Replication Instructions for "A Quantitative Theory of Domestic Outsourcing: The Role of Wage-Proportional Staffing Fees"

Overview

The code in this replication package constructs all figures and tables in "A Quantitative Theory of Domestic Outsourcing: The Role of Wage-Proportional Staffing Fees" using Matlab and STATA.

This replication package relies on publicly available data from the following sources. Additional details are included in the section "Data Availability and Provenance Statements", with data citations included in "References".

- Current Population Survey Annual Social and Economic Supplement (CPS ASEC) data retrieved from IPUMS (ipums.org).
- Current Population Survey Contingent Worker Survey (CPS CWS) retrieved from IPUMS (ipums.org).
- Monthly job-to-job transition rate data computed by Fujita, Moscarini, and Postel-Vinay (2024) (https://doi.org/10.1257/mac.20210076).
- American Community Survey (ACS) 1-year educational attainment estimates, retrieved from data.census.gov.
- Information on the median usual weekly nominal earnings for those with and without a Bachelor's degree and the noncyclical rate of unemployment retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org.

This replication package creates the results from the paper with two master files.

- `run_DO.do` uses STATA to generate empirical estimates the model is compared against and Tables 9 and 10.
- `run_DO.m` uses Matlab to generate all other tables and figures.

Detailed instructions on running these files are included in the section "Instructions to Replicators". The replicator should expect codes to run for about 3.25 hours to reproduce all results.

Data Availability and Provenance Statements

Statement about Rights

☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

The limited IPUMs CPS subsamples required for replication are included with explicit permission from IPUMS. Additional details are provided below.

Summary of Availability

☒ All data **are** publicly available.
☐ Some data **cannot be made** publicly available.
☐ **No data can be made** publicly available.

Details on each Data Source

- The data source used to generate the majority of empirical results is the Current Population Survey Annual Social and Economic Supplement (CPS ASEC) data retrieved from IPUMS (ipums.org).
- Supplementary statistics used for the calibration are derived from the Current Population Survey Contingent Worker Survey (CPS CWS) retrieved from IPUMS (ipums.org), monthly job-to-job transition rate data computed by Fujita, Moscarini, and Postel-Vinay (2024), American Community Survey (ACS) 1-year educational attainment estimates, and information on the median usual weekly nominal earnings for those with and without a Bachelor's degree and the noncyclical rate of unemployment retrieved from FRED, Federal Reserve Bank of St. Louis. This data is publicly available and is provided in the replication package.

The following table summarizes the raw data files used in the analysis and their sources, with further details below.

Data Name	Data Files	Provided	Citation
"CPS ASEC"	DO_ASEC_data.dta	True	Flood (2024)
"CPS CWS"	CPS_CWS_NextMonth.dta	True	Flood (2024)
"Fujita, Moscarini, and Postel-Vinay (2024)"	Raw Data Files/m1_data.xlsx	True	Fujita, Moscarini, and Postel-Vinay (2024)
"NROU"	Raw Data Files/m2_NROU.xlsx	True	U.S. Congressional Budget Office (2025)
"LEU0252918500A"	Raw Data Files/m9_LEU0252918500A.xlsx	True	U.S. Bureau of Labor Statistics (2025a)
"LEU0252917300A"	Raw Data Files/m9_LEU0252917300A.xlsx	True	U.S. Bureau of Labor Statistics (2025b)
"ACS Educational Attainment"	Raw Data Files/EduAttainment_2003.xls	True	U.S. Census Bureau
	Raw Data Files/EduAttainment_2004.xls	True	
	Raw Data Files/EduAttainment_2005.xls	True	
	Raw Data Files/EduAttainment_2006.xls	True	
	Raw Data Files/EduAttainment_2007.xls	True	
	Raw Data Files/EduAttainment_2008.xls	True	
	Raw Data FilesEduAttainment_2009.xls	True	
	Raw Data Files/EduAttainment_2010.xls	True	
	Raw Data Files/EduAttainment_2011.xls	True	
	Raw Data Files/EduAttainment_2012.xls	True	
	Raw Data Files/EduAttainment_2013.xlsx	True	
	Raw Data Files/EduAttainment_2014.xlsx	True	
	Raw Data Files/EduAttainment_2015.xlsx	True	
	Raw Data Files/EduAttainment_2016.xlsx	True	
	Raw Data Files/EduAttainment_2017.xlsx	True	
	Raw Data Files/EduAttainment_2018.xlsx	True	
	Raw Data Files/EduAttainment_2019.xlsx	True	

CPS ASEC

The paper uses data from the Current Population Survey Annual Social and Economic Supplement (CPS ASEC) data retrieved from IPUMS (ipums.org).

Data file: `DO_ASEC_data.dta`

These data are a subsample of the IPUMS CPS data available from cps.ipums.org. Any use of these data should be cited as follows:

> Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Megan Schouweiler, and Michael Westberry. 2024. "Integrated Public Use Microdata Series, Current Population Survey: Version 12.0 [dataset]."

The attached data file is intended only for replication. Individuals are not to redistribute the data without permission. Contact ipums@umn.edu for redistribution requests. For all other uses of these data, please access data directly via cps.ipums.org.

CPS CWS

For some supplemental statistics, the paper uses data from the Current Population Survey Contingent Worker Supplement (CPS CWS) retrieved from IPUMS (ipums.org).

Data file: `CPS_CWS_NextMonth.dta`

These data are a subsample of the IPUMS CPS data available from cps.ipums.org. Any use of these data should be cited as follows:

> Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Megan Schouweiler, and Michael Westberry. 2024. "Integrated Public Use Microdata Series, Current Population Survey: Version 12.0 [dataset]."

The attached data file is intended only for replication. Individuals are not to redistribute the data without permission. Contact ipums@umn.edu for redistribution requests. For all other uses of these data, please access data directly via cps.ipums.org.

Fujita, Moscarini, and Postel-Vinay (2024)

Data file: `Raw Data Files/m1_data.xlsx`

NROU, LEU0252918500A, and LEU0252917300A data from FRED

To obtain the non-cyclical rate of unemployment and relative wages for those with and without a Bachelor's degree reported in Table 2, data from FRED, Federal Reserve Bank of St. Louis are used. These are FRED data series [NROU], [LEU0252918500A], and [LEU0252917300A] for 2003-2019. A copy of the data is provided as part of this archive. The data are in the public domain.

Data files: In the folder `Raw Data Files`, these are the files `m2_NROU.xlsx`, `m9_LEU0252917300A.xlsx`, and `m9_LEU0252918500A.xlsx`.

ACS Educational Attainment Data

To obtain the percent of the population with a Bachelor's degree or more reported in Table 1, data from the American Community Survey (ACS) 1-year estimates for 2003-2019 are used. The data can be downloaded directly from data.census.gov. A copy of the data is provided as part of this archive. The data are in the public domain.

Data files: In the folder `Raw Data Files`, these are the files `EduAttainment_2003.xls`, `EduAttainment_2004.xls`, `EduAttainment_2005.xls`, `EduAttainment_2006.xls`, `EduAttainment_2007.xls`, `EduAttainment_2008.xls`, `EduAttainment_2009.xls`, `EduAttainment_2010.xls`, `EduAttainment_2011.xls`, `EduAttainment_2012.xls`, `EduAttainment_2013.xlsx`, `EduAttainment_2014.xlsx`, `EduAttainment_2015.xlsx`, `EduAttainment_2016.xlsx`, `EduAttainment_2017.xlsx`, `EduAttainment_2018.xlsx`, and `EduAttainment_2019.xlsx`.

Dataset list

The following table summarizes all of the data files, both raw and derived, and their location, broken down by category.

Data file	Source	Provided
`DO_ASEC_data.dta`	Flood (2024)	Yes
`CPS_CWS_NextMonth.dta`	Flood (2024)	Yes
`Raw Data Files/m1_data.xlsx`	Fujita, Moscarini, and Postel-Vinay (2024)	Yes
`Raw Data Files/m2_NROU.xlsx`	U.S. Congressional Budget Office (2025)	Yes
`Raw Data Files/m9_LEU0252918500A.xlsx`	U.S. Bureau of Labor Statistics (2025a)	Yes
`Raw Data Files/m9_LEU0252917300A.xlsx`	U.S. Bureau of Labor Statistics (2025b)	Yes
`Raw Data Files/EduAttainment_2003.xls`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2004.xls`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2005.xls`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2006.xls`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2007.xls`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2008.xls`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2009.xls`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2010.xls`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2012.xls`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2013.xlsx`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2014.xlsx`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2015.xlsx`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2016.xlsx`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2017.xlsx`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2018.xlsx`	U.S. Census Bureau	Yes
`Raw Data Files/EduAttainment_2019.xlsx`	U.S. Census Bureau	Yes

Computational requirements

Software Requirements

- Stata - the code was last run with version 19.5
  - `estout` (as of 2025-05-29)
- Matlab - the code was last run with Matlab Release 2024b

Controlled Randomness

☒ Random seed is set at line 65 of program `DO_sim_func.m`.
Random seed is set at line 71 of program `DO_sim_decomp_func.m`.
Random seed is set at line 436 of program `DO_Appendix_E1.m`.
Random seed is set at line 458 of program `DO_Appendix_E2.m`.
Random seed is set at line 587 of program `DO_Appendix_E3.m`.

Random seed is set at line 582 of program `DO_Appendix_E4.m`.

☐ No Pseudo random generator is used in the analysis described here.

Memory, Runtime, Storage Requirements

Summary Approximate time needed to reproduce the analyses on a standard (2025) desktop machine:

☐ <10 minutes
☐ 10-60 minutes
☐ 1-2 hours
☒ 2-8 hours
☐ 8-24 hours
☐ 1-3 days
☐ 3-14 days
☐ > 14 days

Approximate storage space needed:

☐ < 25 MBytes

☐ 25 MB - 250 MB

☐ 250 MB - 2 GB

☒ 2 GB - 25 GB

☐ 25 GB - 250 GB

☐ > 250 GB

☐ Not feasible to run on a desktop machine, as described below.

Details The code was last run on a **12-core Intel-based laptop with Windows 11 Pro Version 24H2 with 700GB of free space**.

Description of programs/code

- The file `run_DO.do` generates empirical estimates from the CPS ASEC and CPS CWS that the model is compared against and uses CPS ASEC data to create Tables 9 and 10.
- The file `run_DO.m` creates all tables and figures, apart from Tables 9 and 10, by calling on the following files:
  - `DO_combine_moments.m`: This file combines the raw data files found in the folder "Raw Data Files" to compute data moments reported in Tables 1 and 2.
  - `DO_main_func.m`: This file solves the model presented in Section 4. It uses value function iteration to solve for the value and policy functions. Then it uses the policy functions to solve for the steady state distribution of agents across states.
  - `DO_sim_func.m`: This file runs a simulation of 20,000 agents. The results here are needed to generate the model moments regarding wages reported in Table 2.
  - `DO_moments_func.m`: This file combines the results generated by `DO_main_func.m` and `DO_sim_func.m` to generate the model moments reported in Table 2 and other moments of interest reported in Section 5.
  - `DO_sim_decomp_func.m`: This file creates the results reported in Appendix D.1. Specifically, it considers how the estimated model wage penalty associated with outsourcing presented in Table 3 and the wage growth difference in Table 4 would differ in a model without selection into outsourcing.
  - `DO_transitions_func.m`: This file creates the results reported in Appendix D.5. Specifically, it considers how the economy transitions from the steady state with unrestricted outsourcing to the steady state where outsourcing in new jobs is completely banned.
  - `DO_Appendix_E1.m`: This file creates the results reported in Appendix E.1. It shows that equivalent model results can be generated under the assumption that there are complete contracts between employees and their legal employer (which is the staffing agency in outsourced jobs), but not between outsourced workers and client firms.

- `DO_Appendix_E2.m`: This file creates the results reported in Appendix E.2. It reconsiders the results generated using the model described in Section 4, with the additional assumptions that 5% and 10% of human capital is job-specific.
- `DO_Appendix_E3.m`: This file creates the results reported in Appendix E.3. It considers how model results would differ with the extension that jobs may differ in terms of their separation risk.
- `DO_Appendix_E4.m`: This file creates the results reported in Appendix E.4. It considers an extension to the model presented in Section 4 where both outsourced and non-outsourced jobs can be created in sectors that offer different returns to human capital investments.

Instructions to Replicators

- Run the file `run_DO.do`
  - Expect this file to take approximately 3 minutes to run in full.
  - Note that this step must be completed before continuing on to the next step.
- Run the file `run_DO.m`
  - Expect this file to take approximately 3 hours to run in full.

List of tables and programs

The provided code reproduces:

☐ All numbers provided in text in the paper
☒ All tables and figures in the paper
☐ Selected tables and figures in the paper, as explained and justified below.

Note that Figure 2 is a plot of the assumed timeline of events in the model and is created in LaTeX rather than Matlab or STATA.

List of Tables and Figures

Figure/Table	Program	Line	Output file
Figure 1a	`run_DO.m`	32	Figure1a.eps
Figure 1b	`run_DO.m`	58	Figure1b.eps
Table 1	`run_DO.m`	125	Table1.tex
Figure 3a	`run_DO.m`	144	Figure3a.eps
Figure 3b	`run_DO.m`	164	Figure3b.eps
Table 2	`run_DO.m`	212	Table2.tex
Table 3	`run_DO.m`	255	Table3.tex
Table 4	`run_DO.m`	289	Table4.tex
Table 5	`run_DO.m`	332	Table5.tex
Table 6	`run_DO.m`	409	Table6.tex
Figure 4a	`run_DO.m`	419	Figure4a.eps
Figure 4b	`run_DO.m`	429	Figure4b.eps
Figure 5a	`run_DO.m`	445	Figure5a.eps
Figure 5b	`run_DO.m`	473	Figure5b.eps
Figure 6	`run_DO.m`	495	Figure6.eps
Table 7	`run_DO.m`	749	Table7.tex
Figure 7	`run_DO.m`	772	Figure7.eps
Table 8	`run_DO.m`	944	Table8.tex
Figure 8	`run_DO.m`	1038	Figure8.eps
Table 9	`run_DO.do`	555	Table9.tex
Table 10	`run_DO.do`	611	Table10.tex
Table 11	`run_DO.m`	1086	Table11.tex
Table 12	`run_DO.m`	1109	Table12.tex
Figure 9	`run_DO.m`	1249	Figure9.eps
Figure 10a	`run_DO.m`	1262	Figure10a.eps
Figure 10b	`run_DO.m`	1276	Figure10b.eps
Figure 10c	`run_DO.m`	1290	Figure10c.eps
Figure 10d	`run_DO.m`	1304	Figure10d.eps
Table 13	`run_DO.m`	354	Table13.tex
Table 14	`run_DO.m`	369	Table14.tex
Table 15	`run_DO.m`	1399	Table15.tex
Table 16	`run_DO.m`	1422	Table16.tex
Table 17	`run_DO.m`	1462	Table17.tex
Table 18	`run_DO.m`	1514	Table18.tex
Figure 11a	`run_DO.m`	1527	Figure11a.eps
Figure 11b	`run_DO.m`	1541	Figure11b.eps
Figure 12	`run_DO.m`	1564	Figure12.eps
Table 19	`run_DO.m`	1598	Table19.tex
Table 20	`run_DO.m`	1642	Table20.tex
Table 21	`run_DO.m`	1676	Table21.tex
Table 22	`run_DO.m`	1718	Table22.tex
Table 23	`run_DO.m`	1771	Table23.tex
Figure 13a	`run_DO.m`	1791	Figure13a.eps
Figure 13b	`run_DO.m`	1808	Figure13b.eps
Table 24	`run_DO.m`	1842	Table24.tex
Table 25	`run_DO.m`	1886	Table25.tex
Table 26	`run_DO.m`	1936	Table26.tex
Figure 14a	`run_DO.m`	1956	Figure14a.eps
Figure 14b	`run_DO.m`	1973	Figure14b.eps
Table 27	`run_DO.m`	2007	Table27.tex

References

[1] Sarah Flood, Miriam King, Renae Rodgers, Steven Ruggles, J. Robert Warren, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Megan Schouweiler, and Michael Westberry. Integrated Public

Use Microdata Series, Current Population Survey: Version 12.0 [dataset]. Minneapolis, MN: IPUMS, 2024. https://doi.org/10.18128/D030.V12.0

[2] Fujita, S., Moscarini, G., & Postel-Vinay, F. (2024). Time series for the average probability of US workers making employer-to-employer transitions between month t and month t-1 [Data set]. In Measuring Employer-to-Employer Reallocation. American Economic Journal: Macroeconomics, 16(3), 1–51. https://doi.org/10.1257/mac.20210076.

[3] U.S. Census Bureau. (2003–2019). American Community Survey 1-Year Estimates, Table S1501: Educational Attainment [Data set]. data.census.gov.

[4] U.S. Bureau of Labor Statistics, Employed full time: Median usual weekly nominal earnings (second quartile): Wage and salary workers: Bachelor's degree and higher: 25 years and over [LEU0252918500A], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/LEU0252918500A, April 4, 2025.

[5] U.S. Bureau of Labor Statistics, Employed full time: Median usual weekly nominal earnings (second quartile): Wage and salary workers: High School graduates, no college: 25 years and over [LEU0252917300A], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/LEU0252917300A, April 4, 2025.

[6] U.S. Congressional Budget Office, Noncyclical Rate of Unemployment [NROU], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/NROU, April 4, 2025.