Identifying the Multiple Skills in Skill-Biased Technical Change

Seth Gordon Benzell*
Erik Brynjolfsson†
Frank MacCrory‡
George Westerman§

September 4, 2023

Abstract

We characterize the wage and employment growth of occupations by the importance of eight endogenously-identified orthogonal skills. We find supervision-intensive occupations saw significant increases in both wages and employment. Physically-intensive occupations saw significant decreases in occupational wages, and cooperation-intensive occupations saw employment growth. The increase in supervision-intensive occupational wages and decrease in physically-intensive occupational wages is more pronounced for occupations that use IT more intensely. We compare our results with Deming (2017) on social skills and find that wage and employment growth in social skill intensive occupations reflects two distinct trends: increasing wages for *supervision*-intensive occupations and increasing employment for *cooperation*-intensive occupations.

*Chapman University, Argyros School of Business and Economics, MIT Initiative on the Digital Economy, and Stanford Digital Economy Lab. Email: benzell@chapman.edu.
†Stanford and NBER
‡University of California Irvine, Merage School Of Business
§MIT

The authors thank participants at multiple seminars for valuable comments, the MIT Initiative on the Digital Economy for generous support, and Jessie Wang for her tireless research assistance.

1 Introduction

"It is not enough to be industrious; so are the ants. What are you industrious about?"

Henry David Thoreau

Epochal changes in the US occupational wage and employment distribution have taken place over the last forty years. As documented by Autor and Dorn (2013), among others, these changes have tended to polarize employment and wages. A common explanation for these changes in the wage distribution is skill-biased technological changes (SBTC). Under this theory, occupations with weak employment and wage growth are those where new technologies and the relevant skills are substitutes rather than complements.

In prior work, authors typically identify a set of occupational characteristics of interest, *e.g.* routineness, and build an index to measure them, applying economic theory as a lens for interpreting the empirical facts. In this paper, we explore an alternative, data-driven approach to identifying and measuring occupational skills. Specifically, we use an unsupervised iterated approach to a well-established machine learning technique for dimensionality reduction, exploratory factor analysis, to generate a characterization of US occupations. The resulting factors have clear interpretations and intuitive relationships to the wage distribution. By construction, they are also orthogonal, making it easier to interpret the effects of changing an occupation's reliance on one skill while leaving the others constant. The eight factors identified by this approach are characterized as occupational skill intensities in *physicality*, *technical sophistication*, *perception*, *supervision*, *cooperation*, *initiative*, *mathematics*, and *educating* dimensions.[1]

Using these factors, we evaluate how occupations of different types have seen their wages and employment evolve over the course of a decade. We find that physically-intensive occupations saw significant decreases in occupational wage, while supervision-intensive occupations saw significant increases in wage and employment. Cooperation-intensive occupations saw increases in employment, but no wage growth. The fact that supervisory positions saw increases in both

[1] We define these dimensions in detail on page 9.

wages *and* employment suggests that there was an increase in demand for these skills that exceeded supply.

The increase in supervision-intensive occupational wage and relative decrease in physical-intensive occupational wage is more important for jobs using IT capital more intensely. This is consistent with changes in relative demand for these skills due to IT capital deepening. The increase in employment for cooperation-intensive occupations is focused in jobs that use IT capital less intensely. This is consistent with industries and occupations that have been less impacted by technology providing refuge for cooperative workers who have seen their previous positions automated. These results hold whether IT capital use is measured at the occupation or industry level.

Our cooperation and supervision measures overlap significantly with the 'social skills' found to be important to labor demand shifts in previous research. In the final section of our paper, we compare our measures of social skills against those in Deming (2017). Although our supervision and cooperation measures are orthogonal by construction, both are correlated with Deming's measure of social skill. Including Deming's social measure in regressions explaining occupational wage and employment change, the wage effect of supervision and employment effect of cooperation remain large and significant, while social enters approximately neutrally. LASSO regressions confirm this result and provide further evidence that wage growth is concentrated in occupations requiring supervisory skills while employment growth is concentrated in occupations requiring cooperative skills. However, when increasing the LASSO penalty for additional regressors to the level that only a single regressor remains, Deming's indexes are shown to be the single most important summary-level measure. These results confirm the value of both approaches: Deming's theory-driven index identified a skill construct with large explanatory power, and our data-driven approach independently identified measures that could be considered more fundamental underlying components.

2 Background

How does technology affect labor demand? Studies of Skill-Biased Technical Change (SBTC) have used a variety of measures of skill, which have increased in sophistication and explanatory power over time. Early empirical studies of SBTC and task-based labor models focused on the wage premium for high-skilled individuals. The share of workers with a college education increased dramatically in the latter half of the 20th century. Yet, across nations, the wage premium for the educated stayed roughly constant or increased over this interval (Berman et al., 1998). This is consistent with high-skilled individuals being disproportionately favored by technological change. High-skill individuals are typically defined as those with college degrees (such as in the capital skill complementarity literature, see for example Krusell et al. (2000)) and high-skill occupations as those with high wages at the beginning of the period under consideration. There is also theoretical and empirical evidence of technology playing a role in the rise of the 1%'s share of income. Rosen (1981) forecasted that economies of scale enabled by new technologies would increase inequality. The increase in top income shares has affected top earners in all industries (Kaplan and Rauh, 2013). The increasing income of superstar workers can be explained by their increasing importance in output, and may contribute to stagnant median wages, low interest rates, and slow economic growth (Benzell and Brynjolfsson, 2019).

More recent papers have diagnosed labor demand *polarization* as a consequence of technological change. From 1980 to 2005, occupations which were highly compensated in 1980 saw disproportionate growth in both wage and employment. Interestingly, the same was true of occupations in the bottom third of the wage distribution. Those occupations in the middle, in contrast, saw little employment or wage growth. Autor and Dorn (2013) find that areas that specialized historically in industries which use routine tasks intensively (such as manufacturing) saw larger increases in wage and employment polarization. This finding remains after controlling for the offshorability of jobs. They follow Autor et al. (2003) in attributing this to technological advances, in particular information technology, which tend to substitute for people in routine jobs.[2] Michaels et al. (2014) provide cross-country evidence linking IT investment to the polarization of wages, and also utilize (Autor et al., 2003)'s routine/cognitive occupational classifications. Each of these papers uses indexes of occupational characteristics to characterize occupations as routine or non-routine.

While it can be convenient to think of skill as existing on a single dimension, in practice there are multiple dimensions of skill. Accordingly, several papers have tracked employment and wage changes across several different occupational or worker characteristics. To do so, they have used occupational

[2]Other papers looking at the role of technological change in wage polarization in developed countries are Acemoglu (1999), Goos and Manning (2007), and Goos et al. (2010, 2014).

task intensity measures which are typically constructed as averages of a handful of occupational characteristic scores from a government survey or occupational classification system. The data most commonly used for this purpose are the US's Dictionary of Occupational Titles (DOT) which became the Occupational Information Network (O*NET) in 1998.

There are many examples of papers in this vein. For example, Acemoglu and Autor (2011) propose five measures: non-routine cognitive-analytical, non-routine cognitive-interpersonal, routine-cognitive, routine-manual, non-routine manual-physical, and offshorability. Each is an index of three to seven author-selected and normalized O*NET occupational characteristic scores.

One example of this approach is in the definition of Routine Task Intensity (RTI) index in Autor and Dorn (2013). This measure has been used in many papers, including Goos et al. (2010), Cortes et al. (2018) Tuzel and Zhang (2019). Each of these papers measures routineness as the sum $RTI_k = \ln\left(T_k^{Routine}\right) - \ln\left(T_k^{Abstract}\right) - \ln\left(T_k^{Manual}\right)$. $T_k^{Routine}$ is the average of two DOT measures of the occupation: The degree it requires "set limits, tolerances and standards," and "finger dexterity." $T_k^{Abstract}$ is the average of DOT measures "direction, control and planning" and "GED math." The final measure, T_k^{Manual}, is the occupation's DOT score for "eye-hand-foot coordination." Clearly, each of the chosen components of the index are only imprecisely connected to the concept to be investigated.[3]

Deming (2017) uses a different measure of routineness. It is the average of two items in O*Net: (1) "how automated is the job?" and (2) "how important is repeating the same physical activities...over and over, without stopping, to performing this job?" Although item (2) clearly reflects routineness, item (1) is indicative only to the extent that automation standardizes the tasks that remain for people to perform in the occupation.

While these studies have greatly enhanced our understanding of the changing demand for skills, the construction of their underlying indexes is potentially problematic. This is due in part to the difficulty of working with occupational characteristic data. O*NET was not directly designed for this type of analysis, and its extremely high dimensionality (with over 400 separate rating scales cov-

[3]There are also conceptual problems in creating the RTI index as the difference of these logged components. Consider a pair of occupations, where the second occupation has twice as high a score in each of the model's elements (i.e. the second occupation scores twice as high as the first on each of $T_k^{Routine}$, $T_k^{Abstract}$, and T_k^{Manual} skills). Then the second occupation would have a much smaller RTI-index than the first, despite scoring much higher in routineness. This seems a counterintuitive result, even setting aside the loose connection between the intermediate concepts and the underlying measures.

ering aver 1,000 occupations) has daunted many researchers. Leading scholars of the task-based labor model have pointed out that building an index from this data introduces many researcher degrees of freedom. They are therefore only moderately trustworthy "because their complexity and opacity places little discipline on how they are applied and interpreted" (Autor, 2013).[4]

Our paper extends the literature that distinguishes different skills by creating occupational characteristic measures using a well-understood unsupervised machine learning technique for dimensionality reduction. For raw data, we use more than 100 raw occupational characteristic scores from O*NET. The occupational characterization produced by this procedure is reduced to eight dimensions through a method that is repeatable and independent of author intervention or subconscious bias. The dimensions correspond clearly to the physical, technical sophistication, perception, supervision, cooperation, initiative, mathematics, and education [educating or educational] task intensity of an occupation. Notably, some characterizations of an occupation often considered important in the literature, such as routineness, are absent from this list, while others, like social skill intensity, are captured by two factors.

Deming (2017) measures the social and cognitive task intensity of occupations using indices of O*NET characteristics. It measures social orientation of individuals using an index of NLSY survey questions that are related to a person's innate socialness. It finds evidence that returns to social skills have increased in the US, especially for occupations that require a high degree of both social and cognitive skill. To measure individual-level social skill, Deming (2017) uses self-reported sociability, the number of clubs participated in at high school, participation in high school sports, and an extroversion factor from a big five personality inventory (each from the NLSY). In Deming and Kahn (2018) the authors hand-sort a list of words that commonly appear in job postings into one of ten non-exhaustive skill categories. Both within and across occupations they find jobs requiring cognitive and social skills demand higher wages, and they additionally find evidence of complementarity between these skills. Although interesting, these measures can potentially fall prey to subconscious

[4]the full statement in (Autor, 2013) places the quote more fully in context: "In practice, this means that researchers who wish to use these databases as sources for task measures are essentially required to pick and choose among the plethora of scales available, a problem that is much more severe for O*NET than for DOT. Researcher discretion again becomes paramount in this data construction process, and some transparency is inevitably lost. While I have found that task measures distilled from DOT and O*NET can serve as powerful proxies for occupational tasks, I am at best only moderately comfortable with these tools because their complexity and opacity places little discipline on how they are applied and interpreted."

biases or simple cognitive limits and other forms of bounded rationality (Simon 1955) making sense of hundreds of potential items and combinations.

As an alternative, in this paper, we show that a data driven approach distinguishes socialness into supervisory and cooperation tasks. The paper in the literature most similar to ours is Green (2012), which uses exploratory factor analysis on 35 self-assessed skill items from a survey of United Kingdom workers to identify eight skill categories. Two of the categories, Numeracy and Physical Skills line up closely with skills we identify. However, our paper uses a more detailed underlying taxonomy of skills on a US dataset that can be linked directly to the large amount of US labor data used extensively in the economic literature. We focus on the role of skill dimensions in the change in wages and employment of occupations, whereas Green assessed the extent to which the perceived skill content of jobs has changed over time.

Another related paper, Alabdulkareem et al. (2018), normalizes O*NET occupational skill intensity scores using revealed comparative advantage (RCA) and then computes the complementary of each pair of skills as the minimum of the conditional probabilities of the two skills being used by the same occupation. Complementary clusters between pairs of skills reveal a bimodal distribution characterized as more or less socio-cognitive. Using this unidimensional measure, the authors find that more socio-cognitive occupations have higher wages, even after controlling for routineness and education level. They also find that connections between occupational skill usages between occupations predict occupational mobility between these occupations.

3 Data and Skill Measurement

We draw our civilian employment and wage figures from the Bureau of Labor Statistics' Occupational Employment Survey (OES), using the annual statistics published at www.bls.gov/oes for years 2006 through 2016.[5] For each occupation, the OES reports employment, average wage, and median wage by industry. Our main results are based on employment and median hourly wages.[6]

Our underlying skill data is derived from the Department of Labor's O*NET dataset (available at www.onetcenter.org). This database provides empirical

[5]Wages and employment during this time period were punctuated by the Great Recession (Akst, 2013).

[6]All wages are deflated to constant 2006 dollars using the January Consumer Price Index for all urban households (CPI-U) published at www.bls.gov/cpi.

data on the content of occupations in the US economy. It includes information about characteristics of the job itself (*e.g.* its typical tasks, level of responsibility, and exposure to hazards) as well as on the people who perform the job (*e.g.* their abilities, skills, and interests).

Our analysis begins with O*NET's evaluation of the importance of four categories of occupational characteristics: Abilities (1.a), Work Activities (4.a), Skills (2.a and 2.b) and Work Styles (1.c). This encompasses all O*NET importance measures except those categorized as Knowledge (2.c). The 142 elements meeting these criteria reflect highly trained labor experts' assessments of the importance of each skill to each occupation. We focus on occupational characteristics in a base year due to a concern that rankings are incomparable across years.[7] The Standard Occupation Code system was updated in 2006, by which time virtually all occupations in O*NET had Work Style ratings. We therefore chose the December 2006 O*NET release for all occupation characteristics.[8] After cleaning, O*NET has 798 scored occupations in 2006.

We use an unsupervised iterated approach to exploratory factor analysis, itself an unsupervised machine learning technique, to summarize occupations by their skill intensity along several dimensions. The procedure begins by performing a principal-component factoring of the importance scores for retained O*NET questions across occupations. Orthogonal varimax rotation is then applied to the loading matrix. This rotation maximizes the variance of the squared loadings within factors, while making sure all factors are orthogonal. Subsequently, any O*NET score with a loading with an absolute value of less than .50 in all factors is discarded. Any O*NET score with a weighing of more than .40 in at least two factors is also discarded. After these criteria are implemented, the process is repeated, now with a minimum loading of .51 (the maximum cross loading remains fixed at the .40 threshold). The procedure is iterated until the minimum loading reaches .70. This procedure creates several orthogonal factors, with each retained O*NET characteristic contributing primarily to one and only one factor. The iterated nature of raising the threshold allows us to exclude cross-loading items without dropping important items

[7] Up until 2007, skill data was collected primarily from incumbents employed in the focal professions. Updates from 2008 onward also collect skill data from labor analysts. Furthermore, O*Net analysts update only a subset of occupations each year, meaning that longitudinal analysis of changes in the skill content of the full set of occupations is not meaningful from year to year.

[8] Extending our analysis to periods before 2006 is difficult because our analysis leverages improvements introduced that year in how occupational scores were determined, and we measure occupational skill intensity at the *beginning* of the sample period.

due to correlation with unimportant items. The specific thresholds we utilize, *i.e.* the .70 minimum factor loading and .40 maximum cross loading, follow Hair et al. (2013) page 114. This cutoff ensures that the factor analysis has a well-defined structure.

Eight factors are retained which summarize the occupation. These factors are listed from most to least important in terms of explaining variation in occupations' O*NET scores. Alongside each factor are listed their primary constituent O*NET characteristics.

- **Physicality (PHYS):** Arm-Hand Steadiness; Multilimb Coordination; Static Strength; Dynamic Strength; Trunk Strength; Stamina; Extent Flexibility; Gross Body Coordination; Gross Body Equilibrium; Performing General Physical Activities; Handling and Moving Objects; Manual Dexterity
- **Technical Sophistication (TECH):** Repairing and Maintaining Electronic Equipment; Technology Design; Equipment Selection; Installation; Operation Monitoring; Operation and Control; Troubleshooting; Quality Control Analysis; Systems Analysis
- **Perception (PERC):** Speed of Closure;[9] Flexibility of Closure; Perceptual Speed; Selective Attention; Far Vision; Hearing Sensitivity; Auditory Attention
- **Supervision (SUPV):** Scheduling Work and Activities; Coordinating the Work and Activities of Others; Developing and Building Teams; Guiding, Directing, and Motivating Subordinates; Staffing Organizational Units; Monitoring and Controlling Resources
- **Cooperation (COOP):** Cooperation; Concern for Others; Social Orientation; Self Control; Stress Tolerance
- **Initiative (INIT):** Achievement/Effort; Persistence; Initiative; Independence; Innovation
- **Mathematics (MATH):** Number Facility; Mathematical Reasoning; Mathematics
- **Teaching and Education (EDUC):** Learning Strategies; Instructing

[9] In O*NET's terminology, "closure" refers to pattern recognition.

Although we gave labels to each these factors, each factor's contents and the total number of factors arise directly from the raw data.[10] After generating scores for the different skill intensities of different occupations, the data are merged with BLS employment data. BLS employment data are used at the occupation-three digit industry-year level. While the two sources use very similar employment categorizations, a difficulty arises from the fact that O*NET uses a finer level of granularity than BLS. For example, O*NET lists seven kinds of employment officers, while BLS has only one type. For BLS occupations that correspond to more than one O*NET occupation, we take the raw average of the O*NET occupation factor scores when merging into BLS data.

A final source of data used for this analysis is on the use of information technology capital by industry. Our industry IT capital intensity measure is from the BEA current-cost net capital stock of private nonresidential fixed assets.[11] We define the IT capital intensity of an industry as the ratio of IT capital to the current total capital stock. The types of capital considered IT are: Computers, mainframes and accessories (EP1), software (ENS), communications equipment (EP2) and communications structures (SU2). The BEA reports the capital stock for most industries at the three digit level. However, for a large subset of industries for which we have BLS data, the BEA data is at a higher level of aggregation (*e.g.* the total capital stock for a pair of three-digit NAICS industries is reported together). In our main analysis, we assign the same IT capital intensity to sets of industries that are combined in the BEA data.

After this last addition we drop from the data occupation-industry pairs without median wage or employment data in 2006 or 2016. These restrictions produce a final data set with 88 industries and 537 occupations for analysis.

Table 1 gives factor scores for several occupations of interest. Table 2 gives employment weighted percentiles. Dishwasher and CEO are among the lowest and highest compensated occupations, respectively. Physical skill is slightly more important for dishwashers than the average occupation, but the position requires few other skills in abundance. CEOs, on the other hand, need strong skills in all factors except for physical skill and perception. Landscape architecture is similar to CEOs in requiring high supervision skills. However, unlike CEOs, it is less important for landscape architects to develop cooperation, ed-

[10]While the above lists the most important characteristics within each factor (*i.e.* those with loadings of more than .70), all factors are a function of all retained O*NET elements. However, because all elements with cross-loadings of more than .40 are eliminated, non-primary elements contribute relatively little to an occupation's factor score.

[11]Retrieved from https://apps.bea.gov/national/FA2004/Details/Index.htm.

Table 1: Occupational characteristic scores for occupations of interest.

	PHYS	TECH	PERC	SUPV	COOP	INIT	MATH	EDUC
Dish Washer	0.75	0.11	-0.78	-0.15	-0.62	-1.76	-1.08	0.06
Chief Executive	-0.80	0.57	-0.82	1.67	0.94	1.04	1.82	-0.26
Landscape Architect	-0.95	-0.83	0.79	2.00	-2.10	-0.40	0.17	-1.10
Police Officer	1.24	-0.65	1.49	0.30	0.88	1.03	-0.64	0.24
Detective	0.54	-0.39	1.44	-0.33	0.21	0.57	-0.28	0.78
Chemist	-0.74	2.12	-0.89	-1.35	-0.51	1.06	0.19	0.09
Economist	-1.32	-0.82	-0.51	-0.33	-1.32	1.00	1.65	-0.57

Table 2: Occupational characteristic score percentiles for occupations of interest.

	PHYS	TECH	PERC	SUPV	COOP	INIT	MATH	EDUC
Dish Washer	67.7	77.5	32.3	58.2	15.0	7.8	13.4	77.3
Chief Executive	25.5	85.4	30.4	94.0	73.7	93.8	92.5	58.5
Landscape Architect	19.0	37.0	86.4	96.3	.3	42.3	47.1	17.4
Policeman	86.7	40.7	96.8	73.7	70.5	93.1	24.5	81.9
Detective	56.8	57.0	96.7	50.0	42.9	82.8	34.8	90.9
Chemist	26.9	99.6	27.5	10.7	17.4	94.0	47.2	78.5
Economist	3.1	37.1	42.0	50.1	4.0	92.9	87.1	38.5

ucation and initiative skills. This is intuitive. While CEOs must create a new vision while working with near equals, the supervision of landscape architects is more top down and within well-defined constraints. Police and detective work require similar skill sets, except police need to be more physical and detectives require more math, technical and education skills. Chemists and economists are another interesting pair of occupations to contrast. While both types of scientists require a good amount of initiative, a chemist job requires more mastery of equipment and technology while economist jobs require more math.

Table 3 regresses the intensity of different occupational characteristics on the wage by occupation and industry. The specification is

$$Y_{j,i,t} = \sum_{F=1}^{8} \beta_F F_j + X_i + \epsilon_{i,j,t} \tag{1}$$

where F is skill factor, i is industry, j is occupation, t is year, and X_i are industry fixed effects. All years of data, from 2006 to 2016, are included. In the first set of columns, $Y_{j,i,t}$ is the median wage in the occupation-industry. In the second pair, a version of wage skewness is the outcome. It is measured as the average wage less the median wage, divided by the median wage, *i.e.*

$$Skew = (W_{Ave} - W_{Median})/W_{Median} \tag{2}$$

Table 3 on page 12 shows that physical-intensive occupations are significantly[12] lower paid. To give a sense of the size of the effect, the occupation 'lawyer' has a physicality score of -1.52, and stonemasons have a physicality score of +2.53. Based on this factor alone, focusing on the specification without industry fixed effects, we would therefore expect the median lawyer to be paid \$16.93 more per hour than the median stonemason.

Cooperation-intensive occupations are also significantly lower paid when not controlling for industry. The fact that the magnitude of this characteristic's point estimate is significantly reduced when controlling for industry could be due to this form of employment being concentrated in the low paying retail industry.

Occupations intense in other factors are more highly compensated. Occupations which are more intensive in supervision, initiative, math and education are all significantly higher paid even after controlling for industry.

It is critical to remember that the coefficients estimated in table 3 *should not* be directly interpreted as returns to individual ability. It would be absurd to deduce that an individual who saw their physical skills increase should expect to see their wage decrease. Rather, the regression reports how the equilibrium wage of an occupation varies with the occupation's characteristics.

As is typical with hedonic regressions, many forces come to balance in the current equilibrium. To paraphrase Alfred Marshall, asking whether supply or demand sets a market equilibrium is like asking which blade of a scissor does the cutting. Occupations with the highest wage require skills that are rare or are hard to acquire. The relationship between occupational skill intensity and wages mostly corresponds to commonsense intuitions about particularly scarce and valuable skills. Physical, technical operation, and cooperation skills are seemingly abundant or easy to instill. A large percentage of US individuals in the US labor force have sufficient physical strength and dexterity necessary to perform adequately in physical, menial occupations. The basic technical skills involved in most types of troubleshooting, quality control, and installing and repairing equipment are relatively easy to train in the average worker. On the other hand, supervision, education, math and initiative may be more challenging to impart.[13]

[12] Throughout the paper we report standard analytical clustered standard errors. However, our data is comprehensive in the sense that our US occupations covers nearly all US employment. In this setting, Abadie et al. (2020) shows normal analytical standard errors are conservative.

[13] Other factors may also contribute to the correlation between wage and occupational task

	(1) Median Wage	(2) Median Wage	(3) Wage Skewness	(4) Wage Skewness
Physical	-4.181^{***}	-3.521^{***}	-0.003	-0.002
	(0.512)	(0.453)	(0.004)	(0.003)
Technology	0.604	-0.067	-0.019^{***}	-0.020^{***}
	(0.570)	(0.608)	(0.003)	(0.003)
Perception	0.914^{*}	0.486	-0.010^{*}	-0.006^{*}
	(0.371)	(0.340)	(0.004)	(0.003)
Supervision	3.967^{***}	4.063^{***}	0.006	0.007^{*}
	(0.630)	(0.657)	(0.004)	(0.003)
Cooperation	-1.649^{**}	-1.179^{*}	-0.001	0.005
	(0.543)	(0.592)	(0.004)	(0.003)
Initiative	4.685^{***}	4.449^{***}	0.020^{***}	0.018^{***}
	(0.503)	(0.447)	(0.005)	(0.004)
Math	1.802^{***}	1.671^{***}	0.005	0.000
	(0.405)	(0.437)	(0.004)	(0.003)
Educating	2.327^{***}	2.524^{***}	0.005	0.012^{***}
	(0.526)	(0.533)	(0.004)	(0.003)
Industry FE		X		X
Constant	21.508^{***}	20.874^{***}	0.076^{***}	0.078^{***}
	(0.650)	(0.573)	(0.004)	(0.004)
Observations	112,451	112,451	112,451	112,451

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

Table 3: Regression of occupation-industry median hourly wage (real 2006 dollars) and wage skewness on occupational skill intensity with and without industry fixed effects. All years pooled. Observations weighted by employment. Standard errors clustered at the occupation level in parentheses.

Table 3 also reports the skewness of the wage of occupations intensive in different factors. Our measure of skewness of an occupation-industry-year is defined as the average wage less the median wage divided by the median wage. Taking into account the constant term, almost all types of occupations are positively skewed. Occupations intensive in technical and perception skills are less skewed than average. Initiative intensive and, after controlling for industry, supervision and education intensive occupations are more positively skewed than the average. This is the consistent with papers that have found a large right tail for the most productive occupations, including supervision positions (see, for example, Brynjolfsson and Saint-Jacques (2015)).

These results are also consistent with intuitions about which types of occupations may have Pareto returns to excellence, and which have a long tail of mediocre workers with a cap on productivity. When working in a factory or driving a truck – occupations intensive in technical and perception skills respectively – productivity is limited by the technology being operated. It is hard to be two times as productive as the median worker in these tasks. On the other hand, individuals most skilled at education, initiative, and supervision tasks can be dramatically more productive than the median worker. The demand for top workers in these occupations more resembles the market for superstars described by Rosen (1981). Recall also that the highest paid occupations lack median and average wage data in some or all industries. Their wages are censored *because* they are so high. Chief executives are one example. If data were available for these occupations, the median and skewness of wages for high supervision intensity occupations would likely be even higher.

4 Identifying the Multiple Skills in SBTC

We are primarily interested in seeing what role these distinct skills play in SBTC. We do so by measuring how wage and employment in occupations of different skill intensities have changed over time. We begin with the following specifications

intensity. These include things like the geographical dispersion of occupations. Some regions might have a higher concentration of a certain type of occupation but a lower cost of living as well. If occupations of specific tasks are particularly attractive (i.e. have a positive compensating differential) then workers may accept a wage penalty to accept jobs intensive in that task. Future expectations may also play a role. As suggested by Edin et al. (2018) and others, people may sort into occupations based on whether they believe the occupation will be highly compensated in the future. If that is the case, then the wage associated with given skills may not only be a function of their current value, but their future value as well.

$$W_{i,j,2016} - W_{i,j,2006} = \sum_{F=1}^{8} \beta_F F_j + X_i + \epsilon_{i,j} \tag{3}$$

and

$$ln(emp_{i,j,2016}) - ln(emp_{i,j,2006}) = \sum_{F=1}^{8} \beta_F F_j + X_i + \epsilon_{i,j} \tag{4}$$

Table 4 reports the results of these regressions. Both wage and employment increased faster for supervision intensive occupations. Before industry controls, cooperation jobs saw faster employment growth and physical jobs slower wage growth. The specification with industry controls eliminates the significance of the latter results. This could be due to overcontrolling as cooperation and physically intense jobs are concentrated in a small subset of industries.

The point estimates on these effects are moderately sized. An example of an occupation with a low supervision score is credit analyst with -2.08. An occupation with supervision intensity score of $+1.93$ is 'first line supervisors of police and detectives.' Our results indicate that police managers are predicted to have seen \$0.899 dollars per hour in additional median hourly wage growth as a result of their greater supervision intensity and 16.4 percent faster employment growth.

Figure 1 plots the coefficients in Table 4 for the specifications without industry fixed effects. In the bottom panel, supervision and cooperation are the pair of characteristics with positive point estimates for both wage and employment coefficients. This is consistent with an increase in demand for occupations intensive in these tasks. Such an increase in aggregate demand would tend to raise both employment and wages for these types of occupations. However, the wage increase for cooperative occupations is not significant, suggesting that a concomitant increase in supply for this occupation limited wage gains. For supervision intensive tasks, the inference that demand has gone up is particularly strong, as both point estimates are significant.

Figures 2 and 3 report how the wage premium and employment gains have evolved by occupational skill intensity over time, for the four skills with the largest changes. They plot estimates equivalent to those in equation 3 and 4 except that the difference estimated is between wages and log employment in 2006 and different annual end-points. For this reason, the plotted points and confidence intervals for year 2016 are the same as those reported in table 4 for the specification without industry fixed effects.

	(1) ΔWage	(2) ΔWage	(3) Δln(Emp)	(4) Δln(Emp)
Physical	-0.232^{*}	-0.125	0.005	-0.009
	(0.100)	(0.093)	(0.023)	(0.024)
Technology	-0.022	-0.026	-0.025	0.012
	(0.096)	(0.104)	(0.017)	(0.019)
Perception	-0.066	-0.100	0.003	0.019
	(0.066)	(0.055)	(0.013)	(0.013)
Supervision	0.217^{*}	0.176^{*}	0.041^{*}	0.033^{*}
	(0.094)	(0.089)	(0.017)	(0.016)
Cooperation	0.025	-0.037	0.067^{***}	0.029
	(0.084)	(0.096)	(0.018)	(0.022)
Initiative	-0.008	0.010	-0.025	0.002
	(0.086)	(0.080)	(0.019)	(0.018)
Math	-0.009	0.066	-0.011	0.005
	(0.084)	(0.086)	(0.021)	(0.021)
Educating	-0.042	-0.092	0.044	0.034
	(0.128)	(0.125)	(0.032)	(0.028)
Industry FE		X		X
Constant	0.491^{***}	0.429^{***}	-0.025	0.008
	(0.107)	(0.101)	(0.019)	(0.018)
Observations	10,675	10,675	10,675	10,675

Standard errors in parentheses

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

Table 4: Regression of change in median hourly wage and log employment by occupation-industry on 2006 occupational characteristics. Wage observations weighted by 2006 employment. Standard errors clustered by occupation.

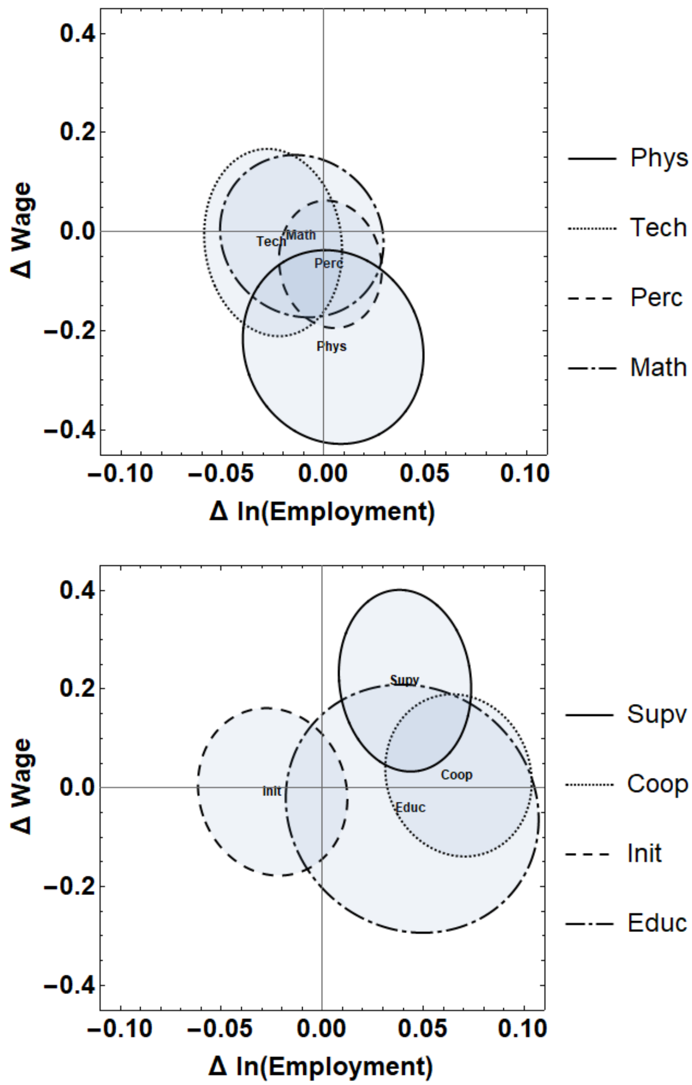

Figure 1: Scatterplot of the coefficients estimated in Table 4 with 95 percent confidence intervals. Tilt of the confidence ellipsoids is due to correlation of the $\epsilon_{i,j}$ across models.

These figures reveal that the decrease in the wage for physical occupations decreased over the entire interval, while the increase in wages for supervision-intensive occupations occurred mostly in the earlier and later years of the sample. These patterns may be driven in part by the impact of the Great Recession as well as more secular trends. The increase in cooperation-intensive employment occurred entirely before 2010, while the increase in supervision-intensive employment increased over the entire period, with a pause (mirroring the pause in wage growth) in the middle of the period.

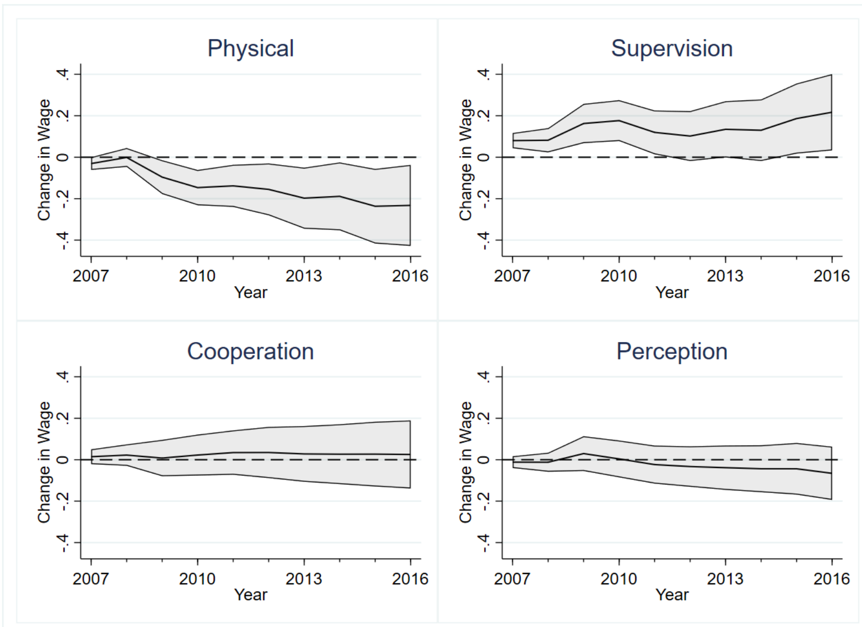

Figure 2: Plots of estimated coefficients from equation 3, with $W_{i,j,2016}$ replaced with the wage in an alternate year indicated on the x-axis. Observations weighted by 2006 employment. 95 percent confidence interval with SEs clustered by occupation displayed.

Table 5 reports how employment changed for occupations of different skill intensities in aggregate. Occupations are binned by quartile, using 2006 employment levels, and then the change in employment controlling for average employment growth was calculated. Occupations in the top quarter of physicality lost about 1.5 million excess net-jobs. Occupations in the bottom quartile of supervision and cooperation added .5 million and 2.3 million fewer net-jobs respectively than unbiased employment growth would have predicted. Top quartile occupations in these characteristics added over 2.2 and 1.1 million excess net jobs, respectively.

While employment growth was not significantly biased towards or against math, tech, or perception skills, the variance of skill intensity moved for these occupations in different ways. Most employment growth was in occupations

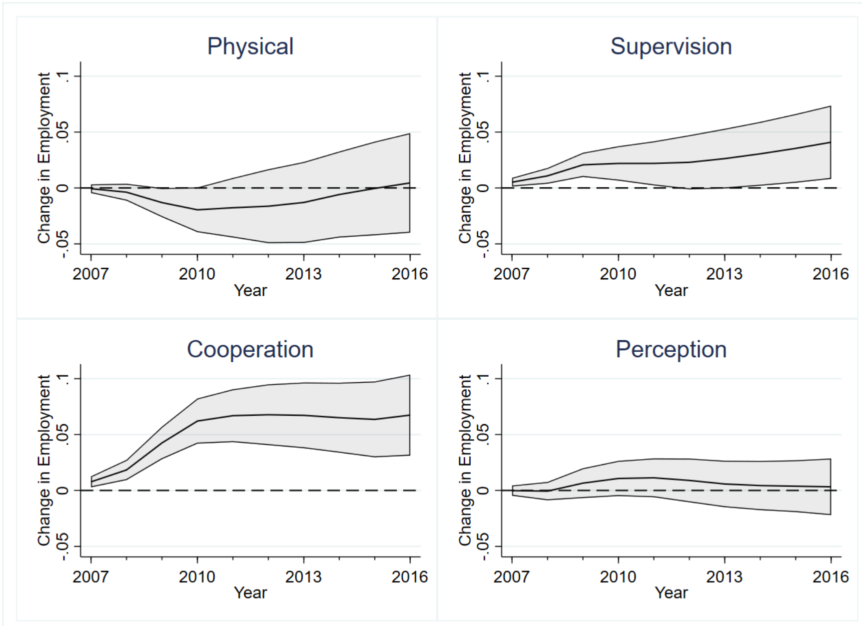

Figure 3: Plots of estimated coefficients from equation 4,with $ln(emp_{j,i,2016})$ replaced with log employment in an alternate year indicated on the x-axis. 95 percent confidence interval with SEs clustered by occupation displayed.

requiring moderate amounts of math and tech. Meanwhile, employment growth was polarizing along the dimension of perception intensity, with disproportionate job growth in the most and least perception intense occupations.

Overall, these changes are consistent with technological change reducing demand for physical-intensive jobs, with some workers moving up the skill chain into supervision intensive occupations, and others sorting into cooperation intensive jobs with less wage growth. That being said, it is important to remember that we cannot rule out alternative interpretations of these coefficients. In particular, sorting of individuals with different levels of productivity into occupations might break the above interpretations. For example, suppose – holding demand for tasks fixed – those with very high latent general productivity began enjoying supervision positions relatively more. The impact on the average wage in high supervision occupations would be the combination of a downward force – a downward sloping demand curve for supervision tasks – and an upward force – the higher average productivity of those who wish to sort into supervision tasks. Similarly, employment change in these occupations would tend to

Excess Employment Growth for Occupations in the Top Quartile of Each Skill	
Physical	-1,533,103
Technology	-889,513
Perception	350,488
Supervision	2,244,428
Cooperation	1,158,518
Initiative	-952,733
Math	-807,213
Educating	594,968

Table 5: Change in occupational employment for occupations in the top quartile of each skill above unbiased employment growth, rounded to the nearest integer. Occupations binned into percentiles weighted by 2006 employment. 4,988,450 additional net-jobs were added from 2016 versus 2006, so total employment growth in occupations in the top quartile of a skill factor can be determined by adding 1,247,112 to the number in the table. Data is restricted to the occupations and industries used in the main regressions (*e.g.* table 4).

increase because of the increased attractiveness of the job but decrease due to the increased productivity of those switching to that form of employment.

4.1 The Changing Importance of Skills and IT

We now turn our attention to how changes in the wages and employment occupational skill characteristics are mediated by technology. The following tables rerun the specification in equation (1) with the modification that occupations or industry be in the bottom/top 40 percentiles of some characteristic.[14]

Table 6 divides occupations by their computer usage intensity. Computer intensive occupations which are high in supervision requirement tended to see faster wage growth, while occupations high in cooperation saw faster wage growth when the occupation involves less computer use. One interpretation is that supervision is complemented by information technologies that allow workers to have better information or extend their influence more broadly, while the basic skills involved in cooperation are less likely to be complemented by deepening IT capital. Most dramatically, while most computer-intensive occupations experienced wage growth (as indicated by the large constant term in the column), physically intensive ones did not keep up with the others. Meanwhile

[14] Weighing by occupational employment, using 2006 occupational employment.

the relative wage for physical skills in less-computerized occupations grew.This is consistent with automation due to robotics or other computer-controlled operations.

Table 7 divides industries by IT capital intensity in 2016. This is defined as the current cost of IT forms of capital as a ratio of total nonresidential fixed private investment in the industry. IT capital includes computer or server hardware and software.[15] More computer intensive industries saw smaller cooperation wage and employment growth.[16]

The main lesson to draw from tables 6 and 7 is how technology differently interacts with supervision and cooperation. The overall increase in wages for supervision-intensive occupations is concentrated in occupations with high computer use. On the other hand, the overall increase in employment for cooperation-intensive occupations is driven by low computer use occupations and low IT capital intensity industries. This is consistent with our hypothesis that technological change is boosting the abilities and wages of managers especially in high-tech industries, while individuals who only have cooperation skills are finding refuge in low-tech industries and occupations.

5 Unpacking Social Skills

Deming (2017) showed that social skills are increasingly important in the labor market. According to his index of social skill intensity, the share of employment in occupations requiring high levels of social interaction grew by nearly 12 percentage points.

Our analysis, however, shows that this measure of socialness is a proxy for at least *two different types* of social skills: Supervision, which is more closely associated with wage increases, and cooperation, which is more closely associated with employment increases.

In this section we juxtapose our results for supervision and cooperation with those of Deming (2017) for social skills. We replicated the Deming (2017) social

[15] Forty-four three digit industries are able to be matched exactly to the rest of the data. The remaining industries were matched many-to-one (the BEA data is coarser) with 2006 employment weighted averages.

[16] Appendix table 13 divides our regression on occupations by the repetitiveness of the occupation as measured by an O*NET question. For technology intensive occupations wage growth is stronger when the occupation is routine. For initiative and supervision intensive occupations this pattern is reversed, with stronger wage gains when the occupation is non-repetitive. Appendix table 14 divides occupations into high and low unstructuredness, and yields similar results. Structured occupations high in initiative and education saw relative wage declines, while those high in technology use saw larger increases.

	(1)	(2)	(3)	(4)
	ΔWage	ΔWage	Δln(Emp)	Δln(Emp)
Computer Use Split	Low	High	Low	High
Physical	0.322^{*}	-0.380^{*}	-0.027	-0.004
	(0.136)	(0.158)	(0.042)	(0.037)
Technology	0.116	-0.195	-0.055^{*}	0.029
	(0.068)	(0.149)	(0.026)	(0.025)
Perception	-0.044	-0.129	0.007	0.033
	(0.039)	(0.110)	(0.019)	(0.024)
Supervision	0.145	0.373^{**}	0.009	0.050^{*}
	(0.078)	(0.134)	(0.045)	(0.021)
Cooperation	0.140^{*}	-0.047	0.098^{***}	0.050
	(0.060)	(0.127)	(0.028)	(0.031)
Initiative	-0.252^{**}	0.310	-0.050	-0.012
	(0.082)	(0.185)	(0.033)	(0.039)
Math	-0.229^{***}	0.135	-0.029	0.010
	(0.043)	(0.120)	(0.020)	(0.034)
Educating	0.009	-0.151	-0.017	0.065
	(0.088)	(0.182)	(0.048)	(0.045)
Constant	-0.091	0.525^{***}	-0.036	-0.024
	(0.181)	(0.111)	(0.046)	(0.022)
Observations	2,877	6,062	2,877	6,062

Standard errors in parentheses

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

Table 6: Regression of change in median hourly wage and log employment by occupation-industry on 2006 occupational characteristics. Wage observations weighted by 2006 employment. Standard errors clustered by occupation. Sample split by occupation computer use, as defined by O*NET element, in 2006. Bottom and top 40 percentile occupation bins with equal total employment in 2006.

skill intensity measure by averaging of four O*NET level measures: social perception, coordination, persuasion and negotiation. This measure is then rescaled to an index lying between one and ten.[17]

[17]We use December 2006 O*NET scores and 2006 occupational employment rather than 1997 data as Deming(2017) does. We thank Deming for making his code available on his website.

	(1) ΔWage	(2) ΔWage	(3) Δln(Emp)	(4) Δln(Emp)
IT Intensity Split	Low	High	Low	High
Physical	−0.193	−0.241*	0.034	−0.013
	(0.106)	(0.121)	(0.029)	(0.022)
Technology	−0.064	−0.000	−0.018	−0.013
	(0.089)	(0.158)	(0.028)	(0.023)
Perception	−0.126*	−0.027	0.008	0.014
	(0.059)	(0.091)	(0.017)	(0.018)
Supervision	0.212*	0.220	0.048*	0.041*
	(0.098)	(0.126)	(0.021)	(0.018)
Cooperation	0.050	−0.017	0.082***	0.028
	(0.074)	(0.132)	(0.023)	(0.026)
Initiative	0.096	−0.030	−0.054*	0.005
	(0.094)	(0.121)	(0.023)	(0.021)
Math	0.035	−0.023	0.005	0.014
	(0.056)	(0.103)	(0.022)	(0.022)
Educating	0.145	−0.150	−0.009	0.030
	(0.124)	(0.176)	(0.034)	(0.033)
Constant	0.707***	0.413**	−0.052	−0.047*
	(0.136)	(0.137)	(0.028)	(0.021)
Observations	4,080	4,923	4,080	4,923

Standard errors in parentheses

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

Table 7: Regression of change in median hourly wage and log employment by occupation-industry on 2006 occupational characteristics. Wage observations weighted by 2006 employment. Standard errors clustered by occupation. Sample split by industry IT capital intensity in 2016 allowing one-to-many and many-to-one industry matches between BEA and BLS data. Bottom and top 40 percentile occupation bins with equal total employment in 2006.

Figure 4 contrasts our measures of social skill with those of Deming (2017). The first panel plots cooperation and supervision intensity scores by occupation. By construction these factors are orthogonal. The remaining panels compare our measures of social skills with Deming's social skill index (hereafter: socialness). It is significantly positively correlated with both. Notably, this is despite the fact that none of the O*NET elements used in the construction of socialness are used in the final calculation of our factors. The adjusted r^2 for regressing socialness on supervision and cooperation are .172 and .047 respectively. This indicates that, of our two orthogonal factors, socialness is more closely related to supervision. This is sensible given that three of the four items comprising Deming's socialness measure are more supervision-related (coordination, persuasion and negotiation) while one (social perception) is related to both.

It is clear that our measures of occupational social skill intensity are correlated with Deming's, but still differ significantly from his. To further provide a sense of how these three measures are distinguished, Table 8 reports relatively high and low supervision and cooperation occupations for occupations with a high socialness score and a low socialness score. Psychiatrists and clergy are both evaluated as high social skill occupations. However, our factors distinguish between the first, which is considered high cooperation and low supervision, while the latter is evaluated as high supervision but requiring only moderate cooperation.

High Deming Social Skill Index Occupations

	High Supervision	**Low Supervision**
High Cooperation	Education administrator	Psychiatrist
Low Cooperation	Clergy	Sales engineers

Low Deming Social Skill Index Occupations

	High Supervision	**Low Supervision**
High Cooperation	Hazmat worker	Gambling & sports book writer
Low Cooperation	Motorcycle mechanic	Sewing machine operator

Table 8: Occupations with relatively high and low cooperation and supervision scores for occupations with high and low Deming Social Skill index measures.

We next wish to juxtapose the importance of our social skill measures to SBTC with Deming's. Table 9 does so by re-running specification (1) with the inclusion of socialness as an additional regressor.

Comparing this table to our original estimates in Table 4, the effects remain broadly similar to our findings without socialness, with one understandable exception. In explaining the change in occupational wage, socialness enters negatively, though the point estimate is not significant. The point estimate of the effect of supervision remains positive and significant. Both these observations are true for the specification with and without industry fixed effects. The point estimates on the effect of supervision on wage growth are actually somewhat in-

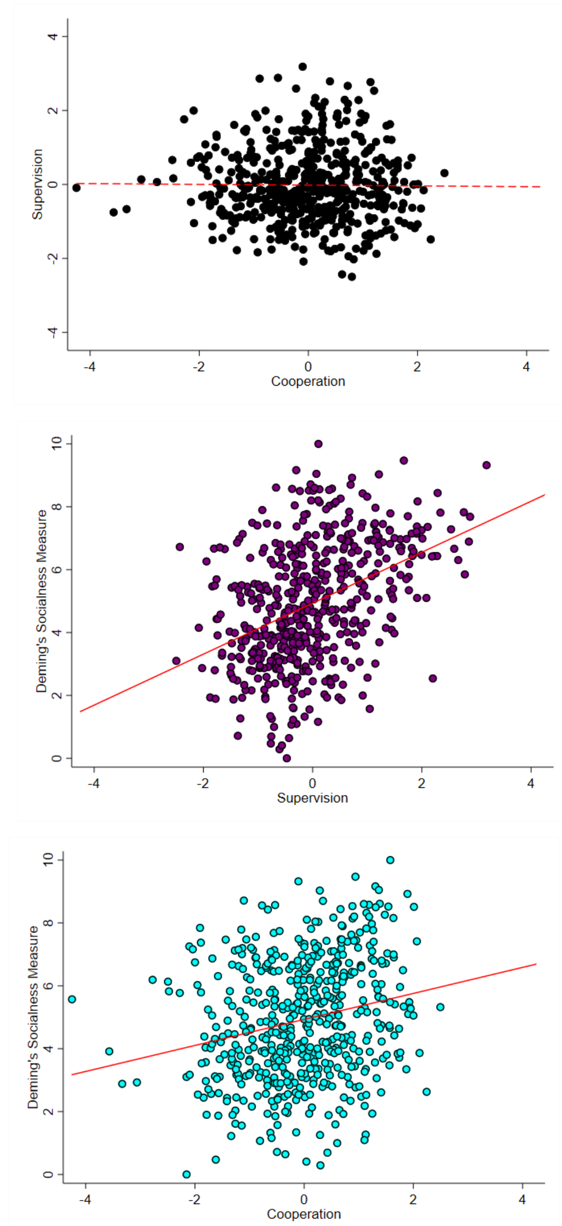

Figure 4: Scatterplots of occupational task intensity and lines of best fit relating Deming's social index and our two primary social skill measures. Socialness is positively associated with both Supervision and Cooperation skills, which are themselves orthogonal by construction. The adjusted r^2s in a regression of the socialness on Supervision and Cooperation are .172 and .047 respectively.

	(1)	(2)	(3)	(4)
	ΔWage	ΔWage	Δln(Emp)	Δln(Emp)
Physical	-0.286^{*}	-0.191	0.033	0.016
	(0.136)	(0.127)	(0.033)	(0.034)
Technology	-0.026	-0.029	-0.022	0.013
	(0.097)	(0.103)	(0.018)	(0.020)
Perception	-0.048	-0.079	-0.006	0.012
	(0.067)	(0.055)	(0.013)	(0.014)
Supervision	0.268^{*}	0.244^{*}	0.013	0.008
	(0.124)	(0.120)	(0.028)	(0.028)
Cooperation	0.044	-0.014	0.057^{**}	0.020
	(0.098)	(0.109)	(0.021)	(0.025)
Initiative	0.052	0.089	-0.057	-0.027
	(0.125)	(0.107)	(0.031)	(0.031)
Math	-0.003	0.075	-0.014	0.001
	(0.081)	(0.081)	(0.019)	(0.020)
Educating	0.038	0.014	0.001	-0.006
	(0.117)	(0.108)	(0.018)	(0.018)
Deming's Socialness	-0.066	-0.087	0.036	0.033
	(0.095)	(0.088)	(0.021)	(0.020)
Constant	0.840	0.890	-0.214	-0.165
	(0.504)	(0.458)	(0.115)	(0.114)
Industry FE		X		X
Observations	10,675	10,675	10,675	10,675

Standard errors in parentheses

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

Table 9: Regression of change in median hourly wage and log employment by occupation-industry on 2006 occupational characteristics and Deming Social Skill intensity. Wage observations weighted by 2006 employment. Standard errors clustered by occupation.

creased from the baseline specification. The significant negative point estimate on the physicality of an occupation, when industry FEs are not included, is also similar to the specification without socialness.

In the explanation of employment change, socialness enters positively but non-signficantly. Including socialness in the specification without industry FEs leaves the significance of cooperation uncahnged versus our earlier estimation without socialness. However, it reduces the significance of supervision, which is understndable given the higher correlation between socialness and supervision.

The results in Table 9 are consistent with cooperation-intensity being most useful for predicting occupational employment growth, supervision and physicality being most predictive for wage growth, and socialness not being particularly important for either, in the presence of our newly derived skill measures. However, one potential concern with this interpretation is over-fitting. Ten thousand observations is a healthy amount, but even with only nine regressors there is the potential that occupational characteristics with a noisier relationship to the data are obscuring robust relationships between some factors and employment and occupation growth.

To address possible concerns of overfitting, we conducted estimation using LASSO, a technique designed to avoid overfitting predictions with many regressors. Table 10 reports the results of a LASSO regression on our eight endogenously determined variables and three indexes constructed following Deming (2017). These are socialness, as well as measures of occupational routineness and non-routine-math intensity. λ, a parameter which governs how coefficient estimates are penalized, is selected using k-fold cross validation to minimize mean-squared error.[18]

This procedure, which maximizes fit in the kth of the data which is held out in each estimate, is a way of determining which how intensely estimates should be penalized to prevent over-fitting. Here λ takes on its minimum possible value. This means that overfitting is not a serious problem when estimating regressions using these eleven skill measures. All variables are useful information in predicting labor market outcomes. The most important regressors, physical, supervision, and cooperation, retain their previous signs and approximate magnitudes. However, the estimate of the effect of socialness on wage growth flips to positive.

IT FEELS LIKE WE NEED AT LEAST ONE MORE PARAGRAPH ON THESE LASSO RESULTS. ALSO NEED DATA ON HOW IT ALL COLLAPSES TO DEMING'S SOCIAL MEASURE IN THE LIMIT. THIS WAS IN

[18] This means that the exact λ, and therefore the variables which are retained, in LASSO regression is dependent on the seed used. However, the presented results are by far the most common. Tables with exogenous λ selections are deterministic.

	(1)	(2)
	ΔWage	Δln(Emp)
Physical	−0.129	−0.008
Technology	−0.044	−0.030
Perception	−0.107	0.014
Supervision	0.210	0.028
Cooperation	−0.035	0.054
Initiative	0.061	−0.006
Math	−0.105	−0.132
Educating	−0.069	0.049
Deming's Routineness	0.042	−0.026
Deming's Socialness	0.064	0.011
Deming's Math	0.194	0.094
Constant	−0.526	−0.308
Observations	10,675	10,675
λ	0.000	0.000
r^2	0.031	0.046

Table 10: LASSO regression of occupation skill intensity and three Deming measures on occupational wage and employment change. λ selected using k-fold cross-validation, $k = 10$

THE LASSO INTERACTIONS TABLE (NOW IN THE APPENDIX) , AND IS REFERRED TO IN THE INTRO. PERHAPS REPLICATE THE FORMAT OF THE INTERACTION TABLE, WITH MULTIPLE TIGHTER-LAMBDA COLUMNS, BUT DON'T INCLUDE THE INTERACTIONS OR THE DEMING ROUTINE / DEMING MATH MEASURES?

6 Conclusion

Previous research has established SBTC as an important force in the evolution of the labor market. A better understanding how individuals can re-skill themselves to deal with these challenges is therefore of utmost importance. Improved understanding can inform governments and educators in guiding individual re-skilling efforts. However, skill has many dimensions. In this paper we introduce a novel approach that uses unsupervised machine learning technique to derive occupational skill dimensions directly from well-established data on occupational requirements, rather than imposing them *ex ante*. We then analyze how occupations of various characteristics were impacted by SBTC.

Compared to earlier work, our approach imposes less prior structure on the data, instead seeking to endogenously identify the factors that jointly characterizing occupational skills.[19] The eight factors identified have intuitive relationships to the wage distribution and partially overlap with the skill categorizations emphasized in earlier research.

We find that occupational supervision, physicality, and cooperation intensity are the most significant predictors of occupational wage and employment growth. Supervision positively predicts employment and wage growth, suggesting an increase in demand. Cooperative jobs primarily experienced increases in employment, consistent with an increase in the supply of individuals seeking these jobs. Physically intense jobs saw decreases in wage, consistent with an decrease in demand for these types of jobs due to automation.

Splitting industries and occupations by IT capital use, we find further support for these hypotheses: the decrease in wages for physical jobs and increase in wages for supervision jobs are driven by high-tech occupations and industries, while the increase in cooperation intensive jobs is concentrated in low-tech occupations and industries.

We identify two measures of social skills that are important to SBTC; therefore, we juxtapose our results with Deming (2017). We show that his finding, an increased role for the importance of social skills, is the result of two distinct trends: (1) the increase in wages for supervision occupations, and (2) an increase in employment for cooperative occupations. This is further confirmed by looking into the underlying elements used to construct supervision, cooperation and socialness as management oriented elements are more closely associated with wage growth, and empathetic oriented elements are more closely associated with employment growth.

While the eight skills that emerged from our analysis provide insights into the nature of recent workforce changes, we make no claims that they will always and everywhere be the most important skills to consider. Instead, data-driven techniques should continue to be applied to reveal not only how any particular skill set can explain changes in employment and wage growth, but also how the skill set itself may change over time.

[19] Our technique does not eliminate human judgment in the construction of occupational measures because it relies on the questions that O*NET decides to ask and the occupations they decide to include.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge (2020) "Sampling-Based versus Design-Based Uncertainty in Regression Analysis," *Econometrica*, Vol. 88, No. 1, pp. 265–296.

Acemoglu, Daron (1999) "Changes in unemployment and wage inequality: An alternative theory and some evidence," *American Economic Review*, Vol. 89, No. 5, pp. 1259–1278, DOI: http://dx.doi.org/10.1257/aer.89.5.125910.1257/aer.89.5.1259.

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

Akst, Daniel (2013) "Automation Anxiety," *The Wilson Quarterly*, Vol. 37, No. 3, pp. 2–14.

Alabdulkareem, Ahmad, Morgan R Frank, Lijun Sun, Bedoor AlShebli, César Hidalgo, and Iyad Rahwan (2018) "Unpacking the polarization of workplace skills," *Science advances*, Vol. 4, No. 7, p. eaao6030.

Autor, David (2013) "The "task approach" to labor markets: an overview," *Journal for Labour Market Research*, Vol. 46, No. 3, pp. 185–199.

Autor, David H. and David Dorn (2013) "The growth of low-skill service job and the polarization of the US labor market," *The American Economic Review*, Vol. 103, No. 5, pp. 1553–1597, URL: http://economics.mit.edu/files/1474.

Autor, David H., Frank Levy, and Richard J. Murnane (2003) "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, Vol. 118, No. 4, pp. 1279–1333, DOI: http://dx.doi.org/10.1162/00335530332255280110.1162/003355303322552801.

Benzell, Seth G and Erik Brynjolfsson (2019) "Digital Abundance and Scarce Genius: Implications for Wages, Interest Rates, and Growth,"Technical report, National Bureau of Economic Research.

Berman, Eli, John Bound, and Stephen Machin (1998) "Implications of Skill-Biased Technological Change: International Evidence," *Source: The Quarterly Journal of Economics*, Vol. 113, No. 49, pp. 1245–127961, URL: http://www.jstor.org/stable/2586980 http://about.jstor.org/terms.

Brynjolfsson, Erik and Guillaume Saint-Jacques (2015) "Information Technology and the Rise of the Power Law Economy," *Working Paper*.

Cortes, Guido Matias, Nir Jaimovich, and Henry E Siu (2018) "The" end of men" and rise of women in the high-skilled labor market,"Technical report, National Bureau of Economic Research.

Deming, David J (2017) "The growing importance of social skills in the labor market," *The Quarterly Journal of Economics*, Vol. 132, No. 4, pp. 1593–1640.

Deming, David and Lisa B Kahn (2018) "Skill requirements across firms and labor markets: Evidence from job postings for professionals," *Journal of Labor Economics*, Vol. 36, No. S1, pp. S337–S369.

Edin, Per-Anders, Georg Graetz, Sofia Hernnas, and Guy Michaels (2018) "The Individual Consequences of Occupational Decline," *Working Paper*.

Goos, Maarten and Alan Manning (2007) "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain," *The Review of Economics and Statistics*, Vol. 89, pp. 118–133, URL: http://www.mitpressjournals.org/doi/pdf/10.1162/rest.89.1.118.

Goos, Maarten, Alan Manning, and Anna Salomons (2010) "Explaining Job Polarization in Europe: The Roles of Technology, Globalization and Institutions," *CEP Discussion Papers*, URL: http://ideas.repec.org/p/cep/cepdps/dp1026.html.

Goos, Maarten, Alan Manning, Anna Salomons, David Autor, Larry Katz, Thomas Lemieux, Stephen Machin, Guy Michaels, John Van Reenen, and Ulrich Zierahn (2014) "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring," *American Economic Review*, URL: http://personal.lse.ac.uk/manning/work/ExplainingJobPolarization.pdf.

Green, Francis (2012) "Employee involvement, technology and evolution in job skills: A task-based analysis," *ILR Review*, Vol. 65, No. 1, pp. 36–67.

Hair, Joseph F, William C Black, Barry J Babin, and Rolph E Anderson (2013) *Multivariate data analysis: Pearson new international edition*: Pearson Higher Ed.

Kaplan, Steven N. and Joshua Rauh (2013) "It's the Market: The Broad-Based Rise in the Return to Top Talent," *Journal of Economic Perspectives*, Vol. 27, No. 3, pp. 35–56, DOI: http://dx.doi.org/10.1257/jep.27.3.3510.1257/jep.27.3.35.

Krusell, Per, Lee E Ohanian, José-Víctor Ríos-Rull, and Giovanni L Violante (2000) "Capital-skill complementarity and inequality: A macroeconomic analysis," *Econometrica*, Vol. 68, No. 5, pp. 1029–1053.

Michaels, Guy, Ashwini Natraj, and John Van Reenen (2014) "Has ICT polarized skill demand? Evidence from eleven countries over twenty-five years," *Review of Economics and Statistics*, Vol. 96, No. 1, pp. 60–77.

Rosen, Sherwin (1981) "The Economics of Superstars," *The American Economic Review*, Vol. 71, No. 5, pp. 845–858.

Tuzel, Selale and Miao Ben Zhang (2019) "Economic stimulus at the expense of routine-task jobs," *Available at SSRN 3072733*.

A LASSO on interactions

Deming (2017) also identifies evidence of increasing complementarity between social and mathematical skills. He finds that wages in occupations requiring both strong mathematical and social skills grew by about 25 percentage points from 1980 to 2012. But are there other important complementarities between skill sets which have emerged due to technological change? In the last regressions of this paper, we use LASSO to examine the role of our 8 endogenously derived factors, Deming's 3 occupational characteristic indexes, and all of their interactions on occupational wage and employment change.

Table 11 reports the estimates of a LASSO regression of change in occupational wage on occupational characteristics. There are potentially 10 non-interacted and 45 (10 choose 2) interacted regressors in all, for a total of 55. In the table, only regressors with non-zero coefficient estimates with some λ are reported.

	(1) ΔWage	(2) ΔWage	(3) ΔWage	(4) ΔWage
Physical	-0.183	-0.118	-0.025	0
Supervision	0.169	0.100	0.044	0
Deming Socialness	0.045	0	0	0
Deming Math	0.055	0	0	0
Physical $\times$ Supervision	-0.008	0	0	0
Physical $\times$ Math	-0.102	-0.044	0	0
Perception $\times$ Supervision	-0.045	-0.016	0	0
Perception $\times$ Cooperation	-0.076	-0.031	0	0
Supervision $\times$ Education	-0.099	-0.081	0	0
Initiative $\times$ Deming Math	0.016	0	0	0
Deming Social $\times$ Deming Math	0	0.019	0.019	0.014
Constant	0.361	0.405	0.420	0.527
Observations	10675	10675	10675	10675
λ	0.075	0.100	0.200	0.300
r^2	0.034	0.031	0.023	0.016

Table 11: LASSO regression of occupation skill intensity and three Deming measures, as well as their forty-five interactions, on occupational wage change. Only regressors that take on non-zero values for some reported λ are presented. λ selected exogenously.

In a series of four regressions we increase λ exogenously and see which regressors are retained. The regressors which are the last to be discarded are the most important individually in explaining wage or employment growth. With

λ at a moderate level, .075, ten regressors are retained. Four of these are raw factors which retain the signs estimated before. For example, supervision and socialness enter positively, while physical enters negatively. The next six terms are interactions. Occupations which are supervision intensive saw wage increases in general, but jobs that are intensive in both this and physicality, perception, or education saw smaller increases or even decreases. Supervision positions that also require physical exertion or perception are likely to be lower level positions, so this result is consistent with that of the previous table. Also entering negatively are the interaction of perception and cooperation and the interaction of physical and math. The interaction of initiative and Deming's non-routine-analytical math measure enters positively.

Increasing λ to the high value of .2, the only regressors retained are physical, supervision and the interaction of Deming's socialness and non-routine-analytical math measures. The fact that this last measure only takes on a non-zero coefficient for high values of λ suggests that this interaction term is a good summary variable for the independent effects of Deming's socialness and math, which take on a coefficient of zero. Increasing the value of λ one more time, only this last interaction term remains. An implication from table 10 is that, including all 45 interactions, physical, supervision and the interaction of Deming's socialness and math are the most important predictors of occupational wage growth.

Table 12 repeats the same exercise with change in log employment as the outcome of interest. For a high value of λ, .08, the only retained term is socialness, marking it as the most important predictor of employment growth. For lower levels of λ, the nine regressors retained include supervision, which enters positively, and cooperation interacted with math, which enters negatively. This last interaction is notable, given that cooperation enters positively as a predictor of occupational employment growth in our OLS specifications. However, it is plausible that jobs requiring both cooperation and the basic numeracy that our math factor measures have decreased in employment due to technological shifts. Overall, these Lasso analyses confirm that Demings' indexes are powerful predictors of wage and employment changes. However, this interpretation should be tempered by the finding that Deming's 'socialness' aggregates two orthogonal concepts, supervision and cooperation, which exhibit distinct trends. Socialness is a good summary measure, but a more predictive model includes these measures.

	(1) Δln(Emp)	(2) Δln(Emp)	(3) Δln(Emp)	(4) Δln(Emp)
Supervision	0.008	0.004	0	0
Deming Routineness	-0.004	-0.002	0	0
Deming Socialness	0.042	0.038	0.035	0.020
Physical $\times$ Supervision	-0.033	-0.026	-0.018	0
Physical $\times$ Education	-0.010	0	0	0
Tech $\times$ Education	-0.004	0	0	0
Perception $\times$ Deming Social	0.005	0.003	0.001	0
Cooperation $\times$ Math	-0.005	0	0	0
Math $\times$ Deming Social	-0.001	0	0	0
Constant	-0.198	-0.197	-0.187	-0.115
Observations	10675	10675	10675	10675
λ	0.030	0.040	0.050	0.080
r^2	0.040	0.033	0.027	0.015

Table 12: LASSO regression of occupation skill intensity and three Deming measures, as well as their forty-five interactions, on occupational employment change. Only regressors that take on non-zero values for some reported λ are presented. LASSO implemented in STATA using the 'elasticregress' package. λ selected exogenously.

B Additional Specifications

	(1) ΔWage	(2) ΔWage	(3) Δln(Emp)	(4) Δln(Emp)
Physical	−0.008	−0.239	−0.056**	−0.009
	(0.144)	(0.140)	(0.019)	(0.029)
Technology	0.160	−0.259	0.002	−0.002
	(0.133)	(0.181)	(0.017)	(0.041)
Perception	−0.136	−0.162	0.052*	0.070
	(0.130)	(0.141)	(0.021)	(0.042)
Supervision	0.367*	0.219	−0.001	0.055*
	(0.154)	(0.151)	(0.032)	(0.025)
Cooperation	0.138	−0.103	0.121***	0.035
	(0.117)	(0.148)	(0.024)	(0.034)
Initiative	−0.027	0.022	−0.050*	0.009
	(0.120)	(0.117)	(0.025)	(0.026)
Math	−0.057	−0.034	−0.022	0.034
	(0.093)	(0.142)	(0.016)	(0.040)
Educating	0.185	−0.160	0.008	0.041
	(0.130)	(0.165)	(0.039)	(0.039)
Constant	0.411*	0.578***	0.051*	−0.074*
	(0.171)	(0.151)	(0.020)	(0.036)
Repetitiveness Split	Low	High	Low	High
Observations	4,057	4,561	4,057	4,561

Standard errors in parentheses

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

Table 13: Regression of change in median hourly wage and log employment by occupation-industry on 2006 occupational characteristics. Wage observations weighted by 2006 employment. Standard errors clustered by occupation. Sample split by occupation repetitiveness, as defined by O*NET element, in 2006. Bottom and top 40 percentile occupation bins with equal total employment in 2006.

	(1) ΔWage	(2) ΔWage	(3) Δln(Emp)	(4) Δln(Emp)
Physical	0.215	−0.234	−0.026	0.000
	(0.113)	(0.177)	(0.036)	(0.029)
Technology	0.221**	−0.059	−0.021	−0.014
	(0.077)	(0.176)	(0.026)	(0.033)
Perception	0.031	−0.210	0.010	0.027
	(0.058)	(0.188)	(0.015)	(0.031)
Supervision	0.121	0.207	0.070*	0.028
	(0.151)	(0.137)	(0.028)	(0.026)
Cooperation	0.163	0.086	0.052*	0.036
	(0.092)	(0.165)	(0.020)	(0.035)
Initiative	−0.327***	−0.045	0.006	−0.027
	(0.088)	(0.183)	(0.027)	(0.045)
Math	−0.137	0.075	0.006	−0.020
	(0.078)	(0.153)	(0.019)	(0.040)
Educating	−0.213*	−0.077	−0.031	0.089*
	(0.085)	(0.186)	(0.040)	(0.045)
Constant	−0.004	0.484***	−0.005	0.016
	(0.160)	(0.125)	(0.044)	(0.034)
Unstructuredness Split	Low	High	Low	High
Observations	2,872	5,843	2,872	5,843

Standard errors in parentheses

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

Table 14: Regression of change in median hourly wage and log employment by occupation-industry on 2006 occupational characteristics. Wage observations weighted by 2006 employment. Standard errors clustered by occupation. Sample split by occupation unstructuredness, as defined by O*NET element, in 2006. Bottom and top 40 percentile occupation bins with equal total employment in 2006.