. Over time, more and more patents are matched to tools because more patents applications are made. The matching rates are stable over time. Example 1 shows an example of sample paragraphs of matched patents and tools. Blue words are the common bigrams in both texts. The text-matching algorithm succeeds in identifying a patent that improves the efficiency of the oxygen sensor.

Table 1: The number of patent matched to each tool

	Mean	Sd	Median	1Q.	3Q.	N.	Matching rate (%)
Patent (1970s)	39.53	94.94	7.92	2.00	30.65	1,802	23.83%
Patent (1980s)	81.93	190.84	17.18	4.23	66.00	1,802	23.87%
Patent (1990s)	152.86	410.81	30.70	8.67	115.23	1,802	23.49%
Patent (2000s)	264.11	806.38	43.90	13.67	175.75	1,802	23.00%

Notes: Matching rate is the number of matched patent divided by the number of total patents in a given period

[8]It is same as Argente et al. (2020). We conducted the same exercises with flexible thresholds but the result roughly stays the same.

Example 1

Patent: System and method for detecting deterioration of oxygen sensor

feedback type air-fuel ratio control system control air-fuel ratio air-fuel mixture fed internal combustion engine accordance information signal issued first oxygen sensor installed exhaust line engine exhaust line catalytic converter position downstream first oxygen sensor provided system control system detects deterioration first oxygen sensor

Wikipedia: Oxygen sensor

oxygen sensor lambda sensor lambda refers air-fuel equivalence ratio usually denoted electronic device measure proportion oxygen gas liquid analysed common application measure exhaust gas concentration oxygen internal combustion engine automobile vehicle order calculate required dynamically adjust air-fuel ratio catalytic converter work optimally.

Next, we aggregate the measure of innovation on tools at the occupation level. One occupation often uses multiple tools. We calculate the average number of patents for each occupation and consider it as occupation-level innovation on tools. To be specific, we sum the number of patents within the occupation and divide by the number of tools that have Wikipedia articles. Table 2 shows an example where the innovation on tools for the baker is (15+10+10+5+5)/5=9.

As a result, we construct innovation on tools for each occupation. Table 3 shows summary statistics of the matching. The occupations without any matched patent is excluded from the analysis below. As a result, only 324 occupations are considered in the regression. The 'Skilled' row means the group of occupations that have skill intensity larger than the median in 1980. The 'Unskilled' row consists of all the other occupations. Skill-intensive occupations have more innovations on tools.

Table 2: Example of counting patents at occupation level

Occupation	Tools	Wikipedia found	Patents
Bakers	Automatic dough presses	Yes	15
Bakers	Bagel dividers	Yes	10
Bakers	Baking frames	Yes	10
Bakers	Baking sheets	Yes	5
Bakers	Balance scales	Yes	5
Bakers	Barquette molds	No	.

Table 3: The number of patents matched to each occupation

	Mean	Sd	Median	1Q.	3Q.	N.
Skilled	3.60	0.67	3.55	3.07	4.06	162
Unskilled	2.88	0.47	2.73	2.56	3.17	162
Total	3.24	0.68	3.09	2.71	3.66	324

Notes: 'Skilled' row is the set of occupations that have skill intensity larger than the median in 1980. 'Unskilled' row contains all the others.

The top five occupations in terms of patent growth are Legislators, Film and Video Editors, Web Developers, First-Line Supervisors of Non-Retail Sales Workers, Desktop Publishers. The bottom 5 occupations are Tailors, Textile Knitting and Weaving Machine Setters, Fabric Menders, Sewing Machine Operators, and Electrical Drafters.

2.4 Labor Market Variables

2.4.1 Occupation-level Information

Data from the Census Bureau is used to construct mean wage and employment level by occupation, year, and skill group. We use the Decennial Census 1980 and the American Community Survey (ACS) from 2015 to 2019[9] for observations in 1980 and 2015, respectively. Mean wage is measured by the average weekly wage earnings and

[9]We use the ACS samples from multiple surveys to increase the size of the samples used in each occupation and skilled labor cell.

computed as the annual labor income divided by the number of weeks worked last year. Employment size at the occupation level is the number of people with the occupation code last year. Only samples with 40 weeks of work or more in the previous year are considered in constructing the average. Each observation is weighted by the individual sampling weight offered by the Census Bureau. Skilled labor is defined by workers who experienced some post-secondary education. The occupation codes are harmonized using the OCC1990 variable provided by the Integrated Public Use Microdata Series (IPUMS) and are switched to the Standard Occupational Classification Code (SOC Code) using correspondence between the OCC1990 and the SOC Code variables in the ACS 2010-2019[10].

The top two rows of Table 4 summarize growth rates of employment and mean wage. across different occupations. When calculating the statistics, each occupation is weighted by its wage bill share in 1980. An average occupation had a 77 percent increase in employment size, and the mean nominal wage almost tripled. However, occupations were heterogeneous in terms of their growth. The bottom two rows of Table 4 show how college premium and skill intensity have changed in the period at the occupation level. The skill premium is defined as the ratio between the mean wage of workers with some college education and the mean wage of workers without any college education, and the skill intensity is the fraction of workers with some college education in the total employment at the occupation level. On average, college premium increased by 9.5 percentage points, and skill intensity increased by 22 percentage points, but the change is heterogeneous across occupations as well. This increase in within-occupation skill premium alone can account for 37% (9.5/26) of the entire increase in skill premium

[10]In a robustness check with within-occupation industry variations, we further decompose the employment size, wage level, and college premium at the occupation level into occupation-by-industry level using the NAICS industry code. For the Decennial census, the NAICS was not reported. As a result, we construct a similar crosswalk mapping between the IND1990 and the NAICS codes using the ACS 2010-2019.

during that period. Some occupations do not have unskilled workers. Thus, one occupation out of 324 has missing values of within-college premium and skill intensity of one.

Table 4: Summary Statistics of Occupation-level Changes

	Mean	Sd	Median	1Q.	3Q.	N.
Employment Growth	0.77	8.38	0.46	-0.13	1.13	324
Wage Growth	2.71	1.03	2.55	2.11	3.22	324
D. College Premium (ppt)	9.50	14.17	8.99	2.96	12.69	323
D. Skill Intensity (ppt)	22.10	6.12	22.23	19.47	26.50	324

Table 5 shows the employment and wage growth across skilled and unskilled occupations at the occupation level[11]. The employment of skilled labor grew while the employment of unskilled labor decreased on average. Moreover, the mean wage of skilled labor increased by more than the mean wage of unskilled labor. Given that the relative wage increased more in favor of skilled labor as well as the relative employment, the demand factor is likely to play a key role in the rise in the skill premium in the 1980-2015 period.

Table 5: Summary Statistics of changes in 1980-2015

	Mean	Sd	Median	1Q.	3Q.	N.
Skilled Employemnt Growth	0.74	1.12	0.47	0.26	0.86	324
Unskilled Employment Growth	-0.34	0.48	-0.49	-0.56	-0.31	324
Skilled Wage Growth	2.67	0.95	2.57	2.18	3.11	324
Unskilled Wage Growth	2.41	0.87	2.36	1.92	2.74	323

[11]Each occupation is weighted by the wage bill share in 1980.

2.4.2 College Graduates Major Information

Later in the structural estimation, an occupation-specific supply shifter is needed to estimate the demand elasticity. We construct a Bartik-type supply shock as an instrumental variable for the demand curve estimation. The composition of college majors in each occupation code comes from the American Community Survey (ACS) from 2010-2019, which asks the major of study for college-educated as well as their current occupation. We use only workers younger than 40 to construct the composition. Then, using the number of Bachelor's degrees conferred by postsecondary institutions in the Digest of Education Statistics (DES), we construct the growth rates of college graduates in each major[12] between 1980 and 2013[13]. Finally, the occupation-specific supply shifter, z^2_{o80-13} is constructed by the following equation.

$$z^2_{o80-13} = \sum_m \omega_{om} \left(\log n_{m,13} - \log n_{m,80} \right) ,$$

where ω_{om} is the share of college major m in skilled labor employment of occupation o, and $\sum_m \omega_{om} = 1$ and $n_{m,t}$ is the number of college graduates with major m in year t (80 for 1980 and 13 for 2013). For some college majors that did not have any graduates in 1980, this measure is undefined thus excluded.

Tables 6 and 7 show the majors and the occupations that had the largest and smallest increases in supply. The number of college graduates in transportation and materials moving increases the most between 1980 and 2013. This is followed by parks, recreation, leisure, and fitness studies and legal professions and studies. Library science, education, agriculture record the smallest increases in graduates among college majors. In Table 7, occupations in healthcare and protective services have the largest supply shocks predicted by the composition of college graduates' majors within each occupation in the

[12] We only consider the first Bachelor's degree.

[13] We manually matched the major codes in the ACS and the DES. Appendix C shows the matching.

ACS, followed by protective and material moving occupations that include commercial pilots and ship and flight engineers. College graduates with the same major end up working in various occupations. As a result, the occupations with the highest supply shock measures do not correspond to the majors with the largest increase in college graduates.

3 Reduced Form Evidence

3.1 OLS result

In this section, we present reduced-form evidence that innovation on tools is associated with a larger increase in skill premium at the occupation level. In particular, we show that 1) tool innovation has been concentrated in skill-intensive jobs, 2) is positively associated with wage and employment, which indicates tool innovation increases the demand of occupations, 3) is positively correlated with changes in skill premium and skill intensity, which implies that it increases the demand for skilled workers relative to the demand for unskilled workers. We set 1980 as our initial year and obtain a correlation between the growth rate of patents in 1980-2015 with other variables including wage, employment, and the skill premium.

Figure 3 shows that tool innovation is biased toward skill-intensive occupations. Skill intensity in the figure is the share of college-graduate (or above) workers in 1980. The patent growth is defined as $\log(\text{Patent}_{o,1980-2015}) - \log(\text{Patent}_{o,1970-1980})$ at the occupation level, and the wage growth is defined as $\log(\text{wage}_{o,2015}) - \log(\text{wage}_{o,1980})$. Each circle represents an occupation, and the size of the circle is proportional to the employment in 1980. The patent growth measure has correlations with task scores at the occupation level. Figures 9, 10, and 11 in Appendix A show that patent growth is positively associated with abstract task scores but negatively associated with routine

Table 6: Majors Ranked by the Supply Growth

Rank	Major	D log
1	Transportation and materials moving	286%
2	Parks, recreation, leisure, and fitness studies	208%
3	Legal professions and studies	176%
4	Homeland security, law enforcement, and firefighting	152%
5	Military technologies and applied sciences	148%
6	Multi/Interdisciplinary studies	131%
7	Computer and information sciences	130%
8	Health professions and related programs	114%
9	Communication, journalism, and related programs	109%
10	Area, ethnic, cultural, gender, and group studies	105%
11	Psychology	105%
12	Communications technologies	99%
13	Biological and biomedical sciences	89%
14	Visual and performing arts	88%
15	Liberal arts and sciences, general studies, and humanities	74%
16	Public administrations and social services	70%
17	Mathematics and statistics	64%
18	Business	58%
19	Philosophy and religious studies	57%
20	Foreign languages, literatures, and linguistics	56%
21	Social Sciences and history	54%
22	Theology and religious vocations	51%
23	Agriculture and natural resources	47%
24	English language and literature/letters	46%
25	Engineering	37%
26	Engineering Technologies	36%
27	Family and consumer sciences/human sciences	30%
28	Physical sciences and science technologies	20%
29	Agriculture and Related Sciences	-3%
30	Education	-9%
31	Library Science	-108%

Table 7: 2-Digit SOCs Ranked by the Supply Growth

Rank	Occupation	Supply Shock
1	Healthcare Support Occupations	97%
2	Healthcare Practitioners and Technical Occupations	96%
3	Protective Service Occupations	96%
4	Transportation and Material Moving Occupations	91%
5	Arts, Design, Entertainment, Sports, and Media Occupations	80%
6	Installation, Maintenance, and Repair Occupations	74%
7	Office and Administrative Support Occupations	73%
8	Building and Grounds Cleaning and Maintenance Occupations	71%
9	Sales and Related Occupations	71%
10	Legal Occupations	70%
11	Computer and Mathematical Occupations	70%
12	Food Preparation and Serving Related Occupations	69%
13	Production Occupations	68%
14	Personal Care and Service Occupations	67%
15	Construction and Extraction Occupations	67%
16	Business and Financial Operations Occupations	65%
17	Management Occupations	64%
18	Life, Physical, and Social Science Occupations	64%
19	Community and Social Service Occupations	62%
20	Farming, Fishing, and Forestry Occupations	60%
21	Architecture and Engineering Occupations	47%
22	Education Instruction and Library Occupations	41%

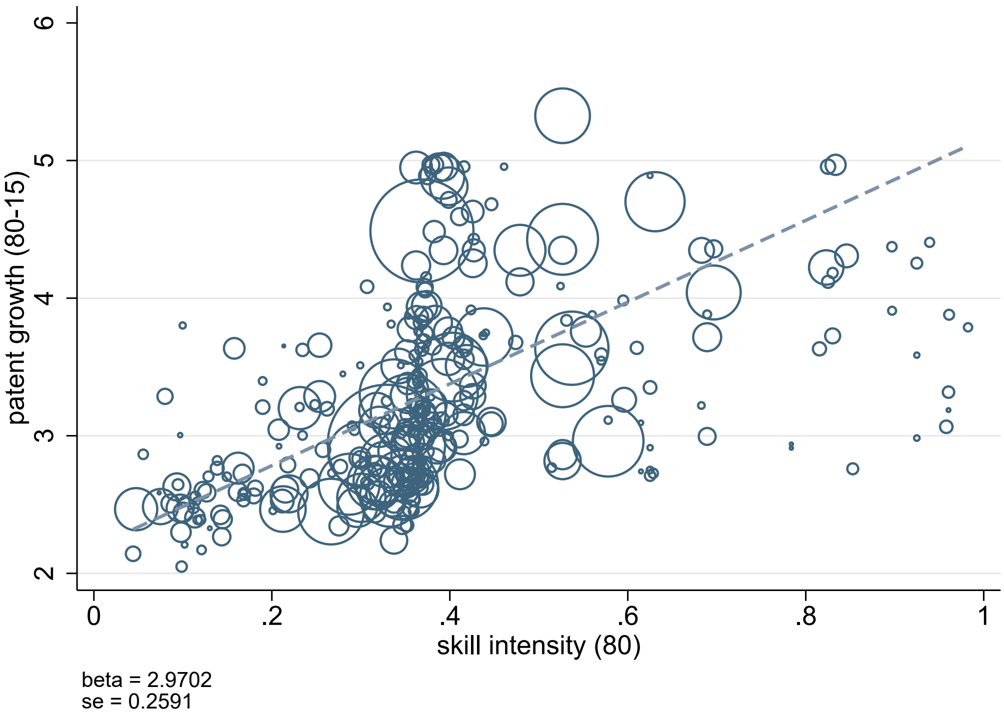


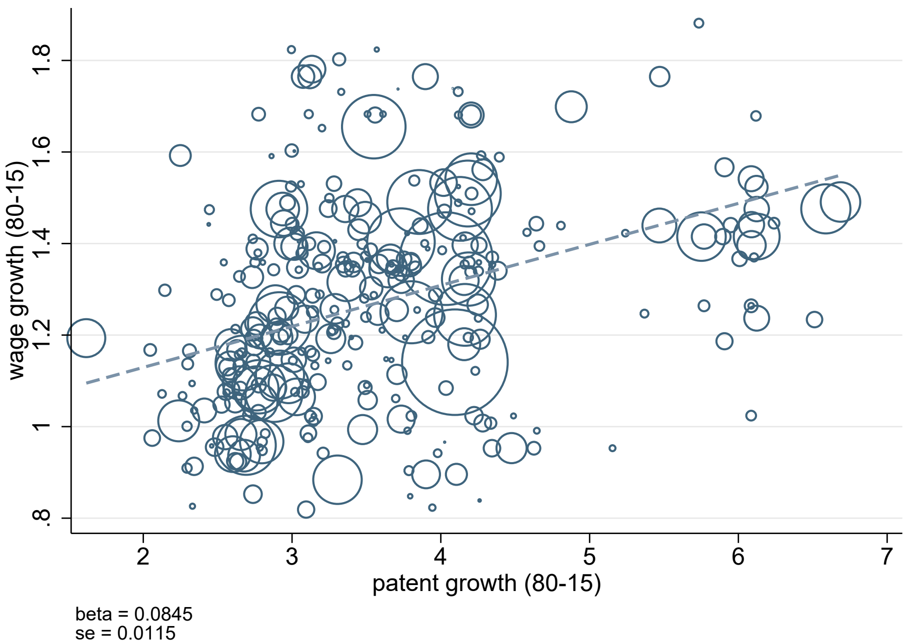
Figure 3: Initial Skill Intensity and Tool Innovation Growth

and manual task scores.

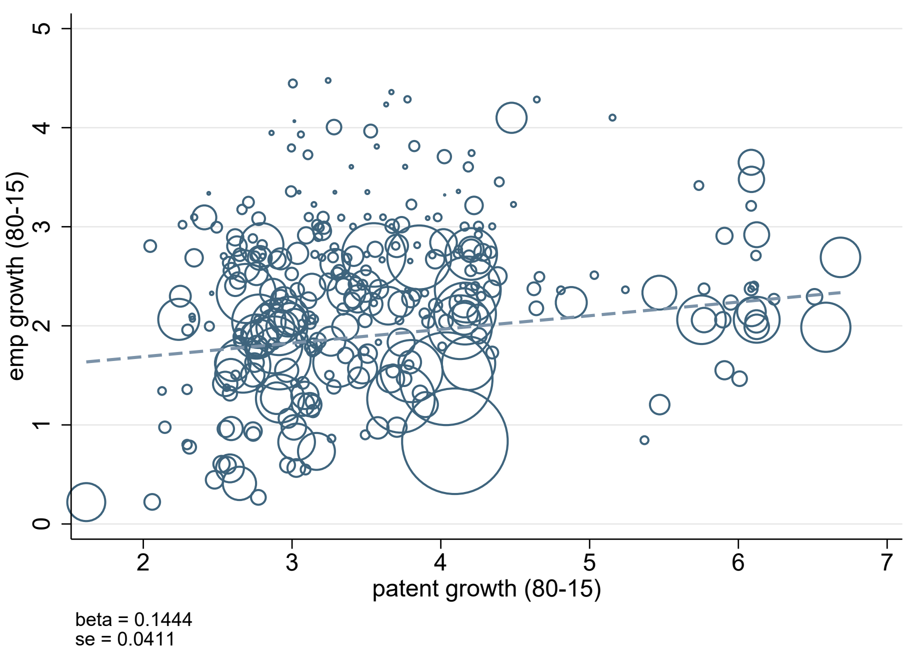
Figure 4a tells that tool innovation growth is positively associated with wage growth, where wage growth is defined by $\log(\text{wage}_{o,2015}) - \log(\text{wage}_{o,1980})$. Figure 4b shows tool innovation growth is positively correlated with employment growth. Employment growth expresses $\log(\text{emp}_{o,2015}) - \log(\text{emp}_{o,1980})$

Figure 5a shows that tool innovation growth is positively associated with skill premium increases within each occupation, where skill premium change is defined as $\text{Skill premium}_{o,2015} - \text{Skill premium}_{o,1980}$. Also, Figure 5b shows that tool innovation growth is positively correlated with skill intensity increases. Skill intensity change means $\text{Skill intensity}_{o,2015} - \text{Skill intensity}_{o,1980}$. The positive correlation suggests that tools are more complementary with skilled workers so the innovation on tools increases

Figure 4: Tool innovation growth and wage, employment growth

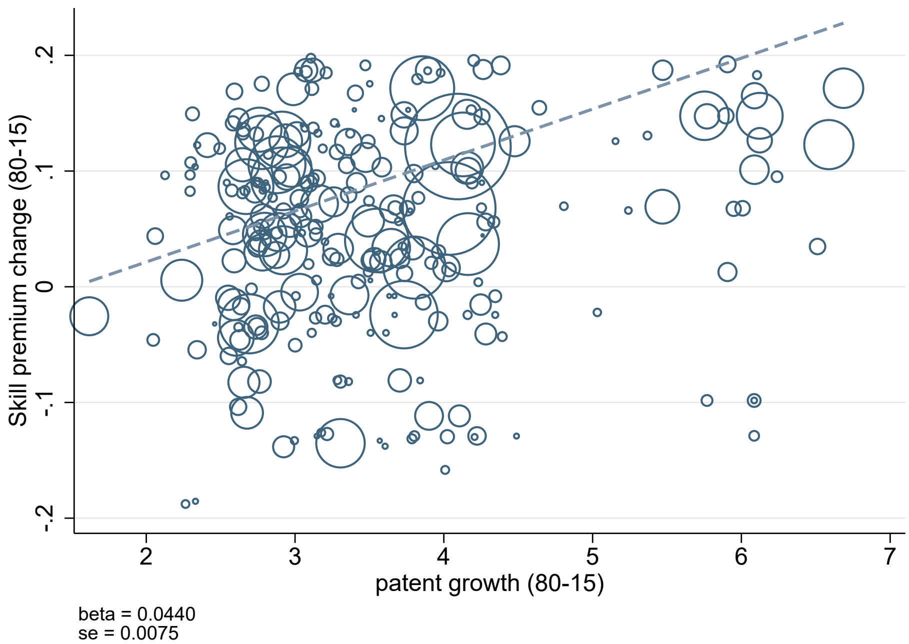


(a)

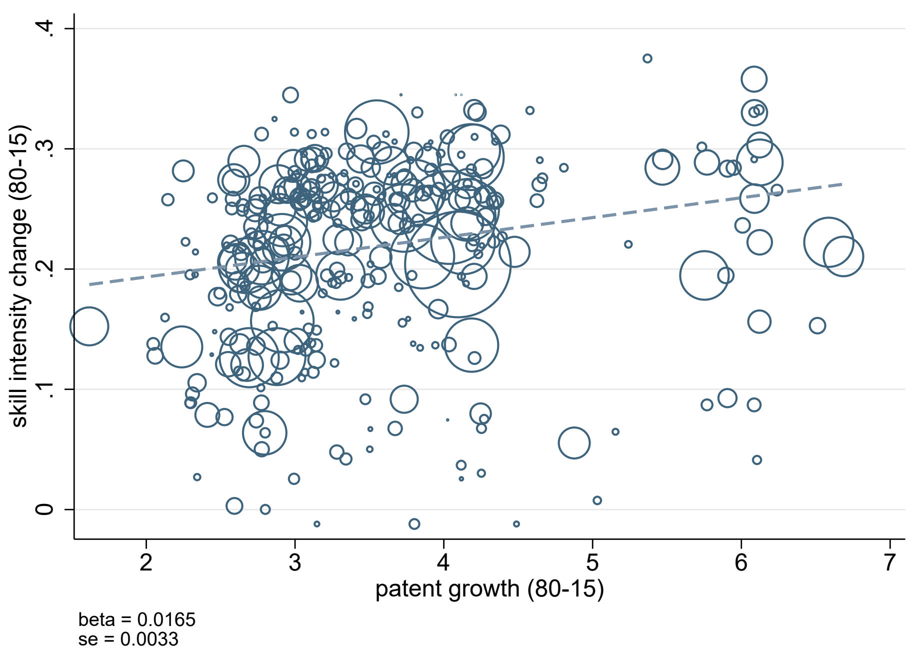


(b)

Figure 5: Tool innovation growth and relative wage, employment growth



(a)



(b)

the relative wage and employment of the skilled worker.

Table 8: OLS result

$$\Delta y_{o,1980-2015} = \beta_0 + \beta_1 \Delta \log p_{o,1980-2015} + \epsilon_o$$

	(1)	(2)	(3)	(4)
	Wage	Emp.	Skill Pre.	Skill Int.
patent	0.0845***	0.144***	0.0440***	0.0165***
	(0.0115)	(0.0411)	(0.00753)	(0.00329)
N	324	324	323	324

Standard errors in parentheses

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

Table 8 shows results from the OLS regression. Columns 1-4 show the baseline result in figure 4a, 4b, 5a, 5b. One standard deviation increase in patent change variable is associated with 8.27% increase in wage, and 14.1% increase in employment. Also, it is associated with 4.31% increase in skill premium, and 1.62% increase in skill intensity. All regressions use wage bills in 1980 at the occupation level for weights.

While we count all patents with the same weight, patents usually have different economic values. Some radical innovations might be more valuable than other marginal innovations. To capture the heterogeneous value of patents, we use results from Kogan et al. (2017), where they measure the monetary value of the US patents from stock market data. We weigh our tool innovation measure by their values and run the same regression as robustness checks. Table 9 shows similar results as the previous table.

For a robustness check, we control industry-specific factors that are related to demand changes. We first decompose workers in the same occupation into different industries of their employers. We then include sectoral fixed effects at the 3-digit NAICS level to control industry-specific shocks. For instance, there has been a structural transformation where the employment share of manufacturing decreases while that of the service

Table 9: OLS result, monetary value weighted patents

$$\Delta y_{o,1980-2015} = \beta_0 + \beta_1 \Delta \log p_{o,1980-2015} + \epsilon_o$$

	(1)	(2)	(3)	(4)
	Wage	Emp.	Skill Pre.	Skill Int.
patent (value)	0.0458***	0.0898*	0.0316***	0.0112***
	(0.0103)	(0.0353)	(0.00650)	(0.00283)
N	324	324	323	324

Standard errors in parentheses

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

sector increases. Table 14 in Appendix B.1 shows the result is similar to the baseline case.

3.2 Endogeneity Concern and IV result

Innovation on tools has an endogeneity problem if the technical change is directed by the demand factor, as suggested by Acemoglu (2002a). For example, when there is another demand shock for IT sector workers, the value of innovation in the IT sector will increase, which leads to the increase in the innovation incentive on tools in the IT sector such as a computer. Innovation on tools can be correlated with this unobserved demand shock which affects wage, employment, and college premium.

To tackle this problem, we construct the "upstream" innovation measure. We exploit the fact that the knowledge production function is interconnected with other sectors in the sense that one technology field uses knowledge from many other fields. For example, innovation in the computer sector uses knowledge from more basic research fields such as electricity. Innovation in the electricity sector would be positively correlated with innovations in the computer sector. However, it is plausible that innovation in electricity is not correlated with demand shock in the IT sector. We use this "upstream" innovation as our instrumental variable.

We use citation network structure to identify upstream sectors in patent data following Cai and Li (2019). To register a patent in the USPTO, the authors need to disclose background knowledge used for the new invention. If the patent cites other patents, we assume that the patent uses knowledge from the cited patents. We use the 3-digit IPC that has 270 unique classes in the citation network. First, we collect all the citations in 1970-2015. Then, for each technology class, we construct the upstream measure as follows.

$$\alpha_{ij} = \frac{c_{ij}}{\sum_{\{j:s(j)\neq s(i)\}} c_{ij}} \tag{2}$$

c_{ij} is the number of citations from class i to j. α_{ij} indicates the degree of dependence of class i on class j. We may concern that some technology classes are very similar and they share the same demand shock. To avoid this problem, we exclude the combination of classes if i and j are in the same 2-digit NAICS. Finally, we construct the following variable for each technology class.

$$\sum_{\{j:s(j)\neq s(i)\}} \alpha_{ij} \Delta \log P_j \tag{3}$$

Each tool has innovations from different technology fields. We aggregate variables by the share of technology class in each tool, then aggregate them into occupations.

Table 10 shows the IV regression results. All regressions are weighted by the wage bills of occupations in 1980. F-statistics from the first stage regression indicate that the instrument is strong enough. The coefficients are similar to the result from the OLS where tool innovation growth is associated with wage and employment of occupations and correlated with the relative wage and employment of skilled workers. Table 15 in Appendix B.2 describes the result where we use value-weighted patents as in Table 9. The results are similar to the baseline case.

Table 10: IV result

$$\Delta y_{o,1980-2015} = \beta_0 + \beta_1 \Delta \log p_{o,1980-2015} + \epsilon_o$$

	(1)	(2)	(3)	(4)
	Wage	Emp.	Skill Pre.	Skill Int.
patent	0.119***	0.0600	0.0406***	0.0297***
	(0.0156)	(0.0551)	(0.0100)	(0.00449)
N	324	324	323	324
F (first stage)	410.8	410.8	409.6	410.8

Standard errors in parentheses

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

There can be a remaining concern that "upstream" sectors are still correlated with the demand shock of the occupations. To avoid this problem, we construct instrumental variables with publication data where we exploit the fact that many patents obtain knowledge spillover from academic papers. Marx and Fuegi (2020) provide citation data from patents to academic papers in Microsoft Academic Graph (MAG hereafter, Sinha et al. (2015)) and 27% of USPTO patents cite academic papers. Similar to the baseline IV, we leverage the fact that different patent classes have different shares of citations to different academic fields. We use the OECD subfield where we have 42 different classifications[14] and calculate citation share using pre-period samples. For the growth rate of papers, we collect all papers across different fields in MAG and calculate the growth rate in 1980-2015. Table 11 shows that the result is both quantitatively and qualitatively similar to the baseline IV result.

[14] We have fields such as mechanical engineering and chemical science

Table 11: IV result with academic publications

$$\Delta y_{o,1980-2015} = \beta_0 + \beta_1 \Delta \log p_{o,1980-2015} + \epsilon_o$$

	(1) Wage	(2) Emp.	(3) Skill Pre.	(4) Skill Int.
patent	0.0963***	0.0562	0.0489***	0.0191***
	(0.0156)	(0.0558)	(0.0102)	(0.00443)
N	324	324	323	324
F (first stage)	390.5	390.5	389.3	390.5
sector fixed	no	no	no	no

Standard errors in parentheses

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

4 Structural Estimation

4.1 Model

For structural estimation, we assume that representative firms have the following aggregate production function with occupational inputs.

$$Y_t = Z_t \left(\sum_o \alpha_{ot} \left(\lambda_{ot} \left(x_{ot} s_{ot} \right)^{\frac{\sigma-1}{\sigma}} + (1 - \lambda_{ot}) \left(x_{ot}^{\delta} u_{ot} \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1} \frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}}$$

This specification is similar to the task-based approach popularized by Acemoglu and Autor (2011), but we substitute tasks with occupations. In this specification, the tasks performed by different occupations are imperfectly substituted with the elasticity of substitution $\rho > 0$. The occupation-specific shock α_{ot} governs the demand for certain occupations and is normalized such that $\sum_o \alpha_{ot} = 1$. Moreover, each occupation task takes skilled and unskilled labor inputs imperfectly with the elasticity of substitution $\sigma > 0$. x_{ot} is the innovation on tools, measured by the number of patents applied to the tools used in that occupation constructed in Section 2.3. The parameter δ is the loading of x_{ot} on unskilled labor relative to skilled labor. $\delta < 1$ means that the productivity of unskilled labor increases by less with a higher x_{ot} than the productivity of skilled labor and vice versa. The other parameter, λ_{ot}, is the demand shock for skilled labor relative to unskilled labor within an occupation. The occupation without any unskilled labor can be understood as $\lambda_{ot} = 1$.

Given the prices, the representative firm solves the following problem.

$$\max_{s_{ot}u_{ot}} \left\{ Z_t \left(\sum_o \alpha_{ot} \left(\lambda_{ot} \left(x_{ot} s_{ot} \right)^{\frac{\sigma-1}{\sigma}} + \left(1 - \lambda_{ot}\right) \left(x_{ot}^{\delta} u_{ot} \right)^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}} - \sum w_{ot}^{u} u_{ot} - \sum w_{ot}^{s} s_{ot} \right\} \tag{4}$$

The first-order conditions of Equation 4, after taking logs, are given as the following equations.

$$\log\left(\frac{w_{ot}^s}{w_{ot}^u}\right) = \log\left(\frac{\lambda_{ot}}{1-\lambda_{ot}}\right) + \frac{\sigma-1}{\sigma}\left(1-\delta\right)\log\left(x_{ot}\right) - \frac{1}{\sigma}\log\left(\frac{s_{ot}}{u_{ot}}\right) \tag{5}$$

$$\log\left(\frac{w_{ot}^s}{w_{1t}^s}\right) = \log\left(\frac{\lambda_{ot}}{\lambda_{pt}}\frac{\alpha_{ot}}{\alpha_{1t}}\right) + \frac{\rho-\sigma}{\rho\left(\sigma-1\right)}\log\left(\frac{\theta_{ot}}{\theta_{1t}}\right) + \frac{\sigma-1}{\sigma}\log\left(\frac{x_{ot}}{x_{1t}}\right) - \frac{1}{\rho}\log\left(\frac{s_{ot}}{s_{1t}}\right) \tag{6}$$

$$\theta_{ot} := \left(\lambda_{ot} x_{ot}^{\frac{\sigma-1}{\sigma}} + \left(1-\lambda_{ot}\right) x_{ot}^{\frac{(\sigma-1)(\delta\sigma-\sigma+1)}{\sigma}} \left(\frac{1-\lambda_{ot}}{\lambda_{ot}} \frac{w_{ot}^s}{w_{ot}^u} \right)^{\sigma-1} \right). \tag{7}$$

Equation 5 defines within-occupation demand for skilled and unskilled labor. The relative demand for skilled labor decreases in within-occupation skill premium with a constant elasticity σ. If $\delta < 1$, the demand curve for skilled labor shifts to the right with a higher x_{ot}, increasing within-occupation skill premium.

Equation 6[15] sets up the relative demand for skilled labor across occupations. The relative demand has the demand elasticity of ρ, and a higher x_{ot} relative to a baseline occupation $o = 1$ shifts the demand curve to the right. This implicitly assumes the technological complementarity between innovation on tools and the skilled labor, motivated by the reduced form evidence in Section 3.

We estimate the structural parameters sequentially using the Generalized Method of Moments (GMM). First, we take differences between 1980 and 2015 over Equations 5 and 6. The purpose of differencing is to minimize time-invariant measurement errors in matching tools with occupations[16]. Second, we estimate the first equation at the occupation level after assuming that $Cov(\Delta_t \lambda_{ot}, \Delta_t z^i_{ot}) = 0$, where $i = 1, 2$[17]. z^1_{ot} is the upstream patent IV constructed in Section 2.3, and z^2_{ot} is the Bartik-style skilled labor supply shock from the number of college graduates for each major constructed in Section 2.4.2.

Next, Equation 6 is estimated using the GMM with $\Delta_t z^1_{ot} - \Delta_t z^1_{1t}$ and $\Delta_t z^2_{ot} - \Delta_t z^2_{1t}$ as two instruments. The composite term in Equation 6, θ_{ot}, can be calculated after the first regression after obtaining λ_{ot} as residuals. This across-occupation regression is over-identified with two instrumental variables and one coefficient. This regression also gives the estimates for $\log(\alpha_{ot}/\alpha_{1t})$ as residuals.

On the supply side, I assume a simple labor market supply function with a constant

[15]The composite term θ_{ot} is defined in Equation 7.

[16]Suppose the text-matching procedure overestimates the number of meaningful innovations on laser cutter relative to other tools. Then, the differencing controls for multiplicative measurement errors on x_{ot}.

[17]To identify two parameters, we need two instruments.

elasticity.

$$s_{ot} = \left(\frac{exp(\log w^s_{ot} + \epsilon^s_{ot})}{\sum_{o'} exp(\log w^s_{o't} + \epsilon^s_{o't})} \right)^{\eta_s} L^s_t$$
$$u_{ot} = \left(\frac{exp(\log w^u_{ot} + \epsilon^u_{ot})}{\sum_{o'} exp(\log w^u_{o't} + \epsilon^u_{o't})} \right)^{\eta_u} L^u_t$$

This supply function can be micro-founded by a discrete choice model á la McFadden (1973) with log indirect utility function, a common supply shock, and the Type 1 Extreme Value idiosyncratic shock across a continuum of households with size L^s_t for skilled labor and L^u_t for unskilled labor. I further assume $\eta_s = \eta_u = \eta$. Notice that we deliberately avoid the option for home production for the tractability of the model. We use observations for skilled labor in 2015 to estimate η. Specifically, the regression equation is described as the following.

$$\log \left(\frac{s_{ot}}{s_{1t}} \right) = \eta \left(\log w^s_{ot} + \epsilon^s_{ot} - \log w^s_{1t} - \epsilon^s_{1t} \right)$$

The supply elasticity is estimated through a similar GMM procedure, with x_{ot} as the instrumental demand shifter. For the counterfactual analysis below, please note that the relative skill supply equation within each occupation is given as

$$\begin{aligned} \log \left(\frac{s_{ot}}{u_{ot}} \right) &= \eta_s \log \left(\frac{w^s_{ot}}{w^u_{ot}} \right) + (\eta_s - \eta_u) \log w^u_{ot} + \eta_s \left(\epsilon^s_{ot} - \epsilon^u_{ot} \right) \\ &\quad + (\eta_s - \eta_u) \left(\epsilon^u_{ot} \right) + \log \tilde{L}^s_t - \log \tilde{L}^u_t \\ \log \left(\frac{s_{ot}}{u_{ot}} \right) &= \eta \log \left(\frac{w^s_{ot}}{w^u_{ot}} \right) + \eta \left(\epsilon^s_{ot} - \epsilon^u_{ot} \right) + \log \tilde{L}^s_t - \log \tilde{L}^u_t, \end{aligned}$$

where $\log \tilde{L}^s_t = \log L^s_t - \eta \log \left(\sum_{o'} exp(\log w^s_{o't} + \epsilon^s_{o't}) \right)$ and $\log \tilde{L}^u_t = \log L^u_t - \eta \log \left(\sum_{o'} exp(\log w^u_{o't} + \epsilon^u_{o't}) \right)$ Along with Equation 5, this supply equation determines within-occupation skill premium and relative employment.

4.2 Results

Table 12: GMM Estimates

Parameter	δ	σ	ρ	η
Estimate	0.94	11.11	1.40	2.40
SE	(0.03)	(10.47)	(0.13)	(0.64)

Table 12 summarizes the structural results. The estimate for δ is 0.94 and significantly smaller than 1, the normalized loading on the demand for skilled workers, at the 5% significance level. This estimate implies that the effect of innovation on tools is larger for skilled workers. As a result, the skill premium within occupations tends to increase with a higher value of $\log x_{ot}$. This tool-skill complementarity is comparable to the capital-skill complementarity pointed out in Krusell et al. (2000). Unlike the model-based accounting exercise in Krusell et al. (2000) that uses only aggregate-level time-series about capital and skill-premium, our exercise exploits the across-occupation variations of technological improvements on tools to show that the occupations with more innovation on tools actually experienced a larger increase in skill premium within each occupation.

Next, the estimate for within-occupation elasticity of substitution, σ, is 11.1, with a large standard error. This estimate is larger than estimates of substitutability between skilled and unskilled labor within an *aggregate* production function[18]. The large standard error of the estimate suggests that occupations are heterogeneous in terms of within-occupation elasticity of substitution between skilled and unskilled labor. Caunedo et al. (2021) recently report that matches tools in O*NET to the capital stock series of the national accounts (NIPA). Moreover, although the large standard error of the estimate for σ makes it difficult to compare the estimate, the estimate,

[18]For example, in Krusell et al. (2000), the demand elasticity between skill groups was estimated to be 1.67.

11.11, is greater in magnitude than the estimate for across-occupation demand elasticity, 1.40. The higher within-elasticity implies that the representative firm tends to substitute unskilled labor with skilled labor rather than put more skilled labor from other occupations when the relative productivity of skilled labor increases in an occupation. As shown in the counterfactual analysis below, this is why the within-occupation channel is more important in generating the rise in skill premium than the across-occupation channel. The across-elasticity estimate being close to one suggests that the aggregate production function is close to the Cobb-Douglas form across occupations, which was assumed in the baseline task-based approach in Acemoglu and Autor (2011).

This estimate of across-occupation demand elasticity is smaller than the number assumed in Hsieh et al. (2019), 3, which do not consider the within-occupation substitution between skilled and unskilled labor. The number is also smaller than the number in Burstein et al. (2019) that estimates the elasticity of substitution to be roughly $1.81 - 2.1$. Their framework, unlike ours, assumes that the comparative advantage of skilled workers in using computers within each occupation is fixed, but the productivity increase in computer equipment is constant across all occupations. In this case, all within-period occupation fluctuations in skill premium are loaded as comparative advantage of skilled workers, and the growth of computer technology tends to be understated. As a result, their estimation equation that regresses the total labor income of a demographic group in each occupation on the computer productivity is likely to overestimate the elasticity of substitution across occupations. Introducing the heterogeneous fluctuations in tools' productivity across occupations, we find that the direct measure of tool productivity is large enough to reduce the magnitude of across-occupation elasticity close to one with smaller bounds.

Lastly, the estimate for supply elasticity, η is 2.40. This number is close to the baseline estimate of extensive elasticity of labor supply reported in Hsieh et al. (2019), 2, but slightly larger than the estimate in Burstein et al. (2019), which ranges from 1.8 to 1.3.

5 Counterfactual

5.1 Procedure

The goal of the counterfactual exercise is to identify the quantitative impact of patents on tools, i.e. changes in x_{ot}, on rising skill premium in contrast to the changes in other demand-side residuals such as λ_{ot} and α_{ot}. The following two equations, one on relative demand and the other one on relative supply, determine within occupation skill premium and relative employment.

$$\log\left(\frac{w_{ot}^s}{w_{ot}^u}\right) = \log\left(\frac{\lambda_{ot}}{1-\lambda_{ot}}\right) + \frac{\sigma-1}{\sigma}(1-\delta)\log\left(x_{ot}\right) - \frac{1}{\sigma}\log\left(\frac{s_{ot}}{u_{ot}}\right) \tag{8}$$

$$\log\left(\frac{s_{ot}}{u_{ot}}\right) = \eta\log\left(\frac{w_{ot}^s}{w_{ot}^u}\right) + \eta\left(\epsilon_{ot}^s - \epsilon_{ot}^u + \log\tilde{L}_t^s - \log\tilde{L}_t^u\right) \tag{9}$$

After estimating σ and δ, estimates for λ_{ot} and $\epsilon_{ot}^s - \epsilon_{ot}^u + \log\tilde{L}_t^s - \log\tilde{L}_t^u$[19] are given as residuals for $t = 1980$ and $t = 2015$. We start from the values of supply and demand shifters in 1980 and substitute x_{o1990} with x_{o2015}. Then, we recalculate the within-occupation skill premium by solving the linear equation system. The relative demand

[19]These four terms are not identified separately.

and supply functions for labor *across* occupations are given by

$$\log\left(\frac{w^s_{ot}}{w^s_{1t}}\right) = \log\left(\frac{\lambda_{ot}}{\lambda_{1t}}\frac{\alpha_{ot}}{\alpha_{1t}}\right) + \frac{\rho-\sigma}{\rho\,(\sigma-1)}\log\left(\frac{\theta_{ot}}{\theta_{1t}}\right) + \frac{\sigma-1}{\sigma}\log\left(\frac{x_{ot}}{x_{1t}}\right) - \frac{1}{\rho}\log\left(\frac{s_{ot}}{s_{1t}}\right) \tag{10}$$

$$\log\left(\frac{s_{ot}}{s_{1t}}\right) = \eta\left(\log w^s_{ot} + \epsilon^s_{ot} - \log w^s_{1t} - \epsilon^s_{1t}\right). \tag{11}$$

As before, the values for α_{ot}/α_{1t} and $\epsilon^s_{ot}-\epsilon^1_{ot}$ are given as the residuals of the regressions. We start from values of $\lambda_{ot}/\lambda_{1t}$, α_{ot}/α_{1t}, θ_{ot}/θ_{1t}, and $\epsilon^s_{ot}-\epsilon^s_{1t}$ for $t = 1980$ and use x_{ot}/x_{1t} for $t = 2015$. Then, we solve the linear equations and compute the counterfactual wage of skilled labor across occupations relative to the skilled labor in the baseline occupation. From there, along with the within-occupation skill premium, we can pin down the counterfactual wage of unskilled labor across occupations relative to the baseline wage. Appendix D shows that this relative wage is sufficient to define the aggregate skill premium.

5.2 Results

Scenario	Pre-1980	Post-1980	Patent Only	Demand	Within only	Across Only
Level	1.32	1.58	1.97	2.38	1.81	1.34
Relative	1	1.20	1.49	1.80	1.37	1.02

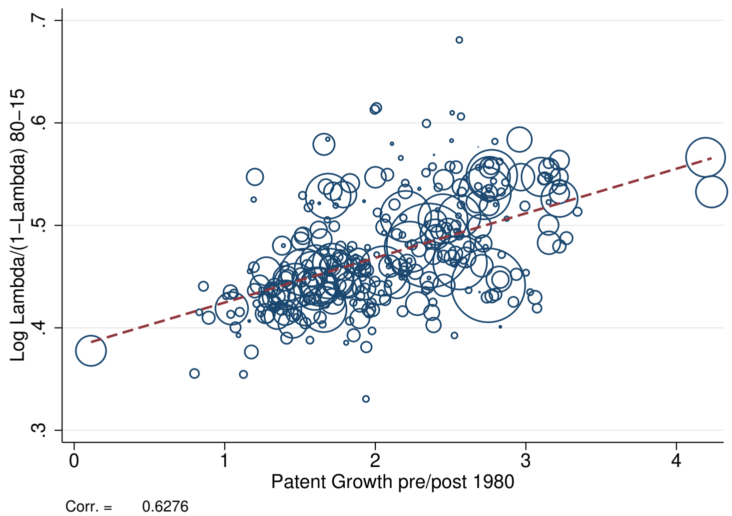
Table 13: Counterfactual Skill Premium

Table 13 shows the values for counterfactual skill premium. Between 1980-2015, the skill premium increased by 26 percentage points. If we substitute the technological improvements measure, x_{ot} with its value of 2015, the skill premium increases by 65 percentage points, which is more than the actual increase in the skill premium. Putting other demand factor changes in this period, α_{ot} and λ_{ot}, the counterfactual premium increases further by 41 percentage points. The innovation on tools accounts for 61% of the total demand factor that contributed to the increase in the skill premium. The supply factor, relative increases in skilled labor supply around occupations with higher demand shocks, suppressed this increase in skilled labor.

To further decompose the impact of innovation on tools between within-occupation and across-occupation margins, we further calculate the counterfactual skill premium when the values of x_{ot} are substituted only in the within-occupation labor demand equation, in the second last column of Table 13, and only in across-occupation skilled labor demand equations in the last column of Table 13. It is the within occupation margin that explains most increases in skill premium from innovation on tools. Since the within-occupation demand elasticity is higher than the across-occupation, the increase in the relative demand for skilled labor is translated into an increase in wage and employment of skilled workers in skill-intensive occupations, which already had a higher within-occupation skill premium, rather than a reallocation of skilled workers from low-productivity occupations to high-productivity ones.

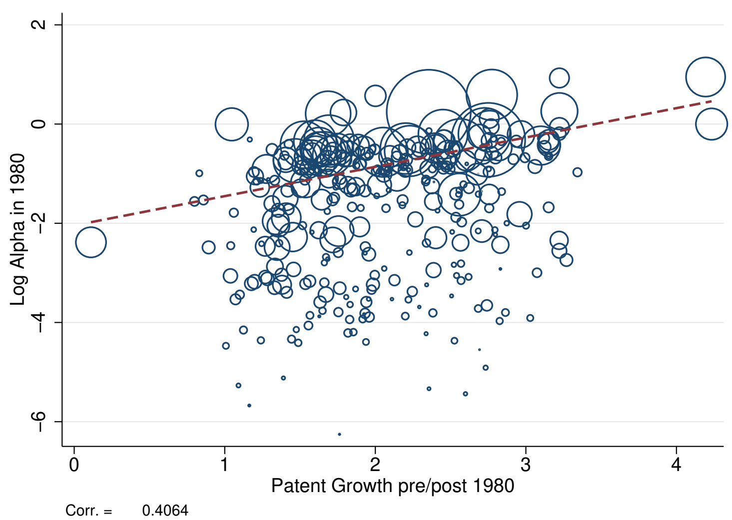
Moreover, the residual demand shock estimates of 1980, α_{o1980} and λ_{o1980}, are positively correlated with the patent growth between 1980 and 2015, as shown in Figures 6 and 7. This correlation supports the idea that knowledge production responds to the increase in demand for occupational tasks. Because of the positive autocorrelation in α_{ot} and λ_{ot} at the occupation-level, the correlation between α_{ot}, λ_{ot}, and x_{ot} is not zero, which rationalizes the use of the instrumental variables, z^1_{ot} and z^2_{ot}, in the GMM estimation above.

Figure 6: Lambda in 1980 and Patent Growth



Note: Log Lambda and patent growth are defined by $\log(\lambda_{ot}/(1-\lambda_{ot}))$ and $\Delta_t \log(x_{ot})$, respectively. The size of circles correspond to wage bill share of occupations in 1980.

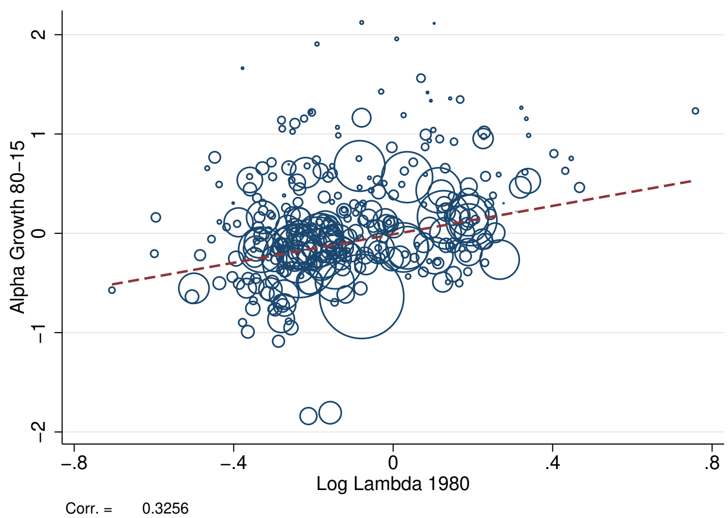
Figure 7: Alpha in 1980 and Patent Growth



Note: Log Alpha and patent growth are defined by $\log(\alpha_{ot}) - \log(\alpha_{1t})$ and $\Delta_t \log(x_{ot})$, respectively. The baseline occupation $o = 1$ is First-Line Supervisors of Non-Retail Sales Workers (SOC code 41-1012). The size of circles correspond to wage bill share of occupations in 1980.

Figure 8 shows why the residual demand components widen the wage gap between skilled and unskilled labor. The within-occupation demand shock for skilled labor, $\log(\lambda_{ot}/(1-\lambda_{ot}))$, for 1980, is positively correlated with the change of α_{ot} between 1980 and 2015. This implies that, even after controlling for the technological changes through innovation on tools, occupations that had a higher demand for skilled labor in 1980 experienced even larger increases in across-occupation relative demand and employment share, which contributes to the historical rise in the skill premium.

Figure 8: Lambda in 1980 and Alpha Growth



Note: Log Lambda and Lambda growth are defined by $\log(\lambda_{ot}/(1-\lambda_{ot}))$ and $\Delta_t \log(\lambda_{ot}/(1-\lambda_{ot}))$, respectively. The size of circles correspond to wage bill share of occupations in 1980.

6 Conclusion

This paper measures the innovation on tools used by different occupations and studies its effect on the aggregate skill premium. We use patent data to measure the technical change on tools. Since patent is not naturally connected to the tools we calculate text similarity between patent abstract and the description of tools from Wikipedia to obtain the number of relevant papers for each tool. With the innovation measure at the occupation level, we first provide reduced form results. We show that tool innovation grew more in skill-intensive occupations. We also show that tool innovation is positively associated with wage, employment, skill premium, and skill intensity. This suggests that tool innovation increases the demand for occupations, more with the skilled workers. To tackle the potential endogeneity problems, we construct instrumental variables using knowledge spillover in patent production and find similar results. Motivated by this reduced form evidence, we build a model where the firm hires skilled and unskilled workers for different occupations and workers choose occupations. We estimate parameters using GMM, and run counterfactual exercises. The result shows that 61% of the total demand factor that contributed to the rise of skill premium can be attributed to the tool innovation.

References

Acemoglu, Daron, "Directed technical change," *The review of economic studies*, 2002, *69* (4), 781–809.

_ , "Technical change, inequality, and the labor market," *Journal of economic literature*, 2002, *40* (1), 7–72.

_ **and David Autor**, "Skills, tasks and technologies: Implications for employment and earnings," in "Handbook of labor economics," Vol. 4, Elsevier, 2011, pp. 1043–1171.

_ **and Pascual Restrepo**, "The race between man and machine: Implications of technology for growth, factor shares, and employment," *American Economic Review*, 2018, *108* (6), 1488–1542.

_ **and** _ , "Robots and jobs: Evidence from US labor markets," *Journal of Political Economy*, 2020, *128* (6), 2188–2244.

Akcigit, Ufuk, John Grigsby, and Tom Nicholas, "The rise of american ingenuity: Innovation and inventors of the golden age," Technical Report, National Bureau of Economic Research 2017.

Argente, David, Salomé Baslandze, Douglas Hanley, and Sara Moreira, "Patents to products: Product innovation and firm dynamics," *FRB Atlanta Working Paper*, 2020.

Autor, David H and David Dorn, "The growth of low-skill service jobs and the polarization of the US labor market," *American Economic Review*, 2013, *103* (5), 1553–97.

— , **Frank Levy, and Richard J Murnane**, "The skill content of recent technological change: An empirical exploration," *The Quarterly journal of economics*, 2003, *118* (4), 1279–1333.

— , **Lawrence F Katz, and Melissa S Kearney**, "The polarization of the US labor market," *American economic review*, 2006, *96* (2), 189–194.

Bloom, Nicholas, Tarek Alexander Hassan, Aakash Kalyani, Josh Lerner, and Ahmed Tahoun, "The Diffusion of Disruptive Technologies," Technical Report, National Bureau of Economic Research 2021.

Burstein, Ariel, Eduardo Morales, and Jonathan Vogel, "Changes in between-group inequality: computers, occupations, and international trade," *American Economic Journal: Macroeconomics*, 2019, *11* (2), 348–400.

Cai, Jie and Nan Li, "Growth through inter-sectoral knowledge linkages," *The Review of Economic Studies*, 2019, *86* (5), 1827–1866.

Caunedo, Julieta, David Jaume, and Elisa Keller, "Occupational Exposure to Capital-Embodied Technical Change," 2021.

Cortes, Guido Matias, Nir Jaimovich, and Henry E Siu, "Disappearing routine jobs: Who, how, and why?," *Journal of Monetary Economics*, 2017, *91*, 69–87.

Dierdorff, EC, DW Drewes, and JJ Norton, "O* NET tools and technology: A synopsis of data development procedures," *North Carolina State University. http://www. onetcenter. org/dl_files/T2Development. pdf*, 2006.

Greenwood, Jeremy and Mehmet Yorukoglu, "1974," in "Carnegie-Rochester conference series on public policy," Vol. 46 Elsevier 1997, pp. 49–95.

Hornstein, Andreas, Per Krusell, and Giovanni L Violante, "The effects of technical change on labor market inequalities," in "Handbook of economic growth," Vol. 1, Elsevier, 2005, pp. 1275–1370.

Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow, "The allocation of talent and us economic growth," *Econometrica*, 2019, *87* (5), 1439–1474.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring technological innovation over the long run," Technical Report, National Bureau of Economic Research 2018.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, "Technological innovation, resource allocation, and growth," *The Quarterly Journal of Economics*, 2017, *132* (2), 665–712.

Krusell, Per, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante, "Capital-skill complementarity and inequality: A macroeconomic analysis," *Econometrica*, 2000, *68* (5), 1029–1053.

Lee, Donghoon and Kenneth I Wolpin, "Intersectoral labor mobility and the growth of the service sector," *Econometrica*, 2006, *74* (1), 1–46.

Marx, Matt and Aaron Fuegi, "Reliance on science: Worldwide front-page patent citations to scientific articles," *Strategic Management Journal*, 2020, *41* (9), 1572–1594.

McFadden, D, "Conditional logit analysis of qualitative choice behavior," *Frontiers in Econometrics*, 1973, pp. 105–142.

Nelson, Richard R and Edmund S Phelps, "Investment in humans, technological diffusion, and economic growth," *The American economic review*, 1966, *56* (1/2), 69–75.

Sinha, Arnab, Zhihong Shen, Yang Song, Hao Ma, Darrin Eide, Bo-June Hsu, and Kuansan Wang, "An overview of microsoft academic service (mas) and applications," in "Proceedings of the 24th international conference on world wide web" 2015, pp. 243–246.

Violante, Giovanni L, "Skill-biased technical change," *The new Palgrave dictionary of economics*, 2008, *2*, 1–6.

Webb, Michael, "The impact of artificial intelligence on the labor market," *Available at SSRN 3482150*, 2019.

Zhestkova, Yulia, "Technology Cluster Dynamics and Network Structure," *Working Paper*, 2020.

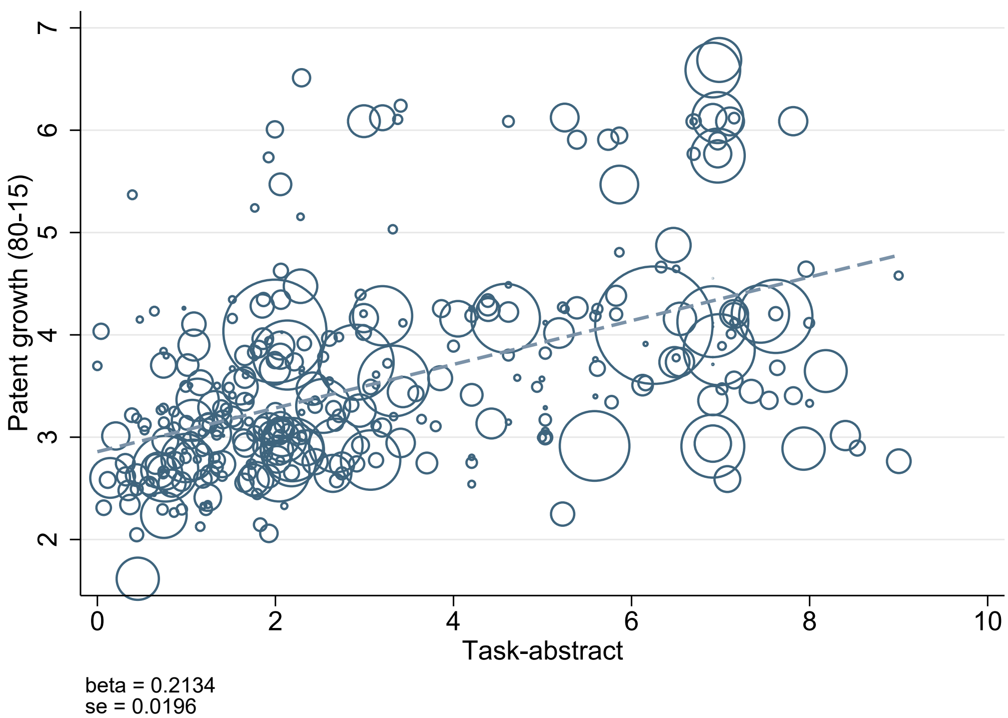
Appendix

A Task-based approaches and Innovation on Tools

The innovation on tools captures a different aspect of occupation-specific shock than task scores in Autor et al. (2006). For example, die makers and payroll clerks have similar skill contents. Both have modest abstract scores (2.65 vs. 2.05), high routine scores (8.17 vs. 7.95), and low manual scores (0.08 vs. 0). However, die makers had a smaller increase in tool innovation, 1.05 log points, while payroll workers had an increase by 1.95 log points. At the same time, while the mean wage for die makers increases by 0.59 log points, the mean wage for payroll clerks grows by 0.85 log points. Skill premium for die makers decreases by 0.93 percentage points, but skill premium for payroll clerks increases by 3.33 percentage points. Similarly, the increases in employment and skill intensity were modestly greater for payroll clerks than for die makers. As such, the tools innovation measure captures a more granular dimension of skill-biased technical changes that task-biased approaches do not.

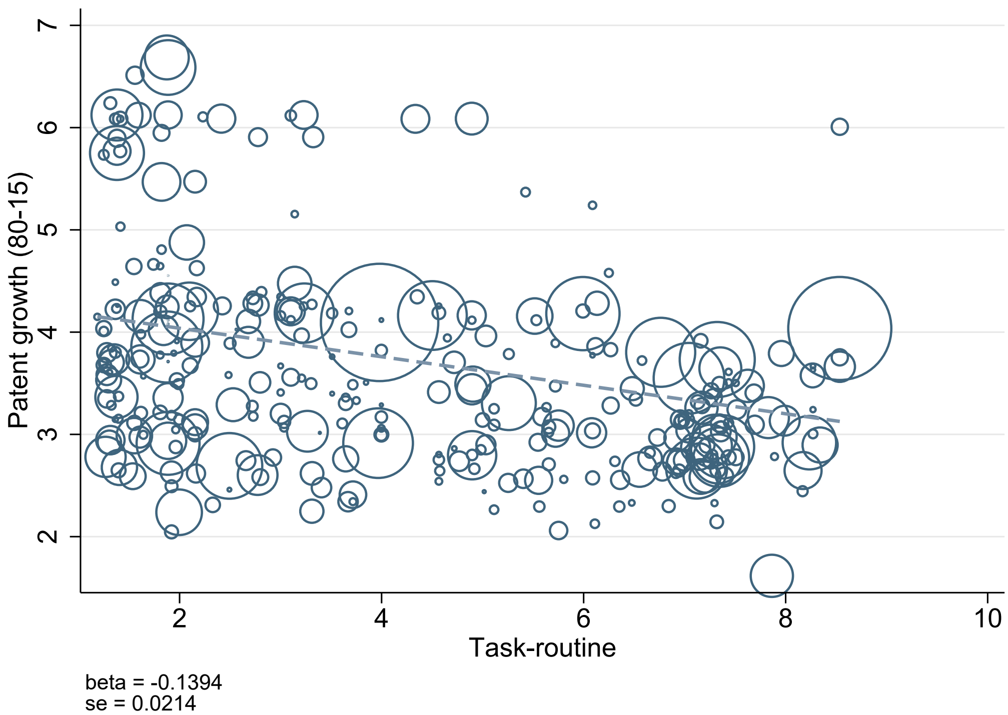
We also study the correlation between task-based measures and our tool innovation measure. Figure 9 shows that the innovation on tools is positively associated with abstract task scores of occupations. Figures 10, 11 show that the innovation on tools is negatively correlated with routine and manual task scores of the occupations. Each circle size represents the wage bill in 1980. The task-based measure is from Autor et al. (2006).

Figure 9: Patent growth and abstract measure



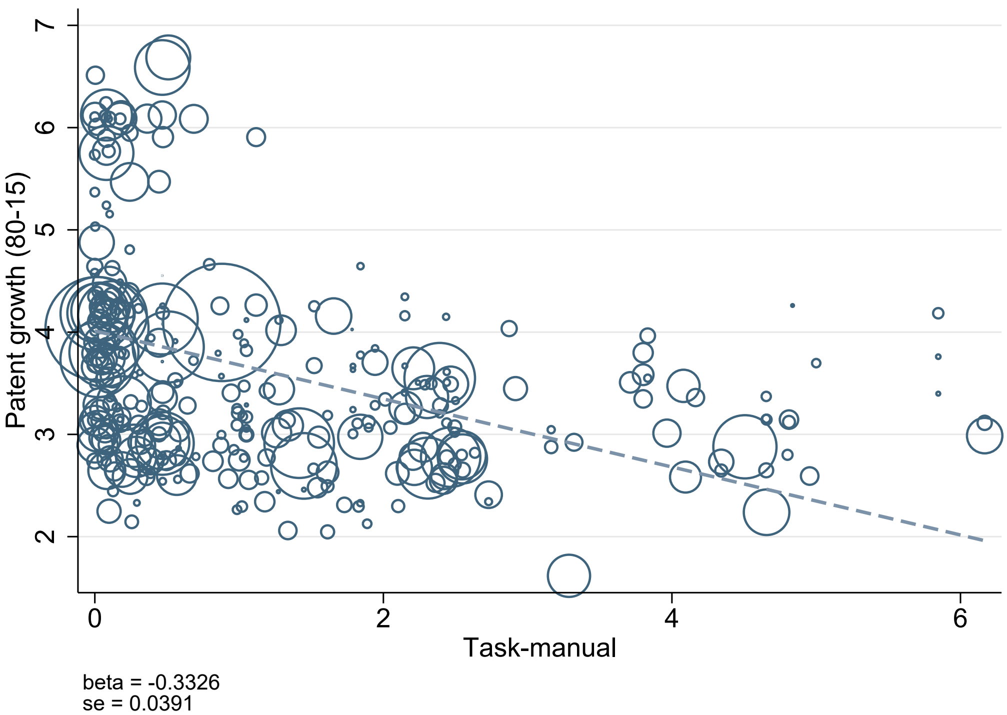
Notes: Task-abstract is occupation score for abstract task from Autor and Dorn (2013).

Figure 10: Patent growth and routine measure



Notes: Task-routine is occupation score for routine task from Autor and Dorn (2013).

Figure 11: Patent growth and manual measure



Notes: Task-manual is occupation score for manual task from Autor and Dorn (2013).

B Robustness check

B.1 OLS with industry fixed effect

We include sector fixed effects at NAICS 3 digit level to control sector-specific shocks. It still has a positive coefficient for wage and employment, which implies that the tool innovation is associated with the demand factor after controlling the sector fixed effects. However, within-occupation skill premium loses statistical significance, and skill intensity exhibits a negative estimate.

Table 14: OLS result with industry fixed effect

$$\Delta y_{o,s,1980-2015} = \beta_0 + \beta_1 \Delta \log p_{o,1980-2015} + \delta_s + \epsilon_{o,s}$$

	(1) Wage	(2) Emp.	(3) Skill Pre	(4) Skill Int.
patent	0.0300***	0.152***	0.00343	-0.00648***
	(0.00131)	(0.00704)	(0.00490)	(0.000710)
N	18,554	18,638	14,539	19,544
sector fixed	yes	yes	yes	yes

Standard errors in parentheses

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

Notes: s is sector at NAICS 3 digit level

B.2 IV with value-weighted patent

Since patents have heterogeneous values, we weight them by the monetary value measured in a way similar to Kogan et al. (2017). To be specific, we use the high-frequency stock market data and estimate the change in stock price around the patent grant day. Value-weighted patent measures do not change the result much. The tool innovation is still positively correlated with the demand for labor at the occupation level and the relative demand for skilled workers within an occupation.

Table 15: IV result with value-weighted patents

$$\Delta y_{o,s,1980-2015} = \beta_0 + \beta_1 \Delta \log p_{o,1980-2015} + \epsilon_{o,s}$$

	(1)	(2)	(3)	(4)
	Wage	Emp.	Skill Pre.	Skill Int.
patent (value)	0.136***	0.0981	0.0407***	0.0307***
	(0.0181)	(0.0556)	(0.0103)	(0.00479)
N	324	324	323	324
F (first stage)	214.0	214.0	213.3	214.0
sector fixed	no	no	no	no

Standard errors in parentheses

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

Notes: Patent is weighted by market value from Kogan et al. (2017)

C College Major Matching between ACS and DES

Each college major code in the ACS is translated into a row of the DES using the following crosswalk in Table 16.

Table 16: College Major Matching between ACS and DES

ACS DegField	ACS DegField Description	DES Description
11	Agriculture	Agriculture and Natural Resources
13	Environment and Natural Resources	Agriculture and Natural Resources
14	Architecture	Architecture and Related Services
15	Area, Ethnic, and Civilization Studies	Area, Ethnic, Cultural, Gender, and Group Studies
19	Communications	Communication, Journalism, and Related Programs
20	Communication Technologies	Communication Technologies
21	Computer and Information Sciences	Computer and Information Sciences
23	Education Administration and Teaching	Education
24	Engineering	Engineering
25	Engineering Technologies	Engineering
26	Linguistics and Foreign Languages	Foreign Languages, Literatures, and Linguistics
29	Family and Consumer Sciences	Family and Consumer Sciences/Human Sciences
32	Law	Legal Professions and Studies
33	English Language, Literature, and Composition	English Language and Literature/Letters
34	Liberal Arts and Humanities	Liberal Arts and Sciences, General Studies, and Humanities
35	Library Science	Library Science
36	Biology and Life Sciences	Biological and Biomedical Sciences
37	Mathematics and Statistics	Mathematics and Statistics
38	Military Technologies	Military Technologies and Applied Sciences
40	Interdisciplinary and Multi-Disciplinary Studies	Multi/Interdisciplinary Studies
41	Physical Fitness, Parks, Recreation, and Leisure	Parks, Recreation,Leisure, and Fitness Studies
48	Philosophy and Religious Studies	Philosophy and Religious Studies
49	Theology and Religious Vocations	Theology and Religious Vocations

ACS DegField	ACS DegField Description	DES Description
50	Physical Sciences	Precision Production and Industrial Arts
51	Nuclear, Industrial Radiology, and Biological Technologies	Biological and Biomedical Sciences
52	Psychology	Psychology
53	Criminal Justice and Fire Protection	Homeland Security, Law Enforcement, and Firefighting
54	Public Affairs, Policy, and Social Work	Public Administration and Social Services
55	Social Sciences	Social Sciences and History
56	Construction Services	Engineering
57	Electrical and Mechanic Repairs and Technologies	Engineering
58	Precision Production and Industrial Arts	Precision Production
59	Transportation Sciences and Technologies	Transportation and Materials Moving
60	Fine Arts	Visual and Performing Arts
61	Medical and Health Sciences and Services	Health Professions and Related Programs
62	Business	Business
64	History	Social Sciences and History

D Skill Premium Decomposition and Normalization

We show that the aggregate skill premium is determined by the wage and employment relative to the wage and the employment of skilled labor in the baseline occupation. Let O^s the set of occupations that are populated only by skilled workers and $O^{u,s}$ the set of occupations with both types of workers. $O^s \cup O^{u,s} = O$ is the entire set of occupations, given that skilled workers are working in all occupations. Section 4.1 contains explains other notations. The skill premium can be decomposed as in the following equation.

$$\begin{aligned}
\frac{\bar{w}^s}{\bar{w}^u} &= \frac{\frac{1}{N_s}\sum_o s_o w_o^s}{\frac{1}{N_u}\sum_o u_o w_o^u} \\
&= \frac{N_u}{N_s}\frac{\sum_{o\in O^{u,s}} s_o w_o^s}{\sum_o u_o w_o^u} + \frac{N_u}{N_s}\frac{\sum_{o\in O^s} s_o w_o^s}{\sum_o u_o w_o^u} \\
&= \sum_{o\in O^{u,s}} \frac{u_o w_o^u}{\sum_{o'} u_{o'} w_{o'}^u}\frac{s_o/N_s}{u_o/N_u}\frac{w_o^s}{w_o^u} + \frac{N_u}{N_s}\frac{\sum_{o\in O^s} s_o w_o^s}{\sum_o u_o w_o^u}
\end{aligned}$$

This tells us that if $O^s = \varnothing$, the overall skill premium is a weighted average of within-occupation skill premium, where the weight is the product between wage bill share of unskilled workers and the relative skill intensity.

Next, I show that the wage and the employment level relative to a wage of skilled labor in an occupation is sufficient in determining the overall skill premium. Let this benchmark occupation $o = 1$.

$$
\begin{aligned}
\frac{\bar{w}^s}{\bar{w}^u} &= \frac{\frac{1}{L_t^s}\sum s_{ot} w_{ot}^s}{\frac{1}{L_t^u}\sum u_{ot} w_{ot}^u} \\
&= \frac{L_t^u}{L_t^s} \frac{s_{1t} w_{1t}^s \sum \frac{s_{ot}}{s_{1t}} \frac{w_{ot}^s}{w_{1t}^s}}{s_{1t} w_{1t}^s \sum \frac{u_{ot}}{s_{ot}} \frac{s_{ot}}{s_{1t}} \frac{w_{ot}^u}{w_{ot}^s} \frac{w_{ot}^s}{w_{1t}^s}}. \\
&= \frac{L_t^u}{L_t^s} \frac{\sum \frac{s_{ot}}{s_{1t}} \frac{w_{ot}^s}{w_{1t}^s}}{\sum \frac{u_{ot}}{s_{ot}} \frac{s_{ot}}{s_{1t}} \frac{w_{ot}^u}{w_{ot}^s} \frac{w_{ot}^s}{w_{1t}^s}}.
\end{aligned}
$$

Thus, in counterfactual exercise, pinning down the relative wage and employment level is sufficient to calculate the counterfactual skill premium.

E Innovation and Price of Tools

This appendix shows that the innovation on tools used in each occupation is likely to reduce the price of tools. This reduction in the price of tools is a plausible channel through which innovation on tools improves the productivity of workers.

We use the commodity-level Producer Price Index (PPI) series from the Bureau of Labor Statistics (BLS) to measure the price of tools. The commodity code in the BLS data is converted to the UNSPSC code used in the O*NET data using the simple following algorithm.

1. Choose a commodity description in the PPI series.
2. From its commodity title, extract the words. Exclude generic terms such as 'Index', 'Data', and 'All'.
3. For each word extracted from step 2, define the specificity of each word as the inverse of the number of occurrences in the entire set from step 1. If a word

appears more than once in a title, count only once.

4. For each term in the set constructed step 2, count how many times it appears in the commodity title, the class title, the family title, and the segment title of each commodity code in the UNSPSC.

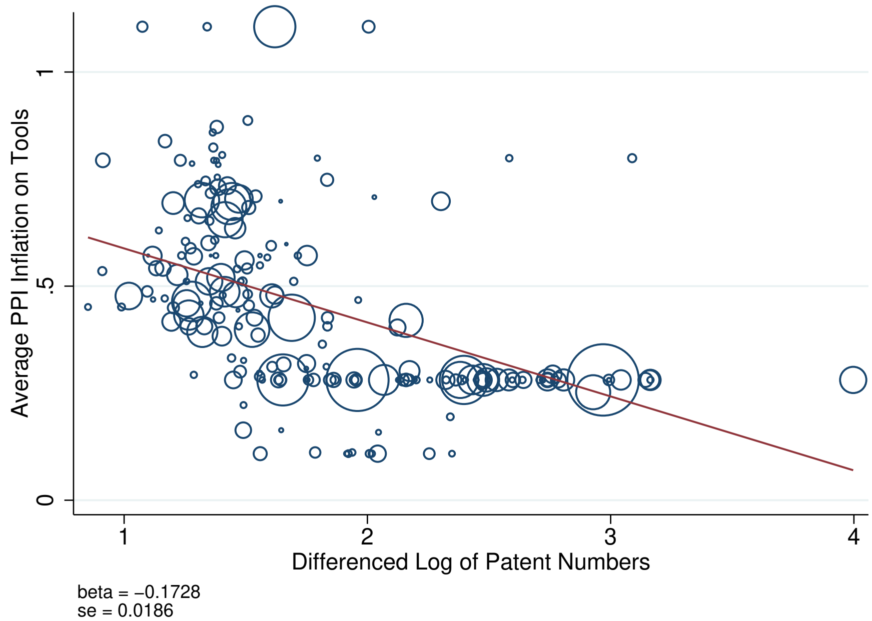
5. Sum the number calculated in step 4 over all the words extracted in step 2 after weighting them using the specificity measure calculated in step 3. This sum becomes the similarity score between the PPI title and the commodity item in the UNSPSC.

6. Repeat 1-5 for all commodities in the PPI series.

7. Match each commodity item in the UNSPSC to the PPI series that gives the highest similarity score.

Figure 12 shows that the number of patents on tools and the average PPI inflation of tools. Please note that each circle corresponds to a SOC occupation code and the size of the circle corresponds to the size of the occupation in 1980. Many occupations have the same average PPI inflation on tools because the matching rate is lower and many tools are matched to the same PPI series. Nonetheless, the number of patents at the occupation level is associated with a lower inflation rate of the PPI, which implies that the innovation on a tool reduces the price of that tool.

Figure 12: Innovation on Tools and Average PPI



We regress dependent variables of wage growth, changes in skill premium, and employment growth on the occupation-level PPI growth. The results from OLS regression are shown in Table 17. The estimates are overall negative, although they are often insignificant because of large measurement errors. We then use the upstream patent instrumental variable in the same regression specification and describe the results in Table 18. The instrumental variable approach gives a consistent estimate even with the presence of measurement errors. The estimate is now more significant and larger in magnitude on skill intensity and skill premium, consistent with the baseline reduced-form regression results in Section 3. A one percent decrease in PPI of tools increases the within-occupation skill intensity and skill premium by 0.36 and 0.18 percentage points in columns (2) and (4) of Table 18, respectively. These effects of lower tool prices from more innovation are robust to expanding the samples into occupations in each sector and adding sector fixed effects on college premium and skill intensity in columns (6) and (8) of Table 18.

Table 17: Impacts of PPI Growth: OLS

	(1) Wage	(2) Col. Pre.	(3) Emp.	(4) Skill Int.	(5) Wage	(6) Col. Pre	(7) Emp.	(8) Skill Int.
PPI	-0.0621	-0.0702	-0.339	-0.111***	-0.194***	-0.0122	-0.252***	-0.00822*
	(0.126)	(0.0478)	(0.230)	(0.0205)	(0.0101)	(0.0334)	(0.0412)	(0.00406)
N	145	145	145	145	13288	9660	13387	13387
sector fixed	no	no	no	no	yes	yes	yes	yes

Standard errors in parentheses

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

Table 18: Impacts of PPI Growth: IV

	(1) Wage	(2) Col. Pre.	(3) Emp.	(4) Skill Int.	(5) Wage	(6) Col. Pre	(7) Emp.	(8) Skill Int.
PPI	-0.266	-0.361***	0.246	-0.181***	-0.533***	-0.313***	0.0680	-0.0208*
	(0.213)	(0.0900)	(0.395)	(0.0358)	(0.0214)	(0.0679)	(0.0836)	(0.00824)
N	145	145	145	145	13288	9660	13387	13387
F (first stage)	77.00	77.00	77.00	77.00	4208.6	3054.8	4241.0	4241.0
sector fixed	no	no	no	no	yes	yes	yes	yes

Standard errors in parentheses

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