**O\*Net Time Series Dataset: Appendix**

**Summary:** This dataset lists detailed SOC 2010 occupation codes by (1) a novel technology and information service (TIS) intensity score and (2) five skill intensity measures based on the work of Acemoglu and Autor[[1]](#footnote-1). The data is available from 2002-2016[[2]](#footnote-2). These measures were complied using the O\*Net (Occupational Information Network) archive. Below are both the methodology and detailed dataset description.

**Sources:** The raw data was obtained from the Occupational Information Networks (O\*Net) archive, available here: <https://www.onetcenter.org/db_releases.html> and the occupation questionnaires were obtained form: <https://www.onetcenter.org/questionnaires.html>

**IT Intensity Sub-Component Selection**

**Description:** The TIS intensity consolidated score was constructed using a vector of 13 job characteristics per occupation. The O\*Net job characteristics, measuring the extent to which a particular skill, knowledge, or activity is applied to a job, are selected based on their ability to capture technology and information occupational components. To determine their relationship to TIS, we (1) read through the questionnaires provided by O\*Net and selected the components based on whether the component’s name, description or question prompt detailed a level of technological or informational expertise. After selecting the subcomponents, we (2) ran a correlation matrix in order to corroborate that the variables are co-moving.

**IT Intensity Score**

**Description**: The TIS score was constructed by adding the 13 sub-components per occupation (using 2010 standard occupation (SOC) codes); the resulting score was then standardized to have mean zero and standard deviation 1.

**Skill Intensity Scores**

**Description:** The skill intensity scores were calculated using the methodology described by Acemoglu and Autor (2011) and the description on the paper “From Occupations to Embedded Skills: A Cross-Country Comparison”. The five skill groups are: ‘non-routine cognitive analytical’, ‘non-routine cognitive interpretative’, ‘routine cognitive’, ‘routine manual’ and ‘non-routine manual physical’.

**Constructing the IT Intensity Time Series**

**Step 1 (Download Raw Data):** First, we download and clean the raw data files for the years 2002-2016 from the O\*Net database. Link: <https://www.onetcenter.org/db_releases.html>

**Step 2 (Extract Key Variables):** After downloading the corresponding raw files for each year- where each year has it’s own folder with text files, we loop through each year to select our relevant variables. We do this by first selecting the text files “Knowledge.txt”, "Skills.txt", "Work\_Activities.txt", and "Work\_Context.txt". Then we extract the variables that belong to each work component (for example: "Computers and Electronics" and Engineering and Technology" are found within the “Knowledge.txt” file). Thus, we end up with four files per year, namely “Knowledge.dta”, “Skills.dta”, "Work\_Activities.dta” and "Work\_Context.dta”, each containing their respective subsets of the 13 subcomponents.

\*Note: the year 2002 doesn’t include the ‘Electronic Mail’ variable.

**Step 3 (Append):** Our next step is appending the four files corresponding to each year, resulting in one file per year. We also add a year variable to each file. Then, we append again, but this time in order to put each yearly file together. From this step, we obtain the file “ONet\_2002\_2016.dta"

**Step 4 (Crosswalk):** Now, we crosswalk our ONet-SOC occupation codes with SOC 2010 occupations in order to be able to link our TIS scores to BLS datasets. We use the crosswalk provided by O\*Net: <https://www.onetcenter.org/crosswalks.html> .After crosswalking we collapse our valuesusing collapse datavalue, by (soc2010 elementname scaleid year). We do this since our crosswalk from ONet-SOC to SOC 2010 is one-to-many, which means that after crosswalking, we get repeated values.

**Step 5 (Clean Dataset):** Once we have our variables uniquely defined by soc2010 codes and year,we reshape our dataset using the following command: reshape wide datavalue, i(soc2010 year) j(unique\_id) string. After reshaping, we generate our final 13 sub-component measures by multiplying their corresponding ‘level’ with their ‘importance’ (e.g. generate COMP\_AND\_ELEC = datavalueComp\_and\_ElectrIM\*datavalueComp\_and\_ElectrLV**).** Our resulting dataset is “ONet\_2002\_2016\_reshaped1.dta"

**Step 6 (Generate Consolidated TIS measure):** Our last step is to first standardize each sub-component (in order to have mean zero and variance 1 (we use the ‘center’ command). Secondly, we add the 13 standardized sub-component scores per occupation and year, and standardized our consolidated score. Our end product is a TIS score per soc2010 occupation per year. We save our results in “ONet\_Time\_Series\_2002\_2016.dta"

**Skill Intensity Variables:** We follow the same process for the skill intensity variables using their respective sub-components. This results in the file “Acemog\_Time\_Series\_2002\_2016\_1.dta". Our final step is to merge these datasets together, producing our final O\*Net TS file: “ONet\_TS\_2002\_2016\_v3.dta"

**O\*NET Variable Selection: IT Intensity Sub Components**

1. **Knowledge Computers and Electronics:** Knowledge of circuit boards, processors, chips, electronic equipment, and computer hardware and software, including applications and programming

**Question:** What level of COMPUTERS AND ELECTRONICS is needed to perform your current job?

**1 (lowest level):** Operate a VCR to watch a pre -recorded training tape

**4 (intermediate):** Use a word processor

**6 (high):** Create a program to scan computer disks for viruses

1. **Work Activity Interacting with computers:** Using computers and computer systems (including hardware and software) to program, write software, set up functions, enter data, or process information.

**Question:** What level of WORKING WITH COMPUTERS is needed to perform your current job?

**1 (lowest level):** Enter employee information into a computer database

**4 (intermediate):** Write software for keeping track of parts in inventory

**6 (high):** Set up a new computer system for a large multinational company

1. **Skills Programming:** Writing computer programs for various purposes

**Question:** What level of PROGRAMMING is needed to perform your current job?

**1 (lowest level):** Write a program in BASIC to sort objects in a database

**4 (intermediate):** Write a statistical analysis program to analyze demographic data

**6 (high):** Write expert system programs to analyze ground radar geological data for probable existence of mineral deposits

1. **Skills System** Evaluation Identifying measures or indicators of system performance and the actions needed to improve or correct performance, relative to the goals of the system

**Question:** What level of SYSTEMS EVALUATION is needed to perform your current job?

**1 (lowest level):** Determine why a coworker has been overly optimistic about how long it would take to complete a task

**4 (intermediate):** Identify the major reasons why a client might be unhappy with a product

**6 (high):** Evaluate the long-term performance problem of a new computer system

1. **Work Context: Email**

**Question:** How frequently does your current job require electronic mail?  (from 1-5)

1. **Skills Quality control analysis** Conducting tests and inspections of products, services, or processes to evaluate quality or performance

**Question:** What level of QUALITY CONTROL ANALYSIS is needed to perform your current job?

**1 (lowest level):** Inspect a draft memorandum for clerical errors

**4 (intermediate):** Measure new part requisitions for tolerance to specifications

**6 (high):** Develop procedures to test a prototype of a new computer system

1. **Skills Operations analysis:** Analyzing needs and product requirements to create a design

**Question:** What level of OPERATIONS ANALYSIS is needed to perform your current job?

**1 (lowest level):** Select a photocopy machine for an office

**4 (intermediate):** Suggest changes in software to make a system more user friendly

**6 (high):** Identify the control system needed for a new process production plant

1. **Work Activities: Updating and using Relevant knowledge**: Keeping up-to-date technically and applying new knowledge to your job

**Question:** What level of UPDATING AND USING RELEVANT KNOWLEDGE is needed to perform your current job?

**1 (lowest level):** Keep up with price changes in a small retail store

**4 (intermediate):** Keep current on changes in maintenance procedures for repairing sports cars

**6 (high):** Learn information related to a complex and rapidly changing technology

1. **Skills Technology design:** Generating or adapting equipment and technology to serve user needs.

**Question:** What level of TECHNOLOGY DESIGN is needed to perform your current job?

**1 (lowest level):** Adjust exercise equipment for use by a customer

**4 (intermediate):** Redesign the handle on a hand tool for easier gripping

**6 (high):** Create new technology for producing industrial diamonds

**Prospective Variables and their respective descriptions/questions:**

1. **Work Activity Analyzing Data or Information:** Identifying the underlying principles, reasons, or facts of information by breaking down information or data into separate parts.

**Question: What level of ANALYZING DATA OR INFORMATION is needed to perform your current job?**

1 (lowest level): Determine the location of a lost order

4: (intermediate): Determine the interest cost to finance a new building

6: (high): Analyze the cost of medical care services for all hospitals in the country

1. **Work Activity: Processing Information:** Compiling, coding, categorizing, calculating, tabulating, auditing, or verifying information or data.

**Question:** What level of PROCESSING INFORMATION is needed to perform your current job?

**1 (lowest level):** Tabulate the costs of parcel deliveries

**4 (intermediate):** Calculate the adjustments for insurance claims

**6 (high):** Compile data for a complex scientific report

1. **Knowledge Engineering and technology:** Knowledge of the practical application of engineering science and technology. This includes applying principles, techniques, procedures, and equipment to the design and production of various goods and services.

**Question:** What level of knowledge of ENGINEERING AND TECHNOLOGY is needed to perform your current job?

**1 (lowest level):** Install a door lock

**4 (intermediate):** Design a more stable grocery cart

**6 (high):** Plan for the impact of weather in designing a bridge

1. **Skills Management of Material Resources:** Obtaining and seeing to the appropriate use of equipment, facilities, and materials needed to do certain work.

**Question:** What level of MANAGEMENT OF MATERIAL RESOURCES is needed to perform your current job?

**1 (lowest level):** Rent a meeting room for a management meeting

**4 (intermediate):** Evaluate an annual uniform service contract for delivery drivers

**6 (high):** Determine the computer system needs of a large corporation and monitor use of the equipment

1. The five skill measures are: Routine Manual (RM), Routine Cognitive (RC), Non- Routine Cognitive Analytical (NRCA), Non-Routine Cognitive Interpretative (NRCI), Non-Routine Manual Physical (NRMP). For more information see Acemoglu, Daron, and David Autor. ”Chapter 12 Skills, Tasks and Technologies: Implications for Employment and Earnings\*.” Chapter 12 - Skills, Tasks and Technologies: Implications for Employment and Earnings. Handbook of Labor Economics, 2011 [↑](#footnote-ref-1)
2. The years 2002-2004 are grouped together in order to account for missing values and measures in those years, especially 2002 ; Is this still true? [↑](#footnote-ref-2)