Responses of International Central Banks to the COVID-19 Crisis
Jacob Haas, Christopher J. Neely, and William R. Emmons

The Economic Effects of the 2018 U.S. Trade Policy: A State-Level Analysis
Ana Maria Santacreu and Makenzie Peake

Why Is the Labor Share Declining?
Sangmin Aum and Yongseok Shin

Fertility and Internal Migration
Pei-Ju Liao, Ping Wang, Yin-Chi Wang, and Chong K. Yip
339
Responses of International Central Banks to the COVID-19 Crisis
Jacob Haas, Christopher J. Neely, and William R. Emmons

385
The Economic Effects of the 2018 U.S. Trade Policy: A State-Level Analysis
Ana Maria Santacreu and Makenzie Peake

413
Why Is the Labor Share Declining?
Sangmin Aum and Yongseok Shin

429
Fertility and Internal Migration
Pei-Ju Liao, Ping Wang, Yin-Chi Wang, and Chong K. Yip
Review

Review is published four times per year by the Research Division of the Federal Reserve Bank of St. Louis. Complimentary print subscriptions are available to U.S. addresses only. 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/).

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.

Subscriptions and Alerts

Single-copy subscriptions (U.S. addresses only) are available free of charge. Subscribe here: https://research.stlouisfed.org/publications/review/subscribe/.

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.

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 760,000 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.

© 2020, Federal Reserve Bank of St. Louis.
ISSN 0014-9187
Responses of International Central Banks to the COVID-19 Crisis

Jacob Haas, Christopher J. Neely, and William R. Emmons

This article reviews and explains the recent policy reactions of the Federal Reserve, the European Central Bank, the Bank of England, and the Bank of Japan to the financial and macroeconomic turmoil caused by the COVID-19 pandemic. The financial and monetary policy actions of major central banks in the most recent crisis have, by some metrics, surpassed their responses to the Global Financial Crisis of 2007-09 in both swiftness and scope. (JEL E58, E59, G01, E44, G15)


1 INTRODUCTION

Concerns about the spread of the COVID-19 virus and its effect on economic activity produced one of the most turbulent periods of financial market activity in history from mid-February to April 2020. In some ways, this episode resembled the events of September-October 2008, during the Global Financial Crisis (GFC) of 2007-09. Severe problems arose suddenly in international financial markets. Uncertainty and fear raised volatility and led investors to broadly sell risky assets, which reduced their prices.

International authorities responded to both the GFC and the recent COVID-19-inspired crisis. Fiscal authorities increased spending, including providing much more generous unemployment benefits, while central banks made credit more widely available in financial markets and supported markets for illiquid securities. Regulators allowed banks to reduce their capital and liquidity buffers and encouraged lenders to work with borrowers.

But there are significant differences between the two events, too. The events of September 2008 resembled a bank run. Investors became uncertain about the value of some types of risky housing assets—mortgage-backed securities (MBS) and other asset-backed securities (ABS) not guaranteed by the government and collateralized debt obligations—and therefore sought to reduce their exposure to all such assets and the firms that might own them or have guaranteed them.
Fear and uncertainty also drove events in 2020, but this anxiety focused not on opaque financial relationships but on the impact of COVID-19—a true negative supply shock but also a demand shock—on economic activity and employment. For 2020, the International Monetary Fund (IMF) projects an 8 percent decline in the gross domestic product (GDP) of advanced economies and a 3 percent decline in that for emerging market economies. If this occurs, the COVID-19 recession will be much deeper than the 2007-09 recession, when the GDP of advanced economies declined by 3.3 percent and that for emerging markets economies grew by 2.8 percent (IMF, 2020b).

Policy responses to the two episodes differed because the shocks were fundamentally different. In the COVID-19 crisis, authorities did not immediately seek to broadly stimulate the economy to put people back to work in March-June because some social isolation was necessary to reduce the spread of COVID-19. Instead, authorities concentrated on maintaining the health of the financial system, supporting individuals in isolating themselves but keeping long-term economic relationships intact. In the United States, these policies included Small Business Administration and Federal Reserve (Fed) support for small businesses.

Benefitting from their previous experience, policymakers responded much faster to the incipient COVID-19 crisis than to the GFC of 2007-09. In particular, they introduced new programs and reintroduced old ones more rapidly. The breadth of the policy response was also, in some cases, beyond any previous crisis response. For example, the introduction of direct Fed lending to businesses, states, and cities in the COVID-19 crisis is unprecedented in the United States, as is the likely eventual size of the Fed’s balance sheet (Timiraos and Hilsenrath, 2020). In a departure from the previous episode, all major central banks felt themselves to be in broadly similar situations. At the start of the GFC, in 2007 to 2008, leaders of the European Central Bank (ECB) and the Bank of Japan (BOJ) thought their economies were somewhat insulated and that the financial crisis was mostly a problem for the United States and United Kingdom. In 2020, the crisis was truly global from the outset.

This article examines how the major central banks—the Fed, the ECB, BOJ, and the Bank of England (BOE)—responded to the financial market turbulence of February-April 2020 and the anticipated plunge in economic activity. We summarize how and why these central banks implemented credit, asset-purchase, and banking support programs and compare and contrast the international policy responses. While future developments may overshadow those we document here, we hope this article will be useful to future readers who wish to understand the policy responses as they were rolled out in the early months of the COVID-19 crisis. We also touch briefly on fiscal and regulatory responses, particularly their interactions with central bank policies.

The next section of the article discusses central bank financial market and regulatory policies, while Section 3 describes the recent COVID-19 financial turmoil that contributed to an unprecedented economic downturn. Section 4 details the policy responses of the four major central banks and compares their efforts. Section 5 draws conclusions.
2 POLICY TOOLS FOR FINANCIAL TURMOIL AND RECESSIONS

The three macroeconomic tools are monetary, fiscal, and financial policies. We briefly review the separate use of these tools by central banks before considering their interaction in specific episodes. In doing so, we often refer specifically to the Fed’s tools. Other central banks possess similar tools.

2.1 Monetary Policy

Former Federal Reserve Chair Ben Bernanke (2012b) described a central bank’s main policy tools as monetary policy and lender-of-last-resort powers. Both types provide credit to the private sector. There is no bright line between the policies, but lender-of-last-resort actions tend to be narrowly focused and short-lived, while monetary policy has broader and longer-lasting impacts.

Central banks conduct monetary policy to stabilize the economy—that is, to decrease fluctuations in economic activity. When short-term interest rates are positive, central banks in developed economies typically conduct monetary policy by adjusting short-term interest rates to influence broad financial conditions, including longer-term interest rates, monetary and credit aggregates, other asset prices, volatility, risk premia, and liquidity. The Fed, for example, normally uses open market operations to maintain an overnight interbank interest rate—the federal funds rate—within a target range that it deems appropriate for economic conditions.3

By influencing asset prices, central banks can affect consumption, investment, production, employment, and inflation.

When short-term interest rates neared the zero lower bound in 2008, central banks, including the Fed, employed unconventional monetary policy tools on a previously untested scale.4 These tools consisted of large-scale asset purchases (also called quantitative easing [QE]) and forward guidance—communications by central bank leaders of their views about the economy and how their policy tools were likely to evolve over time.5 Such tools can be used to influence the broader economy through changes in asset prices, just as short-term interest rates can. To combat the GFC, in 2007-09, central banks expanded their toolkits beyond previous limits and laid the groundwork for many recent policy measures.

Surveying the literature evaluating the international use of new monetary tools, Bhattarai and Neely (forthcoming) find that a variety of types of studies provide strong evidence that broad central bank asset purchases moved domestic asset prices and spilled over to move international prices, while narrow asset purchase programs (in a particular market) and bank lending support programs normalized market functioning and facilitated intermediation. In addition, central banks learned about the degree to which it was possible to impose modestly negative interest rates.6 Central bankers also generally consider that such unconventional programs substantially improved financial conditions, raising growth and avoiding deflation (Bhattarai and Neely, forthcoming).

2.2 Fiscal Policy

Government spending and/or taxes constitute fiscal policy. Fiscal actions that change the current budget balance (i.e., increase or decrease the budget deficit) may affect the economy,
but it is difficult to predict and measure this effect.\textsuperscript{2} Balanced-budget fiscal policy measures—that is, equal increases in spending and tax revenues—are likely to have smaller effects.

Central banks can change the cost of government debt or even the ability of governments to borrow. Central bank purchases of government debt are said to monetize that debt because the central bank creates money to pay for the government debt it retires. Monetization becomes a problem if it is perceived to remove a constraint on government spending and make the debt unsustainable in the longer run.

Another way that central banks may affect fiscal policy is via credit allocation. A central bank purchase of MBS, for example, reduces credit costs for mortgage borrowers relative to other types of borrowers. Similarly, central bank lending to small businesses or purchases of corporate or municipal securities benefit particular borrowers. These are transfers of value to particular agents and therefore are fiscal policy.
2.3 Financial Policy

Government or central bank financial policies can affect individual financial institutions or markets or infrastructure (i.e., trading venues, information networks, and payments and settlement systems) without necessarily affecting broad financial conditions. These policy actions include the following:

- discount window lending: short-term collateralized loans to individual depository institutions,
- emergency lending under Federal Reserve Act Section 13(3) authority as amended in the Dodd-Frank Act: creation and operation of funding, credit, liquidity and loan facilities that are available to a broad set of counterparties; and
- supervision and regulation of financial institutions.

When central banks engage in discount window lending or create specialized credit facilities, they act as lenders of last resort to particular institutions or markets. As the name suggests, a lender of last resort provides short-term loans to banks when private lending is unavailable on reasonable terms.

Walter Bagehot (1873) prescribed similar central bank lending practices in *Lombard Street*: During a panic, central banks should lend freely against good assets at a high interest rate. Bernanke (2012a) argues that lender-of-last-resort powers are the primary tool a central bank has to maintain or restore financial stability and that Bagehot’s dictum on their use remains valid today.

Interventions targeted at particular financial institutions, markets, or utilities may change broad financial conditions. For example, substantial emergency lending could influence interest rates in other sectors, creating a monetary policy action. If that is not intended, the central bank could sterilize (i.e., neutralize) the wider impact of the emergency lending through offsets setting open market operations. For example, if a central bank provided $100 billion to some sector in a special lending program, it could neutralize the action’s broader monetary policy effects by selling $100 billion of Treasury securities in open market operations.

In the United States, the Department of the Treasury typically coordinates and backstops unusual Fed lending, that is, non-discount-window lending, because loan losses are transfers of value to debtors. That is a fiscal-policy action and, therefore, should be decided by Congress.

2.4 A Two-Front War: The Economy and the Financial System

Severe financial crises are rare in U.S. history but have occurred more frequently in other countries (Reinhart and Rogoff, 2009). A nationwide banking panic in 1931-33 extended and deepened the Great Depression. A severe financial meltdown in 2008 likewise turned a mild, incipient downturn into the Great Recession. The current U.S. recession began in March 2020, the same month severe financial turmoil emerged (National Bureau of Economic Research, 2020).

Although recessions can occur in the absence of a financial crisis, the real economy and the financial sector are interdependent. Concern about the real economy can produce volatility that inhibits financial activity, which feeds back to the real economy. Such cycles challenge
President Obama signed the Dodd-Frank Wall Street Reform and Consumer Protection Act into law on July 21, 2010. The act is a complex piece of legislation, consisting of 16 separate sections (see table). The preamble to the act states its purposes:

• to promote the financial stability of the United States by improving accountability and transparency in the financial system,
• to end “too big to fail,”
• to protect the American taxpayer by ending bailouts,
• to protect consumers from abusive financial services practices, and
• for other purposes.

Although many Dodd-Frank provisions were designed to address shortcomings in the financial system demonstrated by the GFC of 2007-09, the overall structure of the law closely followed a proposal laid out in March 2008 by Barack Obama (2008) in a campaign speech in New York. Obama identified these objectives for financial reform:

• expand the Fed’s authority to supervise systemically important financial institutions;
• stiffen existing capital and liquidity rules;
• consolidate the regulatory structure;
• practice functional, rather than institutional, regulation;
• crack down on trading abuses; and
• identify and address systemic risks.

Arguably, the Dodd-Frank Act realized these objectives, with the exception of consolidating the regulatory structure.* Perhaps as a result of some of the changes arising from the act, such as higher bank capital requirements and more disclosure about financial market activity, banks and the broader financial system have remained strong in the early stages of the current recession. As we describe in greater detail in the article, however, some of this resilience also is due to timely large-scale interventions by the Fed and Treasury to preempt emerging credit-market dysfunction in March 2020.

*Although the Office of Thrift Supervision was eliminated, with its authorities and responsibilities transferred primarily to the Office of the Comptroller of the Currency, two new financial regulatory agencies were created: the Bureau of Consumer Financial Protection and the Office of Financial Research. Moreover, the Dodd-Frank Act established the Financial Stability Oversight Council as a framework to bring together existing financial regulators to monitor and address financial-stability risks.
# Dodd-Frank Act: Titles and Major Provisions

<table>
<thead>
<tr>
<th>Title</th>
<th>Purpose</th>
<th>Major provisions</th>
</tr>
</thead>
</table>
• Increases Fed oversight of nonbank financial institutions.  
• Creates Collins Amendment (Section 171) requiring higher bank capital levels.  
• Requires banks to establish “living wills.” |
| II    | Orderly liquidation authority | • The Federal Deposit Insurance Corporation (FDIC) gains non-bank resolution authority. |
| III   | Transfer of powers to the Comptroller of the Currency, the FDIC, and the Board of Governors | • Abolishes the Office of Thrift Supervision (OTS).  
• Transfers OTS powers to the Comptroller of the Currency, the FDIC, and the Board of Governors. |
| IV    | Regulation of advisers to hedge funds and others | • Clarifies requirements for covered investment advisers to provide information to the Securities and Exchange Commission and the FDIC. |
| V     | Insurance | • Establishes a Federal Insurance Office within the Department of the Treasury. |
| VI    | Improvements to regulation of depository institutions | • Closes loopholes such as the use of industrial loan companies.  
• Volcker Rule (Section 619) restricts proprietary trading by banks. |
| VII   | Wall Street transparency and accountability | • Reforms over-the-counter derivatives regulation. |
| VII   | Payment, clearing, and settlement supervision | • Designates certain market utilities as systemically important and subject to supervision. |
| IX    | Investor protections and improvements to the regulation of securities | • Requires risk retention by originating institutions of ABS.  
• Clarifies the liability of rating agencies. |
| X     | Bureau of Consumer Financial Protection | • Creates the Bureau of Consumer Financial Protection  
• Consolidates from other federal agencies and extends existing consumer financial-protection powers. |
| XI    | Federal Reserve System provisions | • Establishes limits on the use by the Board of Governors of its Section 13(3) authority under the Federal Reserve Act. |
| XII   | Improving access to mainstream financial institutions | • Encourages alternatives to payday loans. |
| XIII  | Pay It Back Act | • Establishes procedures to wind down the TARP. |
| XIV   | Mortgage Reform and Anti-Predatory Lending Act | • Prohibits yield-spread payments to mortgage originators.  
• Establishes new rules on high-cost mortgages, appraisals, servicing, and modifications. |
| XV    | Miscellaneous provisions | • Restricts trade in “conflict minerals.” |
| XVI   | Section 1256 contracts | • Tightens the definitions of tax reporting of gains and losses from mark-to-market adjustments and straddles. |

(Source: H.R. 4173 (111th Congress): Dodd-Frank Wall Street Reform and Consumer Protection Act; [https://www.govtrack.us/congress/bills/111/hr4173/text](https://www.govtrack.us/congress/bills/111/hr4173/text).)
Uncertainty about economic events can produce “financial accelerator” effects through its effect on asset prices. Falling asset prices reduce a firm’s creditworthiness by lowering the net worth of the firm’s owners and the value of collateral it can pledge to lenders. Owners with little equity value have an incentive to borrow money to take risky but potentially profitable actions, because the owners keep any gains but cannot lose more than their (small) equity stake in the firm. Lenders are reluctant to lend to such firms (Bernanke and Gertler, 1989, and Calomiris and Hubbard, 1990). Such a sudden retraction of credit or increase in its cost can disrupt the operation of financial firms, such as specialists on Wall Street and some hedge funds, that borrow most of their operating funds. The potential for bankruptcy or disruption of credit for some financial firms jeopardizes the complex system of payments that links all financial institutions. A loan default by a bankrupt firm might render the firm’s creditors insolvent as well, for example.

Disruption to financial market functioning can exacerbate recessions by inhibiting intermediation, that is, credit for investment and trade (Schularick and Taylor, 2012). A large falloff in business investment is virtually synonymous with recession in U.S. data. While disruptions of other industries can be contained, preserving the financial system is crucially important to the broader economy because credit, payments mechanisms, and settlement systems underpin virtually all transactions. Their importance is most evident when they are disrupted, as in September 2008 and March 2020. Despite severe shocks in March, a full-blown financial crisis has not yet resulted. Nonetheless, history may judge that only aggressive central bank interventions prevented this outcome. The Fed rolled out many distinct financial interventions within weeks, several of which were reincarnations from the 2007-09 crisis (see Table 1A). Congress backed these Fed actions by appropriating $454 billion of U.S. Treasury funds to absorb the possibly substantial risk involved. Such actions have often prompted accusations that central banks try to prop up asset prices, protecting stockholders from losses. The shaded insert “Central Bank Responses and the ‘Greenspan Put’” discusses this issue.
### Table 1A

**Central Bank Asset Purchases and Lending Facilities: Federal Reserve Programs**

<table>
<thead>
<tr>
<th>Program</th>
<th>Abbreviation</th>
<th>Targeted sector</th>
<th>Description</th>
<th>Capacity ($)</th>
<th>Initial announcement date (in 2020)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary Dealer Credit Facility</strong>*</td>
<td>PDCF</td>
<td>Primary dealers</td>
<td>Offers term funding to primary dealers with maturities up to 90 days at the discount rate. A broader range of securities are eligible as collateral than are eligible for open market repo operations.</td>
<td>—</td>
<td>March 17</td>
</tr>
<tr>
<td><strong>Paycheck Protection Program Liquidity Facility</strong></td>
<td>PPPLF</td>
<td>Small and medium-sized enterprises (SMEs)</td>
<td>Extends loans at 35 basis points to PPP lenders, taking the PPP loans as collateral at face value. Limited by PPP size</td>
<td>April 6</td>
<td></td>
</tr>
<tr>
<td><strong>Main Street Lending Program</strong></td>
<td>MSLP</td>
<td>SMEs and nonprofits</td>
<td>Offers loans to SMEs at LIBOR plus 3 percent. Loans have a 5-year maturity, and interest and principal payments are deferred for 1 and 2 years, respectively. New-loan facilities allow lenders to originate new loans while keeping a 5 percent stake. An existing-loan facility allows lenders to increase the size of existing loans. $75 billion of support from Treasury’s Exchange Stabilization Fund (ESF). Also includes facilities designed to lend to nonprofit organizations. Up to $600 billion</td>
<td>March 23</td>
<td></td>
</tr>
<tr>
<td><strong>Money Market Mutual Fund Liquidity Facility</strong>*</td>
<td>MMLF</td>
<td>Money market mutual funds</td>
<td>Offers loans to institutions to buy assets that money market mutual funds are selling at rates between the discount rate and the discount rate plus 100 basis points. $10 billion of support from Treasury’s ESF.</td>
<td>—</td>
<td>March 18</td>
</tr>
<tr>
<td><strong>Term Asset-Backed Securities Loan Facility</strong>*</td>
<td>TALF</td>
<td>Asset-backed securities</td>
<td>Offers loans with 3-year maturity to issuers of highly rated asset-backed securities backed by consumer and small-business loans. $10 billion of support from Treasury’s ESF. Up to $100 billion</td>
<td>March 23</td>
<td></td>
</tr>
<tr>
<td><strong>Primary/Secondary Market Corporate Credit Facilities</strong></td>
<td>PMCCF, SMCCF</td>
<td>Corporate debt</td>
<td>PMCCF: Purchases certain new bond and syndicated loan issuances. SMCCF: Purchases certain outstanding corporate bonds, ETFs investing in corporate bonds, and broad market index corporate bond portfolios. $75 billion of support from Treasury’s ESF. Up to $750 billion</td>
<td>March 23</td>
<td></td>
</tr>
<tr>
<td><strong>Municipal Liquidity Facility</strong></td>
<td>MLF</td>
<td>Municipal debt</td>
<td>Purchases short-term notes with maturities less than 3 years issued by U.S. states, large cities and counties, and other designated bond-issuing political entities. $35 billion of support from Treasury’s ESF. Up to $500 billion</td>
<td>April 9</td>
<td></td>
</tr>
<tr>
<td><strong>Commercial Paper Funding Facility</strong>*</td>
<td>CPFF</td>
<td>Commercial paper</td>
<td>Purchases 3-month unsecured and asset-backed U.S. commercial paper. $10 billion of support from Treasury’s ESF.</td>
<td>—</td>
<td>March 17</td>
</tr>
<tr>
<td><strong>Central bank USD liquidity swap lines</strong>*</td>
<td>—</td>
<td>USD liquidity</td>
<td>The Fed, in coordination with the BOE, BOJ, ECB, Bank of Canada, and Swiss National Bank, cuts pricing on standing USD liquidity swaps by 25 basis points. The foreign central banks with regular USD liquidity operations also agreed to increase the frequency of 7-day operations and offer USD with an 84-day maturity weekly. The Fed extended swap lines to 9 more central banks, with $30-$60 billion limits per country.</td>
<td>—</td>
<td>March 15</td>
</tr>
<tr>
<td><strong>Foreign and International Monetary Authorities Repo Facility</strong></td>
<td>FIMA Repo Facility</td>
<td>Foreign and international monetary authorities (FIMA)</td>
<td>Repo facility for FIMA, open to most FIMA account holders, to easily borrow USD at 25 basis points over the interest on excess reserves rate. Facility to run 6+ months.</td>
<td>—</td>
<td>March 31</td>
</tr>
<tr>
<td><strong>Discount window</strong>*</td>
<td>—</td>
<td>Banks</td>
<td>Lowers discount rate by 150 basis points, to 0.25 percent, for loans up to 90 days.</td>
<td>—</td>
<td>March 15</td>
</tr>
<tr>
<td><strong>Repurchase agreements</strong>*</td>
<td>—</td>
<td>Repo market</td>
<td>New York Fed program of expanded repurchase agreement operations.</td>
<td>—</td>
<td>March 9</td>
</tr>
<tr>
<td><strong>Asset purchases</strong>*</td>
<td>—</td>
<td>Treasuries, agency mortgage-backed securities (MBS)</td>
<td>Open-ended asset purchase program of Treasuries and agency MBS. Originally announced to be at least $500 billion in Treasuries and $200 billion in agency MBS before the open-ended announcement.</td>
<td>—</td>
<td>March 15</td>
</tr>
</tbody>
</table>

NOTE: Items with asterisks are similar to analogous programs used to provide support during the GFC of 2007-09 or prior to 2020.

SOURCE: Board of Governors, IHS Markit, and Federal Reserve Bank of New York.
## Table 1B

### Central Bank Asset Purchases and Lending Facilities: European Central Bank Programs

<table>
<thead>
<tr>
<th>Program</th>
<th>Abbreviation</th>
<th>Targeted sector</th>
<th>Description</th>
<th>Capacity (€)</th>
<th>Initial announcement date (in 2020)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Targeted longer-term refinancing operations*</td>
<td>TLTRO III</td>
<td>Bank lending to households and businesses</td>
<td>3-year operations to incentivize bank lending to households and businesses: interest rate reduced to 50 basis points below main refinancing operations rate (–0.5% but can be as low as –1% for some) and borrowing allowance increased.</td>
<td>—</td>
<td>March 12</td>
</tr>
<tr>
<td>Pandemic emergency longer-term refinancing operations</td>
<td>PELTROs</td>
<td>Bank lending</td>
<td>Seven additional lending programs with maturities between 8-16 months: interest rate at 25 basis points below main refinancing operations rate.</td>
<td>—</td>
<td>April 30</td>
</tr>
<tr>
<td>Additional longer-term refinancing operations*</td>
<td>LTROs</td>
<td>Bank lending</td>
<td>Series of long-term lending operations starting in March 2020, with June 2020 maturity designed to fill the gap until June TLTRO operations.</td>
<td>—</td>
<td>March 12</td>
</tr>
<tr>
<td>Eurosystem Repo Facility for Central Banks</td>
<td>EUREP</td>
<td>Euro liquidity</td>
<td>Provides repos to central banks outside the euro area, complementing existing swap and repo lines. The ECB also added new swap lines with the central banks of Croatia and Bulgaria and reactivates swap line with the central bank of Denmark.</td>
<td>—</td>
<td>June 25</td>
</tr>
<tr>
<td>Pandemic Emergency Purchase Programme</td>
<td>PEPP</td>
<td>Public, private securities</td>
<td>€1.35 trillion temporary asset purchase program of private and public sector securities. Expanded on June 4 from its original €750 billion envelope.</td>
<td>€1.35 trillion</td>
<td>March 18</td>
</tr>
<tr>
<td>Asset Purchase Programme*</td>
<td>APP</td>
<td>Public, private securities</td>
<td>€120 billion temporary envelope of asset purchases on top of €20 billion per month pace when the APP restarted in November 2019.</td>
<td>—</td>
<td>March 12</td>
</tr>
</tbody>
</table>

NOTE: Items with asterisks are similar to analogous programs used to provide support during the GFC of 2007-09 or prior to 2020.

SOURCE: ECB.
### Table 1C

**Central Bank Asset Purchases and Lending Facilities: Bank of England Programs**

<table>
<thead>
<tr>
<th>Program</th>
<th>Abbreviation</th>
<th>Targeted sector</th>
<th>Description</th>
<th>Capacity (£)</th>
<th>Initial announcement date (in 2020)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Term Funding Scheme with additional incentives for small and medium-sized enterprises</strong></td>
<td>TFSME</td>
<td>Bank lending to households and businesses</td>
<td>Offers long-term funding at or close to the bank rate, with additional funds available for banks that increase lending, especially to SMEs.</td>
<td>—</td>
<td>March 11</td>
</tr>
<tr>
<td><strong>Contingent Term Repo Facility</strong></td>
<td>CTRF</td>
<td>Repo market</td>
<td>Offers repo operations at 1- or 3-month terms, complementing existing lending facilities. 3-month operations discontinued at the end of May 2020, and 1-month operations discontinued at the end of June.</td>
<td>—</td>
<td>March 24</td>
</tr>
<tr>
<td><strong>Covid Corporate Financing Facility</strong></td>
<td>CCFF</td>
<td>Commercial paper</td>
<td>Purchases commercial paper of up to 1-year maturity at rate based on OIS rate, close to pre-COVID spreads.</td>
<td>—</td>
<td>March 17</td>
</tr>
<tr>
<td><strong>Asset purchases</strong></td>
<td>APF</td>
<td>Public, private securities</td>
<td>Increase of bond holdings by £300 billion, to £745 billion, mostly from purchases of U.K. government bonds but also from nonfinancial corporate bonds.</td>
<td>£300 billion</td>
<td>March 19</td>
</tr>
</tbody>
</table>

NOTE: Items with asterisks are similar to analogous programs used to provide support during the GFC of 2007-09 or prior to 2020. The CCFF is novel, but the BOE has bought commercial paper in the past.

SOURCE: BOE.
Table 1D
Central Bank Asset Purchases and Lending Facilities: Bank of Japan Programs

<table>
<thead>
<tr>
<th>Program</th>
<th>Abbreviation</th>
<th>Targeted sector</th>
<th>Description</th>
<th>Capacity (¥)</th>
<th>Initial announcement date (in 2020)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Securities Lending Facility*</td>
<td>SLF</td>
<td>Repo market</td>
<td>Provides JGS through repo operations. The BOJ increased the number of JGS issues offered in the SLF.</td>
<td>—</td>
<td>March 13</td>
</tr>
<tr>
<td>Special Funds-Supplying Operations to Facilitate Financing in Response to the Novel Coronavirus</td>
<td>SMEs</td>
<td></td>
<td>Provides loans based on the amount of private debt pledged as collateral and SME loans given in response to COVID-19, with maturity up to 1 year. Applies a 0.1% interest rate to current account balances that correspond to amounts of loans provided.</td>
<td>¥90 trillion</td>
<td>March 16</td>
</tr>
<tr>
<td>Corporate bond/commercial paper purchases*</td>
<td>Commercial paper, corporate bonds</td>
<td></td>
<td>Raises limits on holdings of commercial paper and corporate bonds by ¥7.5 trillion each, to ¥20 trillion total.</td>
<td>¥20 trillion</td>
<td>March 16</td>
</tr>
<tr>
<td>Asset purchases of Japanese government bonds and exchange-traded funds/Japanese real estate investment trusts*</td>
<td>JGBs, Treasury discount bills (T-bills), ETFs, and J-REITs</td>
<td></td>
<td>Increases purchases of JGBs and T-bills, with no upper limit on JGB purchases. Purchases ETFs and J-REITs with upper limits at double the previous purchase pace (up to ¥12 trillion and ¥180 billion, respectively).</td>
<td>—</td>
<td>March 16</td>
</tr>
</tbody>
</table>

NOTE: In addition to the programs listed that have been introduced or expanded as a result of COVID-19, other programs are still active and disbursing funds. These extant programs include the BOJ’s Loan Support Program (see https://www.boj.or.jp/en/mopo/measures/mkt_ope/len_b/index.htm/), consisting of the Stimulating Bank Lending Facility (SBLF) and the Growth-Supporting Funding Facility (GSFF). Fawley and Neely (2013) and Neely and Kason (forthcoming) detail these programs. Items with asterisks are similar to analogous programs used to provide support during the GFC of 2007-09 or prior to 2020.

SOURCE: BOJ.
3 THE COVID-19 FINANCIAL TURMOIL

This section first briefly discusses the spread of COVID-19 before reviewing the causes and symptoms of the recent COVID-19-related financial turmoil of February-April 2020.

3.1 The Global Spread of COVID-19

The first cases of COVID-19 were reported in December 2019 in Wuhan, China. The World Health Organization’s (WHO’s) Chinese Office learned of a media statement on the website of the Wuhan Municipal Health Commission on local cases of “viral pneumonia” (Ma, 2020). In the following weeks, the new virus spread to Thailand, Japan, South Korea, the United States, and France. The Chinese government locked down Wuhan and most of Hubei Province in late January 2020, halting outbound travel and public transit. By the end of January there were almost 10,000 confirmed cases of COVID-19 in 19 countries and over 200 deaths (WHO, 2020b).

Countries began implementing travel restrictions and lockdown measures as confirmed cases began to jump in the following weeks. On January 31, the United States suspended travel by foreign nationals from China, then in mid-March likewise suspended travel by foreign nationals from Europe. In late February, Italy locked down its hardest hit towns in the Lombardy region and then moved to a country-wide lockdown by March 9: schools and non-essential businesses were closed (Taylor, 2020). By the end of March, well over 100 countries had instituted a lockdown, and there were almost 800,000 confirmed cases and 40,000 deaths worldwide (BBC, 2020, and WHO, 2020a). The virus continued to spread through the spring and summer of 2020. Some areas saw lower caseload growth and mortality rates and were able to ease some of their lockdown measures, while other regions continued to struggle to contain the spread of the virus.

The loss of life and uncertainty of safety surrounding the pandemic took a devastating human toll but also greatly affected the global economy. Consumers were unable or unwilling to purchase goods and services, and the closure of non-essential businesses, suspension of schools, slowdown of travel and tourism, and cancellation of large public events caused by lockdowns all had major implications on economic activity and employment. As of July 2, 2020, there were over 10 million confirmed cases and 500,000 deaths as a result of COVID-19 (WHO, 2020c).

3.2 The Financial Market Turmoil

Financial crises are periods in which expectations of asset values or economic activity are suddenly revised in ways that dramatically change asset prices and threaten the stability of the economic system through asset price volatility and disrupted financial activity.

The spread of COVID-19 prompted fears of disruptions to supply chains and the labor force that would curtail economic activity and bring on a severe recession. Such fears prompted many investors to sell risky assets, such as high-yield corporate bonds, in favor of safer assets, such as Treasury securities. Prices of many risky assets fell, while volatility, trading volume, and bid-ask spreads rose. That is, there was a flight to safety. Yields on very safe 10-year
Figure 1
Corporate Bond, Mortgage, and Sovereign Spreads

A. U.S. corporate bond and mortgage spreads

B. Sovereign spreads

NOTE: Panel A shows the ICE BofA (Bank of America) AAA U.S. Corporate Index Option-Adjusted Spread (AAA OAS); the ICE BofA FNMA Current Coupon Effective Yield less the 10-year Treasury yield (MBS-Treasury spread); the 30-year mortgage-Treasury spread, which is the 30-year fixed mortgage rate less the 30-year constant-maturity Treasury yield; and the ICE BofA U.S. High Yield Option-Adjusted Spread (High-yield OAS). The vertical line in Panel A indicates the announcements of the PMCCF/SMCCF on March 23, 2020. Panel B shows 10-year sovereign yield spreads versus German bonds. The vertical line in Panel B indicates the announcement of the PEPP on March 18, 2020.

SOURCE: FRED®, Federal Reserve Bank of St. Louis; Haver Analytics; and ICE/Bank of America Merrill Lynch.
Treasuries fell from 1.88 percent in January 2020 to a historic low of 0.54 percent on March 9, and they remained near that level through May. Panel A of Figure 1 shows the rise in risk spreads over Treasury yields. The spread over Treasury securities of government-guaranteed MBS, a default-risk-free security, rose from late February to early March, reflecting heightened interest rate and liquidity risks (Mizrach and Neely 2020a,b,c,d,e). A selloff in high-yield grade (i.e., fairly risky) debt boosted the spreads of those securities to about 11 percent on March 23, in contrast to the 3 to 4 percent range common in early February 2020. Conditions briefly resembled those of September-October 2008 (Fawley and Neely, 2013).

Panel A of Figure 1 shows that most U.S. risk spreads peaked the week of March 16-23. The events of this week merit special explanation. The Federal Open Market Committee (FOMC) had lowered the federal funds target range by 50 basis points at an unscheduled meeting on March 3, but increasing financial market stress prompted repeated credit provision through repurchase agreements (repos) the week of March 9-14. Although the FOMC had a meeting scheduled for March 17-18, financial market dislocation prompted the committee to meet early, on the morning of Sunday, March 15, to be able to announce their plans prior to the opening of Asian markets in the U.S. evening. The FOMC announced purchases of at least $700 billion in Treasuries and MBS, although it would not classify those purchases as open ended until the end of that week. On the same day, six central banks—the Fed, the Bank of Canada, the BOE, the BOJ, the ECB, and the Swiss National Bank—announced new foreign exchange swap arrangements to provide liquidity in U.S. dollars (USD) internationally.

Despite these expansionary moves, however, international stock prices fell sharply—the Dow dropped nearly 3,000 points—on the morning of Monday, March 16 and circuit breakers halted trading in several markets. As discussed previously, falling asset prices reduce the net worth of the firms and individuals that own them, making them less creditworthy borrowers and discouraging lending.

It is not clear exactly how to interpret this negative stock market reaction. Markets may have seen the FOMC announcement as disappointingly timid or as an indication of surprisingly bad economic news. President Trump’s warnings on that Monday of a prolonged shutdown might have exacerbated the stock market reaction.

From March 16 to 23, however, the Fed would release plans for several new lending facilities to stabilize financial markets and make the previously announced asset purchases open ended. These facilities included the Primary Dealer Credit Facility (the PDCF), the Commercial Paper Funding Facility (CPFF), the Money Market Mutual Fund Liquidity Facility (MMLF), the Main Street Lending Program (MSLP), the Term Asset-Backed Securities Loan Facility (TALF), and the Primary and Secondary Market Corporate Credit Facilities (PMCCF and SMCCF). This combination of efforts seemed to stabilize markets in the latter part of March and April.

In Europe, the flight to safety took the form of rising sovereign yields for all countries, including Germany—whose bond yields serve as the low-risk benchmark—and a widening of yield spreads between high-risk countries, such as Italy, and Germany. Panel B of Figure 1 shows the widening of spreads versus Germany in early March, until the announcement of the Pandemic Emergency Purchase Programme (PEPP) on March 18, which is indicated by...
the vertical line. Differences in spread behavior largely reflected the perceived risk of default, which was likely higher for countries with high debt-to-GDP ratios and small size. According to 2018 OECD data, the debt-to-GDP ratios of Italy, Spain, France, the Netherlands, and Germany were 147 percent, 115 percent, 122 percent, 66 percent, and 70 percent, respectively.

The 3-month LIBOR-OIS spread is a popular measure of money market stress, as it measures the willingness of banks to make unsecured loans. Figure 2 shows that the 3-month LIBOR-OIS spread started increasing in early March for the monetary areas of each of the major central banks discussed here. It increased more in the United States and United Kingdom than in the euro area and Japan, indicating that investors might have perceived a greater risk of default by a U.S. or U.K. bank than one in the euro area or Japan. Alternatively, it might have reflected lower expectations of a government rescue of a defaulting bank in the United States or United Kingdom. The U.S. LIBOR-OIS spread was particularly high for the week of March 16-23, which was a particularly difficult week on U.S. financial markets, as discussed previously.

From mid-February to late March, stock prices fell rapidly from fairly high values after a 10-year bull market (Mizrach and Neely, 2020c). Figure 3 illustrates parallel declines in the prices of the S&P 500, Japanese Nikkei 225, German Dax, and U.K. FTSE All-Share indices. Implied volatility, as measured by the VIX (a forward-looking measure of volatility derived...
from options prices), started rising sharply on about February 19 and peaked on March 16 before declining to lower—but still historically high—levels (Panel A of Figure 4). This pattern is consistent with the negative correlation historically observed between uncertainty and stock returns. Increased uncertainty is also consistent with the sharp increase in equity market bid-ask spreads shown in Panel B of Figure 4. Market makers widen their quoted spreads in fast moving, uncertain markets to avoid taking losses to more-informed traders. Equity trading volume (not shown in the figure) began rising on February 20, peaked on February 28, but remained elevated through March (Mizrach and Neely, 2020c).

The flight to safety included a shift to relatively safe USD assets that appreciated the trade-weighted USD by 8 percent from late February to March 23 (Panel A of Figure 5). Such appreciation often occurs in times of crisis. Panel B of Figure 5 shows that the trade-weighted USD also rose after the September 11, 2001, attacks and in September 2008. As in other markets, bid-ask spreads widened and foreign exchange trading volume rose, reaching 540 billion USD per day on FXall, the most active interdealer trading network (Mizrach and Neely, 2020b).

Although recessions and (particularly) financial crises are sudden by nature, the real effects of the current downturn appeared unusually suddenly and with great severity. Weekly U.S. initial unemployment insurance claims surged from 282,000 on March 14 to 3.3 million on March 21 to a completely unprecedented 6.9 million on March 28 (Figure 6). The downturn was not evenly spread out across sectors: The leisure and hospitality and health sectors were
Figure 4
Financial Market Volatility

A. Stock index and Treasury note futures volatility

Volatility, February 3, 2020 = 1

- VIX
- 10-year Treasury volatility futures

B. Bid-ask spreads for individual stocks

Bid-ask spread, February 3, 2020 = 1

- Microsoft
- AstraZeneca PLC
- ASML Holding NV
- Fast Retailing Co.

NOTE: Panel A shows two forward-looking measures of financial market volatility obtained from options markets: the VIX and the volatility index on 10-year Treasury futures contracts. For ease of comparison, the volatility indices are indexed to their values on February 3, 2020. Panel B shows bid-ask spreads for the companies that make up the largest portion of the S&P 500, FTSE 100, Euro Stoxx 50, and Nikkei 225, respectively, as of July 2020.

SOURCE: Bloomberg and FRED®, Federal Reserve Bank of St. Louis.
Figure 5
The Foreign Exchange Value of the USD

A. Normalized foreign exchange value of the USD

Units of foreign currency per USD

- TW broad
- EUR
- JPY
- CAD
- GBP
- MXN
- CNY

B. Trade-weighted value of the USD

Trade-weighted USD index, broad

NOTE: Panel A shows the normalized foreign exchange value of the USD against a broad trade-weighted basket (TW broad), the Japanese yen (JPY), the British pound (GBP), the Chinese renminbi (CNY), the euro (EUR), the Canadian dollar (CAD), and the Mexican peso (MXN). All series are normalized to equal 1 on February 18, 2020. Panel B shows the trade-weighted value of the USD over a much longer sample. Vertical lines indicate September 11, 2001, and September 15, 2008, the day of the Lehman Brothers collapse. The series in Panel B is constructed by splicing two broad currency indexes on January 2, 2006.

SOURCE: Haver Analytics.
Figure 6
Unemployment in the Four Monetary Areas

A. Unemployment rates, 2006-20

B. Unemployment rates, June 2019-June 2020

C. Weekly U.S. unemployment claims, January 2006-June 2020

NOTE: U.K. unemployment data are through April 2020.
SOURCE: FRED®, Federal Reserve Bank of St. Louis.
among the hardest hit (Franck, 2020). Complicating matters, U.S. households had limited savings going into the episode. As of 2018, only 61 percent of U.S. adults could pay off an unexpected expense of $400 by the end of the month (Board of Governors of the Federal Reserve System [Board of Governors], 2019c). Fed Chair Jerome Powell said that recovery could take more than a year (Timiraos, 2020).

Economies around the world were hit hard nearly simultaneously. IMF World Economic Outlook real GDP projections from the June 2020 issue, shown here in Table 2, provide a consistent format with which to compare likely conditions. The second column shows that the advanced economies grew sluggishly in 2019, at a 1.7 percent rate. The third column illustrates the very substantial IMF-projected declines in real GDP in 2020: –4.9 percent growth in the world, –8.0 percent in advanced economies, –8.0 percent in the United States, –10.2 percent in the euro area, –5.8 percent in Japan, and –10.2 percent in the United Kingdom. These GDP projections would be record-breaking negative numbers if they come to pass. The IMF further expects these economies to strongly rebound at (generally) 4 to 6 percent rates in 2021.

### Table 2

<table>
<thead>
<tr>
<th>Area</th>
<th>2019</th>
<th>2020</th>
<th>2021</th>
</tr>
</thead>
<tbody>
<tr>
<td>World</td>
<td>2.9</td>
<td>–4.9</td>
<td>5.4</td>
</tr>
<tr>
<td>Advanced economies</td>
<td>1.7</td>
<td>–8.0</td>
<td>4.8</td>
</tr>
<tr>
<td>United States</td>
<td>2.3</td>
<td>–8.0</td>
<td>4.5</td>
</tr>
<tr>
<td>Euro area</td>
<td>1.3</td>
<td>–10.2</td>
<td>6.0</td>
</tr>
<tr>
<td>Japan</td>
<td>0.7</td>
<td>–5.8</td>
<td>2.4</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>1.4</td>
<td>–10.2</td>
<td>6.3</td>
</tr>
<tr>
<td>Emerging markets</td>
<td>3.7</td>
<td>–3.0</td>
<td>5.9</td>
</tr>
</tbody>
</table>

December 31: Wuhan government confirms treatment of dozens of cases of (later named) COVID-19.


January 21: First confirmed case of COVID-19 in the U.S.

January 23: Chinese government locks down Wuhan, halts outbound flights, and shuts down public transit.

January 24: First cases of COVID-19 reported in Europe (in France).

January 31: Almost 10,000 COVID-19 cases are reported in 19 countries. U.S. suspends travel by most foreign nationals who had been to China in the past 14 days.

February 14: First case of COVID-19 reported in Africa (in Egypt).

February 23: Over 85,000 COVID-19 cases reported worldwide.

February 29: Over 85,000 COVID-19 cases reported worldwide.

March 3: Fed lowers the federal funds target range by 50 basis points, to 1-1.25%.

March 9: Italy goes on lockdown.

March 11: BOE cuts its bank rate by 50 basis points to 0.25%, introduces the TFSME, releases the countercyclical capital buffer, and issues guidance for banks not to increase dividends or other distributions.

March 12: ECB announces €120 billion in additional net asset purchases. ECB also announces relaxed capital and liquidity buffers, additional longer-term refinancing operations, and more favorable terms to the TLTRO III.

March 13: BOJ increases JGS in SLF, conducts unscheduled JGB purchases.

March 13: ECB announces €120 billion in additional net asset purchases. ECB also announces relaxed capital and liquidity buffers, additional longer-term refinancing operations, and more favorable terms to the TLTRO III.

SOURCE: BBC, BOE, BOJ, ECB, Fed, Taylor (2020), and WHO.
March 15: Fed lowers the target range for the federal funds rate, to 0-1/4%, announces at least $500 billion in Treasuries purchases and $200 billion in agency MBS purchases, lowers the discount rate, and drops reserve requirements to zero.

The Fed, in coordination with the BOE, BOJ, ECB, Bank of Canada, and Swiss National Bank, lowers pricing on USD swaps. The foreign central banks with regular USD liquidity operations agree to increase the frequency of 7-day operations and offer USD with an 84-day maturity weekly.

March 16: BOJ continues purchases of JGBs and introduces Special Funds-Supplying Operations to Facilitate Financing in Response to the Novel Coronavirus.

BOJ also increases purchases of commercial paper and corporate bonds by ¥2 trillion total and sets upper limits of ETF and J-REIT purchases at ¥12 trillion and ¥180 billion annually (double previous target purchase pace).

March 17: BOE announces the CCFF.

March 18: ECB announces €750 billion in PEPP purchases, expands the range of eligible corporate sector assets for purchase, and expands the scope of additional credit claims.

March 19: Fed extends limited USD swap lines to nine more central banks.

March 23: U.K. enters lockdown, closing nonessential shops and banning large gatherings.

March 24: BOE announces the CTRF, complementing existing lending facilities.

March 25: ECB increases PEPP flexibility to allow it to buy more than one-third of a country’s eligible bonds, expands criteria for eligible securities.

March 31: Fed establishes the FIMA repo facility.

March 17: Fed announces the PDCF and CPFF.

March 18: Fed announces the MMLF.

March 19: BOE cuts its bank rate from 0.25% to 0.1%, increases the TFSME, and announces £200 billion of net gilt and corporate bond purchases.

March 23: Fed expands asset purchases to agency commercial MBS and makes purchase program open ended.

Fed also establishes the PMCCF, SMCCF, and TALF; announces intentions to create the MSLP; and expands the MMLF and CPFF.

March 24: BOJ extends increase in JGS issues offered in the SLF.

March 27: ECB issues guidance that banks should refrain from share buybacks and not pay dividends for fiscal years 2019 and 2020.

SOURCE: BBC, BOE, BOJ, ECB, Fed, Taylor (2020), and WHO.
Figure 7, cont’d
Timeline of Select Central Bank Actions

April 6: Fed announces the PPPLF.

April 9: BOE announces temporary extension of Ways and Means facility as a source of short-term government liquidity.

April 16: Japan declares a nationwide state of emergency and allows regional governments to urge people to stay inside but without legal force.

April 27: Fed expands the MLF.

April 30: ECB further lowers TLTRO III interest rates and announces non-targeted PELTROs.

April 30: Fed expands the MSLP and PPPLF.

May 22: BOJ introduces fund-provisioning measure to support SMEs, offering loans up to 1 year at a 0% rate and paying financial institutions 0.1% interest on the loans.

May 2: BOE allows TFSME participants to extend the term of some of their funding to align with the U.K. Treasury’s Bounce Back Loan Scheme.

June 8: Fed expands MSLP to allow more SMEs to be eligible and extends loan terms to 5 years.

June 4: ECB increases PEPP envelope by €600 billion, to €1.35 trillion, extends horizon for net purchases to at least June 2021.

June 16: BOJ increases Special Funds-Supplying Operations (including its funds-supplying measure to support SMEs) to ¥90 trillion, from ¥55 trillion.

June 15: Fed expands the SMCCF to buy a diverse set of individual corporate bonds to complement ETF purchases.

June 18: BOE increases its target purchases for U.K. government bonds by an additional £100 billion.

June 25: ECB establishes EUREP to provide repo lines to central banks outside the euro area.


SOURCE: BBC, BOE, BOJ, ECB, Fed, Taylor (2020), and WHO.
4 CENTRAL BANK POLICY RESPONSES TO COVID-19

Figure 7 provides a timeline of events and central bank policy responses, while Table 1 details the programs.

4.1 Policy Rate Cuts

The most common central bank action to stabilize financial markets and protect the functioning of the broader economy is to reduce short-term interest rates. Central banks also responded to the COVID-19 episode with short-rate cuts. Lower interest rates tend to raise asset prices, make credit more affordable, and increase the creditworthiness of borrowers.

Unfortunately, due to relatively low growth and low inflation since the GFC, short rates were already relatively low when the financial turmoil started in February 2020. As central bankers would say, there was not much conventional “policy space” (Bernanke, 2020). ECB and BOJ policy rates were already near zero and remained unchanged. Short-term interest rates were 1.5 to 1.75 percent and 0.75 percent in the United States and United Kingdom, respectively. In March, both the Fed and BOE lowered rates toward zero as the coronavirus spread and economic uncertainty rose. Panel A of Figure 8 provides a long-run perspective on policy rates prior to the current crisis, while Panel B details the specific rate cuts of the Fed and BOE.24

4.2 Broad Asset Purchase Programs

With limited ability to further reduce short-term policy rates, central banks turned to long-term bond purchases to broadly reduce long-term yields, narrow asset purchases to aid the functioning of specific markets, and lending programs to expand credit for specific sectors.25 These actions have expanded central bank balance sheets at record rates in both percentage terms and as a percentage of GDP (Figure 9). Table 1 details these programs.

Broad asset purchases are thought to lower yields through a combination of local supply, duration, and signaling effects. Bhattarai, Eggertsson, and Gafarov (2015, 2019) rationalize how central bank asset purchases can “signal” lower future short-term interest rates in a time-consistent manner and thereby reduce current long rates. Removal of duration risk and local supply can produce portfolio rebalancing that reduces required yields. The literature on portfolio balance effects goes back to Tobin (1958, 1969), while more-recent contributions to this line of thought include Andrés, López-Salido, and Nelson (2004) and Vayanos and Vila (2009).

On March 12, the ECB’s Governing Council announced a €120 ($135) billion set of purchases, which it followed on March 18 by announcing €750 ($844) billion in purchases through the PEPP. Panel B of Figure 1 shows that this announcement appears to have immediately lowered yield spreads versus German bonds. Italian yield spreads fell over 75 basis points within days, and Spanish yield spreads fell almost 50 basis points. The ECB further expanded PEPP purchases on June 4 to €1.35 ($1.52) trillion.

On March 15, 2020, the Fed announced that it would purchase at least $500 billion in Treasuries and $200 billion in agency MBS but soon followed with a March 23 announcement that made the purchase amounts open ended and added agency commercial MBS purchases.
Figure 8
Central Bank Policy Rates

A. Central bank policy rates, 2006-20

B. Central bank policy rates, February 2020-May 2020

NOTE: Panel A: The data are from January 2006 to May 2020. The key policy rates for the BOE, Fed, ECB, and BOJ are, respectively, the official bank rate, the federal funds target rate, the main refinancing operations rate, and the uncollateralized overnight call rate. Starting in December 2008, the Fed began targeting a federal funds range rather than a target rate. Between April 2013 and February 2016, the BOJ did not set a target for the uncollateralized overnight call rate. Starting in February 2016, the BOJ resumed targeting a short-term interest rate, for which we report the BOJ’s basic balance rate, which is part of a tiered system of interest rates. Panel B: bps, basis points. The figure displays interest rate cuts in early 2020. The graph uses the midpoint of the federal funds target range.

Figure 9

Central Bank Assets

A. Central bank assets, 2006-20

Index, January 2007 = 100

B. Central bank assets as percentage of GDP, 2006-20

Percent

NOTE: Panel A shows asset holdings for the Fed, ECB, BOE, and BOJ, normalized to equal 100 in January 2007. Panel B shows asset holdings for the Fed, ECB, BOE, and BOJ, as a percentage of their respective nominal GDP through May 2020. Monthly GDP are interpolated from quarterly values, and 2020:Q2 GDP data are estimated based on OECD forecasts (at https://data.oecd.org/gdp/nominal-gdp-forecast.htm#indicator-chart).

SOURCE: BOE, BOJ, ECB, Fed, Haver Analytics, and OECD.
The BOJ has boosted its Japanese government bond (JGB) purchases, intending to purchase as necessary to keep 10-year yields around 0 percent. The BOJ has also augmented purchases of commercial paper, corporate bonds, exchange-traded funds (ETFs) and Japanese real estate investment trusts (J-REITs). The BOE likewise voted to purchase £200 ($250) billion in public and private bonds on March 19, then raised its purchase target by an additional £100 ($125) billion on June 18.

Figure 10 illustrates that the initial pace of central bank purchases was often unusually fast compared with past episodes. Fed purchases of Treasuries peaked at $75 billion per day in late March, far outstripping the peak pace of $120 billion per month during the previous crisis.
Haas, Neely, Emmons

**Figure 11**

Change in Total Assets, GFC and 2020 COVID-19 Crisis

<table>
<thead>
<tr>
<th>A. Fed</th>
<th>B. ECB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in assets as percentage of GDP</td>
<td>Change in assets as percentage of GDP</td>
</tr>
<tr>
<td>Fed-GFC</td>
<td>ECB-GFC</td>
</tr>
<tr>
<td>Fed-2020</td>
<td>ECB-2020</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>C. BOE</th>
<th>D. BOJ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in assets as percentage of GDP</td>
<td>Change in assets as percentage of GDP</td>
</tr>
<tr>
<td>BOE-GFC</td>
<td>BOJ-GFC</td>
</tr>
<tr>
<td>BOE-2020</td>
<td>BOJ-2020</td>
</tr>
</tbody>
</table>

NOTE: The figures display the monthly change in total assets for each central bank since the beginning of the GFC and the 2020 crisis as a percentage of nominal GDP. August 2008, the month prior to the collapse of Lehman Brothers, is the reference month for the GFC. February 2020 is the reference month for the COVID-19 crisis. Monthly GDP values are interpolated from quarterly values, and 2020:Q2 GDP data are estimated from OECD forecasts: [https://data.oecd.org/gdp/nominal-gdp-forecast.htm#indicator-chart](https://data.oecd.org/gdp/nominal-gdp-forecast.htm#indicator-chart).

SOURCE: BOE, BOJ, ECB, Fed, Haver Analytics, and OECD.

(Haas and Neely, 2020, Figure 10). The BOE purchased £13.5 ($16.9) billion of gilts per week from early April until late June, the fastest pace in the history of the BOE Asset Purchase Facility (APF). The BOJ ramped up corporate bond, commercial paper, and ETF purchases from April through June and increased its Japanese government securities (JGS) holdings, mainly driven by an increase in Treasury bill (T-bill) purchases. The ECB also bought bonds at a faster pace: In 2020:Q2, the PEPP and Asset Purchase Programme (APP) combined conducted more net asset purchases per month than at any point since the start of the APP (Figure 10). These purchases have sharply increased the size of each of the central banks’ balance sheets (Figures 9 and 11).
Philip Lane, chief economist of the ECB, reports that the ECB’s APP has focused on high-risk, high-yield countries, such as Italy, to stifle rising risk premia in its yields lest those increases cancel out stimulative effects (Skolimowski, 2020). Lane emphasizes the importance of flexibility in asset purchases over time and across jurisdictions. The ECB’s need to choose how to allocate purchases across jurisdictions is a unique challenge for that central bank.

A prominent concern about the extensive asset purchases is that central banks may be buying so much of particular markets—or at least so much of the flow of issued assets—that they are distorting the price signals that these markets send. For example, the BOJ has purchased about 45 percent of government debt and has increased its balance sheet to over 120 percent of Japanese GDP due to the strong pace of its asset purchases since 2013 (Panel B of Figure 9 and Figure 12). The BOJ is also authorized to purchase up to 45 percent of outstanding Japanese commercial paper and 15 percent of longer-term corporate bonds (Editorial Board, 2020). Each of the four central banks have purchased over 20 percent of their respective outstanding government debt (Figure 12). At the end of June 2020, the Fed owned about 20 percent of marketable Treasury debt outstanding and 33 percent of MBS issued by Fannie Mae and Freddie Mac (Board of Governors, 2020b; Fannie Mae, 2020; Freddie Mac, 2020; and

![Figure 12](image-url)

**Figure 12**

**Central Bank Ownership of Government Debt Outstanding**

NOTE: The figure shows the percentage of government debt securities outstanding owned by the respective central bank. Fed data are U.S. Treasury securities held outright as a percentage of total marketable publicly held Treasury securities outstanding. ECB data are Public Sector Purchase Programme (PSPP) and PEPP euro area public sector securities held as a percentage of total euro area government debt securities. BOJ data are BOJ JGS holdings (including T-bills) as a percentage of total JGS. BOE data are total APF gilt holdings at market value as a percentage of total stock of British government securities.

U.S. Department of the Treasury, 2020). Given central banks’ objectives to stabilize economies and financial markets, these large holdings may be unavoidable.

4.3 Narrow Asset Purchases

The Fed, BOE, BOJ, and ECB, have all purchased assets to ease functioning in specific markets; that is, they have conducted narrow asset purchases. Table 1 describes the narrow purchase programs and lending facilities the central banks established in the wake of COVID-19. Central banks have tended to focus their narrow purchases on corporate bonds, commercial paper, and MBS.

Theoretical work of De Pooter, Martin, and Pruitt (2018); Pasquariello, Roush, and Vega (2020); Greenwood and Vayanos (2010); and Corsetti and Dedola (2016) has attempted to rationalize price effects from recurring central bank bond market purchases on various grounds.

Corporate Bonds and Commercial Paper. All four major central banks have supported the corporate bond market and provided credit to larger firms during the crisis. The BOE, BOJ, and ECB all announced increased corporate bond purchases, and the BOE and ECB expanded the eligible range of their corporate purchases. The Fed, previously unwilling to purchase corporate bonds, established the PMCCF and SMCCF on March 23. The PMCCF allows the Fed to purchase qualifying corporate bonds at issuance, while the SMCCF was created to purchase investment-grade bonds of U.S. companies and U.S.-listed ETFs that provide broad exposure to those bonds.26 The U.S. Treasury purchased equity in the special purpose vehicle (SPV) established for the PMCCF and SMCCF with resources from the Exchange Stabilization Fund (ESF).27 Panel A of Figure 1 shows that these announcements preceded substantial

![Figure 13 Pending Home Sales](image-url)

**Figure 13**
Pending Home Sales

NOTE: The figure shows the year-over-year percent change in the National Association of Realtors Pending Homes Sales Index from May 2019 to May 2020. Data are seasonally adjusted and indexed to equal 100 in 2001.

SOURCE: Haver Analytics, National Association of Realtors.
declines in corporate yields, as much as 3 percentage points for the high-yield OAS (option-adjusted spread) after the March 23 PMCCF/SMCCF announcements, indicated by the vertical line.

The BOJ and ECB announced increased commercial paper outright purchases, while the Fed and BOE established facilities to purchase commercial paper. On March 17, the Fed reestablished its CPFF, while the BOE announced the Covid Corporate Financing Facility (CCFF) to support U.K. firms by buying commercial paper on terms similar to those prior to the COVID-19 crisis.

**MBS TBA.** The housing market is a cyclically volatile sector and was a particular target of the Fed’s first unconventional asset purchase program in 2008-10, that is, “QE1.” Despite the unusual source of the COVID-19 crisis, housing was one of the first sectors to be affected. Pending home sales slowed significantly during March and April as bad news about the spread of COVID-19 accumulated and the stock market turned down (Figures 3 and 13). To provide liquidity and facilitate trading of agency MBS, the March 15 FOMC announcement included plans to purchase at least $200 billion in agency MBS, concentrated on recently produced coupons in 30-year and 15-year fixed-rate agency MBS in the to-be-announced (TBA) market. The FOMC expanded these purchases on March 20 and then made them open ended on March 23.

**J-REITs and Equity ETFs.** In addition to public securities and corporate debt, the BOJ has also been purchasing J-REITs and equity ETFs. On March 16, the BOJ doubled the annual purchase limits of J-REITs and ETFs and has been accumulating these assets at record rates.

### 4.4 Lending Facilities

Asset purchases greatly increased bank reserves, but these safe liquid assets are in high demand and unevenly distributed among banks; so, particular markets may still lack short-term credit. According to Fed Chair Powell, investor risk aversion rose so sharply that extraordinary interventions by the Fed were required:

As a more adverse outlook for the economy associated with COVID-19 took hold, investors exhibited greater risk aversion and pulled away from longer-term and riskier assets as well as some money market mutual funds. To help stabilize short-term funding markets, we lengthened the term and lowered the rate on discount window loans to depository institutions. The Board also established, with the approval of the Treasury Department, the Primary Dealer Credit Facility (PDCF)…the Commercial Paper Funding Facility, or CPFF, and the Money Market Mutual Fund Liquidity Facility, or MMLF. Both of these facilities have equity provided by the Treasury Department to protect the Federal Reserve from losses. Indicators of market functioning in commercial paper and other short-term funding markets improved substantially and rapid outflows from prime and tax-exempt money market funds stopped after the announcement and implementation of these facilities. (Powel 2020a)

Table 1 details the recently introduced lending facilities from the four central banks and the targeted sectors and provides links to information on the facilities’ purposes and operations. Many of these facilities were directly analogous to initiatives taken in the wake of the 2007-09 crisis.
To ensure the continued availability of consumer loans, on March 23, the Fed established
the TALF under Section 13(3) of the Federal Reserve Act. Through an SPV, the Fed will
initially loan up to $100 billion to issuers of AAA-rated ABS that are backed by receivables,
such as small-business loans, auto loans, credit card loans, and education loans. Although
these loans are nonrecourse, the value of the ABS (collateral) will protect the Fed against
losses to some degree. The U.S. Treasury’s ESF will provide the first $10 billion in capital to
cover any losses.

In addition to rebooting several credit/purchase facilities from the 2007-09 crisis, the
Fed has created new facilities to lend more than $2.3 trillion to municipalities, corporations,
and small and mid-sized enterprises (SMEs). Of these new facilities, the MSLP is one of the
most innovative. This program permits the Fed to use an SPV to purchase bank loans (95
percent of each loan) made to SMEs, which frees up the bank’s capital to make new loans.
The MSLP, along with the Paycheck Protection Program Liquidity Facility (PPPLF), is designed
to stimulate bank lending, similar in purpose to bank lending incentive programs that the
other three major central banks have operated, such as the BOE’s Term Funding Scheme (TFS),
the ECB’s targeted/pandemic longer-term refinancing operations (TLTROs/PELTROs), and
the BOJ’s Stimulating Bank Lending Facility (SBLF) and new Special Funds-Supplying
Operations to Facilitate Financing in Response to the Novel Coronavirus.

The BOE has restarted the TFS with an estimated £100 ($125) billion in funding for bank
lending but has added incentives for lending to SMEs. This TFSME program augments the
BOE’s new commercial paper and repo facilities. Similarly, the ECB has prioritized non-
financial lending by creating incentives for banks to lend to businesses and households in its
long-term lending program (TLTRO III). The BOJ has established new Special Funds-Supplying
Operations programs to aid corporate, SME, and other private lending and expanded the
Securities Lending Facility (SLF) to ensure repo market stability (see Table 1D). These pro-
grams are in addition to their previously established facilities to encourage bank lending to
the nonfinancial sector, the SBLF and the Growth-Supporting Funding Facility (GSFF).

An extraordinary aspect of the current situation is that the ECB and BOJ are actually
charging negative interest rates on bank borrowing. That is, these central banks are paying
banks to lend money (Kihara, Canepa, and Schneider, 2020). The ECB’s TLTRO III loans
banks money at –1 percent on the condition that they don’t reduce their loans. The BOJ is
not so generous but is still paying 0.1 percent to banks that participate in its new Special Funds-
Supplying Operations (Kihara, Canepa, and Schneider, 2020).

4.5 International Facilities

Central banks can lend money and purchase assets indefinitely because they can create
unlimited nominal amounts of their own currencies. The widespread use of the USD in inter-
national financial transactions gives the Fed a special role in international financial stability.
Foreign central banks often would like to make loans in USD to firms doing business in USD,
but they cannot create USD as they can their own domestic currencies. Foreign central banks
could obtain USD by selling their USD-denominated foreign exchange reserves, buying
USD with their domestic currency on foreign exchange markets, or borrowing USD, but
these actions might affect asset prices in undesirable ways, driving up U.S. yields or the value of the USD.

To meet this international demand for the USD, the Fed expanded its foreign exchange swap lines with other central banks and created a repo facility from which foreign and international monetary authorities (FIMA) can borrow USD from the Fed. Providing USD to foreign central banks allows them to function as a lender of last resort in USD and to mitigate the consequences of fire sales of risky assets. This practice indirectly aids the U.S. economy by improving conditions for U.S. trading partners and encouraging the international use of the USD. The Fed has used swap lines to help manage financial crises after September 11, 2001, and during the GFC of 2007-09, but the repo facility is novel. Table 1A details these facilities.

Since the establishment of the euro in 1999, financial markets have increasingly used that currency, and the ECB has also established new arrangements to provide increased access to it. On June 25, 2020, the ECB established the Eurosystem Repo Facility for Central Banks (EUREP), a new facility to provide repo lines to non-euro area central banks, complementing their existing swap and repo lines. This facility will be available until June 2021. The ECB also added new swap lines with the central banks of Croatia and Bulgaria and reactivated their swap line with Denmark’s central bank.

4.6 Regulatory Changes

In addition to their traditional roles as a lender of last resort and monetary authority, central banks usually have substantial regulatory responsibilities with which they can influence financial markets. International regulators quickly made important adjustments when the crisis hit (Board of Governors, 2020a, and IMF, 2020a,c):

- The Fed and other regulatory agencies encouraged U.S. banks to offer loan modifications to bank customers affected by COVID-19, stating that financial institutions generally do not need to categorize these modifications as troubled debt restructurings.
- National regulatory authorities in most European countries and Japan also encouraged banks to restructure loan terms and provide payment holidays when feasible.
- Regulators in all countries allowed banks to defer recognition of credit losses for purposes of regulatory capital computation that resulted from the ongoing introduction of a forward-looking approach to loan-loss provisioning.32
- U.S. supervisory agencies temporarily refocused on off-site monitoring of regulated institutions, suspending most regular examinations for institutions with less than $100 billion in assets.
- U.S. banks were allowed to exclude Treasury securities and deposits held at the Fed when calculating their capital requirements.
- The Fed brought reserve requirements down to zero.

The Fed’s 2020 supervisory stress test, designed to evaluate the resiliency of bank capital to severe economic and financial stress, proceeded as originally announced despite the reservation of some banks about the exercise. To gauge the impact of the unfolding crisis, the Fed required banks to conduct additional sensitivity analyses of three downside scenarios. The Fed
also tightened capital-distribution limits and required large banks to submit new longer-term capital plans (Board of Governors, 2020c). In contrast, the Bank of England cancelled its supervisory stress test (BOE, 2020c).

Like the Fed, Japan’s Financial Services Agency encouraged regulated financial institutions to work with borrowers in light of greatly increased financial stress (Financial Services Agency, 2020).

The BOE reduced the U.K. countercyclical buffer rate to 0 percent to release up to £190 ($238) billion of additional bank liquidity (BOE, 2020b). This buffer had previously been due to increase from 1 percent to 2 percent.  

The ECB focused on reducing banks’ financing costs, including costs of directly borrowing from the central bank and reducing banks’ required capital levels (Schnabel, 2020). Perhaps its most important supervisory action was to temporarily lower banks’ capital requirements for market risk and to allow banks to fully use their capital and liquidity buffers.

4.7 International Policy Comparison

Similarities. In many ways, the four major central banks responded very similarly. They each pushed policy rates toward zero—or held them at negative levels; implemented asset purchase programs; relaxed capital requirements; and instituted initiatives to support SMEs, the repo market, and corporate bond markets. The central banks have also moved together toward buying some nontraditional assets, such as commercial paper and corporate bonds. In general, the central banks have sought to ensure two-way asset markets (i.e., sufficient market liquidity to allow large purchases and sales to take place without moving prices unduly) and access to credit.

The responses of the central banks are notably similar, in contrast to the stark differences in their responses to the GFC. For example, the ECB did not start a large, broad asset purchase program until early 2015, years after other central banks implemented their programs, initially responding with the more narrowly purposed initiatives such as supporting sovereign bond markets (Fawley and Neely, 2013; Karson and Neely, forthcoming; and Altavilla, Carboni, and Motto, 2015). And unlike the other three central banks, the Fed did not buy corporate bonds until the 2020 crisis, nor did they previously use bank lending incentive schemes.

Differences. There are also policy differences, however. Prominently, the BOJ and ECB have embraced slightly negative deposit rates, while the Fed and BOE have not. Former Chair Bernanke (2016) outlines several reasons why the Fed did not turn to negative rates after the GFC of 2007-09: First, lacking much experience with negative rates, Fed economists may have underestimated the potential to push rates below zero. Burke et al. (2010) suggested that short rates below –30 to –35 basis points might induce a widespread conversion of bank reserves to cash. This seems to understate the potential of negative rates, as the costs associated with holding cash have allowed several central banks to push interest rates 50 or 100 basis points below zero. Second, former Fed Chair Janet Yellen said that it was not clear that the Fed can legally impose negative rates on reserve deposits (C-SPAN, 2016). Third, officials worried that negative rates might force (or induce) money market funds to “break the buck” and that ensuing widespread withdrawals would endanger the health of that part of the financial system (Neely, 2020).
More broadly, negative rates tend to compress bank profit margins, functioning as a tax on banks (Waller, 2016; Brunnermeier and Koby, 2018; and Bhattarai and Neely, forthcoming). Indeed, BOE Governor Andrew Bailey cited concerns over banks' net interest margins and a lessening effectiveness of cuts as rates approach their zero lower bound but would not rule out implementation of negative rates entirely (Douglas and Hirtenstein, 2020). Similarly, Schnabel (2020) cites the pressures that more negative rates might put on the financial system in explaining the ECB Governing Council’s decision to not push rates further below zero. And a former BOJ board member, Sayuri Shirai, has spoken out on the costs and limited benefits of negative rates (as quoted in Kihara, Canepa, and Schneider, 2020): "It doesn’t make sense to deepen negative interest rates and hurt banks when you’re actually trying to encourage them to lend more. It’s a tool that is very hard to use at a time like now.” Current BOJ Governor Haruhiko Kuroda agreed that he saw no need for deeper negative rates (as quoted in Kihara, Canepa, and Schneider, 2020): “At this moment, we don’t think it’s necessary.”

Perhaps even more radical than negative deposit rates are the negative interest rates that the ECB and BOJ are paying on borrowing; the Fed and BOE have so far not done so. Interest rates on the ECB’s TLTRO III program were lowered to –0.5 percent and can be as low as –1 percent for banks that meet a 0 percent lending growth threshold, essentially paying banks to not reduce their lending to the nonfinancial sector (European Central Bank, 2020). Similarly, the BOJ offers 0.1 percent interest to banks on the amount that the banks lend using the BOJ’s Special Funds-Supplying Operations, in effect compensating banks to lend to the nonfinancial sector (Kihara, Canepa, and Schneider, 2020, and Dreyer and Nygaard, 2020). In contrast, the Fed’s PPPLF and the BOE’s TFSME are priced at positive rates: at or slightly above the discount rate and the BOE’s bank rate, respectively.

**Lending Program Approaches.** Each central bank has incentivized bank lending to the nonfinancial sector in different ways. Unlike the other central banks, the Fed has directly funded lending to SMEs through the MSLP by purchasing 95 percent of each long-term loan provided by the program. By purchasing loans from banks, the MSLP removes a balance sheet constraint on further loans. Another way to encourage loans is to offer long-term, low-rate credit to banks based on their quantities of nonfinancial lending, as the BOE, BOJ, and ECB have done. The ECB’s TLTRO III and BOE’s TFSME provide interest rate incentives to banks to maintain their net nonfinancial lending (BOE, 2020b, and ECB, 2020). The TFSME also includes quantity incentives for higher SME lending. One difference is that the MSLP’s purchases of SME loans exposes the Treasury/Fed more directly to the risk of losses than do the bank incentive programs of the other central banks.

The central banks have worked with fiscal authorities in different ways. The U.S. Treasury has pledged up to $454 billion to cover losses from Fed lending, allowing the Fed to accept greater risk in lending. Similarly, the U.K. Treasury has pledged to cover any losses from the BOE’s CCFF and APF (Bailey 2020a,b). However, the ECB and BOJ have no explicit fiscal backstop for their lending programs, and the BOJ has already incurred losses on their ETF purchases, though all major central banks enjoy the implicit backing of the sovereign (Sano and Kaneko, 2020).

Differences in financial systems also played a role in why each central bank chose certain policy measures. For example, the Fed has instituted the MMLF to support the uniquely
important U.S. money market mutual fund sector and, as discussed in Section 4.5, the Fed and ECB’s foreign monetary authority repo facilities and foreign exchange swap lines result from the heightened role of the USD and the euro in the world economy.38

**Asset Purchases.** All four central banks purchased their respective government’s bonds as well as corporate bonds and commercial paper. While each of the four central banks purchased long-term bonds to reduce long yields, the potential magnitudes of the purchases differed. While the BOE and ECB set upper limits on their purchases, the Fed and BOJ employed open-ended asset purchase programs contingent on the state of the economy (Bullard, 2010).39 The BOJ implemented such an open-ended JGB-buying scheme despite the fact that it had not come close to hitting its previous annual JGB target of ¥80 trillion ($744 billion) in prior years (Fujikawa, 2019).

As mentioned in Section 4.3, the BOJ and Fed have expanded narrow purchases of additional asset types. The BOJ has been buying equity ETFs and J-REITs, while the Fed has been purchasing agency MBS.40

**4.8 Unwinding Crisis-Era Policies**

As the financial crisis recedes, central banks must exit from their crisis-induced policies. Extended periods of lax financial conditions contribute to the creation of “zombie” banks and other firms that survive only when interest rates are very low. Their survival inhibits the reallocation of resources and reduces growth (Caballero, Hoshi, and Kashyap, 2008, and Lee and Davis, 2020). A belated return to non-crisis monetary policy risks the emergence of fiscal dominance, whereby financial and monetary policies become subordinated to the financing of government expenditures. In other words, very low interest rates may be appropriate in a crisis, but they could also delay or prevent needed adjustments in both the private and public sectors.

To avoid this danger, many emergency credit-market measures are explicitly time-limited, although some can be extended. Some lending facilities are created with terms that will become unattractive as soon as normal market functioning returns, causing them to wind down naturally. The Fed’s MSLP, for example, is designed to be used only when no better market alternatives are available. Open-ended actions and programs, such as interest rate cuts and large-scale asset purchases, on the other hand, require detailed planning to exit.

During its dramatic balance sheet expansion following the GFC of 2007-09, in June 2011, the Fed began to communicate principles underlying its normalization strategies for interest rates and its balance sheet (Board of Governors, 2011). Even earlier, Fed Chair Ben Bernanke discussed how the Fed would wind down its special lending facilities (Bernanke, 2010). Jean-Claude Trichet, president of the ECB, likewise spoke in detail about exit strategies in late 2009 (Trichet, 2009).

Central banks generally significantly altered their initial timetables and targets, for example, as in Board of Governors (2014 and 2019a). But financial markets generally easily digested such changes because they were well communicated. One exception was the so-called “taper tantrum” in 2013, when Chair Bernanke’s public comments about the Fed tapering the pace of its asset purchases triggered a sell-off in bond markets (Neely, 2014).
Given that economic weakness may last longer than financial fragility, central banks must decouple the unwinding of emergency financial policies from expansionary monetary policy. Too-rapid withdrawal of emergency support for financial institutions could destabilize credit flows and market liquidity. On the other hand, failing to prune back financial sector support in a timely manner would distort financial incentives and give advantages to selected banks, other financial institutions, and investors.

From a longer-term perspective, emergency central bank policies that seek to bolster the financial system must be weighed against the aggravation of moral-hazard incentives that will outlast the crisis. Support given in the past crisis may be expected in the next one, affecting asset prices and risk-taking behavior in the interim. The risk of a public backlash against the financial sector and policymakers seen to support it at the expense of the broader public interest can result in legislative changes that reduce the ability of central banks to act in future emergencies.

5 DISCUSSION AND CONCLUSION

This article has reviewed the sudden financial turmoil associated with the outbreak of COVID-19 and detailed the central bank policy responses. Uncertainly about economic activity created financial conditions that were similar to or even worse in some ways than the conditions during the GFC of 2007-09. The initial real consequences of this pandemic were unusually sudden and severe, with U.S. unemployment rising to levels unseen since the Great Depression.

Although the cause of the COVID-19 crisis was quite unlike that of the GFC of 2007-09, many symptoms were similar. For example, both crises spawned flights to safety that produced fire sales of risky assets. Many of the policies to alleviate those symptoms were also similar in kind, but central banks responded with unusual speed and vigor as a result of their post-2008 experience with unconventional policies.

The Fed and BOE cut short-term interest rates, while the ECB and BOJ maintained rates that were already at or below zero. All four central banks introduced or expanded broad asset purchases, special bank-lending facilities, and narrow asset purchase facilities. Extraordinarily, the ECB and BOJ are actually paying negative interest rates on bank borrowing. With the USD and euro playing important roles on international financial markets, the Fed and the ECB expanded swap lines and created repo facilities for international monetary authorities. These measures seemed to be largely successful in maintaining the functioning of financial markets.

The crisis has prompted unprecedented cooperation between fiscal and monetary authorities. Congress appropriated $454 billion for Treasury injections of capital to Fed programs, while the U.K. Treasury offered indemnity on BOE asset purchase and CCFF losses (Timiraos and Hilsenrath, 2020, and Bailey 2020a,b). The Fed, BOJ, and BOE have also designed some of their lending programs to support fiscal authority initiatives. For example, the Fed’s PPPLF and BOJ’s Special Funds-Supplying Operations for SMEs provide funding for banks that use government SME lending programs, encouraging banks to use these fiscal policy programs
and lend more to SMEs. Such cooperation may make central bankers nervous about retaining their independence.

Although the major central banks have taken unusually bold steps to achieve their mandates by supporting financial markets and economic activity, they are not out of ammunition. If they thought it appropriate, all the central banks could make more aggressive use of contingent forward guidance to shape expectations about the future path of short-term interest rates. The Fed could reopen the TAF and provide greater incentives for bank lending. The Fed and the BOE could follow the ECB and BOJ and push rates negative. Otherwise, the other central banks could follow the BOJ in controlling the yield curve. It is not clear, however, that any of these actions would be appropriate to achieve the legally mandated objectives of these central banks, particularly in an unusual crisis in which a certain amount of economic inactivity is desired for public health purposes.

NOTES

1. Brinca, Duarte, and Faria-e-Castro (2020) break sectoral shocks down into supply and demand shocks, finding that about two-thirds of the negative shocks were due to supply factors. Guerrieri et al. (2020) develop a theory of Keynesian supply shocks, that is, supply shocks that trigger larger changes in aggregate demand.

2. See IMF (2020d). For up-to-date information on policy responses of governments around the world to the COVID-19 crisis, see IMF (2020c).

3. The Bank for International Settlements (2020) lists central bank policy rates. Not all central banks influence short-term interest rates with open market operations. The ECB and many other central banks conduct policy by adjusting interest rates for lending to or borrowing from banks. In some countries, particularly emerging markets, the central bank conducts monetary policy by targeting the country’s exchange rate rather than an interest rate.

4. The term “zero lower bound” was commonly used in the past to describe the zero level below which central banks did not push interest rates. As some central banks—not including the Fed—have pushed policy rates below zero, the term “effective lower bound” has become more commonly used. This term refers to the rate (usually negative) below which monetary policy cannot push interest rates or the rate below which such reductions are no longer stimulatory. The effective lower bound is likely to differ across countries.

5. Forward guidance was not a new tool in 2008-09. For example, the Fed had been issuing statements after Federal Open Market Committee (FOMC) meetings since 1995. Similarly, the BOJ had used asset purchases in 2001-06. But central banks would use such methods on unprecedented scales and with much greater aggressiveness after 2008-09.

6. Central banks find it difficult to reduce short-term interest rates much below zero because lenders become increasingly reluctant to lend at negative interest rates when they could hold cash that pays a return of zero. The fact that holding large sums of cash is also costly has allowed some central banks—for example, the ECB and BOJ—to reduce interest rates to modestly negative levels. Molyneux, Reghezza, and Xie (2019) argue that negative interest rates put pressure on the financial system’s health, while many others find no evidence of this (Jobst and Lin, 2016; Basten and Mariathasan, 2018; Lopez, Rose, and Spiegel, 2020; and Arteta et al., 2016 and 2018).

7. See Ramey (2019).

8. The Term Auction Facility (TAF), introduced by the Fed in December 2007, allowed the Fed to loan money to banks for monetary policy purposes (see https://www.federalreserve.gov/monetarypolicy/taf.htm for details). The TAF sought to broadly increase bank credit rather than help particular banks as the discount window does.

9. Prior to the Dodd-Frank Act of 2010, the Fed could lend to individual financial institutions or firms under Federal Reserve Act Section 13(3) authority when the Fed’s Board of Governors deemed circumstances to be “unusual and exigent.” Lending to Bear Stearns and AIG in 2008 occurred under this authority but would not be allowed today. The shaded insert on the Dodd-Frank Act, “Did Post-2008 Financial Reforms Reduce Financial Crisis Risk?,” describes the major changes in the U.S. regulatory structure.
The Fed is only one of many financial regulators, both federal and state, with overlapping responsibilities.

The BOE operated under a gold standard in the 19th century when Bagehot wrote. Bagehot advocated high interest rates during crises, to discourage excessive borrowing that would exhaust the BOE’s gold reserves. Modern central banks can issue fiat currency, however, which allows them to create unlimited amounts of money. Therefore, there is less reason to be concerned about maintaining high lending rates.


For a comparison of Fed actions in these two periods, see Wheelock (2010).

Lacking good information about a particular firm’s prospects or future profitability, lenders see a firm or individual’s net worth as an important, quantifiable determinant of the ability to repay a loan.

A specialist is a member of a stock exchange who is obligated to make a market in a particular stock. That is, the specialist stands ready to buy or sell the asset. Ang, Gorovyy, and van Inwegen (2011) study hedge fund leverage.

Counterparty risk is the danger that one’s counterparty will fail to settle a transaction.

The relationship between gross private domestic investment and U.S. recessionary periods can be seen here: https://fred.stlouisfed.org/graph/?g=r48w.

The secondary shocks likely to hit the economy and the financial system in the months to come still could trigger a financial crisis, however; see Bullard (2020).

The OECD data are available at https://data.oecd.org/gga/general-government-debt.htm.

LIBOR is the London Interbank Offer Rate, applied to unsecured borrowing by a major bank from another bank. OIS is the Overnight Index Swap rate, referring to the fixed-rate leg of an interest rate swap against the overnight policy rate. In the United States, this is the federal funds rate. LIBOR is a default-risky rate and OIS is default-risk free rate because it effectively is secured by the other leg of the swap. Hence, the LIBOR-OIS spread is a clear indicator of perceived risk of default by a major bank.

The USD did not appreciate during the period of the Russian default in August-November 1998.

Financial crises and, to a lesser extent, recessions are very difficult to predict because if they could be predicted well in advance, people and firms would take actions—such as selling risky assets or curtailing investment—that would bring them on immediately. The fourth quarter of 2019 was weak for Germany and Japan. An increase in the national sales tax weakened activity in Japan, and revised German GDP numbers showed a slight contraction (Fairless and Hannon, 2020, and Fujikawa, 2020).

The OECD forecast (available at https://data.oecd.org/gdp/nominal-gdp-forecast.htm#indicator-chart) annualized nominal second quarter GDP declines of 56 percent for the United Kingdom, 40 percent for the United States, 34 percent for Japan, and 48 percent for the euro area. An annualized decline in quarterly GDP of 25 percent, for example, means that second quarter GDP would be almost 7 percent lower than it would otherwise be: 0.75^{1/2} = 0.93 = 1 – 0.07. If a 25 percent annualized drop in second quarter GDP is accompanied by no growth in each of the other quarters, total annual growth would be –7 percent.

In addition, the Fed’s Board of Governors voted to lower the interest rate paid on reserves from 1.6 percent to 1.1 percent on March 3 and then to 0.1 percent on March 15.

The ECB had already been engaged in asset purchases since November 2019. The BOJ was also already conducting asset purchases for years prior to the COVID-19 crisis.

The recent trend toward greater investment by ETFs in corporate bonds is also of concern because this might amplify and transmit price moves through the economy. The BOE wrote about the issue in its July 2019 Financial Stability Report (BOE, 2019a, p. 35): “[L]arge-scale redemptions from funds could result in sales of illiquid assets that may exceed the ability of dealers and other investors to absorb them, amplifying price moves, transmitting stress to other parts of the financial system, and disrupting the availability of finance to the real economy.”

The ESF was established decades ago to allow the Treasury to intervene in foreign exchange markets.

More than 90 percent of MBS trading occurs in the TBA market, which is a forward market for agency MBS (Vickery and Wright, 2013).
ABS are structured securities that deliver coupon and principal payments from the repayment of one of many types of receivables, such as credit card loans, auto loans, or student loans. That is, lenders sell the rights to receive loan (or other) payments in the form of a bond (Agarwal et al., 2010). ABS are not government guaranteed, but they greatly broaden the market for small loans by transforming relatively risky, heterogeneous, low-value assets into safer, homogeneous fixed-income securities that can be sold to savers such as people or pension funds. This securitization increases the availability of loans and decreases borrowing costs. The Securities Industry and Financial Markets Association estimates that ABS issuance peaked in 2007 at $796 billion (see https://www.sifma.org/resources/research/fixed-income-chart/). The current version of the TALF had a similar predecessor during the GFC of 2007-09. The first version of the TALF began operating in March 2009, with outstanding loans peaking at just under $50 billion in 2010. Almost all loans were repaid by the end of 2012.

The TALF is one of several credit facilities that the Fed reincarnated from the GFC of 2007-09, including the PDCF, the CPFF, and the MMLF. The Municipal Liquidity Facility (MLF), PPPLF, PMCCF and SMCCF, and the MSLP are the relevant new programs.

Foreign exchange swaps combine spot and forward transactions. When the Fed lends money through a swap, it sells USD to a foreign central bank and simultaneously contacts to buy those USD back at a specified time in the future for a specified price, which reflects the difference in interest rates on the currencies. Currency swaps entail minimal risk because the swapped currency functions as collateral.

A repo is secured borrowing in which one party purchases an asset and simultaneously sells it forward for a specified price. When the Fed lends money through a repo, it buys an asset from a foreign monetary authority and sells it back through a forward transaction. The difference between the forward and spot prices is the borrowing cost, and the exchanged currency itself acts as collateral that minimizes the risk to the lender.

In the United States, the relevant standard is the current expected credit losses (CECL) methodology; elsewhere, it is the International Financial Reporting Standards (IFRS 9) expected credit loss model.

A capital buffer is the amount of capital a bank has above the required regulatory minimum.

The BOE, ECB, and Fed have also all issued guidelines to banks that restrict buybacks, dividends, and/or bonuses.

A money market fund is said to “break the buck” when its net asset value falls below $1. In such a case, money market fund investors may lose principal.

See https://www.federalreserve.gov/monetarypolicy/mainstreetlending.htm for MSLP facility terms.

The Fed’s PPPLF also incentivizes bank lending by providing funding for banks that originate PPP loans, taking those loans as collateral.

Mizrach and Neely (2020b) cite IMF data in reporting that USD-denominated assets make up 60.9 percent of foreign exchange reserves, while euro-denominated assets make up 20.5 percent.

Central banks had increasingly turned to open-ended asset purchase programs after the Fed shifted to conditioning QE3 asset purchases on economic activity on December 12, 2012 (Fawley and Neely, 2013).

The BOJ is also continuing with its yield-curve control policy, targeting its longer-term 10-year JGB yield to 0 percent.

REFERENCES


Haas, Neely, Emmons


Fawley, Brett W. and Neely, Christopher J. “Four Stories of Quantitative Easing.” *Federal Reserve Bank of St. Louis Review*, 2013, 95(1), pp. 51-88; [https://doi.org/10.20955/r.95.51-88](https://doi.org/10.20955/r.95.51-88).


International Monetary Fund. “IMF DataMapper: Real GDP Growth.” IMF World Economic Outlook, April 2020b; [https://www.imf.org/external/datamapper/NGDP_RPCH@WEO/OEMDC/ADVEC/WEOWORLD](https://www.imf.org/external/datamapper/NGDP_RPCH@WEO/OEMDC/ADVEC/WEOWORLD).


We evaluate, *empirically*, the effect of changes in trade policy during the 2018-19 trade war on U.S. economic activity. We begin by documenting that sectors and states across the United States are heterogeneous in their exposure to international trade. To do that, we construct a measure of exposure that combines the share of a sector’s gross output that is accounted for by trade with the pattern of comparative advantage of each state in that sector. We then exploit cross-state heterogeneity in exposure to international trade and correlate it with measures of economic activity across U.S. states. Our findings suggest that states that were very exposed to trade at the onset of the trade war experienced worse outcomes in terms of employment and output growth. Our analysis is not aimed at concluding any causality effects, but instead focuses on correlations. (JEL F10, F13, F14)

https://doi.org/10.20955/r.102.385-412

1 INTRODUCTION

The U.S. trade policy that started in 2017 has spiraled into a return to protectionism. The trade war began with recommendations from the U.S. administration to implement tariffs on specific goods (e.g., solar panels, washing machines, and steel and aluminum) in an attempt to protect national security and mitigate negative impacts faced by domestic producers of such products.¹ Tariffs on these goods were enacted in 2018 and were followed by retaliatory tariffs from China, along with retaliatory tariffs from the European Union (EU) and Canada on steel, aluminum, and agricultural products. The United States has continued to levy tariffs onto major trading partners such as Canada, Mexico, the EU, and China.² Perhaps the most noteworthy battle in the tariff front exists between the United States and China, given the growing tension between the two countries as the trade war progresses.

In this article, we exploit variation in sector-state exposure to international trade and evaluate, *empirically*, the effect of trade policy on U.S. output growth and employment. Recent papers have explored the effect of the trade war on the United States through the lens of a
quantitative model (see Caldara et al., 2019; Fajgelbaum et al., 2019; and Santacreu, Sposi, and Zhang, 2019). However, empirical studies have been rather limited. We try to fill that gap by providing an empirical analysis of the effects of a trade war on the United States.

We start by constructing a measure of initial sectoral exposure to international trade, from the perspective of both imports and exports. We focus on trade in intermediate goods and abstract from final goods, as tariff announcements have fallen mainly onto intermediate goods. We use data for 19 sectors from the World Input-Output Database (WIOD) 2016 Release for 2014—the latest year available in the WIOD, prior to the trade war. We then compute, for each sector, a measure of trade exposure as the ratio of U.S. total trade of intermediate products with the world to U.S. gross output in that sector. To the extent that U.S. states specialize in the production of different sectors, exposure to trade policy will likely be heterogeneous across states. To capture this idea, we combine our measure of sectoral trade exposure with the production composition of each U.S. state—calculated as the share of value added of each sector in that state—and compute a measure of trade exposure at the state level. It is important to note that this measure of trade exposure does not rely on actual tariff data. We interpret it as a sector’s overall trade exposure that depends only on data of traded intermediate goods by sector, regardless of actual trade policy.

We find that the sectors most exposed to imports—those that rely more heavily on imported intermediate products—are Coke and petroleum and Motor vehicles and trailers. Mining and quarrying and Forestry, fishing, and logging are the sectors least exposed. With respect to exports, the sectors most exposed—those that export more intermediate goods to the world—are Other transport equipment and Computer, electronic, and optical, and the sectors least exposed are Mining and quarrying and Other non-metallic mineral products. Louisiana and Michigan are very exposed to imports from the world. Louisiana specializes in Coke and petroleum, and Michigan specializes in Motor vehicles and trailers. Alaska and North Dakota are the states least exposed to imports, as they both specialize in Mining and quarrying. In the case of exports, Washington and California are very exposed, whereas Alaska and Wyoming are among the states least exposed.

We then follow an approach similar to Mian and Sufi (2009) to exploit the observed cross-state heterogeneity in exposure to international trade and correlate it with measures of economic activity across U.S. states. We focus on quarterly growth rates of employment and output between 2018:Q1 and 2019:Q1. We find the following: (i) There is a negative and significant correlation between the initial exposure to trade and economic activity; (ii) the negative correlation is stronger with employment growth than it is with output growth; and (iii) the negative correlation is stronger with import exposure than it is with export exposure. That is, those states more exposed to trade experienced lower increases or even decreases in output growth and employment growth between 2018 and 2019. These findings reflect that firms operating in states very exposed to trade adjusted their employment and production decisions after announcements of tariff increases. The adjustments were stronger in terms of employment than in terms of output, and they were stronger in states very exposed to U.S. tariffs than in states very exposed to retaliatory tariffs. Obviously, there are economic forces other than trade exposure that could be driving heterogeneous adjustments in employment and production
across U.S. states. Our analysis is not aimed at concluding any causality effects. However, the strong negative correlation found in the data suggests that initial trade exposure could have played a role in those adjustments.

Finally, we analyze the effect of the 2018-19 U.S.-China trade war on U.S. economic activity. We construct exposure measures that use actual U.S. tariffs imposed on China and subsequent retaliatory Chinese tariffs. The measure of import exposure exploits the input-output structure of the economy, as it captures that those sectors that depend more on imports from sectors with larger increases of U.S. tariffs are more exposed to imports from China. These are Motor vehicles and trailers and Machinery and equipment, n.e.c. In the case of export exposure, sectors that the United States exports most or that are subject to higher retaliatory Chinese tariffs will be more exposed to trade with China. These are Other transport equipment and Agriculture. Heterogeneity in sector exposure translates into heterogeneity in state exposure, as states specialize in the production of different sectors. Michigan is very exposed to imports from China, whereas Washington is very exposed to exports. As before, we find a negative correlation—albeit a weaker one than in the case of exposure to the world—between our measures of trade exposure to China and both output growth and employment growth. In contrast to our previous results, the negative correlation is stronger with output growth than it is with employment growth, and it is stronger in the case of retaliatory tariffs.

Our results suggest that cross-state heterogeneity in trade exposure correlates negatively with U.S. economic activity. Although we cannot claim any causality effects, these findings are an indication that the trade war initiated by the United States may have had a stronger impact on U.S. employment and production than what is found through the lenses of standard models of trade. Accounting for this heterogeneity is thus key in capturing the negative impacts.

Our article is related to a very recent strand of literature on the effects of trade policy on the United States. Actual tariff increases, followed by the threat of future raises, have increased uncertainty in the tradable sector. According to several studies, higher uncertainty has already had a negative impact on the United States in terms of investment and production (see Caldara et al., 2019; Handley and Limão, 2015 and 2017; and Bloom, Bond, and Van Reenen, 2007). Recent studies using quantitative models of trade have found small aggregate effects of tariffs on the U.S. economy. These studies, however, do not take into account the effect of uncertainty and focus their analysis on actual tariff increases. Despite small aggregate effects, these models find highly heterogeneous effects of tariffs across both sectors and states (see Fajgelbaum et al., 2019; Amiti, Redding, and Weinstein, 2019; Auer, Bonadio, and Levchenko, 2018; and Santacreu, Sposi, and Zhang, 2019), which could potentially have heterogeneous effects on economic activity across U.S. states. Our article differs from those studies in that we explore empirically—rather than through the lens of a quantitative model—whether heterogeneous exposure to trade had an impact on economic outcomes.

2 HETEROGENEOUS IMPACT OF TRADE POLICY ACROSS U.S. STATES

In this section, we evaluate the effect of changes in trade policy on economic activity across U.S. states. We start by computing a measure of initial sectoral exposure to international
trade with the world. We then combine that measure with the production structure of each state across different sectors to obtain a measure of state trade exposure. Finally, we use an approach akin to Mian and Sufi (2009) and exploit cross-state variation in initial trade exposure to evaluate the impact of trade policy on state employment and output growth.

2.1 Sectoral Trade Exposure

Our first observation is that sectors in the United States are exposed to international trade differently. Using data for 2014 from the WIOD, 2016 Release, we compute a measure of initial trade exposure at the sector level. The idea is to calculate a measure of trade exposure prior to the announcement or the actual change in tariffs, as trade flows may change when tariffs are either announced or actually implemented. By looking at initial trade flows, our exposure measure is not affected by these changes.

Specifically, we compute sector $j$'s exposure to trade between the United States and the world, with $E^j_{US,W}$ as the ratio of total imports or exports of intermediates goods in a sector to U.S. gross output in that sector. That is,

$$E^j_{US,W} = \frac{X^j_{US,W}}{GO^j_{US}},$$

where $X^j_{US,W}$ represents total U.S. trade flows of intermediate goods in sector $j$, and $GO^j_{US}$ is gross output of the United States in sector $j$.

Equation (1) measures the share of U.S. production in sector $j$ that uses traded intermediate goods from the world (in the case of import exposure) and the share of U.S. sales in sector $j$ that is exported to the world (in the case of export exposure). For industries that rely heavily on imported intermediate products, if the United States were to impose uniform tariffs across all sectors and trading partners, these industries would be impacted the most. If trading partners retaliated by increasing tariffs across all U.S. sectors, those industries most exposed to exports would be affected the most.

Panel A of Figure 1 plots the top five U.S. industries with the largest import exposure to the world. These industries are Coke and petroleum, Motor vehicles and trailers, Other transport equipment, Basic metals, and Rubber and plastic products. For instance, the Coke and petroleum industry relies heavily on imports from Mining and quarrying and from Basic metals from Brazil, Mexico, and Canada. Motor vehicles and trailers relies on imports from the Motor vehicles and trailers industry in Germany, Japan, China, Canada, and Mexico. Panel B of Figure 1 plots the top five U.S. industries with the largest export exposure to the world. These are Other transport equipment, followed by Computer, electronic, and optical; Machinery and equipment, n.e.c; Electrical equipment; and Chemicals and pharmaceuticals.

2.2 State-Level Trade Exposure

Sectoral heterogeneity in trade exposure likely translates into state heterogeneity, as states differ in their sectoral composition of production. States that produce more in sectors heavily exposed to trade may be more impacted by changes in international trade policy. We construct a measure of state exposure to international trade by sector that combines sectoral
Figure 1
Sectoral Exposure of the United States to the World

A. U.S. sectoral import exposure to world

B. U.S. sectoral export exposure to world

NOTE: This figure plots sectoral import and export exposure computed using equation (1) for the top five most-exposed sectors.

SOURCE: Authors’ calculations.
exposure from equation (1) with the sectoral composition of production of each U.S. state. In particular, state \( s \)’s trade exposure to sector \( j \) is computed as

\[
E_{US,W}^{j,s} = \frac{X_{US,W}^j}{G_{US}^j} \left( \frac{VA_{US}^{j,s}}{\sum_j VA_{US}^{j,s}} \right),
\]

where \( \frac{VA_{US}^{j,s}}{\sum_j VA_{US}^{j,s}} \) is the value-added share of state \( s \) in sector \( j \).

The following are two contributing factors to state-sector exposure in equation (2):

(i) the initial exposure of each U.S. sector, \( \frac{X_{US,W}^j}{G_{US}^j} \); and (ii) the value-added share of each state in that sector, \( \frac{VA_{US}^{j,s}}{\sum_j VA_{US}^{j,s}} \).

Figure 2 plots the distribution of import exposure to the world across U.S. states for the top two sectors and the bottom two sectors based on their exposure to imports. Figure 3 does the same from the perspective of exports. Darker colors represent higher exposure to trade. Heat maps with the most-exposed sectors have darker colors on average, as we have imposed the scale of the bins to be the same across all sectors. The figures show that, within each sector, there is a lot of heterogeneity across U.S. states in their trade exposure, which is explained by the composition of production. If a state has a high value-added share in a particular sector, the state will be more exposed to changes in trade policy affecting that particular sector (and that state will have a darker color in the graph).

We start by documenting state-sector heterogeneity for import exposure. As we have shown previously, Coke and petroleum and Motor vehicles and trailers are the two sectors most exposed to imports from the world. Forestry, fishing, and logging and Furniture and other manufacturing are among the sectors least exposed. The state most exposed to imports in the Coke and petroleum industry is Louisiana, with 36 percent of its total value added being accounted for by this sector in 2014, followed by Hawaii (30 percent) and Montana (20 percent). In Louisiana, for instance, the extraction and processing of petroleum and natural gas became the state’s largest industrial activity in the twentieth century. One of the largest manufacturers is Odyssea Marine, a gas and oil utility company. In contrast, the state least exposed in this industry is Nebraska, with less than 0.1 percent of its value added in this industry, followed by South Dakota (0.1 percent) and Idaho (0.3 percent). For the Motor vehicles and trailers industry, the state most exposed is Michigan, for which the Motor vehicles and trailers industry comprised 43 percent of its total value added in 2014, followed by Tennessee (19 percent) and Kentucky (17 percent). Michigan has a long history of production in automobiles: Currently, it is home to both Chrysler and General Motors and acts as one of the top manufacturers of automobiles in the United States. In contrast, the state least exposed is Alaska, at a 0.02 percent value-added share, followed by Hawaii (0.03 percent) and Wyoming (0.1 percent). In the Forestry, fishing, and logging industry, the state most exposed is Maine, followed by Oregon and Florida. The state least exposed is Connecticut, followed by Ohio and Texas. In the Food, beverage, and tobacco industry, the state most exposed is Virginia, followed by Georgia and Hawaii. The state least exposed is Wyoming, followed by West Virginia and New Mexico.
With regard to export exposure to the world, Other transport equipment and Computer, electronic, and optical are the sectors most exposed, whereas Other non-metallic mineral products and Wood and cork are the sectors least exposed. The state most exposed to the Other transport equipment industry is Washington, with 56 percent of its total value added in 2014 being accounted for by this sector, followed by Connecticut (25 percent) and Kansas (17 percent). Indeed, one of Washington’s largest manufacturers is the aerospace engineering company Boeing. The state least exposed in this industry is Wyoming, with 0.07 percent of its value added in this industry, followed by Montana (0.09 percent) and Alaska (0.01 percent).

For the Computer, electronic, and optical industry, the state most exposed is Oregon, with a 43 percent value-added share in this industry, followed by Massachusetts (33 percent) and California (27 percent). Major technology moguls such as Apple, Google, and Facebook are...
headed in California, bringing large-scale employment into the area and attracting a continuous stream of innovative companies. The state least exposed in this industry is Alaska, with a 0.09 percent value-added share, followed by Wyoming (0.3 percent) and Louisiana (0.3 percent). In Other non-metallic mineral products, the state most exposed is Hawaii, followed by Vermont and Florida. The state least exposed is Alaska, followed by Connecticut and Wyoming. In the Wood and cork industry, the state most exposed is Oregon, followed by Maine and Vermont. The state least exposed is Wyoming, followed by New Mexico and Connecticut.

Figure 4 shows patterns of specialization and diversification of production for a sample of U.S. states. Louisiana, Michigan, Washington, and California specialize in highly exposed sectors. Wyoming, instead, specializes in Mining and quarrying, which is not very exposed to trade. We also observe sectors that are very diversified, such as Maine.
Analyzing exposure levels of states and sectors is crucial, as states producing a significant portion of their value added in industries that are heavily exposed to trade will be more affected by changes in trade policy. Next, we construct a measure of total trade exposure at the state level, $E_{US,W}^s$, as

$$E_{US,W}^s = \sum_j X_{US,W}^j \frac{VA_{i,W}^s}{GO_{US}^j \sum_j VA_{i,W}^s}.$$  

If a state $s$ specializes in a sector $j$ that is heavily exposed to international trade, then that state will be more exposed to changes in trade policy. Figure 5 depicts total trade exposure to both imports (Panel A) and exports (Panel B) between the United States and the world.

As for imports, the states most exposed are Louisiana and Michigan, followed by Hawaii, Washington, and Ohio. The states least exposed are North Dakota and Alaska, followed by New Mexico, Wyoming, and Idaho. Louisiana’s exposure is a combination of two factors. First, it is one of the main producers in the Coke and petroleum sector, which is the sector most exposed to imports (see Figure 1). Second, Coke and petroleum accounts for around 36 percent of Louisiana’s total value-added share. Similarly, Michigan is one of the top U.S. producers in the Motor vehicle and trailers industry, which accounts for 43 percent of its total value added.
Other states that are highly exposed, such as Hawaii, Washington, and Ohio, are highly specialized in industries in which the United States imports heavily. For example, Hawaii and Ohio are among the top producers in the Coke and petroleum industry, alongside Louisiana, and Washington is heavily specialized in the Other transport equipment industry, which was the third sector most exposed to imports in 2014. The states least exposed are North Dakota and Alaska. Alaska specializes in Mining and quarrying, which accounts for 87 percent of its total value added. North Dakota specializes in Mining and quarrying and Agriculture, which together account for 85 percent of its total value added. These industries are among the least exposed to imports in 2014 (see Figure A1 in the appendix).

For exports, the states most exposed are Washington, Connecticut, Massachusetts, Arizona, and Rhode Island, followed by Oregon and California. The states least exposed are Alaska, Wyoming, North Dakota, Oklahoma, and New Mexico. Washington specializes in Other transport equipment; Connecticut and Massachusetts specialize in Computer, electronic, and optical products and Chemicals and pharmaceuticals; Arizona specializes in Computer, electronic, and optical and Other transport equipment; and Rhode Island specializes in Machinery and equipment, n.e.c. and Other transport equipment. These sectors are heavily exposed to exports. Alaska and Wyoming specialize in Mining and quarrying, of which exports are not very exposed to the world (see Figure A1 in the appendix).

### 2.3 Trade Exposure and Economic Activity

In this section, we exploit the state-level heterogeneity in exposure of imports and exports documented in Section 2.2 and relate it to differences in economic activity across U.S. states. The current U.S. administration began talking about increasing tariffs in October 2017 after
the U.S. International Trade Commission found that imports of certain products were negatively impacting U.S. producers of such goods. These announcements increased uncertainty in the tradable sectors. If firms reacted to this uncertainty by adjusting their employment and production decisions, states that were most exposed to imports would be impacted the most. Talks about retaliation by main trading partners of the United States would instead have a bigger impact on states that were most exposed to exports.

We follow the approach of Mian and Sufi (2009) who exploit within-county variation in credit growth to explain cross-state variations in income growth. We use their methodology to exploit cross-state variation in trade exposure and relate it to cross-state variation in output growth and employment growth. The idea is as follows: Consider two U.S. states—State A is very exposed to trade and State B is less exposed. An announcement of increases in tariffs will have a disproportionately larger impact on State A than on State B.

To analyze the effect on the U.S. economy, we use state-level quarterly data for employment and output for the three main broad industries considered in our analysis (Agriculture, Mining and quarrying, and Furniture and other manufacturing). Employment data are from the Bureau of Labor Statistics (BLS) for all goods-producing industries; output data are from the Bureau of Economic Analysis (BEA)—quarterly GDP data by state in real millions of 2012 chained U.S. dollars—for Agriculture, Mining and quarrying, and Furniture and other manufacturing. For both employment and output, we compute quarterly growth rates from 2018:Q1 to 2019:Q1.

Panels A and B of Figure 6 plot the correlation between the initial state-level import exposure and the following: (i) state-level GDP growth from 2018:Q1-2019:Q1 (Panel A) and (ii) state-level employment growth from 2018:Q1-2019:Q1 (Panel B). We find that states more exposed to imports experienced lower increases in GDP growth (a correlation of –0.18) and even lower increases in employment growth (a correlation of –0.45).

Panels C and D of Figure 6 plot the correlation between the initial state-level export exposure and our measures of economic activity. We find that states more exposed to exports experienced lower increases in GDP growth (a correlation of –0.20) and even lower increases in employment growth (a correlation of –0.23).

The results presented so far suggest the following: (i) There is a negative correlation between initial trade exposure and economic activity across U.S. states; (ii) cross-state variation in trade exposure correlates more negatively with employment growth than it does with output growth; and (iii) import exposure seems to have had a stronger impact on economic outcomes than has export exposure. These points could reflect that firms operating in states that are more exposed to trade adjusted their employment and production decisions more after announcements of tariff increases by the United States.

Finally, we construct a measure of state total trade exposure as the sum of import and export trade exposures. This measure shows which states would be more impacted by trade policy if the United States imposed a uniform tariff across every trading partner and sector and, simultaneously, all trading partners retaliated by increasing tariffs uniformly across every U.S. sector. Panels E and F of Figure 6 plot the correlation between this measure and (i) state-level GDP growth from 2018:Q1-2019:Q1 (Panel E) and (ii) state-level employment growth
Figure 6

Trade Exposure to World and Economic Activity, 2018:Q1-2019:Q1

A. GDP growth

B. Employment growth

C. GDP growth

D. Employment growth

E. GDP growth

F. Employment growth

NOTE: This figure plots, in the x-axis, heterogeneity in import exposure (Panels A and B), export exposure (Panels C and D), and total exposure (Panels E and F), and relates them to GDP growth and employment growth for 2018:Q1-2019:Q1 (y-axis).

SOURCE: Authors’ calculations.
from 2018:Q1-2019:Q1 (Panel F). We observe a negative correlation between total trade exposure and economic activity. Once again the correlation is more negative in the case of employment than in the case of output growth (i.e., –0.35 vs. –0.23, respectively).

Our results suggest that states that were more exposed to trade experienced worse economic outcomes in terms of production and employment. Obviously, there are economic forces other than trade exposure that could be driving heterogeneous adjustments in employment and production across U.S. states. Figure 6 does not imply any causality effects. However, the strong negative correlation between the variables suggests that initial trade exposure could have played an important role in those adjustments.

3 U.S.-CHINA TRADE WAR

In the previous section, we constructed a measure of trade exposure in which we did not use any data on current tariffs. In this section, we focus on one particular episode, the U.S.-China trade war, and use actual data on tariffs. Specifically, we evaluate the impact of the current increase of U.S. tariffs on imports from China, and the subsequent retaliatory measures imposed by the Chinese government on American products. Given the current trade tensions between the United States and China, it is important to narrow our analysis to focus on trade between these two countries. Furthermore, China accounts for the largest share of U.S. exports under retaliatory tariffs among the main U.S. trading partners (China, Canada, Mexico, and the EU).

3.1 Tariff Data

Tariff data in the context of the U.S.-China trade war are from the Iowa State University Center for Agricultural and Rural Development. We look at recent tariff increases by the United States and retaliatory tariffs by China. The database provides tariff data aggregated to 6-digit Harmonized System (HS) code commodity levels. We then classify the products into each of our 19 sectors and compute, for each sector, simple averages of tariff rates across the different 6-digit HS goods that belong to that sector. Figure 7 plots the top 10 sectors with the largest U.S. tariff increases (Panel A) and the largest retaliatory Chinese tariff increases (Panel B). The largest tariff increases imposed by the United States are on the Electrical equipment industry (23 percent), followed by the Other transport equipment industry (just under 23 percent) and the Machinery and equipment, n.e.c. industry (21 percent). China has placed the highest tariffs on the Coke and petroleum industry (20 percent), followed by the Basic metals industry (13 percent) and, finally, the Food, beverage, and tobacco industry (11 percent). China’s retaliatory tariffs are about 10 percent on Furniture and other manufacturing and Agriculture.

3.2 Sectoral Trade Exposure to China

We now compute a measure of initial sectoral trade exposure to China. This measure differs slightly from the one used in the previous section because we now keep track of the
Figure 7
Tariff Levels

A. Tariffs, U.S.

B. Retaliatory tariffs, China

NOTE: The bar charts plot recent tariff changes for the top 10 industries most affected by tariffs, both U.S. tariffs and retaliatory China tariffs.

SOURCE: Authors’ calculations.
Figure 8
U.S. Sectoral Exposure to China

A. U.S. imports from China

B. U.S. exports to China

NOTE: This figure plots sectoral import and export exposure computed using equations (4) and (5) for the top five most-exposed sectors.

SOURCE: Authors’ calculations.
specific sectors affected by actual tariff increases. Specifically, sector \( j \)'s import exposure to China is computed as

\[
M_{US,China}^j = \sum_k M_{US,China}^{jk} \frac{\Delta \tau_k}{GO_{US}^j},
\]

where \( M_{US,China}^{jk} \) represents the United States’ sector \( j \) imports from China’s sector \( k \), and \( \Delta \tau_k \) is the actual increase in tariffs imposed by the United States onto Chinese products from sector \( k \).

Sector \( j \)'s export exposure to China is computed as

\[
X_{US,China}^j = \frac{X_{US,China}^j \Delta \tau_j}{GO_{US}^j},
\]

where \( X_{US,China}^j \) represents the United States’ sector \( j \) exports to China, and \( \Delta \tau_j \) is the actual increase in retaliatory tariffs by China on U.S. products from sector \( j \).

Figure 8 plots the five industries with the largest sectoral trade exposure to China. The five industries most exposed to imports (Panel A) are Motor vehicles and trailers; Machinery and equipment, n.e.c.; Electrical equipment; Other transport equipment; and Computer, electronic, and optical. Note that this measure exploits information from the input-output structure of the U.S. economy and information on actual tariff rate increases. It captures that sectors that import more from sectors experiencing large U.S. tariff increases are more exposed. For instance, the Coke and petroleum sector was very exposed to imports from the world, but not very exposed to imports from China. This is because the Coke and petroleum industry in the United States imports 92 percent of its intermediate inputs from the Mining and quarrying industry, which is not subject to large U.S. tariffs. Because we did not use actual tariff data in our measure of import exposure to the world, the fact that this industry relies heavily on imports—despite being from a sector where the United States has not imposed high tariffs—was sufficient to generate high exposure. In contrast, the Motor vehicles and trailers industry imports 53 percent of its intermediate inputs from that same industry from other countries, which makes this industry bear the brunt of the largest U.S. tariffs. The five industries most exposed to exports (Panel B of Figure 8) are Other transport equipment; Agriculture; Machinery and equipment, n.e.c.; Computer, electronic, and optical; and Motor vehicles and trailers.

### 3.3 State Trade Exposure to China

We now construct a measure of exposure to trade with China by sector and state that combines sectoral heterogeneity at the U.S. level—from equations (4) and (5)—with heterogeneity in the composition of production across U.S. states. Figure 9 plots the state distribution of import exposure to China for the two most- and least-exposed sectors. Figure 10 does the same from the perspective of exports.

With regard to import exposure to China, Motor vehicles and trailers and Machinery and equipment, n.e.c. are the sectors most exposed. In contrast, the Mining and quarrying and the Forestry, fishing, and logging industries are among the industries least exposed in our sample. The state most exposed to imports from China in Motor vehicles and trailers is Michigan, followed by Tennessee and Kentucky. In contrast, the state least exposed in this
industry is Alaska, followed by Hawaii and Wyoming. For Machinery and equipment, n.e.c.,
the state most exposed is Nevada, for which this industry comprised 23 percent of the state’s
total value added in 2014, followed by Rhode Island (14 percent) and South Dakota (11 percent).
In contrast, the states least exposed are Wyoming, Alaska, and North Dakota.

The industries least exposed to imports from China are Mining and quarrying and Forestry,
fishing, and logging. In the Mining and quarrying industry, the state most exposed is Alaska,
with 87 percent of its total value added in that sector, followed by Wyoming (80 percent) and
New Mexico (68 percent). In the Forestry, fishing, and logging industry, the state most exposed
is Maine, followed by Oregon and Florida.

With regard to export exposure to China, Other transport equipment and Agriculture
are the sectors most exposed to exports, while Other non-metallic mineral products and Coke

---

**Figure 9**

**U.S. Import Exposure to China, by State and Sector**

A. Motor vehicles and trailers

B. Machinery and equipment, n.e.c.

C. Mining and quarrying

D. Forestry, fishing, and logging

NOTE: The heat maps plot sector-state import exposure to China for the two sectors most exposed and the two sectors least exposed. Darker
colors represent higher exposure. Bins have the same scale across all sectors.

SOURCE: Authors’ calculations.
and petroleum are the sectors least exposed. The states most exposed to Other transport equipment are Washington, Connecticut, and Kansas. The states least exposed in this industry are Wyoming, Montana, and Alaska. For the Agriculture industry, the states most exposed are South Dakota, Nebraska, and Iowa. South Dakota is a large producer of crops, such as wheat, corn, and soybeans, often ranking within the top 10 national producers of certain crops.\textsuperscript{14} The state least exposed in this industry is Alaska, followed by Massachusetts and Rhode Island. The industries least exposed to exports are Other non-metallic mineral products and Coke and petroleum. In the Other non-metallic mineral products industry, the state most exposed is Hawaii, followed by Vermont and Florida. The state least exposed in this industry is Alaska, followed by Connecticut and Wyoming. In the Coke and petroleum industry, the state most exposed is once again Louisiana, followed by Hawaii and Montana; the state least exposed is Nebraska, followed by South Dakota and Idaho.
Next, we compute a measure of state trade exposure using equation (3). Panel A of Figure 11 shows that the state most exposed to imports from China is Michigan, followed by Connecticut, Washington, Massachusetts, and Rhode Island. The states least exposed are Alaska, Wyoming, Montana, North Dakota, and Oklahoma. As mentioned previously, Alaska and Wyoming specialize in Mining and quarrying, which accounts for very little of the U.S. imports from China.\textsuperscript{15} Panel B of Figure 11 shows that the states most exposed to exports to China are Washington, South Dakota, Connecticut, Nebraska, and Arizona.\textsuperscript{16} The state least exposed is Alaska, followed by Wyoming, Oklahoma, West Virginia, and, finally, New Mexico.

### 3.4 Trade Exposure to China and U.S. Economic Activity

In this section, we exploit cross-state variation in trade exposure to China and relate it to differences in economic activity across U.S. states. The results are plotted in Figure 12 for import exposure (Panels A and B), export exposure (Panels C and D), and total exposure (Panels E and F). We find that, as before, there is a negative correlation—albeit a weaker one than in the case of exposure to the world—between our measures of trade exposure to China and both output growth and employment growth. However, in contrast to the results in Section 2.3, we find the following: (i) This correlation is lower with export exposure than it is with import exposure to China. That is, retaliatory tariffs seem to have had a bigger impact on U.S. economic activity. For instance, in the case of output growth, the correlation with import exposure and export exposure is \(-0.18\) and \(-0.39\), respectively. In the case of employment growth, these correlations are \(-0.18\) and \(-0.27\), respectively. And (ii) the correlation with output growth is lower than it is with employment growth. For instance, in the case of total exposure, these correlations are \(-0.33\) and \(-0.25\), respectively.\textsuperscript{17}
Figure 12
Trade Exposure to China and Economic Activity, 2018:Q1-2019:Q1

A. GDP growth

B. Employment growth

C. GDP growth

D. Employment growth

E. GDP growth

F. Employment growth

NOTE: This figure plots, in the x-axis, heterogeneity in import exposure (Panels A and B), export exposure (Panels C and D), and total exposure (Panels E and F), and relates them to GDP growth and employment growth for 2018:Q1-2019:Q1 (y-axis).

SOURCE: Authors' calculations.
Overall, we find that heterogeneous trade exposure to China translated into variation in economic activity across U.S. states. The impact, however, was weaker than in the case of U.S. variation in trade exposure to the world.18

4 CONCLUSION

We have exploited cross-state variation in trade exposure and related it to U.S. economic outcomes in terms of both employment and output growth during the past year. States that were more exposed to trade with the world have performed worse in terms of employment and output growth than have states that were less exposed. Although we cannot claim any causality effects in this analysis, our findings suggest that the trade war initiated by the United States may have had a stronger impact on the U.S. economy than what standard models of trade have found. Accounting for state heterogeneity is thus key in capturing the negative impacts of tariff increases.

We have abstracted from several interesting channels through which tariffs could impact economic activity. However, we have not taken into account spillovers across U.S. states. If we had observed inter-state trade flows, our measure of state exposure would also account for imports from other highly exposed states. Furthermore, increases in tariffs could reallocate imports away from foreign countries and toward other U.S. states, affecting state exposure to trade. We did not look at this channel because of data limitations on inter-state trade for the sectors in our analysis. This channel, however, has been recently explored, through the lens of a quantitative model of trade, by Santacreu, Sposi, and Zhang (2019). They infer missing trade flows between U.S. states using gravity methods and find that accounting for internal flows has significant effects on welfare. ■
## APPENDIX

### Data Sources

### Table A1

#### Raw Data and Sources

<table>
<thead>
<tr>
<th>Variable</th>
<th>Source</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intermediate, final goods (imports and exports, millions current USD)</td>
<td>WIOD</td>
<td>2014</td>
</tr>
<tr>
<td>Output (value added) by state and sector (GDP, millions current USD)</td>
<td>BEA</td>
<td>2014</td>
</tr>
<tr>
<td>GDP by state (GDP, real millions 2012 chained USD)</td>
<td>BEA</td>
<td>2017-19</td>
</tr>
<tr>
<td>Employment by state (Q1, goods, private industries)</td>
<td>BLS</td>
<td>2017-19</td>
</tr>
<tr>
<td>Goods, tariffs imposed by U.S. and China</td>
<td>PIIE</td>
<td>2018-current</td>
</tr>
</tbody>
</table>


### Table A2

#### Industry List

<table>
<thead>
<tr>
<th>Industry description and code</th>
<th>Industry description and code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>1</td>
</tr>
<tr>
<td>Forestry, fishing, and logging</td>
<td>2</td>
</tr>
<tr>
<td>Mining and quarrying</td>
<td>3</td>
</tr>
<tr>
<td>Food, beverage, and tobacco</td>
<td>4</td>
</tr>
<tr>
<td>Textiles and leather</td>
<td>5</td>
</tr>
<tr>
<td>Wood and cork</td>
<td>6</td>
</tr>
<tr>
<td>Paper, paper products, and printing</td>
<td>7</td>
</tr>
<tr>
<td>Coke and petroleum</td>
<td>8</td>
</tr>
<tr>
<td>Chemicals and pharmaceuticals</td>
<td>9</td>
</tr>
<tr>
<td>Rubber and plastic products</td>
<td>10</td>
</tr>
<tr>
<td>Other non-metallic mineral products</td>
<td>11</td>
</tr>
<tr>
<td>Basic metals</td>
<td>12</td>
</tr>
<tr>
<td>Fabricated metals, except machinery and equipment</td>
<td>13</td>
</tr>
<tr>
<td>Computer, electronic, and optical</td>
<td>14</td>
</tr>
<tr>
<td>Electrical equipment</td>
<td>15</td>
</tr>
<tr>
<td>Machinery and equipment, n.e.c.</td>
<td>16</td>
</tr>
<tr>
<td>Motor vehicles and trailers</td>
<td>17</td>
</tr>
<tr>
<td>Other transport equipment</td>
<td>18</td>
</tr>
<tr>
<td>Furniture and other manufacturing</td>
<td>19</td>
</tr>
</tbody>
</table>
**Figure A1**

Sectoral Import and Export Exposure to the World

A. U.S. imports from world

B. U.S. exports to world

SOURCE: Authors' calculations.
Figure A2
Sectoral Import and Export Exposure to China

A. U.S. imports from world

B. U.S. exports to world

SOURCE: Authors’ calculations.
**Figure A3**


A. GDP growth

B. Employment growth

C. GDP growth

D. Employment growth

E. GDP growth

F. Employment growth

NOTE: This figure plots, in the $x$-axis, heterogeneity in import exposure (Panels A and B), export exposure (Panels C and D), and total exposure (Panels E and F), and relates them to GDP growth and employment growth for 2018:Q1-2019:Q1 ($y$-axis) while weighting by state population as of 2019.

SOURCE: Authors’ calculations.
**Figure A4**


A. GDP growth

B. Employment growth

C. GDP growth

D. Employment growth

E. GDP growth

F. Employment growth

---

NOTE: This figure plots, in the x-axis, heterogeneity in import exposure (Panels A and B), export exposure (Panels C and D), and total exposure (Panels E and F), and relates them to GDP growth and employment growth for 2018:Q1-2019:Q1 (y-axis) while weighting by state population as of 2019.

SOURCE: Authors’ calculations.
NOTES

1 Bown and Kolb (2020).

2 In 2019, they accounted for over 60 percent of total trade with the United States.

3 As noted in Santacreu, Sposi, and Zhang (2019), U.S. states are also exposed differently to trade based on characteristics other than their comparative advantage, such as their geography.

4 This measure of trade exposure captures direct effects of international trade policy on each sector. However, we are not taking into account the complete input-output structure, as we do not look at whether the sector is using intermediate inputs from other U.S. sectors that are heavily exposed to imports or exports.

5 In Figure A1 in the appendix, we plot our measure of sectoral trade exposure for all 19 industries in our sample.

6 A better measure of exposure would rely on trade data at the state and sector level. That is, $\frac{X_{j,s}^W}{GO_j^s}$, where $X_{j,s}^W$ represents trade flows of each state $s$ in sector $j$ with the world, and $GO_j^s$ is gross output of state $s$ in sector $j$. However, because of data limitations, we use equation (2).

7 Kolmar (2020).


9 Figure A3 in the appendix shows the results taking into account each state’s population in 2019. Our findings are robust to using weighted regression with state population as the weight.

10 Parilla and Bouchet (2018).

11 China’s implementation of tariffs on the U.S. Agriculture sector follows a recent report from the Congressional Research Service (Regmi, 2019) claiming that China has retaliated in response to U.S. tariffs by adding tariffs to 94 different U.S. food and agricultural products as of April 2018.

12 In Figure A2 in the appendix, we plot sectoral trade exposure to China for all 19 industries in our sample.

13 A recent article on the industries most impacted by the U.S.-China trade war (Reiff, 2019) lists the U.S. automobile industry as the most impacted by tariffs.

14 Tonneson (2017).

15 These states had among the lowest imports from China in 2018. See Kiersz and Heeb (2019).

16 Further evidence from an article on the trade war (Soergel, 2019) lists Washington as one of the top states exporting to China.

17 A recent article published by Bloomberg (Gongloff, 2019) highlights the dangers ahead for Michigan, as the state’s annual growth rate in manufacturing employment has slowed down significantly.

18 Figure A4 in the appendix shows the results taking into account each state’s population in 2019. Our findings are robust to using weighted regression with state population as the weight.

REFERENCES


Santacreu, Ana Maria; Sposi, Michael and Zhang, Jing. “A Quantitative Analysis of Tariffs across U.S. States.” Unpublished manuscript, 2019.


The fraction of national income accruing to labor (the labor share) had been roughly constant in developed economies for much of the 20th century but has fallen since the 1980s. We review several of the leading explanations in the literature for the declining labor share. We then point to hitherto unexplored dimensions of the data and provide suggestive evidence for a new explanation. In particular, we show that the labor share began a steeper descent in 2000. This more recent break in the labor-share trend coincides with the rapid rise of software investment, which has left a larger impact on service industries (than manufacturing) and on high-skill, cognitive occupations (than middle-skill, routine occupations). (JEL E25, O33)

https://doi.org/10.20955/r.102.413-28

1 INTRODUCTION

For much of the 20th century, the fraction of national income accruing to labor (the labor share, hereafter) remained more or less constant over time, at least in developed economies. Keynes (1939, p. 48) wrote that “the stability of the proportion of the national dividend accruing to labour” was “one of the most surprising, yet best-established, facts in the whole range of economic statistics, both for Great Britain and for the United States.” Phelps Brown and Hart (1952) confirmed this fact, which was then cited by Kaldor (1957) and became widely known as one of the six Kaldor facts of economic growth. Although it is true that the labor share did not have an obvious long-term trend, it did exhibit deviations from its long-run level that lasted over a decade, a fact Kaldor was aware of. For example, he cited the medium-run trends in the labor share in the 1930s, as documented by Kuznets (1952).

The labor share in the United States started to decline following its post-war peak in 1970 but, as Figure 1 shows, occasional reversals make it hard to precisely point out the beginning of the downward trend.
In any case, economists did not pay much attention to the emerging downward trend in the labor share through the 2000s. It was only in the aftermath of the 2007-09 Financial Crisis, with heightened interest in economic inequality, that the falling labor share made headlines. The early contributors were Elsby, Hobijn, and Sahin (2013); Karabarbounis and Neiman (2013); and Piketty (2014), whose work motivated many other researchers to propose various explanations for the phenomenon.

Here we provide an overview of the proposed explanations. It is not meant to be an exhaustive list of all the papers on this topic. Rather, we focus on the three distinct, broad approaches that emerged in the literature.

2 EXPLANATIONS IN THE LITERATURE

2.1 Technological Change: Two-Factor Models

The first set of explanations posits that technological progress that unequally affects labor and capital (also known as biased or directed technological change) caused the decline in the labor share. This argument can be understood more clearly with a simple mathematical representation of the production side of the economy.

Consider a production function with two factors, labor \((L)\) and capital \((K)\):
where $Y$ is the output produced and $F$ is a homogeneous of degree 1 function. The elasticity of substitution between labor and capital ($\sigma$) is defined in the usual way, where $W$ is the wage and $R$ is the user cost of capital:

$$\sigma = \frac{d \ln K / L}{d \ln W / R}.$$  

The elasticity of substitution readily captures how factor income shares respond to changes in factor input prices: It is straightforward to see that

$$\frac{d \ln (RK)/(WL)}{d \ln W / R} \leq 0 \iff \sigma \leq 1.$$  

The labor share ($\ell$) is defined as the share of compensation of employees in total value added:

$$\ell \equiv \frac{WL}{PY},$$

where $P$ is the output price. Because of the homogeneity-of-degree-1 assumption, $PY = WL + RK$, implying that the direction of the change in the labor share $\ell$ is the same as the direction of the change in $\frac{WL}{RK}$.

For example, if technological changes make capital cheaper (that is, $\frac{W}{R}$ increases) and the elasticity of substitution between capital and labor $\sigma$ is greater than 1, then formula (3) implies that the labor income share falls.

This is the argument of Karabarbounis and Neiman (2013). They empirically document that the labor share fell across countries, and their estimated elasticity of substitution between labor and capital is indeed greater than 1 (about 1.25). They conclude that investment-specific technological change, which lowers the price of capital relative to consumption (Figure 2), caused the global decline of the labor share.

Piketty and Zucman (2014) make an analogous argument. If the elasticity of substitution $\sigma$ is greater than 1, an increase in the capital-to-labor ratio ($\frac{K}{L}$) will reduce the labor share. This follows from formula (5):

$$\frac{d \ln (RK)/(WL)}{d \ln K / L} \leq 0 \iff \sigma \leq 1.$$  

Since in the data the capital-to-output ratio increased and the labor share fell, Piketty and Zucman (2014) conclude that the elasticity of substitution between labor and capital is greater than 1, especially in recent years.

Of course, these two papers are not the first to show that the labor share can fall because of the substitution between capital and labor. This point has been made by many earlier papers, such as those by Zeira (1998), Acemoglu (2003), Bentolila and Saint-Paul (2003), and Jones (2003), who emphasize the role of biased technological changes. The contribution of the
Karabarbounis and Neiman (2013) and Piketty and Zucman (2014) papers is that they established the declining labor share as an empirical fact and then explained it with the capital-labor substitution channel.

However, it is not always easy to identify the elasticity of substitution parameter, and whether or not the elasticity is greater than 1 is not a settled matter. The direction and the magnitude of technological changes are difficult to observe and are hence typically constructed as a residual. Micro-level estimates of the elasticity use exogenous variations in factor prices, but a micro-level elasticity is not equal to the macro-level elasticity unless all the plants or firms have the same labor and capital intensities. Oberfield and Raval (2014) estimate the micro-level elasticity of substitution between labor and capital using plant-level data. They infer the aggregate elasticity of substitution using a model and the observed distribution of factor income shares across plants. Their estimated elasticity of substitution between labor and capital is less than 1, contrary to the above papers: They find that investment-specific technological changes or capital deepening cannot explain the declining labor share and that one needs to instead look to labor-augmenting technological change for an explanation.

2.2 Mismeasured Labor Share

Some studies argue that the fall of the labor share had been overestimated, which virtually obviates the need for an economic explanation.

Koh, Santaulalia-Llopis, and Zheng (2019) show that the change in the coverage of capital in the System of National Accounts (SNA) is responsible for the trend decline of the labor share.
share in the data. In an effort to better measure “intangible” capital, the SNA began to recog-
nize software (SNA 1993) and research and development (SNA 2008) as investments rather
than expenses; both of these are now called intellectual property products (IPP) capital. Once
IPP capital was recognized as investment, it had to be included in value added and in the
income account has since been counted as capital income. From the definition of the labor
share (4), with a larger denominator but the same numerator, the labor share decreases mechani-
cally. Because investment in IPP capital has been growing in the United States, as shown in
Figure 3A, the inclusion of IPP in value added results in the fall of the measured labor share
over time. Figure 3B shows the labor share with and without IPP: the green line is the labor
share with IPP treated as investment (the current accounting method), and the orange line is
the labor share with IPP treated as expenses (the pre-1999 accounting method). Since the labor
share without IPP is more or less flat at least until the 2000s, the divergence between the two
lines is evidence that IPP is an important factor in the decline of the labor share.

The argument so far is one of technological change (that is, the rising importance of IPP)
and not mismeasurement. Koh, Santaeulalia-Llopis, Zheng (2019) then question the plausi-
bility of assigning all IPP investment to capital income. Their view is that a significant portion
of investment in intangible capital in reality should accrue to labor, but in many cases the
compensation of workers employed to create intangible investment is in the form of equity
claims and not counted as labor income. In other words, the labor compensation for IPP pro-
duction is likely underestimated and the rising share of IPP has exacerbated this underestima-
tion over time, generating the downward trend in the labor share in the data and the divergence
between the green and the orange lines in Figure 3B.

In the same vein, Eisfeldt, Falato, and Xiaolan (2019) use compensation data for a subset
of firms in the U.S. economy (i.e., those in Compustat) and demonstrate a rise in equity-based
compensation (e.g., stock options) as a fraction of total employee compensation. Since equity-
based compensation is not consistently counted as labor compensation in the National Income
and Product Accounts (NIPA) of the U.S. Bureau of Economic Analysis (BEA), they argue
that this underestimation of labor income could account for up to 60 percent of the observed
decline of the labor share.

Another source of underestimation of labor income is the classification of self-employed
income. To begin with, conceptually, it is not clear what fraction of a business owner’s profit
is compensation for their own labor or return on capital. Using U.S. tax data, Smith et al. (2019)
show that top earners’ entrepreneurial labor income is reported as “pass-through” business
profit for tax reasons and counted as capital income, even though typical top earners are shown
to derive the majority of their income from labor. Clearly, this is another source of underesti-
mation of the labor share.

More broadly, adjusting for the labor income of the self-employed, Gutiérrez and Piton
(forthcoming) show that the labor share has remained stable in most countries, with the
exception of Canada and the United States.

In summary, a careful consideration of the data suggests that the labor share may not have
fallen as much as previously reported. However, we note, even with adjustments, one clear
pattern has not been explicitly mentioned in the literature: The labor share in the United States
Figure 3
IPP and the Labor Share in the United States

A. IPP investment as a share of gross domestic income

IPP investment divided by gross domestic income

B. Labor share with and without IPP

Labor income with and without IPP as share of gross domestic income

NOTE: Panel B: Labor share without IPP is compensation of employees divided by gross domestic income net of IPP investment.

SOURCE: Authors’ calculations based on data from NIPA.
has fallen steadily since the 2000s. (The literature emphasizes the general downward trend since the 1980s.) This decline is shown in Figure 4 by the breaks in the linear trends and can be also seen in Figure 3B, where both the green (IPP as investment) and orange (IPP as expenses) lines show clear downward trends since the 2000s. We will revisit this timing of the trend break in the labor share in Section 3.

2.3 Compositional Effects: Rising Market Concentration

The final set of papers does not so much explain the declining labor share as link it to another dimension of economic inequality: the rising market power of large firms. Departing from the homogeneity assumption of Section 2.1, where value added gets divided into only labor income and capital income, we now allow for markup, or profit. Holding other things equal, a higher markup implies a smaller share of value added accruing to labor.

Unlike the above literature that typically relies on macro data, Autor et al. (forthcoming) use firm-level micro data for the United States and show that large firms have higher markups (and hence lower labor shares) than small firms, and that since the 1980s larger and more-productive firms have gotten even larger and raised their market share. As economic activity has become more concentrated in a small set of “superstar” firms with low labor shares, the aggregate labor share has fallen. They point to globalization and technological changes as the possible causes of the rising market concentration.
Kehrig and Vincent (2018) find similar patterns among manufacturing establishments (plants). They show that since the late 1960s, there has been a massive reallocation toward “hyper-productive” low-labor-share establishments in the U.S. manufacturing sector and that this reallocation accounts for all the decline in the labor share in the manufacturing sector.

This concludes our reading of the leading explanations in the literature for the declining labor share, and we now explore dimensions of the data that have not been carefully studied in the preceding papers.

3 A NEW PERSPECTIVE ON THE DECLINING LABOR SHARE

As summarized in the previous section, many studies have advanced our understanding of the decline of the labor share. However, as we noted for Figure 4, the decline has been more conspicuous since the 2000s, to which the preceding work did not pay particular attention: Again, the literature typically dates the beginning of the labor share decline to the early 1980s. In this section, we investigate three dimensions of the data—sectors, occupations, and types of capital—that could be particularly relevant to the recent trend break.

3.1 Sectors

At the sector level, the labor share declined mostly in the manufacturing sector, as pointed out by Alvarez-Cuadrado, Long, and Poschke (2018). This fact led some researchers to focus on manufacturing industries. For example, Acemoglu and Restrepo (2019) argue that automation caused the labor share to fall particularly in the manufacturing sector.

To be more precise, we write the aggregate labor share ($\ell$) as the average of the labor shares in the manufacturing and service sectors, weighted by the value-added share of each sector:

$$
\ell = \frac{CE_M + CE_S}{VA_M + VA_S} = \frac{VA_M}{VA_M + VA_S} \times \frac{CE_M}{VA_M} + \frac{VA_S}{VA_M + VA_S} \times \frac{CE_S}{VA_S},
$$

where $CE_i$ is the total compensation of employees (labor income) in sector $i = M, S$ and $VA_i$ is the value added of sector $i$. $M$ stands for the manufacturing sector and $S$ for the service sector.

As shown in Figure 5C, the monotonic decline in the value-added weighted labor share of the manufacturing sector (blue line; first term in the right-hand side of equation (6)) coincides with the gentle downward trend in the aggregate labor share since 1980 (see Figure 4). However, we wish to draw attention to the trend break in the value-added weighted labor share of the service sector (red line; second term in the right-hand side of equation (6)) around 2000.

Over the sample period, the decline in labor income in the manufacturing sector relative to aggregate value added had been partly offset by the rising relative importance of the service sector, but this partial offset ceased abruptly around 2000, after which the labor income in the service sector relative to aggregate value added remained more or less constant. This implies that the trend break in the aggregate labor share around 2000, and the most obvious decline in the labor share throughout the sample period, has more to do with what happened to the service sector than the manufacturing sector.¹
To summarize, while the weak downward trend in the aggregate labor share since 1980 is driven by the decline in the value-added weighted labor share of the manufacturing sector $\left( \frac{CE_{ML}}{VA} \right)$, the cause for the break in the trend and the more conspicuous decline of the aggregate labor share in the 2000s may have been caused by changes in the service sector.

**Figure 5**

**Value-Added and Labor Shares for Manufacturing and Services**

**A. Value-added share** $\frac{VA_i}{VA}$

**B. Sectoral labor share** $\frac{CE_i}{VA_i}$

**C. Value-added weighted labor share** $\frac{VA_i}{VA} \times \frac{CE_i}{VA_i}$

NOTE: The dashed lines are fitted linear trends.

SOURCE: Authors’ calculations based on data from NIPA and BEA industry accounts.
3.2 Occupations

At the occupation level, employment has polarized since 1980 (Acemoglu and Autor, 2011; Autor and Dorn, 2013; and Lee and Shin, 2017). The most widely accepted explanation for this polarization is that recent technological changes, especially automation technology, have replaced routine or middle-skill occupations. This replacement is a form of the substitution of labor with capital, discussed in Section 2.1, and can explain the labor share decline. Acemoglu and Restrepo (2019) introduce a task-based model in which technological progress enables capital to replace labor in certain tasks and argue that the displacement effect of auto-

**Figure 6**

**Occupational Contributions to Labor Share: Manual, Routine, and Cognitive**

A. Manual

<table>
<thead>
<tr>
<th>Year</th>
<th>Manual</th>
<th>Routine</th>
<th>Cognitive</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTE: Decomposition is $\ell_j = \frac{W_j L_j}{VA}$ for $j$ = manual, routine, and cognitive. $W_j L_j$ is computed by $CE \times \omega_j$, where $CE$ is the total compensation of employees in NIPA and $\omega_j$ is the earnings share of occupational group $j$ obtained from the U.S. Census Bureau’s American Community Survey. Panel D (All) plots $\frac{CE}{VA}$.

SOURCE: Authors’ calculations based on data from NIPA and the U.S. Census Bureau’s American Community Survey.
information has caused the fall of the labor share since the 1980s, especially through the fall in routine jobs.

In contrast, in order to understand the conspicuous downward trend in the labor share since 2000, we move our focus to another occupational group: cognitive-intensive, or high-skill, occupations. We decompose the aggregate labor share into occupational groups. However, because we do not have value added by occupation, we divide occupation-level total labor compensation by the aggregate value added ($W_jL_j/VA_j$, where $j$ indexes occupations).

From the occupational contributions to the labor share, shown in Figure 6, we see that the trend break in the aggregate labor share around 2000 can be entirely attributed to the stagnation of cognitive-intensive occupations, although routine-intensive occupations have been responsible for the gentle, longer-run downward trend.

### 3.3 Software

Given the above findings, the natural next step is to ask what changed since 2000 for the cognitive-intensive (rather than routine) occupations and the service (rather than manufacturing) sector. Our work suggests that the rise of software investment can potentially explain the sector- and occupation-level labor share trends after 2000.

The rise of information technology (IT) since 1980, both hardware and software, has gotten much attention in the literature. However, it is not as well known that hardware and software have exhibited diverging patterns since 2000. Figure 7 depicts shares of software and computer equipment investment in total non-residential investment. We see a dramatic rise.
The growth of software investment closely tracks the increased demand for cognitive-intensive occupations. Aum (2019) shows that software use is concentrated among cognitive-intensive occupations (shown as Figure 8 here). He argues that, however, software innovation could eventually lead to a lower demand for cognitive-intensive occupations. Intuitively, software innovation makes cognitive-intensive occupations relatively more productive than other occupations, because they use software more intensively. When occupations are complementary to one another, higher relative productivity of cognitive-intensive occupations increases the demand for other occupations. Furthermore, if software substitutes for cognitive-intensive occupations, software innovation will reduce the labor income share, especially for cognitive-intensive occupations, consistent with the evidence in Figure 6. Note that these arguments rely on the complementarity among occupations and the substitutability between cognitive-intensive occupations and software.

Turning to the sector level, we exploit industry-level variations across 60 BEA industry classifications to examine whether the decline in the labor share between 2000 and 2010 is...
### Table 1
Relationship Between Changes in the Labor Share and Potential Explanatory Variables, Across Manufacturing Industries

<table>
<thead>
<tr>
<th></th>
<th>Changes in labor share between 2000 and 2010</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Changes in labor share between 2000 and 2010</td>
<td></td>
</tr>
<tr>
<td>Import competition</td>
<td>-0.023* (0.013)</td>
<td>0.003</td>
</tr>
<tr>
<td>Intermediate offshoring</td>
<td>-0.020 (0.014)</td>
<td>0.003</td>
</tr>
<tr>
<td>Task offshorability</td>
<td>-0.023 (0.015)</td>
<td>-0.010</td>
</tr>
<tr>
<td>Routineness</td>
<td>-0.011 (0.011)</td>
<td>0.007</td>
</tr>
<tr>
<td>Computer intensity</td>
<td>-0.049** (0.018)</td>
<td>-0.024</td>
</tr>
<tr>
<td>Software intensity</td>
<td>-0.134*** (0.022)</td>
<td>-0.149***</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.140</td>
<td>0.089</td>
</tr>
<tr>
<td></td>
<td>0.103</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>0.264</td>
<td>0.451</td>
</tr>
<tr>
<td></td>
<td>0.502</td>
<td></td>
</tr>
</tbody>
</table>

NOTE: Standard errors are in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.
SOURCE: Authors’ calculations based on data from BEA input-output tables, BEA industry accounts, O*NET Tools and Technology, and the U.S. Census Bureau’s American Community Survey.

---

### Table 2
Relationship Between Changes in the Labor Share and Potential Explanatory Variables, Across Service Industries

<table>
<thead>
<tr>
<th></th>
<th>Changes in labor share between 2000 and 2010</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Changes in labor share between 2000 and 2010</td>
<td></td>
</tr>
<tr>
<td>Import competition</td>
<td>-0.023 (0.027)</td>
<td>0.000</td>
</tr>
<tr>
<td>Intermediate offshoring</td>
<td>-0.022 (0.023)</td>
<td>0.002</td>
</tr>
<tr>
<td>Task offshorability</td>
<td>-0.014 (0.018)</td>
<td>0.007</td>
</tr>
<tr>
<td>Routineness</td>
<td>0.015 (0.018)</td>
<td>0.020</td>
</tr>
<tr>
<td>Computer intensity</td>
<td>-0.018 (0.019)</td>
<td>0.012</td>
</tr>
<tr>
<td>Software intensity</td>
<td>-0.134*** (0.022)</td>
<td>-0.149***</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.022</td>
<td>0.027</td>
</tr>
<tr>
<td></td>
<td>0.017</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>0.027</td>
<td>0.524</td>
</tr>
<tr>
<td></td>
<td>0.585</td>
<td></td>
</tr>
</tbody>
</table>

NOTE: Standard errors are in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$.
SOURCE: Authors’ calculations based on data from BEA input-output tables, BEA industry accounts, O*NET Tools and Technology, and the U.S. Census Bureau’s American Community Survey.
associated with six factors (explanatory variables): import competition, intermediate offshoring, task offshorability, routineness, computer (hardware) intensity, and software intensity.  

Tables 1 and 2 show the regression coefficients of each of the six explanatory variable for the manufacturing and service industries, respectively. We normalized the variables (i.e., subtracted the mean and divided by the standard deviation) so that the coefficient estimates are comparable.

For manufacturing industries, import competition, computer intensity, and software intensity show significance. However, when all variables are included in the regression (last column, Table 1), only software intensity remains significant. For services, software intensity is the only variable that shows significance (Table 2). Since the flattening of the labor share in services contributed the most to the sharp drop in the aggregate labor share after the late 1990s, this analysis suggests that software can be an important part of the explanation for the trends in the labor share.

### 4 CONCLUDING REMARKS

The decline of the labor share has been actively studied in recent years. We have reviewed the leading explanations for the decline of the labor share in the literature. However, the literature emphasizes the gentle decline since the 1980s and has been silent about the more stark fall since 2000. Taking a detailed look at sectors, occupations, and types of capital, we provide evidence that service industries and cognitive occupations merit more attention in order to fully understand the trends in the labor share. In particular, separating software investment from other types of capital promises to be a fruitful avenue for research.
NOTES

1 From the decomposition \( \Delta (\omega \LS) = \LS \Delta \omega + \bar{\omega} \LS \), we find that the change in the sectoral value-added share and the change in the sectoral labor share contributed nearly equally to the flattening of the value-added weighted labor share of services.

2 We classify occupational groups into cognitive, routine, and manual occupations based on the one-digit Standard Occupational Classification codes of the U.S. Bureau of Labor Statistics. Cognitive occupations are managers, professionals, and technicians. Routine occupations are machine operators, transportation workers, sales and office personnel, mechanics, and miners and production workers. Manual occupations are low-skill service occupations.

3 Hardware and software facilitate workers' tasks differently. Figure 8 demonstrates that software usage varies across occupations (correlated with cognitive task intensity) much more than computer usage (which is basically whether a computer is used or not).

4 Aum, Lee, and Shin (2018) model computer capital (both hardware and software) separately from other types of capital and allow the elasticity of substitution between labor and computers to differ from that between labor and other types of capital. Their calibration strategy did not target the labor share, but the elasticity of substitution between labor and computers came out greater than 1.

5 Each variable is computed as follows. Import competition for industry \( i \) is \( \text{import}_i / (\text{output}_i + \text{import}_i - \text{export}_i) \). Intermediate offshoring is intermediate use of \( j \) in \( i \) times \( \text{import}_j / \text{output}_j \). Task offshorability is the share of offshorable jobs in industry \( i \), where offshorable jobs are those with an occupational ranking of offshorability greater than the 66th percentile (Autor and Dorn, 2013). Routineness is the share of occupations in industry \( i \) with an occupational ranking of routine-task intensity greater than the 66th percentile (Acemoglu and Autor, 2011). Computer intensity is computer hardware investment divided by industry value added. Finally, software intensity is software investment divided by industry value added.

REFERENCES


While there is a renewed literature connecting internal migration to various issues related to structural transformation such as urban labor and housing markets, the relationship between internal migration and demographic transition is much under-studied despite its importance in the process of economic development. Our article fills this knowledge gap. By constructing a simple dynamic framework in which fertility and rural-urban migration decisions are both determined, we show that more rapid urban productivity advancement can lead to a positive relationship between migration and fertility. Using cross-country data analysis, we support our theory, establishing that both migration and fertility rates are higher in less-developed countries than in advanced economies. Our results imply that policies that may help reduce the cost of urban living or enhance urban benefits would be useful for productive structural transformation. (JEL E24, O15,R23)
productivity (TFP) growth may be viewed as a key driver that may transform a less-developed economy from “Malthus to Solow,” using the terminology in Hansen and Prescott (2002).

Accordingly, we provide a simple dynamic framework in which both fertility and rural-urban migration decisions are endogenously determined. More-rapid advancement in urban TFP relative to rural TFP induces structural transformation from farming to manufacturing. At the balanced-growth spatial equilibrium, there are opposing forces at work for the comparative static outcomes of migration and fertility. We establish conditions under which migration intensity (defined as migration flow divided by the rural population) and the total fertility rate turn out to be positively related, in response to more-rapid urban TFP advancement. The economics are intuitive: When urban TFP continues to rise over time, a higher urban wage encourages migration from rural to urban areas and leads to a fall in the rural-to-urban (rural-urban) population ratio. The decline in this ratio has two opposite effects on migration intensity. First, it causes a direct drop in migration intensity, which we call the population base effect. Second, because the urban fertility rate is lower than the rural fertility rate, the decline in the rural-urban population ratio also leads to a decline in the total fertility rate. Given the same migration flow, the decline in the rural-urban population ratio increases migration intensity, which we call the population growth effect. As long as the direct population base effect on migration intensity dominates the indirect population growth effect, an ongoing rise in relative urban TFP can deliver a positive relation between migration intensity and the total fertility rate. In this case, an economy with a higher relative income is associated with a combination of low migration intensity and low fertility.

The policy implication of our findings is important for economic development. We find that rural-urban migration not only shifts labor from agriculture to more-productive non-agriculture sectors, but is also associated with a decline in fertility accompanied by higher per capita income as a result of labor reallocation across locations. Our results imply that policies that may help reduce the cost of urban living or enhance urban benefits would be useful for productive structural transformation.

Finally, we conclude the article by pointing out the knowledge gap, especially what existing theory may still fail to address and what may be potentially rewarding for future studies.

2 LITERATURE REVIEW

In this section, we provide a critical review of the new strand of the literature connecting internal migration to various issues related to structural transformation. Bencivenga and Smith (1997) and Banerjee and Newman (1998) represent the first generation of this strand of research with thorough modeling of internal migration dynamics in the presence of informational asymmetry. While Bencivenga and Smith (1997) find adverse selection of workers into urban areas as a result of asymmetric information, Banerjee and Newman (1998) show lower credit availability in the process of internal migration, due to higher agency costs. Both articles are interesting, but asymmetric information is not viewed as the primary driver of the great divergence in the speed of urbanization.

Lucas (2004) pioneers the second generation of this strand of research and focuses on explaining the great divergence. More specifically, cities enable new migrants to accumulate
human capital for better earnings, but rural areas remain active due to the presence of a specific factor: land. When more rural workers migrate to urban areas, the remaining rural farmers own more cultivatable land and thus enjoy a rising marginal product and higher returns. Eventually, rural-urban migration ceases when the values of earnings in rural and urban areas equalize. The speed of urbanization thereby depends on the return to human capital accumulation in urban areas relative to land productivity in rural areas. Using a search and matching framework, Laing, Park, and Wang (2005) stress that reductions in labor market frictions may reduce the likelihood of unemployment, induce higher wage offers, and lead to a faster process of urbanization. More recently, Bond, Riezman, and Wang (2016) argue that trade liberalization together with reductions in migration barriers are key drivers of rural-urban migration. While the effect of migration barriers is clear, reduced trade costs induce firms in developing countries to produce more capital-intensive, import-competing goods rather than low-skilled-labor-intensive exportables. The resulting increase in urban productivity thus induces faster urbanization and leads to faster growth.

The third generation of this strand of research not only broadens the spectrum of issues examined but also conducts more comprehensive quantitative studies than the second generation. Following Lucas (2004) in stressing the importance of human capital, Liao et al. (2017) highlight urban education as an incentive for migration, finding that education-based migration could be more crucial than work-based migration in the case of China, where attending colleges mitigates large mobility barriers. Liao et al. (2020) further point out that, despite the lower childrearing cost in rural compared with urban areas, cities provide better opportunities for both economic and non-economic activities. Liao et al. thus develop an internal migration model featuring a locational quantity-quality trade-off of children and find that stricter population control policies in Chinese cities may not be ideal. Particularly, such policies may reduce migration incentives of workers with stronger preferences toward having children, subsequently leading to distorted outcomes in migration and fertility. Focusing on urban housing booms in China, Garriga et al. (2017) find that housing price hikes are largely fundamental, driven by urban TFP-induced rural-urban migration and amplified by continual reductions in migration barriers. Ngai, Pissarides, and Wang (2019) claim that the household registration system prevented labor from moving out of the agricultural sector and subsequently slowed down the process of structural transformation and industrialization in China. The rationale is that the household registration system did not fully secure tenure rights to land, resulting in an inefficient land rental market, underestimation of the effective urban income, and overemployment in the low-productivity agricultural sector. Moreover, the provision of social transfers that were conditional on the area of registration discouraged “floating workers” and subsequently yielded underemployment in the urban non-agricultural sector. Finally, by generalizing the Bond-Riezman-Wang framework to multiple regions, Tombe and Zhu (2019) find that, in the absence of capital, internal reforms on trade and migration are more important for enhancing growth in China than international trade.

While an important factor interacting with internal migration is demographic transition, it is unfortunately under-studied. There are a few exceptions. In Sato and Yamamoto (2005), the decline in infant and child mortality, as typically observed in demographic transition, is a
main driver for urbanization. Sato (2007) further argues that when the substitution effect and 
the income effect on fertility offset each other, the interaction of urban agglomeration economies (Lucas 1988) and urban congestion with the fertility-work trade-off can generate a negative relationship between income and fertility across different regions. More recently, Cheung (2018) proposes rural education reform as a critical force leading to demographic transition accompanied by a shift from rural farming to urban manufacturing. Liao et al. (2020) examine how work-based rural-urban migration and fertility decisions interplay in the process of economic development when the economy exhibits large migration frictions and population controls. Our article complements Sato and Yamamoto (2005) and Sato (2007): While their articles focus on internal migration and local fertility patterns across different regions in a country, our article characterizes rural-urban migration and overall fertility of a country over time.

3 DATA ANALYSIS

In this section, we use real gross domestic product (GDP) per capita from Penn World Table (PWT) 9.0, the total fertility rate from the World Development Indicators (WDI), and migration intensity from Bernard, Bell, and Cooper (2018, Table A4.1) to study the cross-country patterns of internal migration and fertility. To be in line with previous studies cited in this article, our data analysis focuses exclusively on the levels of migration intensity and fertility.

The migration intensity data reported in Bernard, Bell, and Cooper (2018) are based on census data from 1996 to 2011. Migration intensity refers to crude migration intensity of major areas.\(^3\) In an earlier work, Bell et al. (2002) provide more detailed discussion on the definition of crude migration as well as the related adjustments needed for ensuring comparability of the cross-country data.

Accordingly, we focus on cross-country patterns for the period 1996 to 2015. The United States is set as the benchmark country, so relative income is calculated as real GDP per capita relative to that of the United States. We capture different development stages of sample countries based on their initial development stage measured by relative income in 1996 and their development achievement measured by relative income in 2015.

Table 1 provides the summary statistics of total migration intensity and the total fertility rate. Their relationship is illustrated by the scatter plot in Figure 1. Here, the total fertility rate refers to the “middle year” of our sample period, 2006. The main finding is that the total fertility rate is positively correlated with migration intensity, with a correlation coefficient of 0.2411. Moreover, based on the scatter plot, one can conclude that people in developing countries are more likely to have more children and are more likely to migrate than their counterparts in developed countries. This positive correlation between fertility and migration intensity lends empirical support to the model developed in this article.

Liao et al. (2019) find that migration intensity decreases moderately with initial relative income in 1996 (with a correlation coefficient of –0.0767) and that the total fertility rate falls sharply with it (with a correlation coefficient of –0.5965). They also show that when both nega-
Table 1
Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Total migration intensity</th>
<th>Total fertility rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>6.328%</td>
<td>2.956</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>3.947</td>
<td>1.436</td>
</tr>
<tr>
<td>Coefficient of variation</td>
<td>0.624</td>
<td>0.486</td>
</tr>
</tbody>
</table>

SOURCE: Authors’ calculations using data from the WDI and Bernard, Bell, and Cooper (2018).

Figure 1
Migration Intensity and the Total Fertility Rate, 2006

NOTE: The dotted line is the linear regression line.
SOURCE: The WDI and Bernard, Bell, and Cooper (2018).

tive correlations are plotted against final relative income in 2015, the correlations are slightly weakened over the span of 20 years (with the correlation coefficients dropping to −0.0728 and −0.5475, respectively). Taking these results together, we may still infer a positive association between fertility and migration intensity. We turn now to constructing a model to rationalize this relationship.
4 THE MODEL

Time is discrete, indexed by $t$. Consider a two-location economy that extends the locational stratification model of Bénabou (1996) and the surplus labor model of Bond, Riezman, and Wang (2016). One location is called the urban area: It has most of the economic activity and is indexed by superscript $U$. The other is called the rural area: It plays a passive role and is indexed by superscript $R$. The economy is populated with two cohorts of two-period-lived overlapping generations, referred to as the young and the old. Each individual is endowed with one unit of productive time when young. To simplify the setup, we normalize our model to be populated only by females. The population is growing and depends on each individual’s fertility choice at the beginning of the second period of life. Similar to Becker (1960), who was the first to introduce the quantity of children into parents’ preferences, we assume that an individual is altruistic, valuing her own consumption when old ($c$) and number of children ($n$). Furthermore, because the quantity-quality tradeoff of children is not the focus of our underlying mechanism of rural-urban migration, we follow Sato (2007) to abstract from the quantity-quality choice of fertility decisions. As such, the only cost considered here is a resource cost of childrearing at the rate of $\varphi_j > 0$ per child, $j = R, U$.

The economic activity in the rural area is stripped down to the bare necessities. Basically, rural production is just backyard farming, yielding a crop income of $w_R$ in units of the urban good, which is also rural workers’ implicit self-employment wage. The lifetime utility of a rural farmer is given by

$$U_R = \ln c + \beta \ln n^R = \ln \left( w^R - \phi^R n^R \right) + \beta \ln n^R,$$

where $\beta \in (0,1)$ is the altruistic factor (common to all individuals). Thus, the optimization problem is

$$\max_{n^R} \ln \left( w^R - \phi^R n^R \right) + \beta \ln n^R,$$

which implies a first-order condition:

$$\frac{\beta}{n^R} = \frac{\phi^R}{w^R - \phi^R n^R}.$$

Because the marginal benefit (MB) from childbearing is decreasing in $n$ and the marginal cost (MC) is increasing, the second-order condition is met, thus ensuring the solution maximizes lifetime utility. Manipulating this condition gives the solution for fertility:

$$n^R = \frac{\beta w^R}{(1+\beta) \phi^R},$$

which is increasing in the altruistic factor and rural income but decreasing in the childrearing cost.

In the urban area, there is more economic activity. Individuals residing in the urban area are indexed by $i$ and differ in their disutility from work $\delta_i$—that is, a more-able urban worker
suffers less from a utility loss, even though everyone earns the same market wage \( w^U \). Notably, the setting captures the conventional labor-leisure trade-off in which working is costly as a result of reduced leisure time. An agent who resides in the urban area inelastically supplies one unit of labor, which induces a fixed utility loss. Those who would suffer too high a utility loss would thus always prefer to stay in the rural area. For simplicity, we assume that disutility for the urban area \( \delta_i \) is always drawn from the same stationary distribution, which is fixed over time. The lifetime utility of an individual \( i \) living in urban is given by

\[
U^U_i = \ln c_i + \beta \ln n^U_i - \delta_i, \tag{4}
\]

whereas the budget constraint is

\[
c_i + \phi^U n^U_i = w^U, \tag{5}
\]

which can be substituted into lifetime utility to derive

\[
\max_{n_i^U} \left( w^U - \phi^U n^U_i \right) + \beta \ln n^U_i - \delta_i.
\]

The solution is of similar form:

\[
n_i^U = n^U = \frac{\beta w^U}{(1 + \beta) \phi^U}. \tag{6}
\]

Two remarks are in order. First, because disutility from work \( \delta_i \) does not affect the net MB from childbearing, all individuals have the same fertility decision: \( n_i^U = n^U \). Second, urban females bear fewer children than rural females do if the relative childrearing cost is higher in the urban area; that is, \( \frac{\phi^U}{w^U} > \frac{\phi^R}{w^R} \), which is as observed in the real world, as daycare in urban areas is relatively more expansive. We conveniently define the childrearing cost gap as \( \Phi \), so \( \frac{\phi^U}{w^U} / \frac{\phi^R}{w^R} = 1 + \Phi \).

Denote the population in location \( j \) at the end of time \( t \) as \( N^j_t \). With a common solution of fertility within each location and without international immigration, we can write the economy-wide population evolution equations in a parsimonious manner:

\[
N_{t+1} = N^R_{t+1} + N^U_{t+1} = N^R_t \left( 1 + n^R \right) + N^U_t \left( 1 + n^U \right)
\]

\[
= N^R_t \left[ 1 + \frac{\beta w^R}{(1 + \beta) \phi^R} \right] + N^U_t \left[ 1 + \frac{\beta w^U}{(1 + \beta) \phi^U} \right]
\]

\[
= N_t \left[ 1 + \frac{\beta w^R}{(1 + \beta) \phi^R} \right] + N^U_t \frac{\beta w^R}{(1 + \beta) \phi^R} \left[ \frac{\phi^R / w^R}{\phi^U / w^U} - 1 \right],
\]

or

\[
N_{t+1} = N_t \left[ 1 + \frac{\beta w^R}{(1 + \beta) \phi^R} \right] - N^U_t \frac{\beta w^R}{(1 + \beta) \phi^R} \frac{\Phi}{1 + \Phi}. \tag{7}
\]
Now, let us denote \( M_{t+1} \) as net flow migration from the rural to the urban area. Then, the population in each location evolves according to

\[
N_{t+1}^R = N_t^R \left( 1 + n^R \right) - M_{t+1}
\]

\[
N_{t+1}^U = N_t^U \left( 1 + n^U \right) + M_{t+1}
\]

To ensure spatial equilibrium, we specify a locational no-arbitrage condition (LNAC) for the marginal migrant in each period. Under a stationary distribution of disutility types of new borns, the LNAC is given by

\[
U_{i*}^U = U^R,
\]

where \( i^* \) indicates the marginal migrant who feels indifferent between staying in the rural area or migrating to the urban area. That is,

\[
\ln \left( \frac{w^U}{1 + \beta} \right) + \beta \ln \left( \frac{\beta w^U}{(1 + \beta) \phi^U} - \delta_{i^*} \right) = \ln \left( \frac{w^R}{1 + \beta} \right) + \beta \ln \left( \frac{\beta w^R}{(1 + \beta) \phi^R} \right),
\]

or

\[
\delta_{i^*} = \ln \left( \frac{w^U}{w^R} \right) + \beta \ln \left( \frac{w^U}{w^R} \phi^U \phi^R \right) = \ln \left( \frac{w^U}{w^R} \right) - \beta \ln (1 + \Phi).
\]

Intuitively, the larger the urban-rural wage gap \( \frac{w^U}{w^R} \) or the smaller the urban-rural child-rearing cost gap \( \Phi \), the higher the disutility cutoff \( \delta_{i^*} \) and hence the higher the net migration flow from the rural to the urban area \( M \).

To close the model, we solve the labor market equilibrium in both locations. With a rural linear backyard-farming technology \( Y^R = A^R N^R \), the implicit wage is tied to TFP: \( w^R = A^R \). Let there be a continuum of identical and perfectly competitive firms of unit mass in the urban area. The production technology of an urban firm is assumed to take a Romer (1986) form:

\[
Y^U = A^U \left( N^U \right)^{\alpha} \left( \overline{N^U} \right)^{1-\alpha},
\]

where \( \alpha \in (0,1) \) and \( \overline{N^U} \) is the aggregate employment in the urban area, taken as given by each firm but equal to \( N^U \) in equilibrium; that is, \( N^U = \overline{N^U} \) ex post. This is a spatial agglomeration force driven by the Marshallian externality. This simple Romer form implies individual decreasing returns to scale \( \alpha < 1 \) but social constant returns (the powers of \( N^U \) and \( \overline{N^U} \) add up to 1).\(^3\)

Under this setup, the urban wage is simply

\[
w^U = \alpha A^U.
\]

**5 CHARACTERIZATION OF THE SPATIAL EQUILIBRIUM**

We are now ready to characterize the spatial equilibrium. Before establishing the key relationship between fertility and internal migration, we examine several useful urban-rural ratios.
To begin, the urban-rural wage ratio is

\[ \frac{w_U}{w_R} = \alpha \frac{A_U}{A_R}, \]

depending positively on the urban-rural TFP gap and the urban returns-to-scale measure \( \alpha \). The childrearing cost gap and hence the rural-urban fertility differential become

\[ \frac{n_R}{n_U} = 1 + \Phi = \frac{\phi_U}{w_U}, \]

which is decreasing in the urban-rural TFP gap and urban returns to scale, but increasing in the relative childrearing cost in the urban area. Using (13), we can rewrite the LNAC (11) as

\[ \delta_i = \ln \alpha + \ln \left( \frac{A_U}{A_R} \right) - \beta \ln \left( \frac{\phi_U}{\phi_R} \right) \]

\[ = (1 + \beta) \ln \alpha + (1 + \beta) \ln \left( \frac{A_U}{A_R} \right) - \beta \ln \left( \frac{\phi_U}{\phi_R} \right), \]

which is increasing in the urban-rural TFP gap but decreasing in the relative childrearing cost in the urban area, which we conveniently refer to as the spatial equilibrium condition (SEC). The urban-rural output ratio is

\[ \frac{Y_U}{Y_R} = \frac{A_U N_U}{A_R N_R}, \]

whereas the urban-rural per capita income ratio is entirely driven by the TFP ratio:

\[ \frac{y_U}{y_R} = \frac{Y_U / N_U}{Y_R / N_R} = \frac{A_U}{A_R}. \]

We next derive the total fertility rate of the economy as follows. From the population evolution equations (8) and (9), we have

\[ n = \frac{n_R N_R + n_U N_U}{N_R + N_U} = n_R \frac{\frac{N_U}{N_R} + (1 + \Phi)}{N_R + N_U} = n_R \left( \frac{\frac{N_R}{N_U} + \frac{1}{1 + \Phi}}{1 + \frac{N_R}{N_U}} \right). \]

The first equality highlights the fact that the total population growth rate is a simple weighted average of regional population growth rates. Equation (18) also points out that the total population growth rate is increasing in the rural income to childrearing cost ratio via \( n_R \) (income effect) and the rural-urban population ratio (fertility base effect), but decreasing in the rural-urban fertility differential (urban childrearing cost-premium effect). Finally, we define migration intensity as the migration flow divided by the rural population:

\[ m \equiv \frac{M}{N_R} = \frac{N_R}{N_{R+1}} (1 + n_R) - 1. \]
Next, we note that
\[ n_{t+1} = n^R_{t+1} + n^U_{t+1} \]

\[ \left( n - n^U \right) \frac{N_{t+1}}{N_t} = \left( n^R - n^U \right) \frac{N^R_{t+1}}{N^R_t}, \]

or
\[ \frac{N^R_t}{N^R_{t+1}} = \frac{n^R - n^U}{n - n^U} \frac{N^R_t}{N^U_t}. \]

Substituting this relation into (19), we get migration intensity as follows:
\[ (20) \]
\[ m = \frac{\left( n^R - n^U \right) (1+n^R)}{\left( n - n^U \right) (1+n)} \frac{N^R_t}{N^U_t} - 1. \]

Equations (18) and (20) are the key equations for our policy insights on internal migration \((m)\) and fertility \((n)\). While equation (18) focuses on the total fertility rate, equation (20) shows that migration flow affects the migration intensity \((m)\) via two channels: directly through \(N^R_t/N^U_t\) (the population base effect) and indirectly through \(n\) (the population growth effect).

To proceed further, we restrict our attention to a balanced-growth spatial equilibrium (BGSE), where all the growth rates of the level variables are constant. It should be noted that Liao et al. (2019) study the dynamics around the BGSE, thereby enabling a full characterization of the effects of an advance in urban TFP relative to rural TFP. Nonetheless, such dynamic effects turn out to depend on the elasticities that measure the responsiveness of the fertility base and urban childrearing cost, leading to rich outcomes but further complexity. Under the concept of BGSE frequently used in endogenous growth theory, we are able to circumvent such complexity. Accordingly, our aim is to obtain the effects of an improvement in urban TFP relative to rural TFP (a rise in \(A^U\)) on migration intensity \((m)\) and the total fertility rate \((n)\).

Consider that \(\delta_i\) follows a uniform distribution over a compact support \([0, \bar{D}]\), so the density is \(1/\bar{D}\). Denote the two gaps as \(g_A = A^U/A^R\) and \(g_\phi = \phi^U/\phi^R\) and rewrite (15) in dynamic form:
\[ \delta_{i,t} = (1+\beta)\ln(\alpha) + (1+\beta)\ln(g_{A,t}) - \beta\ln(g_{\phi,t}) \]
\[ \delta_{i,t-1} = (1+\beta)\ln(\alpha) + (1+\beta)\ln(g_{A,t-1}) - \beta\ln(g_{\phi,t-1}). \]

Taking differences, we get
\[ \delta_{i,t} - \delta_{i,t-1} = (1+\beta)\ln(1+\gamma_{A,t}) - \beta\ln(1+\gamma_{\phi,t}), \]

where
\[ 1+\gamma_{j,t} = \frac{g_{j,t}}{g_{j,t-1}}, j = A, \phi \]
denote the growth rate of variable \(j\). As a result, the rural-urban migration flow can be solved as follows:
\[ M_t = \frac{1}{D} (\delta_{t,t-1} - \delta_{t,t-2}) = \frac{1}{D} \left[ (1 + \beta) \ln (1 + \gamma_{A,t}) - \beta \ln (1 + \gamma_{\phi,t}) \right]. \]

Recalling (14), we get
\[ \frac{n^R}{n^U} = 1 + \Phi = \frac{g_{\phi,t}}{\alpha g_{A,t}} = \frac{g_{A,t}}{g_{A,t-1}}. \]

At the BGSE, we have constant growth rates so that the \( g_{\phi,t} / g_{A,t} \) ratio must be constant; that is, \( g_t = g \) or
\[ \gamma_{A,t} = \gamma_{\phi,t} = \gamma. \]

This in turn yields
\[ M_t = \frac{1}{D} \ln \left( \frac{g_{A,t}}{g_{A,t-1}} \right) = \frac{1}{D} \ln (1 + \gamma) > 0. \]

As a result, (8) and (9) imply a fall in \( N^R / N^U \). Suppose we consider a continuous increase in urban TFP \( A^U \). Under the restriction of BGSE, it also leads to a proportional increase in \( \phi^U \). Then neither \( n^R \) nor \( (1 + \Phi) \) is affected, so its effect on the total fertility rate \( n \) works only through \( N^R / N^U \). This is the fertility base effect that we highlight in (18):
\[ \frac{\partial n}{\partial A^U} = \frac{\partial n}{\partial \left( N^R / N^U \right)} \frac{\partial \left( N^R / N^U \right)}{\partial A^U} < 0. \]

According to (18), a rise in \( A^U \) increases the urban migration flow, lowers the rural-urban population ratio, and hence lowers the total fertility rate due to the fertility base effect. Thus an ongoing increase in the urban-rural TFP gap reduces the urban fertility rate. The decrease in the urban fertility rate then increases migration intensity indirectly, as shown in equation (20).

We are now prepared to establish the key results of our model based on (20). Consider a long-term trend of structural transformation driven by an ongoing increase in the urban-rural TFP gap \( A^U / A^R \), which is the primary force of the second nature of geography that leads to spatial agglomeration (see Cronon, 1991, and survey articles by Berliant and Wang, 2004 and 2019). From (13), (14), (17), one can see that the urban-rural wage ratio \( (w^U / w^R) \), the urban-rural fertility differential \( (n^U / n^R) \), and the urban-rural per capita income ratio \( (y^U / y^R) \) all rise unambiguously. In spatial equilibrium, the SEC in (15) suggests that a higher urban-rural TFP gap raises the disutility cutoff, thereby encouraging more internal migration from the rural to the urban area. As a result, there is a reduction in the rural-urban population ratio \( N^R / N^U \), which generates the two main effects on migration intensity given by (20). On the one hand, the fall in the rural-urban population ratio lowers the migration intensity via the population base effect. On the other hand, it reduces the fertility base and suppresses the total fertility rate, leading to higher migration intensity via the population growth effect. If the direct population base effect dominates the indirect population growth effect, then (20) yields a positive relation between rural-urban migration and fertility. In particular, the ongoing increase in
in the urban-rural TFP gap \((A^U/A^R)\) yields both low migration intensity and a low total fertility rate thereby lending theoretical support to the empirical correlation between the two measures.

Intuitively, the effect of a rise in urban TFP on the total fertility rate can be observed in equation (18). First, the income effect results in an increase in urban fertility. However, the urban childrearing cost (as a percentage of income) also increases as urban TFP rises. If the later effect dominates the income effect, urban fertility declines. Second, other things equal, a higher urban wage (due to a rise in urban TFP) induces a shift of the population from the rural to the urban area. Because migrants’ fertility is on average lower than that in their original area, total fertility falls. Therefore, a rise in urban TFP leads to a lower total fertility rate. We further use equation (19) to provide the intuition for the overall effect on migration intensity. A rise in urban TFP results in rural-urban migration, so the rural population falls. Given the same migration flow, a lower rural population leads to higher migration intensity. However, a smaller rural population also implies fewer migrants and hence lowers migration intensity. When the later effect dominates, a rise in urban TFP reduces migration intensity. In summary, our result indicates a positive relation between migration intensity and the total fertility rate as urban TFP rises.

Our finding implies that policies that may help reduce the cost of urban living or enhance urban benefits would be useful for productive structural transformation. Such policies include the following: (i) a subsidy for urban childrearing, including provision of low-cost public daycare, (ii) a subsidy for new rural-urban migrants, including public housing assistance, and (iii) better provision of urban benefits to all residents.

But when might this internal migration cease? To address this question, we further examine (15). Let \(\delta_{\min}\) be the minimum support of the stationary distribution of \(\delta_i\); that is, \(\delta_{\min} = \inf \delta_i < \delta^*_i\). Then, internal migration ceases when

\[
\beta \ln \left(\phi^U/\phi^R\right) > (1 + \beta) \ln (\alpha) + (1 + \beta) \ln \left(A^U/A^R\right) - \delta_{\min},
\]

which would happen when urban childrearing becomes unaffordable. That is, a rising urban childrearing cost relative to the rural childrearing cost serves as an anti-agglomeration force in our economy, without any need for other drivers. This complements the literature well. For example, in Lucas (2004), the anti-agglomeration force is rising rural productivity because land is a specific factor only for rural farming. In Sato (2007), the anti-agglomeration force is urban congestion. In Bond, Riezman, and Wang (2016), the anti-agglomeration force is the balance in capital usage between the import-competing and the exporting sectors. In Liao et al. (2020), the anti-agglomeration force is the balance between earnings and child preferences in the presence of heterogeneous altruism. In Garriga et al. (2017), the anti-agglomeration force is housing price hikes. In this article, were we to consider additional spatial diseconomies forces, either internal (say, due to social decreasing returns) or external (say, due to a congestion externality), it is clear that the internal migration process would slow down and rural-urban migration would cease sooner.
Figure 2
Growth Paths of Relative TFP, Relative Childrearing Costs, and Migration Flows

A. Case 1: $\gamma_\phi$ is close to $\gamma_A$ along the transition

B. Case 2: A much faster growing $\gamma_\phi$ along the transition

C. Case 3: Hump-shaped $\gamma_A$ and monotonic $\gamma_\phi$
6 NUMERICAL EXAMPLES

In addition to characterizing BGSE theoretically, we provide three useful numerical examples to illustrate the transitional dynamics for the rural-urban migration flow, $M_t$. In all cases, we begin with setting artificial growth paths for $A$ and $\phi$. Given the growth paths of $A$ and $\phi$ and the parameters $\bar{D} = 1$ and $\beta = 0.52$, we are able to compute the migration flow $M_t$, which may be called migration intensity with a unit mass of total population. The growth paths of $A$ and $\phi$ governed by $\gamma_A$ and $\gamma_\phi$, respectively, and the corresponding migration flows for all three cases are provided in Figure 2.\textsuperscript{6} One model period is equal to one year in the numerical examples.

The three cases begin with the same assumption that the initial growth rate of $A$ is higher than that of $\phi$, $\gamma_{A,0} > \gamma_{\phi,0}$. Then, along the transition, the paths of $\gamma_A$ and $\gamma_\phi$ are all different. The first two cases consider a path where $\gamma_A$ and $\gamma_\phi$ are both hump-shaped and asymptotically approach 1 percent in the BGSE. In particular, in the second case, $\gamma_\phi$ grows faster than $\gamma_A$ in the transition, whereas in the first case, the path of $\gamma_\phi$ is close to that of $\gamma_A$. In the third case, $\gamma_\phi$ remains hump-shaped, approaching 1 percent asymptotically, but $\gamma_A$ rises monotonically to 3 percent asymptotically.

We find that, in all three cases, the transitional dynamics of the migration flows are non-monotone. The migration flows in the first two cases peak at around 8 to 9 percent, before reducing to the long-run BGSE level. In addition, in both cases, it takes more than four decades for the migration flows to reach the BGSE level. However, in the second case, because $\gamma_\phi$ rises much faster than $\gamma_A$, the migration flow does not fall monotonically from the peak to the BGSE level as the first case does. As shown in Figure 2, in the second case, the migration flow overshoots downward from the peak and then moves up toward the BGSE level. In the third case, $\gamma_\phi$ rises much faster to an asymptote higher than $\gamma_A$ and as a result migration ceases after 41 years, yielding a degenerate BGSE.

7 THE WAY FORWARD

In this article, we have developed a simple dynamic model of fertility and internal migration. We have provided conditions to show that, as urban TFP progresses, migration and fertility co-move. Our cross-country data analysis has suggested that migration intensities and total fertility rates are indeed higher in less-developed countries and lower in advanced economies.

Such a dynamic interplay between fertility and migration is interesting, as it can generate a vicious cycle in economic development. Specifically, in a poor country with low manufacture productivity, the total fertility rate is high, due primarily to the rural fertility rate. Without a technology push from the urban modern sector, the economy remains mired in a Malthusian trap. With sufficient technology advancement, however, the incentive to migrate to urban areas rises, starting the urbanization process and raising the urban population. Because fertility rates of urban residents are lower than those of rural farmers, the total fertility rate starts to drop. With the urban cost of living rising over time, the migration rate starts to fall. Thus, the advancement of urban productivity in conjunction with the interplay between fertility and migration can pull a poor country out of a Malthusian trap (a high fertility-migration
nexus) toward modernization (a low fertility-migration nexus). Policies that help reduce rural-urban migration costs may serve the same purpose. In short, rural-urban migration enables a less-developed country to be transformed from Malthus to Solow.

While a simple fertility choice and locational choice framework has been successful in delivering a positive relationship between fertility and rural-urban migration, one may wonder why the relationship between migration intensity and relative income is weaker than that between total fertility and relative income and why both relationships are weakened when relative income is measured by the final value rather than the initial value over the same period. These remain unanswered.

In our work in progress, Liao et al. (2019), we use a richer framework in which households also value children’s education outcomes and their futures. In that framework, we are able to partially address the above unexplored issues by analyzing the dynamic progress of structural transformation. Regardless, we judge that the existing theory lags behind the empirics, particularly in characterizing the long process of demographic transition accompanied by a fairly rapid process of urbanization over the past half a century. It is our belief that addressing these issues would be potentially rewarding.

NOTES

1 As to be discussed in the literature review, there are very few studies connecting fertility and migration in a dynamic setting.

2 Based on a multivariate ordinary least-squares regression analysis on the data from the China Urban Labor Survey of 2001, Werwath (2011) discovers that migrants generally have higher fertility than native urban residents. Thus, there is little doubt that migration and fertility are closely connected.

3 Crude migration intensity of major areas measures the migration between the first subnational geographic levels, such as provinces, out of all of the population 15 years of age and above.

4 In the endogenous fertility literature, there are other ways to model the number of children desired. For example, the dynasty model in Barro and Becker (1989) consider children’s value and Glimm and Ravikumar (1992) value children’s human capital or income in fertility choices. Bloom et al. (2009) develop an endogenous fertility model with a representative female who values the quantity (number) of children but not their quality (education or human capital). Our setup is close to theirs, though our utility function is simply the log-transformation version of that in Sato (2007).

5 This functional form implies that there exists rent, which is assumed to be used to pay for government infrastructure (or given to an absentee landlord), as typically assumed in urban economics. We do not intend to analyze it, because conducting welfare analysis is not the purpose of this article.

6 The three paths of \( \gamma_A \) and \( \gamma_P \) are given by

\[
\begin{align*}
\text{Case 1:} & \quad \gamma_A = 0.01 \left[ 1 + \exp(1 + 0.2 \cdot t - 0.01 \cdot t^2) \right] \\
& \quad \gamma_P = 0.005 \left[ 2 + \exp(1 + 0.26 \cdot t - 0.00995 \cdot t^2) \right], \\
\text{Case 2:} & \quad \gamma_A = 0.01 \left[ 1 + \exp(1 + 0.2 \cdot t - 0.01 \cdot t^2) \right] \\
& \quad \gamma_P = 0.0025 \left[ 4 + \exp(1 + 0.325 \cdot t - 0.00995 \cdot t^2) \right], \\
\text{Case 3:} & \quad \gamma_A = 0.01 \left[ 1 + \exp(1 + 0.15 \cdot t - 0.008 \cdot t^2) \right] \\
& \quad \gamma_P = 0.03 \left[ 1 - \exp(-0.1 \cdot t) \right], \text{ respectively.}
\end{align*}
\]
REFERENCES


FIRST QUARTER

“Worker Diversity and Wage Growth Since 1940”
Makenzie Peake and Guillaume Vandenbroucke

“The College Wealth Divide: Education and Inequality in America, 1956-2016”
Alina K. Bartscher, Moritz Kuhn, and Moritz Schularick

“The Geography of Housing Market Liquidity During the Great Recession”
Matthew Famiglietti, Carlos Garriga, and Aaron Hedlund

“Health and Economic Development from Cross-Country Perspectives”
Ping Wang and Yin-Chi Wang

SECOND QUARTER

“Strategic Review and Beyond: Rethinking Monetary Policy and Independence”
John H. Cochrane

“Is the Phillips Curve Still Alive?”
Brian Reinbold and Yi Wen

“Schools and Stimulus”
Bill Dupor and M. Saif Mehkari

“Taking Stock of the Evidence on Microfinancial Interventions”
Francisco J. Buera, Joseph P. Kaboski, and Yongseok Shin

“On the Aggregate Implications of Removing Barriers to Formality”
Catalina Granda and Franz Hamann

THIRD QUARTER

“A Short Tour of Global Risks”
Carmen M. Reinhart

“Offshoring to a Developing Nation with a Dual Labor Market”
Subhayu Bandypadhyay, Arnab Basu, Nancy Chau, and Devashish Mitra

“Asset Pricing Through the Lens of the Hansen-Jagannathan Bound”
Christopher Otrok and B. Ravikumar

“Reconstructing the Great Recession”
Michele Boldrin, Carlos Garriga, Adrian Peralta-Alva, and Juan M. Sánchez

“The Case of the Reappearing Phillips Curve: A Discussion of Recent Findings”
Asha Bharadwaj and Maximiliano Dvorkin
FOURTH QUARTER

“Responses of International Central Banks to the COVID-19 Crisis”
Jacob Haas, Christopher J. Neely, and William R. Emmons

Ana Maria Santacreu and Makenzie Peake

“Why Is the Labor Share Declining?”
Sangmin Aum and Yongseok Shin

“Fertility and Internal Migration”
Pei-Ju Liao, Ping Wang, Yin-Chi Wang, and Chong K. Yip