Spartan Stata Graphics*

Brendan M. Price
Federal Reserve Board

December 30, 2022

Stata is quite capable of producing slick and effective graphics, but doing so requires both good default settings and (usually) a fair bit of manual tinkering. In this note, I introduce a minimalist scheme called `spartan` and accompanying code to create a set of sample figures using publicly available data. The main contribution is not so much the scheme file—other good schemes are out there—but the worked examples, which span a variety of graph types commonly used in economic research and grapple with numerous practical issues that commonly arise when working with Stata graphics. Adapt freely.

The `spartan` scheme. The enclosed file `spartan.scheme` modifies Stata's built-in `s1color` scheme to reduce clutter, standardize text sizes, accommodate color blindness, and otherwise tweak the graphical defaults in ways I find desirable. Users can simply edit the scheme file to remove or modify any configurations they dislike.

An ideal color palette is aesthetically pleasing, sufficiently high-contrast, color-blind-friendly, and printer-friendly. Rather than reinvent the wheel, `spartan` adapts a palette developed by Paul Tol (`https://personal.sron.nl/~pault`), with the exception that I swap in a shade of blue I prefer to his. You can, of course, swap in your favorite colors.[1]

Both in the scheme file and when producing individual figures, my aesthetic choices are in the spirit of Edward Tufte's *The Visual Display of Quantitative Information*, whose overriding principles are to "show the data" and minimize visual clutter. For a recent treatment targeted at an economics audience, see Jonathan Schwabish's book *Better Presentations*.

Replication code. The enclosed `code` directory contains `.do` files that download and process source data from the US Department of Labor, the NBER-CES Manufacturing Industry Database, and the National Center for Health Statistics, then produce the figures that appear in this document. The `lib` directory contains `spartan.scheme` as well as font files for producing LaTeX-compatible figures.

I tested the code in MacOS, using Stata version 15.1 in console mode. The code should also work on a Mac in GUI mode, except that Stata will use its default font and some

*Email: `brendan.m.price@frb.gov`. The views expressed here are mine and do not necessarily represent the views or policies of the Board of Governors of the Federal Reserve System or its staff.

[1]See William Buchanan's package `brewscheme` and Ben Jann's package `palettes`, both available via `ssc`.

inessential shell utilities won't execute properly.[2] Most of the code should work fine in a Windows or Linux environment, but users will probably have to make some modifications.

Sample figures. The figures represent a broad cross-section of the kinds of graphs commonly produced by economists. I plot data series using Stata's `twoway` and `graph bar` commands, and I plot regression estimates using Ben Jann's superb `coefplot` command, available from `ssc`. I make heavy use of options to make each figure as clear and effective as possible. In doing so, I illustrate solutions to many technical issues and glitches that arise in practice. While I largely opt for stylistic consistency, I present a few stylistic variants (e.g., axis labels and ticks) to showcase some of the available options. The `.do` file headers include some brief commentary on my aesthetic choices and Stata implementations.

Fonts. I produce all figures in this document using the CMU Serif font used natively by LaTeX. Slide decks produced using the LaTeX `beamer` package default instead to CMU Sans Serif, which looks better in presentations. The enclosed package includes both CMU Serif and Sans Serif TrueType font files downloaded from https://fontlibrary.org and redistributed as allowed under the SIL Open Font License. To produce slide-style graphics, run `graph export` with the options `fontdir([path])` and `fontface("CMU Sans Serif")`.

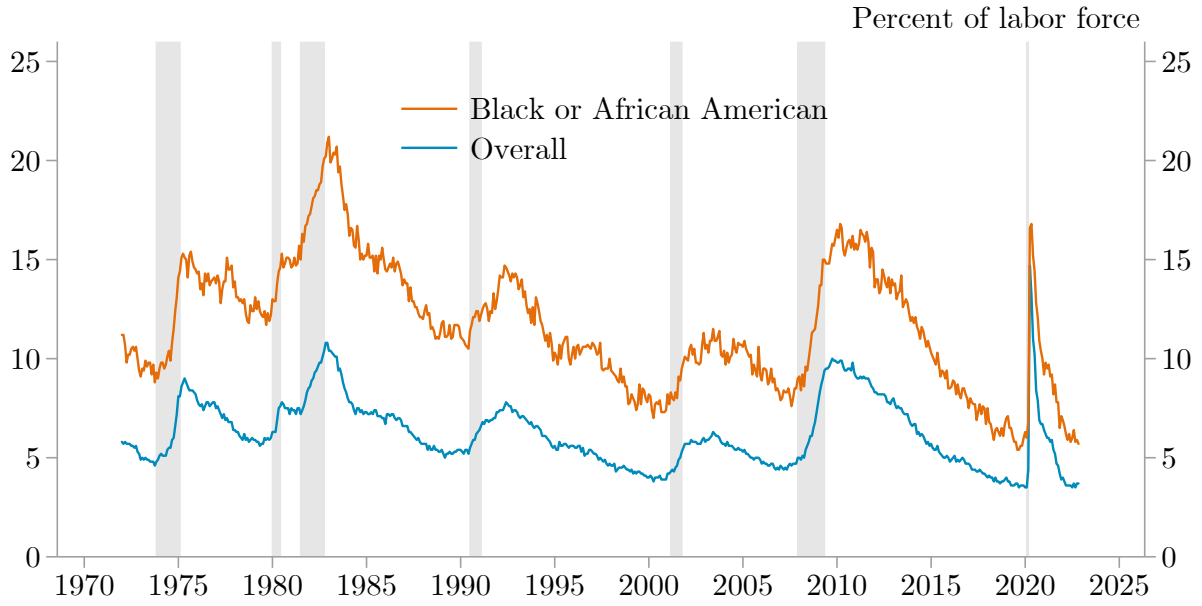
Stata has trouble exporting certain fonts to certain graphical formats; for example, to get CMU Sans Serif working properly, I find I have to export to `.eps` first, then use the MacOS shell utility `epstopdf` to convert to `.pdf`. In addition, some of Stata's SMCL functionality (such as boldface text) malfunctions when using custom fonts in certain formats. Getting the fonts to work on your machine may take some debugging.

LaTeX integration. I write figure titles and notes in LaTeX, both to facilitate cross-referencing and because Stata lacks automatic text-wrapping and is thus ill-equipped to write detailed notes. I create multi-panel figures directly in Stata and write each panel header in Stata. I've tried to get everything sized consistently between Stata and LaTeX text, using a combination of Stata sizing options and LaTeX figure scales, but that's hard to pull off exactly because changing any figure element in Stata—such as legends or axis labels—impacts sizing across the board. Here and in general, I settle for rough parity.

Disclaimer. Per the terms of `readme.txt`, I provide these files without warranty of any kind and without assumption of liability.

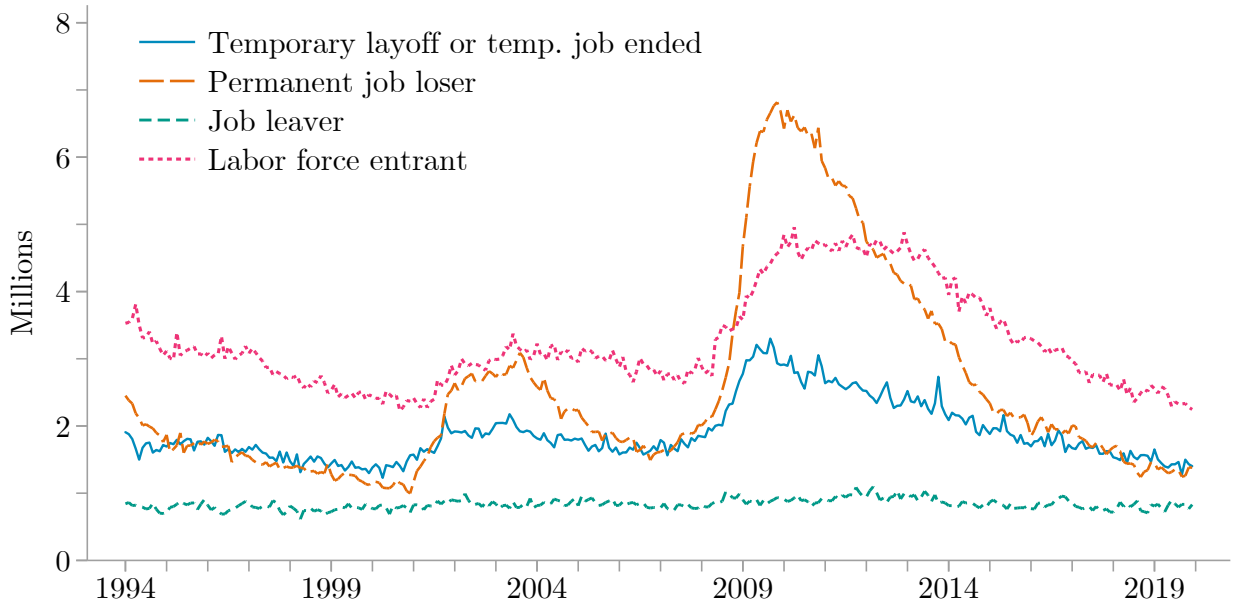
[2] I've used native Stata capabilities wherever possible, but in two places I rely on MacOS shell utilities. First, one figure was prohibitively large in `.pdf` format, so I use `gs` to convert it into `.jpg`. (Newer versions of Stata can export directly to `.jpg`.) Second, I use `pdfcrop` to crop white space, so as to squeeze larger figures into this LaTeX document. Users without these shell utilities available can simply skip those steps.

Figure 1: Black unemployment over the business cycle



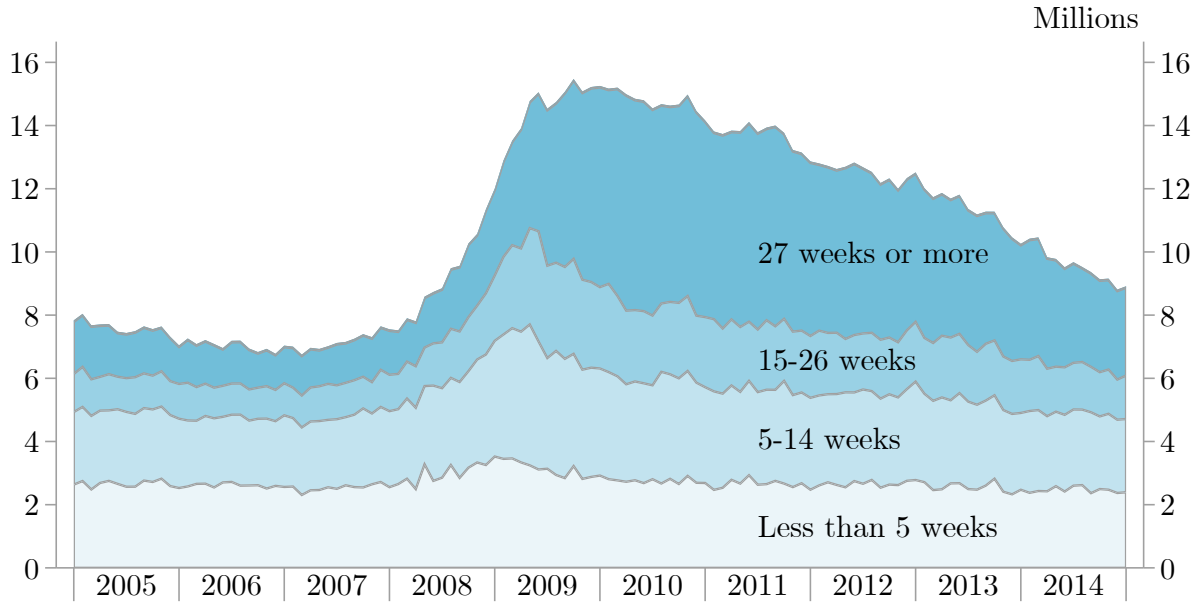
Source: Bureau of Labor Statistics. Produced by `unemp_black.do`.
Note: Shaded regions are recessions as dated by the National Bureau of Economic Research.

Figure 2: Decomposing unemployment by reason



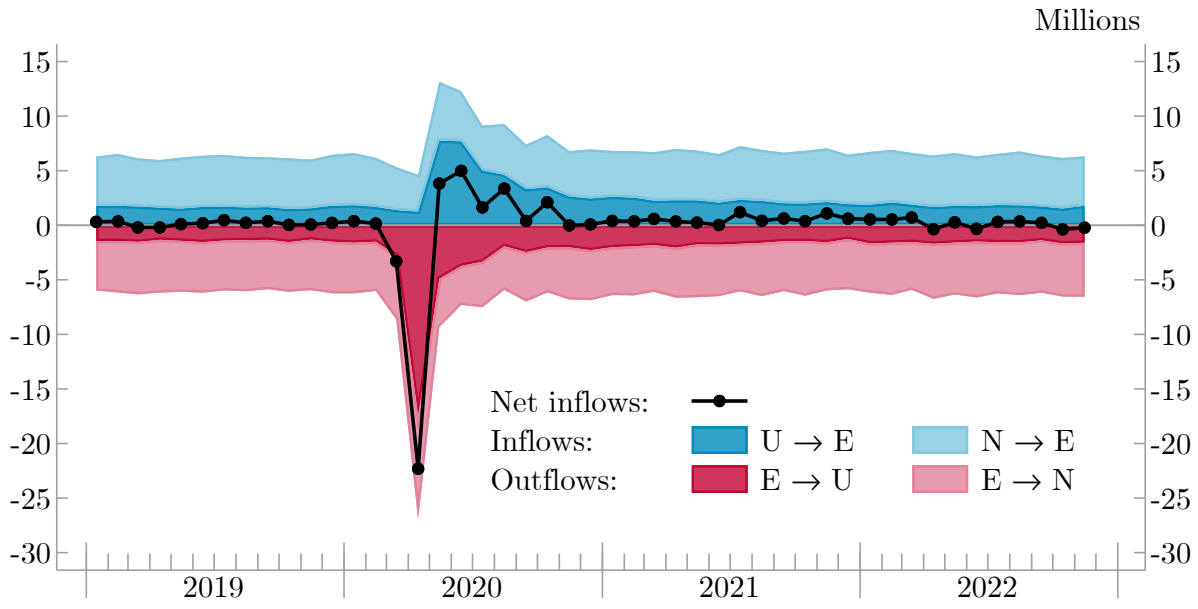
Source: Bureau of Labor Statistics. Produced by `unemp_reasons.do`.
Note: A worker is on temporary layoff if they expect to be recalled to their former employer or have been given a definite recall date. Labor force entrants include both new entrants and reentrants.

Figure 3: Long-term unemployment surged during the Great Recession



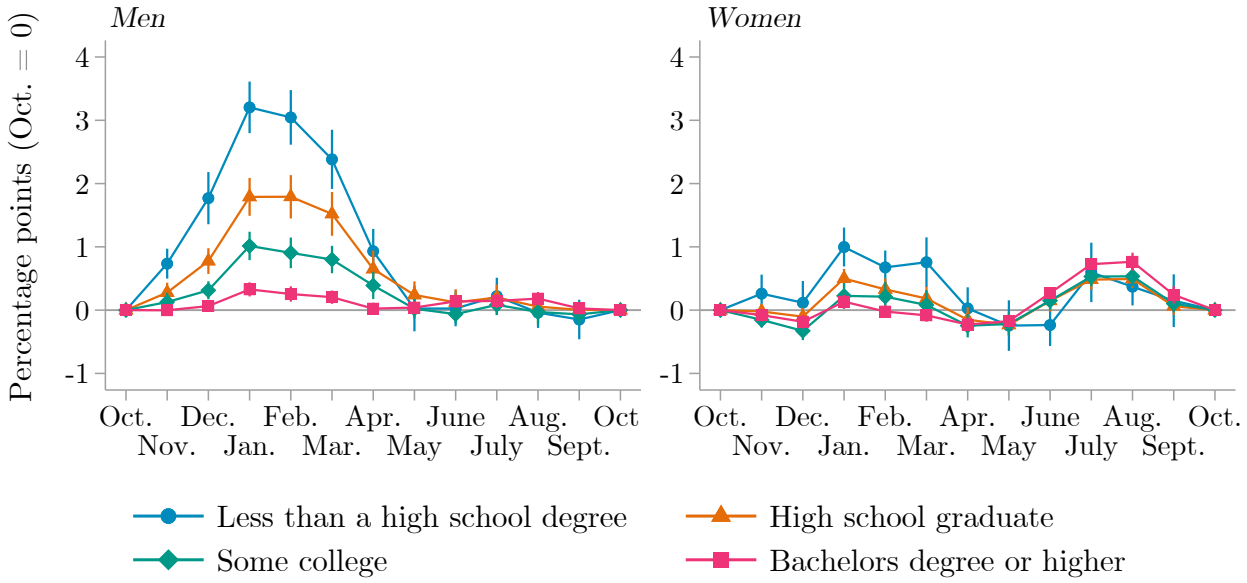
Source: Bureau of Labor Statistics. Produced by `unemp_duration.do`.

Figure 4: Employment flows during the COVID-19 pandemic



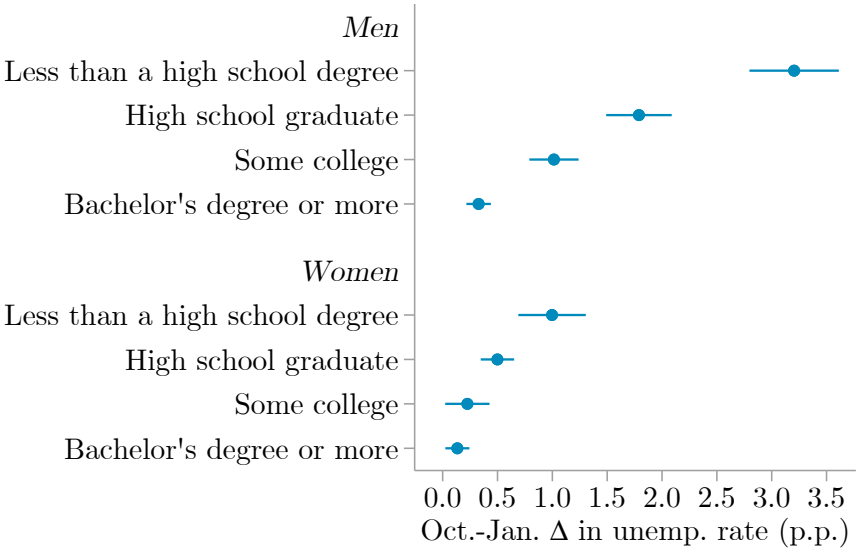
Source: Bureau of Labor Statistics. Produced by `lm_flows.do`.

Figure 5: Seasonal changes in unemployment by sex and educational attainment



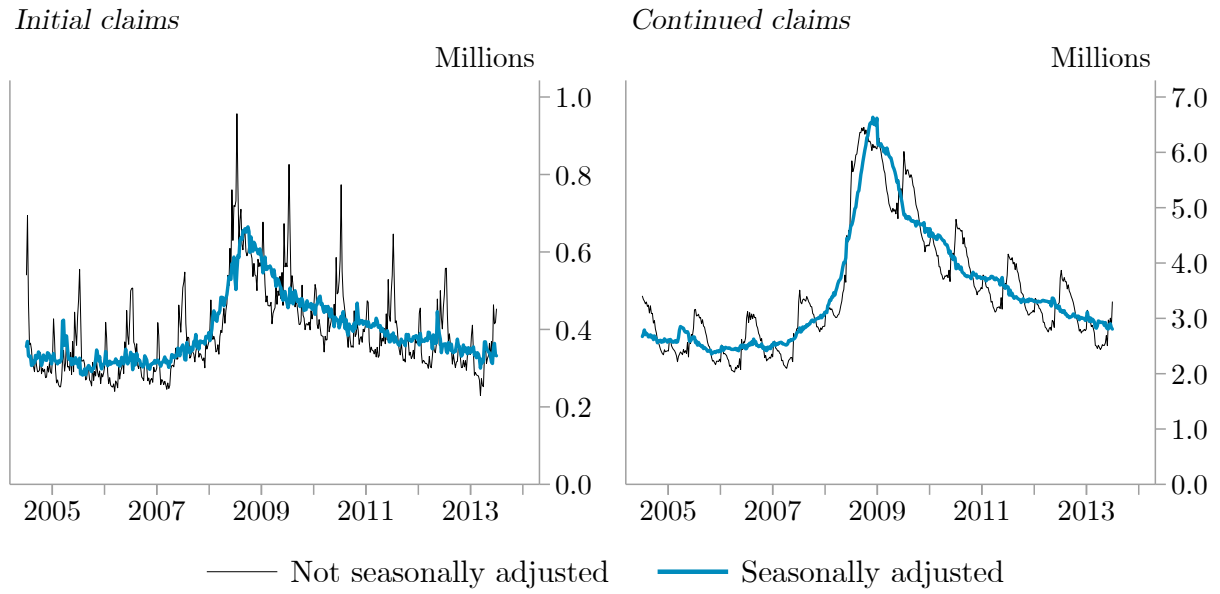
Source: Bureau of Labor Statistics. Produced by `unemp_seasonality.do`.
Note: Coefficients from regressions of each group's unemployment rate on month dummies and a linear spline in calendar time. Bars show 95 percent confidence intervals based on Newey-West standard errors.

Figure 6: The winter surge in unemployment



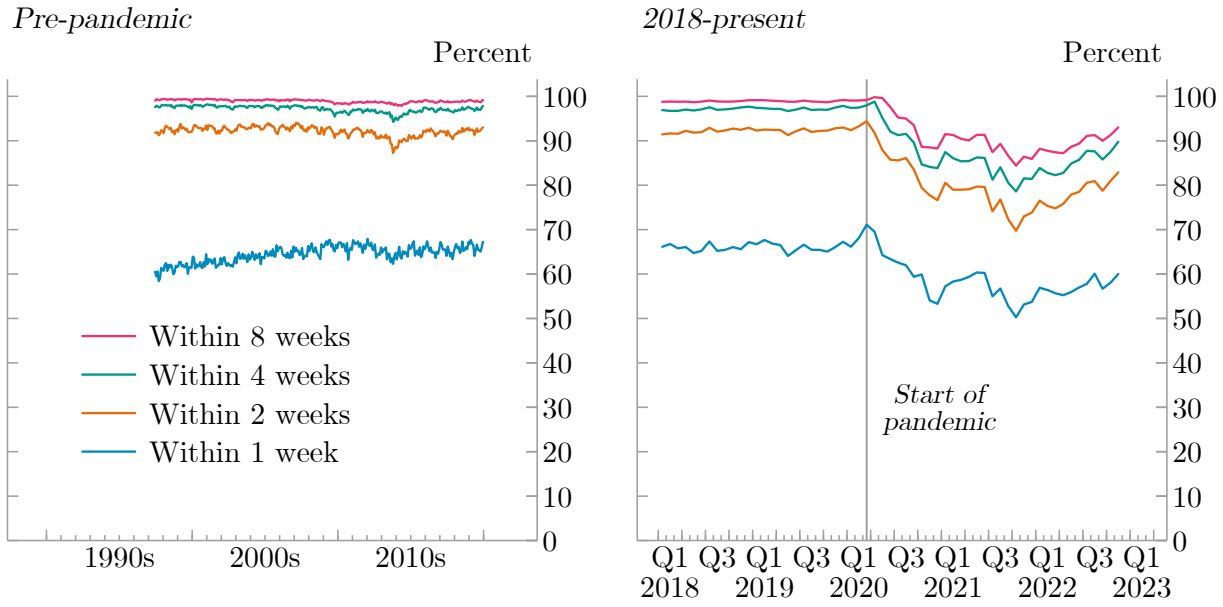
Source: Bureau of Labor Statistics. Produced by `unemp_seasonality.do`.
Note: Estimated changes in the unemployment rate between October and January from a regression of each group's unemployment rate on a set of month dummies and a linear spline in calendar time. Bars show 95 percent confidence intervals based on Newey-West standard errors.

Figure 7: Weekly claims for unemployment insurance benefits



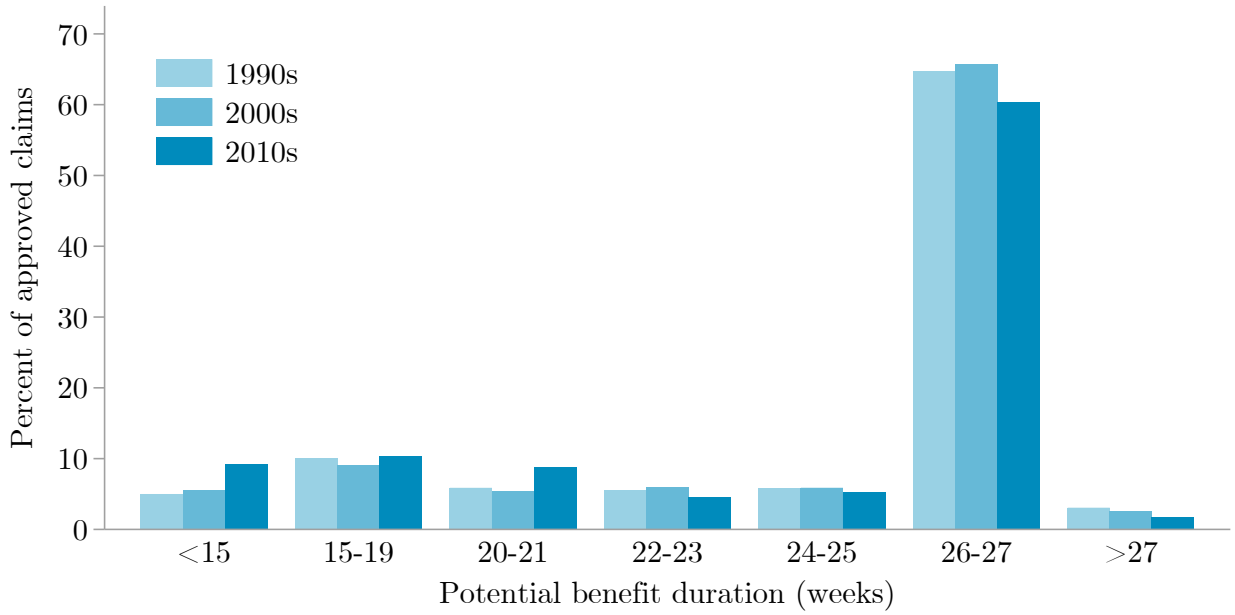
Source: Employment and Training Administration. Produced by `ui_claims.do`.
Note: Series show claims for regular state UI programs and exclude claims under the Emergency Unemployment Compensation and Extended Benefits programs active during this period.

Figure 8: Timeliness of UI benefit payments during the COVID-19 pandemic



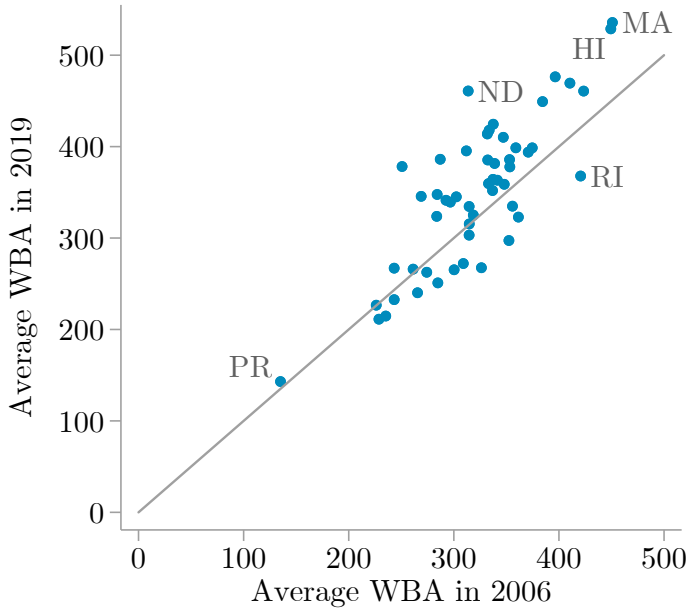
Source: Employment and Training Administration. Produced by `ui_timeliness.do`.
Note: Plotted series reflect time elapsed between the week of compensated unemployment and the date of payment among claimants in regular state UI programs. Data are non–seasonally adjusted.

Figure 9: Potential benefit duration among new UI claimants



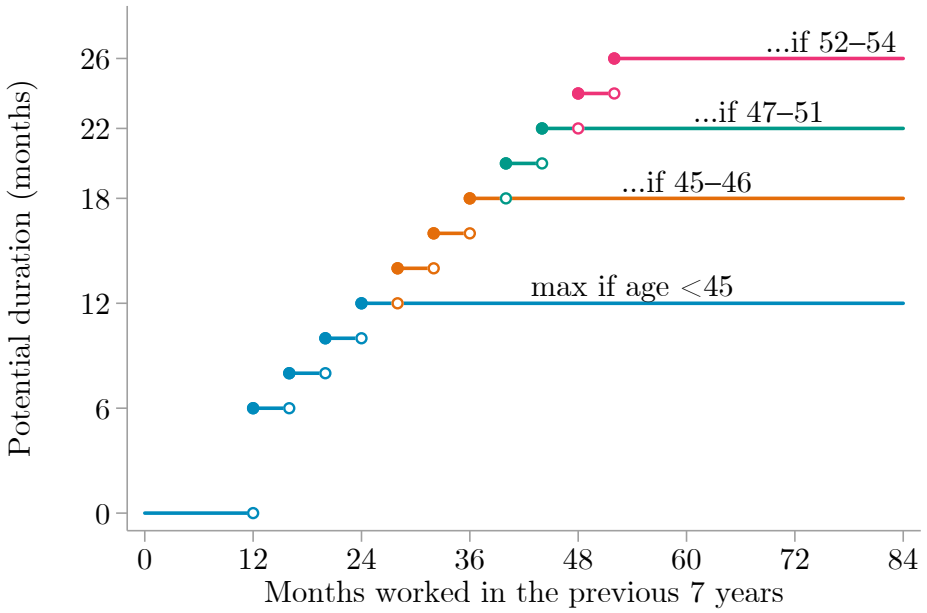
Source: Employment and Training Administration. Produced by `ui_pbd.do`.
Note: Potential duration of claims filed in regular state UI programs. Data are non–seasonally adjusted.

Figure 10: Cross-state variation in average weekly UI benefit amounts



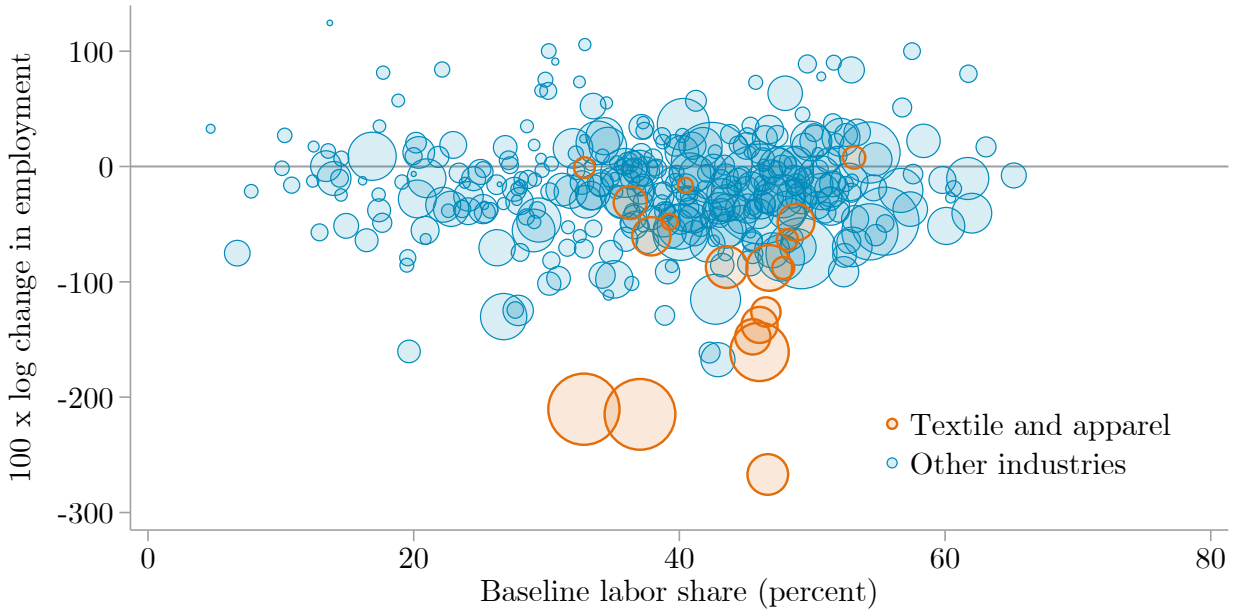
Source: Employment and Training Administration. Produced by `ui_wba.do`.
Note: Data are based on fully unemployed claimants in regular state UI programs. Average weekly benefit amounts reflect both state-level policy and claimants' prior earnings. Data are non–seasonally adjusted and deflated to 2019 USD.

Figure 11: Benefit accrual in the German UI system



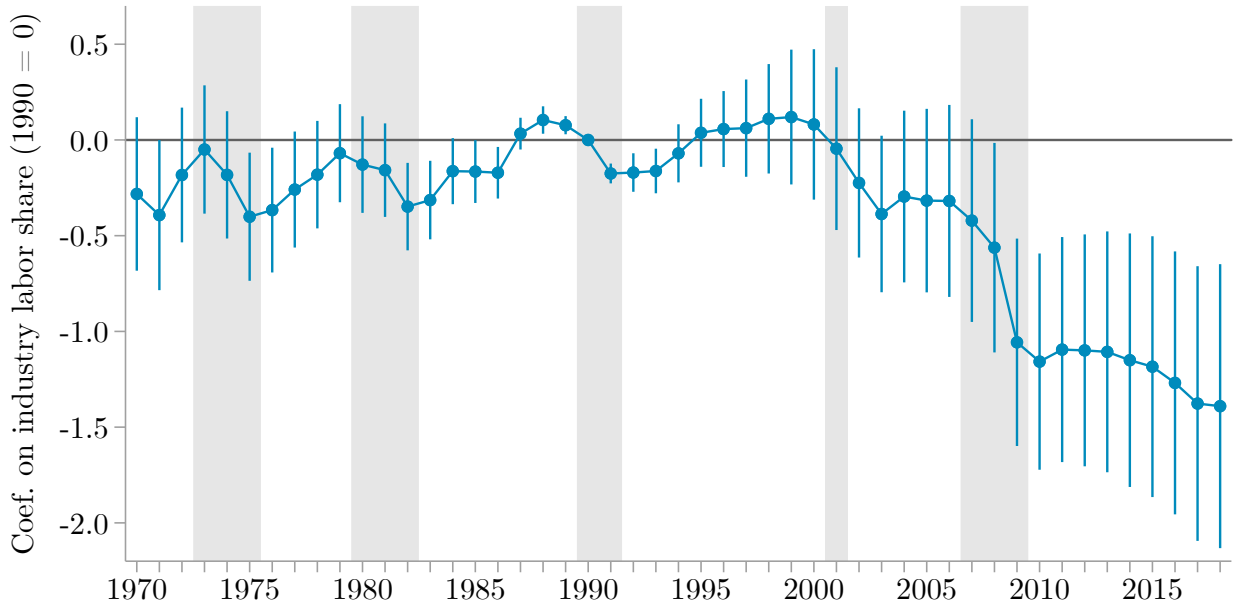
Source: Price (2019), "The Duration and Wage Effects of Long-Term Unemployment Benefits: Evidence from Germany's Hartz IV Reform". Produced by `german_pbd.do`.
Note: Potential duration of "short-term" unemployment insurance benefits (*Arbeitslosengeld*). Claimants who exhaust these benefits may apply for "long-term" unemployment assistance (*Arbeitslosenhilfe*).

Figure 12: Changes in manufacturing employment, 1990–2007



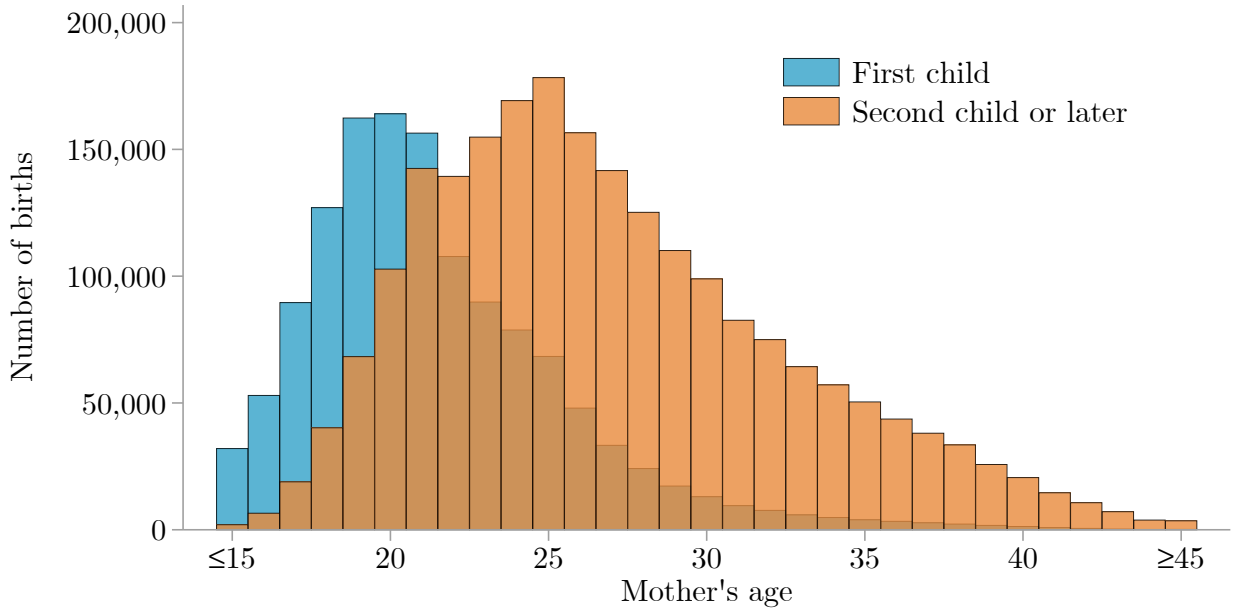
Source: NBER-CES Manufacturing Industry Database. Produced by `manuf_scatter.do`.
Note: Labor share is wage bill divided by value added. Circle sizes reflect 1990 employment levels.

Figure 13: Labor-intensive manufacturing industries are especially procyclical



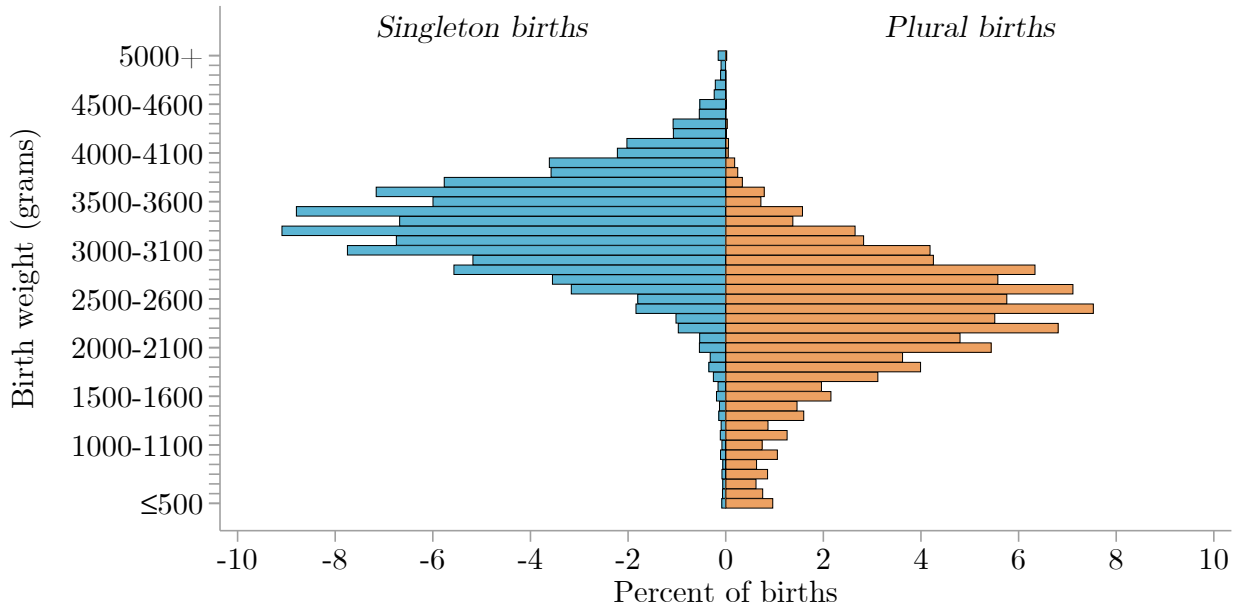
Source: NBER-CES Manufacturing Industry Database. Produced by `manuf_reg.do`.
Note: Coefficients from a regression of log employment on year dummies and year × labor share interactions in a set of 6-digit manufacturing industries, with labor share computed as the average over 1970–2018. Bars show 95 percent confidence intervals clustered on industry. Shaded regions are recessions as dated by the National Bureau of Economic Research.

Figure 14: Maternal age by birth order (1968 births)



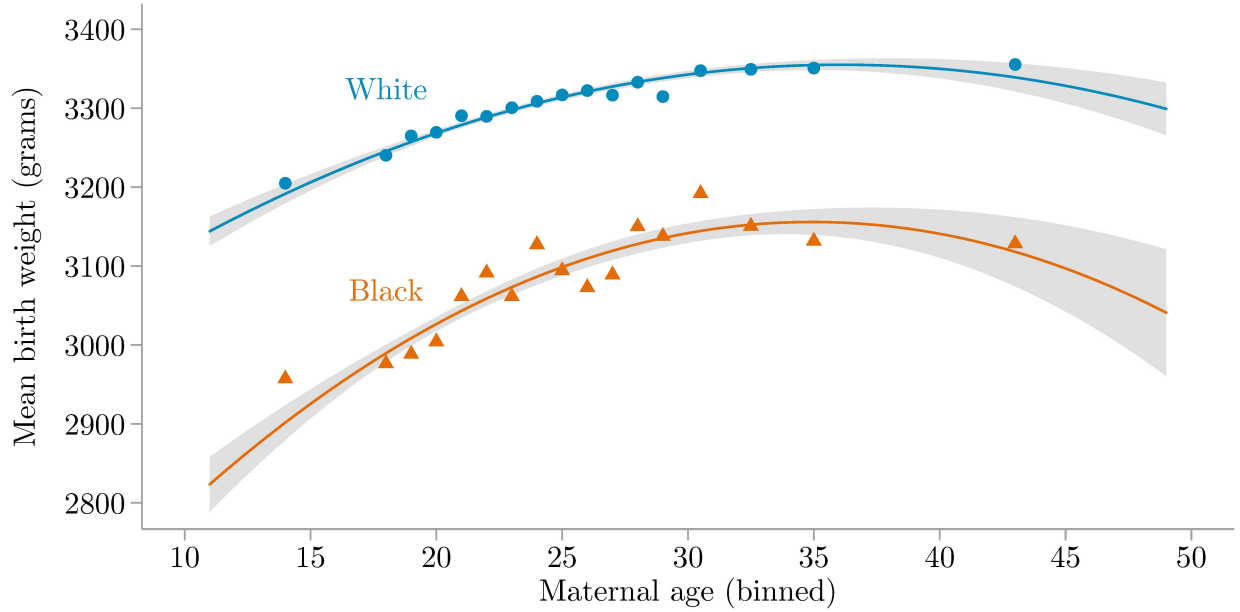
Source: National Center for Health Statistics (1968 natality file). Produced by `birth_age.do`.

Figure 15: Birth weight by plurality (1968 births)



Source: National Center for Health Statistics (1968 natality file). Produced by `birth_weight.do`.

Figure 16: Birth weight by maternal race and age (1968 births)



Source: National Center for Health Statistics (1968 natality file). Produced by `birth_qfit.do`.
Note: Mothers are grouped into 20 age bins based on the age distribution among all white and Black mothers. Plotted curves are best fits from regressions of birth weight on a quadratic in maternal age using the unbinned data; grey bands show 95 percent confidence intervals.

Data citations. (See enclosed code for specific FRED series used.)

1. Becker, Randy A., Wayne B. Gray, and Jordan Marvakov. 2021. "NBER-CES Manufacturing Industry Database (1958-2018, version 2021a)." National Bureau of Economic Research. https://www.nber.org/research/data/nber-ces-manufacturing-industry-database, accessed December 30, 2022.

2. National Center for Health Statistics, National Vital Statistics System, via the National Bureau of Economic Research. 1968 natality file. https://www.nber.org/research/data/vital-statistics-natality-birth-data, accessed December 30, 2022.

3. US Department of Labor, Bureau of Labor Statistics. Numerous series retrieved from FRED, Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org, accessed December 30, 2022.

4. US Department of Labor, Employment and Training Administration. Numerous series retrieved from FRED, Federal Reserve Bank of St. Louis, https://fred.stlouisfed.org, and from https://oui.doleta.gov/unemploy/DataDownloads.asp, accessed December 30, 2022.