This file describes the Replication Package attached to the paper Wage Risk and the Skill Premium" (MS# 20200752) by H. Yazici and C. Slavik, editor in charge: Nezih Guner. There are 2 sets of files, which all run under standard Windows operating system (Windows 10, for example). In addition, Matlab and Microsoft Excel are needed to execute/view these files.

The package contains the file CalibrationData.xlsx, which contains the data used to generate statistics compared to model outcomes in the paper or used as calibration moments. The data files are also contained in .csv format in the folder CSV data files. The package also contains a folder Matlab Codes which contains the codes used to generate model statistics reported in the paper. Below, we provide the details.

1. The file CalibrationData.xlsx contains raw datasets that were manually copied into the xlsx file. We perform basic computations in this file to obtain statistics used in the paper. There are no codes used for these computations. The statistics we computed are reported in Table 1, as calibration targets in Table 2, in Table 4 of the paper, and in Table 11 in the Online Appendix. The details of the contents of the individual sheets follow. All the sheets are also saved separately as .csv files with the same name in the folder 'CSV data files'.

(i) Sheet 'q' contains the FRED series PERIC, which is used to compute the relative price of equipment capital for the 1967 steady state (average over 1967-1976, cell D41, normalized to 1) and the 1980s steady state (average over 1980–1989 relative to the average over 1967-1976 steady state, see cell D54) and 2010 steady states (average over 2001-2010 relative to the average over 1967-1976 steady state, see cell D75). The relative price of equipment in 1967 appears in Table 1 of the Manuscript and in 1980 in Table 11 of the Online Appendix.
Source: Federal Reserve Economic Data. Series: "Relative Price of Equipment."
Link: https://fred.stlouisfed.org/series/PERIC. Accessed: November 2017.
For additional details, see DiCecio (2009) as cited in the paper.

(ii) Sheet 'CPS67(80)' contains the CPS table p.17 used to compute the fraction of skilled agents in 1967 (cell H34 computed as the ratio between the number of people with 4 years of college or more in 1967 - cell B219 - divided by the total number of people in 1967 - cell B34) and 1980 (cell H21 computed at the ratio between the number of people with 4 years of college or more in 1980 - cell B206 - divided by the total number of people in 1980 - cell B21). Share of skilled agents in 1967 appears in Table 1 of the Manuscript and share of skilled agents in 1980 appears in Table 11 of the Online Appendix.

Source: U.S. Census Bureau, Current Population Survey, 1959 to 1991 Annual Social and Economic Supplements (CPS ASEC).
Link: https://www2.census.gov/programs-surveys/cps/tables/time-series/historical-income-people/p17.xlsx. Accessed: 2012-12-21.

(iii) Sheet 'CPS2010' contains the CPS table p.16 used to compute the fraction of skilled agents in 2010 (cell H14 computed at the ratio between the number of people with bachelor's degree or more in 2010 - cell B239 - divided by the total number of people in 2010 - cell B23). Share of skilled agents in 2010 appears in Table 1 of the Manuscript.
Source: U.S. Census Bureau, Current Population Survey, 1992 to 2021 Annual Social and Economic Supplements (CPS ASEC).
Link: https://www2.census.gov/programs-surveys/cps/tables/time-series/historical-income-people/p16.xlsx. Accessed: 2012-12-21.

(iv) Sheet 'GtoY' contains line 1 from NIPA table 1.1.5 (GDP, line 7 in our file) and line 21 from NIPA table 3.1 (Government consumption expenditures, line 35 in our file) which are used to compute government consumption-to-GDP ratio, G/Y. This is computed at the ratio of Government consumption expenditures to GDP, see line 17 and we take the average over 1967-2010. This is cell C18. Government consumption as a fraction of GDP appears in Table 1 of the Manuscript.
Source: U.S. Bureau of Economic Analysis, "National Income Accounts Table 1.1.5 Gross Domestic Product" and "National Income Accounts Table 3.1. Government Current Receipts and Expenditures".
Link: https://apps.bea.gov/iTable/iTable.cfm?reqid=19&step=2#. Accessed 2021-12-21.

(v) Sheet 'Gvt Debt' contains the FRED series FYGFGDQ188S, which is the U.S. Federal Debt Held by the Public as Percent of Gross Domestic Product. Government debt as fraction of GDP in 1967 and I 2010 appear in Table 1 of the Manuscript and in 1980 in Table 11 of the Online Appendix.
Source: Federal Reserve Economic Data.
Link: https://fred.stlouisfed.org/series/FYGFGDQ188S. Accessed: 2021-12-21.

(vi) Sheets 'NIPA 1.7.5' and 'NIPA 1.12' contain the NIPA tables that are used to compute the labor share in sheet 'Labor Share'. To compute the labor share in line 8, we follow the

methodology described in Rios-Rull, J.-V., and R. Santaeulalia-Llopis (2010): Redistributive shocks and productivity shocks," Journal of Monetary Economics, 57(8), 931-948. See Appendix A.2 of that paper for the definitions of the variables calculated on lines 2-7 of Sheet LaborShare. See also page 4 of the Online Appendix to our paper for details.

Source: Rios-Rull, J.-V., and R. Santaeulalia-Llopis (2010): Redistributive shocks and productivity shocks," Journal of Monetary Economics, 57(8), 931-948.

Data source: Source: U.S. Bureau of Economic Analysis, "National Income Accounts Table 1.7.5 Table 1.7.5. Relation of Gross Domestic Product, Gross National Product, Net National Product, National Income, and Personal Income." and "Table 1.12. National Income by Type of Income." Link: https://apps.bea.gov/iTable/iTable.cfm?reqid=19&step=2. Accessed April 2014.

(vii) Sheet 'Real GDP' contains the FRED series A939RX0Q048SBEA, which was used to calculate real U.S. GDP per capital growth between 1967 and 2010, see D55. The number appears in Table 4 of the paper.

Source: Federal Reserve Economic Data. Series: "Real Gross Domestic Product Per Capita." Link: https://fred.stlouisfed.org/series/A939RX0Q048SBEA . Accessed: 2021-12-21.

(viii) Sheet 'MacroStats' contains macro statistics. The first table contained in that sheet is the NIPA table 1.5.5, which is used to compute GDP net of housing investment and services, see line 71. The sheet also contains FAT table 1.1. This Table is used to compute total capital (line 112) and then the capital-to-output ratio in 1967 (cell C117) which is a calibration target and appears in Table 2 of the Manuscript.

Cell AU119 computes the growth rate of the K-to-Y ratio, which appears in Table 4 o the Manuscript. Line 120 of this sheet then computes the investment-to-capital ratio I/K, whose growth rate between 1967 and 2010 is in cell AU122 and Table 4 o the Manuscript. Finally line 114 computes the share of equipment in capital (private) and cell AU116 delivers the growth rate off this variable between 1967 and 2010, see also Table 4 of the Manuscript.

Source: U.S. Bureau of Economic Analysis, "National Income Accounts Table 1.5.5 Gross Domestic Product, Expanded Detail."

Link: https://apps.bea.gov/iTable/iTable.cfm?reqid=19&step=2# and "Fixed Asset Accounts Table 1.1 Current-Cost Net Stock of Fixed Assets and Consumer Durable Goods."

Link: https://apps.bea.gov/iTable/iTable.cfm?reqid=10&step=2#reqid=10&step=2.

Accessed 2021-12-21.

(ix) Sheet 'NFAP' contains the FRED time series used to calculate the NFAP (net foreign asset position). First, column E reports the Private sector wealth which equals U.S. Wealth, series BOGZ1FL892090005Q (column B) minus State and Local Governments; Nonfinancial Assets, series BOGZ1LM212010095Q (column C) minus Federal Government; Nonfinancial Assets, series BOGZ1LM315015005Q (column D). Column F contains the Rest of the World; Total Financial Assets, series ROWTASQ027S and column G contains the Rest of the World; Total Liabilities and Equity, series ROWTLEQ027S. Column H and I then calculate the Rest of the World; Total Financial Assets as a fraction of Private sector wealth and the Rest of the World; Total Liabilities and Equity as a fraction of Private sector wealth, respectively. The difference between the two is the NFAP as a fraction of wealth, see cell L96 for its value in the 1967 steady state, cell L184 for its value in the 1980 steady state and cell L268 for its value in the 2010 steady state. For additional details, please see the Online Appendix of the paper.

Source: Federal Reserve Economic Data. Series: "All Sectors; U.S. Wealth." Link: https://fred.stlouisfed.org/series/BOGZ1FL892090005Q. Accessed: 2021-03-23.

Federal Reserve Economic Data. Series: "State and Local Governments; Nonfinancial Assets." Link: https://fred.stlouisfed.org/series/BOGZ1LM212010095Q. Accessed: 2021-03-23.

Federal Reserve Economic Data. Series: "Federal Government; Nonfinancial Assets." Link: https://fred.stlouisfed.org/series/BOGZ1LM315015005Q. Accessed: 2021-03-23.

Federal Reserve Economic Data. Series: "Rest of the World; Total Financial Assets." Link: https://fred.stlouisfed.org/series/ROWTASQ027S. Accessed: 2021-03-23.

Federal Reserve Economic Data. Series: "Rest of the World; Total Liabilities and Equity." Link: https://fred.stlouisfed.org/series/ROWTLEQ027S. Accessed: 2021-03-23.

(x) Sheet 'Savings rate' reports the FRED series W206RC1Q156SBEA – gross savings as percentage of gross national income. On page 23 of the paper, we mention that the savings rate decreases from about 23% to 18% prior to the Great Recession, i.e. 2007. These two numbers are marked in the excel sheet.

Source: Federal Reserve Economic Data. Series: "Gross savings as percentage of gross national income." Link: https://fred.stlouisfed.org/series/W206RC1Q156SBEA. Accessed: 2021-09-09.

2. The folder Matlab Codes contains Matlab codes used to numerically solve the model. The folder contains various subfolders that contain the files used to solve the various version of the model that we discuss in the paper. The folder also contains a file CodeStructure.xlsx. This file describes in sheet 'codes' the details of the individual models that we solved in these various subfolders, as well where the results are used in the paper. It also contains approximate running time of the codes. Typically, we only use the variable 'sp' which is the skill premium, but in some cases, we use more output variables. All output variables from the codes and always stored in a vector called 'output', which contains the relevant statistics. The .xlsx file CodeStructure.xlsx, sheet 'output vector' contains the description of the variables that appear in the output vector for each and every code provided.

Instructions to run the codes. Select the model you want to replicate from the sheet 'codes' in file CodeStructure.xlsx. Open the file mentioned in column 'Code to run' in Matlab and run it in that folder. There is no need to copy anything anywhere, the folder contains all the inputs and subroutines that are need to run the code. Once the code finishes running, open up the 'output' vector and select the variable of interest. In most cases, the output vector is also stored as a separate .mat file output.mat. We recommend copying the output vector into the excel sheet 'output vector' so that the mapping between the vector and the variable definitions is clear. Some statistics reported in the data were computed using simple manipulations of the output vectors (such as growth rates between the 1967 and 2010 steady states). The tables reported in the paper were generated by manually copying various output from the output vectors into the tables. Typically, one table in the paper contains results from a number of distinct codes.

Example: To replicate model statistics in Table 5 in the paper, run the code 'CalibrationLRIfinal2.m' in subfolder 'Matlab Codes\Benchmark Open Economy\1967BenchmarkCalibration\' and save the output vector in excel sheet 'output vector'. This is the 1967 benchmark calibrated steady state model.

Then run the code 'SimulationLRIExoEduc.m' in subfolder 'Matlab Codes\Benchmark Open Economy\2010BenchmarkSimulation\' and save the output vector in the same excel sheet. Compare the entries in Table 5 to those from Matlab.

System setup used to run the codes

Processor: Intel® Xeon® CPU E3-1246 v3 @ 3.50GHz 3.5GHz

Installed memory (RAM): 16.0 GB (15.39 GB usable)

System type: 64-bit Operating System, x64-based processor

Operating System: Microsoft Windows 10 Enterprise LTSC Version 10.0 (Build 17763)

Java Version: Java 1.8.0_202-b08 with Oracle Corporation Java HotSpot(TM) 64-Bit Server VM mixed mode

MATLAB Version: 9.11.0.1809720 (R2021b) Update 1

MATLAB	Version 9.11	(R2021b)
Simulink	Version 10.4	(R2021b)
Econometrics Toolbox	Version 5.7	(R2021b)
Global Optimization Toolbox	Version 4.6	(R2021b)
Optimization Toolbox	Version 9.2	(R2021b)
Parallel Computing Toolbox	Version 7.5	(R2021b)
Statistics and Machine Learning Toolbox	Version 12.2	(R2021b)
Symbolic Math Toolbox	Version 9.0	(R2021b)

Data citations

(as they appear in the paper, for more details, see above)

Federal Reserve Bank of St. Louis (2021): "Federal Reserve Economic Data," https://fred.stlouisfed.org/. Last accessed: 2021-12-21. For details, see the appendix and the readme.doc file in the replication package.

U.S. Bureau of Economic Analysis (2021): "National Income Accounts," https://apps.bea.gov/iTable/. Last accessed: 2021-12-21. For details, see the appendix and the readme.doc file in the replication package.

U.S. Census Bureau (2021): "Current Population Survey," https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-income-people.html. Last accessed: 2021-12-21. For details, see the appendix and thereadme.doc file in the replication package.