B | Metropolitan Policy Program at BROOKINGS

DIGITALIZATION AND THE AMERICAN WORKFORCE

By Mark Muro, Sifan Liu,
Jacob Whiton, and Siddharth Kulkarni

November 2017

APPENDIX A. DIGITALIZATION METHODOLOGY DOCUMENTATION

This study examines digitalization of the workforce in the United States from 2002 to 2016. The primary data source of occupational digital scores is from the Occupation Information Network (O*NET) database[1], a project funded by Department of Labor's Employment and Training Administration to provide comprehensive information about every occupation in the U.S. economy. The study also uses Occupational Employment Estimates (OES) data for statistics of employment and wages.

DATA

The primary data source of occupational digital scores is the Occupation Information Network (O*NET) database, a project funded by Department of Labor's Employment and Training Administration (ETA) to provide comprehensive information about every occupation in the U.S. economy.[2]

O*NET is critical here because it surveys incumbent workers in every occupation to obtain fine-grained, job-specific information on thousands of workers' education, training, experience, and skill-related work requirements.

At the same time, the study also uses historical Occupational Employment Statistics (OES) data from the Bureau of Labor Statistics (BLS) for statistics on employment and wages, and Current Population Survey (CPS) for statistics on demographic variation.[3]

By collecting, linking, and analyzing these data the present inquiry was able to assess the digital content of hundreds of occupations, analyze their association with particular industries, track their pay and growth rates, map their locations, and consider their distribution across educational and demographic groups.

MEASURING DIGITALIZATION

To identify the digital content for each occupation, O*NET survey data were employed to construct occupation-specific digital scores. O*NET surveys incumbents in every occupation to collect information about the knowledge, skills, tools and technology, education and training, work context, and work activities required for their job and this inquiry drew on selected measures.

Specifically, we are interested in variables describing the digital content of occupations. Therefore, two of O*NET's three technology variables were identified as the most relevant measures of the overall digital tenor of occupations. One of these variables—**Knowledge - Computer and Electronics**—measures the overall knowledge of computers and electronics required by a job, while the other—**Work Activity - Interacting with Computers**—quantifies the centrality of computers to the overall work activity of the occupation. These measures seem to best capture the overall importance of digital knowledge and activity, job by job. As such, the two variables measure the level of digital skills required in each workplace.

In terms of converting these measures to mathematical data, O*NET reports a numerical score for each occupation on each variable. Requisite computer-electronics knowledge levels for each occupation are reported on a 0 to 7 scale and the importance of computers to each job is reported on 1 to 5 scale. In the survey questions, O*NET gives examples (or anchors) of the specific tasks at level 2, 3, and 6. (For detailed descriptions

of each level, consult the Appendix Table A.1).

Because the level and importance scales each have different ranges of possible scores, we have employed O*NET's recommended method to equally weight the two scores by converting the original ratings to a standardized combined score ranging from 1 to 100. Along these lines we use the following equation to construct occupational digital scores as the weighted sum of the two variables. That is:

$$\text{Digital score} = \frac{\sqrt{\text{Knowledge}_{level} \times \text{Knowledge}_{importance}} + \sqrt{\text{Work Activity}_{level} \times \text{Work Activity}_{importance}}}{2}$$

The maximum possible occupational digital score is thus: $\frac{\sqrt{100 \times 100} + \sqrt{100 \times 100}}{2} = 100.$

By dint of these procedures, the least digital occupations rate low scores and the most intensely digital ones obtain high scores.

CONNECTING O*NET OCCUPATIONS WITH EMPLOYMENT STATISTICS

To assess the job market impact of digitalization, we need to connect O*NET occupations with statistics of employment and wage. The Bureau of Labor Statistics' Occupational Employment Statistics Survey (OES) provides the most accurate source of occupational data.

We recognize the limitations of using OES data for time-series analysis, with challenges including changes in survey methodology and changes in occupational, industrial and geographical classification systems. Therefore, this research does not intend to report employment and wage changes in any individual occupations, but to present an overarching trend in the job market. O*NET also uses a different occupation structure

TABLE A1

Representative occupations and their digitalization levels

Digital level	Occupation	Digital score	Education requirements	Mean annual wage
High	Software Developers, Applications	94	Bachelor's degree	$104,300
High	Computer Systems Analysts	79	Bachelor's degree	$91,620
High	Financial Managers	61	Advanced degree	$139,720
Medium	Lawyers	58	Advanced degree	$139,880
Medium	Automotive Service Technicians and Mechanics	55	Some college	$41,400
Medium	Registered Nurses	55	Some college	$72,180
Medium	Office Clerks, General	55	Secondary or below	$33,010
Low	Security Guards	31	Secondary or below	$29,730
Low	Cooks, Restaurant	18	Secondary or below	$25,430
Low	Construction Laborers	17	Secondary or below	$37,890
Low	Personal Care Aides	14	Secondary or below	$22,710

Source: Brookings analysis of O*NET and OES data

with the Standard Occupational Classification (SOC) system, and therefore it's necessary to match O*NET occupation coding scheme with BLS SOC system in order to connect occupational digital scores with employment and wage statistics. We follow Johnathan Rothwell's procedures in linking O*NET to historic census data[4] and calculate average digital scores for SOC occupations with more detailed O*NET occupations for each available year. This procedure has collapsed 964 detailed O*NET occupations into 774 SOC detailed occupations.

Along these lines, the most highly digital band of occupations is exemplified by the high scores of software developers and financial analyst (digital scores = 94 and 73, respectively), while sales managers (60) and industrial machinery mechanics (45) are typical medium digital occupations, and security guards with a digital score of 31 exemplify the low band.

Scores were also developed for industries and states and metropolitan areas.

INDUSTRIAL ANALYSIS

Using occupational digitalization scores and OES industry-specific estimates, we were able to create mean digital scores for each industry. Industrial digital scores are weighted by the occupational employment distribution within the industry.

REGIONAL ANALYSIS

Similarly, we created mean digital scores for each state and metropolitan areas using OES state and metropolitan estimates. Regional digital scores are the mean occupational digital scores, weighted by the occupational employment within the region.

MEASURING CHANGE OVER TIME

Digitalization is not static, meanwhile. Change is proceeding rapidly, driven by the wide adoption of digital devices and processes, with significant implications for workers, firms, and the labor market. Which is why it has been important for this analysis to measure the change of occupations' digital content over time. Fortunately, O*NET permits such measurement.

Starting in 2002, O*NET has employed a database structure consistent with the Office of Management and Budget -approved Data Collection Program in 2002. Since then, O*NET has updated the database periodically, allowing for consistent tracking of change over time.

However, O*NET does not update every occupation on a yearly basis. Therefore, to analyze the change of occupational digital scores over a considerable time span, we have limited our time series to observations of occupations that were first surveyed before 2004 (2004 included), and had been updated at least once since 2009 (2009 included).

This reduces the scope of analysis somewhat—from 774 2010 Standard Occupational Classification (SOC) detailed occupations to 545. Nevertheless, study of those 545 occupations enables significant coverage of the labor market. Using estimates from the 2016 Occupational Employment Statistics (OES), these 545 occupations covered 90.8 percent of total U.S. employment in May 2016.[5]

For this analysis, we take the first score (surveyed between year 2002 and 2004) to represent occupational digitalization score of year 2002, and the latest score (surveyed between year 2009 and 2016) to represent occupational digitalization score of year 2016. Based on the assumption that occupational digital scores have been increasing over the years, using the outer boundary of this time span is a conservative measure of the score from 2002 to 2016.

EDUCATION-ADJUSTED WAGE PREMIUM

We are also interested in the wage premium of digital content among occupations with identical education requirements. O*NET reports the share of surveyed workers in an occupation that falls into 12 different education requirements categories. We use the mode (most frequent) education level to represent the level typically

TABLE A2

Education requirements categories

O*NET level	O*NET category	Analysis level
1	Less than a High School Diploma	
2	High School Diploma - or the equivalent (for example, GED)	
3	Post-Secondary Certificate - awarded for training completed after high school (for example, in agriculture or natural resources, computer services, personal or culinary services, engineering technologies, healthcare, construction trades, mechanic and repair technologies, or precision production)	**Secondary diploma or below**
4	Some College Courses	**Some college**
5	Associate's Degree (or other 2-year degree)	
6	Bachelor's Degree	**Bachelor's degree**
7	Post-Baccalaureate Certificate - awarded for completion of an organized program of study; designed for people who have completed a Baccalaureate degree but do not meet the requirements of academic degrees carrying the title of Master.	
8	Master's Degree	**Advanced degree**
9	Post-Master's Certificate - awarded for completion of an organized program of study; designed for people who have completed a Master's degree but do not meet the requirements of academic degrees at the doctoral level.	
10	First Professional Degree - awarded for completion of a program that: requires at least 2 years of college work before entrance into the program, includes a total of at least 6 academic years of work to complete, and provides all remaining academic requirements to begin practice in a profession.	
11	Doctoral Degree	
12	Post-Doctoral Training	

Source: Brookings analysis of O*NET and OES data

required for each occupation. We further group the 12 categories into four education levels for the purpose of this analysis: S**econdary Diploma or below**, **Some College**, **Bachelor's Degree**, and **Advanced Degree**.

To test if wage is correlated with digitalization score, we regress the average annual wage of 545 occupations on their digitalization scores, controlling for education level required by each occupation. The model is based on the following formula:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

Where y is the average annual wage, x_1 is the digitalization score (0 ~ 100), and x_2 is a categorical variable with four levels representing each education requirement. We used both 2002 data and 2016 data

The model shows that digitalization score has significant and positive effects on real annual wage even when controlling for education level in both years. In 2002, a one-point increase in digitalization score predicts a $166.2 (in 2016 dollar) increase in real annual average wage for occupations with the same education requirements. This wage premium of digitalization has almost doubled to $292.8 in 2016.

STRENGHTS AND LIMITATIONS

The present assessment, in the final analysis, has strengths and weaknesses. A strength is that the methodology exploits a rich source of direct survey data to provide specific, comparable task-level information for 545 occupations as they are changing over time. This allows for the production of a useful new data resource for policymakers and a novel analysis of the impacts of digitalization as they are making themselves felt in wages, across industries, and across metropolitan areas.

A key shortcoming here is that the O*NET occupational data are reported as aggregates and are not available as microdata that provides information at the level of individual respondents. Therefore, the digitalization scores we use for this analysis are not specific to particular industries or metropolitan area. Rather, our analyses—including local ones-- employ digital scores assigned nationally, without regard of the industry and location. That introduces potential inaccuracy into the digital scores of particular industries as well as those of occupations and industries at the local level.

Likewise, the current requirements for a job might be quite different from the O*NET-reported occupational requirements which reflect information from incumbent workers.

*

Overall, though, this analysis provides a fresh look at the magnitude and dynamics of digitalization in the U.S. workplace over the past 15 years, and begins a discussion of how that trend has affected the occupations, industries and metropolitan areas in the United States. For a more in-depth discussion of the data sources and method employed, please see the online appendix accompanying this report.

TABLE A3a

2002

Multiple R-squared: 0.4206, Adjusted R-squared: 0.4163
F-statistic: 97.64 on 4 and 538 DF, p-value: < 2.2e-16

Coefficients	2002 wage	CPI adjusted
(Intercept)	28278***	37609.74
Edu: Some College	6675.05*	8877.82
Edu: Bachelor Degree	20646.54***	27459.90
Edu: Advanced Degree	39190.7***	52123.63
2002 Digitalization Score	124.95**	166.18

TABLE A3b

2016

Multiple R-squared: 0.4406, Adjusted R-squared: 0.4365
F-statistic: 106.1 on 4 and 539 DF, p-value: < 2.2e-16

Coefficients	2016 wage
(Intercept)	29223.5***
Edu: Some College	8515.2*
Edu: Bachelor Degree	30369.5***
Edu: Advanced Degree	55147.3***
2002 Digitalization Score	292.8***

ENDNOTES

1 See https://www.onetcenter.org/overview.html

2 See www.onetcenter.org/overview.html

3 See www.bls.gov/oes/tables.htm and www.bls.gov/cps/tables.htm

4 Jonathan Rothwell, "The Hidden STEM Economy," (Washington: Brookings, 2013)

5 See https://www.bls.gov/soc/classification.htm . The modern Standard Occupational Classification (SOC) system was first developed in 2000 and revised in 2010. We use the official crosswalk to connect 2000 SOC occupations to 2010 SOC occupations. For detailed 2000 occupations that have been divided into multiple new 2010 occupations, we estimated the statistics using the 2010 employment shares as weights.

Learn more

To explore the full report, entitled "*Digitalization and the American workforce,*" visit www.brookings.edu/metro/research/digitalization.

For More Information

Mark Muro
Senior Fellow
Metropolitan Policy Program at Brookings
mmuro@brookings.edu