Hit Harder, Recover Slower? Unequal Employment Effects of the COVID-19 Shock
Sang Yoon (Tim) Lee, Minsung Park, and Yongseok Shin

Work, Leisure, and Family: From the Silent Generation to Millennials
George-Levi Gayle, Mariana Odio-Zuniga, and Prasanthi Ramakrishnan

Monetary Policy and Economic Performance Since the Financial Crisis
Dario Caldara, Etienne Gagnon, Enrique Martinez-Garcia, and Christopher J. Neely

Assessing Labor Market Conditions Using High-Frequency Data
Maximiliano Dvorkin and Maggie Isaacson

Stability and Equilibrium Selection in Learning Models: A Note of Caution
Yili Chien, In-Koo Cho, and B. Ravikumar
Review

Review is published four times per year by the Research Division of the Federal Reserve Bank of St. Louis. Full online access is available to all, free of charge.

Online Access to Current and Past Issues
The current issue and past issues dating back to 1967 may be accessed through our Research Division website: http://research.stlouisfed.org/publications/review. All nonproprietary and nonconfidential data and programs for the articles written by Federal Reserve Bank of St. Louis staff and published in Review also are available to our readers on this website.

Review articles published before 1967 may be accessed through our digital archive, FRASER: http://fraser.stlouisfed.org/publication/?pid=820. Review is indexed in Fed in Print, the catalog of Federal Reserve publications (http://www.fedinprint.org/), and in IDEAS/RePEc, the free online bibliography hosted by the Research Division (http://ideas.repec.org/).

Subscriptions and Alerts
The Review is no longer printed or mailed to subscribers. Our last printed issue was the first quarter of 2020. Our monthly email newsletter keeps you informed when new issues of Review, Economic Synopses, Regional Economist, and other publications become available; it also alerts you to new or enhanced data and information services provided by the St. Louis Fed. Subscribe to the newsletter here: http://research.stlouisfed.org/newsletter-subscribe.html.

Authorship and Disclaimer
The majority of research published in Review is authored by economists on staff at the Federal Reserve Bank of St. Louis. Visiting scholars and others affiliated with the St. Louis Fed or the Federal Reserve System occasionally provide content as well. Review does not accept unsolicited manuscripts for publication.

The views expressed in Review are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

Copyright and Permissions
Articles may be reprinted, reproduced, republished, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. In these cases, there is no need to request written permission or approval. Please send a copy of any reprinted or republished materials to Review, Research Division of the Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166-0442; STLS.Research.Publications@stls.frb.org.

Please note that any abstracts, synopses, translations, or other derivative work based on content published in Review may be made only with prior written permission of the Federal Reserve Bank of St. Louis. Please contact the Review editor at the above address to request this permission.

Economic Data
General economic data can be obtained through FRED®, our free database with over 800,000 national, international, and regional data series, including data for our own Eighth Federal Reserve District. You may access FRED through our website: https://fred.stlouisfed.org.

© 2021, Federal Reserve Bank of St. Louis.
ISSN 0014-9187
The destructive economic impact of the COVID-19 pandemic was distributed unequally across the population. A worker’s gender, race and ethnicity, age, education, industry, and occupation all mattered. We analyze the initial negative effect and its lingering effect through the recovery phase, across demographic and socioeconomic groups. The initial negative impact on employment was larger for women, minorities, the less educated, and the young whether or not we account for the industries and occupations they worked in. By February 2021, however, the differential effects across groups had gotten much smaller overall and had entirely vanished once the different industries and occupations they work in are taken into account. In particular, the differential effects between men and women vanished with or without the industry and occupation compositions taken into account, indicating that women’s progress in the labor market over the past decades has not been wiped out by the pandemic. Across race and ethnicity, Hispanics and Asians were the worse hit but made up most of the lost ground, while the initial impact on Blacks was smaller but their recovery was slower. (JEL J21, J15, J16)
Jobs differ by contact intensity and the ease with which they can be performed remotely, in addition to their essential or nonessential classification (Hensvik et al., 2020; Aum et al., 2020b, 2021). Warnings abounded that the economic toll of the pandemic would be unevenly distributed and exacerbate preexisting inequality across demographic and socioeconomic groups, because women and minorities are more likely to work in more-vulnerable jobs (Alon et al., 2020; Blundell et al., 2020). At the onset of the pandemic, near-real-time data revealed that women lost more jobs and were forced to work less, both in the United States and the United Kingdom (Cajner et al., 2020; Adams-Prassl et al., 2020a,b). It also became apparent that minorities were disadvantaged not only because of the types of jobs they held, but also because they were more likely to face employment reductions even within the same jobs as Whites (Montenovo et al., 2020; Cowan, 2020; Gezici and Ozay, 2020).

In this article, we analyze how the initial economic impact of the pandemic and the subsequent recovery differed along numerous dimensions, including gender, race and ethnicity, age, educational attainment, industry, and occupation, and state-level policies and statewide COVID-19 infection rates. The main contribution to the literature is our analysis of the recovery phase through February 2021, as many researchers have documented the early impact of the pandemic in spring 2020.1

Our main findings can be summarized as follows:

- Minorities were hit harder by the pandemic, largely due to an industry-occupation composition effect—for example, their disproportionate presence in service industries: the leisure and hospitality industry and the other services industry.
- Many demographic and socioeconomic groups that were hit harder initially have also recovered faster, especially once industry and occupation compositions are taken into account.
- More specifically, the pandemic’s differential effects across gender, age, and education vanish by February 2021 once industry and occupation effects are controlled for.
- The differential effects between men and women disappear even when their industry and occupation compositions are not considered, indicating that women’s progress in the labor market over the past few decades has not been wiped out by the pandemic.
- Black workers were the least affected by the initial shock among all racial groups, but their recovery has been the slowest, even when industry and occupation effects are controlled for.
- In April 2020, local employment was hit hard in states that had high infection rates, with containment policies having no significant effect on employment. In February 2021, the severity of the epidemic had no systematic effect on employment.
- Urban areas, especially city centers, were hit the hardest and the effects still remain.

We now describe the data and our methodology (Sections 1 and 2) before discussing the results in more detail (Section 3).
1 DATA

We use the monthly Current Population Survey (CPS) from the Bureau of Labor Statistics (BLS). We limit the sample to those 20 to 65 years of age and consider four variables of interest: (i) unemployment, (ii) jobless unemployment, (iii) furlough, or recall, unemployment, and (iv) nonparticipation (not in the labor force). Unemployment and nonparticipation are directly recorded by the BLS. Jobless unemployment and recall unemployment are subcategories of unemployment. The identification of jobless unemployed and recall unemployed relies on the definition in Hall and Kudlyak (2020). The CPS asks respondents if they are currently laid off. If yes, they are asked whether they were given a return date to work or any indication that they would be called back to work within the next six months. If the answer is again yes, they are asked whether they can return if/when recalled. If the answer is also yes, then the respondent is classified as recall unemployed, that is, one who has a job to return to. On the other hand, if a respondent did not work during the survey week, does not expect to be called back, and has been actively looking for work, then they are classified as jobless unemployed.2

For demographic and socioeconomic characteristics, we consider gender (male and female), race and ethnicity (White, Black, Hispanic, and Asian), age (20 to 35, 36 to 50, and 51 to 65 years of age), educational attainment (a high school diploma or less education, some college but without a four-year degree, and a bachelor’s degree or higher), industry, occupation, and urban/rural residence. We classify industries and occupations into 14 and 11 categories, respectively, based on the major industry recodes and major occupation group recodes provided by the BLS. The CPS has information about whether respondents live in a central city, outside a central city but still in a metropolitan area, or outside a metropolitan area.

We also consider infection rates by state and state government policy responses to the pandemic. Daily case counts from the Centers for Disease Control and Prevention (CDC) COVID Data Tracker are used to calculate the number of cases per 1,000 people.3 We group states into low risk, medium risk, and high risk, with equal numbers of states in each category. In addition, we group states by their policy responses to COVID-19 following the Oxford COVID-19 Government Response Tracker (OxCGRT).4 OxCGRT reports 14 time-varying indicators to measure the policy responses of several governments, including those of the 50 U.S. states and Washington, D.C. Each indicator is classified as “containment and closure,” “economic response,” “health systems,” or “miscellaneous” and is used for creating a score for the overall government response (Hale et al., 2020).5 Based on these scores, states are grouped into three policy-response categories: (i) robust-response states, which adopted and maintained robust containment, testing, and contact tracing policies; (ii) rapid-rollback states, which adopted a robust response initially but then rolled back policies relatively quickly; and (iii) low-response states, which never adopted particularly restrictive containment measures or robust testing and contact tracing systems.
2 ESTIMATION

The panel dimension of the CPS is short, so it is not possible to track individuals over the course of a year. We instead estimate the following individual-level linear regression model to capture the factors correlated with the labor market impact of the pandemic:

\[
Y_{it} = \alpha + \alpha_1 \chi_{t'=t} + X_{si} [\beta + \beta_1 \chi_{t'=t}] + \epsilon_{si}.
\]

We run the regression separately for \( s = 4 \) (April 2020) and \( s = 14 \) (February 2021), where \( t = 2019 \) or 2020 for \( s = 4 \), and \( t = 2020 \) or 2021 for \( s = 14 \). For each \( s \), \( t' \) indicates the latter year.

April 2020 was when the pandemic’s economic impact was at its peak, and February 2021 was the most recent sample available from the CPS to gauge the recovery process; it is also the last month we can compare annual differences between pre- and post-COVID-19 months. Comparing the same months for 2019 and 2020 or for 2020 and 2021 is informative about the economic effect of the pandemic, seasonally adjusted. The dependent variable \( Y_{it} \) is a binary variable of individual \( i \)’s employment status in month \( s \) in year \( t \), and we run separate regressions for jobless unemployment, recall unemployment, unemployment, and nonemployment (unemployment plus nonparticipation).

The vector of regressors \( X_{si} \) includes group dummies on gender, race and ethnicity, education, age, industry, occupation, and geographic location. The location variables include (i) urban/rural residence, (ii) statewide new COVID-19 cases per 1,000 people during the preceding month (to be precise, cumulative counts through April 15, 2020, for the April 2020 regression, and January 15 to February 15, 2021, for the February 2021 regression, since CPS interviews are conducted during the week that contains the 19th of each month), and (iii) the state government’s policy response. For April 2020 the state policy responses include only the robust and low-response categories (because no state had a rapid rollback), while for February 2021 they also include the rapid-rollback category.

For each pair of years for the same month, the indicator function \( \chi_{t'=t} \) equals the latter year (February 2021 or April 2020) and zero otherwise. In this specification, \( \beta_1 \) is the parameter of interest, which captures the differential effects of the pandemic on each demographic and socioeconomic group.

3 RESULTS

3.1 Unemployment by Gender, Race and Ethnicity, Age, and Education

Before we report the estimation results, we first show the evolution of labor market outcomes as a whole and then by gender, race and ethnicity, age, and educational attainment.

Figure 1 plots the nonparticipation rate, unemployment rate, jobless unemployment rate, and recall unemployment rate from January 2019 onward. The pandemic hit the economy hard in April 2020, when unemployment peaked at 14.8 percent. The economy has since been recovering toward the pre-pandemic levels. Note that the unemployment jump is almost entirely accounted for by recall unemployment, which came down fast in the following months.
(but was still 1 percentage point higher in February 2021 than in the same month of the previous year). This is broadly consistent with the findings of Hall and Kudlyak (2020). On the other hand, the jobless unemployment rate began to rise in July 2020 and more than two-fifths of the year-over-year rise in the unemployment rate as of February 2021 is explained by higher jobless unemployment (1.2 percentage points out of 2.9 percentage points). The pace of recovery has slowed markedly since October 2020. At the time of writing, the unemployment rate in June 2021 was 5.9 percent, down only 1 percentage point from 6.9 percent in October 2020.

Figure 1 also shows that some workers dropped out of the labor force (instead of entering unemployment) when the pandemic hit. The nonparticipation rate increased by 3.1 percentage points between March and April 2020. This is the largest monthly increase ever recorded. For comparison, after the onset of the Great Recession, it took nearly six years for the nonparticipation rate to rise by 3.1 percentage points (from December 2007 to October 2013). The recovery in the nonparticipation rate stalled starting in June 2020 and was still 1.7 percentage points higher in February 2021 than in February 2020.

Figure 2 shows the impact of the pandemic on the employment status of men and women, not controlling for any other variable. In Panel A, for women, the first four bars show the changes in the jobless unemployment, recall unemployment, unemployment, and nonparticipation rates between April 2019 and April 2020, capturing the peak impact of the pandemic. The next four bars show the four rates between February 2020 and February 2021. Panel B is for men. Comparing the two panels, we see that women were hit harder by the pandemic than men: Year-over-year April unemployment rose by 12.7 percentage points for women
vs. 9.9 percentage points for men, all driven by the rise in recall unemployment. This was a unique phenomenon: Typically men are more adversely affected by recessions than women (Alon et al., 2020). Year-over-year nonparticipation rose slightly more for men than for women (3.0 percentage points vs. 2.4 percentage points). But in February 2021, this gender gap completely disappeared—and even reversed: For men, the unemployment rate rose by 2.9 percentage points relative to February 2020 but for women by 2.7 percentage points. (The year-over-year change in the nonparticipation rate was the same for women and men in February 2021: 1.9 percentage points.) In summary, the pandemic hit women harder initially, but by February 2021 the pandemic’s effect on employment was the same for men and women. We again see that the initial impact and the ensuing recovery were driven by recall unemployment.

Figure 3 shows the employment impact across race and ethnicity. Comparing the year-over-year change in the unemployment rate, it is clear that in April 2020 Hispanics were hit harder than any other group (15.1 percentage points), followed by Asians (12.0 percentage points). The unemployment rose the least for Blacks among all groups, including Whites (10.0 percentage points vs. 10.2 percentage points), but their nonparticipation rate rose by 5 percentage points, double the increases for Whites and Hispanics. For Whites, the unemployment rate in February 2021 was only 2.1 percentage points higher than in February 2020, a smaller negative effect compared with 4.0 percentage points for Blacks, 4.4 percentage points for Hispanics, and 2.7 percentage points for Asians. It is clear that minorities were hit harder economically by the pandemic and also have been recovering more slowly. The remaining effect on the nonparticipation rate was also larger for minorities, although the magnitude was smaller. The year-over-year changes in the February 2020 nonparticipation rates were 2.2 percentage points, 2.0 percentage points, and 2.0 percentage points for Hispanics, Asians, and Blacks, respectively, compared with 1.8 percentage points for Whites.
Figure 4 shows how the employment outcomes of different age groups were affected by the COVID-19 shock. Clearly, the young (20 to 35 year olds) were hit the hardest in April 2020: The year-over-year increases in their unemployment rate and nonparticipation rate were 12.9 percentage points and 4.6 percentage points, respectively. However, in February 2021, the unemployment effects of the pandemic were fairly similar across all three age groups, except the youngest group’s nonparticipation rate had not recovered as much, 2.7 percentage points compared with 1.2 percentage points and 0.9 percentage points for the two older groups.

Figure 5 shows the negative employment effects by educational attainment: those with a high school diploma or less education (high school or less), some college but without a four-year degree (some college), and a bachelor’s degree or higher. Consistent with the general findings in the labor literature (e.g., Lee et al., 2015), the patterns for those with high school
or less and those with some college are broadly similar. The unemployment rate for those with high school or less was higher by 15.0 percentage points in April 2020 than in April 2019 and for those with some college by 13.6 percentage points, while for those with a bachelor’s degree or higher it was only 6.7 percentage points higher. (The magnitudes are smaller but the patterns are similar for the nonparticipation rates.) By February 2021, all groups had experienced significant recoveries, again due to the drop in recall unemployment. The unemployment rates were 3.8 percentage points and 3.0 percentage points higher in February 2021 than in February 2020 for those with high school or less and those with some college, respectively, and 2.0 percentage points higher for those with a bachelor’s degree or higher. The picture is clear that those with more education were economically less affected by the pandemic.
3.2 Estimation Results

We now turn to the estimates from equation (1). Although the figures in the previous section offer a snapshot of the unequal employment effects of the pandemic across demographic and socioeconomic groups, the effects shown there were confounded by the overlapping compositions across those groups as well as their distributions across industries, occupations, and geographic areas—with all hit differently by the pandemic. Regression (1) can isolate the effect specific to each group, which is captured by the coefficient $\beta_1$.

The estimated $\beta_1$ for each group (other than the reference group, by construction) is reported in Table 1 for the year-over-year change in April 2020. A significant positive estimate means that the employment outcome of the given group was worse than the reference group’s.

Columns 1 to 3 are the estimates for when the outcome variable $Y_{st}$ is jobless unemployment, recall unemployment, and unemployment, respectively, and include industry and occupation fixed effects. Since the majority of CPS individuals who do not participate in the labor market do not record their previous industry or occupation, we cannot include such fixed effects for nonemployment (again, unemployment plus nonparticipation) in Column 5. So for ease of interpretation, in Column 4 we also estimate (1) for unemployment without industry and occupation fixed effects. The coefficients on the industry and occupation fixed effects for Columns 1 to 3 are relegated to the tables in the appendix.

This is a saturated regression, and the excluded groups are males, Whites, people with a high school diploma or less education, people between 20 and 35 years of age, and people living in a city center of a state with a robust COVID-19 response and low COVID-19 risk. For regressions with industry and occupation fixed effects, the added excluded groups are the public administration industry and the management, business, and financial occupation category.

The year-over-year increase in the aggregate unemployment rate for April 2020 was 11.1 percentage points. The magnitude of the estimated coefficients in the table can be interpreted relative to this number.

Consistent with the results in Section 3.1, we find that the negative employment effects at the peak of the pandemic were larger for women (than men), for Hispanics and Asians (than Whites or Blacks), for the less educated, and for young workers, controlling for all other factors. These differential effects are smaller with industry and occupation fixed effects (Column 3) than without (Column 4), but they do exist even within occupations and industries. There are two remarkable findings: First, the unemployment effects for Blacks were significantly smaller than those for Whites, with or without industry and occupation fixed effects. Second, despite the larger point estimate for Hispanics, their unemployment rate did not rise significantly more than that for Whites (at the 10 percent significance level) once industry, occupation, and other effects are controlled for, implying that Hispanics were economically exposed to the pandemic by virtue of the types of jobs they held.8

Table 1 also shows how state-level policy responses and the extent of the pandemic in the preceding month are correlated with employment outcomes. Somewhat surprisingly, state-level containment policies had no significant effect on employment. On the other hand, the number of newly confirmed COVID-19 cases leads to (or “Granger causes”) more unemploy-
### Table 1
COVID-19 Shock: April 2020

<table>
<thead>
<tr>
<th></th>
<th>Gender [Males]</th>
<th>Furlough</th>
<th>Unemployment</th>
<th>Unemployment without FE</th>
<th>Nonemployment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender [Males]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females × 20/4</td>
<td>−0.0011 (0.0028)</td>
<td>0.0332*** (0.0042)</td>
<td>0.0333*** (0.0050)</td>
<td>0.0353*** (0.0047)</td>
<td>0.0052 (0.0063)</td>
</tr>
<tr>
<td><strong>Race [Whites]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blacks × 20/4</td>
<td>−0.0113** (0.0052)</td>
<td>−0.0127* (0.0067)</td>
<td>−0.0230*** (0.0085)</td>
<td>−0.0188** (0.0087)</td>
<td>0.0017 (0.0109)</td>
</tr>
<tr>
<td>Hispanics × 20/4</td>
<td>0.0054 (0.0036)</td>
<td>0.0040 (0.0059)</td>
<td>0.0127* (0.0070)</td>
<td>0.0251*** (0.0072)</td>
<td>0.0228** (0.0093)</td>
</tr>
<tr>
<td>Asians × 20/4</td>
<td>0.0030 (0.0045)</td>
<td>0.0107 (0.0070)</td>
<td>0.0165** (0.0084)</td>
<td>0.0252*** (0.0087)</td>
<td>0.0033 (0.0126)</td>
</tr>
<tr>
<td><strong>Education [High school or less]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some college × 20/4</td>
<td>0.0017 (0.0036)</td>
<td>0.0062 (0.0059)</td>
<td>0.0078 (0.0069)</td>
<td>−0.0133* (0.0070)</td>
<td>−0.0060 (0.0086)</td>
</tr>
<tr>
<td>College × 20/4</td>
<td>0.0034 (0.0037)</td>
<td>−0.0301*** (0.0058)</td>
<td>−0.0252*** (0.0070)</td>
<td>−0.0856*** (0.0061)</td>
<td>−0.0657*** (0.0078)</td>
</tr>
<tr>
<td><strong>Age [Ages 20 to 35]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ages 36 to 50 × 20/4</td>
<td>−0.0013 (0.0031)</td>
<td>−0.0160*** (0.0046)</td>
<td>−0.0205*** (0.0056)</td>
<td>−0.0350*** (0.0058)</td>
<td>−0.0465*** (0.0077)</td>
</tr>
<tr>
<td>Ages 51 to 65 × 20/4</td>
<td>0.0011 (0.0031)</td>
<td>−0.0083* (0.0048)</td>
<td>−0.0117** (0.0058)</td>
<td>−0.0297*** (0.0060)</td>
<td>−0.0556*** (0.0080)</td>
</tr>
<tr>
<td><strong>Policy [Robust-response states]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low-response states × 20/4</td>
<td>−0.0022 (0.0038)</td>
<td>−0.0012 (0.0064)</td>
<td>−0.0054 (0.0075)</td>
<td>−0.0008 (0.0078)</td>
<td>−0.0136 (0.0107)</td>
</tr>
<tr>
<td><strong>COVID-19 cases [Low-risk states]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium-risk states × 20/4</td>
<td>0.0010 (0.0032)</td>
<td>0.0119** (0.0048)</td>
<td>0.0119** (0.0059)</td>
<td>0.0146** (0.0060)</td>
<td>0.0228*** (0.0082)</td>
</tr>
<tr>
<td>High-risk states × 20/4</td>
<td>−0.0013 (0.0030)</td>
<td>0.0304*** (0.0046)</td>
<td>0.0286*** (0.0056)</td>
<td>0.0317*** (0.0057)</td>
<td>0.0401*** (0.0077)</td>
</tr>
<tr>
<td><strong>City [Central city]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outside central city × 20/4</td>
<td>0.0050* (0.0030)</td>
<td>−0.0009 (0.0044)</td>
<td>0.0035 (0.0053)</td>
<td>−0.0010 (0.0055)</td>
<td>0.0000 (0.0073)</td>
</tr>
<tr>
<td>Not in metropolitan area × 20/4</td>
<td>−0.0020 (0.0039)</td>
<td>−0.0240*** (0.0062)</td>
<td>−0.0270*** (0.0074)</td>
<td>−0.0320*** (0.0074)</td>
<td>−0.0286*** (0.0101)</td>
</tr>
<tr>
<td><strong>FE</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Observations</td>
<td>78,051</td>
<td>78,051</td>
<td>78,051</td>
<td>78,179</td>
<td>103,457</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.015</td>
<td>0.123</td>
<td>0.107</td>
<td>0.063</td>
<td>0.083</td>
</tr>
</tbody>
</table>

**NOTE:** FE, industry and occupation fixed effects. Reference groups are in brackets. Robust standard errors are in parentheses. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. 

**SOURCE:** Authors’ calculations from the CPS.
### Table 2
**COVID-19 Shock: February 2021**

<table>
<thead>
<tr>
<th></th>
<th>(1) Jobless unemployment</th>
<th>(2) Furlough</th>
<th>(3) Unemployment</th>
<th>(4) Unemployment without FE</th>
<th>(5) Nonemployment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender [Males]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females × 21/2</td>
<td>0.0035 (0.0032)</td>
<td>0.0047** (0.0019)</td>
<td>0.0088** (0.0040)</td>
<td>0.0040 (0.0037)</td>
<td>−0.0008 (0.0062)</td>
</tr>
<tr>
<td><strong>Race [Whites]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blacks × 21/2</td>
<td>0.0022 (0.0063)</td>
<td>0.0035 (0.0028)</td>
<td>0.0150** (0.0075)</td>
<td>0.0164** (0.0076)</td>
<td>0.0109 (0.0109)</td>
</tr>
<tr>
<td>Hispanics × 21/2</td>
<td>0.0017 (0.0042)</td>
<td>0.0013 (0.0027)</td>
<td>0.0083 (0.0054)</td>
<td>0.0141*** (0.0055)</td>
<td>0.0083 (0.0087)</td>
</tr>
<tr>
<td>Asians × 21/2</td>
<td>−0.0146*** (0.0048)</td>
<td>0.0058** (0.0029)</td>
<td>−0.0054 (0.0063)</td>
<td>−0.0029 (0.0066)</td>
<td>−0.0189 (0.0122)</td>
</tr>
<tr>
<td><strong>Education [High school or less]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Some college × 21/2</td>
<td>0.0049 (0.0044)</td>
<td>0.0028 (0.0027)</td>
<td>0.0032 (0.0056)</td>
<td>−0.0053 (0.0055)</td>
<td>−0.0074 (0.0085)</td>
</tr>
<tr>
<td>College × 21/2</td>
<td>0.0065 (0.0043)</td>
<td>−0.0021 (0.0025)</td>
<td>0.0048 (0.0054)</td>
<td>−0.0183*** (0.0048)</td>
<td>−0.0268*** (0.0076)</td>
</tr>
<tr>
<td><strong>Age [Ages 20 to 35]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ages 36 to 50 × 21/2</td>
<td>−0.0009 (0.0036)</td>
<td>0.0023 (0.0021)</td>
<td>−0.0000 (0.0045)</td>
<td>−0.0031 (0.0046)</td>
<td>−0.0194*** (0.0073)</td>
</tr>
<tr>
<td>Ages 51 to 65 × 21/2</td>
<td>0.0052 (0.0036)</td>
<td>0.0032 (0.0022)</td>
<td>0.0065 (0.0046)</td>
<td>0.0020 (0.0047)</td>
<td>−0.0174* (0.0078)</td>
</tr>
<tr>
<td><strong>Policy [Robust-response states]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rapid-rollback states × 21/2</td>
<td>0.0007 (0.0035)</td>
<td>−0.0049** (0.0021)</td>
<td>−0.0058 (0.0044)</td>
<td>−0.0061 (0.0045)</td>
<td>−0.0146** (0.0074)</td>
</tr>
<tr>
<td>Low-response states × 21/2</td>
<td>0.0087* (0.0049)</td>
<td>0.0022 (0.0026)</td>
<td>0.0075 (0.0060)</td>
<td>0.0092 (0.0061)</td>
<td>−0.0080 (0.0107)</td>
</tr>
<tr>
<td><strong>COVID-19 cases [Low-risk states]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium-risk states × 21/2</td>
<td>−0.0103** (0.0041)</td>
<td>−0.0042* (0.0023)</td>
<td>−0.0139*** (0.0051)</td>
<td>−0.0125** (0.0052)</td>
<td>−0.0132 (0.0093)</td>
</tr>
<tr>
<td>High-risk states × 21/2</td>
<td>0.0013 (0.0039)</td>
<td>0.0050** (0.0023)</td>
<td>0.0073 (0.0048)</td>
<td>0.0092* (0.0050)</td>
<td>0.0133 (0.0085)</td>
</tr>
<tr>
<td><strong>City [Central city]</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outside central city × 21/2</td>
<td>−0.0066* (0.0034)</td>
<td>−0.0027 (0.0019)</td>
<td>−0.0107*** (0.0042)</td>
<td>−0.0125*** (0.0043)</td>
<td>−0.0112 (0.0070)</td>
</tr>
<tr>
<td>Not in metropolitan area × 21/2</td>
<td>−0.0146*** (0.0050)</td>
<td>−0.0056* (0.0031)</td>
<td>−0.0266*** (0.0062)</td>
<td>−0.0302*** (0.0061)</td>
<td>−0.0296*** (0.0103)</td>
</tr>
<tr>
<td><strong>FE</strong></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Observations</td>
<td>73,926</td>
<td>73,926</td>
<td>73,926</td>
<td>74,031</td>
<td>96,511</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.021</td>
<td>0.017</td>
<td>0.036</td>
<td>0.020</td>
<td>0.071</td>
</tr>
</tbody>
</table>

**NOTE:** FE, industry and occupation fixed effects. Reference groups are in brackets. Robust standard errors are in parentheses. *p < 0.10, **p < 0.05, ***p < 0.010.

**SOURCE:** Authors’ calculations from the CPS.
ment and nonemployment, suggesting that people’s voluntary reduction of economic activities out of fear is an important channel through which the pandemic has hampered the economy.9

The final few rows show that those living outside metropolitan areas sustained fewer job losses, even controlling for all other factors. One explanation is that in April 2020, urban areas on average had more stringent lockdowns. (Our policy variables are constructed at the state level.) In addition, with a lower population density in rural areas, even the same social distancing measures represent less of a restriction on economic activities in those areas.

Table 2 shows the estimation results for the changes between February 2020 and February 2021, 10 months into the recovery and the last month we can compare year-over-year with the pre-pandemic statistics. The year-over-year increase in the aggregate unemployment rate was 2.7 percentage points, or only a quarter of the increase in April 2020 over April 2019. This number can help interpret the magnitudes of the estimated coefficients in Table 2.

Consistent with Figure 2, the differential effects of the pandemic on men’s and women’s unemployment had all but disappeared by February 2021. Controlling for industries and occupations, the differential effects on women’s furloughs and unemployment are statistically significant but the magnitudes are small.

As for minorities, only Blacks exhibit a larger shock to unemployment when industry and occupation fixed effects are controlled for (Column 3). Since Blacks were hit less hard than even Whites in April 2020, this finding shows that Blacks have been the slowest to recover. Hispanics still have a somewhat higher unemployment effect in Column 4, implying that they tend to work for industries and in occupations that have been recovering more slowly. However, in terms of nonemployment, there is no difference across these groups even when industry and occupation compositions are not considered.

We also see that by February 2021, differential effects across education groups and across age groups evaporate once industries and occupations are controlled for.10 The larger effects on the nonemployment of the young and the less educated (column 5) show that these are compositional effects: The industries and occupations that are over-represented by these groups have been recovering more slowly than other industries and occupations.11

Table 2 also shows that state-level policies do have some effect on employment outcomes in February 2021, but the differences are negligible once industries and occupations are controlled for. At the same time, somewhat surprisingly, medium-risk states had better outcomes and high-risk states had only slightly worse outcomes than low-risk states. This suggests that the fear effect evident in April 2020 may have lessened thereafter, possibly because people reassessed infection risks or adopted other ways of mitigating the risk (e.g., wearing masks).12

Finally, the employment of city center residents has been the slowest to recover. The most likely explanation is that remote work reduced not only the number of workers in city centers but also these workers’ demand for local consumer service businesses, further worsening the employment prospects of city center residents (Eckert et al., 2020).

4 CONCLUDING REMARKS

The economic impact of the pandemic was unequal across demographic and socioeconomic groups. The initial shock hit women harder than men, but the differential effects
disappeared by February 2021. Similarly, Hispanics and Asians were hit harder than Blacks and Whites in April 2020, but both groups have recovered quite a bit, especially Hispanics. Blacks on the other hand, in spite of having a smaller initial shock than all other racial groups, have experienced slower recoveries in their employment outcomes. These results remain even after controlling for all other factors, including industries, occupations, state-level policies and statewide infection rates, and urban/rural residence. In this context, it is not clear what explains the slower reduction of Black unemployment. One possibility is that our industry and occupation classifications are not detailed enough (a choice we made in recognition of the sample size of the CPS) and we are not fully capturing the compositional effects. We leave this question for future research but note that the remaining effect on the Black nonemployment rate is not significantly different from that for Whites in February 2021.

By educational attainment, the less educated were hit worse than the more educated in April 2020. By February 2021, the differential effects across education groups had gotten smaller but still remained due to the different industries and occupations they work in.

In addition, while the young were harder hit initially, by February 2021, there was no systematic difference in the employment impact of the pandemic across age groups, except that the young in certain industries and occupations had left the workforce altogether.

Our findings call for a careful investigation of the mechanism through which different demographic and socioeconomic groups have been affected unequally by the pandemic, not only on impact but also during the recovery.

APPENDIX

Tables A1 and A2 report the coefficients on the industry and occupation fixed effects in regressions (1) to (3) of Tables 1 and 2 in the main text.

By industry, we see that the leisure and hospitality industry and the other services industry were hit the hardest, and they still had not recovered even by February 2021. However, taking into account the large, negative initial hit the other services industry took, it showed the fastest recovery. On the other hand, the public administration industry and the agriculture industry not only suffered less but also recovered faster than most other industries. Since most agricultural work can be done outdoors with ample room for social distancing, this finding is not surprising. Although the initial impact of the shock was hard on the manufacturing industry and the education and health services industry, their speeds of recovery surpassed even that of the public administration industry. The financial activities industry was relatively safe, which contrasts with the effect on that industry from the 2008-09 Financial Crisis.

Looking into occupations, we find that service and construction occupations were the hardest hit and slowest to recover. Production occupations also suffered initially but recovered faster than other occupations. Management, business, and financial occupations; office and administrative support occupations; and professional and related occupations suffered the least and also recovered the fastest. This may have been thanks to their ability to work remotely, which is in line with results from Montenovo et al. (2020).
### Table A1
COVID-19 Shock: April 2020

<table>
<thead>
<tr>
<th>Industry [Public administration]</th>
<th>(1) Jobless unemployment</th>
<th>(2) Furlough</th>
<th>(3) Unemployment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mining × 20/4</td>
<td>0.0268</td>
<td>0.0170</td>
<td>0.0559**</td>
</tr>
<tr>
<td></td>
<td>(0.0174)</td>
<td>(0.0163)</td>
<td>(0.0258)</td>
</tr>
<tr>
<td>Construction × 20/4</td>
<td>−0.0138*</td>
<td>0.0731***</td>
<td>0.0609***</td>
</tr>
<tr>
<td></td>
<td>(0.0074)</td>
<td>(0.0108)</td>
<td>(0.0134)</td>
</tr>
<tr>
<td>Manufacturing × 20/4</td>
<td>0.0041</td>
<td>0.0660***</td>
<td>0.0728***</td>
</tr>
<tr>
<td></td>
<td>(0.0053)</td>
<td>(0.0081)</td>
<td>(0.0102)</td>
</tr>
<tr>
<td>Wholesale and retail trade × 20/4</td>
<td>−0.0007</td>
<td>0.0747***</td>
<td>0.0756***</td>
</tr>
<tr>
<td></td>
<td>(0.0060)</td>
<td>(0.0086)</td>
<td>(0.0109)</td>
</tr>
<tr>
<td>Transportation and utilities × 20/4</td>
<td>−0.0076</td>
<td>0.0619***</td>
<td>0.0539***</td>
</tr>
<tr>
<td></td>
<td>(0.0067)</td>
<td>(0.0100)</td>
<td>(0.0124)</td>
</tr>
<tr>
<td>Information × 20/4</td>
<td>0.0145</td>
<td>0.0494***</td>
<td>0.0706***</td>
</tr>
<tr>
<td></td>
<td>(0.0104)</td>
<td>(0.0119)</td>
<td>(0.0171)</td>
</tr>
<tr>
<td>Financial activities × 20/4</td>
<td>0.0004</td>
<td>0.0186***</td>
<td>0.0220**</td>
</tr>
<tr>
<td></td>
<td>(0.0055)</td>
<td>(0.0070)</td>
<td>(0.0094)</td>
</tr>
<tr>
<td>Professional and business services × 20/4</td>
<td>0.0041</td>
<td>0.0384***</td>
<td>0.0444***</td>
</tr>
<tr>
<td></td>
<td>(0.0048)</td>
<td>(0.0068)</td>
<td>(0.0087)</td>
</tr>
<tr>
<td>Education and health services × 20/4</td>
<td>0.0048</td>
<td>0.0620***</td>
<td>0.0678***</td>
</tr>
<tr>
<td></td>
<td>(0.0040)</td>
<td>(0.0065)</td>
<td>(0.0079)</td>
</tr>
<tr>
<td>Leisure and hospitality × 20/4</td>
<td>0.0146**</td>
<td>0.2450***</td>
<td>0.2740***</td>
</tr>
<tr>
<td></td>
<td>(0.0071)</td>
<td>(0.0120)</td>
<td>(0.0136)</td>
</tr>
<tr>
<td>Other services × 20/4</td>
<td>0.0074</td>
<td>0.1340***</td>
<td>0.1500***</td>
</tr>
<tr>
<td></td>
<td>(0.0061)</td>
<td>(0.0125)</td>
<td>(0.0142)</td>
</tr>
<tr>
<td>Agriculture, forestry, fishing, and hunting × 20/4</td>
<td>0.0025</td>
<td>0.0153</td>
<td>0.0206</td>
</tr>
<tr>
<td></td>
<td>(0.0134)</td>
<td>(0.0132)</td>
<td>(0.0186)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Occupation [Management, business, and financial]</th>
<th>(1) Jobless unemployment</th>
<th>(2) Furlough</th>
<th>(3) Unemployment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professional and related × 20/4</td>
<td>−0.0062*</td>
<td>0.0332***</td>
<td>0.0260***</td>
</tr>
<tr>
<td></td>
<td>(0.0034)</td>
<td>(0.0045)</td>
<td>(0.0059)</td>
</tr>
<tr>
<td>Service × 20/4</td>
<td>−0.0072</td>
<td>0.1150***</td>
<td>0.1120***</td>
</tr>
<tr>
<td></td>
<td>(0.0048)</td>
<td>(0.0079)</td>
<td>(0.0093)</td>
</tr>
<tr>
<td>Sales and related × 20/4</td>
<td>−0.0114*</td>
<td>0.0668***</td>
<td>0.0643***</td>
</tr>
<tr>
<td></td>
<td>(0.0059)</td>
<td>(0.0079)</td>
<td>(0.0101)</td>
</tr>
<tr>
<td>Office and administrative support × 20/4</td>
<td>−0.0105**</td>
<td>0.0395***</td>
<td>0.0330***</td>
</tr>
<tr>
<td></td>
<td>(0.0048)</td>
<td>(0.0064)</td>
<td>(0.0082)</td>
</tr>
<tr>
<td>Farming, fishing, and forestry × 20/4</td>
<td>−0.0492**</td>
<td>0.0622***</td>
<td>0.0096</td>
</tr>
<tr>
<td></td>
<td>(0.0223)</td>
<td>(0.0256)</td>
<td>(0.0330)</td>
</tr>
<tr>
<td>Construction and extraction × 20/4</td>
<td>0.0147*</td>
<td>0.0881***</td>
<td>0.1110***</td>
</tr>
<tr>
<td></td>
<td>(0.0086)</td>
<td>(0.0122)</td>
<td>(0.0152)</td>
</tr>
<tr>
<td>Installation, maintenance, and repair × 20/4</td>
<td>−0.0132*</td>
<td>0.0573***</td>
<td>0.0462***</td>
</tr>
<tr>
<td></td>
<td>(0.0071)</td>
<td>(0.0123)</td>
<td>(0.0146)</td>
</tr>
<tr>
<td>Production × 20/4</td>
<td>−0.0163**</td>
<td>0.1020***</td>
<td>0.0922***</td>
</tr>
<tr>
<td></td>
<td>(0.0069)</td>
<td>(0.0112)</td>
<td>(0.0134)</td>
</tr>
<tr>
<td>Transportation and material moving × 20/4</td>
<td>−0.0014</td>
<td>0.0738***</td>
<td>0.0784***</td>
</tr>
<tr>
<td></td>
<td>(0.0073)</td>
<td>(0.0100)</td>
<td>(0.0125)</td>
</tr>
<tr>
<td>Observations</td>
<td>78,051</td>
<td>78,051</td>
<td>78,051</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.015</td>
<td>0.123</td>
<td>0.107</td>
</tr>
</tbody>
</table>

Note: Reference groups are in brackets. Robust standard errors are in parentheses. *p < 0.10, **p < 0.05, ***p < 0.010. SOURCE: Authors’ calculations from the CPS.
## Table A2
COVID-19 Recovery: February 2021

<table>
<thead>
<tr>
<th>Industry [Public administration]</th>
<th>(1) Jobless unemployment</th>
<th>(2) Furlough unemployment</th>
<th>(3) Unemployment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mining × 21/2</td>
<td>0.0360</td>
<td>0.0208</td>
<td>0.1070**</td>
</tr>
<tr>
<td></td>
<td>(0.0310)</td>
<td>(0.0155)</td>
<td>(0.0415)</td>
</tr>
<tr>
<td>Construction × 21/2</td>
<td>0.0069</td>
<td>0.0064</td>
<td>0.0257**</td>
</tr>
<tr>
<td></td>
<td>(0.0074)</td>
<td>(0.0055)</td>
<td>(0.0103)</td>
</tr>
<tr>
<td>Manufacturing × 21/2</td>
<td>−0.0009</td>
<td>−0.0019</td>
<td>0.0022</td>
</tr>
<tr>
<td></td>
<td>(0.0060)</td>
<td>(0.0037)</td>
<td>(0.0079)</td>
</tr>
<tr>
<td>Wholesale and retail trade × 21/2</td>
<td>0.0155**</td>
<td>−0.0002</td>
<td>0.0245***</td>
</tr>
<tr>
<td></td>
<td>(0.0068)</td>
<td>(0.0037)</td>
<td>(0.0085)</td>
</tr>
<tr>
<td>Transportation and utilities × 21/2</td>
<td>0.0135*</td>
<td>0.0061</td>
<td>0.0315***</td>
</tr>
<tr>
<td></td>
<td>(0.0080)</td>
<td>(0.0050)</td>
<td>(0.0104)</td>
</tr>
<tr>
<td>Information × 21/2</td>
<td>0.0276**</td>
<td>0.0046</td>
<td>0.0365***</td>
</tr>
<tr>
<td></td>
<td>(0.0111)</td>
<td>(0.0049)</td>
<td>(0.0135)</td>
</tr>
<tr>
<td>Financial activities × 21/2</td>
<td>0.0103*</td>
<td>−0.0002</td>
<td>0.0127</td>
</tr>
<tr>
<td></td>
<td>(0.0063)</td>
<td>(0.0033)</td>
<td>(0.0079)</td>
</tr>
<tr>
<td>Professional and business services × 21/2</td>
<td>0.0086</td>
<td>0.0018</td>
<td>0.0144*</td>
</tr>
<tr>
<td></td>
<td>(0.0059)</td>
<td>(0.0038)</td>
<td>(0.0076)</td>
</tr>
<tr>
<td>Education and health services × 21/2</td>
<td>−0.0022</td>
<td>−0.0013</td>
<td>−0.0024</td>
</tr>
<tr>
<td></td>
<td>(0.0046)</td>
<td>(0.0029)</td>
<td>(0.0061)</td>
</tr>
<tr>
<td>Leisure and hospitality × 21/2</td>
<td>0.0294***</td>
<td>0.0132***</td>
<td>0.0610***</td>
</tr>
<tr>
<td></td>
<td>(0.0085)</td>
<td>(0.0050)</td>
<td>(0.0107)</td>
</tr>
<tr>
<td>Other services × 21/2</td>
<td>0.0159**</td>
<td>0.0067</td>
<td>0.0326***</td>
</tr>
<tr>
<td></td>
<td>(0.0077)</td>
<td>(0.0046)</td>
<td>(0.0102)</td>
</tr>
<tr>
<td>Agriculture, forestry, fishing, and hunting × 21/2</td>
<td>−0.0201*</td>
<td>0.0281*</td>
<td>0.0165</td>
</tr>
<tr>
<td></td>
<td>(0.0116)</td>
<td>(0.0159)</td>
<td>(0.0203)</td>
</tr>
<tr>
<td>Occupation [Management, business, and financial]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Professional and related × 21/2</td>
<td>0.0009</td>
<td>−0.0003</td>
<td>0.0008</td>
</tr>
<tr>
<td></td>
<td>(0.0038)</td>
<td>(0.0018)</td>
<td>(0.0046)</td>
</tr>
<tr>
<td>Service × 21/2</td>
<td>0.0181***</td>
<td>0.0069***</td>
<td>0.0386***</td>
</tr>
<tr>
<td></td>
<td>(0.0058)</td>
<td>(0.0034)</td>
<td>(0.0073)</td>
</tr>
<tr>
<td>Sales and related × 21/2</td>
<td>−0.0056</td>
<td>0.0059*</td>
<td>0.0023</td>
</tr>
<tr>
<td></td>
<td>(0.0066)</td>
<td>(0.0033)</td>
<td>(0.0079)</td>
</tr>
<tr>
<td>Office and administrative support × 21/2</td>
<td>0.0044</td>
<td>−0.0003</td>
<td>0.0115*</td>
</tr>
<tr>
<td></td>
<td>(0.0053)</td>
<td>(0.0026)</td>
<td>(0.0066)</td>
</tr>
<tr>
<td>Farming, fishing, and forestry × 21/2</td>
<td>0.0350</td>
<td>−0.0349</td>
<td>0.0004</td>
</tr>
<tr>
<td></td>
<td>(0.0257)</td>
<td>(0.0243)</td>
<td>(0.0363)</td>
</tr>
<tr>
<td>Construction and extraction × 21/2</td>
<td>0.0156*</td>
<td>0.0096</td>
<td>0.0266**</td>
</tr>
<tr>
<td></td>
<td>(0.0091)</td>
<td>(0.0073)</td>
<td>(0.0128)</td>
</tr>
<tr>
<td>Installation, maintenance, and repair × 21/2</td>
<td>0.0022</td>
<td>0.0016</td>
<td>0.0134</td>
</tr>
<tr>
<td></td>
<td>(0.0074)</td>
<td>(0.0044)</td>
<td>(0.0102)</td>
</tr>
<tr>
<td>Production × 21/2</td>
<td>0.0021</td>
<td>0.0018</td>
<td>0.0109</td>
</tr>
<tr>
<td></td>
<td>(0.0086)</td>
<td>(0.0051)</td>
<td>(0.0108)</td>
</tr>
<tr>
<td>Transportation and material moving × 21/2</td>
<td>0.0135*</td>
<td>0.0114***</td>
<td>0.0267***</td>
</tr>
<tr>
<td></td>
<td>(0.0079)</td>
<td>(0.0049)</td>
<td>(0.0100)</td>
</tr>
<tr>
<td>Observations</td>
<td>73,926</td>
<td>73,926</td>
<td>73,926</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.021</td>
<td>0.017</td>
<td>0.036</td>
</tr>
</tbody>
</table>

NOTE: Reference groups are in brackets. Robust standard errors are in parentheses. *p < 0.10, **p < 0.05, ***p < 0.010. SOURCE: Authors’ calculations from the CPS.
NOTES

1 The paper most closely related to ours is Couch, Fairlie, and Xu (2020), which compares the experiences of Blacks, Hispanics, and Asians relative to those of Whites from April to June 2020. Our results complement theirs with data from later months and show new evidence for the recovery phase.

2 When added together, the total for recall unemployment and jobless unemployment is smaller than total unemployment, but the difference is small.

3 The CDC COVID Data Tracker is available at https://covid.cdc.gov/covid-data-tracker.

4 The OxCGRT is available at www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker.

5 The online repository provides detailed coding information and is available at https://github.com/OxCGRT/covid-policy-tracker/blob/master/documentation/codebook.md.

6 The CPS has outgoing rotation samples, and the BLS interviews each household for four consecutive months. The household leaves the sample for the next eight months and returns for another four months. The sample-collecting process happens every month, so only a quarter of the sample can be tracked from one month to the next.

7 We have analyzed all months from October 2020 to February 2021. From November 2020 onward, there are almost no differences in our estimates. Lee, Park, and Shin (2021), the earlier version of this article, has the results through November 2020.

8 Table A1 in the appendix, for April 2020, shows that by industry, the leisure and hospitality industry and the other services industry were hit the hardest, while the service, construction, and production occupations suffered more than other occupations.

9 This is consistent with evidence from other countries. See Aum et al. (2020a), for example.

10 The results show that by February 2021 the impact on more-educated and less-educated workers was consistent with that in Forsythe et al. (2021), which shows that labor market tightness converged for college-educated and high-school workers.

11 Among industries, as of February 2021, the leisure and hospitality industry had not recovered from the shock. There is not much of a pattern across occupations, except that service occupations still showed a significantly higher unemployment rate from their February 2020 level.

12 These estimates are different from the November 2020 estimates. In the November data, states that rolled back containment policies or implemented less-restrictive policies had a smaller year-over-year rise in unemployment than states with more-restrictive policies. Furthermore, statewide infection rates in the preceding month were uncorrelated with employment outcomes.

REFERENCES


1 INTRODUCTION

Over the past half-century, some of the most striking socioeconomic changes in developed countries have been the radical changes in the family structure, fertility behavior, and the division of labor within the household. These changes have consequences for labor market productivity and the viability of Social Security and other programs. For example, the decline in fertility below the replacement rate has led to a major concern for pay-as-you-go social security in Organisation for Economic Co-operation and Development countries, including the United States. Moreover, the significant increase in the percentage of women in the workforce has had a considerable effect on the type of benefits employers offer, specifically regarding a family-friendly workplace, parental leave, and other work-family balance policies.
Several articles analyze the changes in the family structure, fertility behavior, and the division of labor within the household for older generations (see Eckstein, Keane, and Lifshitz, 2019; Kong, Ravikumar and Vandenburgroucke, 2018; and Ramey and Francis, 2009); however, to the best of our knowledge, no study has analyzed these changes in the latest generation—Millennials—and how their family, fertility, and labor market behavior compares with that of the previous generations. The scarcity of studies analyzing the behavior of Millennials is mainly because of a lack of data. Using data from the Panel Study of Income Dynamics (PSID) from 1968 to 2015, this article presents the first read on the behavior of Millennials as they complete their education, form their families, and transition into adulthood.

This article focuses on three key aspects: work, leisure, and family. For each of these, trends over time as well as life-cycle profiles over generations are presented. In addition to this, an Oaxaca-Blinder decomposition of wages into explained and unexplained components is estimated to understand the changes in the gender wage gap. Given the trends in education and the recent convergence in the wage gap, the returns to the labor market are also estimated and the changes are analyzed over generations by race and gender. Along with the changes in education and hours worked, there has been a major change in housework hours as well. Using a linear regression framework, as a first cut, the differential effects of education, race, gender, marital status, and the presence of children on housework hours are estimated. Lastly, with declining marriage rates, the key question of what predicts partner choice is analyzed using a multinomial logit model.

In the article, we assign the generations the following names and birth cohorts:

- **Silent generation**: 1940-49
- **Baby Boomers-1 (Boomers-1)**: 1950-59
- **Baby Boomers-2 (Boomers-2)**: 1960-69
- **Generation X (GenX)**: 1970-79
- **Millennials**: 1980-89

Focusing on labor markets and education decisions, we find that the wage-age profile has been shifting down over generations, with Millennials men having the lowest real wages over the life cycle (up to age 33) so far. However, women’s wages have not decreased but have stagnated, with Millennials women earning lower wages than GenX for the part of the life cycle that can be analyzed (up to age 33). This finding indicates that most of the documented rise in wage inequality has come from men and not women. For both Black men and White men, returns to college graduates (those with a four-year degree or higher) have increased for all generations, with the most significant increase for Millennials, specifically Black men.

Despite the increasing returns to college for men, Black women have always graduated from college at higher rates than Black men; and this gap has only increased over the birth cohorts. On the other hand, up to Boomers-1, White men graduated from college at higher rates than White women; this gap reversed in Boomers-2 and has continued to increase over generations (see Goldin, Katz, and Kuziemko, 2006; Murnane, 2013; Blau and Kahn, 2017; and Eckstein, Keane, and Lifshitz, 2019), with Millennials having the biggest gap between White men and White women. While the college graduation gap between White men and
White women is more than the college graduation gap between Black men and Black women, this outcome reverses if individuals with some college and college graduates are combined.

There has been a significant increase in college graduation rates for Boomers-2 and GenX for both Blacks and Whites, a pattern that has continued to accelerate for White Millennials. This is in sharp contrast to the graduation rates for the Silent and Boomers-1 generations, which were generally stable across the two. Another striking feature of the change in the education distribution over the generations is the significant reduction in high school dropout rates; this has been most pronounced for Blacks. The percentage of Black high school dropouts fell from 27.5 percent for the Silent generation to 7.1 percent for the Millennial generation.

A few studies in the existing literature analyze leisure over generations and the life cycle. Three notable exceptions are Aguiar and Hurst (2007); Ramey and Francis (2009); and Aguiar, Hurst, and Karabarbounis (2012). This article finds that over generations, while the amount of leisure enjoyed by women has been increasing, the amount of leisure enjoyed by men shows no clear pattern. However, for both men and women, Millennials enjoy a higher level of leisure than previous generations. This rise in leisure for women is primarily coming from married women. Splitting the sample by education, this article finds that women who are college graduates enjoy less leisure than those who are high school graduates, and the same is true for men.

The finding that the amount of leisure enjoyed by men is stable over generations masks different dynamics in the components of leisure. There has been a significant reduction in hours worked by men, with Millennials working the least. At the same time, there has been an increase in the housework hours for men, with Millennial men devoting more hours to housework than men of any previous generation. The opposite is true for women: Their reduction in hours devoted to housework over generations has been more than offset by the increase in hours devoted to market work. More importantly, most of this movement has happened within married couples.

The findings of this article are in contrast to Aguiar and Hurst’s (2007), who find that leisure is increasing for everyone. Several factors may account for these conflicting results. First, different data sources are used in this article. Second, this article compares leisure over generations and the life cycle, while they do not. Third, the measure of leisure used here may not coincide with the measure of leisure they use, as they are using self-reported data from time diaries, whereas the measure in this article is the residual of hours worked in the market and in household production. Because of these differences, the measure of leisure used in this article is more comparable with the measure used in Ramey and Francis (2009), and results here confirm the life-cycle analysis reported in that article.

The media has often conjectured that the marriage, cohabitation, divorce, and fertility behaviors of Millennials are radically different from previous generations. A detailed analysis of these behaviors is provided, which has been lacking so far. Marriage (and, thus, fertility) has seen many drastic changes over generations. Lower-educated individuals have retreated from marriage, and this finding is the most pronounced among Millennials. However, while Millennial college graduates are delaying marriage significantly, they do catch up with previous generations by ages 30 to 33. Thus, it is not clear whether Millennials would be marrying at an overall rate lower than previous generations, as the composition of education has also
changed, with a much higher number of college graduates among Millennials. No clear trends in assortative matching are found over generations. Although this is in contrast to the findings of Greenwood and Guner (2008) and Santos and Weiss (2016), it is in line with those of Gihleb and Lang (2016). However, Millennials are cohabiting at a much higher rate than previous generations. For GenX and Millennials, cohabitation rates are much higher early on in life relative to previous generations; however, the rates drop off significantly after that, indicating cohabitation itself is transitory. This was not true for the Silent and Boomer generations. With respect to divorce, the Silent generation has a fairly flat profile over the number of years since marriage. The subsequent generations present rising divorce rates and more pronounced duration dependence; that is, as the length of the marriage increases, the likelihood of divorce decreases, with Millennials having the highest probability of divorce after five years of marriage. The major changes across generations in divorce rates are primarily coming from Blacks and couples where at least one spouse has completed education of less than or equal to a high school diploma. Couples where both spouses have completed some college or more have significantly lower divorce rates, as also documented by Lundberg, Pollak, and Stearns (2016).

Finally, concerning fertility trends, the significant decline in completed fertility is confirmed in this article, with Boomers-2 and GenX fertility rates falling below the replacement rate. Similar trends exist across race. There has been a steady decline across generations at each age parity of the proportion of births to married women, although this decline has accelerated among Millennials. For example, the proportion of births to married women ages 31 to 35 for the Silent generation is 88 percent, which fell to 83 percent for GenX and even further to 75 percent for Millennials. The age-specific fertility rate has declined for every generation up to GenX. Fertility rates for Millennial women ages 18 to 30 are below every previous generation’s; however, for Millennial women ages 31 to 35, the fertility rate is higher than for GenX women of the same ages. This finding suggests that the completed fertility rate for Millennials may not necessarily be below that for GenX; although Millennials are delaying fertility, their fertility rates at later ages might be higher than those of the immediate previous generation.

The remainder of the article proceeds as follows. Section 2 describes the construction of our primary dataset from the PSID. Section 3 presents trends related to education, hours worked, and wages over time as well as by generation. Section 4 details the changes in housework hours and leisure hours over time and by generation. Section 5 delves into time and generation trends for fertility, marriage, and divorce. Finally, Section 6 concludes.

2 DATA

The main dataset used in this analysis is constructed using various files from the PSID. The PSID is a nationally representative household panel survey that includes economic, social, and health information from 1968 to 2015. This survey was conducted annually from 1968 to 1997 and biannually thereafter. To compile the dataset, we used the family-individual files, marriage history files, childbirth and adoption history files, and the T-2 income and transfers files. The latter helps to complete the information regarding labor income and annual hours.
worked in the intervening years during the periods when the survey is conducted biannually. The final sample used in this study contains data on individuals ages 18 to 65, resulting in 38,958 unique individuals and 505,496 individual-year observations.

The employment rate is defined as the fraction of individuals with annual labor market hours greater than zero. Hourly wages are defined as the sum of labor income, farming income, and business income divided by the annual labor market hours worked. All nominal values are deflated to 2015 U.S. dollars using the consumer price index deflator.

The marriage history file contains information regarding individuals’ marital status from 1901 to 2015. The variable “legally married” is taken directly from this file. The PSID separately documented from 1983 onward if the head of the household has a “wife” (cohabiting but not married) for more than a year. Prior to 1983, both a legally married wife and a cohabiting “wife” were grouped without distinction. This presents an issue in constructing a consistent measure for cohabitation. Hence, the marriage information from the marriage history file is used and compared with the marital pairs indicator—whether or not there exists a marital pair in the household from the family-individual file. If the head is not legally married and there exists a marital pair in the household, it is assumed that they are cohabiting. This measure is fairly comparable with the measure constructed from the post-1983 data.

3 WORK

3.1 Education

Throughout the article, two measures of the level of education are used. The first is the years of completed education, and the second is a discrete measure based on the highest level of education completed. To create the second measure, years of completed education are divided into four groups: high school dropouts, high school graduates, those with some college, and college graduates (with a four-year degree or higher). Figure 1 presents years of completed education over time and by generation for individuals ages 30 to 35.

Panels A to E of Figure 1 present the years of completed education over time by gender, marital status, and race. Panel A shows the well-documented reversal in the gender education gap, where prior to the early 1990s, men were more educated than women; however, post 1990s, women have not only caught up to but have also overtaken men. This reversal has occurred because, while the years of completed education have significantly increased over time for both men and women, the rate of increase for women has been significantly faster.

Panels B and C of Figure 1 show that marriage is becoming more concentrated among the highly educated. This is more pronounced among women than among men. Prior to the 1990s, single men were more educated than married men; by the late 1990s, this trend had reversed. However, by 2015, single and married men had similar completed years of education. A similar pattern is observed for women prior to the early 1990s (Panel C); however, married women continued to be more educated than single women by significant margins in 2015. It is important to note that the composition of these groups has also changed as marriage rates have fallen over the past few decades.
Figure 1
Education: Trends and Generations

A. Education: Gender
B. Marital status: Men
C. Marital status: Women
D. Race: Men
E. Race: Women
F. Silent: Whites
G. Boomers-1: Whites
H. Boomers-2: Whites
I. GenX: Whites
J. Millennials: Whites
K. Silent: Blacks
L. Boomers-1: Blacks
M. Boomers-2: Blacks
N. GenX: Blacks
O. Millennials: Blacks

NOTE: All graphs are restricted to head and spouse. The trend graphs are restricted to ages 30 to 35, and the relevant variable is the years of education. For the generation graphs, the unweighted proportions are plotted for ages 30 to 35. <HS, high school dropouts (<12 years of education); HS, high school graduates (12 years of education); SC, those with some college (>12 and <16 years of education); and College, college graduates (with a four-year degree or higher, 16 or more years of education).

SOURCE: PSID and authors’ calculations.
Panels D and E of Figure 1 show that the significant convergence of the racial education gap stalled and since the early 2000s has reversed. As shown in Panel D, the racial education gap between Black men and White men trended toward convergence; however, post 2004, the trend reversed and started diverging. A similar trend is seen for Black women and White women as well (Panel E); however, the divergence occurs at a much later point in time—post 2007.

Table 1 presents the years of completed education by gender, marital status, and race over generations. The table pins down the reversal of the gender education gap: It occurred with the Boomers-2 generation, where men have only 13.10 years of education compared with 13.40 years for women. For the Millennials generation, the levels have increased for men and women but the gap persists (13.91 years for men versus 14.44 years for women). Similar to Figure 1, Table 1 also shows reversals in the education gaps by marital status—for both men and women in the Boomers-2 generation. However, the convergence for men is not seen. Interestingly, the racial gap in years of education follows a similar trend for both men and women: It narrows until the Boomers-2 generation and then starts widening again.

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th></th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Silent</td>
<td>Boomers-1</td>
<td>Boomers-2</td>
</tr>
<tr>
<td></td>
<td>(2.76)</td>
<td>(2.31)</td>
<td>(2.81)</td>
</tr>
<tr>
<td>Observations</td>
<td>6,981</td>
<td>11,707</td>
<td>9,143</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.86)</td>
<td>(2.31)</td>
<td>(2.43)</td>
</tr>
<tr>
<td>Observations</td>
<td>618</td>
<td>2,247</td>
<td>2,268</td>
</tr>
<tr>
<td></td>
<td>(2.49)</td>
<td>(2.31)</td>
<td>(2.91)</td>
</tr>
<tr>
<td>Observations</td>
<td>5,200</td>
<td>2,247</td>
<td>6,676</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.61)</td>
<td>(2.32)</td>
<td>(2.77)</td>
</tr>
<tr>
<td>Observations</td>
<td>4,912</td>
<td>7,262</td>
<td>5,860</td>
</tr>
<tr>
<td></td>
<td>(2.51)</td>
<td>(1.87)</td>
<td>(2.12)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,942</td>
<td>4,025</td>
<td>2,432</td>
</tr>
</tbody>
</table>

**NOTE:** The data are calculated for ages 30 to 35. Marital status is defined as the marital status at the time of the observation. **SOURCE:** PSID and authors’ calculations.
Gayle, Odio-Zuniga, Ramakrishnan

Focusing on years of education only tells half the story. The transition between education groups also needs to be understood. There is certainly no doubt that everyone is getting more educated—but where are the gains coming from? Is it from a rise in the college graduates category or some other category? Panels F to O of Figure 1 present the discrete measures of completed education by gender and race. For the Silent and Boomers-1 generations, White men had much higher rates of college completion (college graduates category) than White women; however, the reversal started with the Boomers-2 generation and has increased by a significant margin for the Millennials generation. Namely, 43 percent of Millennial men and 54 percent of Millennial women are college graduates compared with 40 percent and 29 percent of the Silent generation, respectively. The similarities between GenX and Millennials in terms of the proportions educated in each age group is striking, with Millennials being more educated, as a higher proportion are college graduates.

While there has not been much progress at the top of the education distribution for Blacks, there has been some progress for them at the bottom of the distribution, with a fall in high school dropouts from the Silent to the Boomers-2 generation (Panels K to O of Figure 1). For example, 29 percent and 26 percent of Black men and Black women, respectively, of the Silent generation were high school dropouts, but those numbers fell to 8.5 percent and 6.0 percent, respectively, for Black Millennials. The college graduation rate for Blacks rose significantly from the Silent generation to GenX; however, for Black Millennials, the college graduation rate has fallen.

Most articles in the literature tend to focus on the stock of educated individuals, that is, years of education for individuals ages 25 to 64, for example, Blau and Kahn (2017). While these numbers are not directly comparable to our numbers, the same trends are seen for the reversal of the gender gap. As documented by Goldin, Katz, and Kuziemko (2006), there has been a sharp rise in college graduation rates for women as well as a slower rise for men. They suggest that one of the reasons for this rise for women is changing social norms and expectations about work, marriage, and motherhood. Similar to Murnane (2013), we find that Blacks are less likely to graduate from college, while women are more likely to. Eckstein, Keane, and Lifshitz (2019) also document the reversal in the trend of education of married women versus single women. They find that while in 1962 only 7 percent of married women and 10 percent of unmarried women had a bachelor’s degree or higher, by 2015 this pattern had reversed: 36 percent of married women but only 28 percent of unmarried women had a college degree (or higher). This finding is consistent with the trend shown here.

3.2 Employment and Annual Hours Worked

In this section, the employment rate is defined as the fraction of the total working-age population employed. The working-age population refers to individuals ages 18 to 65, and an individual is classified as employed if they work a positive number of hours during the calendar year.

Figure 2 presents the patterns of the employment rate for the sample from 1968 to 2015 for different demographic groups and the life-cycle profiles across generations. Panel A of Figure 2 shows that although the employment rate for women increased from 1968 to 2015, the rate of increase slowed considerably from 2000 to 2015. At the same time, the employment
Figure 2
The Employment Rate: Trends and Generations

NOTE: An individual is classified as employed if a positive number of hours are spent working and as unemployed if zero hours are worked. All graphs are restricted to head and spouse of the family unit. The trend graphs are restricted to ages 18 to 65. The generation graphs are also restricted to ages 18 to 65 and are plotted for three age-group intervals for smoothing of the trend. For example, age 18 refers to the 18-20 age group.

SOURCE: PSID and authors' calculations.
rate for men decreased from 1968 to 1993, with a slight recovery thereafter. The employment rate decreased during the Great Recession for both genders; however, the fall was greater for men than for women. For example, the employment rates in 2008 were 92 percent and 80 percent for men and women, respectively; in 2010, they dropped to 89 percent and 79 percent, respectively; and by 2015 they were 87 percent and 80 percent, respectively.

After controlling for marital status, the employment gap between married and single men widened after 1990, while the employment gap between married and single women shrunk but persisted throughout the whole period (Panels B and C of Figure 2, respectively). On the other hand, the employment rates for Black men fell more sharply than those for White men and the racial employment gap between them has continued to widen (Panel D). In the early 1980s, Black women had higher employment rates than White women. The gap was the smallest in the early 2000s, but by 2008, the recession affected Black women more, spreading the gap again (Panel E).

Panels G to O of Figure 2 analyze employment rates using the life-cycle profiles over generations. The increases in employment rates by generation are shown in Panel G, with GenX and Millennials having fairly similar rates. Separating the trends by gender shows that the rises in the overall employment rates have primarily come from women (Panel I) and that men’s employment rates have steadily declined (Panel H), with Millennials men working the least of all generations. In contrast, Millennials women have higher employment rates than all other generations of women except GenX. For women ages 30 to 32, the employment rate rose from 67.6 percent for the Silent generation to 84.2 percent for Millennials; for men of the same ages, it fell from 98.0 percent to 94.4 percent. These patterns for men and women are robust; they persist even after conditioning on marital status and race (Panels J to O). The shape of the age profile for women flattens from the Silent generation to Millennials. This is possibly due to changes in patterns of fertility, which are analyzed in Section 5.1.

Other articles tend to use different definitions of the employment rate and working-age population for two reasons: (i) the legal working age has changed over time and (ii) data sources differ. For example, Ramey and Francis (2009) use data from the U.S. Census and the Bureau of Labor Statistics to calculate the employment rate using the same definition as in this article. However, they provide three age measures for the working-age population: (i) age 10 and older, (ii) age 14 and older (14+), and (iii) ages 14 to 64. They report that for the period 1900-2005, the employment rate rose from 51 percent to 60 percent for measure (i), from 55 percent to 64 percent for measure (ii), and from 56 percent to 73 percent for measure (iii). For comparability, using the PSID dataset, the employment rate is constructed for ages 14+ and ages 14 to 64 (Panel F of Figure 2). The pattern found by Ramey and Francis (2009) is confirmed. In addition, the time series is extended until 2015 and the increasing trend continues until 2009. Due to the impact of the recession, the rate fell from 75 percent in 2009 to 70 percent in 2015 for working ages 14+ and from 87 percent to 84 percent over the same period for working ages 14 to 64. Similarly, Blau and Kahn (2017) find that the employment rates for women increased from 1947 to 2013 and that the gender gap has narrowed due to a steady decline in employment rates for men over this period. In Panel A of Figure 2, the same trend is seen for men and women; however, the findings of this article have higher employment
**Figure 3**

Annual Hours Worked: Trends and Generations

NOTE: All graphs are restricted to head and spouse. The trend graphs are restricted to ages 18 to 65. The generation graphs are also restricted to ages 18 to 65 and are plotted for three age-group intervals for smoothing of the trend. For example, age 18 refers to the 18-20 age group. <High school, high school dropouts; High school, high school graduates; Some college, those with some college; and College, college graduates (with a four-year degree or higher).

SOURCE: PSID and authors’ calculations.
rates than those presented by Blau and Kahn (2017). Greenwood (2019) also finds that over time, women’s labor supply has risen and suggests this may be from a decline in the amount of time women spend on housework, a topic discussed in detail in Section 4.1.

Figure 3 presents the trends and life-cycle profiles for annual hours worked for different demographic groups, which are taken directly from survey data. The results show that annual hours worked mirror the trends of employment rates over time and across generations. The one exception is the racial gap for women, which is non-existent in this case. The long-run trends presented in Figure 3 are supported by the findings of Aguiar and Hurst (2007). They find that hours worked in the labor market decreased significantly for men but increased for women over the period 1965-2003. However, they also find that the average time men and women spent on total market work dropped from 35.9 hours per week to 31.7 hours per week over the period 1965-2003, despite women increasing their time in market work.

3.3 Wages

In this article, wages are calculated in hourly terms, with the total labor income of an individual divided by the total annual hours worked and then deflated to 2015 U.S. dollars values.

Any discussion about wages has to begin with the wage gap. Panel A of Figure 4 shows there has been convergence in men’s and women’s median hourly wages, with the women-to-men median-wage ratio increasing from 59 percent to 83 percent—inchning closer to parity. Although wages have been stagnant for men and women college graduates (Panels B and C, respectively), there has been a clear decline in the wages of all other education groups for men. This decline has resulted in a polarization of wages across education groups. For women, wages have stagnated for all education groups. This finding would imply that the college premium has risen more for men than women. Regarding racial wage gaps (Panel D), the Black-to-White wage ratio has declined for men, whereas for women the 2015 level is similar to the 1968 level.

Panel E of Figure 4 disaggregates the trends by worker type and shows convergence in the women-to-men hourly wage ratio both for all workers and for full-time workers. This finding implies that this ratio for part-time workers is catching up with that of full-time workers. The same exercise for annual earnings shows a similar convergence trend; however, the initial gaps in earnings in 1968 are close to 61 percent for all workers and 45 percent for full-time workers—these fall to 33 percent and 25 percent, respectively, by 2015 (Panel F). The convergence in the annual earnings of men and women is a function of hours worked as well, and as stated earlier, there has been a sharp rise in the hours worked by women and a marginal decline for men.

Real hourly wages for men by generation show a sharp fall from the Silent to the Boomers-2 generation (Panel G). Although there was a rise in real wages over the life cycle for GenX relative to previous generations, these gains have not been seen by Millennials—who have the lowest real wages over the life cycle so far. Specifically, real hourly wages for men ages 30 to 32 were $23.50 for the Silent generation but have fallen to $17.80 for Millennials. A flattening of the wage-age profiles over the generations is also seen, confirming the trend documented...
Figure 4
Wages: Trends and Generations

A. Hourly wage and gender gap
B. Education: Men
C. Education: Women
D. Race wage gap
E. Worker type: Hourly wages
F. Worker type: Annual earnings
G. Men
H. Women
I. Overall
J. Married
K. High school
L. Some college
M. College
N. Wage gap (ages 18-65)
O. Wage gap (ages 20-34)
P. Returns to education

NOTE: All of the graphs are restricted to head and spouse. The trend graphs are restricted to ages 18 to 65. Graphs E and F also present a fitted time trend. Panel P represents the college wage premium as presented in Chiappori, Salanié, and Weiss (2017) but with different definitions of generations. See the main text for more details. The wages used in Panels N and O are log wages and truncated at the bottom 1 and top 99 percentiles. The bar colors on Panel P correspond to the colors/labels on Panel G. The generation graphs are also restricted to ages 18 to 65 and are plotted for three age-group intervals for smoothing of the trend. For example, age 18 refers to the 18-20 age group. <HS, high school dropouts; High school (HS), high school graduates; Some college (SC), those with some college; and College (Coll), college graduates (with a four-year degree or higher). S, Silent generation; B-1, Boomers-1; B-2, Boomers-2; X, GenX; and M, Millennials.

SOURCE: PSID and authors’ calculations.
by Kong, Ravikumar, and Vandenbroucke (2018). It does appear as though men are being left behind, with the stagnation of men’s education as well as a fall in hours worked. Interestingly, this trend is not seen for women. The rise in women’s wages over previous generations begins with GenX and is carried forward by Millennials, although it is not as high for Millennials after age 27 (Panel H). However, in levels, men’s wages are still higher than women’s (Panels G and H).

The median women-to-men wage gaps clearly show that Millennials have made the largest strides toward gender equality in pay (Panels N and O). Moreover, there has been a change in the profiles over the life cycle as well: Earlier they had an inverted-U shape, but the profiles over the life cycle have become much flatter in the recent generations. The trend is similar across married and unmarried individuals. Splitting the wage gap by education group shows that it is much smaller for college graduates than high school graduates or those with only some college (Panels K to M).

**Framework for Analyzing Wages.** Following the human capital accumulation literature (Altuğ and Miller, 1998; Gayle and Miller, 2006; Gayle and Golan, 2012; and Chiappori, Salanié and Weiss, 2017), the following specification is used to analyze wages:

\[
\log(w_{it}) = \sum_{r=1}^{4} \gamma_{1r}d_{i,t-r} + \sum_{r=1}^{4} \gamma_{2r}h_{i,t-r} + \gamma_3a_{it} + \gamma_4a_{it}^2 + \eta_i + \epsilon_{it},
\]

where \(w_{it}\) denotes the hourly wages of individual \(i\) in calendar year \(t\). The return to experience is captured by two components: \(d_{i,t-r}\), the indicator for labor force employment of individual \(i\) in calendar year \(t-r\), and \(h_{i,t-r}\), the hours worked by individual \(i\) in calendar year \(t-r\). The standard age-earnings profile is captured by age, denoted by \(a_{it}\), and age squared. Generically, an individual-specific effect is included and denoted by \(\eta_i\).

**Decomposition of Wage Gaps.** As in Blau and Kahn (2017), the framework in (1) is used to perform an Oaxaca-Blinder decomposition of the different wage gaps into an unexplained and an explained component. This estimation of (1) is done separately by gender, race, and generation. The coefficient on the hours worked is restricted to be the same across \(t-r\); that is, \(\gamma_2 = \gamma_2 \forall r = \{1,2,3,4\}\). The individual-specific component is specified as a linear function of completed education. Completed education is discretized into four categories: high school dropouts, high school graduates, those with some college, and college graduates.

Panel N of Figure 4 shows that there has been a significant decline across generations in the women-to-men mean wage gap for those ages 18 to 65 (where the data for a specific age group are available). However, since all age groups are not available for all generations, Panel O is restricted to ages 20 to 34 for direct comparison and shows the same sharp decline. Ages 18 to 65 have a sharp fall in the unexplained component of the wage gap over the generations, from 30.5 percent for the Silent generation to 8.5 percent for Millennials. However, the proportion in the total wage gap accounted for by the unexplained component has also fallen from 76 percent for the Silent generation to 60 percent for Millennials. Because education and hours worked have risen over generations, this finding makes intuitive sense because experience and education are the variables most predictive of the wage level. However, there appears to be a marginal rise in the explained component for GenX, which then falls for Millennials. A similar trend is observed for ages 20 to 34; the explained component appears to rise for
Boomers-1 relative to the Silent generation and then starts to fall. For GenX, 48 percent of the wage gap is accounted for by the explained component, which falls to 39 percent for Millennials.

**Returns to Education.** Has the return to education changed over generations? To answer this question, a decomposition exercise is performed and is similar to the one in Chiappori, Salanié, and Weiss (2017). First, equation (1) is estimated by gender, education, race, and generation. In this regression, the individual-specific component is unrestricted as a fixed effect. For comparability with the results presented in Chiappori, Salanié, and Weiss (2017), it is assumed that $h_{i,t-r}$ represents the proportion of total time endowment hours worked by individual $i$ in calendar year $t-r$, instead of the actual amount of hours worked. The returns to education are calculated by subtracting the predicted log wage of a college graduate from that of a high school graduate. The college wage premia presented in Panel P of Figure 4 are for an individual who is age 35 and has worked full-time in the past four periods ($(40 \times 52)/ (365.25 \times 24)$).

As Panel P shows, labor market college premia have increased for men (both Black and White) over generations. On the other hand, the trend for women has not been that straightforward. For Black women, it fell from the Silent to the Boomers-2 generation and then rose (representing a U shape) for GenX. For White women, it fell for Boomers-1, increased for Boomers-2, and then continued to fall for the subsequent generations. It is important to note that the differences in these numbers are not statistically significant from zero. A similar fall was reported in Chiappori, Salanié, and Weiss (2017); however, the recent rise in the labor market premia for Black women has not been documented before.

### 4 LEISURE

#### 4.1 Housework Hours

There has been much talk about technological progress helping in the reduction of time spent on housework (Greenwood and Seshadri, 2005). Figure 5 presents the trends for annual housework hours over the years and by generation. Housework hours are calculated on an annual basis by multiplying weekly housework hours (as reported by the PSID) by 52. The PSID does not report weekly housework hours in the years 1975 and 1982 and does not ask this question in the T-2 years. Figure 5 shows some very striking trends. Over the years, men have increased their housework hours, while women have drastically reduced theirs. On an overall level, housework hours have declined, implying that the fall in these hours for women has not been offset by the rise in these hours for men (Panel A of Figure 5). However, no conclusions can be drawn from simply focusing on the hours, as there might have been significant technological progress in home production and, therefore, equal time spent in 1968 and 2015 would produce different levels of output.

Before the 1990s, married men did not spend as much time on housework as single men (not shown on Figure 5); after that, the trend has been similar by marital status. However, single women spend significantly less time on housework than married women (and always have). Most of the decline in housework hours appears to come from married women, who
Figure 5
Annual Housework Hours: Trends and Generations

NOTE: All graphs are restricted to head and spouse. The trend graphs are restricted to ages 18 to 65. Annual housework hours are calculated by taking the weekly housework hours, as reported by the PSID, and multiplying by 52. The generation graphs are also restricted to ages 18 to 65 and are plotted for three age-group intervals for smoothing of the trend. For example, age 18 refers to the 18-20 age group. High school, high school graduates, and College, college graduates (with a four-year degree or higher).

SOURCE: PSID and authors’ calculations.
saw a drastic fall from 1,641 hours in 1968 to 709 hours in 2015 (Panel B of Figure 5). Married women have also increased their labor supply. One argument put forward in the literature that would be consistent with this observation would be a significant improvement in household technology. Since technology costs more, women might need to increase their labor supply to be able to afford it. However, technology might also reduce the time women spend on housework (Greenwood, 2019).

There is not much difference in housework hours by race for men (Panel C), while Black women put in fewer housework hours than White women (Panel D). It is important to note that Black married women spend more hours working in the labor market than White married women; however, since there are more Black single women in our data set, the aggregated trend of Black women is more representative of the single ones.

Examining the data by generation paints an even clearer picture about the change in housework hours. While Millennial men do spend more hours on housework than the previous generations of men, the rise over the generations has not been that dramatic (Panel E). On the other hand, there is a clear fall with each generation for women (Panels F); and as seen earlier, most of this fall comes from married women. When looking at disaggregation by education, it is important to note that across all generations, housework hours decrease as educational attainment increases: Women college graduates put in the least number of hours relative to the other education groups. Of Millennial women ages 30 to 32, the high school graduates spent 883 hours on housework (Panel J), while the college graduates spent 605 hours (Panel L). However, this trend is not seen for the men (Panels I and K). The same patterns for race are seen as in the time-trend panels.

The findings of this article on the fall in the housework hours for women and rise of them for men is confirmed by Ramey and Francis (2009) and Aguiar and Hurst (2007). Both of those articles use different datasets from the PSID. Ramey and Francis (2009) compute housework hours using data from the American Heritage Time Use Survey and American Time Use Survey of the Bureau of Labor Statistics, while Aguiar and Hurst (2007) link five major time-use surveys to get their results. Thus, the trends of rising housework hours for men and falling hours for women appear to be robust to the source of data and method of measurement.

**Framework for Analyzing Housework Hours.** In the previous section, aggregate housework hours were examined over time and by generation; however, it is important to understand what predicts housework hours at the individual level in order to understand the driving force behind these changes. Is it education, marital status, or the number of children? There have been changes over time and generations in education, family structure, and the number of children. These variables are known to be correlated with the number of hours spent on housework. Can they statistically explain the change in housework hours over generations? To answer this question, a statistical decomposition exercise is conducted using the following regression for Blacks and Whites together:
Gayle, Odio-Zuniga, Ramakrishnan

\[ h_{it}^w = \alpha_0 + \sum_{r=2}^{5} \alpha_{1r} g_{ir} + \sum_{r=2}^{5} \beta_{r} g_{ir} Z_{it} + \sum_{r=2}^{5} \delta_{r} g_{ir} Z_{it} + \epsilon_{it}, \]

where \( h_{it}^w \) denotes annual housework hours for individual \( i \) in calendar year \( t \) and \( g_{ir} \) is an indicator equal to 1 if individual \( i \) is from generation \( r \) and zero otherwise. In this specification, the Silent generation is set as the baseline (\( r = 1 \)). The regression includes controls for education, gender, race, marital status, the number of young children (less than age 6), the number of old children (ages 6 to 18), age, and age squared. These are all represented in equation (2) by the vector \( Z_{it} \), whose effects are allowed to vary across generations.

What Explains Housework Hours? Two possible explanations are explored in this analysis of housework hours: the roles of education and children. To isolate the effects of these separately, the following approach is used. First, using the regression model in equation (2), the effect of education is isolated by focusing on a baseline case of an age 35 high school dropout with no children and analyzing the change in housework hours as educational attainment is increased. Figure 6 presents the predicted housework hours by education group. Second, two additional specifications of the regression model in equation (2) are estimated by imposing restrictions on the interactions of children with each gender and with each generation, to disentangle the role of children. Specification (1) assumes that each of these interactions is restricted to zero. Specification (2) relaxes this restriction on the interactions between children and each generation but maintains it on the interactions between children and each gender from Specification (1). Specification (3) is the full regression, as in equation (2). Table 2 presents the regression results for young and old children for these three specifications.

As shown in Figure 6, the most striking features of the role of education are the increase in housework hours for men over generations and the even larger decline in housework hours for women over generations for the baseline case. As educational attainment increases, the housework hours for women decline, thus indicating that part of the reason for the large decline over generations is the significant increase in the educational attainment of women over generations. However, educational attainment has a marginal effect on the increase in housework hours for men over generations. Another notable feature of housework hours is that both married men and married women spend more time doing housework than single men and women.

On the other hand, examining the results from Specification (3), it is clear that young and old children have a significant impact on housework hours, with young children having a higher impact than old children. Compared with the Silent generation, each generation spends less time in housework when young children are in the household. However, for households with old children, the Silent generation and Boomers-1 spend the same amount of time in housework, with a fall in housework hours for each successive generation.

After controlling for children, the estimated coefficients on the generation dummies are all positive and statistically significant. Additionally, more children are associated with a large increase in housework hours for women. This finding implies that a large part of the decline in women’s housework hours over generations is explained by the significant decline in the number of children. However, there is a possible role for home production technological
### Table 2

*The Effects of Young and Old Children on Housework Hours, by Gender and Generation*

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Women</strong></td>
<td>1,097.031***</td>
<td>1,098.919***</td>
<td>860.877***</td>
</tr>
<tr>
<td></td>
<td>(9.9872)</td>
<td>(10.0121)</td>
<td>(10.2262)</td>
</tr>
<tr>
<td><strong>Young children</strong></td>
<td>101.702***</td>
<td>111.478***</td>
<td>25.210***</td>
</tr>
<tr>
<td></td>
<td>(1.6538)</td>
<td>(3.8219)</td>
<td>(3.3746)</td>
</tr>
<tr>
<td><strong>Old children</strong></td>
<td>67.786***</td>
<td>71.717***</td>
<td>8.636***</td>
</tr>
<tr>
<td></td>
<td>(1.2849)</td>
<td>(2.4840)</td>
<td>(2.3159)</td>
</tr>
<tr>
<td><strong>Women × Young children</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>199.714***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.7305)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Women × Old children</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>118.341***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.2853)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Boomers-1</strong></td>
<td>145.832***</td>
<td>146.919***</td>
<td>166.766***</td>
</tr>
<tr>
<td></td>
<td>(13.7775)</td>
<td>(14.1102)</td>
<td>(13.7341)</td>
</tr>
<tr>
<td><strong>Boomers-2</strong></td>
<td>183.643***</td>
<td>194.738***</td>
<td>216.899***</td>
</tr>
<tr>
<td></td>
<td>(16.4685)</td>
<td>(16.8341)</td>
<td>(16.4876)</td>
</tr>
<tr>
<td><strong>GenX</strong></td>
<td>132.431***</td>
<td>172.314***</td>
<td>202.435***</td>
</tr>
<tr>
<td><strong>Millennials</strong></td>
<td>180.901***</td>
<td>201.002***</td>
<td>242.025***</td>
</tr>
<tr>
<td></td>
<td>(31.5501)</td>
<td>(31.6822)</td>
<td>(31.7629)</td>
</tr>
<tr>
<td><strong>Boomers-1 × Young children</strong></td>
<td></td>
<td>−2.234</td>
<td>−15.615***</td>
</tr>
<tr>
<td></td>
<td>(4.3540)</td>
<td>(4.1247)</td>
<td></td>
</tr>
<tr>
<td><strong>Boomers-2 × Young children</strong></td>
<td></td>
<td>−9.170*</td>
<td>−33.499***</td>
</tr>
<tr>
<td></td>
<td>(4.7507)</td>
<td>(4.5933)</td>
<td></td>
</tr>
<tr>
<td><strong>GenX × Young children</strong></td>
<td></td>
<td>−36.152***</td>
<td>−57.261***</td>
</tr>
<tr>
<td></td>
<td>(5.3375)</td>
<td>(5.2006)</td>
<td></td>
</tr>
<tr>
<td><strong>Millennials × Young children</strong></td>
<td></td>
<td>−16.078**</td>
<td>−41.113***</td>
</tr>
<tr>
<td></td>
<td>(6.6208)</td>
<td>(6.4994)</td>
<td></td>
</tr>
<tr>
<td><strong>Boomers-1 × Old children</strong></td>
<td></td>
<td>0.164</td>
<td>−0.782</td>
</tr>
<tr>
<td></td>
<td>(3.0906)</td>
<td>(3.0280)</td>
<td></td>
</tr>
<tr>
<td><strong>Boomers-2 × Old children</strong></td>
<td></td>
<td>−6.713*</td>
<td>−14.092***</td>
</tr>
<tr>
<td></td>
<td>(3.6640)</td>
<td>(3.6290)</td>
<td></td>
</tr>
<tr>
<td><strong>GenX × Old children</strong></td>
<td></td>
<td>−17.966***</td>
<td>−22.054***</td>
</tr>
<tr>
<td></td>
<td>(4.3959)</td>
<td>(4.3530)</td>
<td></td>
</tr>
<tr>
<td><strong>Millennials × Old children</strong></td>
<td></td>
<td>−11.047</td>
<td>−20.751***</td>
</tr>
<tr>
<td></td>
<td>(7.1474)</td>
<td>(7.0130)</td>
<td></td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>257,055</td>
<td>257,055</td>
<td>257,055</td>
</tr>
</tbody>
</table>

**NOTE:** Column (1) includes race; education group; gender; marital status (married or single); number of young children; number of old children; age; age squared; and the interactions between gender and education group, between gender and generation, between race and generation, between generation and education group, and between marriage and generation. Column (2) adds the interactions between generation and young and old children. Column (3) adds the interaction between gender and young and old children.

**SOURCE:** PSID and authors’ calculations.
**Figure 6**

What Predicts Housework Hours? By Sex and Marital Status, Changing Education

<table>
<thead>
<tr>
<th>Sex</th>
<th>Marital Status</th>
<th>Education</th>
<th>Predicted Housework Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. White unmarried men</td>
<td>B. Black unmarried men</td>
<td>C. White married men</td>
<td>D. Black married men</td>
</tr>
<tr>
<td>E. White unmarried women</td>
<td>F. Black unmarried women</td>
<td>G. White married women</td>
<td>H. Black married women</td>
</tr>
</tbody>
</table>

NOTE: Baseline, high school dropouts; HS, high school graduates; SC, those with some college, and College, college graduates (with a four-year degree or higher).

SOURCE: PSID and authors’ calculations.
progress, as the estimated amount of housework hours needed for the number of children decreases in each subsequent generation relative to the Silent generation.

However, neither education nor children appear to contribute significantly to explaining the rise in housework hours for men. This finding implies that most of the effects are primarily cohort effects; that is, the change in household technology and gender roles result in men investing marginally more in housework hours than they used to.

4.2 Leisure

The small empirical literature that has studied the change in leisure hours over time and generations uses several different measures of leisure. Following Aguiar and Hurst (2007) and Aguiar, Hurst, and Karabarbounis (2012), the measure of leisure used in this study is calculated as the residual of annual hours worked and annual housework hours from total time available of 5,840 hours (where eight hours per day are allocated to sleep and personal care). An alternative measure used in the literature (see, e.g., Ramey and Francis, 2009), defines weekly leisure as the residual time after subtracting time spent in nonleisure activities (work, school, housework, commuting, and personal care) from the time available.

Using their measure of leisure, Ramey and Francis (2009) conclude that individuals ages 25 to 54 have the lowest amount of leisure time regardless of gender. Panels A to C of Figure 7, which use the same age-group definitions as Ramey and Francis (2009), show individuals ages 25 to 54 enjoy the least amount of leisure hours, confirming the findings of Ramey and Francis (2009). This pattern holds for both genders. These panels also show that the annual leisure hours for individuals ages 25 to 54 increased between 1968 and 2015 for both genders. Aguiar and Hurst (2007) find similar patterns for the period 1965-2003. In particular, they find that leisure has increased significantly for men and women. However, using their “Leisure Measure 1” definition, which is their narrowest measure, men enjoy more leisure than women. They also show that their result is robust to any of the four leisure measures they propose—the significant rise in leisure persisted.

Additionally, Panels D to O of Figure 7 present the life-cycle profiles for the five generations by gender, marital status, race, and education group. In general, there is a distinct U-shaped curve across profiles, indicating that more hours for leisure are enjoyed during youth and retirement. Millennial men enjoyed higher leisure levels earlier in life than previous generations (Panel D). In contrast, for women, there was a consistent rise with each generation until GenX (Panel E). Leisure levels for Millennial women are similar to those for GenX women. Further, by marital status, married men enjoy less leisure than married women, especially, early in the life cycle (Panels F and G). For the past two generations, this is because married men have worked longer hours than married women, even though the latter have increased their market-work hours by a significant amount. For the earlier generations, housework hours dominated the effect and thus women had lower leisure. Similar patterns are found for White men and White women. Of note, the profile of Black Millennial men shifts upward with respect to the other generations, especially before age 25. Finally, for high school graduates and college graduates, the trend of rising leisure holds for both genders (Panels L to O).
Figure 7
Annual Leisure Hours: Trends and Generations

NOTE: All graphs are restricted to head and spouse. The trend graphs are restricted to ages 18 to 65. Leisure is calculated as the difference of the sum of annual work and household hours from total hours. The generation graphs are also restricted to ages 18 to 65 and are plotted for three age-intervals for smoothing of the trend. For example, age 18 refers to the 18-20 age group. High school, high school graduates, and College, college graduates (with a four-year degree or higher).

SOURCE: PSID and authors’ calculations.
5 FAMILY

5.1 Fertility

A well-known fact in the academic literature is that fertility has declined over time. Panels A to C of Figure 8 present the time trends for completed fertility by marital status (if ever married), race, and education. Completed fertility corresponds to the average number of children born per woman who has reached the end of her childbearing years. Therefore, only women between ages 45 and 50 are taken into account. On average, in 1968, women had 3.04 children and by 2015 that number had decreased to 2.01 (Panel A of Figure 8). Notice the shape of the curve: There is a hump before the 1990s that then flattens out. This article does not find a big difference between women who have ever married and single women because there is little difference between “All” and “Ever-Married Women” on Panel A. Over time there has been convergence by race, although Black women have tended to have a higher average number of children than White women (Panel B). Disaggregating by education group shows a significant decline in births for women with a high school diploma or less, which may be due to a decline in teenage pregnancy. Although completed fertility for women college graduates reached a low of 1.64 in 1999, the trend has been rising since 2012, reaching a high of 1.90 in 2015 (Panel C).

A second measure used in this article to analyze fertility is the age both of men and of women when they had their first child. This is calculated by restricting the sample to individuals ages 35 to 40. Panels D to I of Figure 8 present the trends for age at first birth by gender, marital status, race, and educational group. The main findings are as follows: (i) In general, there was an increase in age at first birth, for both men and women, from 1968 until the late 2000s; the age then started to decline in 2011 and 2013 for men and women, respectively (Panel D). (ii) The same pattern is found for men and women ever married (by age 35); therefore, there is no difference in the age at first birth between never and ever-married individuals (Panel E). (iii) By race, Black men generally have been younger at first birth than White men; although this gap had been closing over the early part of the sample up to 2000, the trend reversed, with the gap widening (Panel F). For women, the age at first birth increased from 1968 to 2013, with a slight decline after 2013 for White women; however, the age at first birth for Black women has remained essentially unchanged over the entire sample period 1968-2015 (Panel G). (iv) By completed education, while the age at first birth differs by education level, the disaggregated trends are similar to the overall trends (Panels H and I).

These results confirm the long-run trends documented in the literature. In particular, Lundberg and Pollak (2007), using the number of births per 1,000 women as their measure of fertility, find that as the postwar baby boom waned, birth rates for women ages 15 to 44 fell from 118 births per 1,000 women in 1960 to 68 births per 1,000 women in 1980. They also found that women have continued to delay childbearing, confirming the findings in this article on age at first birth. In addition, they find that completed fertility was approximately at the replacement rate of 2.1 children per woman by 2005, consistent with the results in this article. More recently, Greenwood (2019) suggests that this long-run decline in fertility is explained by the increase in women’s wages, which raised the opportunity cost of having children. The
Figure 8
Trends in Completed Fertility and Age at First Birth

A. Completed fertility: Marital status

B. Completed fertility: Race

C. Completed fertility: Education

D. Age at first birth: Gender

E. Age at first birth: Marital status

F. Age at first birth: Race (men)

G. Age at first birth: Race (women)

H. Age at first birth: Education (men)

I. Age at first birth: Education (women)

NOTE: All graphs are restricted to head and spouse. The completed fertility graphs are restricted to ages 45 to 50 for completed fertility graphs and the age-at-first-birth graphs to ages 35 to 40. Marital status refers to individuals ever married by age 45 for completed fertility and by age 35 for age at first birth. <=HS, those with education less than or equal to a high school diploma; SC, those with some college; and College, college graduates (with a four-year degree or higher).

SOURCE: PSID and authors’ calculations.
### Table 3
**Completed Fertility by Generation**

<table>
<thead>
<tr>
<th>Generation</th>
<th>Silent</th>
<th>Boomers-1</th>
<th>Boomers-2</th>
<th>GenX*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>2.55 (1.79)</td>
<td>2.12 (1.44)</td>
<td>1.93 (1.41)</td>
<td>1.63 (1.34)</td>
</tr>
<tr>
<td>Observations</td>
<td>11,586</td>
<td>23,376</td>
<td>18,814</td>
<td>325</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Marital status</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Not married</td>
<td>2.15 (1.77)</td>
<td>1.94 (1.51)</td>
<td>2.08 (1.54)</td>
</tr>
<tr>
<td>Observations</td>
<td>2,709</td>
<td>3,507</td>
<td>2,670</td>
</tr>
<tr>
<td>Married</td>
<td>2.72 (1.78)</td>
<td>2.22 (1.39)</td>
<td>2.17 (1.34)</td>
</tr>
<tr>
<td>Observations</td>
<td>6,705</td>
<td>7,620</td>
<td>4,953</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>High school dropouts</td>
<td>3.52 (2.26)</td>
<td>2.83 (1.65)</td>
<td>2.70 (1.55)</td>
</tr>
<tr>
<td>Observations</td>
<td>2,544</td>
<td>3,258</td>
<td>2,574</td>
</tr>
<tr>
<td>High school graduates</td>
<td>2.51 (1.54)</td>
<td>2.21 (1.33)</td>
<td>1.93 (1.37)</td>
</tr>
<tr>
<td>Observations</td>
<td>3,912</td>
<td>8,430</td>
<td>6,578</td>
</tr>
<tr>
<td>Some college</td>
<td>2.25 (1.42)</td>
<td>2.02 (1.36)</td>
<td>1.81 (1.35)</td>
</tr>
<tr>
<td>Observations</td>
<td>2,478</td>
<td>6,390</td>
<td>5,561</td>
</tr>
<tr>
<td>College graduates</td>
<td>1.85 (1.27)</td>
<td>1.65 (1.31)</td>
<td>1.62 (1.24)</td>
</tr>
<tr>
<td>Observations</td>
<td>2,316</td>
<td>4,860</td>
<td>3,755</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Race</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Whites</td>
<td>2.29 (1.53)</td>
<td>2.00 (1.33)</td>
<td>1.80 (1.27)</td>
</tr>
<tr>
<td>Observations</td>
<td>7,668</td>
<td>12,942</td>
<td>10,208</td>
</tr>
<tr>
<td>Blacks</td>
<td>2.86 (1.98)</td>
<td>2.20 (1.47)</td>
<td>1.99 (1.52)</td>
</tr>
<tr>
<td>Observations</td>
<td>2,730</td>
<td>8,232</td>
<td>6,222</td>
</tr>
</tbody>
</table>

**NOTE:** *The generation has right censoring in terms of age. The table is calculated for women ages 45 to 50.
SOURCE: PSID and authors’ calculations.*
The author calculates completed fertility in the year 1800, where the average White woman had 7 children; yet, by 1990, this number had dropped to just 2. Finally, he argues that there was a significant recovery in fertility in the mid-1960s, as seen in Panel A of Figure 8 as well. He explains that advances in medicine led to both younger and older women having more children. Fertility then reverted back to its common trend, and the “baby bust” resumed.

Table 3 presents completed fertility by generation for women ages 45 to 50. There are no results for Millennials, as this generation had not reached ages 45 to 50 by 2015 (when the sample ends). For GenX, there is some right censoring in terms of age; however, the numbers for GenX are still reported in the table. As the table shows, there has been a decreasing trend in completed fertility, with it falling below the replacement rate for Boomers-2. This trend is seen after disaggregating by marital status, education, or race, with the only exception being unmarried women. The gap in completed fertility between high school dropouts and college graduates has narrowed from 1.67 for the Silent generation to 1.08 for Boomers-2. A similar decline is observed for the difference in completed fertility between Blacks and Whites from the Silent (0.57) to the Boomers-2 (0.19) generation.

In order to analyze the fertility behavior of Millennials, Table 4 presents data on parity of births as well as the proportion of births to married women across generations. It is evident that non-marital fertility, as measured by the proportion of children born to unmarried women, has risen for all four Millennial age groups (spanning ages 18 to 35). This finding is in line with the rise in non-marital fertility reported by Lundberg and Pollak (2007), who state that 37 percent of U.S. births were out-of-wedlock in 2005. The parity of births presents the evolution of fertility over the life cycle and is defined as the number of live births to a woman so far. At the youngest age group, Millennials start at a much lower parity than any other generation and remain so for the other three age groups.

### Table 4

<table>
<thead>
<tr>
<th>Age group</th>
<th>Proportion of births to married women</th>
<th>Parity of births</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Silent</td>
<td>Boomers-1</td>
</tr>
<tr>
<td>18-20</td>
<td>71.64</td>
<td>60.55</td>
</tr>
<tr>
<td>21-25</td>
<td>83.47</td>
<td>74.57</td>
</tr>
<tr>
<td>26-30</td>
<td>87.05</td>
<td>80.87</td>
</tr>
<tr>
<td>31-35</td>
<td>88.40</td>
<td>82.77</td>
</tr>
<tr>
<td>36-40</td>
<td>86.12</td>
<td>83.71</td>
</tr>
<tr>
<td>41-45</td>
<td>84.98</td>
<td>83.41</td>
</tr>
</tbody>
</table>

NOTE: Marital status refers to the marital status of a woman at the time of birth.
SOURCE: PSID and authors’ calculations.
5.2 Marriage

This article uses two different measures of marital status: (i) legally married, defined as stated—individuals who are legally married, and (ii) PSID married, defined as those individuals who are either legally married or have a cohabiting partner. The marriage rate corresponds to the proportion of individuals ages 18 to 65 who are legally married. A similar definition is used for PSID marriage rates using PSID married. Figure 9 presents the marriage rates by gender, race, and educational level.

There is an active literature that has documented the decline in marriage rates over time (see Greenwood and Guner, 2008; Lundberg and Pollak, 2015; Lundberg, Pollak, and Stearns, 2016; Santos and Weiss, 2016; and Greenwood, 2019). Panel A of Figure 9 also confirms the decline of the overall marriage rates for the period 1968-2015. Panels B and C of Figure 9 present the marriage rate by race. It shows that Whites marry at a higher rate than Blacks and that this racial gap in marriage rates has widened over time. Panels B and C of Figure 9 also show a larger decline in marriage rates over time for Blacks relative to Whites. For example, in 1968, the gap between White and Black marriage rates was 3.3 percentage points for men; by 2015, this gap had increased to 23.6 percentage points. Similarly, that gap for women rose from 6.4 percentage point to 31.2 percentage points over the same period. Panel D of Figure 9 shows an increase in cohabitation; however, this rate tends to be small: It was 0.6 percent in 1968 and by 2015 had increased to 5.9 percent. This finding is in line with the discussion by Lundberg, Pollak and Stearns (2016), where they point out that over time there has not only been a “retreat from marriage” but also an increase of cohabitation by many “single” Americans. Finally, for both definitions of marriage (legally and PSID), there is a similar pattern by educational level, implying that cohabitation does not have a significant impact on the overall trend of household formation (Panels E to H). While all education groups have seen a decline in marriage rates, the rate has declined more slowly for college graduates. Moreover, there was a crossover in marriage rates for both men and women college graduates in the years 1987 and 1998, respectively; that is, before 1987 and 1998, college graduates married at a lower rate than all other education groups but after 1987 and 1998 the opposite has been true. Greenwood (2019) also pointed out that the decline in marriage rates are greater for the non-college graduates than for the college graduates, while Lundberg and Pollak (2007) found that the marriage-rate trajectories of the more and less educated began to diverge in the mid-1980s.

The decline in marriage rates is clear, yet it raises the question as to what the driving forces are behind the decline. Greenwood (2019) discusses that there exist three possible explanations in the literature. First, as previously mentioned, there has been a rise in wages, which makes a one-person household more affordable. Second, the labor-saving technological progress in the home has led to less need for specialization (Greenwood and Guner, 2008). Finally, there has been a fast drop in the prices of time-saving goods used at home. The authors argue that these three forces reduced the importance of scale economies in household consumption/production; hence, single households are more common in the current time. In addition, Santos and Weiss (2016) propose that because the childcare costs that parents have to incur are fixed and difficult to avoid during tough economic times, people delay marriage when there is instability in the labor market.
**Figure 9**

**Trends in Marriage Rates and Age at First Union**

**A. Married: Gender**

**B. Married: Race (men)**

**C. Married: Race (women)**

**D. Cohabiting**

**E. Married: Education (men)**

**F. Married: Education (women)**

**G. PSID married: Education (men)**

**H. PSID married: Education (women)**

**I. Age at first union: Overall**

**J. Age at first union: College**

**K. Age at first union: Blacks**

**L. Age at first union: Blacks with college**

NOTE: All graphs are restricted to head and spouse and to ages 18 to 65. Married refers to legally married, while PSID married refers to married or cohabiting individuals. In Panels E to H, “All” includes high school dropouts plus the following: HS, high school graduates; SC, those with some college; and College, college graduates (with a four-year degree or higher).

SOURCE: PSID and authors’ calculations.
**Figure 10**

**Marriage Rates by Generation**

NOTE: All graphs are restricted to head and spouse. Married refers to legally married. The generation graphs are also restricted to ages 18 to 65 and are plotted for three age-group intervals for smoothing of the trend. For example, age 18 refers to the 18-20 age group. High school, high school graduates; Some college, those with some college; and College, college graduates (with a four-year college degree or higher).

SOURCE: PSID and authors’ calculations.
To complement this analysis, this article also presents the “age at first union,” which is defined as the age when the individuals got married or started cohabiting with their partner for the first time. Panels I to L of Figure 9 present the trends for this variable by gender, educational group, and race. The age at first union has increased over time for men and women. Moreover, college graduates tend to further delay their first union, possibly due to human capital accumulation. For example, in 2015, men (women) high school graduates had their first union on average at age 24.7 (23), while men (women) college graduates had their first union on average at age 27.4 (26.1). Finally, Panel K shows that over time Black women have delayed their age at first union and, thus, the gap in the age at first union between Black men and Black women has narrowed.

Figure 10 presents the life-cycle trends of marriage rates over generations. It shows a sharp decline in marriage rates over generations as well as a flattening over the life cycle, with Millennials having the lowest marriage rates among all generations at all comparable ages (Panel A). Cohabitation rates have risen over each generation, with Millennials having the highest rates. The rate rose from 2.3 percent (2.8 percent) to 11 percent (12 percent) for men (women) ages 21 to 23 from Boomers-1 to Millennials. In line with Lundberg and Pollak (2007), the less-educated groups (some college and less) have retreated from marriage. This is because the life-cycle trends for each generation are clearly below the previous one, with Millennials being the most distinct. It could be the case the less-educated individuals are ruling themselves out of the marriage market (Lundberg and Pollak, 2007). However, for college graduates, Figure 10 shows a delay in marriage instead of a retreat from marriage, since by ages 30 to 33, Millennials have caught up to the marriage rates of the previous generations. The patterns are similar across Black men and White men; however, Black Millennial women see a catching up of the marriage rate by ages 30 to 33.

Framework for Analyzing Education and Marriage. It is clear from the previous trends and is well documented in the literature (Chiappori, Salanié, and Weiss, 2017, and Gayle and Shephard, 2019, among others) that education plays an important role in the marriage market. This section presents an empirical framework for how an individual chooses a partner of a certain education level or remains single in a frictionless marriage market. The framework is the multinomial logit empirical analog of the equilibrium marriage market model in Choo and Siow (2006), Chiappori, Salanié, and Weiss (2017), and Gayle and Shephard (2019) and is summarized by the log-odd ratio of marrying a partner of a particular type and singleness:

\[
\log \frac{P(m_{it} = j)}{P(m_{it} = 0)} = \alpha_0 + \sum_{r=1}^{4} \alpha_r d_{i,t-r} + \sum_{m=1}^{4} \alpha_m h_{i,t}^m + \delta' Z_{it},
\]

where \(P(m_{it} = j)\) refers to the probability that individual of gender \(i\) chooses a partner of type \(j\) in calendar time \(t\). The index \(j\) indexes the type of partner, where choosing to remain single is the baseline \((j = 0)\). In the empirical implementation, partners are indexed by three levels of educational attainment as follows: high school diploma or less \((j = 1)\), some college \((j = 2)\), and college graduate (with a four-year degree or higher) \((j = 3)\). The regression also includes controls \((Z_{it})\), which are age, age squared, race, education, and number of children. Educational
attainment of the individual is categorized the same as the education attainment of the partners. All other variables are as defined in equation (1). The marriage transition analysis is done separately for each gender and cohort, and the data are restricted to Blacks and Whites between ages 20 and 40, as most marriages take place during these years.

What Predicts Partner Choice? This analysis focuses on the role that education plays in partner choice. Panels A, B, E, and F of Figure 11 plot the observed marriage and education transitions for women ages 30 to 35 by generation. Although the regression is run for the 20–40 age group, we focus on the 30–35 age group for this analysis because most education is completed by then. It is clear from these panels that irrespective of race or education, singlehood has risen over the generations. While this has been a monotonically increasing trend for women with a high school diploma or less (Panel A), White GenX women and Black Boomers-1 women college graduates each had a reduction in singlehood relative to the previous generation (Panels B and F, respectively). Another striking feature to note is the differences in the rates of singlehood across race and education. For White GenX women, while 27.3 percent of those with a high school diploma or less are single, the number shrinks to 16.7 percent for college graduates. The equivalent numbers for Black women are 65 percent and 48.5 percent, respectively, which are significantly larger.

Panels I, J, M, and N of Figure 11 plot the model-predicted marriage and education transitions for women age 35 by generation, using the specification, as defined in equation (3), where the remaining variables (unless stated otherwise) are set to their average values. While the model underpredicts the singlehood rates for White women, specifically for those that are college graduates in the generations beyond Boomers-1, it overpredicts the rate for Black women, specifically for those that are Boomers-1 college graduates. A reason for this could be the lack of observations in these categories, specifically college graduates in the earlier generations.

Table 5

<table>
<thead>
<tr>
<th></th>
<th>Silent</th>
<th>Boomers-1</th>
<th>Boomers-2</th>
<th>GenX</th>
<th>Millennials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of children</td>
<td>2.05</td>
<td>1.70</td>
<td>1.63</td>
<td>1.78</td>
<td>1.58</td>
</tr>
<tr>
<td></td>
<td>(1.44)</td>
<td>(1.24)</td>
<td>(1.26)</td>
<td>(1.32)</td>
<td>(1.33)</td>
</tr>
<tr>
<td>Participation in past four periods</td>
<td>2.66</td>
<td>2.69</td>
<td>2.80</td>
<td>3.19</td>
<td>3.42</td>
</tr>
<tr>
<td></td>
<td>(1.54)</td>
<td>(1.49)</td>
<td>(1.48)</td>
<td>(1.33)</td>
<td>(1.15)</td>
</tr>
<tr>
<td>Hours worked in past four periods / 10</td>
<td>377.62</td>
<td>394.62</td>
<td>443.36</td>
<td>540.73</td>
<td>574.13</td>
</tr>
<tr>
<td></td>
<td>(309.76)</td>
<td>(310.74)</td>
<td>(324.22)</td>
<td>(329.40)</td>
<td>(300.56)</td>
</tr>
<tr>
<td>Observations</td>
<td>12,764</td>
<td>34,804</td>
<td>25,625</td>
<td>16,542</td>
<td>6,582</td>
</tr>
</tbody>
</table>

NOTE: A multinomial logit regression is run where the dependent variable categories are based on the marital status of the individual: Single (unmarried); <=HS, married to a partner with high school diploma or less; SC, married to a partner with some college; and College, married to a partner that is a college graduate (with a four-year degree or higher). The key explanatory variable is on the education of the woman (man) and the following controls are included: age, age squared, race, number of children, employment in the past four periods, and hours worked in the past four periods.

SOURCE: PSID and authors’ calculations.
Figure 11

Predicted Probability of Choosing a Partner, by Educational Attainment

NOTE: All graphs are restricted to head and spouse. A multinomial logit regression is run where the dependent variable categories are based on the marital status of the individual: Single; <=HS, married to a partner with high school diploma or less; SC, married to a partner with some college; and College, a married to partner that is a college graduate (with a four-year degree or higher). The key explanatory variable is on the education of the woman (man) and the following controls are included: age, age squared, race, number of children, employment in the past four periods, and hours worked in the past four periods. S, Silent generation; B-1, Boomers-1; B-2, Boomers-2; X, GenX; and M, Millennials.

SOURCE: PSID and authors’ calculations.
To tease out the assortative mating patterns, the analysis now focuses on women who do get married. Panels C, D, G, and H of Figure 11 show the empirical patterns, while Panels K, L, O, and P show the model-predicted patterns. Assortative mating is defined here as individuals of the same education level marrying each other, for example, college graduates marrying each other. While there is no clear trend in terms of a rise or fall in assortative mating over generations, there is no doubt that assortative mating exists in the data and is predicted by the model as well. Averaging over generations, White women with a high school diploma or less have a 63.5 percent likelihood of assortative mating; for those that are college graduates, the number rises to 69.2 percent. On the other hand, averaging across generations, assortative mating is more pronounced for Black women with a high school diploma or less (74.1 percent), but less pronounced for those that are college graduates (41.8 percent). However, note that singlehood rates are significantly higher for Black women as well. Focusing on the model predictions, note that the model under predicts assortative mating for women with a high school diploma or less and over predicts for those that are college graduates. The overpredictions are more notable for Black women.

5.3 Divorce

Any discussion of marriage without talking about divorce would be incomplete. The divorce rate in this article is defined as the proportion of individuals ages 18 to 65 who are not legally married in the current period but were married in the previous period.

Panels A and B of Figure 12 plot a smoothed version of the divorce rate by fitting a quadratic time trend to the raw data. An inverted-U-shaped trend is seen for “All” for the years 1968 to 2015, with the divorce rate falling post 2000, reaching 1.69 percent in 2015. This trend is directly affected by the declining marriage rate, as discussed in Section 5.2. Lundberg and Pollak (2015) find that the divorce rate peaked in the 1980s. However, they use the number of divorces per thousand married couples and the data are from the U.S. Census and the American Community Survey, which are repeated cross-sections and, therefore, can analyze marriage and divorce patterns only for a point in time. They also argue that the increase from 1960 to 1980 is in part explained by liberalized divorce laws, where unilateral divorce became universal across the country, as well as the decrease in the social and legal costs of exiting a marriage. Disaggregating by education groups, it is clear that college graduates divorce the least relative to the other education categories. Analysis by race shows that Blacks divorce at significantly higher rates than Whites (Panel B).

To complement these findings, a survival analysis is performed by estimating a kernel regression of divorce on years of marriage (or marriage tenure). Three additional restrictions are imposed: First, the survival analysis is conducted at the household level; second, only marriages where the ages upon marriage of both the husband and wife were between 18 and 65 are considered; and third, marriage tenure is plotted to 35 years to circumvent the possibility of the marriage ending due to the death of a spouse. Panels C to J of Figure 12 present the survival analysis by race, education, and generation. Overall, the probability of divorce peaks in the first eight years of marriage and then decreases monotonically (Panel C). This decrease indicates strong duration dependence; that is, the longer the duration of the marriage,
Figure 12
Divorce Rates: Trends and Generations

NOTE: The graphs are restricted to head and spouse for the trend graphs and to households for the kernel regressions of divorce (defined as divorce or death). The generation graphs are also restricted to ages 18 to 65 and are plotted for three age-group intervals for smoothing of the trend. For example, age 18 refers to the 18-20 age group. In Panel A, “All” includes high school dropouts plus the following: HS, high school graduates; SC, those with some college; and Coll, college graduates (with a four-year degree or higher). Prov(Divorce), probability of divorce. In Panel D, “low” refers to completed education of less than or equal to a high school diploma and “high” refers to completed education of some college or above. In Panel E, “low” refers to completed education less than a four-year college degree and “high” refers to completed education of a four-year college degree or higher. Then, for each definition of education, the four groups of couples are low-low, high-low, low-high, and high-high, where the first component corresponds to the education type of the husband and the latter to the wife. Panel I, “Cross-Edu,” combines the high-low and low-high education groups of Panel D since their patterns do not differ.

SOURCE: PSID and authors’ calculations.
Gayle, Odio-Zuniga, Ramakrishnan

the lower the likelihood of divorce. Although this pattern holds by race, Blacks have a significantly higher likelihood of divorcing than Whites. For example, after nine years of marriage, Blacks have a 2.4 percent probability of getting divorced compared with 1.6 percent for Whites.

Education is an important factor in marriage rates, as seen in the previous section, where college graduates delayed marriage but non-college graduates retreated from it. The education of the couples is classified according to two different definitions of education, each with “low” and “high” education types: Education 1, where low refers to completed education of less than or equal to a high school diploma and “high” refers to completed education of some college or above, and Education 2, where “low” refers to completed education less than a four-year college degree and “high” refers to completed education of a four-year college degree or higher. Then, for each definition of education, the four groups of couples are low-low, high-low, low-high, and high-high, where the first component corresponds to the education type of the husband and the latter to the wife. Panels D and E of Figure 12 break out the probability of divorce by the education definitions and types.

Using the Education 1 definition, the assortatively matched pairs (low-low and high-high) have the lowest probability of divorce during the first 10 years of marriage. However, couples where the woman “marries down” (low-educated husband with a high-educated wife) have the highest probability of divorce. For example, at nine years of marriage, the divorce rate for the low-high group is 2.8 percent, which is more than double the rate for the low-low group (1.3 percent). One explanation is that when the wife has more education than the husband, she will have a higher outside option, which can lead to divorce. This finding is in line with Bertrand, Kamenica, and Pan (2015), who argue that households where the wife earns more than the husband tend to have a higher probability of divorce, in part due to gender identity norms such as the husband should be the breadwinner.

Panel E of Figure 12 uses the Education 2 definition. The key difference is that high-high couples have the lowest divorce rates, which is significant in the early years of the marriage. However, low-high couples still stand out due to their significantly higher rates relative to the other education pairs. These patterns confirm the results in Lundberg, Pollak, and Stearns (2016), who find that college-graduate parents have the lowest divorce rates and use marriage as a commitment device to facilitate joint investment in their children.31

Using the generation of the husband as the generation of the household, Panels F to J present the divorce rates by generation. It is important to note that long-term marriage analysis for the younger generations, especially Millennials, is not yet possible. The Silent generation has a very flat profile with minimal duration dependence, indicating that their probability of divorce is stable and does not change significantly as marriage tenure increases (Panel F). For Boomers-1, the divorce rate increased relative to the Silent generation and has more pronounced duration dependence. Although there is not much statistical difference in the probability of divorce for Boomers-2 and later generations, Millennials are the most likely to divorce after five years of marriage. Stevenson and Wolfers (2007) find similar results using the retrospective marriage history from the Survey of Income and Program Participation. They show that the proportion of marriages that end in divorce increases for each successive marriage cohort33 from 1950-59 to 1970-79 and then subsequently falls for the 1980-89 and
1990-99 cohorts. The pattern is similar across Blacks and Whites, although Blacks have higher divorce rates. Therefore, the major changes in divorce rates across generations are primarily coming from Blacks and not necessarily Whites (Panel G).

Finally, Panels H to J of Figure 12 show divorce rates by generation and education type using the Education 1 definition. High-high couples (Panel J) have significantly lower divorce rates than couples where at least one spouse is of low education (Panel I). Thus, most of the changes in divorce rates across generations come from these three couple types: low-low, low-high, and high-low. This finding cannot be viewed in isolation, as the marriage rate for less-educated individuals has changed over time. As discussed in Section 5.2, the singlehood rate is what has changed and not the sorting pattern; that is, fewer less-educated individuals are getting married, implying that the composition of the education types of married couples has been changing over time.

6 CONCLUSION

This article analyzes the changes in family structure, fertility behavior, and the division of labor within the household over generations. Using PSID data, it documents time trends and life-cycle profiles over generations on three aspects—work, family, and leisure. This article provides a first cut on the behavior of Millennials and how they compare with previous generations.

Focusing on work, it finds that the wage-age profile has been shifting down over generations, especially for Millennial men (up to age 33). However, women’s wages have instead stagnated, with Millennial women earning lower wages than Generation X after age 30. Therefore, rising wage inequality has come from men and not women. To understand the decomposition of the gender gap in wages, an Oaxaca-Blinder decomposition is estimated, which decomposes the gender wage gap into explained and unexplained components. For ages 18 to 65, there is a sharp fall in the unexplained component over the generations, from 30.5 percent for the Silent generation to 8.5 percent for Millennials.

The reversal in the gender gap in education is well documented; however, this switch occurred in the Boomers-2 generation, since up to Boomers-1, men were graduating from college at higher rates than White women. Focusing on race, there has also been a significant increase in college graduation rates from Boomers-2 to GenX for both Blacks and Whites, a pattern that has continued to accelerate for Millennials. This is in sharp contrast to the generally stable graduation rate from the Silent generation to Boomers-1. The article also finds a significant reduction in high school dropout rates, which has been most pronounced for Blacks. The percentage of Black high school dropouts fell from 27.5 percent for the Silent generation to 7.1 percent for Millennials. The article also estimates the returns to education over generations, especially given the changing educational attainment and narrowing of gender wage gaps. For both Black men and White men, the returns to college have increased for all generations, with the most significant increase for Black Millennial men. For women, the trend is less clear.

With respect to leisure, the article finds that over generations, while the amount of leisure enjoyed by women has been increasing, the amount of leisure enjoyed by men shows no clear
pattern. However, for both men and women, Millennials enjoy a higher level of leisure than previous generations. This rise in leisure for women is primarily coming from married women. The finding that the amount of leisure enjoyed by men has been stable over generations masks different dynamics in the components of leisure. There has been a significant reduction in hours worked by men, with Millennials working the least. At the same time, there has been an increase in the hours men devote to housework, with Millennial men devoting more hours than men in any previous generation. The opposite is true for women: Their reduction in housework hours over generations has more than offset the increase in hours devoted to market work. More importantly, most of this movement for women has occurred within married couples. To understand what predicts housework hours, the article controls for race, education, gender, marital status, the numbers of young and old children, and age. It finds that young and old children in a household have a significant impact on housework hours, especially young children.

The article finds that marriage (and thus, fertility) has seen many drastic changes over generations. Less-educated individuals have retreated from marriage, especially those that are Millennials. Yet it is not clear whether Millennials would be marrying at an overall rate lower than previous generations, as the composition of education has also changed, with a much higher number of Millennial college graduates. Using a multinomial logit to predict partner choice, no clear trends in assortative matching are seen over generations. Although this finding is in contrast to those of Greenwood and Guner (2008) and Santos and Weiss (2016), it is in line with the those of Gihleb and Lang (2016). With respect to cohabitation, for GenX and Millennials, cohabitation rates are much higher early on in life relative to previous generations; nevertheless, later on, it drops off significantly, indicating cohabitation has been transitory—a pattern not seen for the Silent and Boomer generations.

Any changes in marriage patterns will also affect divorce rates. For the period 1968-2015, divorce rates follow an inverted-U shape, with college graduates divorcing the least. As argued by Lundberg, Pollak, and Stearns (2016), some of the reasons for this trend are the declining marriage rate, the transition to unilateral divorce laws, and the decrease in the social and legal costs of exiting a marriage. The present article also finds that couples where the woman marries down (a low-educated husband with a high-educated wife) have the highest probability of getting a divorce. A possible explanation is due to gender identity norms such as the husband should be the breadwinner, which is in line with the findings of Bertrand, Kamenica, and Pan (2015). Analysis across generations finds that the Silent generation has a very flat divorce profile with minimal duration dependence, whereas the remaining generations have a subsequent increase in divorce as well as more pronounced duration dependence, with Millennials having the highest probability of divorce after five years of marriage. The majority of the changes in divorce across generations primarily come from Blacks and couples where at least one spouse has completed education of less than or equal to a high school diploma.

Finally, the results in the article confirm the significant decline in completed fertility, with the fertility rates of Boomers-2 and GenX falling below the replacement rate. Similar trends exist across races (see Lundberg and Pollak, 2007, and Greenwood, 2019). There has been a steady decline at each age parity of the proportion of births to married women, although this decline has accelerated among Millennials. The article also finds that the completed Millennial fertility rate will most likely be below those of all generations analyzed in this article.
NOTES

1. Millennials are normally defined as a person reaching young adulthood in the early twenty-first century. For the analysis in this article, Millennials are defined as a person born in the birth cohort from 1980 to 1989.

2. See Kong, Ravikumar, and Vandenbroucke (2018) for similar results in the flattening of the age-earnings profile for previous generations.

3. This is similar to the result in Blau and Kahn (2017), who also documented that women are doing better than men.


7. These are unweighted numbers.

8. Married men are employed at a higher rate than single men while married women are employed at a lower rate than single women.

9. Greenwood (2019) argues that the “decline in fertility, improvements in household technologies, advances in obstetric and pediatric medicine have reduced the time off of work that a woman needs to bear and raise children.”

10. For the PSID the data regarding the employment status is not available for ages 10 to 13.


12. The data sources are different: They use Current Population Survey data, whereas this article uses PSID data.

13. Aguiar and Hurst (2007) define total market work as total time spent working in the market sector on main jobs, second jobs, and overtime, including any time spent working at home plus commuting and break times.

14. Prior to 1993, farm income and the labor portion of business income were included in individual income, by construction. Post 1993, these are reported as separate amounts. However, individual income is created by adding up the business and farm income so that it is consistent across years.

15. This includes all variables that vary by individuals but are time invariant, for example, completed education, race, and gender.

16. As this article uses log hourly wages, the difference between the men and women log wages is taken.

17. Total time endowment is set at 365.25 × 24 hours.

18. While Reid et al. (1934) was the first person to introduce the notion of household production, Becker (1965) was the first to formalize it.

19. Ramey and Francis (2009) define home production as the time spent in planning; purchasing goods and services (except medical and personal care services); care of children and adults; general cleaning; care and repair of the house and grounds; preparing and clearing food; and making, mending and laundering of clothing and other household textiles.


21. Core non-market work is defined by Aguiar and Hurst (2007) as “any time spent on meal preparation and cleanup, doing laundry, ironing, dusting, vacuuming, indoor household cleaning, and indoor design and maintenance (including painting and decorating).”

22. This measure of leisure corresponds to Leisure Measure 1 from Aguiar and Hurst (2007), which includes activities related to “entertainment/social activities/relaxing” and “active recreation.”

23. For more details, see Table 1.2 in Aguiar and Hurst (2007).

24. All of this generation had not reached age 50 by the end of the sample; however, conclusions can be drawn based on averages.
The construction of cohabitation is explained in Section 2.

Using the PSID, Schneider, Harknett, and Stimpson (2018) analyze the driving forces of the decline in the rate of entry into first marriage for the 1969-2013 period for different cohorts. They argue that part of the decline is due to a decrease in economic prospects in the labor market and higher probabilities of incarceration for men.

This way of categorizing education is used in this section of the article because there are very few women who are high school dropout in the data. Hence, in order to obtain enough data to analyze marriage transitions, high school dropouts and high school graduates are placed in the same category.

The averages are presented in Table 5.

A similar pattern was found in Pessin (2018).

See Eckstein, Keane, and Lifshitz (2019) as well.

See Pessin (2018) as well.

The data are restricted to the years since marriage, where the number of observations is greater than or equal to 100.

The results in this article are for birth cohort, or generation.

Panel I of Figure 12 combines the high-low and low-high groups since their patterns do not differ. This joint group is called “Cross-Edu” in the graph.

REFERENCES


Reid, Margaret Gilpin. Economics of Household Production. J. Wiley and Sons, 1934.


Monetary Policy and Economic Performance Since the Financial Crisis

Dario Caldara, Etienne Gagnon, Enrique Martínez-García, and Christopher J. Neely

The analysis in this article was presented to the Federal Open Market Committee as background for its discussion of the Federal Reserve’s review of monetary policy strategy, tools, and communication practices. The Committee discussed issues related to the review at five consecutive meetings from July 2019 to January 2020. References to the Federal Open Market Committee’s current framework for monetary policy refer to the framework articulated in the “Statement on Longer-Run Goals and Monetary Policy Strategy” first issued in January 2012 (FOMC, 2021).

We review the macroeconomic performance during the Global Financial Crisis and subsequent economic expansion, as well as the challenges in the pursuit of the Federal Reserve’s dual mandate. We characterize the use of forward guidance and balance sheet policies after the federal funds rate reached the effective lower bound. We also review the evidence on the efficacy of these tools and consider whether policymakers might have used them more forcefully. Finally, we examine the post-crisis experience of other major central banks with these policy tools. (JEL E31, E32, E52, E58)

https://doi.org/10.20955/r.103.425-60

1 INTRODUCTION

In this article, we summarize macroeconomic outcomes during the Global Financial Crisis (GFC) and subsequent economic expansion from the point of view of the Federal Reserve’s dual mandate. Unemployment rose sharply during the crisis and declined steadily thereafter, whereas inflation persistently fell short of the symmetric 2 percent longer-run inflation goal adopted in January 2012. We highlight that departures from mandated goals reflect structural changes—some preceding the GFC, others brought about by the shock of the GFC—that took time to recognize and may have inhibited the policy response. We then review the evolving policy response through the increasingly forceful use of balance sheet

Dario Caldara is the chief of the Trade and Financial Studies Section and an economist and Etienne Gagnon is the assistant director of the Program Direction Section and an economist at the Board of Governors of the Federal Reserve System. Enrique Martinez-Garcia is a senior research economist and policy advisor at the Federal Reserve Bank of Dallas. Christopher J. Neely is a vice president and economist at the Federal Reserve Bank of St. Louis.

© 2021, Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.
policies (BSPs) and forward guidance (FG), and we assess their efficacy, costs, and risks. We consider how perceptions of the benefits and potential costs likely shaped the deployment of these policies. We explore to what extent more-forceful use of these policies within the then-prevailing framework could have mitigated the constraints imposed by the effective lower bound (ELB) on the attainment of policy objectives.

The labor market recovery from the GFC was within the range of historical experience, with monetary policy supporting steady job gains despite impairment of some transmission channels. With respect to price stability, longer-run inflation expectations generally proved well anchored during the crisis, but inflation subsequently ran below 2 percent prior to the onset of the pandemic, and some measures of long-run inflation expectations softened to undesirably low levels.

We contend that, under the policy framework at the time, policymakers could have employed accommodation more forcefully. The fact that policymakers did not judge more-forceful policy to be appropriate, however, especially in the early years of the post-crisis period, was not a shortcoming of the framework but arguably reflected the challenges of conducting monetary policy in an uncertain economic environment using largely untested policy tools. In particular, we describe several structural transformations that were difficult to discern in real time, including a diminished sensitivity of inflation to resource slack, a decline in the natural rate of unemployment, and a decline in the neutral federal funds rate ($r^*$. These transformations limited the scope of federal funds rate policies, weakened the effect of monetary policy on inflation, and revealed the labor market gap to be larger than once thought. Recognition of these changes would have strengthened the case for even greater accommodation.

The evidence shows that the BSPs and FG deployed at the ELB eased financial conditions, supported employment, and helped raise inflation toward 2 percent in a manner roughly consistent with expectations at the time, though much uncertainty remains about the size and persistence of these effects. By contrast, worries that BSPs would disrupt market functioning, induce excessive risk-taking, or fuel inflation did not materialize.

A number of foreign central banks responded to the GFC with strategies and tools that were similar to those used by the Federal Open Market Committee (FOMC). Their experiences highlight the importance of anchoring longer-term inflation expectations, the risk of potentially inconsistent policy actions, and the possibility of pursuing BSPs on a larger scale than the FOMC did.

Our article is organized as follows. In Section 2, we review macroeconomic performance during the GFC and subsequent economic expansion and discuss the challenges in recognizing structural transformations. In Section 3, we explore the extent to which the ELB constrained policymakers’ ability to support the economy. We also review the evidence on the benefits and costs of BSPs and FG as well as the implications of these assessments for the amount of accommodation that policymakers could provide under the policy framework at the time. In Section 4, we draw lessons for the U.S. monetary policy framework from the experience of foreign central banks. Section 5 concludes.
2 U.S. MACROECONOMIC PERFORMANCE IN A CHANGING ECONOMY

The GFC had multiple causes and aggravating factors, including negative foreign shocks, notably the European debt crisis, that hindered the ensuing recovery. Thus, in assessing the effectiveness of the policy measures taken, the question is not whether economic performance was unsatisfying—it clearly was—but rather what lessons the episode taught about the uses and risks of monetary policy tools. With the benefit of hindsight, we judge that the policies deployed achieved mixed success.

2.1 U.S. Macroeconomic Performance in the Aftermath of the Global Financial Crisis

Figure 1 shows that the GFC led to an acute rise in the unemployment rate and a marked step-down in inflation in the fall of 2008. The unemployment rate peaked at 10 percent in 2009, 5 percentage points above the median longer-run value in the Summary of Economic Projections (SEP).1 Sharp drops in energy and food prices dragged headline PCE (personal consumption expenditures) inflation well below the median SEP longer-run estimate of 2 percent.

In the recovery phase, the economy absorbed the recessionary labor market slack at a pace within the range experienced in the past few recoveries. In particular, the unemployment rate declined 0.75 percentage points per year, on average, from its peak in late 2009 until it reached estimates of its longer-run level around 2015. This pace was faster than the corresponding averages in the previous two labor market recoveries, at about 0.50 percentage points per year, but slower than the average in the recovery from the 1981-82 recession, at about 1 percentage point per year.2 Whether the labor market could have recovered faster—say, as fast as during the early 1980s—is unclear because the crisis probably impaired some of the transmission channels of monetary policy. In any case, limitations of, and lags in, the transmission of monetary policy would have precluded the economy from quickly and fully absorbing the labor market slack created by the GFC. The unemployment rate dropped to its lowest level in mid-2018 and remained there until early 2020, with labor force participation moving above its trend.3

Figure 2 shows that longer-run inflation expectations were stable during the GFC and the early years of the recovery. Survey-based measures—such as the University of Michigan Surveys of Consumers (Michigan Surveys) median inflation projections for the next 5-to-10 years, and the Survey of Professional Forecasters (SPF) median 6-to-10-years ahead—remained near pre-GFC levels. Although the measure of inflation compensation based on Treasury Inflation-Protected Securities slid during the depths of the crisis, it quickly retraced its losses at the end of the recession. The anchoring may have helped maintain inflation closer to the 2 percent target than historical experience would suggest, given ample resource slack.4 It also helped support real activity and employment because reductions in nominal interest rates passed through, almost one for one, to lower expected real interest rates.

On the negative side, real activity and the productive capacity of the economy grew modestly for many years. Moreover, PCE inflation ran below 2 percent for most of the decade after the GFC, raising concerns that longer-run inflation expectations could have become unanchored.
Figure 1

Macroeconomic Outcomes

A. Unemployment rate

B. PCE inflation

NOTE: Headline and core PCE inflation are shown on a 12-month basis and the unemployment rate on a monthly basis. The gray bars indicate a recession as determined by the National Bureau of Economic Research.

SOURCE: Federal Reserve Board; FRED®, Federal Reserve Bank of St. Louis.
or anchored at too low a level. Some survey-based measures of longer-run inflation expectations (such as the Michigan Surveys measure shown in Figure 2) and measures of inflation compensation ran below their pre-GFC trend and possibly below levels consistent with the 2 percent goal. In mid-2014, the Board of Governors of the Federal Reserve System’s staff Tealbook projection became conditioned on the explicit assumption that “underlying inflation”—defined as the level of PCE inflation that would prevail in the absence of slack or other shocks—was below 2 percent.5 In short, longer-run inflation expectations appeared to be lower than the FOMC’s target. Letting inflation expectations remain below target can hamper achievement of the dual mandate. As we discuss later, the Bank of Japan (BOJ) struggled to raise inflation expectations, which fueled weak inflation and exposed the economy to adverse shocks.

2.2 Did the U.S. Economy Behave as in the Past?

A number of structural transformations occurred over the past decade that could only be recognized over time.

2.2.1 Fall in the Natural Rate of Unemployment. The paucity of price and wage pressures as slack disappeared suggests that the labor market had more room to run than previously thought. From 2015 to 2019, the median SEP value for the unemployment rate in the longer...
Figure 3
Long-Run Estimates of Real Neutral Interest Rates

NOTE: All estimates are one sided, with the exception of Del Negro et al. (2019), which is two sided. The statistics in Panel A are based on eight time-series models maintained by Federal Reserve System staff. The gray bars indicate recessions as defined by the National Bureau of Economic Research. Mean Blue Chip (6-to-10-year) values are deflated by corresponding values for the GDP deflator.

SOURCE: Federal Reserve Board; National Bureau of Economic Research; and Wolters Kluwer Legal and Regulatory Solutions U.S., Blue Chip Economic Indicators.
Figure 4
Forecast Revisions

A. Unemployment rate
Percent

B. Real GDP growth
Four-quarter percent change

C. Headline inflation
Four-quarter percent change

D. Federal funds rate
Percent

NOTE: SEP medians correspond to projections made in the first quarter of each year. Where unavailable, we approximate SEP medians with the midpoints of central tendencies. In January 2012, the FOMC established a longer-term inflation goal and began reporting participants’ policy rate assumptions. Market participants’ values are from the Survey of Professional Forecasters for the unemployment rate and PCE inflation, the Blue Chip Economic Indicators for real GDP growth, and the Survey of Primary Dealers for the federal funds rate.

SOURCE: FRED®, Federal Reserve Bank of St. Louis; Federal Reserve Board; Survey of Professional Forecasters, Federal Reserve Bank of Philadelphia; Survey of Primary Dealers, Federal Reserve Bank of New York; and Blue Chip Economic Indicators, Wolters Kluwer Legal and Regulatory Solutions U.S.
run fell from 5.5 percent to 4.2 percent (see Figure 1). Factors such as population aging, rising educational attainment, or other aspects of human capital formation may have reduced the natural rate.\textsuperscript{6}

2.2.2 Step-Down in the Trend Rate of Productivity Growth. Real output per hour in the business sector grew a little less than 1.25 percent annually during the post-GFC economic expansion, half its pace during the previous two economic expansions.\textsuperscript{7}

2.2.3 Decline in $r^\circ$. As Figure 3 shows, time-series estimates of $r^\circ$, both in the United States and abroad, declined notably from their pre-GFC levels. The median SEP estimate fell from 2.25 percent in the first quarter of 2012 (when this information was first gathered) to only 0.5 percent during the second half of 2019. A decline in $r^\circ$ could reflect factors such as an aging population, the step-down in the pace of productivity growth, and lower risk tolerance.\textsuperscript{8}

2.2.4 Diminished Sensitivity of Inflation to Resource Slack. Inflation has become less sensitive to contemporaneous movements in domestic resource slack and less persistent, so that a given movement in resource slack today will result in a smaller cumulative price response than previously. A decline in sensitivity need not imply a structural change: A monetary policy that stabilizes inflation weakens the correlation between inflation and resource slack.\textsuperscript{9} However, policymakers’ continued difficulties in raising inflation to 2 percent despite extraordinary policy actions suggest that structural change partly explains the diminished sensitivity.

Identifying the transformations in real time is an inherently challenging task, and there remains substantial uncertainty about these phenomena and their evolution. Policymakers and market participants learned only slowly about these transformations through their forecast errors. Figure 4 illustrates that, during the economic recovery, policymakers and market participants systematically underpredicted the speed at which the unemployment rate fell and overpredicted real gross domestic product (GDP) growth—that is, their projections implied overly optimistic views of labor productivity growth.\textsuperscript{10} It also shows that market participants were repeatedly disappointed by the failure of inflation to rise to 2 percent over the medium term as the labor market tightened. By contrast, FOMC participants generally saw medium-term inflation falling short of 2 percent under appropriate policy during much of the recovery, including for a few years after they adopted 2 percent as their longer-run goal.\textsuperscript{11}

Structural transformations can call for changes in the conduct of monetary policy to achieve the dual mandate. Earlier recognition of these structural transformations might have strengthened the case for more accommodative policies. Consistent with this conjecture, Figure 4 shows that market participants and policymakers repeatedly deferred the projected liftoff date, as they concluded that the labor market had greater room to run and that $r^\circ$ had fallen more than they had previously assumed.
3 DEPLOYMENT OF THE FEDERAL OPEN MARKET COMMITTEE’S POLICY TOOLS

The rapid worsening of the economic outlook in the fall of 2008 led the FOMC to slash the target for the federal funds rate to a range of 0 to 25 basis points. The simple rules in Yellen (2017) prescribed lowering the policy rate to between –1.5 and –9 percent during the GFC. The FOMC judged using a cost-benefit analysis, including practical and legal considerations, that negative deposit rates were unappealing. Accordingly, the FOMC used BSPs and FG to provide additional monetary stimulus. The Federal Reserve’s BSPs comprised three large-scale asset purchase (LSAP) programs (henceforth LSAP1, LSAP2, and LSAP3) and a maturity extension program. These programs focused on purchases of long-term Treasury bonds, agency mortgage-backed securities (MBS), or both, with gross purchases totaling $4.65 trillion. The FOMC also attempted to manage expectations about future policy with FG. From December 2008 to the summer of 2011, FG was of a qualitative nature, conveying the FOMC’s anticipation that short-term interest rates would remain low “for some time” or “for an extended period.” In August 2011, the FOMC switched to a calendar-based approach that emphasized its expectation that the funds rate would remain exceptionally low at least until some preannounced date. Then, in December 2012, the FOMC began using a threshold-based approach that signaled a low policy rate for at least as long as unemployment and inflation stayed above or below preannounced values.

3.1 Did Balance Sheet Policies and Forward Guidance Lead to Better Outcomes?

There is broad consensus that BSPs and FG helped at least partially overcome the ELB constraint. For example, Eberly, Stock, and Wright (2020) argue that BSPs and FG reduced the unemployment rate by as much as 2 percentage points and raised inflation a few tenths of 1 percentage point in the first years of the recovery. That said, researchers debate the effectiveness of BSPs and FG. Some observers argue that BSPs and FG fully substituted for the shortfall, so that the ELB did not really constrain the response to the GFC, while other observers see alternative policy tools as having had little, if any, positive macroeconomic effects (see our review of estimates below). The majority view is that BSPs and FG made up for some, though not all, of the shortfall.

3.2 Were the Efficacy and Costs of Balance Sheet Policies and Forward Guidance as Expected?

In memos sent to the FOMC in late 2008, the Federal Reserve Board’s staff judged that BSPs would likely support the economy by reducing borrowing costs for the private sector and by exerting downward pressure on the exchange value of the dollar. The staff also noted that only very large purchases would likely have the desired effects and that FG could lessen policy uncertainty and reduce longer-term interest rates. However, the paucity of historical precedents and exceptionally stressed financial conditions left great uncertainty about these judgments. FOMC participants expressed similar judgments that BSPs would likely have positive effects, in addition to raising several concerns (to be discussed). Our review of the
evidence on the financial and macroeconomic effects of BSPs and FG suggests that early estimates of the likely effects were in the right ballpark and that adverse side effects have been less severe than feared.

3.2.1 Financial Effects. Researchers have most often evaluated the financial effects of BSPs and FG on asset prices with event studies, to exploit the rapid reaction of asset prices to shocks, such as unexpected announcements. Market participants were surprised by the advent of LSAP1, and event studies around LSAP1 communications show large, immediate financial effects.18

Depending on the study, LSAP1 reduced the 10-year Treasury yield by 50 to 91 basis points. Table 1 presents results from Gagnon et al. (2011), showing that LSAP1 announcements reduced 10-year Treasury yields by 55 to 91 basis points, depending on the event set. Corporate yields fell even more. The magnitude of the effects is sensitive to the details of the specification but is typically large.19 Larger event sets tend to imply smaller policy effects. Using alternative model-based approaches, Ihrig et al. (2018) assess that, altogether, past BSPs had reduced the 10-year term premium 100 basis points by early 2015. Other studies indicate that these declines stimulated bond issuance.20 Appendix D summarizes the key findings from selected studies of financial market effects of unconventional monetary policy.

Panels B and C of Table 1 show that LSAP1 announcements also reduced 10-year government yields in the advanced foreign economies (AFEs) by more than 40 basis points on average and depreciated the U.S. dollar relative to AFE currencies by almost 6 percent. BSP announcements also reduced market estimates of tail risk but only modestly raised equity prices (Wright, 2012).

A number of studies have found that the type of assets purchased mattered: MBS purchases particularly pushed down mortgage rates and had a greater effect on lending for banks that owned greater amounts of MBS.21 Table 2 describes results from selected studies of effects of U.S. BSPs and FG on financial markets.

BSPs affect yields through several channels. Signaling reduces expected future interest rates, while duration-risk effects reduce term premiums by lowering the quantity of duration risk held by the public. A third channel, “local supply effects,” reduces the yields of bonds with durations similar to those of bonds that were actually purchased. Studies disagree about the relative importance of these channels, but there is evidence that all three had economically significant effects.

Regarding FG, the qualitative guidance provided by the FOMC until mid-2011 did not induce market participants to expect a much easier path of monetary policy: As Panel D of Figure 4 makes clear, private agents continued to expect the federal funds rate to rise after four quarters. The introduction of calendar-based FG in August 2011, when the FOMC conveyed its expectations that the policy rate would stay near zero “at least through mid-2013” in its post-meeting statement (FOMC, 2011), led market participants to push back the expected liftoff from three quarters to seven quarters. Yields on 1- and 2-year Treasury securities became less responsive to economic news.22 That said, despite these developments, Del Negro, Giannoni, and Patterson (2012) argue that the August 2011 date-based FG failed to be stimulative because it was interpreted as predicting low future GDP growth, in contrast to the September
Table 1
The Effect of LSAP1 Events on U.S. and Foreign Yields and on the Foreign Exchange Value of the Dollar

<table>
<thead>
<tr>
<th></th>
<th>Baseline events</th>
<th>p-value</th>
<th>All FOMC events</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Effect on U.S. yields (basis points)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-year U.S. Treasury</td>
<td>−34</td>
<td></td>
<td>−1</td>
<td></td>
</tr>
<tr>
<td>10-year U.S. Treasury</td>
<td>−91</td>
<td></td>
<td>−55</td>
<td></td>
</tr>
<tr>
<td>10-year agency</td>
<td>−156</td>
<td></td>
<td>−134</td>
<td></td>
</tr>
<tr>
<td>Agency MBS</td>
<td>−113</td>
<td></td>
<td>−114</td>
<td></td>
</tr>
<tr>
<td>10-year term premium</td>
<td>−71</td>
<td></td>
<td>−47</td>
<td></td>
</tr>
<tr>
<td>10-year swap</td>
<td>−101</td>
<td></td>
<td>−75</td>
<td></td>
</tr>
<tr>
<td>Baa index</td>
<td>−67</td>
<td></td>
<td>−72</td>
<td></td>
</tr>
<tr>
<td><strong>B. Effect on foreign yields (basis points)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia (10 year)</td>
<td>−63</td>
<td>(0.01)</td>
<td>−26</td>
<td>(0.49)</td>
</tr>
<tr>
<td>Canada (10 year)</td>
<td>−50</td>
<td>(0.00)</td>
<td>−47</td>
<td>(0.06)</td>
</tr>
<tr>
<td>Germany (9 to 10 year)</td>
<td>−39</td>
<td>(0.01)</td>
<td>−5</td>
<td>(0.84)</td>
</tr>
<tr>
<td>Japan (10 year)</td>
<td>−18</td>
<td>(0.04)</td>
<td>−9</td>
<td>(0.54)</td>
</tr>
<tr>
<td>United Kingdom (10 year)</td>
<td>−43</td>
<td>(0.02)</td>
<td>−21</td>
<td>(0.46)</td>
</tr>
<tr>
<td>Average effect</td>
<td>−42.6</td>
<td>(0.00)</td>
<td>−21.6</td>
<td>(0.23)</td>
</tr>
<tr>
<td><strong>C. Effect on exchange rate (percent)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUD/USD</td>
<td>−6.16</td>
<td>(0.01)</td>
<td>−13.00</td>
<td>(0.00)</td>
</tr>
<tr>
<td>CAD/USD</td>
<td>−7.76</td>
<td>(0.00)</td>
<td>−9.74</td>
<td>(0.01)</td>
</tr>
<tr>
<td>EUR/USD</td>
<td>−6.70</td>
<td>(0.01)</td>
<td>−5.20</td>
<td>(0.19)</td>
</tr>
<tr>
<td>JPY/USD</td>
<td>−3.54</td>
<td>(0.12)</td>
<td>−9.68</td>
<td>(0.01)</td>
</tr>
<tr>
<td>GBP/USD</td>
<td>−5.98</td>
<td>(0.00)</td>
<td>−10.15</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Average effect</td>
<td>−5.73</td>
<td>(0.12)</td>
<td>−13.11</td>
<td>(0.03)</td>
</tr>
</tbody>
</table>

NOTE: Panel A is excerpted from Table 1 in Gagnon et al. (2011). It shows sums of 1-day nominal U.S. yield changes, in basis points, for two event sets: a "baseline events" set of eight LSAP1 news events (11/25/2008, 12/1/2008, 12/16/2008, 1/28/2009, 3/18/2009, 8/12/2009, 9/23/2009, and 11/4/2009) and an "all FOMC events" set adding all FOMC meetings and minutes releases from November 2008 through January 2010. Panels B and C are excerpted from Tables 2 and 3, respectively, in Neely (2015). They show 1-day nominal foreign long-term yield changes and 1-day exchange rate (foreign exchange per U.S. dollar [USD]) changes for the same event sets. AUD/USD, CAD/USD, EUR/USD, JPY/USD, and GBP/USD, respectively, denote the exchange rates of Australia, Canada, the euro area, Japan, and the United Kingdom against the USD. The p-values show the proportions of 8-day or 13-day changes from July 2007 through January 2010 that were larger in absolute value than the actual change in the corresponding 8-day or 13-day event windows.

SOURCE: Authors’ calculations.
2012 FOMC statement, which was interpreted as a commitment to prolonged policy accommodation. The rollout of threshold-based FG in December 2012 elicited little, if any, market reaction, consistent with the FOMC’s indication that the new FG was in keeping with the date-based FG it replaced.

U.S. Treasury debt issuance may have lessened the macroeconomic effects of the Federal Reserve’s BSP and FG policies over time. Greenwood et al. (2014) compare the effect of the Treasury’s maturity extensions on 10-year yields with those of the Federal Reserve’s BSP and FG programs, using data from the end of 2007 to mid-2014. During that period, the Treasury increased the average duration of publicly held debt from 3.9 years to 4.6 years, which is consistent with its desire to reduce refinancing and rollover risk when the debt-to-GDP ratio rises. That increase in the duration of the publicly held debt offset the Federal Reserve’s efforts to reduce yields. Greenwood et al. (2014) calculate that the Treasury’s duration increase raised 10-year yields by about 48 basis points, which offset about 35 percent of the Federal Reserve’s BSPs and FG.

23 Greenwood et al. (2014) further argue that the Treasury and Federal Reserve should coordinate their efforts when the economy is at the ELB.

The results on the effects of BSPs and FG on financial markets come with several caveats. First, although monetary policy announcements can have immediate, sometimes large effects, calculating the effects of a whole program is hazardous because policy announcements are generally partially expected or generate expectations of further actions. Such leakage can render the market reaction—as judged by event studies—an inaccurate guide to the actual policy effect. Second, announcements on LSAPs probably contain signals about future short rates; hence, disentangling the effects of BSPs and FG is difficult. 24 Third, event studies estimate the reaction of asset prices in a narrow window around policy announcements, typically one day, but are agnostic about the persistence of the effects. Looking over a longer horizon, Wright (2012) finds relatively short-lived effects of BSPs and FG, while Swanson (2021) finds they are very persistent. Neely (2020) criticizes Wright’s (2012) model, however, arguing that it is unreliable. Finally, event studies generally rely on few key announcements, making the estimates sensitive to individual observations and obscuring statistical significance (Kuttner, 2018).

### Table 2

<table>
<thead>
<tr>
<th>Study</th>
<th>Key finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walentin (2014)</td>
<td>LSAP1 increased consumption and GDP by 3.2 percent and 3.8 percent at peak, respectively, largely by increasing residential investment.</td>
</tr>
<tr>
<td>Baumeister and Benati (2013)</td>
<td>LSAP1 prevented inflation dropping by 1 percentage point and output growth reaching a trough of −10 percent.</td>
</tr>
<tr>
<td>Wu and Xia (2016)</td>
<td>Without the Federal Reserve’s BSPs and FG, the unemployment rate would have been 1 percentage point higher at peak than actually observed. The macro effects of these policies are modest but nontrivial.</td>
</tr>
<tr>
<td>Swanson (2021)</td>
<td>FG, asset purchases and policy rate adjustments all have persistent effects on different parts of the yield curve.</td>
</tr>
</tbody>
</table>
3.2.2 Macroeconomic Effects. In contrast to the many empirical estimates of the effects of BSPs on financial conditions, relatively fewer studies quantify the effects on economic activity and inflation. The available macro studies find support for the notion that BSPs have important macroeconomic effects, but there is much uncertainty about their magnitude. Differences in estimates can be ascribed mainly to alternative modeling frameworks and assumptions about the channels through which unconventional policies stimulate the economy. Table 2 reports results from selected studies of the effects of U.S. BSPs and FG on the macroeconomy.

There are three broad classes of models used to estimate the effects of BSPs on output and inflation: dynamic stochastic general equilibrium (DSGE) models, structural vector auto-regressions (SVARs), and large-scale “semi-structural” models such as the Federal Reserve’s FRB/US model. Among DSGE models, Gertler and Karadi (2013) find that the peak effect of a balance sheet intervention along the lines of LSAP2 is about an additional 1 percent on the level of real output. By contrast, Chen, Curdia, and Ferrero (2012) find LSAP2 had only modest effects on output, a 0.1 percent increase. Using a SVAR, Baumeister and Benati (2013) find that the level of GDP was about 3 percent higher at its peak than it would have been absent LSAP1, while the inflation rate was about 1 percentage point higher. Weale and Wieladek (2016) find similar effects using different identifying assumptions. Engen, Laubach, and Reifschneider (2015) use the FRB/US model and find that the collective effect of all of the Federal Reserve’s asset purchase programs and FG subtracted 1.2 percentage points from the unemployment rate at its peak in early 2015 and would have had a peak effect of raising inflation by 0.5 percentage points in 2016.

3.2.3 Costs and Risks. Policymakers and analysts expressed a range of concerns regarding BSPs and FG during the debate on their use. Some worried that a huge increase in the monetary base would lead to an inflation breakout. Others worried that plentiful bank reserves might make it difficult to raise the funds target when that eventually would become necessary. Another concern was that unusual accommodation could create incentives for excessive risk-taking—that is, reaching for yield—and so undermine financial stability. Adverse effects of large purchases on market functioning were also feared. For example, if the Federal Reserve were to buy most or all Treasury issues, liquidity in that market would be adversely affected. Yet another concern was that BSPs might permanently ratchet up the size of the Federal Reserve’s balance sheet if subsequent economic expansions did not provide enough time for normalization. With respect to FG, FOMC participants extensively discussed the risk that their communications might be misinterpreted as unconditional commitments, might not convey the complexities of the economy and the policy process, or might downplay the data-dependent nature of their policy communications.25

If we look back, many of those potential risks raised by policymakers did not materialize. Worries that inflation might run above 2 percent, or that longer-run inflation expectations might become unmoored, proved to be unwarranted. Elevated reserves did not prevent raising the policy rate when it was deemed appropriate. Thanks to the Federal Reserve’s ability to pay interest on excess reserves, and the creation of an overnight reverse repurchase agreement facility, the FOMC successfully raised the federal funds rate from its ELB in a context of abundant reserves starting in December 2015. Furthermore, BSPs supported market function-
ing rather than impaired it. For example, Federal Reserve purchases of agency MBS reduced agency spreads, and LSAP announcements trimmed corporate credit risks.\textsuperscript{26} Moreover, the FOMC reduced its balance sheet largely uneventfully through runoffs starting in October 2017.\textsuperscript{22} While the “taper tantrum” episode illustrates that communicating with the public can be challenging, it is not clear that the Committee’s FG created much confusion. Evidence accumulated since the crisis indicates that the costs and risks of the BSP and FG actions were probably overstated.

3.3 Could the Committee Have Used Existing Policy Tools to Support More Stimulative Policy Under the Prevailing Framework?

Overall, the evidence suggests that the Committee could have provided greater accommodation through BSPs and FG under its framework at the time. Specifically, the Federal Reserve’s balance sheet peaked at 25 percent of GDP, a ratio lower than the European Central Bank (ECB), BOJ, and Bank of England (BOE) achieved through their LSAP programs.

Moreover, the FOMC made only moderate use of FG during the recovery, mostly to clarify the path of monetary policy. For many of the years during which the Committee employed FG, most FOMC participants projected that inflation would fall short of 2 percent and that the unemployment rate would exceed its longer-run level over the medium term. Thus, as with BSPs, FG could have been used without creating a conflict between the legs of the dual mandate. In sum, our view is that the Committee could have done more under the prevailing framework—had it chosen to do so—though further policy action would have been outside of historical experience and thus subject to considerable uncertainty regarding efficacy and risks.

We note, however, that there were limits to the extent that policymakers could have used BSPs and FG under its framework. The Federal Reserve’s review of its monetary policy framework takes as given that, under the prevailing framework, policymakers use their tools to achieve the dual mandate, but they never seek to deliberately overshoot or undershoot the longer-run inflation objective, which the framework did not clarify.\textsuperscript{28} Alternative “makeup” strategies would attempt to improve near-term macroeconomic conditions by making greater use of BSPs and FG, including overshooting the objective. With the exception of the BOJ’s (so far unsuccessful) attempt at lifting inflation above its inflation goal, the ability to use BSPs, FG, and other tools to engineer a substantial easing of financial conditions and create expectations of future inflation running above the stated goal is largely untested.

3.4 How Much Benefit Would Additional Accommodation Have Provided?

Next, we discuss several counterfactual model simulations that support the argument that the FOMC might have improved inflation and employment outcomes by pursuing more accommodative policies during the recovery.

By some metrics, the Committee patiently deferred policy normalization. Figure 5 shows the results from three counterfactual simulations of the FRB/US model in which policymakers followed three Taylor-type rules, starting from 2012:Q1. The simulations suggest that following such rules would have produced worse economic outcomes. The well-known Taylor (1993) rule called for raising the federal funds rate in early 2012, which would have delayed the return
Figure 5
Counterfactual Simple Rule Policies and Outcomes

A. Nominal federal funds rate

B. Unemployment rate

C. Real 10-year Treasury yield

D. PCE Inflation four-quarter change

NOTE: We simulate the FRB/US model under the assumptions that agents form VAR-based expectations and that the intercepts and gaps in each policy rule are consistent with the median longer-run projections of FOMC participants over time. See Appendix A for details.

SOURCE: Authors’ calculations.
of the unemployment rate to its longer-run level by several years and led to a more pronounced undershooting of the 2 percent inflation goal. Adherence to the balanced-approach rule or its inertial version would have similarly delayed achievement of the dual mandate relative to the strategy pursued by the FOMC, though not by as much as adhering to the Taylor (1993) rule.29

Given the realized evolution of the U.S. economy, other model simulations suggest that the provision of even more policy accommodation during the recovery may have improved outcomes. Eberly, Stock, and Wright (2020) use a SVAR model to compute outcomes under alternative historical paths for the federal funds rate and the 10-year term spread. They treat this spread as a policy variable that captures the combined effects of BSPs and FG on the slope of the term structure of interest rates. Their simulations suggest that policies that would have flattened the yield curve an extra 1 percentage point from late 2008 to late 2013 would have shaved a bit over 1 percentage point off the unemployment rate. However, these policies may not have pushed inflation sustainably to the 2 percent target. FRB/US simulations suggest that very large asset purchases would have been required to raise medium-term inflation even modestly.30

4 THE INTERNATIONAL EXPERIENCE: CAUTIONARY TALES FROM EUROPE AND JAPAN

4.1 Macroeconomic Performance in Europe and Japan in the Aftermath of the Global Financial Crisis

Figure 6 shows that, during the GFC, the euro area and Japanese economies suffered large contractions in economic activity and increases in unemployment, with euro area labor markets deteriorating further during the European sovereign debt crisis.

Unemployment rates eventually fell to near or below pre-GFC levels in these economies. Despite tightening labor markets, the ECB and, particularly, the BOJ have struggled to raise inflation to their targets on a sustained basis. In Japan, core inflation (excluding food and energy) averaged only 0.7 percent from the adoption of an explicit inflation target in January 2013 until mid-2019.31 In the euro area, core inflation averaged only 1.2 percent from May 2009 until mid-2019—a level short of the ECB’s mandate at the time of maintaining inflation below, but close to, 2 percent over the medium term. Longer-run Japanese inflation expectations had been well below the target level and even retraced the gains registered following the adoption of the inflation target. Euro area survey-based measures of longer-run inflation expectations were more consistent with the stated objective, but market-based inflation compensation softened to about 1.2 percent by mid-2019. By contrast, U.K. inflation averaged about 2 percent from the GFC until mid-2019, and longer-run inflation expectations were near pre-GFC levels in mid-2019. Temporary import price pressures from sterling depreciation supported U.K. inflation.

The BOJ, BOE, and ECB took substantial policy steps to raise inflation to mandated levels and to support employment and economic activity. Figure 7 illustrates that the ECB cut its policy rate from 3.25 percent to 1 percent in response to the GFC and to 0.25 percent in 2014...
Figure 6
Macroeconomic Performance of Japan, the Euro Area, and the United Kingdom

A. Real GDP growth and unemployment rate
Four-quarter percent change
Percent

B. Inflation
12-Month percent change

C. Long-run inflation expectations
Percent

D. Real GDP growth and unemployment rate
Four-quarter percent change
Percent

E. Inflation
12-Month percent change

F. Long-run inflation expectations
Percent

G. Real GDP growth and unemployment rate
Four-quarter percent change
Percent

H. Inflation
12-Month percent change

I. Long-run inflation expectations
Percent

NOTE: The BOJ implemented an inflation target of 2 percent in January 2013, replacing its explicit “positive range of 2 percent or lower” in place since February 2012 (Bank of Japan, 2012). The ECB’s inflation target was “close to but below” 2 percent over the medium term, which we approximate with a black line at 1.85 percent. The headline inflation measure plotted is the consumer price index (CPI) for Japan and the harmonized index of consumer prices (HICP) for the euro area and the United Kingdom. The core inflation measure plotted is the CPI excluding fresh food and energy (with and without a staff adjustment for consumption tax changes) for Japan; the HICP excluding energy and unprocessed food for the euro area; and the HICP excluding food, energy, tobacco, and alcohol for the United Kingdom. Consensus Forecasts are for the expected CPI percent change over previous year, 6 to 10 years ahead. 5yr/5yr is the market-based long-run inflation expectations of the 5-year, 5-year-forward implied inflation rate from Bloomberg. Shaded bars indicate recessions based on the chronology for each economy.

SOURCE: Bloomberg, Consensus Economics, Economic Cycle Research Institute, and Haver Analytics.
following the deepening of the European sovereign debt crisis. The BOJ entered the GFC with its policy rate very near its ELB. With policy rates near zero, these central banks deployed BSPs, FG, and other tools to provide additional policy accommodation.33 The ECB expanded its balance sheet from about 13 percent of GDP in 2007 to 40 percent of GDP in 2018. The BOJ has boosted its balance sheet even more, from 21 percent of GDP in 2007 to 100 percent of GDP in 2018. Before the Brexit referendum in 2016, the BOE expanded its balance sheet in a similar proportion to the economy as did the Federal Reserve. The additional asset purchases conducted in response to Brexit raised holdings to as much as 30 percent of GDP in 2017.

As with FOMC policy actions, market participants and academics generally agree that BSPs, FG, and other measures implemented by the BOJ, BOE, and ECB eased financial conditions, supported economic activity and inflation, and put downward pressure on currencies, although these effects are imprecisely estimated.34

4.2 What Are the Lessons from the European and Japanese Experiences for the U.S. Monetary Policy Framework?

First, the struggles of the BOJ and ECB to achieve their mandates illustrate the difficulties in raising inflation once longer-run inflation expectations become entrenched at too low of a level. Most notably, in Japan, inflation dropped from a range of 2 to 4 percent in the early 1990s to essentially zero by 1995 and was then mildly negative in most years until 2012 (Figure 8). Low inflation realizations likely eroded the public’s longer-run inflation expectations, reducing incentives to raise prices and wages and thus creating a vicious circle. The BOJ arguably allowed deflation to become entrenched by not acting promptly and forcefully

![Figure 7](image-url)
enough, lowering its key policy rate to zero only in 1999 and initiating asset purchases in 2001—several years after the onset of deflation. The task of gauging the appropriate amount of monetary stimulus required was likely complicated by policymakers’ slow recognition of a fall in $r^*$ and by impaired balance sheets in the banking sector that hindered the transmission of monetary policy.\(^{35}\) BOJ communications and actions at the time also conveyed a lack of confidence in its tools and uneasiness with their deployment.\(^{36}\) Some observers have argued that proximity to the ELB has created a situation in which low inflation and low GDP growth is self-fulfilling.\(^{37}\) BOJ officials have further suggested that a rapidly aging population and a tendency of labor and management to prioritize employment stability over wage increases have contributed to the entrenched perception that wages and prices will not rise.\(^{38}\) It is nonetheless understood that monetary policy and low long-run inflation expectations in Japan have played a major role in creating that perception.

Second, the foreign experience with BSPs suggests that the Federal Reserve could have expanded its balance sheet further without adversely affecting market functioning and still have positive financial and macroeconomic effects at the margin. The ECB, BOE, and, especially, BOJ increased their balance sheets uneventfully to higher shares of their GDPs than did the FOMC. That said, there is some evidence that, while positive, the marginal macroeconomic effects of BSPs abroad were smaller for later programs than earlier programs.\(^{39}\) Hence, setting aside complications regarding the subsequent balance sheet reduction, scaling up BSPs seems to have limited adverse effects, even if also perhaps limited efficacy.
Third, half-hearted commitments, or the perception that a central bank would tolerate persistent deviations from its objectives, can undermine the efficacy of both current and future policy actions. For example, the ECB’s open-ended asset purchases since 2014 have been seen as more potent than its earlier long-term refinancing operations, perhaps because open-ended purchases and the associated large balance sheet expansion convey a clearer commitment to maintaining accommodative conditions for an extended period. ECB president Mario Draghi’s remarks in July 2012, that the institution was ready to do whatever it took to preserve the euro immediately and persistently, calmed financial markets because they conveyed a credible commitment—even though he provided no policy details (Draghi, 2012). As previously noted, the BOJ’s timid use of FG and BSPs in the 1990s and 2000s may have undermined the credibility of the post-2013 policies in support of the 2 percent inflation target.

Finally, limited space to cut policy rates abroad, along with the challenges of deploying untested tools, likely left foreign economies more exposed to the GFC and created negative spillovers for the United States. For instance, the BOJ entered the GFC with its policy rate barely above its ELB, which may have exacerbated Japan’s slump and created the need for greater reliance on BSPs and other tools. Figure 9 illustrates simulated paths of U.S. macro variables in a GFC-like scenario in which the ELB constrains both U.S. and AFE policy rates (labeled “Baseline”) and a counterfactual scenario that relaxes the ELB in AFEs.
lations suggest that the ELB on AFE policy rates depressed U.S. real GDP as much as 0.6 percent, lowered U.S. core inflation by as much as 0.15 percentage point, and delayed U.S. liftoff from the ELB by one quarter.\(^{42}\)

**CONCLUSION**

The FOMC has operated in a challenging environment from the GFC through 2019. The financial crisis produced a sharp rise in unemployment, and inflation persistently fell short of 2 percent during the post-GFC recovery and subsequent economic expansion. Structural transformations that could not be quickly recognized likely weakened the transmission channels of monetary policy and revealed the labor market gap to be larger than previously thought. With the ELB constraining the federal funds rate, the FOMC employed two novel tools, BSPs and FG, to further provide accommodation. These tools effectively facilitated the return to full employment and helped mitigate the ELB, though inflation ran somewhat below 2 percent prior to the onset of the pandemic. With the benefit of hindsight, even bolder use of these tools to achieve mandated goals seems feasible in the future and might have been helpful in the past. However, the experiences of the BOJ and ECB suggest that more-forceful deployment of these tools might still fail to return longer-run inflation expectations to target once they have slipped. Instead, these international episodes point to the importance of prompt action with a clear and sustained commitment. ■
APPENDIX A

Description of Counterfactual Policy Rate Simulations

This appendix describes how we compute the counterfactual historical policy rates and outcomes shown in Figure 5 in the article, under the assumption that policymakers strictly followed the prescriptions of either the Taylor (1993) rule, the balanced-approach rule, or an inertial version of the balanced-approach rule. Table A1 describes the rules.

Table A1
Simple Rules

<table>
<thead>
<tr>
<th>Rule</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taylor (1993) rule</td>
<td>( R_t = r_{t}^{LR} + \pi_t + 0.5(\pi_t - \pi_{t}^{LR}) - \text{ugap}_t )</td>
</tr>
<tr>
<td>Balanced-approach rule</td>
<td>( R_t = r_{t}^{LR} + \pi_t + 0.5(\pi_t - \pi_{t}^{LR}) - 2\text{ugap}_t )</td>
</tr>
<tr>
<td>Inertial balanced-approach rule</td>
<td>( R_t = 0.85R_{t-1} + 0.15(r_{t}^{LR} + \pi_t + 0.5(\pi_t - \pi_{t}^{LR}) - 2\text{ugap}_t) )</td>
</tr>
</tbody>
</table>

Consistent with the FOMC’s annual “Statement on Longer-Run Goals and Monetary Policy Strategy” (FOMC, 2021), we set the inflation goal (\( \pi_{t}^{LR} \)) to 2 percent. As intercept of the rules (\( r_{t}^{LR} \)), we use the median SEP projection for the real federal funds rate in the longer run. For the unemployment gap (\( \text{ugap}_t \)), we use the percentage-point deviation of the unemployment rate from the median longer-run estimate in the SEP.

We perform the simulations using an approach broadly similar to that described by Kashkari (2017). For all simulations, we use the data and model equations in the public release of the FRB/US model that are consistent with the March 2019 SEP. We assume that agents form VAR-based expectations so that they do not anticipate events such as the European sovereign debt crisis. For each simple rule, we begin by calculating the equation residuals of the model such that it perfectly replicates the historical data. Next, we iteratively calculate the model’s counterfactual solution under each rule, zeroing out the residuals in the monetary policy rule equation. This procedure ensures that monetary policy strictly follows the assumed policy rule, while all other shocks in the model are held constant. It also means that the effects of the Federal Reserve’s BSPs and FG are subsumed in the equation residuals and thus held constant across all simulations.
APPENDIX B

Description of Mandates and Strategies of the Bank of Japan, European Central Bank, and Bank of England

This appendix describes the current monetary policy frameworks of the BOJ, ECB, and BOE. Unlike the Federal Reserve, these central banks’ mandates define price stability as their primary objective. However, like the Federal Reserve, these central banks have pursued flexible inflation-targeting strategies in practice (Bernanke, 2003). Moreover, in responding to the GFC and other shocks, these central banks have also used BSPs and FG (as well as a number of other tools that we will describe). Appendix C provides a chronology of these central banks’ policy actions since the GFC.

Bank of Japan. In 1997, the Japanese government significantly increased the BOJ’s independence and established its price-stability objective, which the BOJ defined as a situation in which inflation rates do not affect economic decisions. In February 2012, the BOJ clarified that its price-stability goal meant aiming for annual inflation “within a positive range of 2 percent or lower” and set a goal of “1 percent for the time being” (BOJ, 2012). In January 2013, the Japanese government and BOJ jointly announced a 2 percent consumer price index inflation target.

In the 1990s, the BOJ fought low inflation through low policy rates. In April 1999, it introduced its zero interest rate policy, which was supported by new FG, and, in 2001, launched its first quantitative easing (QE) program, purchasing short-term securities; these measures continued until 2006. Following the GFC, the BOJ launched a comprehensive monetary easing (CME) strategy in 2010 that included FG and a loan support program. The BOJ redoubled its anti-deflationary efforts in early 2013 with the adoption of its 2 percent inflation target and the announcement of a package of stimulative measures called quantitative and qualitative monetary easing (QQME). This package included large-scale, open-ended purchases of Japanese government bonds (JGBs), exchange-traded funds (ETFs), and Japanese real estate investment trusts (JREITs). These purchases expanded the BOJ’s balance sheet to a much larger extent than its previous BSPs. In 2016, the BOJ expanded its QQME strategy further, introduced negative interest rate policies (NIRPs), and soon after added yield curve control (YCC). Under YCC, the BOJ explicitly targets both short- and long-term interest rates, setting the overnight deposit rate at negative 0.1 percent and conducting asset purchases to target a yield on 10-year JGBs at around zero percent.

European Central Bank. The Treaty on the Functioning of the European Union establishes price stability as the primary objective of Eurosystem monetary policy. It also provides that the ECB should avoid excessive fluctuations in employment and output as it pursues price stability. In 1998, the ECB’s Governing Council defined price stability “as a year-on-year increase in the Harmonized Index of Consumer Prices for the euro area of below 2%” (ECB, 1998). In 2003, the ECB’s Governing Council clarified that it seeks to maintain inflation rates below, but close to, 2 percent over the medium term.

In response to the GFC, the ECB lowered the rate on its main refinancing operations to a bit above 0 percent (see Figure 7 in the article), conducted longer-term refinancing operations...
(LTROs) of 3- and 6-month maturities, and eventually complemented these measures with the first covered bond purchase programme (CBPP1) in 2009 and the Securities Market Programme (SMP) in 2010. The resulting expansion of the ECB’s balance sheet was more modest than that of the Federal Reserve’s or the BOE’s at the time. As economic conditions improved, the ECB raised its policy rate in 2011 while keeping BSPs in place, but it then reversed course before the end of the year with the onset of the European sovereign debt crisis. The crisis led the ECB to deploy further accommodative measures. The ECB lengthened the maturity of its LTROs up to three years and launched a new round of the CBPP (CBPP2) in November 2011. President Draghi’s “whatever it takes” remarks in July 2012, which had a calming effect on financial markets, foreshadowed the introduction of qualitative FG and further BSPs—notably the replacement of the SMP with the Outright Monetary Transactions Programme (OMTP) in September 2012 (Draghi, 2012). Despite these measures, the ECB’s balance sheet expansion was temporary in nature—particularly from LTROs—and shrank by 2014 to its 2009-11 levels as banks made early repayments and financial conditions started to ease.

By mid-2014, the ECB’s third phase started to reverse course and the ECB began to expand its balance sheet again: It made a first round of targeted LTROs (TLTRO I) in June 2014, launched a new round of its CBPP (CBPP3) in October 2014, and added its asset-backed securities purchase programme (ABSPP) in September 2014 and its public sector purchase programme (PSPP) in January 2015. In June 2014, the ECB introduced negative interest rates, and, in 2016, lowered the deposit rate to –0.4 percent. The ECB further strengthened its BSP actions later on with another round of TLTROs (TLTRO II) in March 2016, the introduction of its corporate sector purchase programme (CSPP) in June 2016, and an emergency liquidity assistance facility in June 2017. An additional round of TLTROs (TLTRO III) was announced in June 2019. As of mid-2019, the size of the ECB’s balance sheet stood at around 40 percent of euro area GDP, a footprint about twice as large as the Federal Reserve’s (see Figure 7 in the article).

**Bank of England.** The BOE seeks to achieve consumer price index inflation at the annual rate of 2 percent. In response to the GFC, the BOE cut its main policy rate from 5.75 percent to an ELB of 0.5 percent. In March 2009, it embarked on the first of three LSAP phases. Its purchases totaled £200 billion and comprised mostly gilts, along with residual amounts of commercial paper and corporate bonds to support private debt issuance. In response to the European sovereign debt crisis, the BOE launched a second phase of LSAPs in October 2011, introduced a funding for lending scheme in 2012, and used threshold-based FG. In August 2016, following the Brexit referendum, the BOE cut its policy rate to 0.25 percent, announced a third phase of asset purchases, and enhanced its liquidity provision.
# APPENDIX C

## Summary of Policy Actions and Communications in Selected Major Advanced Economies at the Effective Lower Bound

### Panel A: Balance sheet policies

<table>
<thead>
<tr>
<th>Federal Reserve</th>
<th>Bank of Japan</th>
<th>European Central Bank</th>
<th>Bank of England</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LSAP</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSAP1: November 2008 to March 2010</td>
<td>CME$^1$</td>
<td>LTRO</td>
<td>APF: Gilt (QE1)</td>
</tr>
<tr>
<td>LSAP2: November 2010 to June 2011</td>
<td></td>
<td>3 months: Starting in August 2007</td>
<td>January 2009 to February 2010</td>
</tr>
<tr>
<td>LSAP3: September 2012 to October 2014</td>
<td>QQME (JGB purchases)</td>
<td>6 months: March 2008 to May 2010</td>
<td>APF: Commercial paper</td>
</tr>
<tr>
<td>Maturity Extension Program (MEP)</td>
<td>April 2013 to present</td>
<td>12 months: June 2009 to October 2011</td>
<td>March 2009 to November 2011</td>
</tr>
<tr>
<td>September 2011 to June 2012</td>
<td>QQME (ETFs and JREIT purchases)</td>
<td>CBPP1</td>
<td>APF: Gilt (QE1)</td>
</tr>
<tr>
<td></td>
<td>April 2013 to present</td>
<td>July 2009 to June 2010</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>SMP</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>May 2010 to September 2012$^3$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CBPP2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>November 2011 to October 2012</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>LTRO</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3 years: In December 2011 and March 2012</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>OMTP</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>September 2012 to present$^4$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>TLTRO</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TLTRO I: June 2014 to present</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ABSPP</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>September 2014 to December 2018$^4$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CBPP3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>October 2014 to December 2018$^4$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>PSP</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>January 2015 to present</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CSPP</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>June 2016 to December 2018$^5$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>TLTRO</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TLTRO II: March 2016 to present</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TLTRO III: September 2019 to present</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C, cont’d

Summary of Policy Actions and Communications in Selected Major Advanced Economies at the Effective Lower Bound

Panel B: Forward guidance: Qualitative, calendar-based, and threshold-based

<table>
<thead>
<tr>
<th>Type of FG</th>
<th>Federal Reserve</th>
<th>Bank of Japan</th>
<th>European Central Bank</th>
<th>Bank of England</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qualitative</td>
<td>December 2008 to July 2011</td>
<td>July 2013 to present†</td>
<td>Early 2015 to present†</td>
<td></td>
</tr>
<tr>
<td>Calendar-based</td>
<td>August 2011 to November 2012</td>
<td>October 2014 to early 2015†</td>
<td>August 2013 to early 2015**</td>
<td></td>
</tr>
<tr>
<td>Threshold-based</td>
<td>December 2012 to February 2014</td>
<td>With QE</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>March 2001 to March 2006</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>With CME</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>October 2010 to March 2013</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>**With QQME with price-stability</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>target of 2 percent†</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>April 2013 to present</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Panel C: Negative interest rates

<table>
<thead>
<tr>
<th>Type of negative rates</th>
<th>Federal Reserve</th>
<th>Bank of Japan</th>
<th>European Central Bank</th>
<th>Bank of England</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short rates</td>
<td>QQME with NIRP</td>
<td>Negative deposit facility rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>January 2016 to August 2016</td>
<td>August 2014 to present</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short and long rates</td>
<td>QQME with YCC</td>
<td>September 2016 to present</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE:† The Federal Reserve’s communication practices evolved significantly with the introduction, in January 2012, of an explicit inflation target of 2 percent on headline PCE inflation and of a depiction of the FOMC’s assumptions for the federal funds rate (informally known as the “dot plot”) in the SEP.

‡ President Draghi’s “whatever it takes” remarks on July 26, 2012, preceded the formal implementation of FG and the deployment of the OMTP (Draghi, 2012). The ECB’s SMP was discontinued at the same time as that program was introduced, which would then allow purchases (“outright transactions”) in secondary, sovereign bond markets, under certain conditions, of bonds issued by euro area member states.

§ The Federal Reserve used qualitative FG before the GFC during the “deflation scare” episode (August 2003 to December 2005).

¶ Another policy tool considered by the BOJ is raising the inflation target—Japanese policymakers did this while implementing QQME with a 2 percent price stability target in April 2013 (up from a midpoint of 1 percent set a year earlier, and a goal of positive inflation before that).

** The BOE committed not to raise rates until the unemployment rate fell to 7 percent, which happened sooner than expected. It then revamped its FG strategy in February 2014, promising to focus on 18 measures of spare capacity instead, but it stopped mentioning threshold-based FG in its quarterly economic updates by early 2015. The BOE has continued to use qualitative FG to signal the path of interest rates but only incidentally.
### APPENDIX D

**Selected Studies of Financial Market Effects of Balance Sheet Policies and Forward Guidance**

<table>
<thead>
<tr>
<th>Source</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gagnon et al. (2011)</td>
<td>10-year Treasury yield fell 91 basis points over baseline event set and 55 basis points over the “all-event” set.</td>
</tr>
<tr>
<td>D’Amico and King (2013)</td>
<td>Asset purchases produce local supply effects.</td>
</tr>
<tr>
<td>Di Maggio, Kermani, and Palmer (2020)</td>
<td>MBS purchases, but not Treasury purchases, depressed mortgage rates with particularly strong effects on conforming mortgages.</td>
</tr>
<tr>
<td>Rodnyansky and Darmouni (2017)</td>
<td>Institutions with relatively large holdings of MBS expanded lending after LSAP1 and LSAP3, but not after LSAP2.</td>
</tr>
<tr>
<td>Kurtzman, Luck, and Zimmermann (2018)</td>
<td>Banks with greater MBS holdings reduced lending standards and made riskier loans.</td>
</tr>
<tr>
<td>Chakraborty, Goldstein, and MacKinlay (2020)</td>
<td>Banks that owned greater amounts of MBS increased mortgage lending at the expense of commercial and industrial lending.</td>
</tr>
<tr>
<td>Neely (2015)</td>
<td>BSPs spilled over to foreign exchange markets and foreign bond markets, reducing the value of the dollar and foreign bond yields.</td>
</tr>
<tr>
<td>Kiley (2014)</td>
<td>BSPs had modest effects on equities because they moved the medium and long yields but not short yields.</td>
</tr>
<tr>
<td>Hamilton and Wu (2012)</td>
<td>A segmented-markets model implies that a $400 billion purchase will reduce 10-year Treasury yields by 13 basis points.</td>
</tr>
<tr>
<td>Greenwood and Vayanos (2014)</td>
<td>Maturity-weighted debt-to-GDP increases long-term yields. The Federal Reserve’s LSAP1 and LSAP2 together reduced yields 40 basis points.</td>
</tr>
<tr>
<td>Altavilla and Giannone (2017)</td>
<td>The unconventional policy effects on the Federal Reserve Bank of Philadelphia’s SPF forecasts are persistent and consistent with those from event studies.</td>
</tr>
<tr>
<td>Greenwood et al. (2014)</td>
<td>U.S. Treasury lengthening the maturity structure of U.S. debt may have offset one-third of the effects of Federal Reserve asset purchases on long yields.</td>
</tr>
<tr>
<td>Raskin (2013)</td>
<td>The FOMC’s promise to keep rates low until mid-2013 had a much bigger effect than the promise to keep rates low until mid-2014.</td>
</tr>
<tr>
<td>Femia, Friedman, and Sack (2013)</td>
<td>The Federal Reserve Bank of New York’s Survey of Primary Dealers indicates that FG provides market participants with economic information.</td>
</tr>
</tbody>
</table>
The SEP added the central tendency and range of longer-run estimates for the unemployment rate, headline PCE inflation, and real gross domestic product (GDP) growth in April 2009, and corresponding median estimates in September 2015. Before the latter date, we derive median estimates from declassified individual SEP contributions where possible and report the midpoint of the central tendency otherwise.

These statistics are based on the peak-to-NAIRU analysis of Eberly, Stock, and Wright (2020). Comparisons across recoveries are sensitive to the reference period. For example, the unemployment rate rose modestly in the first two years of the 1990s and early 2000s economic expansions, making the labor market recovery from the GFC look especially strong relative to those episodes when measured from the start of the economic expansion rather than from the peak in the unemployment rate. By contrast, the unemployment rate fell 3.5 percentage points in the first year and a half following the 1981-82 recession, making the initial labor market recovery from the GFC look especially weak relative to that episode.

More precisely, the labor force participation rate held steady, on net, from late 2013 until 2019, in contrast with the trend decline projected by Aaronson et al. (2014).

For evidence on the role of longer-run inflation expectations and slack in determining realized inflation, see Ball and Mazumder (2011); Del Negro, Giannoni, and Schorfheide (2015); Duncan and Martínez-García (2015); Coibion and Gorodnichenko (2015); Yellen (2015); and Kabukcuoglu and Martínez-García (2018).

As of the writing of this article, 2014 is the latest calendar year for which Tealbook forecasts and FOMC memos are publicly available. In mid-2014, the staff also projected longer-run inflation expectations to eventually drift higher, pushing underlying inflation toward 2 percent, but acknowledged that such an upward drift was highly uncertain. See, for example, Aaronson et al. (2015) and Cairó and Cajner (2018).

To some extent, a step-down in productivity growth following the information technology boom of the mid-1990s to early 2000s is unsurprising, though the extent and timing of that step-down were difficult to predict. See Fernald (2015) and Gordon (2015).

All estimates but one in Figure 2 are “one sided,” meaning that, at each date, they use historical data only up to that date. Accordingly, the fact that many series show a decline around the GFC need not imply that it itself fell as a result of the crisis, but rather that the crisis marked the moment when models began to identify the fall.

For evidence suggesting such a change in monetary policy, see Clarida, Gali, and Gertler (2000); Boivin and Giannoni (2006); and Boivin, Kiley, and Mishkin (2010).

Based on pre-2020 information, Figure 4 reports, for the first quarter of each calendar year, median SEP projections at various yearly horizons, along with corresponding medium-term projections from market participants. Individual SEP projections are conditional on each policymaker’s assessment of appropriate monetary policy, whereas market participants’ projections are modal forecasts. For this reason, the two projection concepts reported in Figure 3 differ. Evidence of policymakers’ slow recognition of the fall in productivity growth is corroborated by the significant downward revisions to real output growth in the longer run: The median SEP estimate fell from 2.6 percent in April 2009 to 1.9 percent by mid-2019.

Policymakers expected inflation to be weaker than did market participants during the first years of the recovery, even though policymakers held relatively upbeat views of future real GDP growth and unemployment. This observation is consistent with policymakers assuming that the necessary reduction in slack to lift inflation was larger than market participants expected. We also note that policymakers’ inflation forecast errors averaged about zero, even though the unemployment rate fell faster than policymakers had anticipated, suggesting that inflation proved even less responsive to movements in slack than policymakers had assumed.

Burke et al. (2010) review the considerations raised by Board staff at the time.

The FOMC’s use of BSPs and FG in response to the GFC was unprecedented. Though the FOMC purchased longer-term Treasury securities using proceeds from sales of shorter-term holdings in the early 1960s (the so-called Operation Twist), the total size of this program, at 1.7 percent of GDP, was much smaller than total purchases during the GFC and its aftermath (Swanson, 2011). Also, before the GFC, the FOMC used FG regarding the likely path of interest rates on occasion, but that guidance was usually confined to a relatively short time frame.
FG can serve dual purposes: It can convey policymakers’ predictions about future economic conditions and policies or their commitment to future policy. These two purposes are called “Delphic” and “Odyssean,” respectively, by Campbell et al. (2017). Del Negro, Giannoni, and Schorfheide (2015) argue that the FOMC’s FG became more of an Odyssean commitment to stimulative policies with the introduction of the 2 percent inflation target and the release of FOMC participants’ assessments of appropriate monetary policy in 2012.

The mean DSGE estimates in Gust et al. (2017) suggest that a binding ELB accounted for about 30 percent (roughly 2 percentage points) of the 6 percent contraction in GDP in 2009 relative to its peak in 2007 and was responsible for an even larger fraction of the ensuing slow recovery. The estimates of Gust et al. (2017), which embed the effects of FG and LSAP, are more consistent with the more immediate negative effects of the ELB on outcomes than are those of Eberly, Stock, and Wright (2020).

See, for example, Swanson (2018); Sims and Wu (2019); and Debortoli, Gali, and Gambetti (2020).

For example, Cabana et al. (2008) estimate that “a purchase of $50 billion of longer-term Treasury securities… would lower the 10-year Treasury yield somewhere between 2 and 10 basis points.” See also Gagnon and Holscher (2008) and Erceg, Kiley, and Levin (2008a,b).

One exception is Greenlaw et al. (2018), who take a skeptical view of the effects of Federal Reserve’s asset purchases. Bhattarai and Neely (forthcoming) and Kuttner (2018) review research on BSPs and FG.

See Di Maggio, Kermani, and Palmer (2020).

See Rodnyansky and Darmouni (2017); Chakraborty, Goldstein, and MacKinlay (2020); and Kurtzman, Luck, and Zimmermann (2018).

See Swanson and Williams (2014).

Using data from 2007 to mid-2014, Greenwood et al. (2014) report that the Federal Reserve’s BSP programs reduced the publicly held 10-year equivalent universe (Treasury securities, MBS, and agencies) by approximately 15.6 percent of GDP, while Treasury maturity extensions increased the outstanding 10-year equivalents by 5.5 percent of GDP. Greenwood et al. (2014) then calculate that the Treasury’s duration increase offset more than 35 percent (≈ 5.5/15.6) of the effect of the Federal Reserve’s BSP and FG policies on 10-year yields. Greenwood et al. (2014) cite meta estimates from Williams (2014) to put the total effect of Federal Reserve policies on 10-year yields at 137 basis points, so the offset was 48 basis points.

Swanson (2021) calculates that the point estimate of the half-life of FG effects is 6.5 months, which is statistically indistinguishable from complete persistence.

See, for example, the minutes to the September 2012 FOMC meeting, available on the Federal Reserve Board’s website at https://www.federalreserve.gov/monetarypolicy/files/fomcminutes20120913.pdf.

See Gilchrist and Zakrjašček (2013).

See Kiley (2018) and Chung et al. (2019) for stochastic simulations of the U.S. economy under an endogenous balance sheet response.

Clarida (2019a) discusses the Federal Reserve’s monetary policy framework articulated in 2012 and the scope of the Federal Reserve’s 2019-2020 review of its strategy, tools, and communication practices.

Taylor (1993) shows that the rule now bearing his name fits policy rate settings well during the 1987-92 period. See Yellen (2017) for an application of the balanced-approach rule. The inertial version of the balanced-approach rule has been featured in Federal Reserve Board staff analysis; see, for example, Erceg et al. (2012).

The median SEP projection for core inflation three years out was only 1.8 percent when the FOMC announced its 2 percent objective in 2012. The simulations in Engen, Laubach, and Reifschneider (2015) and Chung et al. (2019) suggest that policymakers would have needed to boost asset purchases an extra couple trillion dollars to meet their inflation objective over the medium term.
Although headline inflation exceeded the targets in Japan and the euro area on occasion, that achievement typically reflected temporary factors, for example, the effects of the 2014 value-added tax, or VAT, hike in Japan and of global oil price rebounds in the euro area.

Appendix B describes the policy frameworks of the BOJ, ECB, and BOE; Appendix C summarizes their key policy actions and communications in the wake of the GFC.

Some policy tools used by the ECB and BOJ—such as negative policy rates and yield curve control—are not part of the FOMC’s current framework.

For a review of financial and macroeconomic effects in the United States and abroad, see Andrade et al. (2016); Lombardi, Siklos, and St. Amand (2018); Kuttner (2018); and Dell’Ariccia, Rabanal, and Sandri (2018), who focus on the effect on yields and bank lending. On the international experience and spillovers, see Clarida (2019b), Martínez-García (2019), and the evidence in Haldane et al. (2016), Chen et al. (2016), and Martínez-García (2018).


Japanese asset purchases during this period were probably unsuccessful partly because the BOJ purchased bonds with relatively short remaining maturity (McCauley and Ueda, 2009). The average maturity of the BOJ’s portfolio actually declined from 2001 to 2005.

See Krugman (1998); Benhabib, Schmitt-Grohé, and Uribe (2011); Eggertsson (2010); and Bullard (2010) for a discussion of self-fulfilling deflationary expectations and policy options in a liquidity trap (when policy rates are constrained at the ELB).


Hesse, Hofmann, and Weber (2018) revisit the macroeconomic and financial effects of the asset purchase programs launched by the Federal Reserve and the BOE from 2008 and suggest that the early programs “had significant positive macroeconomic effects, while those of the subsequent ones were weaker and in part not significantly different from zero.” See also Joyce and Tong (2012) and Churm et al. (2021).

See Andrade et al. (2016); Doehr and Martínez-García (2015); and Weale and Wieladek (2016) for a related discussion of the macroeconomic effects of news and announcements about asset purchase programs in the United States and the United Kingdom.

We use a three-country version of SIGMA, a DSGE model maintained by Federal Reserve Board staff. In this simulation, AFE policymakers cut their policy rate 350 basis points at the recession’s onset.

In the current low interest rate environment, the ELB would bind faster in the AFEs for a given recessionary shock than in our scenario, leading to larger downturns abroad and greater negative spillovers for the U.S. economy, unless AFE central banks provide accommodation through other tools.

REFERENCES


Gertler, Mark and Karadi, Peter. “QE 1 vs. 2 vs. 3…: A Framework for Analyzing Large-Scale Asset Purchases as a Monetary Policy Tool.” International Journal of Central Banking, January 2013, 9, pp. 5-53; https://www.iijcb.org/journal/ijcb13q0a1.pdf.


Kuttner, Kenneth N. “Outside the Box: Unconventional Monetary Policy in the Great Recession and Beyond.” Journal of Economic Perspectives, Fall 2018, 32, pp. 121-46; https://doi.org/10.1257/jep.32.4.121.


When the COVID-19 pandemic struck in March 2020, the U.S. economy experienced a sharp, unexpected recession with large employment losses. The information on employment available from traditional data sources arrives with a lag and does not promptly reflect sudden changes in labor market conditions. In this article, we discuss how new high-frequency data from Homebase and Ultimate Kronos Group can offer critical information on the state of labor markets in real time. Using these datasets, we construct coincident employment indices to assess employment at a high frequency. Employment during the pandemic reacted to changes in the number of infections and the restrictions imposed by government officials (see, e.g., the discussion in Dvorkin and Bharadwaj, 2020). Our latest data suggest that employment has recently increased and will continue to increase as the pandemic wanes. (JEL J2, E27)
In normal times, labor market conditions change slowly, and monthly information from the CPS and CES on employment and unemployment is sufficient to gauge the state of labor markets and predict their evolution. However, the rapidly changing conditions during the pandemic called for higher-frequency data to evaluate the state of labor markets in real time. In a recent working paper, Chetty et al. (2020) developed a granular real-time publicly available dataset to address this challenge. We also use high-frequency employment data from two companies with ample coverage of U.S. labor markets, Homebase and Ultimate Kronos Group (UKG). Kurmann, Lalé, and Ta (2020) use the Homebase data to demonstrate how business closures by firm size impacted employment throughout the pandemic. In this article, we use the Homebase and UKG data to create and track useful measures of aggregate and regional employment conditions.

HIGH-FREQUENCY DATA

Homebase is a private company delivering payroll, scheduling, and timesheet tools to businesses. They offer both a free version of their software with basic capabilities and a fee-based version with premium services. Homebase currently provides services to over 100,000 businesses, per their website, across many industries. However, the bulk of Homebase customers are centered in leisure and hospitality, especially in the food and drink industry. After the pandemic started in March 2020, Homebase began releasing daily information on key labor market variables for the companies using their products. These updates took two forms: (i) a summary dataset detailing how much national- and state-level employment, hours, and businesses changed relative to January 2020 and (ii) firm- and worker-level microdata.

The summary datasets, published daily, use a cohort of roughly 60,000 firms that were active on the Homebase platform in January 2020 to create benchmark numbers of employees, the hours they log, and active businesses in the data. For example, on July 12, 2020, across the entire United States, the number of employees working was down 26.3 percent relative to January, the number of hours worked was down 27.0 percent, and the number of active businesses was down 22.0 percent. These numbers are available also by state and industry.

The microdata contain similar information on hours, employees, and businesses, but this information is not limited to the cohort of firms active in January 2020—it is available for all firms in the sample. For 2020, there are about 100,000 unique companies represented in the data, with an average of 12 employees and about 4,800 hours logged per company between January and December.

Figure 1 compares the distribution of employment across industries for the Homebase data with that of the CPS data. Homebase uses its own industry classification system, which we translate into nine broad sectors according to the North American Industry Classification System (NAICS) to be able to compare the Homebase data with the CPS data. Some major industries, such as manufacturing and construction, are not present in the Homebase data at all, while professional employment is severely underrepresented. The figure shows that the Homebase data are heavily weighted toward establishments that serve food and drink and provide recreational activities. Employees in the food, drink, and leisure industry make up over half of the entire dataset.
Figure 1
Distributions of Employment in the Homebase and the CPS Data, by Industry

![Bar chart showing distributions of employment by industry.](chart1)

SOURCE: Homebase, CPS, and authors’ calculations.

Figure 2
Distributions of Employment in the Homebase and the CPS Data, by State

![Bar chart showing distributions of employment by state.](chart2)

SOURCE: Homebase, CPS, and authors’ calculations.
Figure 2 compares the distribution of employment in the Homebase data and in the CPS data across U.S. states. We can see that states such as California, Florida, and Texas are overrepresented in the Homebase data compared with the CPS data, and states such as New York and Illinois are underrepresented. However, all states and Washington, D.C., are present in the Homebase data. The overly large number of firms in the food, drink, and leisure industry might overrepresent certain areas of the country. Other key industries, such as manufacturing, are not present in the Homebase data at all, which would underrepresent areas of the country in which those industries are prevalent.

Many Homebase clients are small firms, which is reflected in the Homebase sample. Almost 90 percent of the businesses in the Homebase data have fewer than 100 employees. Large firms, those with more than 500 employees, represent less than 1 percent of the firms in the Homebase sample. Figure 3 compares the shares of employment by firm size for the Homebase data with those for the U.S. economy in the first quarter of 2020. As the figure shows, most of the employment in the Homebase data corresponds to firms with fewer than 100 employees, while for the U.S. economy this share is close to 37 percent. This is a caveat of the data that we need to consider in our analysis.

The second dataset of interest comes from UKG, a company that also provides payroll and time management tools to businesses and their employees. UKG provides services to over half of the Fortune 1000 firms and has data on millions of employees. The UKG sample, used for...
January 2020 to March 2021, contains weekly information on total employee “clock punches” at the national, state, and firm levels. It contains information on over 30,000 businesses across all 50 states. The punch data record each time an employee clocks in or clocks out. For most companies, one employee working one shift would record one punch in and one punch out. However, some require their employees to clock out and back in again for breaks. For those companies, a single employee shift would generate four punches. While the punch data cannot perfectly capture the exact number of employees working in a week, the changes in the number of punches are indicative of the changes in employment in a state or industry.

Figure 4 compares the distribution of employment by firm size for the UKG sample with that for the U.S. economy. As the figure shows, the UKG data include firms of all sizes and the distribution of employment is similar to that for the U.S. economy, with only a slightly larger share of employment in smaller firms and slightly smaller share in very large firms in the UKG sample.

UKG classifies industries into six major sectors: healthcare; manufacturing; retail, hospitality, and food service; services and distribution; public sector and non-profit work; and unclassified. They also categorize industries by NAICS, which in Figure 5 we group into nine broad sectors and compare with the CPS data. Unlike the Homebase dataset, most of the major industries in the UKG dataset have positive employment, including natural resources, financial occupations, and construction, and the distribution of employment is similar to that for the U.S. economy.
Figure 5
Distributions of Employment in the UKG Dataset and the U.S. Economy, by Industry

![Bar chart showing the distribution of employment by industry for the UKG dataset and the U.S. economy.](chart)

SOURCE: UKG, CPS, and authors’ calculations.

Figure 6
Distributions of Employment in the UKG Dataset and the U.S. Economy, by State

![Bar chart showing the distribution of employment by state for the UKG dataset and the U.S. economy.](chart)

SOURCE: UKG, CPS, and authors’ calculations.
Last, Figure 6 compares the distribution of employment by state for the UKG sample with that for the U.S. economy. The UKG sample has all states represented in the data, and the distribution of employment by state is similar to that for the U.S. economy. However, some individual states, notably Florida, Texas, and Missouri, are over or underrepresented by a few percentage points.

**EVOLUTION SINCE MARCH 2020**

We focus on employment dynamics in the two datasets since the beginning of the pandemic in March 2020 and compare them with the dynamics in the U.S. economy as measured by the CPS. In order to compare similar time periods, we use data from Homebase and UKG for the week in which the CPS survey was conducted. The scatterplot in Figure 7 shows employment changes relative to January 2020 in the Homebase dataset versus in the CPS data, by month and state. Each circle represents the change in employment in a given month for a U.S. state, the size of each circle corresponds to the state’s share of employment in January 2020, and the color of the circle represents the given month. The April 2020 data (teal circles) are among the lowest data on the chart, indicating that the level of employment in April 2020 was much lower relative to January 2020 than in other months. Employment in May 2020 and June 2020 was also low relative to January 2020, but as employment recovered, the values become less negative on average. The points since June 2020 have stayed relatively close together.
In a similar way, the scatterplot in Figure 8 shows employment changes relative to January 2020 in the Homebase dataset versus the CPS data, by month and industry. Each marker color represents one of the NAICS sectors, the size of the marker corresponds to the size of the industry in January 2020, and the shape of the marker indicates the month. April 2020 and May 2020 employment across almost all industries dropped sharply relative to January 2020, just like in the state data. Employment in the food, drink, and leisure industry dropped the most out of all the industries in the data, as the restrictions imposed due to the pandemic affected these industries (which include restaurants, hotels, and so on) more than others.

The UKG punch data can be used to map national, state, and industry employment across UKG customers starting in January 2020. In this case, the UKG data track the overall number of punches across a week instead of the number of active employees each day. Figure 9 compares the indexed values of UKG employment to those of CPS employment, by state and by month, using the same markings as in Figure 7. The UKG dataset is highly correlated with the CPS index, and while the correlation is lower than that for the Homebase data with the CPS index, both datasets display similar patterns. The levels of employment in April 2020 and May 2020 compared with those in January 2020 are consistently lower than those for other months. Some small states in terms of employment in the UKG data display large departures from the average changes in the data, which suggests that for these cases the employment changes in the UKG data do not align well with the changes in the CPS data. This could be
**Figure 9**

Employment Changes in the UKG Dataset vs. in the CPS Data, by Month and State

![Graph showing correlation between UKG and CPS employment changes by month and state.](image)

Correlation = 0.56

SOURCE: UKG, CPS, and authors’ calculations.

**Figure 10**

Employment Changes in the UKG Dataset vs. in the CPS Data, by Month and Industry

![Graph showing correlation between UKG and CPS employment changes by month and industry.](image)

Correlation = 0.62

SOURCE: UKG, CPS, and authors’ calculations.
the result of a small sample (a small number of firms and employees in these states in the UKG dataset) leading to an outsized impact by a small number of firms.

Figure 10 compares employment changes in the UKG dataset and the CPS for the major industries, using the same markings as in Figure 8. The UKG data have a larger set of industries than the Homebase data, and employment in the UKG data is not heavily concentrated in a small set of sectors. As the figure shows, for both datasets, the food, drink, and leisure industry displays the sharpest decrease in employment early in the pandemic.

An advantage of the UKG and Homebase datasets is the high frequency of their data, which is available almost in real time. Figure 11 shows the weekly evolution of aggregate employment in both the UKG and Homebase datasets. They follow similar patterns across the given period, although the magnitudes of the changes vary, with the Homebase dataset having a much larger contraction in employment in April 2020. This finding reflects differences in the compositions of the sectors and firms by size in these datasets. As mentioned, the Homebase dataset has a large number of firms in the food, drink, and leisure industry. In the early weeks of the pandemic, restaurants and bars quickly shut down, leaving many employees with no work. Once the economy began to slowly open again in May and June, many restaurants reopened and were able to bring their employees back to work or hire new staff. Industries in the UKG data are less concentrated in terms of employment, thus the evolution of employment in that dataset does not present such a large initial drop.

In addition, both series present sudden temporary drops in employment, which are due to holidays, such as the Fourth of July, Thanksgiving, and Christmas, when a large number of employees are not at work and thus not clocking in and out.
The Homebase and UKG employment datasets are strongly correlated with the CPS employment data, suggesting that we can use these datasets to predict national employment at a higher frequency and in real time. In this section we discuss how we do this.

As we showed previously, there are potential caveats with the use of these datasets since the composition of the sample of firms in terms of geography, industry, and firm size may differ in some dimensions from that of the U.S. economy. There is one additional problem we must tackle. Every period, new firms are added to the dataset. These new firms are new customers of Homebase and UKG that start using their products, but they are not necessarily newly created firms. While we would like to capture employment growth due to new firm creation in our indicators, using these new firms in the datasets would be inappropriate. At the same time, completely excluding these new firms would bias downward our estimates since the sample would have business closures and firm shutdowns but not any new businesses.

The number of firms in the datasets dropped precipitously in April 2020, as seen in Figure 12. The blue portions of the bars represent the firms that were in the dataset for at least one month prior, while the pink portions of the bar represent any firms entering the dataset for the first time. If we were to use a fixed cohort of firms in the construction of our indicators, our measures of employment would be affected by having a reduction in the number of firms but no firms entering to make up for the loss. Unless the remaining firms in the dataset significantly increased their employment, it would take a long time for any measure of employment to fully return to the levels of January 2020, if it returned fully at all. All of the firms in pink
enter the dataset throughout the months we analyze, and any increase in employment that they provide goes uncounted if a fixed set of firms is used. Instead, our approach incorporates information from the new firms entering the dataset, but with a lag. In this way, we allow for the cohort of firms to vary over time, and compute changes in employment from each cohort by period. In other words, we construct a chain index for each of the Homebase and UKG datasets.

We construct a chain index for each state as well as the overall United States on a weekly basis. The Homebase data is first collapsed to weekly frequency and measured from Friday to Friday, while the UKG data is already reported at a weekly level from Sunday to Sunday. The first value in the index in January 2020 is normalized to 100, while the values of the index for the weeks in February use the employment growth between January and February for all the firms present in the dataset in January. Similarly, the index for March is computed using employment growth for the firms present in the dataset in February and so on throughout the year. The base value of employment for each month is the median value of weekly employment in each state in the given month. The raw value for the number of employees in a week is divided by the base value drawn from the month before and multiplied by 100. This creates the relative links for each week. The relative links are then multiplied by the previous month’s index and divided by 100 to create the chain index values.

Then, using our chain indices for UKG and Homebase at the state level, we regress a state-level index of employment from the CPS for each month on the state values of each of our indices separately, where the regression is weighted by the level of employment in the CPS in January 2020. In this regression, we use the weekly data from UKG and Homebase that correspond to the survey week of the CPS. The coefficients from the regression, found in Table 1, are used to predict values of employment for the overall United States for each chain index. The R-squared of the regressions are 0.636 for Homebase and 0.316 for UKG, so our indices have substantial forecasting ability. The coefficient of the UKG index is slightly smaller than that of the Homebase index, which implies that our forecasts of the CPS values tend to react more to changes in the Homebase data when compared with similar changes in the UKG data.

Table 1

Regression Output from Regressing CPS Index on Homebase Chain Index and UKG Chain Index

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1) Homebase index</th>
<th>(2) UKG index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Homebase index</td>
<td>0.183*** (0.006)</td>
<td></td>
</tr>
<tr>
<td>Change in UKG index</td>
<td></td>
<td>0.166*** (0.011)</td>
</tr>
<tr>
<td>Observations</td>
<td>765</td>
<td>765</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.636</td>
<td>0.316</td>
</tr>
</tbody>
</table>

NOTE: Each regression also includes a constant term. Standard errors are in parentheses. ***p < 0.01.
Using these regression coefficients, we project the employment indices from the Homebase and UKG data onto CPS values and create a real-time predictor of CPS employment. Figure 13 shows the values of these two predictors over time and compares them to the actual CPS values for employment as deviations from January 2020. We label our predictors the coincident employment indices (CEI).

As the figure shows, both indices track well the evolution of the CPS. This is not surprising since the regression is meant to deliver that. However, over time, the regression coefficients tend to change only slightly; thus, it is the underlying data from Homebase and UKG that are responsible for the good in-sample fit.

CEI FORECASTING PERFORMANCE

We now ask whether our CEI are good at predicting labor market conditions and compare them with another indicator. We focus on the ADP Research Institute (ADP; 2021) employment forecasts from May 2020 onward (published in their “ADP Employment Reports”). Panel A of Figure 14 compares the Homebase and UKG CEI forecasted changes to the actual changes in employment as measured by the CPS (not seasonally adjusted). To generate the values for the CEI forecasted changes in employment, we computed smoothed versions of each index by taking the average between the value of the index a week before the survey and a week after the survey. The reason for this is that some holidays tend to lower the number of clock punches in the Homebase and UKG data, as some employees take some time off. Whether
we use the smoothed version in our forecast depends on when holidays fall. For most of the
data in Figure 14, we used the smoothed version of the index and then multiplied the index
depends on when holidays fall. For most of the
data in Figure 14, we used the smoothed version of the index and then multiplied the index value by the value for employment that month.

ADP forecasts U.S. nonfarm private sector employment as measured by the CES (seasonally adjusted). Thus, Panel B of Figure 14 compares ADP forecasts to the CES employment numbers as published in the first release. The ADP value is the value for the current month minus the value for the previous month (reported in the same report) to generate the predicted change.

While all three indices displayed above are not perfect predictors, the ADP forecast and the CEI values all do a good job predicting CPS employment levels. Toward the beginning of the pandemic, the Homebase and UKG indices do a better job than the ADP forecast.
PRE-RECESSION INDICES

Although the different high-frequency indices generate accurate predictions for employment in the current recession, that does not necessarily imply that the same methodology will work for non-pandemic and non-recession times. Using Homebase microdata from 2018 onward, we generate chain index predictions starting in March 2018. Indexing all time periods to January 2018 produces a relatively large constant and a relatively small coefficient on index changes. Thus, instead of indexing everything on January 2018, we separate the data into time periods from March to March of each available year, resulting in three separate indices for three blocks of time. The data from March 2018 to March 2019 are indexed to January 2018, the data from March 2019 to March 2020 to January 2019, and the data from March 2020 to March 2021 to January 2020. Then we replicate the chain index process from above in each time period. Splicing the separate indices together, we generate Figure 15, which compares the separated indices to the CPS unemployment number in each month.

The predictions track the CPS unemployment values extremely well for both pre-recession and recession employment values. The high-frequency nature of the data even tracks the initial employment drop, which is difficult to track with less-frequent datasets.

CONCLUSION

In the second quarter of 2020, the U.S. economy was facing one of the most severe and sudden recessions, with unprecedented employment loss. Yet, the information available up to the early days of May about employment and unemployment only reflected labor market conditions in mid-March before much of the restrictions and lockdowns were set.
Policymakers need accurate and timely data to adopt optimal policy actions. Traditional data sources for labor market data, such as the CPS and CES, provide valuable information but face important lags. In this article, we discuss how new high-frequency data can offer critical information on the state of labor markets in real time. Using data from Homebase and UKG, we construct coincident employment indices to assess labor market conditions at a high frequency.

To a large extent, the number of COVID-19 cases has been steadily falling since January 2021, and state and local governments have lifted most restrictions on daily activities. The recent evolution of our indices suggest that employment has increased and most likely will continue to do so as the pandemic wanes.

NOTES
1 The CPS is co-sponsored by the Census Bureau and the Bureau of Labor Statistics (BLS).
2 The CES is sponsored by the BLS.
3 The Homebase data are available at https://joinhomebase.com/.
4 The Ultimate Kronos Group data are available at https://www.ukg.com/.
5 Eleven percent of the firms in the Homebase data do not have a classification and are not included in the distribution of firms by industry.
6 The distribution of employment by firm size for the U.S. economy comes from BLS data and the authors’ calculations.
7 In addition, the UKG sample contains information on payrolls, which provides a count on the number of employees receiving paychecks in a given week. The payroll data record the number of paychecks issued, which, for most companies, occurs every two weeks. The payroll data keep weekly records; as such, many companies have payroll levels that oscillate between zero employees recorded and the actual number of paychecks issued.
8 The distribution of employment excludes the 16.7 percent of employees who work for a firm without any industry classification. Of the 29,642 firms in the data, 11,415 firms, or 38.5 percent, do not have an industry classification.
9 The BLS conducts the CPS every month on the week containing the 12th day of the month. We use the same week in our comparison using Homebase and UKG data.
10 The CES employment numbers are subject to revisions. Here we use the first release, which is the first (unrevised) number the BLS publishes, typically on the first Friday of each month.

REFERENCES
1 INTRODUCTION

Under the rational expectations (RE) hypothesis, the expectations of agents are consistent with and always confirmed by equilibrium outcomes. This hypothesis often significantly simplifies the analysis of complicated economic problems. However, the RE hypothesis is silent on how agents form their expectations and, hence, provides no guidance on selecting an equilibrium when multiple RE equilibria occur. In contrast, a learning model specifies the agent’s learning rule for forming expectations and is used to select one from multiple RE equilibria. The advantage of using the learning approach for equilibrium selection is noted by Evans and Honkapohja (2001). The standard criterion for equilibrium selection in the learning approach is stability of the learning dynamics. This consideration is quite intuitive: If a learning equilibrium is not stable, then it is unlikely to be the long-run equilibrium outcome.

Examples of using stability of the learning dynamics to eliminate some RE equilibria include Lucas (1986), Marcet and Sargent (1989), Woodford (1990), and Bullard and Mitra (2002). Lucas (1986) argues that adaptive behavior of economic agents may narrow the set of equilibria in some economic models. By using the stability criterion under learning, Marcet...
and Sargent (1989) eliminate the hyperinflation equilibrium of Sargent and Wallace (1985).
In a monetary model with multiple RE equilibria, Woodford (1990) shows that the economy could converge to a stationary sunspot equilibrium under learning. Bullard and Mitra (2002) argue that the stability under learning criterion is necessary for monetary policy evaluation, especially in situations where multiple RE equilibria could be induced by policy.

The focus of our article is on equilibrium selection using the stability criterion under different specifications of learning. We conduct our exercise using an example in Bullard (1994), which is a simplified version of the model in Sargent and Wallace (1981). Our choice of the Sargent-Wallace framework is deliberate. The framework is simple and admits two steady states under RE, so we can explore the issue of selection. Furthermore, the key features of the model have been used repeatedly in the learning literature; see, for instance, Marcet and Sargent (1989), Bullard (1994), and Marcet and Nicolini (2003). We employ the model to examine equilibrium selection in different learning specifications via the stability criterion, a criterion that is common in the learning literature.

Our model is an overlapping generations endowment economy where money is the only store of value. The optimal decision of agents depends on their inflation forecasts, and so do the equilibrium outcomes. Under RE, the model has two steady states: high inflation and low inflation. To select one of the two steady states, we consider a learning model where agents forecast inflation using a rule that is a convex combination of past expected inflation and actual inflation. We examine the learning dynamics under two specifications: (i) Agents know only the past prices, and (ii) agents know the current price in addition to past prices. The two specifications imply different values for actual inflation in the learning rule. Under (i), agents use last period’s inflation rate, whereas under (ii), they use the current inflation rate. Thus, the only difference between the two specifications is that the value of actual inflation is current in one specification and lagged by just one period in the other.

Our main result is that the stability criterion selects qualitatively different equilibria even when the differences in the learning specifications are minor. In particular, the learning rule using last period’s inflation rate implies that the low-inflation RE steady state is the only stable learning equilibrium. Thus, using the stability criterion to select an RE equilibrium implies that the low-inflation steady state would be the long-run outcome. However, under the learning rule using the current inflation rate, both RE steady states are stable. Thus, the stability criterion does not offer useful guidance for equilibrium selection. In other words, our simple model shows that the stability of the learning equilibrium is sensitive to the specification of the learning rule. The learning dynamics may not be robust against seemingly minor differences in the learning rule.

Earlier work by Marcet and Sargent (1989) demonstrated how the stability criterion under RE dynamics selects the equilibrium that validates the “unpleasant monetarist arithmetic” in Sargent and Wallace (1981) but that the same criterion under least-squares learning selects another equilibrium that invalidates the unpleasant monetarist arithmetic. Our investigation of the sensitivity issue differs from that in Marcet and Sargent (1989) in an important way: Marcet and Sargent (1989) compare the equilibrium selection under two substantially different specifications—RE and learning. Under RE, the decisionmaker is a forward-looking rational
agent, while under the learning dynamics, the decisionmaker is a backward-looking bound-
edly rational agent. We, on the other hand, compare outcomes under two learning specifica-
tions. Under both learning specifications, we maintain the assumption that the decisionmaker
is a least-squares learner, but we change the timing of the observation of one variable by just
one period.

Our exercise is in the same spirit as the exercises by Hansen and Sargent (2007) that exam-
ine robustness of an equilibrium to small changes in specifications. Hansen and Sargent (2007)
endow the agents with a set of models instead of just one. The agents have a reference model
and entertain a small neighborhood of models around the reference model and respond by
choosing the model that performs the best against the worst possible state. Similarly, Cho and
Kasa (2017) consider a set of “nearby” learning models close to a benchmark model, but the
agent uses model averaging between the benchmark and the nearby models. In Cho and Kasa
(2015), the agents choose a model based on a specification test. In contrast to these articles,
our agents do not evaluate multiple models simultaneously and do not choose one based on
a pre-specified performance criterion. Instead, our agents consider only one model at a time.
We examine the equilibrium selected by the stability criterion in each model.

Section 2 sets up an overlapping generations model and characterizes the equilibrium
outcome under RE. In Section 3, we investigate the issue of equilibrium selection under learn-
ing. In particular, we demonstrate that the stability properties change with minor changes in
the learning specifications. Section 4 contains concluding remarks.

2 AN OVERLAPPING GENERATIONS MODEL OF INFLATION

To demonstrate our idea, we adopt a simplified version of the model in Sargent and
Wallace (1981). Time is discrete and indexed by \( t = 0, 1, 2, 3 \ldots \) In each period \( t \), a new gener-
ation of households is born and lives for two periods: \( t \) and \( t+1 \). An old generation also exists
at \( t = 0 \). Thus, at any point in time \( t \geq 0 \): There is a young and an old generation of households.
The population size of each generation remains constant over time and is normalized to unity.
Each agent in generation \( t \) has a logarithmic utility function with no discounting:

\[
U_t = \ln c_{1,t} + \ln c_{2,t},
\]

where \( c_{i,t} \) is consumption of the generation-\( t \) agent in \( i = 1, 2 \) period of their life. Each
generation-\( t \) agent is endowed with 2 and \( 2\lambda, \lambda \in (0,1) \), units of perishable consumption goods
when young and old, respectively. The only asset available is fiat money, which is denoted by
\( m_t \). The flow budget constraints of a generation-\( t \) agent in periods \( t \) and \( t+1 \) are given, respec-
tively, by

\[
p_t c_{1,t} + m_t \leq 2 p_t,
\]

and

\[
p_{t+1} c_{2,t} \leq 2\lambda p_{t+1} + m_t,
\]

where \( p_t \) is the price level in period \( t \).
The government finances its expenses by issuing fiat money. The nominal government budget constraint is given by

\[ \xi p_t = M_t - M_{t-1}, \]

where \( \xi > 0 \) is the constant real government expenditure and \( M_t \) is the aggregate money supply in period \( t \).

The timing is as follows. In each period, the old agents enter with the nominal balances from the previous period. The young agents make their consumption and saving decisions. The government purchases goods by injecting money. Finally, consumption takes place based on realized prices at the end of the period.

The consumption of the initial old generation is determined by their budget constraint:

\[ c_{2,0} = 2\lambda + \frac{m_0}{p_0}, \]

where the initial nominal money balance, \( m_0 \), is exogenously given.

Note that aggregate gross domestic product (GDP) in each period is pinned down by the endowments and is \( 2 + 2\lambda \).

### 2.1 RE Equilibria

We first demonstrate that there are two steady states under RE. Given the deterministic setup, agents of each generation know the entire sequence of prices. Given the prices in \( t \) and \( t+1 \), using the flow budget constraints of a generation-\( t \) agent, we can write the lifetime budget constraint as

\[ c_{1,t} + \frac{p_{t+1} c_{2,t}}{p_t} \leq 2 + 2\lambda \frac{p_{t+1}}{p_t}. \]

The problem of a generation-\( t \) agent is

\[ \max \ln c_{1,t} + \ln c_{2,t} \]

subject to (2). The flow budget constraints and the choice of consumption determine the real balances of generation-\( t \) agents.

An RE equilibrium is a sequence of quantities and prices—\( c_{1,t}, c_{2,t}, M_t, p_t, \Pi_t \)—for \( t = 0, 1, \ldots, \infty \) such that agents in each generation choose consumption and real balances optimally, and the asset market and goods market clear in every period.

Note that goods market clearing implies that aggregate consumption in each period equals aggregate GDP and is \( 2 + 2\lambda \). However, as we show below, the distribution of consumption across generations and real balances in each period are affected by the path of inflation.

The first-order conditions for the above maximization problem can be summarized as

\[ \frac{c_{1,t}}{c_{2,t}} = \frac{p_{t+1}}{p_t}, \]

which together with (2) imply the optimal (interior) choices are
\[ c_{1,t} = 1 + \frac{\lambda}{P_{t+1}} \text{ and } c_{2,t} = \frac{P_t}{P_{t+1}}. \]

Therefore, the saving of generation \( t \) is \( \frac{m_t}{P_t} = 2 - c_{1,t} = 1 - \frac{\lambda}{P_{t+1}}. \) Since fiat money is the only store of value, that saving must be in the form of real money balances:

\[ \frac{M_t}{P_t} = 1 - \frac{\lambda}{P_{t+1}} = 1 - \lambda \Pi_{t+1}, \]

where \( \Pi_{t+1} \equiv \frac{P_{t+1}}{P_t} \) denotes the actual inflation rate between periods \( t \) and \( t+1 \).

If the inflation rate exceeds \( \frac{1}{\lambda} \), the real rate of return on money, \( \frac{P_t}{P_{t+1}} \), is "too low" and the young agent would like to borrow, not save. However, this is impossible in a two-period overlapping generations setup. Therefore, for inflation rates greater than or equal to \( \frac{1}{\lambda} \), the young agent would just consume their endowment. Hence, for money to be held (i.e., for real balances to be positive) the inflation rate must be less than \( \frac{1}{\lambda} \). Thus, the demand for money can be written as

\[ (3) \quad \frac{M_t}{P_t} = \max \left( 0, 1 - \lambda \Pi_{t+1} \right). \]

The asset-market-clearing condition implies that money supplied in each period must equal the money demand in that period. Rewriting equation (1) as

\[ \frac{M_t}{P_t} = \frac{M_{t-1}}{P_{t-1}} \frac{1}{\Pi_t} + \xi, \]

we can substitute money demand (3) into the above equation and get

\[ (4) \quad \max \left( 0, 1 - \lambda \Pi_{t+1} \right) = \max \left( 0, 1 - \lambda \Pi_t \right) + \xi. \]

Consider the case where \( \Pi_t < \frac{1}{\lambda} \) for all \( t \). The real balances are positive, so equation (4) can be simplified as

\[ \Pi_{t+1} = 1 + \frac{1 - \xi}{\lambda} - \frac{1}{\lambda} \Pi_t. \]

In a steady state, \( \Pi_{t+1} = \Pi_t = \Pi \) so the above equation can be rewritten as

\[ (5) \quad \lambda \Pi^2 - \left( \lambda + 1 - \xi \right) \Pi + 1 = 0, \]

which is a quadratic equation in \( \Pi \). Figure 1 illustrates equation (4).

If \( \xi < \lambda - 2\lambda^2 + 1 \), then the equation has two real distinct solutions, which represent two steady-state values of \( \Pi \). Denote them as \( \Pi_1^{RE} \) and \( \Pi_2^{RE} \) and let \( \Pi_1^{RE} < \Pi_2^{RE} \) without loss of
Figure 1
RE Equilibrium

NOTE: The steady-state inflation values are 1.02 and 1.08. The parameter values are $\xi = 0.0015$ and $\lambda = 0.9078$.

Figure 2
Real Balances: Path to the Stable RE Steady State

NOTE: The high-inflation steady state value is 1.08. The parameter values are $\xi = 0.0015$ and $\lambda = 0.9078$. The real money balance is 0.196 in the stable high-inflation steady state. For the simulation path, the initial condition is $\Pi_0 = 1.05$. 
generality. It is straightforward to show that \(1 < \Pi_1^{RE} < \Pi_2^{RE} < \frac{1}{\lambda}\). Clearly, as shown by Figure 1, the low-inflation steady state, \(\Pi_1^{RE}\), is unstable and the high-inflation steady state \(\Pi_2^{RE}\) is stable.

In Figure 2, we illustrate the path of real balances converging to the stable high-inflation steady state. In the next section, we investigate the issue of equilibrium selection using the stability criterion under learning.

### 3 EQUILIBRIUM SELECTION UNDER LEARNING

Suppose that agents do not have perfect foresight and they forecast prices based on information available to them. Generation-\(t\) agents make their optimal choices based on their expectation of \(p_{t+1}\). Let \(\Pi_{t+1}^{e} = \frac{p_{t+1}^{e}}{p_t}\) denote the expected inflation rate between periods \(t\) and \(t+1\), where the superscript “\(e\)” denotes the expected value. The learning rule we consider is a convex combination of past expected inflation and actual inflation:

\[
(6) \quad \Pi_{t+1}^{e} = \Pi_t^{e} + \gamma \left[ \text{Actual inflation} - \Pi_t^{e} \right],
\]

where \(\gamma \in (0,1)\) is the gain parameter. The value of actual inflation used in the learning rule depends on the information available to the agents. We examine two specifications in the next two subsections. Given the learning rule, the lifetime budget constraint of a generation-\(t\) agent is

\[
c_{1,t} + \Pi_{t+1}^{e} c_{2,t} \leq 2 + 2\lambda \Pi_{t+1}^{e}.
\]

A learning equilibrium is a sequence of quantities, prices, and forecasts—

\(c_{1,t}, c_{2,t-1}, \frac{M_t}{p_t}, p_t, \Pi_t, \Pi_t^{e}\) for \(t = 0, 1, \ldots, \infty\) such that agents in each generation choose consumption and real balances optimally based on their forecast of inflation, the asset market and goods market clear in every period, and the dynamics of expected inflation satisfy equation (6).

Aggregate GDP and consumption in each period in a learning equilibrium are the same as in an RE equilibrium: \(2 + 2\lambda\).

Similar to (3), the demand for real balances under learning is

\[
(7) \quad \frac{M_t}{p_t} = \max \left(0, 1 - \lambda \Pi_t^{e} \right).
\]

Again, the real money demand has a zero lower bound: If expected inflation exceeds \(\frac{1}{\lambda}\), then the expected real return on money is too low and the young agents will not hold any money.

As in Section 2.1, using equation (1), we get \(\frac{\delta L_t}{p_t} = \frac{M_{t-1}}{p_{t-1}} \frac{1}{\Pi_t} \Pi_t^{e} + \xi_t\). Substituting money demand into this equation, we get the asset-market-clearing condition:
\[ \max\left(0, 1 - \lambda \Pi_{t+1}^e\right) = \max\left(0, 1 - \lambda \Pi_t^e\right) \frac{1}{\Pi_t^e} + \xi. \]

Under the assumption of \( \Pi_t^e < \frac{1}{\lambda} \) for all \( t \), the above relationship yields the law of motion for inflation under learning:

\[ \Pi_t^e = \frac{1 - \lambda \Pi_t^e}{1 - \lambda \Pi_{t+1}^e - \xi}. \]

### 3.1 Actual Inflation Using Past Prices

Assume that the agents do not know the price \( p_t \), so they do not know the actual inflation \( \Pi_t \) between periods \( t - 1 \) and \( t \). When the agents forecast the inflation between periods \( t \) and \( t+1 \), the actual inflation used in forecasting \( \Pi_{t+1}^e \) is \( \Pi_{t-1}^e \), which is the most recent data available to the agents. The learning rule then is as follows:

\[ \Pi_{t+1}^e = \Pi_t^e + \gamma \left[ \Pi_{t-1}^e - \Pi_t^e \right]. \]

Note that lack of knowledge of \( p_t \) does not affect the demand for real balances, since equation (7) implies that the demand depends on the ratio of prices.

Using equation (8) to substitute for \( \Pi_{t-1}^e \) in the learning rule above, we get the law of motion for expected inflation \( \Pi_{t+1}^e \) under learning:

\[ \Pi_{t+1}^e = \Pi_t^e + \gamma \left[ 1 - \lambda \Pi_t^e \right] \left[ \frac{1}{1 - \lambda \Pi_{t+1}^e - \xi} - \Pi_t^e \right]. \]

Three features of equation (10) are worth noting. First, the steady-state version of (10) is identical to the RE steady-state equation (5). Hence, the learning dynamics imply the same two steady states as under RE: \( \Pi_1^{RE} \) and \( \Pi_2^{RE} \). Second, equation (10) is not tractable since it is a second-order nonlinear difference equation, which does not have analytical solutions. Finally, we can show that, under this learning rule and conditional on a small enough \( \gamma \), the low-inflation steady state, \( \Pi_1^{RE} \), becomes stable and the high-inflation steady state, \( \Pi_2^{RE} \), becomes unstable (see Chien, Cho, and Ravikumar, 2021).

Hence, selection using the stability criterion implies that the low-inflation steady state should be the long-run equilibrium outcome of this model. This result is also true in Marcet and Sargent (1989). Furthermore, since the stable steady state under RE is different from that under learning rule (9), the paths of real balances converging to the respective stable steady states would be different. Figure 3 illustrates the path of real balances converging to the low-inflation steady state under learning rule (9). Starting from “similar” initial conditions, the real balances increase monotonically to the stable state under learning, but decline monotonically under RE.

In the next subsection, we show that the equilibrium selection using stability is sensitive to the learning specification.
3.2 Actual Inflation Using Current and Past Prices

Now suppose that when agents forecast inflation \( \Pi_{t+1} \) they do know the price \( p_t \) and, hence, are able to use the actual inflation \( \Pi_t \) to update their forecast. Hence,

\[
\Pi_{t+1} = \Pi_t^e + \gamma \left[ \Pi_t - \Pi_t^e \right].
\]

Note that the only difference between equations (9) and (11) is the actual inflation used in the formula: \( \Pi_{t-1} \) in (9) and \( \Pi_t \) in (11).

Using equation (8) to substitute for \( \Pi_t \) in (11), we get an alternative law of motion for expected inflation \( \Pi_{t+1}^e \):

\[
\Pi_{t+1}^e = \Pi_t^e + \gamma \left[ \frac{1 - \lambda \Pi_t^e}{1 - \lambda \Pi_{t+1}^e - \xi} - \Pi_t^e \right].
\]

In the steady state, equation (12) is reduced to RE steady-state equation (5). Hence, \( \Pi_{1,RE}^e \) and \( \Pi_{2,RE}^e \) remain the two possible steady states under learning rule (11). However, other features of equation (12) are quite different from those of equation (10). First, equation (12) is a first-order nonlinear equation and its solution is analytically tractable. Second, despite having the same set of steady states, both steady states are stable. To see this, Figure 4 illustrates learning dynamics (12). For a given \( \Pi_t^e \), equation (12) is a quadratic equation in \( \Pi_{t+1}^e \) and, hence,
Figure 4

Learning Dynamics Using Current Inflation

NOTE: The steady-state inflation values are 1.02 and 1.08. The parameter values are $\xi = 0.0015$, $\lambda = 0.9078$, and $\gamma = 0.05$. The values of steady-state inflation are identical to those in the RE equilibrium.

Figure 5

Real Balances: Path to the Stable Steady States Under Learning Using Current Inflation

NOTE: The steady-state inflation values are 1.02 and 1.08. The parameter values are $\xi = 0.0015$, $\lambda = 0.9078$, and $\gamma = 0.05$. The values of steady-state inflation are identical to those in the RE equilibrium. The steady-state values for real balances are 0.0741 and 0.0196. For the simulation path, the initial condition is $\Pi_0 = 1.05$. 
has two possible solutions. For any \( \Pi'_e \), the red and yellow lines of Figure 4 illustrate the solutions of high and low expected inflation rates, \( \Pi'_{e+1} \), respectively. As indicated by the figure, the slopes of both the red and yellow lines are less than 1 around their corresponding steady-state values. Therefore, both steady states are stable and the stability criterion under learning does not help select an equilibrium outcome.

Figure 5 illustrates the paths of real balances to the two steady states. These outcomes are qualitatively different from the path in Figure 3. While the stability criterion implies a unique path for real balances under learning specification (9), the same criterion implies multiple paths under specification (11).

4 CONCLUDING REMARKS

Stability under learning is often considered an important criterion for equilibrium selection, especially when there are multiple RE equilibria. In a simple overlapping generations model with high and low RE steady-state inflation rates, we use stability of the learning equilibrium as a criterion for equilibrium selection. Our results show that the stability of learning dynamics is sensitive to the specifications of the learning rule. When agents use last period’s inflation in the learning rule, the low-inflation RE steady state is stable, while the high-inflation steady state is not. When agents use current inflation in the learning rule, both steady states are stable. The notion that such a seemingly minor variation in the specification could lead to qualitatively different learning equilibria questions the validity of using stability as a criterion for equilibrium selection.

NOTES

1. According to Evans and Honkapohja (2001, pg. 13-14), “Another advantage of the learning approach arises in connection with the issue of multiple equilibria...Throughout the book the multiplicity issue will recur frequently, and we will pay full attention to this role of adaptive learning as a selection criterion.”

2. Esponda and Pouzo (2016) formally extend the notion of misspecification from exogenous to endogenous data-generating environments.

3. For equation (10) to describe a learning equilibrium, we have to impose an additional restriction that \( \Pi'_{e+1} < \frac{1}{\lambda} \). If any element in the sequence of \( \Pi'_e \)'s exceeds \( \frac{1}{\lambda} \), then (10) cannot be used to recover future expected inflation.

REFERENCES


