README for Task-Based Discrimination

Introduction

This document describes the contents of the replication package for "Task Based Discrimination" (Hurst et al. (Forthcoming)), forthcoming in the *American Economic Review*. We discuss the data sources, the associated data sets included in the replication package, and the files included that reproduce all tables and figures in the paper.

Statement about Rights

We certify that the authors of the manuscript have legitimate access and permission to use the data used in this manuscript. We also certify that the authors of the manuscript have documented permission to redistribute/publish the data contained within this replication package.

Overview of File Structure

The associated replication file package has a root directory that includes this "README.PDF" document as well as four subdirectories. The first subdirectory – entitled "data" – includes all of the primitive data sources used for our analysis. The second subdirectory – entitled "code" – includes all of the Stata and MATLAB code used to make all figures and tables in the paper. The third subdirectory – entitled "figures" – includes all the figures and tables embedded in the main paper and the online appendix, which are created running the files in the "code" subdirectory. The final subdirectory – entitled "output" – stores all the output generated by running the files in the "code" subdirectory. These outputs are either datasets that are used as intermediate inputs for other analyses or are the final outputs from which tables and figures are made. The following sections expand on some of these subdirectories in greater detail. Lastly, the root directory also includes an Excel file entitled "*README_supplement.xlsx*" that summarizes (i) the location and name of each of the programs in the replication kit, (ii) a list of which programs call other programs when they are running, (iii) a list of figures and tables produced by each program, (iv) the data files that are inputs into each program, and (v) the intermediate outputs created by each program.

Computational Requirements

Below are the software and memory/runtime requirements needed to replicate all of our analysis.

Software Requirements

- Stata (code was last run with version 18)[1]
  - estout
  - reghdfe
  - ftools
- MATLAB (code was last run with version R2023a)[2]
  - Optimization Toolbox
  - Global Optimization Toolbox
  - Parallel Computing Toolbox
  - Statistics and Machine Learning Toolbox

Memory and Runtime Requirements

The code was last run on a 6-core Intel-based laptop (16GB RAM) with Windows 11 Version 22H2. The approximate run-time on a desktop machine is 2 hours.

Details on each Data Source

The section reviews all primitive data sources used in the paper. Each one of the primitive data sources are saved in the subdirectory of the replication kit entitled "**data/**". We use five primitive data sources for the main analysis in the paper. First, we use data from IPUMS USA for historical Censuses and the American Community Surveys. Second, we use files created by David Deming (in his replication kit) to map tasks to occupations; we also take the skill measures Deming constructed from his files. Third, we create our own crosswalk that maps the occupational task measures in David Deming's work to the Census harmonized occupation codes. Fourth, we use data from the Bureau of Labor Statistics's National Longitudinal Survey of Youths. Finally, we use the measures of state-level discrimination measures created by Kerwin Charles and Jon Guryan. For the analysis in the Online Appendix, we use two additional data sources. First, we use files created by David Autor (available on his website) to compare task measures in the Dictionary of Occupational Titles (DOT) across different years. Second, we download data from the O*NET Database to compare the O*NET task measures in different years. The data files in "**data/**" are organized into subdirectories by the data source, as described below. We also provide additional details on these data sources, including how they were accessed. A detailed discussion of the variable creation within these data sources can be found in the Online Data Appendix accompanying the main paper.

[1]The program automatically downloads the required Stata packages.

[2]Download these toolboxes by clicking Home >Add-Ons >Get Add-Ons on the top bar and then searching for the required toolboxes.

- **Census/ACS Data**
  - Location: **data/acs/**
  - Data file name: "census_acs_extract.dat"
  - Data description: This file contains the raw data directly extracted from the IPUMS website. This file includes data from the 1960, 1970, 1980, 1990, and 2010 U.S. *Census* long forms. This file also includes data from the 2010, 2011, 2012, 2016, 2017, and 2018 *American Community Surveys (ACS)*. The file is quite large given that it is pooling individual observations, with many variables per observation, for each of the above survey years. The data serves as the backbone from almost all of the empirical analysis in the main paper.
  - Data source: This data was downloaded directly from https://usa.ipums.org/usa/ (Ruggles et al., 2021).
  - Associated Stata "do file": This data is converted into a Stata data set using *1_import_datafile.do* found in the "acs" folder of the "code" subdirectory in the replication kit (*/code/acs/1_import_datafile.do*).

- **Task Measures**
  - Location: **data/deming/**
  - Data file names: "dot77-occ1990dd-Tr.dta" and "onet98_occ1990dd.dta"
  - Data description: These two files were downloaded directly from the replication kit for Deming (2017b). These files map a variety of occupational task measures to various occupation codes. The former file includes Deming's mapping of task measures in the 1977 Dictionary of Occupational Titles (DOT) to various occupations. These data are used to create our measures of *Abstract*, *Manual*, and *Routine* tasks. The latter file includes Deming's mapping of task measures in the 1998 O*NET data to various occupations. These data are used to create our measure of *Contact* tasks. Both data sets also contain alternate task measures that we use in various robustness exercises in the Appendix.
  - Data source: This data was downloaded directly from David Deming's replication kit found at https://doi.org/10.7910/DVN/CYPKZH (Deming, 2017a).
  - Associated Stata "do file": This data is merged into our primary Census/ACS Stata data set using 2*_create_datasets.do* found in the "acs" folder of the "code" subdirectory in the replication kit (*/code/acs/*2*_create_datasets.do*). This data is merged into our primary NLSY data set using 5*_merge_datasets.do* found in the

"nlsy" folder of the "code" subdirectory (*/code/nlsy/5_merge_datasets.do*) in the replication kit.

- **Crosswalk to Map Task Measures Into Census/ACS Data**
  - Location: **data/crosswalk/**
  - Data file name: "cross_walk_census_dd_occs_final.txt"
  - Data description: This crosswalk was created by authors to map the occupation code in the Deming task files (occ1990dd) to the harmonized occupation code in the Census/ACS data (occ1990). We need to have a modified crosswalk given that Deming's analysis starts in 1970 and our analysis starts in 1960. We use the harmonized occupation codes (occ1990) to link occupations consistently from 1960 through 2018. As seen from our crosswalk, most of the Deming occupation codes in occ1990dd are exactly the same as the Census/ACS codes in occ1990. In particular, only about 65 of the 384 harmonized Census/ACS occupation codes (occ1990) differ from the occupation codes in Deming (2017b) (occ1990dd). Most of this is due to the fact that the task codes in the Deming file are slightly more aggregated than the Census harmonized occupation codes (occ1990). For example, there are 15 distinct occupations identifying teachers of different subject areas in the Census harmonized occupation codes (occ1990 with codes between 113 and 154). These all get mapped to code 154 in Deming's occupation code (occ1990dd) which is for a broader "Subject instructors (HS/College)" occupation.
  - Data source: Created by authors. See discussion above.
  - Associated Stata "do file": This data is used to facilitate the merge between the Deming task measures and the Census/ACS measures in *2_create_datasets.do* found in the "acs" folder of the "code" subdirectory in the replication kit. This data is used again in *7_skills_by_occ.do* found in the "nlsy" folder of the "code" subdirectory in the replication kit to aggregate the NLSY skill measures by broad occupation categories based on the Census harmonized occupation codes.

- **NLSY Data**
  - Location: **data/nlsy/**
  - Data file name: "AER_NLSY1979.dct" and "AER_NLSY1997.dct"
  - Data description: These files contain the raw data directly extracted from the NLSY websites. The data in the first file comes from the National Longitudinal Survey of Youth, 1979 Cohort (NLSY-79). The data in the second file comes from the National Longitudinal Survey of Youth, 1997 Cohort (NLSY-97). These

panel data allow us to measure how pre-labor market factors are correlated with subsequent occupational sorting along task dimensions. In particular, the results are used for Table 3 in the main text and Appendix E in the online appendix.

  - Data source: The data was downloaded directly from the BLS website hosting the NLSY data: https://www.nlsinfo.org/content/getting-started/accessing-data. See, U.S. Department of Labor, Bureau of Labor Statistics (2019a) and U.S. Department of Labor, Bureau of Labor Statistics (2019b) for data citation.
  - Associated Stata "do file": The 1979 and 1997 datafiles are converted into Stata data sets respectively by *1_import_datafile_1979.do* and *2_import_datafile_1997.do* found in the "nlsy" folder of the "code" subdirectory in the replication kit (*/code/nlsy/1_import_datafile_1979.do, /code/nlsy/2_import_datafile_1997.do*).

- **Crosswalk between NLSY and Deming (2017b)'s occupation codes**
  - Location: **data/deming/**
  - Data file names: "occ1970_occ1990dd.dta", "occ1980_occ1990dd_update.dta", and "occ2000_occ1990dd_update.dta"
  - Data description: These three files were downloaded directly from the replication kit for Deming (2017b). These files map the Census Occupational codes used in the NLSY data to the occupation codes used by Deming (2017b) (occ1990dd) in his task files, thus allowing us to merge his task measures into the NLSY data. The three files contain the crosswalk with the 1970, 1980, and 2000 Census Occupational codes, respectively.
  - Data source: These files were downloaded directly from David Deming's replication kit found at https://doi.org/10.7910/DVN/CYPKZH (Deming, 2017a).
  - Associated Stata "do file": The crosswalks are used in *3_create_dataset_1979.do* and *4_create_dataset_1997.do* found in the "nlsy" folder of the "code" subdirectory in the replication kit.
- **Skill Measures Constructed by Deming (2017b) based on NLSY**
  - Location: **data/deming/**
  - Data file name: "nlsy_merged.dta"
  - Data description: This file was downloaded directly from the replication kit for Deming (2017b). This is the primary data NLSY set used for analysis in Deming's paper, and it contains various pre-labor market skill measures he constructed. We merge Deming's skill measures into our NLSY data set above. The results are used for Table 3 in the main text and Appendix E in the online appendix.

- Data source: This data was downloaded directly from David Deming's replication kit found at https://doi.org/10.7910/DVN/CYPKZH (Deming, 2017a).
- Associated Stata "do file": This data is merged into our primary NLSY data set using 5_*merge_datasets.do* found in the "nlsy" folder of the "code" subdirectory in the replication kit.

- **Charles-Guryan Measures of State-Level Discrimination**
  - Location: **data/charles_guryan/**
  - Data file name: "charles_guryan_data_raw.txt"
  - Data description: This data was created in Charles and Guryan (2008). The text file has four columns. The first column has state two-digit postal code identifiers, while the second column contains the state's fips number. This column allows merging with the Census/ACS data. The third and fourth columns include the average and marginal discrimination measures made by Charles and Guryan for 44 U.S. states. These measures were created based on survey answers provided by respondents in the General Social Survey (GSS).
  - Data source: This data file was sent to us directly by Kerwin Charles. Charles and Guryan do not have an online replication kit for their paper. However, the authors shared this data with us and noted we can use it our paper and include it as part of our replication file. A description of how the data was created can be found in Charles and Guryan (2008).
  - Associated Stata "do file": This data was used to create Figure 11 in the main text. In particular, Figure 11 compares our estimates of state-level discrimination (as measured by the racial task gaps in Contact and Abstract tasks) to the Charles-Guryan survey-based measures of discrimination. The creation of Figure 11 can be found in *3_main_analysis.do* found in the "acs" folder of the "code" subdirectory in the replication kit.

- **Task Measures based on DOT constructed by Autor et al. (2003)**
  - Location: **data/autor/**
  - Data file names: "dot77-8090.dta", "dot91-8090.dta"
  - Data description: These files are available on David Autor's website as supplemental materials to Autor et al. (2003). They contain task measures from the Dictionary of Occupational Titles (DOT) released in 1977 and 1991, respectively. We use the data to examine the persistence of DOT task measures over time, whose results are presented in Appendix Table R2 Panel A.

- Data source: These files were downloaded directly from Data Archive on David Autor's website at https://economics.mit.edu/people/faculty/david-h-autor/data-archive (Autor, 2024). They are part of the supplemental materials to Autor et al. (2003) and are titled "DOT means by occupation".
- Associated Stata "do file": This data is used in *6_dotonet.do* found in the "acs" folder of the "code" subdirectory in the replication kit.

- **O*NET 98 Data**
  - Location: **data/onet/onet1998/**
  - Data file names: "means_ab.sas7bdat", "means_kn.sas7bdat", "means_sk.sas7bdat", and "means_wc.sas7bdat".
  - Data description: These files contain the raw O*NET 98 data directly downloaded from the O*NET website. The Occupational Information Network (O*NET) data contains measures of various skills demanded in each occupation; O*NET 98 is the version released in December 1998. We use this data (together with O*NET 26.1 released in 2021, see below) to examine the persistence of task measures over time; the results are shown in Panel B of Appendix Table R2.[3]
  - Data source: These SAS Database files were downloaded directly from O*NET Resource Center at www.onetcenter.org/db_transitional.html (U.S. Department of Labor, Employment and Training Administration, 1998).
  - Associated Stata "do file": This data is used in *6_dotonet.do* found in the "acs" folder of the "code" subdirectory in the replication kit.

- **O*NET 26.1 Data**
  - Location: **data/onet/onet2021/**
  - Data file names: "Abilities.xlsx", "Knowledge.xlsx", "Skills.xlsx", and "Work Content.xlsx".
  - Data description: These files contain the raw O*NET 26.1 data directly downloaded from the O*NET website. The Occupational Information Network (O*NET) data contains measures of various skills demanded in each occupation; O*NET 26.1 is the version released in November 2021. We use this data (together with O*NET 98 released in 1998, see above) to examine the persistence of task measures over time; the results are shown in Panel B of Appendix Table R2.

[3]For the analysis in the main text, we use task measures taken from the replication kit in Deming (2017b) as discussed above.

- Data source: These SAS Database files are taken from the O*NET 98 database downloaded directly from O*NET Resource Center at www.onetcenter.org/db_releases.html (U.S. Department of Labor, Employment and Training Administration, 2021).
- Associated Stata "do file": This data is used in *6_dotonet.do* found in the "acs" folder of the "code" subdirectory in the replication kit.

- **Crosswalk between O*NET 98 Occupation Units and O*NET-SOC codes**
  - Location: **data/onet/crosswalk/**
  - Data file names: "socxonet.sas7bdat".
  - Data description: This file is downloaded from the O*NET website. It provides a crosswalk between the occupational classification used in O*NET 98 and the O*NET-SOC occupation codes used in O*NET releases from 2000 onwards. The crosswalk is used together with the O*NET-SOC crosswalks below.
  - Data source: The SAS Database file is taken from the O*NET 3.1 database downloaded directly from O*NET Resource Center at www.onetcenter.org/db_transitional.html (U.S. Department of Labor, Employment and Training Administration, 2001).
  - Associated Stata "do file": This data is used in *6_dotonet.do* found in the "acs" folder of the "code" subdirectory in the replication kit.

- **O*NET-SOC Occupation Classification Crosswalks**
  - Location: **data/onet/crosswalk/**
  - Data file names: "2000_to_2006_Crosswalk.xlsx", "2006_to_2009_Crosswalk.xlsx", "2009_to_2010_Crosswalk.xlsx", and "2010_to_2019_Crosswalk.xlsx".
  - Data description: The files are downloaded from the O*NET website. They contain crosswalks between different revisions of the O*NET-SOC occupation classifications, used in O*NET releases from 2000 onwards. We use these crosswalks, together with the crosswalk between O*NET 98 occupational units and the 2000 O*NET-SOC codes above, to harmonize the occupation classifications in the O*NET 98 and O*NET 26.1 databases.
  - Data source: The files are downloaded directly from the O*NET-SOC Taxonomy page in the O*NET Resource Center website at www.onetcenter.org/taxonomy.html (U.S. Department of Labor, Employment and Training Administration, 2023).

– Associated Stata "do file": This data is used in *6_dotonet.do* found in the "acs" folder of the "code" subdirectory in the replication kit.

Instruction to Replicators

To replicate the analyses in the paper, please run the following four programs (in the order listed). Note: You will have to adjust the default directory path at the beginning of each file.

- code/acs/0_acs_run_all.do (Analysis with Census/ACS)
- code/model/m1_model_run_all.m (Model Estimation and Analysis, MATLAB)
- code/model/m2_figures.do (Model Estimation and Analysis, Stata)
- code/nlsy/0_nlsy_run_all.do (Analysis with NLSY)

The replicated figures and tables are saved under the directory **figures/**. For information on which program produces which figures and tables, refer to the file "README_supplement.xlsx". Descriptions of what each program does are given below.

Naming Conventions for Figures and Tables

As noted above, the replicated figures and tables are saved under the directory **figures/**. The conventions for filenames are as follows. All figures are saved with filename "fig_X.eps" or "fig_X_panel_Y.eps" where X is the figure number and Y is the panel letter. The data underlying the plots is saved as a CSV file with the filename "fig_X_data.csv" or "fig_X_panel_Y_data.csv". If the text discusses any other information about the figures, it is saved as a text file with the filename "fig_X_notes.txt" or "fig_X_panel_Y_notes.txt". Similarly, tables are saved with the filename "table_X.tex" where X indicates the table number, with any supplemental information saved in "table_X_notes.txt" (if any). Finally, the motivational occupational statistics mentioned in Section 1 are saved in "intro_occ_shares.csv".

Description of programs

Analysis with Census/ACS Programs under **code/acs/** perform the analyses with Census/ACS.

- code/acs/0_acs_run_all.do: will configure the default path and other settings, and run all the programs under this directory.
- code/acs/1_import_datafile.do: will import the Census/ACS data extract.
- code/acs/2_create_datasets.do: will create the data set for analysis. In particular, it will merge the task measures from Deming (2017b) with the Census/ACS dataset and also define variables needed for the analysis.

- code/acs/3_main_analysis.do: will produce figures and other numbers (that are based on Census/ACS) in the main text of the paper, namely Figures 1-4, 9, and 11.
- code/acs/4_appendix.do: will produce figures and tables (that are based on Census/ACS) in the Online Appendix, namely Appendix Figures R1-R11 and Appendix Table R1.
- code/acs/5_create_moments.do: will calculate moments needed for the model estimation.
- code/acs/6_dotonet.do: will import the DOT and O*NET task measures in different years and produce Appendix Table R2.

Log files are saved under **output/acs/logs/**.

Model Estimation and Analysis Programs under **code/model/** perform the model estimation and other analyses based on the estimated model. There are two sets of programs under this directory: programs in MATLAB and a program in Stata.

MATLAB programs: [4]

- code/model/m0_model_run_all.m: will configure the default path and other settings, and run all the MATLAB programs under this directory.
- code/model/depends/estimate_model.m: will estimate the model using the data moments created by "code/acs/4_appendix.do" above.
- code/model/depends/create_tables.m: will prepare CSV tables summarizing estimation results, calculate additional moments, and perform the sensitivity analysis described in the Online Appendix.
- code/model/depends/decompose_trends.m: will perform the decomposition of trends in moments over time as described in Section 5.2 of the paper.
- code/model/depends/save_tables.m: will prepare the formatted TEX tables in the paper, namely Table 2 and Appendix Tables R7-R10.

STATA program:

- code/model/m1_figures.do: will retrieve model results produced by MATLAB programs above and create figures in the paper, namely Figures 5-8, 10, and R12-R18.

Analysis with NLSY Programs under **code/nlsy/** perform analysis using NLSY in the Online Appendix.

[4]The last four MATLAB programs are *functions* (not scripts). These functions are called with the path of the output directory ('savedir_root') as well as values of ψ and (optionally) θ. This allows us to run the same estimation code with alternate ψ and θ values for the robustness exercise in the appendix. The main program ("code/model/m0_model_run_all.m") will handle all this and run the functions with appropriate arguments.

- code/nlsy/0_nlsy_run_all.do: will configure the default path and other settings, and run all the programs under this directory.
- code/nlsy/1_import_datafile_1979.do: will import the NLSY1979 data extract.
- code/nlsy/2_import_datafile_1997.do: will import the NLSY1997 data extract.
- code/nlsy/3_create_dataset_1979.do: will prepare the NLSY1979 data set.
- code/nlsy/4_create_dataset_1997.do: will prepare the NLSY1997 data set.
- code/nlsy/5_merge_datasets.do: will merge the NLSY1979 and NLSY1997 datasets, as well as the task and skill measures from Deming (2017b). This yields the main data set for analysis.
- code/nlsy/6_analysis.do: will produce Table 3 in the main text of the paper, as well as Tables R3 and R6 in the appendix.
- code/nlsy/7_skills_by_occ.do: will aggregate the NLSY skill measures by the broad occupation categories used in model estimation, in preparation for the next step.
- code/nlsy/8_decompose_delta_eta.do: will produce Appendix Tables R4 and R5, combining the NLSY skill measures and the model estimates.

Log files are saved under **output/nlsy/logs/**.

Dependencies between programs

For more information on the dependency structure between programs (i.e., which programs call which other programs), refer to the Excel file "README_supplement.xlsx". The file also contains information on which figures are produced by which programs.

Structure of Output Directory

As noted above, the figures and tables in the paper are saved under **figures/**. However, should the replicators wish to inspect the code more closely, they may look at intermediate output files saved under the directory **output/**. The directory **output/acs/** and **output/nlsy/** save outputs from the analyses based on Census/ACS and NLSY, respectively. On the other hand, the directory **output/model/** saves intermediate outputs from the model estimation and analysis. The structure of this latter directory requires some explanation.

After running the programs, there will be five sub-directories under **output/model/**:

- **baseline_psi_d_4.5/**: will contain baseline results in the paper.
- **baseline_psi_d_3.5/**: will contain robustness results with $\psi = 3.5$.
- **baseline_psi_d_5.5/**: will contain robustness results with $\psi = 5.5$.
- **fixtheta_psi_d_4.5_theta_2.8/**: will contain robustness results with $\theta = 2.8$.
- **fixtheta_psi_d_4.5_theta_4.5/**: will contain robustness results with $\theta = 2.8$.

The robustness results are for the Online Appendix.

Next, within each of the sub-directories above, there will be three main folders:

- **region_all/**: will contain baseline results using the data for all regions of the US.
- **region_nonsouth/**: will contain results using the data for the South region of the US.
- **region_south/**: will contain results using the data for the Non-South region of the US.

The latter two folders are for the regional analysis in Section 7.1.

Each of the above folders will contain csv tables summarizing parameter estimates and model-implied moments. Those files are produced by code/model/depends/create_tables.m. Furthermore, it will nest two sub-folders:

- **sensitivity/**: will contain results of the sensitivity analysis. (Produced by code/model/depends/create_tables.m)
- **decompose_trends/:** will contain results from decomposing trends in moments over time. (Produced by code/model/depends/decompose_trends.m)

References

Autor, David, "David Autor's Data Archive," 2024. MIT Economics, Cambridge, MA https://economics.mit.edu/people/faculty/david-h-autor/data-archive.

_ , **Frank Levy, and Richard Murnane**, "The Skill Content of Recent Technological Change: An Empirical Exploration," *The Quarterly Journal of Economics*, 2003, *118*, 1279–1333.

Charles, Kerwin Kofi and Jonathan Guryan, "Prejudice and Wages: An Empirical Assessment of Becker's The Economics of Discrimination," *Journal of Political Economy*, 2008, *116* (5), 773–809.

Deming, David, "Replication Data for: The Growing Importance of Social Skills in the Labor Market," 2017. Havard Dataverse, Version 1, https://doi.org/10.7910/DVN/CYPKZH.

Deming, David J., "The Growing Importance of Social Skills in the Labor Market," *Quarterly Journal of Economics*, 2017, *132* (4), 1593–1640.

Hurst, Erik, Yona Rubinstein, and Kazuatsu Shimizu, "Task Based Discrimination," *American Economic Review*, Forthcoming.

Ruggles, Steven, Sarah Flood, Sophia Foster, Ronald Goeken, Jose Pacas, Megan Schouweiler, and Matthew Sobek, "IPUMS USA," 2021. Minneapolis, MN, Version 11.0, https://doi.org/10.18128/D010.V11.0.

U.S. Department of Labor, Bureau of Labor Statistics, "National Longitudinal Survey of Youth 1979 Cohort, Rounds 1-27," 2019. Distributed by the Center for Human Resource Research (CHRR), The Ohio State University, Columbus, OH, https://www.nlsinfo.org/content/cohorts/nlsy79/get-data.

_ , "National Longitudinal Survey of Youth 1997 Cohort, Rounds 1-18," 2019. Distributed by the Center for Human Resource Research (CHRR), The Ohio State University, Columbus, OH, https://www.bls.gov/nls/nlsy97.htm.

U.S. Department of Labor, Employment and Training Administration, "O*NET 98 Database," 1998. Used under the CC BY 4.0 license.

_ , "O*NET 3.1 Database," 2001. Used under the CC BY 4.0 license, www.onetcenter.org/db$_t$*ransitional.html*.

_ , "O*NET 26.1 Database," 2021. Used under the CC BY 4.0 license, www.onetcenter.org/db$_r$*eleases.html*.

_ , "O*NET-SOC Taxonomy," 2023. Used under the CC BY 4.0 license, www.onetcenter.org/taxonomy.html.