Estimates using the hazard function estimator

The "format_LFS" programs assemble the mini-panels from the raw monthly masterfiles of the Labour Force Survey (LFS) available through the Statistics Canada Research Data Centre (RDC). The variables needed to match records across months are not available in the public use microfiles (PUMFs). Information on how to access this data is available at: https://www.statcan.gc.ca/en/microdata/data-centres/access

Additional data sources:

Historical provincial minimum wages are from

https://www.edsc-esdc.gc.ca/ouvert-open/labour-travail/WD-datatable-histminwage-20180724-ENG.csv

and were downloaded on March 7, 2019.

This data is also available at:

https://minwage-salairemin.service.canada.ca/en/since1965.html

The monthly consumer price index (CPI) data is from Statistics Canada Table 18-10-0004-01 ("Consumer Price Index, All-items monthly, not seasonally adjusted.") and was downloaded on March 7, 2019.

The minimum wage data and the CPI data was combined in a data set with the monthly CPI and provincial minimum wages.

Programs (in the order to be run to replicate our results):

master.do
=========

This program will require some modification by the user to run, as many of the directories are set within the subsequent programs. However, it does list the programs sequentially as they are required to replicate the results in the paper.

format_LFS_part1.do
===================

-assembles the "raw" monthly LFS data into yearly STATA files. The format of the raw data has changed over the years (either ASCII files or STATA files). The code will need to be changed to reflect a researcher's own directory structure as well as the current location of the master files. Collection of wage data began in September 1996. However, we do not use the LFS data from 1996 in our analysis.

Input files: The raw LFS monthly data files available through the RDC.
Output files: The above data appended into annual STATA files named LFS`year'.dta.

format_LFS_part2.do
==========

- uses the LFS data from part 1 to put together mini-panels following the first four months after joining the in-rotation group (including the initial month).
-One output file per in-rotation group (e.g. the panel for those individuals entering the LFS in January of 1997) is saved in a file called "panel_6M_`month1'`year1'

Input files: The annual STATA files named LFS`year'.dta
(generated by format_LFS_part1.do.)
Output files: The individual data files containing the "mini-panels for each in-rotation group: panel_6M_`month1'`year1'.dta
(where `month1' indicates the first month in the LFS and `year1' indicates the first year.)

format_LFS_part3.do
==========

-Appends the files generated by format_LFS_part2.do into a single file.

Input files: panel_6M_`month1'`year1'.dta
Output file: panel_6M_allyears.dta

format_LFS_part4.do
==========

-Applies criteria to limit the sample (e.g. ages 15 to 59, not in military in inrotation month)
-creates various "flags" that indentify potential mismatches in the mini-panels (for example, is the reported
tenure consistent with what was reported in the previous month) and fully imputed records.
Resulting file is used as the input (i.e. data) file used to estimate the densities.

Input file: panel_6M_allyears.dta
Output file: panel_6M_sample.dta

joiners_main_program.do
==========

-Estimates the following models reported in Tables 1 and 3: "All Workers", "Stayers" and "Joiners"

Input files: panel_6M_sample.dta, min_wag_1996_2017.dta
Output files: Saved as joiners_`sex'_`k'_`j'_v1_5.ster, where `sex'=1 (male), 2(female), `k'=1 (no restrictions
based on type of wages) and `j'=1 (all), 2 (joiners) and 4 (stayers).
Released output: The worksheet Model_estimates_November_27_2023_RELEASED.xlsx contains the parameter estimates for
the models listed above, using the convention indicated above.

leavers_main_program.do
==========

-Estimates the following models reported in Tables 1 and 3: "Leavers"

Input file: panel_6M_sample.dta, min_wag_1996_2017.dta
Output files:
Model estimates aved as joiners_`sex'_`k'_`j'_v1_5.ster, where `sex'=1 (male), 2(female), `k'=1 (no restrictions
based on type of wages) and `j'=2 (leavers) and 3 (stayers; not reported in the paper).
dollar_bins.dta (location of wage bins containing integer values of nominal wages e.g. $4/hr in any given month).
Released output: The worksheet
Model_estimates_November_27_2023_RELEASED.xlsx contains the parameter estimates for
the models listed above, using the convention indicated above.

joiners_main_program_interact.do
==========

-Estimates various models in which the minimum wage effects are interacted with the magnitude of the real minimum wage.

Input files: panel_6M_sample.dta, min_wag_1996_2017.dta
Output files: Saved as joiners_`sex'_`k'_`j'_spec_4_int_a.ster, where `sex'=1 (male), 2(female) and 3(males and females pooled), `k'=1 (no restrictions
based on type of wages) and `j'=1 (all).
Released output: The workbook "Model_dube_and_interact.xlsx" contains the parameter estimates for
the models listed above, in sheets using the naming convention indicated above. The results in Table 2 are for the pooled model (i.e. estimated using both males and females).

joiners_main_program_covariates.do
==========

-Estimates variants of the models that compare difference-in-difference estimates similar to the Cengiz et al. () specifications to the triple difference-in-difference estimates reported in the paper.
Variants also exclude the dummy indicating that a wage bin contains wages at nominal dollar intervals (e.g $5, $6, $7, etc.) to control for wage bunching. None of these are directly reported in the paper. However, the baseline hazard
from one of the Cengiz specification is used to generate Figures 4 and Figures 6. These are joiner_3_1_1_dube_2 results listed below and were used as a means of smoothing the wage distribution.

Input files: panel_6M_sample.dta, min_wag_1996_2017.dta
Output files: Saved as joiners_`sex'_`k'_`j'_`suffix'.ster, where `sex'=1 (male), 2(female) and 3(males and females pooled), `k'=1 (no restrictions

based on type of wages) and `j'=1 (all). `Suffix' is as follows:

dube_2 - Provincial, year and quarter dummies for all five wage partitions. No "dollar" dummy indicating that bin contains wages.
dube_3 - As above, but also includes the "dollar" dummy.
spec_4 - This is the baseline model (see "joiners_main_program.do" above).

Released output: The workbook "Model_dube_and_interact.xlsx" contains the parameter estimates for
the models listed above, in sheets using the naming convention indicated above.

joiners_main_program_no_covariates.do
==========

-These models were estimated for comparative purposes and exclude the individual controls for age, and educational attainment of individuals.Estimates variants
of the models that compare difference-in-difference estimates similar to the Cengiz et al. (2019) specifications to the triple difference-in-difference estimates reported in the paper.
Variants also exclude the dummy indicating that a wage bin contains wages at nominal dollar intervals (e.g $5, $6, $7, etc.) to control for wage bunching. None of these results are used in the final paper.

Input files: panel_6M_sample.dta, min_wag_1996_2017.dta
Output files: Saved as joiners_`sex'_`k'_`j'_`suffix'.ster, where `sex'=1 (male), 2(female) and 3(males and females pooled), `k'=1 (no restrictions based on type of wages) and `j'=1 (all). `Suffix' is as follows:

dube_1 - Provincial, year and quarter dummies for all five wage partitions. No "dollar" dummy indicating that bin contains wages.
spec_5 - This is the baseline model (see "joiners_main_program.do" above) but without the individual controls.

Released output: The workbook "Model_dube_and_interact.xlsx" contains the parameter estimates for
the models listed above, in sheets using the naming convention indicated above.

fitted_densities_interact.do
==========

Uses the baseline hazard from the "dube_2" model and the minimum wage effect estimates from the "spec_4_int_a" model to generate fitted densities around a minium wage of $6.75 replacing the "level" interaction terms with minimum wages of $7.75 and $8.75. (Figure 4 in the paper uses both the actual minimum wage at $6.75 and the effect at $6.75 setting the interaction term to $7.75. This is Figure 8 in the paper.

Input files: sheets "joiners_3_1_1_dube_2" and
"joiners_3_1_1_spec_4_int_a" from the workbook

"Model_dube_and_interact.xlsx". dollar_bins.dta (indicators for wage bins containing integer values of the nominal wage).

Output file: Saved as worksheets named "joiners_`sex'_1_`j'_spec_4_int_a.xlsx", where `sex'=1 (male), 2(female) and 3(males and females pooled) and `j'=1 (all). These are compiled as separate work sheets in
the workbook"fitted_densities_interact_775.xlsx".

joiners_main_program_1_year.do
==========

Estimates models for joiners as well as models with extended dynamics (i.e. interaction terms for a minimum wage change in the last 3-12 months.

Input file: panel_6M_sample.dta, min_wag_1996_2017.dta
Output files: Saved as joiners_`sex'_`j'_original.ster, where `sex'=1 (male), 2 (female) and 3 (pooled); `k'=1 (no restrictions based on type of wages) and `j'=1 (full sample), `j'=2 (new hires; <3 months of job tenure),
`j'=3 (up to one year of job tenure), `j'=4 (stayers), `j'=5 (stayers with up to 12 months of job tenure),
`j'=6 (stayers with 1-5 years of job tenure), `j'=7 (stayers with more than 5 years of job tenure).

Results for women for `j'=5 and `j'=7 are reported in Figure 15.

The results with the "_1_lag are experimental results used for robustness testing and are only reported graphically in Figure 10 of the paper.

Released output: The worksheet joins_1_lag_WTD_RELEASED.xlsx contains the parameter estimates for
the models listed above, using the convention indicated above.

fitted_densities_lag.do
==========

Uses the baseline hazard from the "dube_2" model and the minimum wage effect estimates from the "joiners_1_lag" model to generate fitted densities from a minium wage of $6.75 to $7.25 using the dynamic "switches" (1-2 months after
an increase, 3-12 months after an increase. (This is Figure 10 in the paper.)

Input files: sheets "joiners_`sex'_1_1_dube_2" from the workbook "Model_dube_and_interact.xlsx" and sheets "joiners_`sex'_1_4_1_lag from the workbook "joiners_1_lag.xlsx". dollar_bins.dta (indicators for wage bins containing integer values of the nominal wage).

Output file: Saved as worksheets named Saved as worksheets named "fitted_densities_`sex'_1_`j'_1_lag.xlsx", where `sex'=1 (male), 2(female) and 3(males and females pooled) and `j'=1 (all). The numbers underlying Table 10 are in the workbook "fitted_densities_3_1_4_1_lag.xlsx".

leavers_case_1_1_sep_12_2024_1_forward.do
==========

-Estimates models for leavers for the case of layoffs. Results are reported in Figure 11.

Input file: panel_6M_sample.dta, min_wag_1996_2017.dta
Output files: Saved as layoffs_`sex'_`j'_layoff_original_forward.ster, where `sex'=1 (male), 2 (female) and 3 (pooled); `k'=1 (no restrictions based on type of wages) and `j'=1 (no restrictions on duration of pre-layoff job tenure or) `j'=2(less than one year of job tenure prior to layoff). Only the results for women are reported (Figure 11).

Released output: The worksheet quitters_1_lag_forward_WTD_RELEASED.xlsx contains the parameter estimates for the models listed above, using the convention indicated above.

leavers_case_2_1_sep_12_2024_1_forward.do
==========

-Estimates models for leavers for the case of quits. Results are reported in Figure 11.

Input file: panel_6M_sample.dta, min_wag_1996_2017.dta
Output files: Saved as quits_`sex'_`j'_quit_original_forward.ster, where `sex'=1 (male), 2 (female) and 3 (pooled); `k'=1 (no restrictions based on type of wages) and `j'=1 (no restrictions on duration of pre-quit job tenure or) `j'=2(less than one year of job tenure prior to quitting) . Only the pooled results are reported (Figure 12).

Released output: The worksheet quitters_1_lag_forward_WTD_RELEASED.xlsx contains the parameter estimates for the models listed above, using the convention indicated above.

joiners_main_program_oct_29_2024_seven.do
==========

-Estimates models for joiners extending minimum wage effects up to $7 above the minimum wage.. Results are reported in Figure A2 and A3.

Input files: panel_6M_sample.dta, min_wag_1996_2017.dta

Output files: Saved as joiners_`sex'_`k'_`j'_seven.ster, where `sex'=1 (male), 2(female), `k'=1 (no restrictions
based on type of wages) and `j'=1 (all), 2 (joiners) and 4 (stayers).
Released output: The worksheet joiners_1_seven_WTD_RELEASED.xlsx contains the parameter estimates for
the models listed above, using the convention indicated above.

Results in Section F of the On-line appendix

Note: several programs that appear before are based on programs made below were developed by modifying code that were made available by Cengiz et al. (2019) under a Creative Commons license (see citation below).

format_LFS_part_4_mod_Cengiz.do
==========

The density estimator only uses workers that were employed in a given month. Cengiz et. al use employment rates, so individuals that are not employed are now included.

-Applies criteria to limit the sample (e.g. ages 15 to 59, not in military in in-rotation month)
-creates various "flags" that identify potential mismatches in the mini-panels (for example, is the reported
tenure consistent with what was reported in the previous month) and fully imputed records.
Resulting file is used as the input (i.e. data) file used to estimate the densities.

Input file: panel_6M_allyears.dta
Output file: panel_6M_sample_mod.dta

1 Minimum Wage Bins.do
==========

-merges the labour force data with the minimum wage (and CPI) data.
-Use the full sample to generate estimates of the population and the employed by month and province.
These are used to renormalize the overall employment rates at the province/month level after the flag identifying
potential mismatches in the panels is applied.
-Assigns employed individuals to wage bins using the conventions used in Cengiz et al.

Input files: min_wag_1996_2017.dta, panel_6M_sample_mod.dta
Output file: Y_variable

2 Minimum Wage Assign.do
==========

-Uses code from the Cengiz et al. program "state_panels_tercile1979_QJE" to create the minimum wage variables and the province-quarter panel data used to implement the

Cengiz et al. estimator. Note that unlike the U.S., in Canada all provinces set their own minimum wages and although there is a federal minimum wage, very few workers are affected.

Input file: Y_variable.dta

Output files: new_y2.dta

create_programs.do
==========

This is a modified version of the Cengiz et al. program "create_programs.do"
It defines a number of programs that are called in the subsequent program used to estimate the model underlying Figure 2 in Cengiz et al. (2019) and Figures A13-A15 in Brochu et al.

It is called upon in the subsequent program "3_Figure2.do"

Input file: new_y2.dta

3 Figure2.do
============

Estimates the treatment effects for the minimum wage. Note that "create_programs.do" is called within the program, which creates necessary programs, including the one that loads the data set "new_y2_dta" used in the estimation. Presumably due to the large number of fixed effects in the U.S., the fixed effects estimation command is run three times, with different treatment variables included in the absorbed portion of the model each time.

Output files:
forplacebofig_true`limitsample'`limittreat'_`s'_`Y'_`b'_w`w'_`temp1'_`temp2'_${Tmax}${Tmin}

forplacebofig_pl1`limitsample'`limittreat'_`s'_`Y'_`b'_w`w'_`temp1'_`temp2'_${Tmax}${Tmin}

forplacebofig_pl2`limitsample'`limittreat'_`s'_`Y'_`b'_w`w'_`temp1'_`temp2'_${Tmax}${Tmin}

Released output: Cengiz_joiners_stayers_WTD_RELEASED_with_diagrams.xlsx (first three worksheets; details below)
Cengiz_5_year_with_std_error.xlsx (contains the cumulative 5-year post-treatment effects, along with standard erros, for the models below).

The sheets "estimates_1", "estimates_2" and "estimates_10" in the spreadsheet "cengiz_joiners_stayers_WTD_RELEASED_with_diagrams.xlsx" contain the estimated minimum wage effects for three specifications, where the final number indicates:
1 (basic specification reported in Figure A13),
2 (additional province-bin trends; not reported in Appendix)

10 (additional province-bin trends AND province-bin quadratics; not reported in Appendix).

The first 10 columns are the underlying parameter estimates (and associated statistics). The remaining columns are the cumulative effects on
employment over the five years following the change. These are calculated again in the program "plot.do" using the estimates directly.

1 Minimum Wage Bins_together.do
==========

-merges the labour force data with the minimum wage (and CPI) data.
-Use the full sample to generate estimates of the population and the employed by month and province.
These are used to renormalize the overall employment rates at the province/month level after the flag identifying
potential mismatches in the panels is applied.
-Assigns employed individuals to wage bins using the conventions used in Cengiz et al.
-Constructs bins separately for recent hires (up to one year of job tenure) and more established workers (>1 year of job tenure). Groups are indicated by the dummy variable "recent_emp" (0=established, 1= recent hire).

Input files: min_wag_1996_2017.dta, panel_6M_sample_mod.dta
Output file: Y_variable_joiners.dta

2 Minimum Wage Assign_together.do
==========

-Uses code from the Cengiz et al. program "state_panels_tercile1979_QJE" to create the minimum wage variables and the province-quarter panel data used to implement the
Cengiz et al. estimator. Note that unlike the U.S., in Canada all provinces set their own minimum wages and although there is a federal minimum wage, very few workers are affected.

Input file: Y_variable_joiners.dta

Output files: new_y2_`recent_emp'.dta where `recent_emp= 0 (more than 12 months of job tenure) and 1 (up to one year of job tenure).

create_program_together.do
==========

This is a modification of the Cengiz et al. program "create_programs.do" It defines a number of programs that are called in the subsequent program used to estimate the model underlying Figure 2 in Cengiz et al. (2019) and Figures A13-A15 in Brochu et al. (2025).

It is called upon in the subsequent program 3_Figure_2_together.do

Input file: Depending on which value the local variable `recent_emp' takes in the following program, the input data set is either `new_y2_0.dta'
or `new_y2_1.dta' (see description of program "2 Minimum Wage Assign_together.do" for more details.

3 Figure2_together.do
==========

Estimates the treatment effects for the minimum wage. Note that "create_programs_together.do" is called within the program, which creates necessary programs, including the one that loads the data set "new_y2_dta" used in the estimation. Presumably due to the large number of fixed effects in the U.S., the fixed effects estimation command is run three times, with different treatment variables included in the absorbed portion of the model each time.

Output files:
forplacebofig_true`limitsample'`limittreat'_`s'_`Y'_`b'_w`w'_`temp1'_`temp2'_${Tmax}${Tmin}

forplacebofig_pl1`limitsample'`limittreat'_`s'_`Y'_`b'_w`w'_`temp1'_`temp2'_${Tmax}${Tmin}

forplacebofig_pl2`limitsample'`limittreat'_`s'_`Y'_`b'_w`w'_`temp1'_`temp2'_${Tmax}${Tmin}

Released output: Cengiz_joiners_stayers_WTD_RELEASED_with_diagrams.xlsx (last two worksheets; details below)

The sheets "stayers", "joiners" and "estimates_10" in the spreadsheet "cengiz_joiners_stayers_WTD_RELEASED_with_diagrams.xlsx" contain the estimated minimum wage effects for models with province fixed effects for joiners (or "entrants" using Cengiz et al.'s terminology) and stayers (or "incumbants" using Cengiz et al.'s terminology).

plot.do
=======

Program that uses estimates to generate the figures A13-A15.

Input: Cengiz_5_year_all.dta, Cengiz_stayers_1_year.dta, Cengiz_joiners_1_year.dta (these files were created using the worksheets “joiners” and “stayers” respectively).

Notes: Cengiz_5_year_all.dta are the five-year cumulative post-treatment effects and include upper and lower bounds of the confidence error calculated using the relevant standard errors.

Cengiz_stayers_1_year.dta and Cengiz_1_year.dta are the one-year post-treatment effects and include upper and lower bounds of the confidence error

Output: Cengiz_5_year_all.pdf (Figure A13), Cengiz_joiners_1_year.pdf, and Cengiz_stayers_1_year.pdf (Figure A15)

Citation for Cengiz et al. Replication package
==========

Cengiz, Doruk; Dube, Arindrajit; Lindner, Attila; Zipperer, Ben, 2019, "Replication Data for: 'The Effect of Minimum Wages on Low-Wage Jobs'", https://doi.org/10.7910/DVN/TJCTC7, Harvard Dataverse, V2