Economic Insights

First Quarter 2020
Volume 5, Issue 1

Banking Trends
No More Californias
Regulating Consumer Credit and Protecting (Behavioral) Borrowers
Research Update
Data in Focus
Contents
First Quarter 2020  Volume 5, Issue 1

Economic Insights
A publication of the Research Department of the Federal Reserve Bank of Philadelphia

Economic Insights features nontechnical articles on monetary policy, banking, and national, regional, and international economics, all written for a wide audience.

The views expressed by the authors are not necessarily those of the Federal Reserve.

The Federal Reserve Bank of Philadelphia helps formulate and implement monetary policy, supervises banks and bank and savings and loan holding companies, and provides financial services to depository institutions and the federal government. It is one of 12 regional Reserve Banks that, together with the U.S. Federal Reserve Board of Governors, make up the Federal Reserve System. The Philadelphia Fed serves eastern and central Pennsylvania, southern New Jersey, and Delaware.

Patrick T. Harker
President and
Chief Executive Officer
Michael Dotsey
Executive Vice President and
Director of Research
Adam Steinberg
Managing Editor, Research Publications
Brendan Barry
Data Visualization Manager

ISSN 0007–7011

Connect with Us
We welcome your comments at: PHIL.EI.Comments@phil.frb.org
Twitter: @PhilFedResearch
E-mail notifications: www.philadelphiafed.org/notifications
Facebook: www.facebook.com/philadelphiafed/
Linkedin: https://www.linkedin.com/company/philadelphiafed/

Previous articles: www.philadelphiafed.org/research-and-data/publications/economic-insights

About the Cover
Production Possibility Frontier
This issue’s cover depicts the production possibility frontier, a graphical representation of how much an economy can produce given existing resources. The horizontal axis represents the production of services, the vertical axis the production of goods. Once the economy reaches the curved line, or frontier, running through the graph, there is a tradeoff between the two—it is impossible, at this point on the graph, for the economy to produce more goods without cutting services, and vice versa. If the economy is producing goods and services on a point along this frontier, it is Pareto efficient. If it’s producing goods and services anywhere beneath the curve, it is Pareto inefficient, because it could produce more goods and services, if it so chose.

1
Banking Trends: Do Stress Tests Reduce Credit Growth?
In ensuring that large banks can provide credit during the downturn, stress tests may limit the credit supply today—or so some critics claim. Edison Yu examines both sides of the debate.

8
No More Californias
As American mobility declines, some wonder if we’ve lost our pioneer spirit. Kyle Mangum finds that the situation is more complicated than it appears.

14
Regulating Consumer Credit and Protecting (Behavioral) Borrowers
As we consider how best to protect individual borrowers, Igor Livshits asks: Why and from whom are we protecting these borrowers, and what policies may be most effective?

20
Research Update
Abstracts of the latest working papers produced by the Philadelphia Fed.

25
Data in Focus
Partisan Conflict Index.

1
Banking Trends: Do Stress Tests Reduce Credit Growth?
In ensuring that large banks can provide credit during the downturn, stress tests may limit the credit supply today—or so some critics claim. Edison Yu examines both sides of the debate.

8
No More Californias
As American mobility declines, some wonder if we’ve lost our pioneer spirit. Kyle Mangum finds that the situation is more complicated than it appears.

14
Regulating Consumer Credit and Protecting (Behavioral) Borrowers
As we consider how best to protect individual borrowers, Igor Livshits asks: Why and from whom are we protecting these borrowers, and what policies may be most effective?

20
Research Update
Abstracts of the latest working papers produced by the Philadelphia Fed.

25
Data in Focus
Partisan Conflict Index.
Banking Trends

Do Stress Tests Reduce Credit Growth?

Stress tests are supposed to ensure your access to credit during the next downturn, but some critics claim that they also limit your access to credit today. We test that theory.

BY EDISON YU

As we approach the 10th anniversary of the nation’s first supervisory stress test, some analysts argue that stress tests have gone too far and that large banks have inefficiently restricted credit. This article explores the preliminary evidence about the effects of stress tests on the credit supply. However, before considering the evidence, we need to know how the stress tests work in the U.S. and why the stress tests might reduce credit growth.

What Is a Stress Test?
The goal of supervisory stress tests is to ensure that systemically important banking institutions are adequately capitalized under even very adverse economic conditions. Stress tests use models to estimate a bank’s need for capital under these conditions. Among other benefits, stress tests ensure that large banks can provide credit to households and firms in a downturn, thus reducing the severity of the downturn.

To restore public confidence in the largest financial institutions at the height of the financial crisis in 2009, the Federal Reserve and other banking supervisors implemented the first stress test, the Supervisory Capital Assessment Program (SCAP), which estimated the potential losses that would be incurred by the largest U.S. banks if economic and financial conditions worsened.

Under SCAP, supervisors determined whether the largest financial institutions in the U.S. had sufficient capital to weather the recession and worsening financial conditions. They assessed 19 financial institutions’ capital buffers based on potential macroeconomic scenarios in 2009 and into 2010. Building on SCAP, the U.S. implemented two related stress test programs: the Dodd–Frank Act Stress Test (DFAST) and the Comprehensive Capital Analysis and Review (CCAR) program.

DFAST was created by the 2010 Dodd–Frank Wall Street Reform and Consumer Protection Act (“Dodd–Frank”), which required annual supervisory stress tests for all financial institutions that met two criteria. First, the institution had to have total consolidated assets of more than $50 billion. And second, its primary regulator had to be federal. In addition to the supervisory tests, large
banking organizations, or bank holding companies ( BHCS), are also required to run internal stress tests.

Congress raised the threshold of the supervisory tests to $100 billion in 2018. As of that change, BHCS with consolidated assets between $100 billion and $250 billion are now only subject to periodic supervisory stress tests. (Banks with total consolidated assets of more than $250 billion are still subject to annual supervisory stress tests.)

The Federal Reserve conducts DFAST with its own independent models to project a bank’s income, loan loss, and capital level over a nine-quarter planning horizon under three different hypothetical scenarios of the aggregate economy. The three scenarios—baseline, adverse, and severely adverse—hypothesize future economic outcomes, including recessions of different magnitudes. For example, in the severely adverse scenario, the U.S. falls into a deep recession with a large increase in unemployment and sharp declines in asset prices.

Each bank subject to the supervisory stress tests submits detailed information about its balance sheet to the Federal Reserve. For each hypothetical scenario, the Federal Reserve forecasts the bank’s pre-provision net revenue and the potential amount of losses due to adverse economic conditions. After calculating taxes and capital distributions such as dividends, the Federal Reserve projects banks’ regulatory capital ratios over the nine quarters of the test. The Dodd-Frank Act requires the Federal Reserve to publicly disclose the DFAST results, but it does not require any supervisory actions for banks whose projected capital falls below regulatory minimums.

The more comprehensive CCAR program applies to the biggest and most complex financial institutions, with assets of at least $100 billion. Through 2019, CCAR, like DFAST, has been conducted annually by the Federal Reserve to ensure that the largest and most complex financial institutions have sufficient capital to continue normal operations in times of economic and financial distress. In 2019, the 18 largest financial institutions were subject to CCAR.

CCAR includes both a quantitative assessment and a qualitative assessment. The quantitative assessment starts with banks submitting financial information and their capital plans to the Federal Reserve. The assessment includes tests run by the banks and the supervisory tests run by the Federal Reserve. The quantitative assessment uses the projections of income

Changes in DFAST Thresholds

The thresholds of stress test requirements have changed more than once. In 2009, banks with consolidated assets over $100 billion were subject to the SCAP. Nineteen banks underwent the 2009 supervisory stress test.

Originally, Dodd-Frank required all financial institutions with total consolidated assets of more than $50 billion and whose primary regulator is a federal financial agency to be subject to annual supervisory stress tests. In addition, banks with assets over $10 billion are required to run internal stress tests. In May 2018, Congress passed the Economic Growth, Regulatory Relief, and Consumer Protection Act, which increased the asset thresholds for the stress tests. Effective from the 2019 stress test cycle, banks with assets less than $100 billion are no longer subject to stress tests. Banks with assets between $100 billion and $250 billion are subject to periodic supervisory stress tests, while banks with assets of over $250 billion are subject to annual supervisory stress tests and are required to conduct periodic internal company-run stress tests. As a result, the number of banks tested in the DFAST program decreased from 35 in 2018 to 18 in 2019.

This article focuses on the effects of supervisory stress tests, but some of the cited articles use information about the internal stress test results for their statistical analysis.
from DFAST and incorporates banks’ planned capital actions, such as dividend payments and stock repurchases. A quantitative objection is based on whether a bank maintains capital ratios above regulatory minimums under both the projections by the Federal Reserve and the bank’s own projections. In the qualitative assessment, the Federal Reserve evaluates how the banks identify, measure, and determine capital needs for their material risks. Until 2019, the Federal Reserve could issue an objection to the banks’ capital plan based on either the quantitative or the qualitative assessment, but as of 2019 the Federal Reserve has eliminated the qualitative component for most banks. Unlike under DFAST, supervisory actions can be taken if the Federal Reserve objects to a bank’s capital plan under CCAR. When this happens, the bank may not make any capital distribution without the Federal Reserve’s permission. (See Figure 2.)

Unlike a point-in-time capital requirement, the supervisory stress tests look to the future. Financial regulations such as Basel III typically require banks to maintain a sufficient current percentage of their balance sheet as capital. The stress tests, on the other hand, focus on future capital planning, ensuring banks have sufficient capital to maintain lending during a major shock to the economy or firms.

**How Do Stress Tests Affect Lending?**
To avoid receiving a CCAR objection from the Federal Reserve, a bank needs to hold more capital or reduce its assets to keep its capital ratio above regulatory minimums. A bank can increase its capital holdings by either selling more stock, reducing capital distribution, or increasing retained earnings. Alternatively, a bank can reduce its total assets by making fewer and smaller loans and buying fewer and smaller securities. If a bank chooses not to increase its capital holding, then it must reduce the size of its assets to avoid a CCAR objection, potentially reducing lending to households and firms. (See Figure 3.)

But stress tests may also prompt a bank to shift the composition of its portfolio. In an economic downturn or during financial distress, banks typically lose more money on riskier loans. Thus, banks that have riskier loans on their portfolio must keep more capital on hand in order to pass the stress test. Since holding more capital is costly, stress tests encourage banks to avoid risky borrowers and make safer loans even in good times.

One important goal of stress tests is to ensure that banks can continue their normal operations in a time of distress, when higher loan losses reduce bank capital. The higher capital provision during good times takes into account the potential capital needed due to loan losses in a time of distress. This can help a bank absorb the larger losses and smooth the credit supply during an economic downturn.

Some critics argue that the stress tests have gone too far and inefficiently limit the credit supply, especially to risky but profitable borrowers. After all, banks are in the business of taking and managing risks, not just making ultrasafe loans. Other critics argue that the stress tests might reduce credit to riskier borrowers. In addition, due to the higher capital requirement of the stress tests, banks may search for higher-interest returns by making riskier loans in order to compensate for the higher capital costs.

So far we have focused on the impact of stress tests on bank lending. But not all loans are made by banks subject to the stress tests, or, for that matter, by banks. The overall aggregate impact of stress tests on lending depends on the extent to which borrowers can obtain credit from smaller banks or nonbank lenders instead of from larger banks. For example, if borrowers could get all their mortgages from tech companies or Quicken Loans rather than from banks, mortgages overall may be unaffected even as banks make fewer mortgage loans.

Recent empirical work tests these claims.

**Empirical Evidence**
A fast-growing body of empirical literature studies the impact of stress tests on bank lending. And many of these studies try to find out whether stress tests impede credit growth. These papers use different methods and focus on different loan markets, such as mortgages, commercial and industrial lending, and small-business loans.

However, regardless of method or focus, it is challenging to study the effects of stress tests on bank lending. Supervisory...
stress tests were first implemented right after the financial crisis, when many banks were losing money and the economy and regulations were changing, so it is difficult to isolate the effects of the stress tests by simply comparing lending outcomes before and after they were implemented. Furthermore, regulators only stress-test larger banks, making it difficult to identify which differences in lending outcomes are due to stress tests and which are due to the different sizes of these banks.

### Comparison of Stress-Tested and Non-Stress-Tested Banks

Despite these empirical challenges, some papers compare lending growth and loan characteristics between stress-tested banks and non-stress-tested banks.

In their 2018 paper, Viral Acharya and his coauthors compared banks subject to stress tests with those that were not. They focused on the syndicated loan market and used DealScan data on syndicated loan origination from 2004 to 2014. They found that banks subject to stress tests reduced their credit supply (particularly credit to relatively risky borrowers) and that banks subject to stress tests extended smaller loans, shortened loan maturities, and charged higher spreads. This is all consistent with banks lowering the risk of their loan portfolios. They found similar results using the bank-level data from the Call Reports. In addition, by using the data on small-business loans collected under the Community Reinvestment Act (CRA), they found that stress-tested banks originated fewer small-business loans. Because small-business loans are riskier, they argued, the stress-tested banks' decision to reduce small-business lending was evidence that stress tests reduce the supply of risky lending. In the last empirical exercise of the paper, the authors showed that bank-level measures of risk, such as the tier 1 capital ratio, improved after a bank was subjected to stress tests.

In their 2017 working paper, Paul Calem and his coauthors also compared stress-tested banks to non-stress-tested banks, but they focused on mortgage markets. They used Home Mortgage Disclosure Act (HMDA) data and studied jumbo-loan origination activities of banks from 2009 to 2014. From the banks’ perspective, jumbo mortgage loans are riskier because they cannot be sold to government sponsored enterprises (GSEs) such as Freddie Mac and Fannie Mae. (By definition, a jumbo loan is larger than what a GSE is willing to buy.) Accordingly, they are not subject to the GSEs’ underwriting standards and are usually held in the bank’s loan portfolio. They found that, immediately following the 2011 CCAR stress test, banks subject to supervisory stress tests originated fewer jumbo mortgages as a total share of the banks’ mortgages and had lower jumbo mortgage approval rates. In particular, the paper estimated that stress-tested institutions’ share of jumbo mortgage origins was 5 to 7 percentage points lower in 2011. But the effects are not statistically significant for the other years. They argued that the subsequent effects were small because banks had become better capitalized and hence the supervisory stress tests were no longer binding.

In his 2018 paper, Francisco Covas explicitly addressed the concern that the stress-tested banks are also the largest banks, which are subject to a range of capital requirements. He showed that, for most banks, the capital requirements imposed by the stress tests are higher than other capital requirements, such as the point-in-time risk-based capital requirement imposed by Basel III for some classes of loans. In particular, the capital charges imposed under the stress tests are particularly stringent for small-business loans and residential mortgages, so Covas suggested that stress-tested banks might shift lending away from small-business loans and mortgages. By using Call Report data from 2011 to 2016, he found that growth in small-business lending was significantly slower for banks after they were subject to stress tests. In particular, he estimated that the U.S. supervisory stress tests led to a 4 percentage point reduction in the annual growth of small-business loans secured by nonfarm, nonresidential properties.

### Using an Instrument to Isolate the Effects of Stress Tests

Although it seems intuitive to compare lending outcomes of stress-tested and...
non-stress-tested banks, drawing accurate conclusions can be
difficult because other factors are at play. Banks subject to stress
tests are primarily very large, and it is possible that these big
banks differ from smaller banks in other aspects that also affect
lending growth. The different lending outcomes between large
and small banks may thus be due to those other factors and
not to the stress tests. Simply comparing stress-tested and non-
stress-tested banks without accounting for these other factors
may lead to biased estimates.

Papers that use this comparison approach attempt to deal
with this problem by taking into account a host of observable
factors. However, the statistical problem may persist if their
statistical analysis fails to capture unobserved variables. For
example, larger banks are subject to other, stricter regulatory
requirements, such as higher leverage requirements and living-
will requirements. Some of these stricter requirements are
difficult to measure and quantify, but they could affect the lend-
ing supply, making it difficult to isolate the effects of stress tests.

To address this concern, a second group of papers constructed
an instrument that measures how strongly the regulations
pressured each stress-tested bank to adjust its lending behavior. In
their 2018 working paper, William Bassett and Jose Berrospide
constructed a measure called the capital gap, which is the differ-
ence between the capital level required according to the
supervisory stress tests and the level of capital from the bank’s
own stress-test model. The larger the capital gap, the more
additional capital banks need to hold to pass the supervisory
stress tests.

Note that this measure avoids the problem of comparing
the largest banks to smaller banks and is also quite specific to the
stress-testing exercise, so the effect of the shortfall is plausibly
distinct from other supervisory requirements. The authors also
argued that banks have a limited ability to manage this gap,
because the models used for the supervisory tests by the Federal
Reserve are not disclosed to the banks. Hence the capital
shortfall is likely to be random and not correlated with other
confounding factors, such as the size of the bank, which might
affect lending outcomes. The randomness of the capital gap
that a bank faces is thus useful in statistical analysis for isolating
the effects of stress tests on lending growth.

Bassett and Berrospide used balance sheet data from the Call
Reports from 2013 to 2016 and found no significant relationship
between loan growth and the capital gap. This does not support
the notion that the supervisory stress tests are reducing loan
growth. In addition, they found a small effect of the capital gap
on improving lending standards, as measured by the Senior
Loan Officer Opinion Survey on Bank Lending Practices. Thus,
the authors also found no evidence for greater risk-taking.

Kristle Cortés and her coauthors use a similar approach in
their forthcoming article. They calculate the stress-test exposure
of a bank as the difference between the starting capital level of
a test period and the lowest capital level implied by the severely
adverse scenario of the supervisory stress test. They argue that
a larger value of the exposure indicates a bigger expected decline
in a bank’s equity capital should an economic downturn occur,
and that this would increase the likelihood that the regulators
will pressure the bank to hold more capital. Then they examine
the effects of the stress-test exposures on small-business loan
growth. They argue that the exposure measure is unlikely to be
correlated with unobserved factors, as the exposure measure
is driven by a bank’s entire loan portfolio, and small-business
lending is a small fraction of a bank’s portfolio.

Using the 2012–2015 data on small-business lending provided
under the CRA, Cortés and her coauthors find that banks with
larger stress-test exposure reduced the subsequent supply of the
riskier small-business loans in counties with more employment
risk. But they do not find evidence that stress tests affected
the supply of small-business loans in safer counties with less
employment risk. The paper then investigates the characteristics
of small-business loans, using data from the Survey of Terms of
Business Lending (STBL) from 2013 to 2016. They show that banks
with larger stress-test exposure charged higher interest rates
and shortened the maturity of riskier small-business loans,
evidence that the tested banks reduced the riskiness of their
small-business loans.

### Aggregate Effects on Credit Supply

With the exception of Bassett and Berrospide, the papers above
found evidence that banks more affected by the supervisory
stress tests reduced their credit supply, and none of the papers
found evidence that these banks increased risk-taking. These
banks, however, are not the only bank lenders—the vast majority
of medium-size and small banks are not subject to the stress
tests. Indeed, banks are not the only lenders—for example, firms
may borrow from finance companies or sell bonds that are held
by insurance companies and other intermediaries. Perhaps the
stress tests have simply shifted borrowing away from stress-tested
banks to other banks and to nonbank lenders.

To examine the impact of stress tests on the overall credit
supply, the last group of papers studied the impact of stress
tests on lending in a geographic area in which large banks, small
banks, and nonbank lenders compete to provide loans to both
businesses and households. Studying the impact of stress tests in
a county, for example, allows the researchers to capture the sub-
stitution across types of lenders within the county. If they find
that a bank subject to the supervisory stress test reduces the credit
supply in the county, but that the overall credit supply in the
county does not change, they can infer that borrowers are able to
obtain credit through non-stress-tested banks or other lenders.

Cortés and her coauthors found no reduction of small-business
lending by banks in counties with more exposure to stress tests,
while small banks not subject to stress tests increased their market
share among all banks.7 So the total quantity of small-business
loans made by banks did not appear to decrease.

The data used by Cortés and her coauthors don’t permit an
examination of substitution from bank lending to nonbank
lending. In some markets, particularly for residential mortgages,
nonbank lenders have taken a significant market share in the
postcrisis years. Although they did not isolate the effects of the
stress tests from other factors affecting the largest banks, Brian
Chen and his coauthors were able to provide some evidence
about this margin by using a unique dataset of nonbank loans
through PayNet Inc. They found that the share of originations of
small-business loans by the four largest banks fell from 2010 to 2014, while the market shares for both smaller banks and nonbanks increased relative to those four largest banks.  

Taken together, the evidence suggests that small-business lending has shifted from larger banks to smaller banks or nonbanks while not affecting the overall credit supply at the county level. This implies that the overall vulnerability of the market hasn’t changed but has shifted, although further research is needed to test this hypothesis.

Conclusion
So far, empirical work in the literature has shown post-financial-crisis stress testing tends to reduce the credit supplied by banks more affected by the tests, with the reduction mostly in riskier loans. In addition, there is evidence that the reduction in the credit supplied by the large banks is mostly offset by smaller banks or nonbanks, leading to no overall reduction in the credit supply.

Whether this is optimal for financial stability depends on whether increasing the smaller banks’ or nonbanks’ share of the loan market reduces systemic risk. Stress tests are supposed to bolster the financial stability of the banking system by increasing the capital buffer of the largest banks. If we believe that smaller banks and nonbanks pose less systemic risk to the financial system, shifting credit or riskier lending from large to smaller institutions may improve financial stability. We have not experienced an economic downturn since the stress tests were implemented, so all the empirical work so far uses data collected during an economic expansion. Stress-testing’s effectiveness in ensuring financial stability and lending during a downturn will be tested in the next recession. Future research is needed to examine the efficacy of the stress tests during an economic downturn.

Notes
1 Bank holding companies are the entities subject to the supervisory stress tests. I will call them banks for the remainder of the article.

2 Pre-provision net revenue (PPNR) is defined as net interest income (interest income minus interest expense) plus noninterest income minus noninterest expense. The projection of PPNR includes projected losses due to operational-risk events and expenses related to the disposition of real-estate-owned properties. See “Dodd–Frank Act Stress Test 2019: Supervisory Stress Test Results” for more details.

3 Before publishing the quantitative test results, the Federal Reserve provides each bank with a onetime opportunity to adjust its planned capital distributions after it receives the Federal Reserve’s preliminary estimates of the bank’s poststress capital ratios. The original submitted capital plan, the adjusted capital plan, and the decision of an objection on the final capital plan are published after the adjustment. See “Comprehensive Capital Analysis and Review 2019: Assessment Framework and Results” for more information.

4 The qualitative component still exists for some banks and in some circumstances. For example, if a bank becomes subject to supervisory stress tests for the first time and has not been subject to a qualitative assessment before, the bank would still have to be reviewed by the Federal Reserve through the CCAR qualitative assessment.

5 A bank that receives an objection from the Federal Reserve on its capital plan is colloquially described as “failing” the stress test.

6 Capital ratio is defined as capital divided by its risk-weighted assets. To increase that capital ratio, the bank needs to either increase the numerator (capital) or reduce the denominator (assets).

7 See the 2017 Clearing House report, for example.

8 See the 2018 paper by Viral Acharya and his coauthors for a detailed discussion of the potential impacts of stress tests on credit supply.

9 Syndicated loans are large corporate loans to large corporations. They are often funded by a group of lenders, hence the name. For more information, see Edison Yu’s 2018 article.

10 The quarterly Consolidated Report of Condition and Income (or Call Report) is a report filed with regulators by banks in the U.S. The report summarizes a bank’s financial information, including its balance sheet, regulating ratios, and loan portfolios.
The tier 1 capital ratio is the ratio of a bank's core capital, such as equity and retained earnings, to its risk-weighted assets. It is a key measure of a bank's financial health.

These years include 2009, when SCAP was conducted, and 2011–2014, when CCAR was carried out.

For example, the largest banks are subject to extra capital charges because they are systemically important, the so-called SIFI surcharge.

He estimated the stress-test models used by the Federal Reserve and found that post–stress-test capital requirements are more stringent than the point-in-time capital requirements of Basel III. The models used by the Federal Reserve are not publicly released and hence needed to be approximately estimated in the paper.

Formally, a regression has an endogeneity problem if the explanatory variable is correlated with the error term of the regression (or unobserved variables). The regression-with-endogeneity problem can lead to biased estimators. An instrument can be used to solve this problem. An instrumental variable is one that is not correlated with the error term of the regression but is correlated with the explanatory variable of interest.

Employment risk is measured as the sensitivity of the county unemployment rate to the national unemployment rate.

The exposure variable is the average bank exposure in a given county.

For example, Greg Buchak and his coauthors, in their forthcoming article, find that the nonbank share of the U.S. mortgage market nearly doubled from 2007 to 2015.

The four largest banks are Bank of America, Citigroup, JPMorgan Chase, and Wells Fargo.

See Kohn and Liang (2019) for more details.


References


As American mobility declines, some wonder if we've lost our pioneer spirit. A closer look at the data suggests that the situation is less dire—and more complicated—than it at first appears.

BY KYLE MANGUM

The modern world moves fast, as the cliché goes, but in the U.S. today, people move less frequently than their parents did a generation ago. The decline in mobility is much more than an academic curiosity. Economists widely view labor mobility as the principal mechanism by which regions adjust to local economic shocks. If local industries fall on hard times, workers can leave; in places where labor demand is high, new residents flow in. The decline has therefore generated concern that the economy is less adaptable to local shocks, ultimately resulting in labor misallocation, unrealized output, and lower productivity.

More broadly, the decline runs counter to widely held notions of American culture. The U.S. is a nation of immigrants and pioneers, always on the move in search of better opportunities. Paradoxically, in a time of easy transportation and information access, this nation of pioneers has parked its wagons.

Before we identify a proper policy response, we need to understand why mobility has declined. But to do that, we need to consider the history of population expansion across the North American continent. Since European settlers landed on the East Coast, the population of the U.S. has spread to the West and South. This trend continued well into the 20th century, when sparsely populated outpost towns in places such as California, Florida, and Arizona burgeoned into the major metropolitan areas known today.

This geographic expansion of population throughout the continent was mostly complete by the 1980s. Recent population growth is still far from uniform, but the regional component has diminished; a city’s presence on the West Coast, for example, is no longer a sufficient predictor of its population growth. So the regional reallocation of population has declined, but rarely is that what people mean when they talk about the decline in migration.

There are two senses of migration, a word meaning generically the movement of population from one settlement to another. Net migration is the difference between inflows and outflows of population, whereas gross migration is the total turnover resulting from those inflows and outflows. It is gross, not net, migration.
that has notably declined in recent years. By differentiating between the two, we can better understand why mobility has declined and, if we are to design policy, at least craft it for the right object.

**Go West, Young Man!**

The first clue to understanding the causes of the gross migration decline is its spatial pattern. The decline is substantially different across regions of the country—and not randomly so. The decline has predominantly occurred in cities with typically high rates of turnover, while many low-turnover places have shown no change at all. With high-turnover cities being major sources of inflows to other places, total flows across the system have declined. Thus, the national decline is really the sum of fast-turnover cities slowing down and slow-turnover cities holding steady.

The population of the U.S. has aged during roughly the same period that migration has declined. Older households tend to move less than younger households, making aging an obvious candidate for explaining the decline. It is true that the increase in average household age has contributed to the reduction in the aggregate average rate of migration decline. Aging, however, cannot be the whole story. Researchers have shown that typical aging differences are not quantitatively big enough to generate the observed national decline.1 Perhaps more importantly, the decline is present within age groups, so that young people today, for instance, are also moving less than their parents did at the same age. Moreover, aging has occurred at similar rates across cities, so there is no scope for aging to explain the spatial differences in the decline.2

Instead, what’s important is that the country itself, not just its population, has aged. Cities with high turnover were the population growth destinations of the 20th century in newly developing regions. This growth was the real-world manifestation of the famous 19th century advice, “Go West, young man.” The cities of the Northeast, already well established at the founding of the country, have effectively grown at rates below the national average since then (with a modest bump during industrialization). As the country pushed west and south, newly formed cities grew explosively—Chicago and Cleveland in the late 1800s; Los Angeles, Miami, and San Diego in the early 1900s; Phoenix, Las Vegas, and Orlando in the postwar period.3 (See Figure 1.)

Major technological innovations caused—or at least facilitated—the development of these new regions. Transportation underwent a revolution. Railroads in the 1800s connected the coasts, crisscrossing the continent and making its far reaches accessible for the first time. Automobiles

---

**Gross vs. Net Migration**

<table>
<thead>
<tr>
<th>City</th>
<th>Migration Gross</th>
<th>Net</th>
</tr>
</thead>
<tbody>
<tr>
<td>City A</td>
<td>200</td>
<td>0</td>
</tr>
<tr>
<td>City B</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>City C</td>
<td>700</td>
<td>100</td>
</tr>
</tbody>
</table>

Now imagine a third city, City C. In 1980, 400 people moved into City C and 300 moved out, so in 1980 its turnover was 700 and its net migration was 100. Last year, however, only 200 people moved in and 100 moved out. Now its gross migration is just 300, but its net migration is still 100. That’s what we observe in many formerly fast-growing cities throughout the West and South.

**FIGURE 1**

The Boom Moves South and West

After booming first in the Northeast and Midwest, metro population is booming in the South and West.

Metro area populations, millions of people

<table>
<thead>
<tr>
<th>Northeast Cities</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>New York</td>
<td>15</td>
</tr>
<tr>
<td>Philadelphia</td>
<td>10</td>
</tr>
<tr>
<td>Boston</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Midwestern Cities</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicago</td>
<td>15</td>
</tr>
<tr>
<td>Detroit</td>
<td>10</td>
</tr>
<tr>
<td>Cleveland</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Southern and Western Cities</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Los Angeles</td>
<td>15</td>
</tr>
<tr>
<td>Miami</td>
<td>10</td>
</tr>
<tr>
<td>Phoenix</td>
<td>5</td>
</tr>
<tr>
<td>Las Vegas</td>
<td>5</td>
</tr>
</tbody>
</table>

*Source:* Jonathan Schroeder, Minnesota Population Center, University of Minnesota.
soon followed, along with an expanding highway system that substantially enhanced regional connections. In the later 20th century, air travel further closed the gaps, turning a transcontinental trip into less than a day’s affair.

Developing water technologies made these new regions viable. Water delivery systems (such as the aqueduct serving Los Angeles) were vital to large-scale population growth in the arid West. On the other coast, in damp South Florida, for instance, stormwater control and swamp draining significantly enabled development.

Finally, almost all of these newly developing regions were hot (and sometimes also humid), so the expansion of air conditioning was critical. Besides enhancing household comfort, air conditioning was essential for making viable large-scale buildings like apartment and office towers and manufacturing plants.

The 20th century was then essentially the last movement in the long transition of population expansion across the American continent. Aided by new technologies, unpopulated areas filled with residents relocating from older, colder areas. As the technological shocks abated, and as development blanketed the once-vacant land, rates of population change slowly converged across space. Today, the growing areas are not new cities in unpopulated regions but rather the established midsize, interior cities throughout all regions of the county.

There’s No Place Like Home

Once the westward expansion was complete, an older and arguably more primal tendency became more apparent: On average, all types of people show a preference for their initial locations—an attachment to home. Social scientists have explored this phenomenon by looking closely at trust-based social ties to family and friends. These ties offer nonpecuniary benefits such as the pleasure of close relationships, but also pecuniary benefits such as informal childcare arrangements and financial support in times of personal distress. Moreover, place familiarity—the benefit of “knowing your way around”—can offer myriad pecuniary and nonpecuniary benefits as well.
In principle, home attachment is straightforward and intuitive, but empirically it is difficult to measure what a person considers his or her “home.” One somewhat crude but readily available measure is the U.S. Census question about state of birth. For some people, one’s birth state has little connection to one’s sense of home. Some respondents may not even remember their birth state. Even so, it is a remarkably strong predictor of one’s propensity to migrate. People living near their birthplace show a strong proclivity to remain in their location compared with people born out of state.\textsuperscript{5}

A transplanted population, by contrast, is more transient and more subject to various idiosyncratic changes in circumstance. For example, if someone moved to a new place for a job, and the job dissolves for whatever reason, they are likely to move away. Someone with strong local ties whose job dissolves is more inclined to search locally. Hence, turnover rates are high in growing locations. (See Figures 2 and 3.)

This propensity explains why the end of westward expansion could lead, a generation later, to a decline in mobility. High gross migration was an echo effect following population change. Cities with a large share of out-of-state residents lost a lot of their new arrivals, resulting in high turnover rates. Then, as the major shifts in regional population dissipated, an increasing share of people in newly formed locations were “from there” and less susceptible to leaving, and rates of gross migration fell. So the gross migration decline attracting attention today is actually the secondary effect of population shifts that slowed several decades ago. (See Figure 4.)

**The New Normal**

So perhaps the U.S. is finally in a “long-run spatial equilibrium,” as some have suggested.\textsuperscript{6} The term suggests that households’ incentives to relocate have diminished, either because places are more similar than they used to be,\textsuperscript{7} or structural changes in the economy have caused real estate and labor prices to rationalize spatial differences,\textsuperscript{8} so that, in either case, relative population adjustments across space are no longer necessary.

It is difficult to know whether the country is (or ever will be) truly in such a state, but there is reason to expect that massive population changes across regions—of the degree seen from colonization to westward expansion—will no longer be business as usual. The major differences in regional habitability have diminished. Transportation has crisscrossed the continent, water delivery-and-control infrastructure has been put in place, and air conditioning is ubiquitous. Technologies today focus on speed and efficiency within cities, not on developing new cities. And in the digital age, new technologies are less spatial.\textsuperscript{9}

Population growth today is more balanced across locations compared to the skewness of the early and middle 20th century. Some recently established locations, such as Las Vegas, Phoenix, and Orlando, are still growing at above-average rates, but not at the extreme rates of a generation ago. For the most part, population growth is highest in well-established places with space to accommodate more residents. For example, cities like Atlanta, Charlotte, Dallas, Denver, and Nashville were long-important regional centers that recently achieved major city status on the national stage. Some smaller cities near major metropolitan areas, such as Port St. Lucie, FL, Olympia, WA, and Stockton, CA, are also growing above the national rate.\textsuperscript{9} (See Figure 5.)

And this population growth is occurring more within regions than across regions. To the extent that imbalances exist, growing places are established cities rising in the urban hierarchy, leaving the rest of their home region behind and largely drawing people from within their region.\textsuperscript{9}

**On the Road Again**

Now that we understand why mobility has declined, we can ask, what if anything should policymakers do about it?

If decreasing turnover is the result of more people rationally deciding to remain in place, the decline could be evidence of increasing welfare across the economy. Households no longer have to incur the costs of relocation to find suitable locations for themselves. Deepening family and social capital, especially in

---

*Source: U.S. Census Bureau.*

**FIGURE 5**

Sunbelt Cities Boom

Metros in the West and South have seen much bigger growth in population.

Percent change, 1990–2018

Las Vegas
Atlanta
Charlotte
Dallas
Denver
Nashville
Stockton
Monmouth
Philadelphia
Steubenville
once high-turnover locations, could have a wide range of benefits individually and socially. So maybe policymakers shouldn’t do anything about the decline in mobility.

However, these individually optimal decisions could have negative aggregate consequences. For instance, workers may choose less-productive jobs in their home city so they can be near family, which would be optimal for them but would reduce their labor market output. If such cases are pervasive, it could add up to a knock on aggregate productivity.

It is notable that the migration decline out of high-turnover places has not been seen in older cities developed in previous industrial transitions. Indeed, most examples of struggling labor markets, such as postindustrial cities in the Northeast and Midwest, show no trend at all. To some observers, there is a natural inclination to presume the migration decline as one more force pummeling already-beleaguered cities, but as we have seen, this is not actually how the trend plays out. Those places (as well as some older cities with strong labor markets) show little mobility, and little mobility change, because they already had well-established populations.

In many of these cases, in light of the advantages of personal place attachments, the ideal policy response would not be an incentive to move but rather an enhancement to the productivity in the local job market. Such place-based policies become more appropriate as an economy becomes more locally tied.

Would such place-based policies be sufficient? Or should we also encourage the population shifts America once experienced?

There are two perspectives on this question. One is that the expansion of population across the continent was simply a phase in the life cycle of American development. Unsettled land was available, new technologies made it productive and habitable, and then the land filled with settlement and fixed investment until regions converged to an equilibrium size. Maybe there was nothing uniquely American about high mobility (besides, perhaps, open land) and no reason to desire it now. The wagons reached the coast, and there were no more Californias to settle. In this case, there is no problem for policy to fix.

The second perspective is that population change is unduly restricted by policy failures that create congestion in desirable, productive places. Regulations that make it hard to build new homes increase costs and prevent cities, especially those offering high incomes or many amenities, from adding new residents. Suboptimal urban planning could lead cities to be overly congested and below capacity. This is the more pessimistic perspective, suggesting that restrictions on population growth restrain productivity growth and exacerbate inequalities by prohibiting access to the best spaces. In this case, policy (or perhaps the removal thereof) has more scope to improve welfare. But the goal of these policies is not to encourage people to move more frequently per se but rather to enable desirable cities to accommodate more residents.

These two perspectives are not mutually exclusive, and the reality likely combines the two. The regional transition is mostly complete (subject to the caveat that there is always potential for new shocks), and the new trend in population growth is in the expansion of existing cities (especially those away from the coasts) across various regions. This should assuage the fears raised by the interregional migration decline, and there is really no clear role for policy here anyway. The real question is whether this natural new phase of population growth is producing the optimal distribution of population across cities, especially across cities within each region.

This issue needs to be analyzed carefully. There is nothing inherently good or bad about rates of population growth being similar; indeed, they should be different if some places are better than others. To the extent that there are market failures inhibiting population growth in some places, however, there is a need for a policy intervention. If housing regulations are the result of rent-seeking on the part of current residents, or if additional population would enhance worker productivity, or if poor urban planning leads to unproductive (and unenjoyable) travel congestion, then a “benevolent social planner” would design the infrastructure (physical and legal) to accommodate more people. In many cases, local interests may oppose this (for individually rational reasons), but such “growth positive” policy may nonetheless benefit society. If we are out of Californias—if, that is, there are fewer new places to settle—we must manage the urban frontier with great care.
Although rates of telecommuting have increased, it is still a relatively small form of commuting; by census estimates, 5.3 percent of employed persons "worked from home" in 2018, up from 3.3 percent in 2000 (2018 American Community Survey and 2000 U.S. Census, respectively). Besides, telecommuting cannot sufficiently explain migration trends from workplace, and job relocation is frequently a reason for relocation. This pattern holds within slower-growing regions as well. For example, in the mid-Atlantic, Monmouth, NJ, is growing at about the national rate but decidedly above the rates of nearby New York City and Philadelphia.

Notes
1 See Molloy, Smith, and Wozniak (2011) and Kaplan and Schulhofer-Wohl (2017).

2 The cities with the largest declines have, if anything, aged less than those with smaller or negligible declines.

3 Local industrial booms generated some off-path geographic patterns. For example, Detroit grew later than Chicago—the automobile industry took off in the early 20th century, after Chicago was well-established—and San Francisco grew as a gold rush town before most of the rest of California was populated. However, the common pattern was explosive growth as each city was established and then tapering growth as the city matured.

4 Compared with the middle 20th century, domestic natural increase in population has slowed, and a greater share of new population comprises arrivals from foreign countries. Thus, while local population change in the middle 20th century consisted of relocating Americans born in this country, in the late 20th and early 21st centuries local population change substantially consists of immigration from abroad.

5 Return rates are also high. Those living away from their birthplace are far more likely to return there than are other similar people. This is evidence that initial locations are “special places” to most people. If not for this evidence, the observed inclination to stay put could merely be the result of those people having a stronger distaste for moving (anywhere, ever).

6 See Partridge et al. (2012).

7 See Kaplan and Schulhofer-Wohl (2017).

8 See, for example, Gyourko et al. (2013) and Ganong and Shoag (2017). Partridge et al. (2012), in raising the prospect of a new "long-run spatial equilibrium," found evidence of a reduced population response to observed regional differences in labor markets or amenities.

9 Among new technological advances, telecommuting may be a contributing factor to a migration decline because it detaches residence from workplace, and job relocation is frequently a reason for relocation. Although rates of telecommuting have increased, it is still a relatively rare form of commuting; by census estimates, 5.3 percent of employed persons "worked from home" in 2018, up from 3.3 percent in 2000 (2018 American Community Survey and 2000 U.S. Census, respectively). Besides, telecommuting cannot sufficiently explain migration trends across regions or among occupations with limited scope for working from home.

10 This pattern holds within slower-growing regions as well. For example, in the mid-Atlantic, Monmouth, NJ, is growing at about the national rate but decidedly above the rates of nearby New York City and Philadelphia.

11 It would be naïve to assume that nothing will ever change. Climate change, as one prominent example, could produce new dramatic shocks to habitability, causing a new phase of shifts in population that renders the population weakly attached all over again.

References


Regulating Consumer Credit and Protecting (Behavioral) Borrowers

Public policy debate around consumer credit has focused on consumer protection. But from whom are we protecting these borrowers?

BY IGOR LIVSHITS

Since the financial crisis of 2007–2008, consumer credit has gotten a lot of attention, especially as it relates to consumer protection. And the attention is not just academic: The Consumer Financial Protection Bureau (CFPB) and the Credit Card Accountability Responsibility and Disclosure (CARD) Act, both instituted after the crisis, have dramatically altered the regulatory landscape of the consumer credit industry. A guiding principle behind the creation of this new regulatory environment is that consumers need protection from predatory lending practices.1 This article highlights some of the key considerations underlying the design of such policies and possible pitfalls that arise in implementing them.

In designing any regulation to protect consumers, we need to first answer three questions. First, why do (some) consumers need to be protected? The most basic answer is that (some) consumers make “mistakes,” that is, they make decisions the regulator deems suboptimal. There is a range of causes of these mistakes, including various behavioral biases and a lack of information or attention on the part of the consumer. I argue below that the details of this answer are very important for policy design, as they affect how we answer the next two questions.

Second, whom do the consumers need to be protected from? We consider three possible answers: lenders, more sophisticated borrowers, and themselves.

Third, which policies offer effective protection? Here, the range of answers includes financial education and restrictions on pricing and contracts. The answer depends on the answers to the previous two questions. If the regulations are based on a “wrong” model, well-intentioned policies may backfire, causing harm even to the borrowers they aim to protect. To complicate matters further, protecting some (less sophisticated) borrowers may come at the expense of limiting the (informed) choices of others. As John Campbell put it in his 2016 Ely Lecture, “Financial regulators face a difficult tradeoff between the benefits of regulation to households that make mistakes, and the cost of regulation to other financial market participants.”2

This article briefly reviews the recent and ongoing research on these issues. It is this rigorous economic research that allows
us to formulate effective policies and evaluate the tradeoffs associated with the regulation of consumer finance.

**Why Do Borrowers Need Protection?**

The most conventional insight in standard economics is that well-functioning markets deliver efficient allocations. Economists call this the first welfare theorem, and it assumes that economic agents are fully rational and perfectly informed. If that were true for all households in the consumer credit marketplace, they wouldn’t need protection.

But the data (and common sense) suggest that borrowers are not always fully rational. Many empirical observations may evidence of mistakes (from the point of view of a perfectly rational and informed borrower). These observations include the so-called “debt puzzle”: Laibson et al. (2003) pointed out that 60 percent of all credit card holders carry a balance and pay interest, whereas a standard model of rational borrowers predicts that only 20 percent should do so.

An even more dramatic observation is the “credit card debt puzzle,” documented by Gross and Souleles (2002): Many credit-card borrowers have liquid wealth they could use to fully pay the balance on their credit cards, thus avoiding high borrowing interest rates. The use of other, even more costly borrowing outlets, such as payday loans, is also hard to reconcile with the model of fully rational borrowers, especially when one considers how often these presumably very short-term loans turn into extended indebtedness. Even the failure of many heavily indebted households to utilize personal bankruptcy, as documented by White (1998), may be evidence of limited rationality (or limited information).

Interventions in consumer credit markets are thus typically motivated and justified by the idea that borrowers make “wrong decisions,” or “mistakes.” These mistakes may arise from either limits to borrowers’ rationality, their incorrect beliefs, or lack of information. Behavioral economics is the study of these deviations from the assumptions of standard (neoclassical) economics. Three behavioral deviations have received the most attention: the “present bias,” temptation preferences, and incorrect beliefs. All three apply to consumer finance.

Because these three behavioral deviations help explain the empirical puzzles, they are a natural starting point for answering the question “Why do borrowers need protection,” and for designing consumer protection in credit markets.

**The Three Behavioral Biases**

The classic example of behavioral deviation in consumer credit is the idea that borrowers do not fully value or plan for the future, which economists refer to as “time-inconsistent preferences.” Individuals subject to this bias fail to obey their own financial plans when those plans are optimal from the rational perspective. Or at least they want to deviate from these best-laid plans. This essentially defines the time-inconsistency of preferences. The so-called “present bias” is a typical manifestation of time-inconsistent preferences. It refers to consumers’ elevated desire to consume instantly rather than postponing consumption even by a single period.

Experimental evidence supports the conclusion that present-bias preferences shape human behavior. More importantly for our purposes, present-bias preferences help explain a number of aggregate phenomena in consumer credit markets. Laibson et al. (2003) argued that present-bias preferences are needed to reconcile an otherwise standard model with the “debt puzzle”—the fact that 60 percent of credit card holders used their cards to borrow, far more than a model with standard time preferences would imply. Skiba and Tobacman (2019) argued that the present bias (which naïve borrowers are unaware of) is essential for explaining consumer behavior in the payday-loan market.

Another behavioral deviation that justifies interventions in consumer credit markets is temptation preferences. Models that incorporate these preferences assume that individuals suffer from temptation and have to exercise costly self-control to resist it. Temptation preferences help explain a number of otherwise puzzling observations. Gathergood and Weber (2014) used survey data to argue that self-control problems (for example, impulsive spending behavior) are the driving force behind the “co-holding puzzle.” As documented by Gross and Souleles (2002), many individuals carry balances (and pay interest) on credit cards while having liquid funds in low- or no-interest bank accounts.

This form of behavioral bias has a distinct set of policy implications. Nakajima (2017) pointed out that policies that restrict consumers’ ability to borrow may benefit them by limiting their temptation to consume early. Nakajima (2012) also pointed out that by considering temptation preferences, we may dramatically alter how we think of the secular increase in consumer credit over the last half-century. In the presence of temptation, rising indebtedness is not a sign of better consumption smoothing but rather of overborrowing as individuals succumb to temptation.

The third deviation is incorrect beliefs or information. This category bundles together such behavioral biases as overconfidence, overoptimism, and “cognitive limitation” in assessing prospective contract terms or the market environment. These biases’ key common feature is that they directly lead borrowers to make financial “mistakes”—decisions that their fully rational, fully informed selves would disagree with. The justification for an intervention from a (better-informed) regulator is thus clear.

**From Whom Do Borrowers Need Protection?**

Politicians and consumer advocates often portray lenders as culprits, and regulatory responses and proposals certainly take aim at lenders’ practices (see, for example, the Credit CARD Act). One illustrative quote comes from Bar-Gill and Warren (2008): “Sellers of credit products have learned to exploit the lack of information and cognitive limitations of consumers.”
From the point of view of economic modeling, this presumes that lenders have monopoly power that allows them to exploit behavioral borrowers. Indeed, Ausubel (1991) argued that the credit card market displays signs of collusion among lenders, and Herkenhoff and Raveendranathan (2020), pointing to the profitability of transaction services, proposed a model of limited competition. But I view the consumer credit market in its current state as highly competitive.

Even so, contracts offered by competitive lenders may still be predatory. Competitive lenders can offer exploitative contracts in equilibrium, if borrowers are willing to accept such contracts. Bar-Gill (2012) made the important observation that, in a competitive environment, lenders have little choice but to cater to borrowers’ tastes, with all their biases and miscalculations. This reasoning implies that policymakers need to protect borrowers from themselves.

But there’s someone else who may take advantage of behavioral borrowers: other, “sophisticated” borrowers. That point was well illustrated by Heidhues and Köszegi (2010). Sophisticated borrowers benefit from favorable prices that are subsidized by the mistakes made by their behavioral peers. As a modeling approach, this answer offers a helpful alternative to blaming lenders (and demonstrates that policies benefiting one group of borrowers may disadvantage another).

What Policies Offer Effective Protection of Behavioral Borrowers?
The choice of policy instruments should be informed by a specific market failure or behavioral bias. Furthermore, it has to take into account (equilibrium) market responses of both lenders and borrowers, which may undo or offset the intended effects. Failure to do so may result in policy backfiring—doing more harm than good.

Available policies include restrictions on pricing (for example, interest caps or restrictions on teaser rates), restrictions on the set of available contracts (for example, limiting payday loans or the lock-in features of long-term contracts), information provision and counseling, and various wedges (for example, restricting which mortgages qualify as conforming).

Interest rate caps (also known as usury laws) are widely adopted though often sparsely enforced. These restrictions can be justified either as limiting the ability of lenders to exploit their monopoly power or as protecting behavioral borrowers from undertaking excessively costly (that is, excessively risky or excessively large) loans.

Restricting the kinds of contracts allowed in the marketplace is another popular policy measure. The Credit CARD Act, for example, is one set of such restrictions for credit cards. These policies are often motivated by the (perceived) lack of accurate information on the part of consumers, who may misunderstand either details of the contract they are offered or the probability of triggering certain aspects of the contract, such as late fees.

Another policy that can address such lack of understanding is financial education, regarding both contract details and the propensity of borrowers to be subject to penalty clauses. This is the kind of policy prescription that arises from Heidhues and Köszegi (2010).

Lastly, rather than prohibiting certain contracts, policymakers can use price wedges to make some contracts more or less attractive. These wedges can range from taxes on certain activities (making them more expensive) to de facto subsidies for more desirable contracts. One example of the latter is the de facto subsidy from government-sponsored enterprises (such as Freddie Mac and Fannie Mae) that applies only to conforming (desirable) mortgages.

Cautionary Tales: How Well-Intentioned Policies Can Backfire
Not all policies designed to protect the consumer actually do so. These well-intentioned policies are more likely to fail if they misidentify the underlying behavioral friction or ignore markets’ reaction to the policy. Unfortunately, these failures are not unusual.

Cuesta and Sepulveda (2019) convincingly argued that the introduction of interest rate caps in Chile led to a dramatic decline in consumer welfare. The reduction in the interest rates induced by the policy was not enough to compensate for the dramatic reduction in the number of loans issued, even in the most monopolistic submarkets.

Limiting the set of contracts is definitely a double-edged sword. Restricting lock-in clauses in contracts may help protect behavioral borrowers who are unaware of their biases. But the same policy harms behavioral borrowers who are aware of their bias and thus may want to use lock-in features (such as large penalties for missing or adjusting payments) to discipline their behavior by preventing themselves from overconsuming in the future.

Even financial education requirements are not necessarily a slam-dunk policy prescription. Allcott et al. (2019) documented that the majority of borrowers take on seemingly exploitative contracts (payday loans) with their eyes wide open, fully aware not only of the costs but also the likelihood that they will have to roll these debts into yet another round of payday loans. And financial counseling may be costly to prospective borrowers, especially in terms of the time they would need to devote to it. Kilborn (2016) argued that mandatory counseling for bankruptcy filers, implemented in Canada in 1992 and in the U.S. in 2005, is ineffective and misguided. While well intended, it seems to have only made bankruptcy more costly for the most vulnerable segment: single parents who had to not only make time and pay for the counseling sessions, but also find and pay for child care.
When it comes to addressing borrowers’ overoptimism, Exler et al. (2019) argued that none of the basic policies improves the well-being of behavioral borrowers.² Although overoptimistic individuals borrow too much and default too little or too late, policies that bluntly discourage borrowing or encourage default backfire and make all (even behavioral) borrowers worse off. Surprisingly, even “financial literacy” intervention can be counterproductive, including for behavioral borrowers—it helps these borrowers avoid mistakes, but it also shuts down cross-subsidization from rational borrowers to their behavioral peers.

Despite such examples of policy failures, other policies do protect consumers.

Agarwal et al. (2015) found that implementation of the Credit CARD Act yielded a substantial decline in fees paid by borrowers, especially those with low credit scores, with no evidence of an off-setting increase in interest rates or a reduction in access to credit.

In an example from a different type of intervention, Carlin et al. (2019) documented how an introduction of a mobile app, which facilitated individuals’ access to their financial information, led to a significant reduction in high-interest debt and bank fees. This suggests that some form of financial education may indeed be effective. It also points to the effectiveness of subtle “nudge” policies.

### Details Matter

Behavioral borrowers’ awareness of their own biases is critical for the design of policy remedies. Although unaware behavioral borrowers may be made better off (from a paternalistic perspective) by a regulation that limits the set of contracts available to them, behavioral borrowers who are aware of their biases are more likely to be hurt by such regulations. An “aware” behavioral borrower may choose a credit card with high financing charges (or a mortgage with high refinancing costs) over more flexible products specifically in order to address their own behavioral bias, by, for example, preventing their future selves from indulging in excessive consumption. On the other hand, these same contracts can be simply predatory when borrowers are unaware of their biases.

One aspect of the consumer credit market makes it distinct from other markets, such as cellphone contracts, where exploitation of behavioral consumers is a concern: the possibility of default. In many settings—including those with overoptimistic households, for example—behavioral borrowers are more likely not to repay their debts than are their sophisticated, fully rational peers. This difference in default rates implies that when the two types of borrowers take on the same contract, rational borrowers tend to subsidize behavioral borrowers, and not the other way around.

This point makes all the difference in policy prescriptions resulting from Heidhues and Köszegi (2010), who abstracted from the possibility of default, versus those from Exler et al. (2019), who treated default explicitly as a possible outcome.³ For example, financial education is unequivocally beneficial in Heidhues and Köszegi (2010) but may backfire in Exler et al. (2019). Indeed, Exler et al. (2019) argued that, rather than being exploited by their rational peers, behavioral borrowers may instead benefit from being pooled with less risky, rational borrowers.

### Conclusion

Policy prescriptions depend critically on the details of the economic environment. Specifics of the behavioral biases that motivate the intervention, borrowers’ awareness of their biases, the extent of competition in the marketplace, the presence of fully rational borrowers, and the prevalence of default—they all matter when identifying the right regulation or intervention. This points to the importance of both empirical analysis of borrowers’ behavioral biases and theoretical analysis of the equilibrium responses of all market participants to any potential market intervention.

### Notes

1 One of the key objectives of the Credit CARD Act was the elimination of so-called “gotcha” clauses in the fine print of credit card contracts. I am not too proud to admit that I got “caught” by at least four of the credit card features subsequently outlawed by the Credit CARD Act. And I tend to think of myself as a sophisticated and attentive consumer.

2 According to the Financial Crisis Inquiry Commission (2011), this argument is one reason why federal regulators didn’t rein in mortgage market excesses in the run-up to the crisis (p. 93).

3 Admittedly, Telyukova and Wright (2008) and Telyukova (2013) offered a resolution of this puzzle without relying on behavioral assumptions.

4 See Carter et al. (2011) and Skiba and Tobacman (2019).

5 Less prominent but still interesting puzzles include “credit smoothing” (as opposed to consumption smoothing), documented by Hundtofte et al. (2019), and overborrowing in response to windfalls, documented by Olafsson and Pagel (2019).

6 Lack of competition may also justify policy interventions as it distorts allocations, leads to inefficiencies, and allows lenders with monopoly power to take advantage of borrowers. However, arguments by Ausubel (1991) and Herkenhoff and Raveendranathan (2020) notwithstanding, the consumer credit market is quite competitive, as discussed later in this article. This is not an exhaustive list of reasons for regulation. See Elul and Gottardi (2015) for an example of a very different motivation.

7 Gathergood (2012) provided survey evidence that behaviors associated with these biases (namely, impulsive spending, heavy discounting, and financial illiteracy) are associated with overindebtedness and financial distress.

8 O’Donoghue and Rabin (1999) offered the accepted formal definition: “When considering trade-offs between two future moments, present-biased preferences give stronger relative weight to the earlier moment as it gets closer.”
9 See Benhabib et al. (2010), Meier and Sprenger (2010), Balakrishnan et al. (2017), and Bisin and Hyndman (2020), just to name a few.

10 Also known as “the credit card debt puzzle.” See Telyukova and Wright (2008) and Telyukova (2013).

11 Nakajima (2012) focused on the staggering increase in the revolving debt from practically zero in 1969 to 7 percent of GDP in 2009. (Today, credit card debt amounts to about 5 percent of GDP.) Increases in total consumer debt (which excludes mortgages) and total household debt were less dramatic but still substantial (from 12 percent of GDP to 19 percent today for consumer debt, and from 43 percent in 1982 to almost 100 percent at the peak for total household debt).

12 See Grubb’s (2015) discussion of the distinction between over-confidence, which he calls “overprecision,” and overoptimism.

13 A more promising approach to studying this aspect of the market could be a search model of limited competition along the lines of Drozd and Nosal (2008), Nosal and Galenianos (2015), Drozd and Serrano-Padial (2013, 2017), and Raveendranathan (2019). But this branch of the literature is still nascent.

14 The Credit CARD Act prohibits “universal default” (increasing the interest rate on one card in response to a delinquency on another one) and retroactive interest-rate increases. It also restricts “two-cycle billing,” the marketing of credit cards on university campuses, credit limits offered to young borrowers (under 21 years of age), and changes to interest rates and other fees (for credit cards and gift cards). Under the Act, lenders must also apply payments to the balance with the highest interest rate, and they must disclose how long it would take to repay the balance by making only minimal payments.

15 Even a mortgage prepayment penalty (or closing fee) may serve as such a commitment device by making cash-out refinancing less attractive.

16 Overoptimism has been documented in various forms and settings. Overoptimism regarding individuals’ income is documented by Arabsheibani et al. (2000), Dawson and Henley (2012), and Balasuriya et al. (2014). Gathergood (2012) offered evidence of unforeseen expenditures, which amounts to overoptimism regarding expenses. Weinstein (1980) found that people generally underestimate the probability of negative events for themselves.

17 In other words, the model in Heidhues and Kőszegi (2010) ruled out default by assumption: All debts are always repaid. Exler et al. (2019), by contrast, explicitly modeled default as a possibility, thus reversing some key forces, such as the direction of cross-subsidization.

References


Research Update

These papers by Philadelphia Fed economists, analysts, and visiting scholars represent preliminary research that is being circulated for discussion purposes.

Home Equity in Retirement

Retired homeowners dissave more slowly than renters, which suggests that homeownership affects retirees’ saving decisions. We investigate empirically and theoretically the life-cycle patterns of homeownership, housing and nonhousing assets in retirement. Using an estimated structural model of saving and housing decisions, we find, first, that homeowners dissave slowly because they prefer to stay in their house as long as possible but cannot easily borrow against it. Second, the 1996–2006 housing boom significantly increased homeowners’ assets. These channels are quantitatively significant; without considering homeownership, retirees’ net worth would be 28–44 percent lower, depending on age.


Bayesian Estimation and Comparison of Conditional Moment Models

We provide a Bayesian analysis of models in which the unknown distribution of the outcomes is specified up to a set of conditional moment restrictions. This analysis is based on the nonparametric exponentially tilted empirical likelihood (ETEL) function, which is constructed to satisfy a sequence of unconditional moments, obtained from the conditional moments by an increasing (in sample size) vector of approximating functions (such as tensor splines based on the splines of each conditioning variable). The posterior distribution is shown to satisfy the Bernstein-von Mises theorem, subject to a growth rate condition on the number of approximating functions, even under misspecification of the conditional moments. A large-sample theory for comparing different conditional moment models is also developed. The central result is that the marginal likelihood criterion selects the model that is less misspecified, that is, the model that is closer to the unknown true distribution in terms of the Kullback-Leibler divergence. Several examples are provided to illustrate the framework and results.

Financial Constraints of Entrepreneurs and the Self-Employed

Growth-oriented entrepreneurial start-ups generate more economic growth than other self-employed businesses, yet they only constitute a small fraction of start-ups. We examine whether financial constraints impede these types of start-ups by exploiting lottery wins as exogenous wealth shocks. We find that lottery-win magnitude increases winners’ subsequent incorporation, implying that entrepreneurs face financial constraints, but not business registration, implying that financial constraints do not bind as much for the self-employed. Our results, that financial constraints bind for incorporations among men, for serial entrepreneurs, during economic booms, and in neighborhoods without local lenders, are important for understanding the financial impediments to entrepreneurial start-ups.


Financial Consequences of Health Insurance: Evidence from the ACA’s Dependent Coverage Mandate

We study the financial effects of health insurance for young adults using the Affordable Care Act’s dependent coverage mandate as a source of exogenous variation. Using nationally representative, anonymized credit report and publicly available survey data on medical expenditures, we exploit the mandate’s implementation in 2010 and its automatic disenrollment mechanism at age 26. Our estimates show that increasing access to health insurance lowered young adults’ out-of-pocket medical expenditures, debt in third-party collections, and the probability of personal bankruptcy. However, most improvements in financial outcomes are transitory, as they diminish after an individual ages out of the mandate at age 26.

Supersedes Working Paper 18-03.


Owner-Occupancy Fraud and Mortgage Performance

We use a matched credit bureau and mortgage dataset to identify occupancy fraud in residential mortgage originations, that is, borrowers who misrepresented their occupancy status as owner-occupants rather than residential real estate investors. In contrast to previous studies, our dataset allows us to show that—during the housing bubble—such fraud was broad based, appearing in the government-sponsored enterprise market and in loans held on bank portfolios as well, and increases the effective share of investors by 50 percent. We show that a key benefit of investor fraud was obtaining a lower interest rate, particularly for riskier borrowers. Mortgage borrowers who misrepresented their occupancy status performed substantially worse than otherwise similar owner-occupants and declared investors, and constituted one-sixth of the share of loans in default by the end of 2008. We show that these defaults were also significantly more likely to be “strategic,” further highlighting the contribution of fraud in the housing bust.


Population Aging, Credit Market Frictions, and Chinese Economic Growth

We build a unified framework to quantitatively examine population aging and credit market frictions in contributing to Chinese economic growth between 1977 and 2014. We find that demographic changes together with endogenous human capital accumulation account for a large part of the rise in per capita output growth, especially after 2007, as well as some of the rise in savings. Credit policy changes initially alleviate the capital misallocation between private and public firms and lead to significant increases in both savings and output growth. Later, they distort capital allocation. While contributing to further increase in savings, the distortion slows down economic growth. Among factors that we consider, increased life expectancy and financial development in the form of reduced intermediation cost are the most important in driving the dynamics of savings and growth.


The Trade-Comovement Puzzle

Standard international transmission mechanism of productivity shocks predicts a weak endogenous linkage between trade and business cycle synchronization: a problem known as the trade-comovement puzzle. We provide the foundational analysis of the puzzle, pointing to three natural candidate resolutions: i) financial market frictions; ii) Greenwood–Hercowitz–Huffman preferences; and iii) dynamic trade elasticity that is low in the short run but high in the long run. We show the effects of each of these candidate resolutions analytically and evaluate them quantitatively. We find that, while i) and ii) fall short of the data, iii) goes a long way toward resolving the puzzle.

Appendix


Capital Income Taxation with Housing

This paper quantitatively investigates capital income taxation in the general-equilibrium overlapping generations model with household heterogeneity and housing. Housing tax policy is found to affect how capital income should be taxed, due to substitution between housing and nonhousing capital. Given the existing U.S. preferential tax treatment for owner-occupied housing, the optimal capital income tax rate is close to zero (1 percent), contrary to the high optimal capital income tax rate found with overlapping generations models without housing. A low capital income tax rate improves welfare by narrowing a tax wedge between housing and nonhousing capital; the narrowed tax wedge indirectly nullifies the subsidies (taxes) for homeowners (renters) and corrects overinvestment to housing. Naturally, when the preferential tax treatment for owner-occupied housing is eliminated, a high capital income tax rate improves welfare as in the model without housing.


Self-Fulfilling Debt Crises, Revisited

We revisit self-fulfilling rollover crises by exploring the potential uncertainty introduced by a gap in time (however small) between an auction of new debt and the payment of maturing liabilities. It is well known (Cole and Kehoe, 2000) that the lack of commitment at the time of auction to repayment of imminently maturing debt can generate a run on debt, leading to a failed auction and immediate default. We show that the same lack of commitment leads to a rich set of possible self-fulfilling crises, including a government that issues more debt because of the crisis, albeit at depressed prices. Another possible outcome is a “sudden stop” (or forced austerity) in which the government sharply curtails debt issuance. Both outcomes stem from the government’s incentive to eliminate uncertainty about imminent payments at the time of auction by altering the level of debt issuance. In an otherwise standard quantitative version of the model, including such crises increases the default probabilities by a factor of five and the spread volatility by a factor of 25.


Concentration in Mortgage Markets: GSE Exposure and Risk-Taking in Uncertain Times

When home prices threaten to decline, lenders bearing more of a community’s mortgage risk have an incentive to combat this decline with new lending that boosts demand. We test whether this incentive drove the government-sponsored enterprises (GSEs) to guarantee riskier mortgages in early 2007, as the chance of substantial declines grew from small to significant. To identify the effect we relate new risky lending to regional variation in the GSEs’ exposure and the interaction of this variation with home-price elasticity. We focus on the GSEs’ discretion across potential purchases by reference to the credit-score threshold that triggers manual underwriting. We conclude that this incentive helps explain the GSEs’ expansion of risky lending shortly before the financial crisis.


Health Insurance as an Income Stabilizer

We evaluate the effect of health insurance on the incidence of negative income shocks using the tax data and survey responses of nearly 14,000 low-income households. Using a regression discontinuity (RD) design and variation in the cost of nongroup private health insurance under the Affordable Care Act, we find that eligibility for subsidized Marketplace insurance is associated with a 16 percent and 9 percent decline in the rates of unexpected job loss and income loss, respectively. Effects are concentrated among households with past health costs and exist only for “unexpected” forms of earnings variation, suggesting a health-productivity link. Calculations based on our fuzzy RD estimate imply a $256 to $476 per year welfare benefit of health insurance in terms of reduced exposure to job loss.


The Role of Startups for Local Labor Markets

There are substantial differences in startup activity across U.S. local labor markets. We study the causes and consequences of these differences. Startup productivity shocks are found to drive much of these cross-city differences in startup activity. They explain half of the forecast error variance of startup job creation, accounting for 40 percent of population growth and long-run changes in employment. Shocks to barriers to firm entry have economywide effects similar to those of startup productivity shocks but operate largely through the number of startups, rather than their size. We use a novel spatial panel VAR, identifying shocks using shift-share external instruments.

Appendix

**Debt Collection Agencies and the Supply of Consumer Credit**

This paper finds that stricter laws regulating third-party debt collection reduce the number of third-party debt collectors, lower the recovery rates on delinquent credit card loans, and lead to a modest decrease in the openings of new revolving lines of credit. Further, stricter third-party debt collection laws are associated with fewer consumer lawsuits against third-party debt collectors but not with a reduction in the overall number of consumer complaints. Overall, stricter third-party debt collection laws appear to restrict access to new revolving credit but have an ambiguous effect on the nonpecuniary costs that the debt collection process imposes on borrowers.


---

**“Don’t Know What You Got Till It’s Gone” — The Community Reinvestment Act in a Changing Financial Landscape**

This study provides new evidence on the impact of the Community Reinvestment Act (CRA) on mortgage lending by taking advantage of an exogenous policy shock in 2014, which caused significant changes in neighborhoods’ CRA eligibility in the Philadelphia market. The loss of CRA coverage leads to an over 10 percent decrease in purchase originations by CRA-regulated lenders. While nondepository institutions replace approximately half, but not all, of the decreased lending, their increased market share was accompanied by a greater involvement in riskier and more costly FHA lending. This study demonstrates how different lenders respond to the incentive of CRA credit.


---

**Supply Shock Versus Demand Shock: The Local Effects of New Housing in Low-Income Areas**

We study the local effects of new market-rate housing in low-income areas using microdata on large apartment buildings, rents, and migration. New buildings decrease nearby rents by 5 to 7 percent relative to locations slightly farther away or developed later, and they increase in-migration from low-income areas. Results are driven by a large supply effect—we show that new buildings absorb many high-income households—that overwhelms any offsetting endogenous amenity effect. The latter may be small because most new buildings go into already-changing areas. Contrary to common concerns, new buildings slow local rent increases rather than initiate or accelerate them.


---

**Debt Collection Agencies and the Supply of Consumer Credit**

This paper finds that stricter laws regulating third-party debt collection reduce the number of third-party debt collectors, lower the recovery rates on delinquent credit card loans, and lead to a modest decrease in the openings of new revolving lines of credit. Further, stricter third-party debt collection laws are associated with fewer consumer lawsuits against third-party debt collectors but not with a reduction in the overall number of consumer complaints. Overall, stricter third-party debt collection laws appear to restrict access to new revolving credit but have an ambiguous effect on the nonpecuniary costs that the debt collection process imposes on borrowers.


Partisan Conflict Index

The Philadelphia Fed collects, analyzes, and shares useful data about the Third District and beyond. Here's one example.

The Philadelphia Fed’s Partisan Conflict Index (PCI) tracks the degree of political disagreement among U.S. politicians at the federal level by measuring the frequency of newspaper articles reporting disagreement in a given month. Higher index values indicate greater conflict among political parties, Congress, and the President. The horizontal axis represents the years 1981 to 2020. The vertical axis measures partisan conflict, with the solid line representing the average level of conflict in 1990. Stony Brook University Professor Marina Azzimonti, who created the PCI while working for the Philadelphia Fed in 2014, writes in her 2018 *Journal of Monetary Economics* article that partisan conflict is associated with lower capital investment rates at the firm level, even when she controls for economic policy uncertainty and macroeconomic conditions. She estimates that about 27 percent of the decline in corporate investment between 2007 and 2009 can be attributed to a rise in partisan conflict. The Philadelphia Fed updates this index monthly to allow researchers to observe how the indicator moves in relation to the salient economic policy issues of the day.

Note: Average of 1990 = 100. Shaded areas indicate NBER recessions.

Source: Federal Reserve Bank of Philadelphia

1 https://doi.org/10.1016/j.jmoneco.2017.10.007

Learn More
Online: philadelphiafed.org/research-and-data/real-time-center/partisan-conflict-index
E-mail: PHIL.PCI@phil.frb.org

Connect with Us

Read previous articles
Economic Insights
www.philadelphiafed.org/research-and-data/publications/economic-insights

Business Review
www.philadelphiafed.org/research-and-data/publications/business-review

Sign up for e-mail notifications
www.philadelphiafed.org/notifications
You can find Economic Insights via the Research Publications part of our website.