FEDERAL RESERVE BANK OF ST. LOUIS

REVIEW

SECOND QUARTER 2020
VOLUME 102 | NUMBER 2

Strategic Review and Beyond: Rethinking Monetary Policy and Independence

John H. Cochrane

Is the Phillips Curve Still Alive?

Brian Reinbold and Yi Wen

Schools and Stimulus

Bill Dupor and M. Saif Mehkari

Taking Stock of the Evidence on Microfinancial Interventions

Francisco J. Buera, Joseph P. Kaboski, and Yongseok Shin

On the Aggregate Implications of Removing Barriers to Formality

Catalina Granda and Franz Hamann

REVIEW

Volume 102 • Number 2

President and CEO

James Bullard

Director of Research

Christopher J. Waller

Chief of Staff

Cletus C. Coughlin

Deputy Directors of Research

B. Ravikumar

David C. Wheelock

Review Editor-in-Chief

Carlos Garriga

Research Economists

David Andolfatto

Subhayu Bandyopadhyay

Serdar Birinci

YiLi Chien

Riccardo DiCecio

William Dupor

Maximiliano Dvorkin

Miguel Faria-e-Castro

Victoria Gregory

Sungki Hong

Kevin L. Kliesen

Julian Kozlowski

Fernando Leibovici

Oksana Leukhina

Fernando M. Martin

Michael W. McCracken

Amanda M. Michaud

Alexander Monge-Naranjo

Christopher J. Neely

Michael T. Owyang

Paulina Restrepo-Echavarría

Hannah Rubinton

Juan M. Sánchez

Ana Maria Santacreu

Guillaume Vandenbroucke

Yi Wen

Christian M. Zimmermann

Managing Editor

George E. Fortier

Editors

Jennifer M. Ives

Lydia H. Johnson

Designer

Donna M. Stiller

FEDERAL
RESERVE
BANK *of*
ST. LOUIS

CENTRAL TO AMERICA'S ECONOMY®

99

Strategic Review and Beyond: Rethinking Monetary Policy and Independence

John H. Cochrane

121

Is the Phillips Curve Still Alive?

Brian Reinbold and Yi Wen

145

Schools and Stimulus

Bill Dupor and M. Saif Mehkari

173

Taking Stock of the Evidence on Microfinancial Interventions

Francisco J. Buera, Joseph P. Kaboski, and Yongseok Shin

203

On the Aggregate Implications of Removing Barriers to Formality

Catalina Granda and Franz Hamann

Review

Review is published four times per year by the Research Division of the Federal Reserve Bank of St. Louis. Complimentary print subscriptions are available to U.S. addresses only. Full online access is available to all, free of charge.

Online Access to Current and Past Issues

The current issue and past issues dating back to 1967 may be accessed through our Research Division website:

<http://research.stlouisfed.org/publications/review>. All nonproprietary and nonconfidential data and programs for the articles written by Federal Reserve Bank of St. Louis staff and published in *Review* also are available to our readers on this website.

Review articles published before 1967 may be accessed through our digital archive, FRASER: <http://fraser.stlouisfed.org/publication/?pid=820>.

Review is indexed in Fed in Print, the catalog of Federal Reserve publications (<http://www.fedinprint.org/>), and in IDEAS/RePEc, the free online bibliography hosted by the Research Division (<http://ideas.repec.org/>).

Authorship and Disclaimer

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

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

Subscriptions and Alerts

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

Our monthly email newsletter keeps you informed when new issues of *Review*, *Economic Synopses*, *Regional Economist*, and other publications become available; it also alerts you to new or enhanced data and information services provided by the St. Louis Fed. Subscribe to the newsletter here: <http://research.stlouisfed.org/newsletter-subscribe.html>.

Copyright and Permissions

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

STLS.Research.Publications@stls.frb.org.

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

Economic Data

General economic data can be obtained through FRED®, our free database with over 670,000 national, international, and regional data series, including data for our own Eighth Federal Reserve District. You may access FRED through our website: <https://fred.stlouisfed.org>.

© 2020, Federal Reserve Bank of St. Louis.

ISSN 0014-9187

Strategic Review and Beyond: Rethinking Monetary Policy and Independence

John H. Cochrane

This article is based on the Homer Jones Memorial Lecture delivered at the Federal Reserve Bank of St. Louis, March 4, 2020.

I survey monetary policy strategy, regulation, and central banks' mandates and independence. I do not think strongly negative interest rates, vastly expanded quantitative easing, or extensive forward guidance can or should stimulate in the next recession. I advocate a price-level target and that the Fed fix the spread between indexed and nominal debt. I argue for a large balance sheet of interest-paying reserves, achieved via a flat supply curve, though I argue that the Treasury should issue reserve-like debt as well to take up much of that role. Central banks should avoid the temptation toward ever-expanding roles including "macroprudential" policy, discretionary credit cycle management, asset price targeting, and using their regulatory power to advance social and political goals such as climate change and inequality. Only limited scope of action to areas of agreed technocratic competence will salvage central banks' and international institutions' useful independence. (JEL E52, E58, E61, G28)

Federal Reserve Bank of St. Louis *Review*, Second Quarter 2020, 102(2), pp. 99-119.
<https://doi.org/10.20955/r.102.99-119>

This is an exciting moment to think about central banks and monetary policy. The Fed and many other central banks are undergoing a formal strategy review. More broadly, the community that studies monetary policy has entered a period of fundamental rethinking. I'm grateful for this opportunity to put together my thoughts and to reflect together on the answers.

I'll start with the easy question, monetary policy. Then I'll move on to regulation, and finally we'll think about central bank mandates and independence. I will end up contradicting just about every tenet of current consensus opinion, well articulated, for example, by Bernanke (2020). Well, examine the logic and evidence and decide for yourself.

John H. Cochrane is the Rose Marie and Jack Anderson Senior Fellow, Hoover Institution, Stanford University with affiliations at the National Bureau of Economic Research, Cato Institute, and University of Chicago. He thanks Richard Anderson, David Andolfatto, Bill Dupor, and Max Gillman for helpful comments.

© 2020, Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

A theme underlies my thoughts: We have learned a lot in the 12 years since the financial crisis. Some of these lessons are slowly percolating into policy. I mostly take these lessons to their logical conclusion.

In 2007, the question was wide open: What happens if interest rates hit zero and the Fed cannot lower interest rates? What happens if the Fed pays interest on reserves—deposits that banks hold at the Fed—and increased reserves by 3000 percent, from \$10 billion to \$3 trillion? Every mainstream school of thought—monetarist, old ILSM Keynesian, new-Keynesian—predicted massive deflation or inflation or great volatility in these events. Yet nothing of the sort happened. The long quiet slow recovery decisively disproved these ideas.¹ Yet the vestiges of these ideas still guide central banks' policy thinking and review. How we think going forward must incorporate the lessons of this clear experiment.

Twelve years of reflection about regulation in the wake of the crisis have taught us that we can stop financial panics forever, without turning the financial system into a sclerotic regulated utility, featuring lots of leverage, tight regulation, and inevitable rounds of crisis and bailout. I'll show you how. It's time to complete that agenda.

NEGATIVE RATES, QUANTITATIVE EASING, FORWARD GUIDANCE

The narrow question facing the Fed is simple. A recession will come. Interest rates will hit zero again. Now what? (As I write, the Covid-19 virus is spreading, and the economy is screeching to a halt. The event may have already happened by the time you read these words.)

Should the Fed prepare to set interest rates substantially below zero? You pay the bank to hold your money, and they pay you to take out a mortgage? Should the Fed prepare for even more massive Treasury bond purchases, increasing the supply of reserves by additional trillions? Should the Fed add purchases of mortgages, stocks, and corporate bonds, as the Bank of Japan and European Central Bank have done? Should it prepare for aggressive direct lending, or directly try to prop up asset prices? Should it prepare an arsenal of speeches, or announce new, and binding, inflation targets and interest rate rules, hoping to stimulate today by promises of or commitments to future actions?

In my opinion, no.

Substantially negative interest rates would require big structural changes to our banking and financial systems. Someone has to reprogram every computer to accept a minus sign! We have to get rid of cash and solve the consequent delicate balance of privacy versus law and tax enforcement in a fully cashless world. We have to change lots of accounting and tax payment conventions that allow people to pre-pay without penalty. It's a big and disruptive job.

Evidence and logic on asset purchases—quantitative easing (QE)—tell us that they had no prolonged effect. Long-term rates have been on a downward trend since the 1980s. In a graph,² you cannot see any sign of QE.

Remember what QE is. If the Fed just gave people money, that might make them spend. But the Fed exchanges money for bonds. And when money and bonds pay the same interest, they are perfect substitutes. More interest-bearing reserves and fewer bonds is only a slight readjustment of the maturity structure of government debt. If the Fed takes a \$20 bill and

gives you two \$5s and a \$10 in return, how does that make you spend more? It doesn't. If the Fed takes your red M&Ms and gives you green M&Ms in return, how does that help your diet? It doesn't.

Moreover, the Treasury was selling faster than the Fed was buying, to fund massive deficits.³ Overall people held more, not less, Treasury debt and more long-maturity Treasury debt. And we have seen vast changes in the maturity structure of the debt over history with no visible effect on interest rates.⁴ The notion that people have a fixed desire for specific maturities of government debt fits neither experience nor simple economic logic.

The economics literature concludes that QE was at best a signal—if the Fed is going to do something this wild, the Fed must think the problem is really bad and interest rates are likely to be low for a long time.⁵

Now, by the same token, if QE doesn't really change anything, it isn't doing any harm. But don't count on massive Fed purchases of Treasury securities to stop the next recession.

Buying mortgages, bonds, and stocks might have some effect on those prices. (At least for a while. Demand and supply curves for assets are flatter than you think, especially after a month or two.) But this kind of direct intervention in asset markets has enormous risks. Does the Fed really want to get involved with propping up the stock market? Or deciding which business should get cheap credit and which should not?

My judgment reflects a deeper issue that we should state and debate. In my view, negative interest rates and asset purchases are not hugely stimulative. So there are no great benefits to balance against the costs.

If you think a 2 percent negative overnight rate and another trillion dollars of Treasury purchases is all it would have taken for the agonizing 8-year recovery from the Great Recession to happen in 1 or 2 years, then the case is much stronger. But if these actions stimulate, then it was a crime for the Fed not to do far more.

I don't think that's true, and I don't think the Fed does either. If that were true the Fed would have done more. If that were true, Japan and Europe, who did go negative and bought more assets than we did, would have grown like gangbusters. They did not. But face the logic: Either the Fed had a powerful tool but failed to use it, or it had a mostly symbolic tool that we should not count on to do much next time.⁶

Why did the Fed choose a finite *quantity* of QE, leaving us in an eternal debate about whether it did or did not lower interest rates? Why did the Fed not just say, "We want the 10-year Treasury rate to be 2 percent. (Or 1 percent. Or 0 percent). We will buy bonds until that price is achieved. Nay, we will buy any quantity of bonds at a 1 percent yield and set the darn price." I think the answer is obvious: The Fed was worried about just how many bonds it would have to buy, or that it could not change market rates even by buying all the bonds. Then the powerlessness of QE would have been revealed. Better to keep the smoke and mirrors going.

"Forward guidance" is the idea that even though the interest rate is stuck at zero today, the Fed can stimulate the economy by giving speeches about what it will do in the future. Once the recession ends, the Fed will keep interest rates low for a substantial period of time, or the Fed will allow inflation above target for a few years. This change in expectations is supposed to boost the economy today.

The difficulties of this argument are apparent.

First, pre-commitment. What Fed chair will ever go to Congress and say, “Yes, inflation is growing and the economy is healthy again. By all normal indications now is the time to raise interest rates. But we promised 5 years ago that we would keep rates low and allow inflation in order to stimulate back then, so now I have to keep that promise even though it would be better for the economy today to raise rates.” Ex ante, will anyone really expect this sort of behavior?

The Fed is considering changes to its inflation target that would institutionalize and attempt to pre-commit to this sort of response. But the Fed has never really pre-committed ex ante to do something it does not want to do ex post, or adopted any formal and binding rules.

Second, the case for forward guidance is based on economic models in which promises about actions further in the future have *greater* effects today. In any sensible model, expectations about events further in the future have *lesser* effects today, which dramatically lowers the power of forward guidance.⁷

Third, I think we all live in a bit of a bubble. We all read Fed speeches closely, as do bond traders. But the average businessperson and the average worker don’t really even know what the Fed is, let alone parse FOMC statements for hints about interest rates and inflation targets 5 years from now. The immense dispersion in survey expectations gives you some hint about just how different from ours most people’s expectations are.

This logic points to a deeper lesson. The Fed really has limited power to stimulate our way out of recessions.

A prolonged recession reflects something wrong in the economy. Effective policy should fix what’s wrong. Cappuccinos stimulate, but if you’re having a heart attack, a cappuccino is a terrible substitute for going to the emergency room. Lots of things were wrong in the U.S. economy in the Great Recession. Much of the literature sees financial frictions, limitations on credit supply. I see that recessions are a feature of *risk premiums*, willingness to take risk, not of overnight interest rates. I see a lot of regulation, social program, and tax disincentives that held back the supply side of the economy.

Really, of all the things holding back growth, were the central underlying causes an overnight interest rate of a quarter of a percent rather than negative 2 percent, that the Treasury sold too much 10-year debt and not enough short-term debt, and that central bankers had not made enough hearty speeches making promises about what they might do in 5 or 10 years time when they might normally start to raise rates?

The Fed’s main job is not to screw up—not to cause a recession, not to cause or worsen a financial panic. That the Fed did not repeat the mistakes of 1929-1933, and many since then, is praiseworthy.

But money is oil in the economic car, not gas. Once the car is full of oil, adding more oil does not help. Monetary and financial frictions are inequality constraints. Once an inequality constraint is slack, pouring more on the slack side does no good.

We face a similar situation as I write. The long-awaited slowdown seems to be happening as the economy slows down due to a spreading disease. The Fed has cut the overnight rate. But is demand side stimulus of any use? Should people look at low rates, borrow some money,

and go on a cruise? Will they? No. Neither low overnight rates, nor Treasury purchases, nor promises about interest rates after the virus has passed will raise output now. Supply, public health, and avoiding a cascade of failures in businesses that must stay shut down for a while so the economy can restart easily are the challenges for policy.

A counterargument I hear from many people at the Fed amounts to this: Sure, we know negative rates, QE, and forward guidance don't do much. But we have to be seen to do something in hard times. We have to keep up the appearance that we're in charge. It's a symbolic, political PR move.

I think there is a good deal of truth to this analysis of the Fed's actions, but questionable courage and wisdom. It would indeed have been politically difficult for the Fed to state the facts—this recession is going on far too long but there is nothing we can do about it. But I also think that being honest about central banks' limited ability to fix everything would be good for central banks, and for our politics and economic policy. Stimulus has limits. "Structural adjustment," fixing broken microeconomics, needs greater emphasis. And sooner or later they will pull aside the curtain. Having pretended to be so powerful will come back to haunt central banks.

INFLATION TARGET

What should the inflation target be? Right now the answer is 2 percent, which really grew out of habit rather than any particular analysis. More importantly, it is an eat-your-mistakes target: If inflation really comes out at 1 percent, the Fed keeps the target going forward at 2 percent, not 3 percent to correct the past shortfall.

Lots of alternative proposals abound, including nominal-GDP targets, average-inflation targets, price-level targets, and complex asymmetric targets (forgive inflation, but make up deflation).

I favor a symmetric price-level target. The consumer price index shall be 250 forever. If it falls, the Fed (and Treasury) are committed to getting it back up; if it rises, they are committed to bringing it back down.⁸

We do not shorten the yard by 2 percent each year to try to goose the economy. Should not the standard of value be as constant over time as every other measurement? A constant long-run price level would make the financial system more efficient in lots of little ways. For example, we pay capital gains taxes on inflation-induced gains. Sure, says the perfectionist, fix the stupid tax code. But it's easier to fix the inflation. It would lower long-term bond yields, by shielding from inflation risk, lowering both the government's and private borrowers' interest costs. Sure, says the perfectionist, fix TIPS (Treasury inflation-protected securities) so they are more popular and index contracts better. But nominal contracts are surprisingly popular, hence, evidently efficient.

A constant price-level target is simple and visible. Quick: If the Fed is following a one-sided 2 percent price-level or nominal-GDP target that makes up for past undershoots but not past overshoots, just what is inflation supposed to be this year?

Others like a price-level or nominal-GDP target because those targets commit the Fed to promise a period of higher future inflation if there has been a period of below-target inflation,

as in forward guidance. As I explained above, I'm dubious of the idea that expected inflation in the distant future stimulates much today. But one can like the same proposal for different reasons.

Milton Friedman wrote⁹ that slight deflation and a zero nominal interest rate are optimal. Then money and bonds become perfect substitutes, and people do not economize on cash. At a zero nominal rate we know the economic car is full of oil. I favor the price-level target because it's simpler, clearer, and close enough, especially in these days of near-zero or negative real rates. Contracting costs are arguably minimized at zero inflation not a zero nominal rate. Interest-paying electronic money obeys the optimal quantity at any interest rate, and the small interest costs of physical cash are minor and perhaps a reasonable tax on all of cash's illegal uses.

Some people dislike deflation or zero inflation, on the notion that wages are sticky downward. I think the sticky wage story is overblown. Remember, even with no price inflation, each individual's wage can rise as they age and gain experience and as individual and aggregate productivity increases. Wages are certainly less sticky when people change jobs, and many people experience large income fluctuations due to hours, contractor status, piece rate, and bonus compensation. The increased number of actual individual wage cuts demanded by 2 percent versus 0 percent inflation is tiny.

Moreover, if sticky wages are the single salient economic problem that causes recessions, depressions, and unemployment, why is so much law, regulation, and policy devoted to making wages *stickier*, especially downward, and making labor markets less competitive, which adds to wage rigidity. Why are macroeconomists, and the Fed, silent on this great economic tragedy?

There is an apparent incoherence if not hypocrisy here, an unwillingness to consider the logical conclusions of one's assumptions. Macroeconomists habitually adduce a friction that causes enormous economic pain, but then they advise the Fed to cleverly exploit the friction to manage the economy. They never even *mention* "Let's get rid of that friction" or think about whether the friction is due to policy or directly remediable by policy. Now in many cases, frictions such as sticky wages and prices, enormous markups, and credit constraints are just simple parables, used to fit aggregate dynamics and generate views about desirable monetary policy, not serious descriptions of the microeconomy. OK, such simplified parables are important parts of good macroeconomic modeling. But if we don't trust the modeling assumption enough to say loudly that policy should make wages less sticky, then we are on shaky ground to object to permanent zero inflation on the grounds that wages are too sticky.¹⁰ (There is also an increasing amount of microeconomic investigation of issues such as sticky prices and wages, which finds nothing so simple as the uniform stickiness of aggregated models.)

Others like a higher inflation target to prompt higher nominal interest rates and give the Fed "headroom" to stimulate more in the next recession. That's like wearing shoes that are two sizes too small so it feels better when you take them off at the end of the day. The theory that the path of interest rates matters—that going up to 3 percent and then cutting to 0 percent is more stimulative than arriving at the same 0 percent from 1 percent—may sound intuitive, but I don't know of any model with such path-dependence. The idea that getting inflation up to 2 percent gives you a lower real rate when you cut nominal rates to 0 percent presumes inflation expectations are mechanical and backward looking.

What about Japan? That's the inevitable question when one favors low inflation and low nominal interest rates. Japan has had zero interest rates and steady deflation for three decades. But in most of that time, Japan's unemployment rate has been around 3 percent. That does not sound like deficient demand. Japan is not growing badly per demographically adjusted capita. It has not completed catch-up to U.S. productivity and per-capita GDP levels because of greater microeconomic distortions.¹¹ But after three decades, money is surely neutral and prices and wages adjust. Japan has a Friedman-optimal monetary policy—zero interest rates, slow deflation—and people use all the money they need.

THE LOWER BOUND

The question remains: What *should* the Fed do at the zero bound, or the effective lower bound of negative 1 percent or so? There are two parts to this question.

First, how will the Fed stop a deflation spiral next time interest rates hit zero? To answer that question, we need to know *why* the dreaded inflation spiral did not break out last time in the U.S., Europe, or Japan and just why the models that predicted spirals failed so catastrophically.

Here's my answer to that question: All the deflation-spiral models leave out a crucial detail. If deflation were to cut the price level in half, then nominal income would fall by half, and tax revenues would fall by half. Congress would have to sharply raise taxes or cut spending to pay off the debt.

Now ask yourself, what is the chance that in a sharp deflation, with a huge recession, that the U.S. Congress passes draconian tax increases or spending cuts, all to pay for an unexpected, undesired, and surely, it will be argued, undeserved windfall profit to fat-cat Wall Street bakers, billionaires, and the Chinese who hold our Treasury bonds? It's not happening. And that is exactly why deflation can't happen.¹²

This is the mechanism that needs strengthening in the event of zero interest rates and emerging deflation: a clearer set of monetary *and* fiscal arrangements. There is some hint of it in calls for fiscal stimulus at the zero bound, but if people expect borrowed money to be paid back, then that stimulus is ineffective. Running an unbacked fiscal expansion, persuading people that a fraction of the debt will be inflated away—but not all of it—is very tricky.¹³

In my view, the answer is an explicit price-level (or inflation, for this purpose) target that is as binding on Congress as on the Fed. This fiscal policy rule works much like John Taylor's monetary policy rule and links primary deficits to the price level.¹⁴ But that step represents a fundamental restructuring of fiscal as well as monetary policy, so let us return to what the Fed can and should do on its own.

The second question is what should the Fed do to exit the zero bound. Here the lesson of the recent past is crucial: Inflation is, in fact, stable when interest rates are fixed at or near zero. We do not see deflation or inflation spirals. This prediction is also a robust feature of models with forward-looking expectations. It follows that if the Fed were to announce ahead of time that it will slowly start raising interest rates, and stick to that plan, inflation will rise. This prediction is consistent with the U.S.'s slow and preannounced liftoff relative to Europe and Japan. This “neo-Fisherian” prediction can coexist with the more common experience and VAR evidence that a *surprise and transitory* interest rate rise lowers inflation.¹⁵ Though con-

sistent with theory and evidence, this prediction is still hard to swallow as it contradicts so many stories told and retold. I agree that policy should be based on well-tested models and not the latest clever idea in one's most recent paper. So I leave that observation as an indication of just how little we know confidently, scientifically, and in a consensus about how interest rates actually do affect inflation, but also as a serious possibility we should consider.

But I do not think it vital that the Fed *should* take this step and deliberately raise inflation. It is also true in forward-looking and, hence, stable models that, if the economy needs a negative real interest rate at a zero nominal rate, inflation will eventually take care of that on its own, as Japan's deflation took care of a needed positive real interest rate all on its own.

INTEREST RATES

How should the Fed set interest rates, or move inflation back to target? Should the Fed keep (or return to, John Taylor might argue) the Taylor rule, raising rates when the GDP gap and inflation rise? If we want a price-level target, should the Fed just raise rates when the price level is above target and vice versa? Should it follow a first-difference rule, raising or lowering the rate in response to events rather than a rule describing the level of rates? Should rates move gradually or instantly?¹⁶

Let me advance an out-of-the box alternative. On first principles, the idea of the Fed setting the interest rate is a poor policy design. To set an interest rate, on top of an inflation target, the Fed must somehow divine the appropriate real interest rate. (Interest rate = real interest rate plus expected inflation.) But nobody sitting in a desk in Washington could possibly figure out the right price of a 2×4 , let alone the right real interest rate.

Here is an alternative. Since the Fed wants to target inflation, why not target inflation? Target the spread between indexed and non-indexed debt, which represents expected inflation. If the Fed wants, say, 2 percent inflation, then make the following offer to markets: Bring in a 1-year indexed Treasury bill and get back 1.02 1-year non-indexed Treasury bills and vice versa.¹⁷

Targeting this *spread*, the Fed can directly target the inflation rate or price level and get out of the business of divining the true natural rate of interest. Interest rates themselves will vary, potentially a lot, reflecting market forces, as stock, bond, and foreign exchange rates vary.

The natural worry is stability: We know that in the long run higher inflation must come with a higher indexed vs. nonindexed spread. But which is the chicken and which is the egg? If the Fed targets the spread, does inflation converge to the target or does inflation spiral away? This is exactly the same question as whether inflation is stable when rates are stuck at zero or at a peg. The zero bound era answered that question in favor of stability, as do all our forward-looking models. This is also a policy the Fed can experiment with gingerly in addition to normal interest rate targeting.

THE GOLD STANDARD

Many commenters write longingly of a return to the gold standard. The gold standard simply will not work for a modern economy. The gold standard period had sharp inflations

and deflations around a long-run steady price level, as the value of gold varied relative to other goods and services, culminating in the deflation of the Great Depression, which led to the abandonment of the gold standard. That problem would be much worse today. In the 19th century, gold coins were still widely held and used for transactions. That fact linked the value of gold to the value of everything else, at least eventually. But the relative price of gold can wander away from other prices today. Targeting the price of gold would have no more stabilizing effect on overall inflation than targeting the price of tungsten.

But gold standard advocates' hearts are in the right place. My proposals embody much of the spirit and logic of the gold standard, applied to a modern economy. The gold standard promises a long-run steady price *level*. So does my price-level target. The gold standard is at heart a fiscal rule. If the dollar devalues and people demand gold, the government must raise taxes or cut spending enough to get the gold. My expected CPI target and fiscal rule embody those ideas. The promise to trade indexed for non-indexed debt acts like the promise to trade dollars for gold. My CPI target is better, as it eliminates inflation and deflation (as in the 1930s) when both gold and the dollar gain or lose value relative to goods and services.

BALANCE SHEET

How big should the Fed's balance sheet be, and how should the Fed run the balance sheet?

The Fed is essentially a huge money market fund invested in U.S. Treasury and agency securities. The "size of the balance sheet" means the size of reserves, banks' accounts at the Fed, plus currency, and the corresponding size of the assets that the Fed holds to back those accounts.

The Fed should have a large balance sheet, with reserves that pay full interest. Don't deliberately drive the car with less than full oil. Don't try to speed up or slow the car down by adding or draining oil. Full interest on reserves, and giving banks as much reserves as they want, allows the monetary system all the "liquidity" it needs.

For decades, this regime was widely recognized as providing the "optimal quantity of money," along with efficiency and stability of the financial system. The debate was whether adopting it might unleash unstable inflation or multiple-equilibrium volatility. Once more, the past 12 years—the past 30 in Japan—settle that debate. We can be awash in interest-paying reserves with no inflation. Let's do it.

The Fed is still hesitant, limiting the quantity of reserves as well as paying full price, and limiting who can get reserves. Perhaps some ghost of the quantity theory keeps the Fed up at night. The result is last summer's gyrations in money markets and the unseemly apparent subsidy to big banks who can get higher interest rates than anyone else.

No, if you want to target interest rates, target interest rates. Offer a flat supply. Any qualified financial institution can bring in some Treasury securities and earn 1.99 percent on reserves. Or they can borrow against Treasury collateral at 2.01 percent. Done. The Fed was founded in 1914 to furnish an elastic currency. Provide 21st century electronic money elastically.

This step is made harder by the fact that the Fed is a bank and legally designed to serve banks. It would be much easier for this and many reasons if the Treasury offered the same security: The Treasury should offer fixed-value, floating-rate, electronically transferable

(instantaneous, date 0 settlement) Treasury debt.¹⁸ If the Treasury offered the same security directly, then the Fed could keep a much smaller balance sheet, and we wouldn't have to get in to the thicket of whether the Fed should allow money market funds, narrow banks, fintech innovators, government agencies, large corporations, and you and me to have reserves accounts.

It's as if the Treasury had only minted \$10,000 coins, so the Fed had to buy those and issue nickels, dimes, and quarters. Well, let the Treasury issue the kind of debt people want in the first place.

Also, if it wants a small balance sheet, the Fed should fix the resolution liquidity and other regulations that are inflating demand for reserves, and the Fed should provide date 0 reserves for Treasury securities via widely available stigma-free purchase or repo. Then Treasury securities would be exactly as liquid as reserves.

BANK REGULATION¹⁹

Interest rates and monetary policy are actually a small fraction of what the Fed does. The Fed is primarily the gargantuan financial regulator. How should the Fed approach this task?

We have learned a lot in 12 years here, too. The central problem of the financial crisis was not how banks invested their money. Even portfolios of subprime mortgages are a lot less risky than the assets of run-of-the-mill stock mutual funds, whose losses cause little pain to the system overall, or the income streams of non-financial corporations.

The central problem was that financial institutions *got* their money far too much by short-term borrowing and far too little by issuing equity or retaining earnings to build up equity capital. When every day you have to find new lenders to pay off the previous day's borrowing, you're open to a run, which is what we had, pure and simple. A bank that funds its risky investments by equity simply cannot fail or suffer a run. Bankruptcy happens if you can't pay debt. No debt, no bankruptcy.

We have learned that more capital is not *socially* costly. Especially when we count getting rid of private sector financial crises forever, more capital and less short-term debt is a great benefit to society. There will still be plenty of money to fund mortgages and business loans, and it won't cost more. More capital is *privately* costly to banks, as they lose the value of debt subsidies and bailout guarantees. That's why they fight so hard to lower capital requirements. But we who pay the debt subsidies should value more capital.

We can now live in a financial system that never has a private-sector crisis again. Banks get money for risky investments by issuing equity and long-term debt. Many investments, like mortgages, are sold into pools held in mutual funds at fluctuating values. Institutions that want to offer deposits can do it, backed 100 percent by Treasury securities or reserves. (We're still open to sovereign default, which would be catastrophic, but that's another lecture entirely.)

Financial regulation could be very simple. My favorite version would be a simple tax on short-term debt, along with a sliding scale of regulations based entirely on liabilities—the ratio of the face value of debt to the market value of equities. We would no longer need clairvoyant regulators to see “bubbles” building and prick them, or courageous bailer-outers to save the system in crisis. There would be fewer statues to heroic Fed chairs, perhaps, but a much better financial system for the rest of us.

Modern technology allows this system. We could make transactions by swiping a card that sells an S&P500 index. *Liquidity* no longer requires *fixed-value* run-prone securities. And the vast expansion of government debt means we can fully back \$20 trillion of deposit accounts. We no longer need banks to “create money” from mortgages. Though this idea stems from the “Chicago plan” of the 1930s, one can argue it was infeasible then. No longer.

It took a long time to figure this out. But I think now the economic point is pretty much acknowledged by people who study these issues. They say only “It’s not politically feasible.” Big banks, who profit from the many subsidies to debt, and enjoy the protection from competition that extensive regulation gives them, don’t like it. Central banks don’t want to hear it either, so nobody will invite you back if you offer this impolite simple truth. They would rather continue the path of expanding their regulatory power. And with the financial crisis now in the rearview mirror and bank regulation a sure topic to put anyone to sleep except those with money or power at stake, nobody is really paying attention. What good does it do a Fed governor to stir up a hornet’s nest of trouble with big banks and his or her regulatory staff?

But we needn’t give up. We don’t have to focus at all on today’s big banks. *Just allow an on-ramp*. Answer the question: What does a new financial firm need to do, so that it needs no regulation? The answer is, finance it by equity and long-term debt and no short-term run-prone debt. In the tens of thousands of pages of financial regulation, that on-ramp is missing. Allow the on-ramp, allow equity-financed lightly regulated innovated competitors, add a little tax on short-term debt, and the problem will solve itself.

Technology will eventually also force change. Narrow banks, fully backed private currencies like Libra, and fintech lenders will grow like weeds in the cracks, or rather like Ubers in the taxi system. But the process would be a lot smoother and faster if the Fed would allow that on-ramp and not just continue protecting big banks from competition and blocking narrow banks.²⁰

MANDATES

What is a central bank’s job, scope, or mandate? The Fed was founded as lender of last resort and to provide an elastic currency. That mandate expanded to “price stability,” which somehow evolved to a 2 percent inflation target²¹ and “maximum employment,” along with a lot of bank regulation. “Financial stability” is a new mandate, involving detailed direction of the financial system.

There is a strong push for “macroprudential” policy, that central banks should direct and manage the “credit cycle,” deploying a wide variety of discretionary tools to directly regulate lending in the booms, prick asset price “bubbles,” and then by regulation open the spigots and prop up prices in recession.²²

In international affairs, the IMF used to urge countries to keep trade and capital open. In a crisis, the IMF required a commitment to micro deregulation, cutting subsidies, and getting the fiscal house in order before offering a bridge loan. This is like borrowing from your grumpy uncle: Get a job, stop drinking and gambling, here is some money to tide you over, but I’ll be watching.

In 2012, the IMF moved to an “institutional view,” advocating that central banks “manage” capital flows and exchange rates, along with extensive macroprudential direction.²³ The IMF’s new “integrated policy framework” promises an even more ambitious “integrated” approach to “monetary policy, macroprudential policy, exchange rate interventions, and capital flow measures,” tailored to disparate “country circumstances.”²⁴

It is all very tempting. Central bankers like to feel important. Interest rates are either stuck at zero or don’t seem to do a heck of a lot. Well, take on broad new powers to run things and do good as you see it.

But like discretionary monetary policy, central banks have never been able to time credit and asset price cycles, or micromanage dozens of interacting policy levers to offset poorly understood (and country-specific) “frictions” and “imperfections.” How do you tell a boom from a bubble in real time? How and why will central banks get it right this time after so many abject failures—2007 being the most recent and screaming example? How will they avoid repeating the endless problems of managed exchange rates and extensive capital controls that finally blew up in the 1970s? Central bankers are only human, just like the rest of us—and just as prey to the fallacy that we’re the smart ones and everyone else is behavioral. In the crisis, as monetary policy committees were begging banks to lend, regulators were telling banks to cut back lest the crisis get worse. Through the 12th year of the subsequent expansion, the U.S. has been if anything loosening capital and credit standards, despite great increases in credit. So much for macroprudence.

Rather than try to stop anyone from ever borrowing too much or losing money ex post, we should make the financial system robust so that people can make and lose money without burning down the house. That’s the equity-financed banking approach.

The current trend is even more ambitious. Now, the International Monetary Fund, the Bank for International Settlements,²⁵ and the Financial Stability Board are advocating and the Bank of England²⁶ is starting to implement climate and inequality²⁷ policies. Central banks should demand extensive disclosures of “climate risk” and contributions to “sustainable investing.” Those lending to, say, fracking companies will have an army of regulators descend on them. The European Central Bank is buying “green” bonds. Fed Chair Jay Powell has so far been a courageous resister to the climate side of this movement,²⁸ but we’ll see how long that lone voice of resistance can hold out.

The IMF is now advocating, along with climate, a full range of policies including increased “social spending,” progressive taxation, income redistribution, and social-justice policies far beyond anything traditionally monetary or financial.

Requirements for “sustainable accounting” (see Finley, 2020), “disclosure” of environmental, social, and corporate governance (ESG) blessings, “stakeholder capitalism,” divestiture, and de-financing more disfavored industries are already being advanced.

There is a reasonable risk that climate change may be, in 50 or 100 years, a big economic problem. But the risk that unforeseen changes—risk—in *climate* threatens the financial system with another run is essentially zero on the 5-year-or-so timeline of honest risk assessment. (Except maybe risks induced by the same regulators!) Repeating the contrary assertion over

and over in speeches does not make it so. That, say, coal company stock investors may lose money when regulators shut down their businesses is not a systemic risk, unless we debase “systemic” to mean anyone ever losing money on anything. Bringing inequality into the financial mandate by claiming that inequality causes systemic runs, as the IMF is doing, is a similar flight of fancy. And once you cook the books to advance climate and inequality, the books are cooked for everything else, too.

As I write, the chance of a systemic crisis induced by a pandemic is a strong possibility. That none of this scenario-building and stress-testing even considered pandemic risk, in the wake of SARS, MERS, Ebola, and HIV, exposes just how much groupthink and virtue-signaling and how little quantifiable prescience any of this effort has—and how utterly this whole project for a regulatory elite to foresee risk has failed. The possibility of advanced country sovereign default is similarly absent from these exercises, though it has happened many times before and would be a calamity to our system built on the sanctity of such debt and its ability to bail others out in crisis.

My objection has nothing to do with the importance or not of climate and inequality or the worthiness or not of these (regulate, de-fund, redistribute) policy approaches to climate and inequality. The main problem is that these are, obviously, highly partisan and deeply political actions on which people disagree rather strongly, at least outside of the bubbles in which international central bankers and NGO staff seem to operate.

Maybe climate change and inequality are the existential problems our economies must address. Perhaps green new deal controls, highly progressive taxation, universal basic incomes, and wealth taxes, rather than a carbon tax and a focus on opportunity—my favorites—are necessary means to fight them. But should *central banks* and their supporting alphabet soup institutions coerce financial institutions and governments to these causes, especially by such transparently dishonest means?

INDEPENDENCE

That question leads me to my final thought: How independent should central banks be? And the obvious answer: If they are going to wade this far into politics, they will quickly lose that independence and we all will lose the benefits of independent central banks.

We coo that the Fed should be independent and free from political interference. But why should the Fed be independent and not, say, the Consumer Financial Protection Bureau? Or the Justice Department? Or the Environmental Protection Agency? Or even the president? What’s different about the Fed?

All agencies live in a balance. We in the U.S. believe in democracy, in accountability, in facing the voters every so often. We do not believe in perpetual rule by independent technocrats—and for good reasons.

On the other hand, we distrust political meddling. The mantra of independent central banking really came of age in 1972, when President Nixon got his Fed to stoke inflation and then clamped it down with disastrous wage and price controls in order to get reelected. We don’t want that to happen again.

Independence is a useful device for our government to pre-commit *ex ante* to not meddle *ex post*, as Ulysses had himself tied to the mast. Our system of government is full of such pre-commitments, from the basics of property rights to the Constitution.

But here's the deal: *An agency in a democracy gains independence only if it accepts and respects sharply limited scope of action.*

The more political an agency's actions, and the more power it exercises to enrich or destroy individual businesses and people, the more it must trade the annoyance of political interference for the reality of political accountability. The IRS, the EPA, the border patrol, tariffs, and the green new deal are all terrible candidates for independence. Judges have great independence—but may only rule on cases before them.

Central bank example A is helicopter money. For a decade now, and three in Japan, central banks have been trying unsuccessfully to raise inflation. Yet central banks are legally forbidden from the one step that most clearly would raise inflation: drop money from helicopters. Why not? Well, in our economy, that means write checks to voters. But in our democracy, only the Treasury and Congress can write checks to voters. And then face those voters over whether it was a good idea.

The Fed may not lend directly to businesses, but only to banks. The Fed may not buy private securities, such as stocks and bonds. Boy, could the Fed stimulate more in bad times with bigger tools. But all of these are obviously political tools.

Now, monetary policy is somewhat political. Changing interest rates and inflation moves money from savers to borrowers, stimulates some industries at the expense of others, and helps some candidates get elected and hurts others. And as a result, the Fed is not totally independent, either in its structure or in its traditions. Presidents appoint board members, subject to confirmation, and the Fed must report to Congress. Congress can rewrite the Fed's founding legislation any time.

More importantly, as elsewhere in our fraying democracy, independence is sustained by a set of norms of behavior: The Fed does not take actions it could legally take, at least outside of extreme crises, and in return politicians do not interfere or demand actions, no matter how desirable politicians may feel those to be.

Independence has a second prerequisite: *technocratic competence*. The Fed keeps its independence so long as the public is convinced the Fed is competent, that there is some science, rule, and order to its policy decisions and not just gut decisions, likely to be influenced by political preferences. A lot of our political moment comes down to the fact that a vast swath of the electorate has lost faith that elites have any idea what they're doing. Listen to that.

So, before we complain about President Trump's tweets or congressional hectoring, how are the Fed, other central banks, the BIS, FSB, IMF, and so forth doing on their part of the deal?

Independence in setting interest rates in the interest of national inflation and unemployment is a settled question, despite its political implications.

Though other central banks are buying stocks and corporate and green bonds, our Fed has tasted mortgage-backed securities and wisely stepped back from the buffet. I sense in this, and the Fed's reluctance to pursue a larger balance sheet, a well-placed fear that Congress would see Fed assets as a kitty for unwise spending. The Fed is limiting its scope of action to

preserve its independence. (This fear is one more reason for the Treasury to provide its own version of reserves.)

But if the Fed takes on macroprudential policy, decreeing that San Francisco real estate is in a bubble or the “credit cycle” is too hot, and uses regulatory power to curtail lending, a lot of homebuyers and builders will be hopping mad and call their congresspeople. If the Fed routinely starts managing stock and bond prices—in both directions—it will not stay independent long.

And if central banks force financial institutions to implement green new deal climate and inequality policies, the result must be, and will be, an end to central bank independence. The minute anyone in the Trump administration reads the IMF’s new policy guidance, expect a storm of protest and a huge reduction in the IMF’s ability to act as a politically neutral technocratic institution. These are not only intensely partisan political issues, they are issues where central banks and associated NGOs have no special technocratic competence.

There is a larger trend in our government of institutions overstepping norms—presidents ruling by executive order, agencies issuing regulations beyond statutory authority, judges passing nationwide injunctions on policy issues they don’t agree with—and it is tearing the country and those institutions apart. Please, Fed, central banks, and NGO cheerleaders, do not follow down this path.

Independence is a great thing. But central banks must buy it by resisting the external call and the internal temptation to ever expand into these politically charged waters. The more often a central bank says, “That is a huge problem, but it’s not our job to fix it,” the more it can preserve its independence to actually fix things it knows how to fix.

I have a lot of political opinions, too, but I do not want the central bank wading in to enforce my political preferences. My proposals go in the other direction. If the Fed’s monetary policy had a simple price-level target, the Fed could be as politically independent as the Bureau of Weights and Measures.²⁹ It would also be about as frequently in the headlines.

Regulation is much more political than monetary policy. Regulation has more ability to transfer money directly to and from specific parties. If we are going to keep the large discretionary regulatory apparatus going, it is worth considering whether the regulatory role should be less independent than the monetary policy role. Bank regulation seems a natural candidate for as much accountability, at the cost of more interference, as the IRS, EPA, or other regulatory agencies.

The unpleasant tradeoff between unaccountable technocrats moving billions from one citizen or business to another, versus politicians doing the same but getting kicked out of office when they do too much, is one more reason to escape detailed regulation in favor of equity-financed banking and narrow deposit taking.

CONCLUSION

With this, I close. We have covered a lot of territory, in the Fed’s review and beyond. Should, and can, the Fed stimulate with strongly negative rates, immense QE asset purchases, and an arsenal of forward guidance speeches? I think not. What sort of target should it follow?

A price-level target. The Fed should get out of the business of setting the level of nominal rates and target the price level directly. Price-level control will be much more effective with fiscal policy coordination. The Fed should offer a flat supply curve of interest-paying reserves, open basically to anyone, though the Treasury should take up much of that role directly.

Going forward, the Fed and its international counterparts should disavow the temptation toward ever-expanding mandates and economic and financial dirigisme that would take them to “macroprudential” policy, discretionary credit cycle management, asset price targeting, and exploiting regulatory power to embrace social and political goals... today on climate change and inequality, perhaps tomorrow on immigration, trade restriction, China-isolation, or whatever the smart set at Davos wants to see. Only limited scope of action to areas of agreed technocratic competence will salvage the Fed’s, other central banks’, and international institutions’ useful independence.

POSTSCRIPT

Naturally, this lecture summarizes a lot of my recent writing on these issues. The underlying citations noted here explain in detail and provide answers to many what-abouts and what-ifs that I do not cover in this short lecture. If you have any objections, these are good places to start. ■

NOTES

- ¹ Cochrane (2018a) is devoted to this point.
- ² Cochrane (2018a, Figure 1).
- ³ Greenwood, Hanson, Rudolph, and Summers (2015).
- ⁴ Cochrane (1998).
- ⁵ Greenlaw, Harris, Hamilton, and West (2018) and Hamilton (2018).
- ⁶ Discussion after the Homer Jones lecture centered on the possibility that stimulus is nonlinear—the first \$3 trillion had a great stimulus effect, but after that no more. There are a lot of speculative nonlinearities and stories that, but for us, the world would have ended. The theme here is that policy should stick to relatively tried-and-true documented principles.
- ⁷ Cochrane (2017).
- ⁸ The CPI has many flaws. I do not mean here to endorse the CPI versus another measure, or to disparage improvements on the CPI. Choose the best price index, and then let it be constant.
- ⁹ Friedman (1969).
- ¹⁰ Admittedly, the same literature that likes forward guidance also says prices and wages should be made more sticky, because in those models the recession gets worse as prices and wages are made less sticky, all the way to the limit of an infinite recession—and then at the flexible price limit point, the economy has no recession. Limits that are not the same as limit points are another easily fixed pathology of such models, in my view. See Cochrane (2017) for this literature and the response.
- ¹¹ See Anderson (2016).
- ¹² This is a central point of Cochrane (2017 and 2018a).
- ¹³ Jacobson, Leeper, and Preston (2019) have a beautiful description of the carefully structured unbacked fiscal expansion that stopped 1933 deflation.
- ¹⁴ Cochrane (2019) describes the idea in more detail. Basically, let the real primary surplus follow $s_t = s(p_t, x_t)$, where p_t is the price level, x_t represents other variables like unemployment, and $s(p_t, x_t)$ is a rising function of p_t . Now even in the case of real debt b_t , the government debt valuation equation $b_t = E_t \sum_{j=0}^{\infty} M_{t,t+j} s(p_{t+j}, x_{t+j})$ implies a unique price level, where $M_{t,t+j}$ represents the stochastic discount factor.
- ¹⁵ Schmitt-Grohé and Uribe (2014) and Uribe (2018) advocate higher interest rates to escape the lower bound. Cochrane (2018a) analyzes the issue extensively. Cochrane (2020) gives a model in which unexpected interest rate rises lower inflation, but expected interest rate rises raise inflation.
- ¹⁶ The last two questions concern the role of lagged interest rates, ρ in $i_t = \rho i_{t-1} + \phi_{\pi} \pi_t + \phi_y y_t + v_t$. Should the Fed follow $\rho = 0$, a pure level rule, $\rho = 1$, a difference rule, or $\rho \in (0, 1)$, a slow-adjustment rule? Many of these options are set out in Federal Reserve Board (2019). Cochrane, Taylor, and Wieland (2020) summarize the rules and analyze many of the rules in the context of a variety of models.
- ¹⁷ Targeting the interest rate spread is the same as targeting the spread between CPI spit and futures. Sumner (1995) gives an extensive analysis of this proposal, including literature review.
- ¹⁸ Cochrane (2015) explains this proposal in detail.
- ¹⁹ This section summarizes Cochrane (2014). It owes a great debt to Admati and Hellwig (2013). See also Miles, Yang, and Marcheggiano (2012).
- ²⁰ See Cochrane (2018b) and Konig (2019).
- ²¹ Gillman (2018) argues that the law requires a price-level target.
- ²² See, for example, the review in “Moving Forward with Macroprudential Frameworks,” Chapter IV of Bank for International Settlements (2018).

²³ For example, IMF (2018) writes

CFMs ...are designed to limit capital flows. These can include administrative and price-based restrictions on capital flows, for instance bans, limits, taxes, and reserve requirements.

See also IMF (2013). The BIS (2019) Annual Report Chapter II chimes in enthusiastically as well. The IMF and BIS reports make clear that they are following emerging common practice at central banks around the world, rather than leading a new agenda. Whether jumping in front of a bandwagon is wise is a good question to ask. If the ambitious but maddeningly vague dirigisme of these reports drives you batty, I recommend re-reading Lucas (1979).

²⁴ Georgieva (2020a).

²⁵ See, for example, Bolton et al. (2020), whose abstract states central banks should step up to

coordinating actions among many players including governments, the private sector, civil society and the international community. ... Those include climate mitigation policies such as carbon pricing, the integration of sustainability into financial practices and accounting frameworks ...

In his foreword, BIS general Manager Augustín Carstens starts reasonably by also advocating carbon taxes—though this has nothing to do with central banks under usual readings of their mandates. But, since carbon taxation “requires consensus building” and is “difficult to implement,” central banks should plow forward to

raising stakeholders’ awareness and facilitating coordination among them. Central banks can coordinate their own actions with a broad set of measures to be implemented by other players (governments, the private sector, civil society and the international community)

...there are many practical actions central banks can undertake (and, in some cases, are already undertaking). They include... environmental, social and governance (ESG) criteria in their pension funds; helping to develop and assess the proper taxonomy to define the carbon footprint of assets more precisely (eg “green” versus “brown” assets); working closely with the financial sector on disclosure of carbon-intensive exposure...; ...examining the adequate room to invest surplus FX reserves into green bonds.

In a separate preface, François Villeroy de Galhau, Governor of the Banque de France, advocates that “more holistic perspectives become essential to coordinate central banks’, regulators’, and supervisors’ actions with those of other players, starting with government.”

²⁶ Carney (2019) is a good place to start. The first step is “disclosure.” The FSB instigated a “task force on climate-related financial disclosures” (TCFD):

four-fifths of the top 1,100 Group of Twenty companies now disclose climate-related financial risks as some TCFD recommendations advise.

The next step is to make disclosure mandatory, as the United Kingdom and European Union have already signaled.

The third step is regulation and de-financing unpopular industries:

Banks ...are taking steps to assess exposure to transition risks in anticipation of climate action. This includes exposure to carbon-intensive sectors, consumer loans for diesel vehicles, and mortgages for rental properties, given new energy efficiency requirements.

The approach is clear: Nice bank you’ve got there. It would be a shame if something should happen to it.

The Bank of England is ...setting out our expectations with respect to the following:

Governance: Firms will be expected to embed the consideration of climate risks fully into governance frameworks, including at the board level...

Risk management: Firms must consider climate change in accordance with their board-approved risk appetite...

Appropriate disclosure of climate risks: Firms must develop and maintain methods to evaluate and disclose these risks.

The Bank of England will be the first regulator to stress-test its financial system under various climate pathways...This stress test will.. make the heart of the global financial system more responsive to changes to both the climate and to government climate policies.

The Bank of England will develop the approach in consultation with ...other informed stakeholders, including experts from the Network of Central Banks and Supervisors for Greening the Financial System....

(Yes, the quotations are selective, so you can see what’s going on in the otherwise sleep-inducing verbiage. Read the originals if you’re unhappy about that.)

²⁷ For example, IMF Managing Director Kristalina Georgieva (2020b) writes

Progressive taxation is a key component of effective fiscal policy...Gender budgeting is another valuable fiscal tool in the fight to reduce inequality...The ability to scale up social spending is also essential... Active labor market policies...job search assistance, training programs, and in some instances, wage insurance...Geographically-targeted policies and investments can complement existing social transfers...

During the implementation of the IMF-supported program, Egypt more than doubled its coverage of cash transfers,...we are working to implement our social spending strategy by weaving it into the fabric of our work...

The IMF (2019) “Strategy for Engagement on Social Spending” goes into details.

...concerns about rising inequality and the need to support vulnerable groups,... a global commitment to continue support for inclusive growth, as expressed in the 2030 Sustainable Development Goals (SDGs)... Social spending is viewed as a key policy lever for addressing these issues.

The Fund has concomitantly increased its work on social spending. ...The growing emphasis on inclusive growth is also reflected in operational activities, including the use of social spending “floors” in IMF-supported programs. There has been enhanced engagement on inequality issues in surveillance, as well as increased technical assistance to expand fiscal space for social spending.

On climate, see Georgieva (2019).

²⁸ *Wall Street Journal* (2019).

²⁹ Technically, the Weights and Measures Division is part of the National Institute of Standards and Technologies, a division of the Department of Commerce, whose head is a political appointee. Independence is as much earned as formal.

REFERENCES

- Admati, Anat R. and Hellwig, Martin. *The Bankers' New Clothes: What's Wrong with Banking and What to Do about It*. Princeton, NJ: Princeton University Press, 2013.
- Anderson, Richard G. “Japan’s Economic Performance ‘Lost Decade’: Myth, Reality, or Role Model?” *Journal of Asian Finance, Economics and Business*, 2016, 3(1), pp. 15-32; <https://doi.org/10.13106/jafeb.2016.vol3.no1.15>.
- Bank for International Settlements. *2018 Annual Report*. 2018; <https://www.bis.org/publ/arpdf/ar2018e4.htm>.
- Bank for International Settlements. *2019 Annual Report*. 2019; <https://www.bis.org/annualeconomicreports/index.htm>.
- Bernanke, Ben. “The New Tools of Monetary Policy.” Manuscript (AEA Presidential Speech), 2020; <https://www.brookings.edu/blog/ben-bernanke/2020/01/04/the-new-tools-of-monetary-policy/>.
- Bolton, Patrick; Despres, Morgan; Awazu, Luiz; Da Silva, Pereira; Samana, Frédéric and Svartzman, Romain. *The Green Swan: Central Banking and Financial Stability in the Age of Climate Change*. Basel: Bank for International Settlements, 2020; <https://www.bis.org/publ/othp31.pdf>.
- Carney, Mark. “Fifty Shades of Green.” *Finance and Development*, 2019, 56(4), pp. 12-15; <https://www.imf.org/external/pubs/ft/fandd/2019/12/a-new-sustainable-financial-system-to-stop-climate-change-carney.htm>.
- Cochrane, John H. “A Frictionless View of U.S. Inflation.” *NBER Macroeconomics Annual*, 1998, 13, pp. 323-84; <https://doi.org/10.1086/ma.13.4623752>.
- Cochrane, John H. “Toward a Run-Free Financial System,” in Martin Neil Baily and John B. Taylor, eds., *Across the Great Divide: New Perspectives on the Financial Crisis*. Stanford: Hoover Press, 2014, pp. 197-249; <http://faculty.chicagobooth.edu/john.cochrane/research/papers/across-the-great-divide-ch10.pdf>.
- Cochrane, John H. “A New Structure for U.S. Federal Debt,” in David Wessel, ed., *The \$13 Trillion Question: Managing the U.S. Government’s Debt*. Washington, DC: Brookings Institution Press, 2015, pp. 91-146.
- Cochrane, John H. “The New-Keynesian Liquidity Trap.” *Journal of Monetary Economics*, 2017, 92, pp. 47-63; <https://doi.org/10.1016/j.jmoneco.2017.09.003>.

- Cochrane, John H. "Michelson-Morley, Fisher, and Occam: The Radical Implications of Stable Inflation at the Zero Bound." *NBER Macroeconomics Annual*, 2018a, 32, pp. 113-226; <https://doi.org/10.1086/696050>.
- Cochrane, John H. "Fed Nixes Narrow Bank." *The Grumpy Economist* (blog), September 5, 2018b; <https://johnhcochrane.blogspot.com/2018/09/fed-nixes-narrow-bank.html>.
- Cochrane, John H. *The Fiscal Theory of the Price Level*. Unpublished manuscript, 2019.
- Cochrane, John H. "A Fiscal Theory of Monetary Policy with Partially-Repaid Long-Term Debt." Unpublished manuscript, 2020.
- Cochrane, John H.; Taylor, John B. and Wieland, Volker. "Evaluating Rules in the Fed's Report and Measuring Discretion," in John B. Taylor and John H. Cochrane, eds., *Strategies for Monetary Policy*. Hoover Institution Press, 2020.
- Federal Reserve Board. *Monetary Policy Report*. February 2019; <https://www.federalreserve.gov/monetarypolicy/2019-02-mpr-summary.htm>.
- Finley, Allisia. "Bloomberg Sells 'Sustainability,' but Buyer Beware." *Wall Street Journal*, March 2, 2020; <https://www.wsj.com/articles/bloomberg-sells-sustainability-but-buyer-beware-11583193439>.
- Friedman, Milton. *The Optimum Quantity of Money and Other Essays*. Chicago: Aldine, 1969.
- Georgieva, Kristalina. "The Adaptive Age." *International Monetary Fund Finance & Development*, December 2019, 56(4), pp. 20-21; <https://www.imf.org/external/pubs/ft/fandd/2019/12/climate-change-and-the-age-of-adaptation-georgieva.htm>.
- Georgieva, Kristalina. "IMF Chief: We Are Rethinking Our Advice to Emerging Markets." *Financial Times*, February 17, 2020a; <https://www.ft.com/content/b53468f0-4e46-11ea-95a0-43d18ec715f5>.
- Georgieva, Kristalina. "Reduce Inequality to Create Opportunity." *IMFBlog*, January 7, 2020b; <https://blogs.imf.org/2020/01/07/reduce-inequality-to-create-opportunity/>.
- Gillman, Max. "The Welfare Cost of Inflation with Banking Time." *B.E. Journal of Macroeconomics*, 2018; <https://doi.org/10.1515/bejm-2018-0059>.
- Greenlaw, David; Harris, Ethan; Hamilton, Jim and West, Kenneth. "A Skeptical View of the Impact of the Fed's Balance Sheet." Manuscript prepared for the 2018 U.S. Monetary Policy Forum, 2018.
- Greenwood, Robin; Hanson, Samuel G.; Rudolph, Joshua S. and Summers, Lawrence H. "Government Debt Management at the Zero Lower Bound," in David Wessel, ed., *The \$13 Trillion Question: Managing the U.S. Government's Debt*. Washington, DC: Brookings Institution Press, 2015.
- Hamilton, James. "The Efficacy of Large-Scale Asset Purchases When the Short-Term Interest Rate Is at Its Effective Lower Bound." *Brookings Papers on Economic Activity*, 2018; <https://doi.org/10.1353/eca.2018.0016>.
- International Monetary Fund. "Key Aspects of Macro Prudential Policy." 2013; <https://www.imf.org/external/np/pp/eng/2013/061013b.pdf>.
- International Monetary Fund. "The IMFs Institutional View on Capital Flows in Practice." 2018; <https://www.imf.org/external/np/g20/pdf/2018/073018.pdf>.
- International Monetary Fund. "A Strategy for IMF Engagement on Social Spending." 2019; <https://www.imf.org/en/Publications/Policy-Papers/Issues/2019/06/10/A-Strategy-for-IMF-Engagement-on-Social-Spending-46975>.
- Jacobson, Margaret M.; Leeper, Eric M. and Preston, Bruce. "Recovery of 1933." NBER Working Paper 25629, National Bureau of Economic Research, 2019; <https://doi.org/10.3386/w25629>.
- Konig, J.P. "The Fed Wants to Close the Window on Narrow Banks." *American Institute for Economic Research* (blog), May 28, 2019; <https://www.aier.org/article/the-fed-wants-to-close-the-window-on-narrow-banks/>.
- Lucas, Robert E. Jr. "A Report to the OECD by a Group of Independent Experts OECD, June 1977: A Review." *Carnegie-Rochester Conference Series on Public Policy*, 1979, 11, pp. 161-68; [https://doi.org/10.1016/0167-2231\(79\)90039-3](https://doi.org/10.1016/0167-2231(79)90039-3).
- Miles, David; Yang, Jing and Marcheggiano, Gilberto. "Optimal Bank Capital." *The Economic Journal*, 2012, 123(567), pp. 1-37; <https://doi.org/10.1111/j.1468-0297.2012.02521.x>.

Schmitt-Grohé, Stephanie and Uribe, Martín. "Liquidity Traps: An Interest-Rate-Based Exit Strategy." *The Manchester School*, 2014, 81, pp. 1-14; <https://doi.org/10.1111/manc.12065>.

Sumner, Scott. "The Impact of Futures Price Targeting on the Precision and Credibility of Monetary Policy." *Journal of Money Credit and Banking*, 1995, 27, pp. 89-106; <https://doi.org/10.2307/2077852>.

Uribe, Martín. "The Neo-Fisher Effect: Econometric Evidence from Empirical and Optimizing Models." NBER Working Paper 25089, National Bureau of Economic Research, 2018; <https://doi.org/10.3386/w25089>.

Wall Street Journal. "Bankers Aren't Climate Scientists." Editorial, November 15, 2019; <https://www.wsj.com/articles/bankers-arent-climate-scientists-11573861841>.

Is the Phillips Curve Still Alive?

Brian Reinbold and [Yi Wen](#)

A.W. Phillips's discovery that inflation is negatively correlated with unemployment served as a heuristic model for conducting monetary policy; but the flattening of the Phillips curve post-1970 has divided debate on this empirical relation into two camps: "The Phillips curve is alive and well," and "The Phillips curve is dead." However, this dichotomy oversimplifies the issue. In this article, we apply spectral analysis to the U.S. inflation rate and unemployment rate to conduct a comprehensive analysis of the Phillips curve in the frequency domain. We find that in the very short run, there is no systemic relationship between inflation and unemployment; in the intermediate run, which includes the business cycle frequency, they are strongly negatively correlated; and in the very long run the Phillips curve is strongly positively sloped. Such an analysis of the frequency domain provides a natural demarcation of frequency bands that allows us to recover the Phillips curve in the time domain by applying band-pass filters. Most importantly, we show how spectral analysis can be used to identify a "supply" (permanent) and a "demand" (nonpermanent) shock in the context of a vector autoregression and that demand shocks drive the Phillips curve. Finally, the phase spectral analysis also shows that despite the existence of the Phillips curve at the business cycle frequency under a demand shock, the monetary policy implications are not obvious, due to the unclear lead-lag relationship between inflation and unemployment. (JEL C100, E240, E310, E520, E580)

Federal Reserve Bank of St. Louis *Review*, Second Quarter 2020, 102(2), pp. 121-44.
<https://doi.org/10.20955/r.102.121-44>

1 INTRODUCTION

In 1958, economist A.W. Phillips discovered a strong negative correlation between the money wage rate and the unemployment rate in the United Kingdom.¹ Shortly after his findings were published, numerous studies confirmed that this relationship held in many developed economies. For example, Samuelson and Solow (1960) demonstrated that the Phillips curve held in U.S. data, and they began to explore its policy implications.

The profession holds that the inverse relationship between unemployment and inflation implies a tradeoff between the two: low unemployment at the cost of higher inflation or low

Brian Reinbold is a research associate and Yi Wen is an assistant vice president and economist at the Federal Reserve Bank of St. Louis.

© 2020, Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

inflation at the cost of higher unemployment. This tradeoff provides policymakers a menu of monetary policy prescriptions and also shows them how influencing nominal variables can affect the real economy. Monetary policy, for example, can adjust the money supply or nominal interest rates to affect the price level and then through the Phillips curve affect employment. Because of the explanatory power of the Phillips curve, after its introduction, economists immediately incorporated it into structural models and this literature flourished and became an indispensable part of Keynesian economics.

However, the 1970s saw the Phillips curve breakdown, and the correlation in fact became positive. The U.S. experienced higher oil prices, and these adverse supply shocks caused the Phillips curve to disappear. Economists then worked on alternative explanations to rectify this experience. One branch of research incorporated rational expectations under supply shocks and long-run neutrality of money (meaning that the Phillips curve is flat in the long run in the absence of supply shocks). Another branch had a much different approach, where prices are sticky but monetary policies are endogenous responses to output gaps and inflation (Gordon, 2011). This split in analyzing the Phillips curve led to two very different conclusions on the Phillips curve: “The Phillips curve is alive and well,” and “The Phillips curve is dead.” Since the 1970s, a plethora of theoretical models and regression techniques, ranging from vector autoregression (VAR) to instrumental variable models, have been developed to study the existence of the Phillips curve.

Despite the numerous econometric specifications economists have used for that purpose, very few have investigated the Phillips curve in the frequency domain using spectral analysis. The issue with pure time-domain methods is that it is difficult to distinguish short-run, intermediate-run, and long-run relationships between the inflation rate and the unemployment rate, especially when the time series are full of noise that can mask the underlying dynamics of the data. Although filters such as a band-pass filter can be used to isolate specific components of the data according to the specified frequency of fluctuation, this becomes arbitrary since there are an infinite number of specifications of the frequency bands. Furthermore, the lead-lag relations between inflation and unemployment can also be masked by noise, which makes systematic regression analyses in the time domain challenging.

Spectral analysis, on the other hand, provides a clear way to decompose a time series and its relationships with other time series into movements and comovements across a continuum of cyclical frequencies and leads and lags—all at once. That way the short-run, the intermediate-run, and the long-run behaviors of a vector of time series and their mutual cross lead-lag correlations (or covariances) can be studied simultaneously without resorting to filters where one must specify arbitrarily the frequency intervals a priori. Thus, spectral techniques provide a more compact and complete picture of the joint dynamic behaviors of a vector of time series and ultimately provides economists an additional tool to better characterize any systematic relationships at any cyclical frequencies at any leads and lags among any specified number of economic variables.

The main purpose of this article is to demonstrate how spectral techniques can be used to analyze the U.S. Phillips curve and extract information not easily seen in the time domain. Using such techniques, we find that (i) in the very long run (such as fluctuations at frequencies

lower than 0.02 cycles per quarter or 50 up to infinity quarters per cycle) the Phillips curve is positively sloped, except in the 1950s and 1960s when the Phillips curve became popular; (ii) however, in the intermediate run (i.e., around frequencies of 6 to 50 quarters per cycle), the Phillips curve is alive and well—the correlation between unemployment and inflation is always negative and significant throughout the entire postwar U.S. history, including in the 1970s and 1980s when the Phillips curve was thought to have broken down; and (iii) in the very short run (fluctuations at frequencies higher than 0.17 cycles per quarter or less than 6 quarters per cycle), there is zero correlation between unemployment and inflation throughout the entire sample. We also find the Phillips curve at the conventional business cycle frequency to be highly stable over time.²

These findings explain why in the time domain it is hard to detect the existence of the Phillips curve (especially since the 1970s), because the long-run, intermediate-run, and short-run movements are mixed and thus offset each other in the time domain; in addition, the large amount of noise in the inflation rate has dominated and masked any systematic relationships the rate has with unemployment.

Perhaps equally important and interesting is that the monetary policy implications (justified by the Keynesian models that use the Phillips curve) are not obvious, based on the phase spectrum. The conventional wisdom is that a negative correlation between inflation and unemployment automatically implies a trade-off between the two and a causal link running from inflation to unemployment. But the phase spectrum shows that in the long run, high inflation tends to “cause” high unemployment instead of low unemployment, and in the intermediate run high inflation tends to follow (instead of lead) low unemployment. Therefore, it is not clear how monetary policy that directly affects inflation could affect unemployment during the business cycle despite the strong evidence of a negatively sloped Phillips curve at the business cycle frequency.

The rest of the article is organized as follows. Section 2 provides a brief literature review of the Phillips curve. Section 3 provides a brief technical review of spectral analysis. Section 4 presents the data and empirical findings. Finally, Section 5 provides some concluding remarks.

2 LITERATURE REVIEW

To conserve space, we review only the works most closely related to ours. Scully (1974) was one of the first papers to employ spectral analysis on the U.S. Phillips curve. Using data from 1900 to 1969 on changes in money wages, the inverse of the unemployment rate, and changes in the price level, Scully (1974) found that unemployment lags wage inflation and follows a distributed lag structure. Importantly, he found that most of the variance of each time series is captured in the business cycle frequency range of 4 to 16 years. Furthermore, the relationships between (i) wage inflation with the inverse of the unemployment rate and (ii) price level changes with the inverse of the unemployment rate are not independent of frequency. The systematic shifting of the gain estimate suggests that the inflation-unemployment tradeoff is transitory and not permanent.

King and Watson (1994) apply a band-pass filter to recover the U.S. Phillips curve using monthly data on the consumer price index inflation rate and the unemployment rate from

1950 to 1992. They extracted long-run movements in the time series by applying a low-pass filter (isolating periodicities greater than 8 years) and find that the Phillips relation experiences a change in 1970. Before 1970, there was a strong negative correlation between the inflation rate and the unemployment rate (-0.62) but there is no consistent relation afterwards. The overall correlation is 0.50 . These long-run dynamics of the Phillips relation became more important after 1970, contributing to the flattening of the Phillip's curve in the full sample of the raw data. Applying a band-pass filter that isolates components at periodicities between 18 months and 8 years, they find a stable Phillips curve at the business cycle frequency. The correlation was -0.69 from 1954 to 1969, -0.67 from 1970 to 1987, and -0.66 for the whole sample. Despite the inflation process becoming more volatile post 1970, they find that the Phillips relation remains strong and stable at the business cycle frequency. Finally, at high-frequency components (i.e., short-term fluctuations), they find inflation has much larger variation than unemployment and that these components have a slightly negative correlation.

Iacobucci (2005) applies cross-spectral analysis to the U.S. Phillips curve and finds that the Phillips curve is negatively sloped at the frequency band between 3 and 14 years, with a -0.38 correlation. Furthermore, they find that unemployment leads inflation.

Gallegati et al. (2011) use wavelet analysis to study the U.S. Phillips curve. Wavelet analysis allows time and frequency to be studied simultaneously, as it provides an estimate of the frequency structure of a signal locally at a given point in time. Thus, the frequency resolution is allowed to vary across time for the given time series. Intuitively, we can think of wavelet analysis as subsample spectral analysis over a moving window of observations. They regress wage inflation on the unemployment rate, price inflation, and labor productivity growth at different time scales. They find the following: At high-frequency scales, there is a negative relation between wage inflation and the unemployment rate, but it is not statistically significant. However, there is a strong, statistically significant negative relationship at the business cycle frequency. At scales greater than 8 years, the long-run components still exhibit a negative relationship but less so than at the business cycle frequency. Furthermore the low-frequency components explain a substantial component of the total variation, so they are a large driver of the Phillips relation. The varying estimates at different time scales is a symptom of the non-linearity in the wage-unemployment relationship. In addition, they find a stable Phillips curve relation pre-1993, but this relationship breaks down afterwards.

3 KEY CONCEPTS OF SPECTRAL ANALYSIS

We provide a brief review of some basic concepts in spectral analysis to help us understand the estimated spectral density functions based on U.S. data in the next section. A more complete and self-contained technical review of spectral analysis is provided in the appendix.

Given a white-noise process, we can construct a new time series y_t as distributed leads and lags of the white noise. This new time series has its autocovariance function defined at any lag $\tau \in (-\infty, \infty)$, which can be denoted as $c_y(\tau)$. Then, applying the Fourier transform to the sequence $\{c_y(\tau)\}_{\tau=-\infty}^{\infty}$ gives the power spectrum (or spectral density function) of y_t denoted by $g_y(e^{-i\omega})$.

There are several basic properties of the spectral density function over the domain $\omega \in [-\pi, \pi]$: (i) It is real valued and nonnegative; (ii) it is symmetric about $\omega = 0$, or $g_y(e^{-i\omega}) = g_y(e^{i\omega})$; and (iii) a spectral peak (local maximum) at ω_0 reflects relatively large contributions from cyclical fluctuations of around $2\pi/\omega_0$ periods per cycle.

Using the inverse Fourier transform, we can recover the autocovariance at any lead or lag $\tau \in (-\infty, \infty)$:

$$c_y(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g_y(e^{-i\omega}) e^{i\omega\tau} d\omega$$

$$c_y(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g_y(e^{-i\omega}) d\omega,$$

where $c_y(0)$ is simply the variance of the original time series at $\tau = 0$.

The above two equations show clearly that the spectrum $g_y(e^{-i\omega})$ is a moment-generating function for autocovariance at any leads or lags and that the total area underneath the spectrum is proportional to $c_y(0)$, which is the total variance of y_t . It is in this sense that we view the spectrum as a distribution of variance across cyclical frequencies in the domain $\omega \in [-\pi, \pi]$.

Analogously, given another time series x_t , the covariance between the two time series y_t and $x_{t-\tau}$ at any lead or lag $\tau \in (-\infty, \infty)$ can be defined as $c_{yx}(\tau)$. Applying the Fourier transform to the sequence $\{c_{yx}(\tau)\}_{\tau=-\infty}^{\infty}$ gives the cross spectrum of y_t and x_t , denoted by $g_{yx}(e^{-i\omega})$.

The cross spectrum is a “cross-covariance” generating function; namely, given $g_{yx}(e^{-i\omega})$, we can recover the covariance of y_t and $x_{t-\tau}$ at any lead or lag $\tau \in (-\infty, \infty)$:

$$c_{yx}(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g_{yx}(e^{-i\omega}) e^{i\omega\tau} d\omega.$$

In particular, setting $\tau = 0$ we have

$$c_{yx}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g_{yx}(e^{-i\omega}) d\omega.$$

So the covariance of y_t and x_t (at $\tau = 0$) is proportional to the area underneath the cross spectrum.

There are several basic properties of the cross-spectral density function (cross spectrum): (i) It is complex-valued, (ii) its real part is symmetric about $\omega = 0$: $g_{yx}(e^{-i\omega}) = g_{xy}(e^{i\omega})$, and (iii) its imaginary part has rotational symmetry about $\omega = 0$.

The real part of the cross spectrum is called the “cospectrum” and the imaginary part the “quadrature spectrum.” Given these symmetry properties of the spectrum and cross spectrum, in spectral analysis we need only to consider the domain $\omega = [0, \pi]$.

Notice that (i) the sign of the cospectrum (the real part) reflects the sign of the covariance between y_t and x_t —it can be either positive or negative—and (ii) the quadrature spectrum captures the phase differences of cycles in the two series (discussed more below).

As in the time domain where we obtain correlation by normalizing the covariance of two time series by the product of their respective variances, the frequency-domain analog is the coherence function, which is essentially the (spectral) distribution of the absolute value of the correlation between y_t and x_t across frequencies.

Also, since any complex number has a polar form representation, we can also express the cross spectrum in its polar form:

$$g_{yx}(\omega) = r(\omega)e^{i\theta(\omega)},$$

where the function $r(\omega)$ is called the gain and the function $\theta(\omega)$ is called the phase, which has a maximum of $\pi/2$ and a minimum of $-\pi/2$.

The meaning of the gain and the phase can be illustrated in the following simple example: Let $y_t = x_{t-k} = L^k x_t$, where L^k is the lag operator to the k^{th} power. Clearly, shifting the original time series x_t by k units of time does not change the amplitude of the cyclical components of the original time series in the frequency domain at any frequency. Therefore, the gain function is 1 across all frequencies. However, there exist serious phase effects. Since y_t can be predicted by x_{t-k} k units ahead of time, it lags x_t by k units of time, suggesting that the events in x_t are delayed in y_t or that the cyclical phase of fluctuations in y_t is simply the same as in x_t but shifted backward in time. Thus, it can be predicted by the past history of x_t . This phase effect is captured by the phase function $\theta(\omega)$. In particular, y_t lags (leads) x_t if the phase $\theta(\omega)$ is negative (positive) at frequency ω .

However, caution must be exercised when using the phase function to gauge lead-lag relations. The reason is that, for pure *sin* waves, a lead can also be interpreted as a lag if the lead is too big (larger than half the length of the full cycle or 180° in phase angle). Since the phase function switches sign for every 90° or $\frac{\pi}{2}$, it is better to use the sign of the quadrature spectrum to gauge the lead-lag relationship.

4 DATA ANALYSIS AND FINDINGS

4.1 Data in the Time Domain

We use quarterly, seasonally adjusted consumer price index and unemployment rate data from 1948 to 2018 (from the Bureau of Labor Statistics [BLS]). The annualized inflation rate is calculated as

$$400 * \left(\frac{CPI_t}{CPI_{t-1}} - 1 \right).$$

Panel A of Figure 1 shows both the full sample of unemployment and the inflation rates in the time domain. It's immediately evident that inflation is significantly more volatile than unemployment and contains more variations at the short-run (high) frequencies as well. Their relationship in the time domain is revealed in Panel B, which plots the unemployment rate on the horizontal axis against the inflation rate on the vertical axis. Over the sample, no significant relationship exists between the two series; the sample correlation is 0.05 with a standard error of 0.059. When we break the sample into two subsamples (red dots for 1948 to 1969 and green dots for 1970 to 2018), the correlation is -0.39 for the first subsample and 0.03 for the second subsample. These simple statistics suggest the well-known “facts”: The Phillips curve exists in the earlier period and then breaks down in the latter, suggesting that the Phillips curve is unstable or even an artifact of the past when viewed in the time domain.

Figure 1**Unemployment Rate and Inflation Rate**

SOURCE: BLS and authors' calculations.

4.2 Data in the Frequency Domain

We now look at the two time series in the frequency domain by estimating a two-variable VAR for the inflation rate and unemployment rate, and we apply the Fourier transform on the VAR to produce the power spectrum and cross spectrum of the two variables.³

Figure 2 plots the power spectrum of the unemployment rate (Panel A) and the inflation rate (Panel B). The horizontal axes measure frequency as the number of cycles per quarter (the inverse being the number of quarters it takes to finish a full cycle). For example, frequency zero literally means zero cycles per quarter, while its inverse means that it takes infinite periods to finish a cycle, thus movements at this frequency capture the long-run trend of the time series. Frequency 0.5 means a half cycle per quarter or that it takes two periods to finish a full cycle and thus is the highest frequency (or shortest cycle) possible for any cyclical movement. The vertical axes measure the density (power) of the spectrum, and the total area underneath the spectral density function is proportional to the total variance of the time series, which is normalized to 1 for convenience (the normalization does not change the shape of the spectral function).

Panel A shows that for unemployment, movements at low frequencies (i.e., at or near zero) and business cycle frequencies (i.e., 6 to 40 quarters per cycle, shaded in gray) accounted for more than 95 percent of its time-domain variations (variance). Similarly, Panel B shows that these movements contributed most (more than 70 percent) of the variance in the inflation rate, but that movements at high frequencies (2 to 6 quarters per cycle) also contribute to the variance.

Next, we investigate the relationships between unemployment and inflation in the frequency domain. In Figure 3, Panel A plots the coherence function, which is analogous to the

Figure 2**Power Spectrum of the Unemployment Rate and Inflation Rate**

NOTE: Gray bars indicate business cycle frequencies of 6 to 40 quarters.

SOURCE: BLS and authors' calculations.

absolute value of the correlation in the time domain. At low and business cycle frequencies, the two series are strongly correlated (with peak correlations of 0.57 at frequency zero and frequency 0.05, respectively, which each correspond to a long-run trend and periodicity of 20 quarters per cycle, respectively). Such strong correlations are completely missed by the time-domain plots in Panel A of Figure 1. The reason is that too much noise exists in the time domain and masks any systematic relationships. Indeed, the coherence function shows that at high frequencies, the coherence is mostly near zero except for a local peak at frequency 0.35. Hence, the coherence function suggests that the two seemingly uncorrelated time series (when viewed in the time domain) are actually strongly correlated at low and business cycle frequencies. But the noise at high frequencies (especially in the inflation rate) has masked such strong relationships.

However, note that the coherence does not give us any information on the sign of the correlation or the lead-lag relationships between the two variables. Hence, we need to explore the spectral information further. To highlight the details, we have truncated the spectral functions for frequencies higher than 0.18 in Panels B-D of Figure 3.

The cospectrum that captures the covariance between unemployment and inflation across cyclical frequencies is shown in Panel C. It suggests that the two variables are positively correlated in the very long run (for frequencies lower than 0.02). However, they are negatively correlated at frequencies above 0.02 and have a trough at around 30 quarters per cycle. Their covariance becomes very small for frequencies higher than the conventional business cycle frequency interval (i.e., $\omega \geq 1/6$).

Figure 3**Cross-Spectral Analysis**

NOTE: Gray bars indicate business cycle frequencies of 6 to 40 quarters.

SOURCE: BLS and authors' calculations.

For lead-lag relationships, we consider the phase spectrum (Panel B) and the quadrature spectrum (Panel D). The phase spectrum fluctuates between $\frac{\pi}{2}$ and $-\frac{\pi}{2}$, or between 0.5 and -0.5 after normalizing the vertical axes by π . The phase function is negative in the low-frequency interval $[0, 0.02]$, becomes positive in the low-frequency interval $[0.02, 0.05]$, and then switches sign again in the business cycle frequency interval $[0.05, 0.17]$. In principle, a positive value of the phase function implies that unemployment leads inflation and a negative value implies that it lags inflation.

However, as explained in the technical analysis, since the phase function switches sign for every 90° before a full cycle is finished and since the first sign switch at frequency 0.02 is not a continuous movement, we can interpret the phase function in the interval $[0, 0.05]$ as an indication that unemployment lags inflation until the second sign switches around fre-

quency 0.05. That is, for low-frequency movements (with periodicity longer than 20 quarters per cycle), unemployment lags inflation, but for business cycle frequencies in the interval $[0.05, 0.125]$ (8 to 20 quarters per cycle), unemployment slightly leads inflation (although the maximum lead is only 0.1π , or 20° , and is not very significant).

Panel D of Figure 3, the quadrature spectrum, shows that unemployment indeed lags the inflation rate in the frequency interval $[0, 0.05]$ (20 up to infinite quarters per cycle). The lead-lag relation is slightly reversed but not significant in higher-frequency intervals such as the frequency interval $[0.05, 0.125]$. Thus, at low frequencies or in the very long run and part of the intermediate run, inflation significantly leads unemployment; at the shorter end of the intermediate run, inflation slightly lags unemployment, but the relationship is not significant; and in the high-frequency interval $[0.17, 0.5]$ (not shown in Panels B-D), no significant lead-lag relationship exists between inflation and unemployment.

However, even though the quadrature spectrum in Panel D indicates that unemployment significantly lags inflation in the low-frequency interval $[0, 0.05]$, which suggests that inflation “causes” unemployment at low frequencies, the cospectrum in Panel C shows that the sign of the cospectrum switches once from positive to negative in the middle of this low-frequency interval at around frequency 0.02 (50 quarters per cycle), suggesting that high inflation “causes” (leads to) high unemployment in the very long run (cycles longer than 50 quarters) but “causes” (leads to) low unemployment in frequency interval $[0.02, 0.05]$ cycles per quarter (20 to 50 quarters per cycle).

To summarize, the Phillips curve is positively sloped in the very long run in frequency interval $[0, 0.02]$ and negatively sloped in the low-frequency interval $[0.02, 0.025]$. At the conventional business cycle frequency interval $[0.025, 0.17]$ (6 to 40 quarters per cycle), the cospectrum remains negative, suggesting that the Phillips curve is negatively sloped, although the lead-lag relationships change directions slightly, from inflation strongly leading unemployment in frequency interval $[0, 0.05]$ to weakly lagging unemployment in part of the business cycle frequency interval $[0.05, 0.125]$.

4.3 Time-Domain Filtering

The cospectrum in Panel C of Figure 3 tells us how to design band-pass filters to extract cyclical fluctuations from unemployment and inflation in the time domain: We should specify three frequency bands, $[0, 0.02]$, $[0.02, 0.17]$, and $[0.17, 0.5]$ (50 quarters to infinity, 6 to 50 quarters, and 2 to 6 quarters per cycle, respectively).⁴ We should then be able to find that the correlation between the filtered unemployment and inflation rates is strongly positive in the first frequency band, strongly negative in the second frequency band, and near zero in the third frequency band.

Notice that the second frequency band includes the conventional business cycle frequency interval of $[0.025, 0.17]$ (6 to 40 quarters per cycle) and part of the low-frequency movements in the interval $[0.02, 0.025]$ (40 to 50 quarters per cycle). Our findings from spectral analysis suggest that the conventional definition of the business cycle frequency interval $[0.025, 0.17]$ may be too restrictive for analyzing the Phillips curve.

Indeed, these predictions are confirmed in the time domain, as shown in Figure 4. In particular, Panel A of Figure 4 shows the high-frequency components of the data obtained

Figure 4**Band-Pass-Filtered Unemployment and Inflation Rates**

SOURCE: BLS and authors' calculations.

by applying a band-pass filter that isolates fluctuations in the third frequency band that correspond to cycles with periodicity between 2 and 6 quarters per cycle. At this high-frequency band, inflation is extremely volatile compared with unemployment and their correlation is essentially zero, with a correlation of -0.06 with a standard error of 0.062 , which is not significant. This correlation is slightly positive before 1970 $[0.10]$ and slightly negative after 1970 $[-0.16]$.

Panel B shows time-domain movements in the second frequency band (which includes the conventional business cycle components of the series) obtained by applying a band-pass filter that isolates components in the frequency band that correspond to cycles with periodicity between 6 and 50 quarters per cycle. As the panel shows, as inflation rises, unemployment falls over the business cycle. The Phillips curve is also remarkably stable in this frequency band, with a correlation of -0.45 for the whole sample, -0.41 before 1970, and -0.47 after 1970. Interestingly, the unemployment rate and the inflation rate are actually more negatively correlated post-1970, in contrast to the nonfiltered raw data where the Phillips curve reverses.

Panel C of Figure 4 shows the movements in unemployment and inflation at the low-frequency band of 50 to infinite quarters per cycle (which is the residual component after subtracting the movements from the first two frequency bands). Obvious change is evident in the correlation between the long-run components of inflation and unemployment around 1970. Pre-1970, there is a strong negative correlation of -0.53 , but that reverses to 0.26 after 1970. The overall correlation is 0.38 .

These results suggest that the Phillips curve is alive and well over the intermediate run (including the conventional business cycle frequency interval) but that short-run irregularities and the long-run positive correlation have masked the Phillips curve relation in the time domain in the raw data. The low-frequency movements post-1970 have especially flattened the Phillips curve.

4.4 Shock Decomposition

The VAR-based spectral analysis of unemployment and inflation can be exploited further in both the frequency domain and the time domain. For example, the spectral density of both unemployment and inflation have a spectral peak at frequency zero. Using the shock-identification method in the frequency domain proposed by Wen (2001, 2002), we conduct a structural analysis in the frequency domain to decompose the spectrum (or cross spectrum) into two orthogonal components: (i) a spectrum (or cross spectrum) generated from a long-run shock that is responsible for the peak spectrum of unemployment at frequency zero (or in a frequency band near zero) and (ii) a spectrum (or cross spectrum) generated from an orthogonal shock that is not responsible for movements of unemployment at very low frequencies.

As explained by Wen (2001, 2002), the time domain analog of this frequency-domain decomposition is the method of Blanchard and Quah (1989), but Wen's spectral decomposition method is far more general since it can be applied in the frequency domain to any frequencies or frequency intervals. Also, it can be applied to either the power spectrum or the cross spectrum (cospectrum or quadrature spectrum) or any combination of them, whereas the method of Blanchard and Quah (1989) is only a special case of Wen's generalized shock-identification method in the frequency domain. Wen's generalized spectral method is useful

here because we can use it to find structural shocks responsible for the existence and stability of the Phillips curve.

We now apply the method of Wen (2001, 2002) to the spectrum (or cross spectrum) of unemployment and inflation. In particular, for the sake of demonstration, we simply assume that two orthogonal structural shocks are responsible for the movements in unemployment and inflation, as in the spirit of Blanchard and Quah (1989) in the traditional structural VAR literature. One shock is called the “permanent shock,” and the other is called the “nonpermanent shock,” and intuitively we can think of the permanent shock as a “supply” shock and the nonpermanent shock as a “demand” shock. The permanent shock is assumed to be responsible for the peak spectrum of unemployment at frequency zero, and the nonpermanent shock is assumed to be not responsible for this movement. Notice that these identifying assumptions do not impose any restrictions regarding how much the first and second shocks influence the spectrum of unemployment at nonzero frequencies—or how these two shocks influence the spectrum of inflation or the cross spectrum at any frequencies (including frequency zero). Instead, based on the limited identifying assumption, we let the data tell us how these two identified shocks determine the joint behaviors of unemployment and inflation in the frequency domain.

Once the structural identification and spectral decomposition are achieved in the frequency domain, the results can be easily converted to the time domain through the inverse Fourier transform, as the following exercises show.

Figure 5 demonstrates the structural decomposition of the spectrum and cospectrum under our simple identifying assumptions in the spirit of Blanchard and Quah (1989). Panel A shows the power spectrum of unemployment and its structural decomposition, where the blue solid line is the data power spectrum of unemployment, the red-dashed line is the contribution of the permanent shock to the spectrum of unemployment, and the green-dashed line is the contribution of the nonpermanent shock to the spectrum of unemployment. The identifying assumption is that the permanent shock explains the maximum amount of the power spectrum of unemployment at frequency zero, while the other shock explains the minimum, but no further assumptions or restrictions are made regarding their respective contributions to unemployment at nonzero frequencies or their contributions to inflation at any frequency (including frequency zero). As proved by Wen (2001, 2002), the identification is unique, so these two shocks’ contributions add up to the total spectral densities at each frequency between 0 and π . This panel shows that the permanent shock is largely responsible for the bulk of movements in unemployment at low frequencies, but its contribution declines toward higher frequencies, while the nonpermanent shock is largely responsible for the movements at the conventional business cycle frequency and high-frequency intervals.

However, with respect to inflation, Panel B shows that the two shocks split the data spectrum (blue line) almost evenly across frequencies, or each shock contributes to about half of the power spectrum, with the permanent shock slightly dominating the nonpermanent shock at frequency zero.

The most interesting results are with respect to the cross spectrums shown in Panels C and D. Panel C shows that the data cospectrum (blue solid line) switches signs from positive

Figure 5**Spectral Decomposition of Unemployment and Inflation**

NOTE: Gray bars indicate business cycle frequencies of 6 to 40 quarters.

SOURCE: BLS and authors' calculations.

to negative at frequency 0.02. The structural decomposition reveals this switching occurs because of the contributions of the nonpermanent shock (green-dashed line), which generate negative covariance between unemployment and inflation across all frequencies, thus is essentially 100 percent responsible for the existence of the Phillips curve. In other words, without the nonpermanent shock, there would be no Phillips curve, since the permanent shock alone would have generated the cospectrum (red-dashed line), which is positive across the entire frequency domain.

Panel D shows the quadrature spectrum that reflects the lead-lag relationships between unemployment and inflation. It reveals that the permanent shock (red-dashed line) is entirely responsible for the negative part of the quadrature spectrum, so that unemployment lags

Figure 6**Impulse Response Function of Unemployment and Inflation Under Structural Shocks**

SOURCE: BLS and authors' calculations.

inflation significantly under this shock, which explains the effects of the two major oil shocks in the 1970s and early 1980s, during which high inflation caused by high oil prices generated high unemployment and stagnation in the U.S. economy. The nonpermanent shock (green-dashed line), however, is responsible for the slightly positive quadrature spectrum in the data (blue solid line) because under this type of shocks there is a sign switch near frequency 0.03, suggesting that under a nonpermanent shock, unemployment tends to lag inflation only at frequencies below 0.03 but tends to lead inflation at frequencies above 0.03.

These findings have important monetary policy implications. If monetary policy mainly affects nonpermanent shocks or the demand side of the economy, then policymakers may be able to use high-inflation policy to lower unemployment in the long-run frequency interval $[0.00, 0.03]$, which does not include much of the conventional business cycle frequency interval $[0.025, 0.17]$. In the conventional business cycle frequency interval, inflation lags unemployment instead, so monetary policy does not “cause” unemployment through inflation targeting, although such a “causal” effect is not very strong based on quarterly data.

Since the spectrum and cospectrum are moment-generating functions and they contain information about autocovariance and cross variance of a vector of time series at any leads and lags, the above structural decomposition can be converted back to the time domain by applying the inverse Fourier transform.

Figure 6 shows the impulse response functions of unemployment and inflation under the permanent and nonpermanent shocks, respectively, identified in the frequency domain. Panel A shows that unemployment (green solid line) and inflation (red-dashed line) move in

Figure 7**Unemployment vs. Inflation under Permanent and Nonpermanent Shocks**

SOURCE: BLS and authors' calculations.

opposite directions under a one-standard-deviation nonpermanent shock, and both series are highly persistent. Panel B shows that unemployment and inflation move together under a permanent shock because both respond positively to the shock.

Next, we simulate time series based on the uncovered autocoefficients from the inverse Fourier transform and the two identified structural shocks (using information from the residuals of the VAR in the time domain upon which the Fourier transform is based). The uncovered time series under each structural shock process and their relationships are presented in Figure 7.

Panel A of Figure 7 shows the raw unemployment data (vertical axis) against the raw infla-

tion data (horizontal axis). The straight regression line shows that the slope is weakly positive, with a value of 0.03, which is not statistically different from zero. Hence, according to this panel, the Phillips curve does not exist, as the literature often claims.

Panel B shows the same two time series but under only the permanent shock. The regression line has a strong, positive slope, with a value of 0.33, which is highly statistically significant. Hence, according to this panel, the Phillips curve is turned on its head under a permanent shock.

Panel C shows the same two time series under only the nonpermanent shock. The regression line is negatively sloped, with a value of -0.35 , which is highly statistically significant. Hence, according to this panel, the Phillips curve definitely exists under a nonpermanent shock.

Finally, Panel D mixes the decomposed time series together under both shocks and then plots the mixed unemployment series against the mixed inflation series. It shows the same result as Panel A based on the raw data—namely, the Phillips curve disappears.

These results suggest that the very reason that the literature cannot find the existence of the Phillips curve is simply because permanent and nonpermanent shocks lead to the opposite slopes of the Phillips relationship, so the true Phillips curve is masked by permanent shocks in the time domain, as our structural analyses clearly reveals.

5 CONCLUSION

We employ spectral techniques to analyze the U.S. Phillips curve in the frequency domain. We find that (i) the Phillips curve is strongly positively sloped at the low-frequency interval $[0, 0.02]$ (50 or more quarters per cycle) and (ii) the Phillips curve has the correct negative slope in the frequency interval $[0.02, 0.17]$ (6 to 50 quarters per cycle), which includes the conventional business cycle of 6 to 40 quarters and a portion of the low-frequency interval. The negative relationship is also stable over time despite the oil shocks during the 1970s. There do not exist systemic relationships between the two series at frequencies higher than 0.17 (6 quarters or less per cycle).

We use our spectral analysis as a guide for specifying the windows for band-pass filters. Again, we find that there is a negative and stable Phillips curve at the business cycle frequency. We conclude that short-run fluctuations and long-term trend components of the data mask the Phillips curve in the data.

The phase spectrum analysis shows that despite the existence of the Phillips curve at the business cycle frequency, the monetary policy implications are not obvious. In order for monetary policy to be effective in exploiting the trade-off between inflation and unemployment, it would be best for inflation to lead unemployment, so that a policy-induced increase in the inflation rate can lead to lower unemployment. However, we cannot find systematic and significant lead-lag relationships in the quadrature spectrum between unemployment and inflation at frequency intervals that monetary policy could exploit. If anything, unemployment weakly leads inflation, or inflation weakly lags unemployment, at the conventional business cycle frequency. This finding casts doubt on the so-called tradeoff between unemployment and inflation even if the Phillips curve is negatively sloped.

We also use the generalized structural identification method of Wen (2001, 2002) in the frequency domain to identify two structural shocks: a permanent shock that causes long-run movements in unemployment at frequency zero and a nonpermanent shock that is orthogonal to the permanent shock. We reveal that nonpermanent shocks are responsible for almost the entire negative relationship between unemployment and inflation in the frequency interval $[0.02, 0.17]$, while the positive relationship in the frequency interval $[0, 0.02]$ is entirely due to the permanent shock. We also illustrate the time-series properties of the two shocks and their implied movements in unemployment and inflation in the time domain, using the inverse Fourier transform. We show that in the time domain, permanent shock-implied movements and nonpermanent shock-implied movements in unemployment and inflation move in opposite directions and lead to a mixture that masks any significant and systemic relationships between unemployment and inflation. In addition, when we split the sample into two subsamples, we obtain the same results. So the policy implication is that for monetary policy to be at all effective in exploiting the trade-off between inflation and unemployment, it must be conducted only under nonpermanent shocks and not under permanent shocks (such as the oil shocks). In fact, it may be counterproductive under permanent shocks.

We conclude that the Phillips curve is alive and well, at least at the business cycle frequency and under nonpermanent shocks (which are often interpreted as demand-side shocks). However, short-term changes in the inflation rate will not necessarily lead to immediate changes in the unemployment rate. Instead, it is important to look at the intermediate run. In the long run, inflation leads to high unemployment instead of low unemployment. ■

APPENDIX 1: A BRIEF TECHNICAL REVIEW OF SPECTRAL ANALYSIS

Given a square-summable sequence $\{c_n\}_{n=-\infty}^{\infty}$ $\left(\sum_{n=-\infty}^{\infty} |c_n|^2\right) < \infty$, its Fourier transform is defined as

$$f(\omega) = \sum_{j=-\infty}^{\infty} c_j e^{-i\omega j},$$

where $\omega \in [-\pi, \pi]$ is angular frequency, $i \equiv \sqrt{-1}$ is the imaginary unit, and $e^{\pm ix} = \cos(x) \pm i\sin(x)$ is the Fourier operator. Given the function $f(\omega)$, each element c_k in the original sequence $\{c_n\}_{n=-\infty}^{\infty}$ can be recovered from the inverse Fourier transform:

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\omega) e^{+i\omega k} d\omega.$$

To apply the Fourier technique to time-series analysis, consider a white-noise process ϵ_t with mean zero and variance σ_ϵ^2 . Based on the white-noise process, we can construct a new time series:

$$y_t = \sum_{j=0}^{\infty} b_j \epsilon_{t-j} = B(L) \epsilon_t,$$

where L is the lag operator. The covariogram (autocovariance) at any lag $\tau \in (-\infty, \infty)$ of this new time series is given by

$$c_y(\tau) = E y_t y_{t-\tau} = \sigma_\epsilon^2 \sum_{j=-\infty}^{\infty} b_j b_{j-\tau} \quad (\text{Note: } b_j = 0 \text{ for } j < \tau).$$

Applying the Fourier transform to the sequence $\{c_y(\tau)\}_{\tau=-\infty}^{\infty}$ gives the power spectrum (or spectral density function) of y_t :

$$g_y(e^{-i\omega}) = \sum_{\tau=-\infty}^{\infty} c_y(\tau) e^{-i\omega\tau} = B(e^{-i\omega}) B(e^{i\omega}) \sigma_\epsilon^2.$$

There are several basic properties of the spectral density function over the domain $\omega \in [-\pi, \pi]$: (i) It is real valued and nonnegative; (ii) it is symmetric about $\omega = 0$, or $g_y(e^{-i\omega}) = g_y(e^{i\omega})$; and (iii) a spectral peak (local maximum) at ω_0 reflects relatively large contributions from cyclical fluctuations of around $2\pi/\omega_0$ periods per cycle.

Using the inverse Fourier transform, we can recover the autocovariance at any lead or lag $\tau \in (-\infty, \infty)$:

$$c_y(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g_y(e^{-i\omega}) e^{+i\omega\tau} d\omega$$

$$c_y(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g_y(e^{-i\omega}) d\omega,$$

where $c_y(0)$ is the variance of the original time series. That is, the spectrum $g_y(e^{-i\omega})$ is a moment-generating function and the total area underneath the spectrum is proportional (2π times) to the total variance of y_t . It is in this sense that we view the spectrum as a distribution of variance across cyclical frequencies.

To analyze the dynamic relationships between any two time series, consider the VAR

$$Y_t = \begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} A(L) & B(L) \\ C(L) & D(L) \end{pmatrix} \begin{pmatrix} \epsilon_t \\ \mu_t \end{pmatrix}.$$

Then, the spectrum of the vector Y_t is

$$\begin{aligned} \equiv g(e^{-i\omega}) &= \begin{pmatrix} g_y(e^{-i\omega}) & g_{yx}(e^{-i\omega}) \\ g_{xy}(e^{-i\omega}) & g_x(e^{-i\omega}) \end{pmatrix} \\ &= \begin{pmatrix} A(e^{-i\omega}) & B(e^{-i\omega}) \\ C(e^{-i\omega}) & D(e^{-i\omega}) \end{pmatrix} \begin{pmatrix} \sigma_\epsilon^2 & 0 \\ 0 & \sigma_\mu^2 \end{pmatrix} \begin{pmatrix} A(e^{i\omega}) & B(e^{i\omega}) \\ C(e^{i\omega}) & D(e^{i\omega}) \end{pmatrix}, \end{aligned}$$

where the cross spectrum $g_{yx}(e^{-i\omega}) = g_{xy}(e^{i\omega}) \equiv \overline{g_{xy}(e^{-i\omega})}$ (complex conjugation) is complex-valued in general.

Since

$$g_{yx}(e^{-i\omega}) \equiv \sum_{\tau=-\infty}^{\infty} c_{yx}(\tau) e^{-i\omega\tau},$$

the cross spectrum is a “cross-covariance” generating function; namely, given $g_{yx}(e^{-i\omega})$, we can recover the covariance of y_t and $x_{t-\tau}$ at any lead or lag $\tau \in (-\infty, \infty)$:

$$c_{yx}(\tau) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g_{yx}(e^{-i\omega}) e^{i\omega\tau} d\omega.$$

In particular, we have

$$c_{yx}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g_{yx}(e^{-i\omega}) d\omega.$$

So the covariance of y_t and x_t (at $\tau = 0$) is proportional to the area underneath the cross spectrum.

There are several basic properties of the cross-spectral density function: (i) It is complex-valued; (ii) its real part is symmetric about $\omega = 0$: $g_y(e^{-i\omega}) = g_y(e^{i\omega})$; and (iii) its imaginary part has rotational symmetry about $\omega = 0$. Specifically,

$$\begin{aligned} g_{yx}(e^{-i\omega}) &= g_{yx}(\cos(\omega) - i\sin(\omega)) \equiv \text{Re}_{yx}(\omega) - i\text{Im}_{yx}(\omega) \\ g_{xy}(e^{+i\omega}) &= g_{xy}(\cos(\omega) + i\sin(\omega)) \equiv \text{Re}_{xy}(\omega) + i\text{Im}_{xy}(\omega), \end{aligned}$$

where the real part $\text{Re}(\omega)$ is called the “cospectrum” and the imaginary part $\pm\text{Im}(\omega)$ (*sign included*) is called the “quadrature spectrum.” Since $g_{yx}(e^{-i\omega}) = g_{xy}(e^{i\omega})$, we have (for all ω) that

$$\text{Re}_{yx}(\omega) = \text{Re}_{xy}(\omega) \text{ and } \text{Im}_{yx}(\omega) = -\text{Im}_{xy}(\omega).$$

Hence, given these symmetry properties of the spectrum and cross spectrum, in spectral analysis we need only to consider the domain $\omega \in [0, \pi]$.

Since the covariance is real valued, we also have the following important property:

$$c_{yx}(0) = c_{xy}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \text{Re}(\omega) d\omega \pm \underbrace{\frac{i}{2\pi} \int_{-\pi}^{\pi} \text{Im}(\omega) d\omega}_{=0},$$

where the integral of the quadrature spectrum over $[-\pi, \pi]$ is always zero, which again suggests that the quadrature spectrum has “rotational symmetry” around frequency zero. This also implies (i) that the sign of the cospectrum (the real part) reflects the sign of the covariance between y_t and x_t —it can be either positive or negative—and it reflects the density distribution of the covariance across cyclical frequencies and (ii) that the quadrature spectrum captures the phase differences of cycles in the two series (discussed below).

As in the time domain where we can normalize the covariance of two time series by the product of their respective variances to obtain the correlation, the frequency-domain analog is the coherence function defined by

$$\text{Coh}(e^{-i\omega}) \equiv \sqrt{\frac{|g_{yx}(\omega)|^2}{g_y(\omega)g_x(\omega)}} \in [0, 1],$$

which is essentially the (spectral) distribution of the absolute value of contemporaneous correlations between y_t and x_t across frequencies. Since the cross spectrum is complex-valued, it needs to be multiplied by its conjugate to yield a real-valued function.

Also, since any complex number has a polar form representation, we can also express the cross spectrum in its polar form:

$$g_{yx}(\omega) = r(\omega)e^{i\theta(\omega)},$$

where

$$r(\omega) = \sqrt{\text{Re}^2(\omega) + \text{Im}^2(\omega)} \text{ and } \theta(\omega) = \tan^{-1} \left[\frac{\text{qu}(\omega)}{\text{co}(\omega)} \right] (+180^\circ \text{ if } \text{co} < 0).$$

The function $r(\omega)$ is called the gain, and the function $\theta(\omega)$ is called the phase, which has a maximum of $\pi/2$ and a minimum of $-\pi/2$.

The meanings of the gain and the phase can be illustrated in the following simple example: Let $Y_t = x_{t-k} = L^k x_t$, where L^k is the lag operator to the k^{th} power. Then the cross spectrum is given by

$$\begin{aligned}
g_{yx}(e^{-i\omega}) &= e^{-i\omega k} g_x(e^{-i\omega}) \\
&= [\cos(\omega k) + i(-\sin(\omega k))] g_x(e^{-i\omega}) \\
&= r(\omega) e^{i\theta(\omega)} g_x(e^{-i\omega}),
\end{aligned}$$

where

$$r(\omega) = \sqrt{\cos^2(\omega k) + \sin^2(\omega k)} = 1 \text{ and } \theta(\omega) = \tan^{-1} \left[\frac{-\sin(\omega k)}{\cos(\omega k)} \right] = -k\omega.$$

So, shifting the original time series x_t by k units of time does not change the amplitude of the cyclical components of the original time series in the frequency domain at any frequency. Therefore, the gain function is 1 across all frequencies. However, there exist serious phase effects. Since y_t can be predicted by x_{t-k} k units ahead of time, thus it lags x_t by k units of time, suggesting that the events in x_t are delayed in y_t or that the cyclical phase of fluctuations in y_t is simply the same cyclical fluctuations in x_t shifted backward in time. Thus, it can be predicted by the past history of x_t . This phase effect is captured by the phase function $\theta(\omega)$. y_t leads (lags) x_t if the phase $\theta(\omega)$ is negative (positive) at frequency ω .

More specifically, suppose $x_t = \cos(\omega t)$, which peaks at $t = 0$. Then

$$y_t = x_{t-k} = \cos(\omega(t-k)) \equiv \cos(\omega t + \theta),$$

where $\theta = -\omega k$, so y_t peaks at $t = k = -\frac{\theta(\omega)}{\omega}$. Thus, output y_t lags input x_t by k units of time.

Since the function $\tan^{-1}(x)$ fluctuates in the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ and changes sign suddenly whenever x changes sign, it is not smooth whenever the phase shift is larger than $\frac{\pi}{2}$; yet, because the full phase of a cycle is π or twice as long, we cannot rely simply on the phase function to gauge the lead-lag relationship. Instead we can use the sign of the quadrature spectrum to determine the lead-lag relationship.

In the above example, suppose $k = 1$. We have

$$\begin{aligned}
g_{yx}(e^{-i\omega}) &= e^{-i\omega} g_x(e^{-i\omega}) \\
&= [\cos(\omega) + i(-\sin(\omega))] g_x(e^{-i\omega}) \\
&= g_x(\omega) \cos(\omega) - i g_x(\omega) \sin(\omega) \\
&= co(\omega) + iqu(\omega),
\end{aligned}$$

where the power spectrum $g_x(\omega) > 0$ and the function $\sin(\omega) > 0$ for $\omega \in (0, \pi)$. Thus it must be true that $qu(\omega) < 0$ over the interval $\omega \in (0, \pi)$. Since by design y_t lags x_t by 1 unit of time, it suggests that the sign of the quadrature spectrum $qu(\omega) = -\sin(\omega)$ can also reveal the lead-lag relationship: y lags x_t if $qu(\omega) < 0$. In contrast, if $k = -1$, then the quadrature spectrum would be positive, so y leads x_t if $qu(\omega) > 0$.

In contrast, the phase spectrum shows that

$$\theta(\omega) = \tan^{-1} \left[\frac{-\sin(\omega)}{\cos(\omega)} \right] = \begin{cases} -\omega & \text{if } \omega \in [0, \pi/2] \\ \omega & \text{if } \omega \in [\pi/2, \pi] \end{cases},$$

where we know that when $\omega \in \left(\frac{\pi}{2}, \pi\right)$, $\cos(\theta) < 0$, so $\theta(\omega) > 0$; thus, the phase spectrum

suggests that y_t leads x_t once $\omega > \frac{\pi}{2}$, which cannot be right since by design y_t lags x_t . Hence, caution must be exercised when using the phase function to gauge lead-lag relations. The reason is as follows: For pure sin waves, a lead can also be interpreted as a lag if the lead is too big (larger than half the length of the full cycle or 180° in phase angle). Since the phase function switches sign for every 90° or $\frac{\pi}{2}$, it is better to use the sign of the quadrature spectrum to gauge the lead-lag relationship.

APPENDIX 2: UNIT ROOT TEST

Table A2.1
Unit Root Test

	<i>t</i> -Statistic	<i>p</i> -Value
Inflation rate	−3.683	0.004
Unemployment rate	−2.936	0.041

The augmented Dickey-Fuller test fits a model of the form

$$\Delta y_t = \alpha + \beta y_{t-1} + \sum_{j=1}^p \phi_j \Delta y_{t-j} + \epsilon_t,$$

where y is the variable under consideration and p is the maximum number of lags (we set $p = 4$). The null hypothesis is $\beta = 0$, or, equivalently, that y_t follows a unit process. From Table A2.1, we reject the null hypothesis at the 1 percent level for the inflation rate and at the 5 percent level for the unemployment rate. Therefore, we can conclude that the inflation rate and the unemployment rate are stationary.

NOTES

- ¹ Phillips's original analysis was based on the wage inflation rate, but later research replaced the wage inflation rate simply with the general price inflation rate.
- ² We define the conventional business cycle frequency as fluctuations between 6 and 40 quarters per cycle.
- ³ We find that four lags in the VAR are optimal, and our results are robust to other lag specifications.
- ⁴ We use the Baxter-King (1999) approximation to the band-pass filter. Baxter and King (1999) derived their method for nonstationary time series by imposing that the filter coefficients sum to zero. Dropping this condition yields a filter for stationary time series. We assume that the inflation rate and unemployment rate are stationary, and we verify this assumption by running Dickey-Fuller tests. See the appendix.

REFERENCES

- Baxter, Marianne and King, Robert G. "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series." *Review of Economics and Statistics*, 1999, 81(4), pp. 575-93; <https://doi.org/10.1162/003465399558454>.
- Blanchard, Olivier Jean and Quah, Danny. "The Dynamic Effects of Aggregate Demand and Supply Disturbances." *American Economic Review*, 1989, 79(4) pp. 655-73; <https://www.jstor.org/stable/1827924>.
- Gallegati, Marco; Gallegati, Mauro; Ramsey, James Bernard and Semmler, Willi. "The US Wage Phillips Curve across Frequencies and over Time." *Oxford Bulletin of Economics and Statistics*, 2011, 73(4), pp. 489-508; <https://doi.org/10.1111/j.1468-0084.2010.00624.x>.
- Gordon, Robert J. "The History of the Phillips Curve: Consensus and Bifurcation." *Economica*, 2011, 78, pp. 10-50; <https://doi.org/10.1111/j.1468-0335.2009.00815.x>.
- Iacobucci, Alessandra. "Spectral Analysis for Economic Time Series," in Jacek Leskow, Lionello F. Punzo, and Martin Puchet Anyul, eds., *New Tools for Economic Dynamics*. Lecture Notes in Economics and Mathematical Systems. Volume 551. Springer, 2005, pp. 203-19; https://doi.org/10.1007/3-540-28444-3_12.
- King, Robert G. and Watson, Mark W. "The Post-War U.S. Phillips Curve: A Revisionist Econometric History." *Carnegie-Rochester Conference Series on Public Policy*, 1994, 41, pp. 157-219; [https://doi.org/10.1016/0167-2231\(94\)00018-2](https://doi.org/10.1016/0167-2231(94)00018-2).
- Phillips, A.W. "The Relation Between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957." *Economica*, 1958, 25(100), pp. 283-99; <https://doi.org/10.2307/2550759>.
- Samuelson, Paul A. and Solow, Robert M. "Analytical Aspects of Anti-Inflation Policy." *American Economic Review*, 1960, 50(2), pp. 177-94.
- Scully, Gerald W. "Static vs. Dynamic Phillips Curves." *Review of Economics and Statistics*, 1974, 56(3), pp. 387-90; <https://doi.org/10.2307/1923978>.
- Wen, Yi. "A Generalized Method of Impulse Identification." *Economics Letters*, 2001, 73(3), pp. 367-74; [https://doi.org/10.1016/S0165-1765\(01\)00509-2](https://doi.org/10.1016/S0165-1765(01)00509-2).
- Wen, Yi. "The Business Cycle Effects of Christmas." *Journal of Monetary Economics*, 2002, 49(6), pp. 1289-314; [https://doi.org/10.1016/S0304-3932\(02\)00150-2](https://doi.org/10.1016/S0304-3932(02)00150-2).

Schools and Stimulus

Bill Dupor and M. Saif Mehkari

This article analyzes the impact of the education funding component of the American Recovery and Reinvestment Act of 2009 (Recovery Act) on public school districts. We use cross-sectional differences in district-level Recovery Act funding to investigate the program's impact on staffing, expenditures, and debt accumulation. To achieve identification, we use exogenous variation across districts in the allocations of Recovery Act funds for students with special needs. We estimate that \$1 million in grants to a district had the following average effects: Expenditures increased by \$570,000, employment changed little to none, and debt increased by \$370,000. Moreover, 70 percent of the increase in expenditures was in the form of capital outlays. Next, we build a dynamic, decision-theoretic model of a school district's budgeting problem, which we calibrate to district-level expenditures and staffing data. The model can qualitatively match the employment and capital expenditure responses from our regressions. We also use the model to conduct policy experiments. (JEL D21, D24, E52, E62)

Federal Reserve Bank of St. Louis *Review*, Second Quarter 2020, 102(2), pp. 145-71.
<https://doi.org/10.20955/r.102.145-71>

1 INTRODUCTION

The American Recovery and Reinvestment Act of 2009 (the Recovery Act) was signed into law with a primary goal of creating and saving millions of jobs during and following the 2007-09 recession. A large share of the appropriations from the act consisted of grants. Public school districts constituted one of the largest groups of these recipients, receiving \$64.7 billion in Department of Education Recovery Act funds.^{1,2}

The act's education component has been touted as one of the success stories by the law's supporters. Shortly after its passage, Vice President Joe Biden stated that funds from the act would "help to keep outstanding teachers in America's schools."³ According to the Executive Office of the President of the United States (2009, p. 4), "the rapid distribution of SFSF [State Fiscal Stabilization Funds] funding helped fill the gaps and avert layoffs of essential personnel

Bill Dupor is an assistant vice president and economist at the Federal Reserve Bank of St. Louis. M. Saif Mehkari is an associate professor of economics at the University of Richmond. The authors thank Peter McCrory for helpful research assistance.

© 2020, Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

Figure 1**Education and Noneducation Employment Growth in the United States, 1991-2016**

NOTE: Shaded bars indicate recessions as determined by the National Bureau of Economic Research. Growth rates are demeaned.

SOURCE: Bureau of Labor Statistics.

in school districts and universities across the nation.” The act’s official website, Recovery.gov, used surveys of recipient organizations to track the number of jobs payrolled by the act’s funds. The Council of Economic Advisers (various quarterly reports) used the job-count data from these surveys as evidence of the act’s success.⁴ According to these reports, Department of Education Recovery Act dollars *alone* directly created and saved over 750,000 jobs during the first two school years following the act’s passage.⁵ For context, Figure 1 plots the demeaned growth rates of education employment and noneducation, that is, total net of education employment between 1991 and 2016. Noneducation employment is the much more cyclical of the two.

This article analyzes the act’s impact on schools, using cross-sectional differences in district-level Recovery Act grants and expenditures, staffing, and debt accumulation. We compare the behavior of districts receiving relatively little grant money with that of districts receiving generous grants. From this comparison, we infer what districts would have done without the grants.

To address the potential endogeneity of spending, we employ two instruments. Our first instrument is the ratio of the number of special-needs students to the total number of students in each district. Our second instrument is the Recovery Act dollars received by a district through the act’s Special Education Fund (SEF). The SEF was one category of the Recovery Act

education component and constituted one-fifth of the education grants. Its allocation across districts was determined primarily by the requirement that districts finance their special-needs programs. Although each instrument is highly correlated with overall Recovery Act education spending, each is plausibly uncorrelated with the short-run business cycle and tax revenue situations faced by school districts.

We have four main findings. First, the grants had either no or only a small impact on education jobs. Each \$1 million of aid to a district resulted in roughly an average of 1.5 additional jobs within that district. The point estimate implies that, in the first two school years following its passage, the act increased education employment by 95,000 persons nationwide. Moreover, this estimate is not statistically different from zero.

We find no evidence that the grants increased the number of classroom teachers. Intuitively, district administrators may have shown a strong preference for maintaining teacher-to-student (teacher-student) ratios and, to a lesser extent, staff-to-student (staff-student) ratios. As such, school officials may have found margins other than firing or hiring with which to cover shortfalls or spend surpluses.

Second, each \$1 million of grants to a district increased its expenditures by \$570,000. Because districts already had substantial funds from local and state sources, the additional Recovery Act funds were effectively fungible. Thus, upon receipt of Recovery Act funds, state and local funding sources may have reduced their own contributions to district funding, thereby offsetting the act's grants.

Third, districts that received grants tended to accumulate more debt. Roughly 70 percent of the spending increases were capital expenditures, that is, spending on construction, land, or existing structures (CLS) and equipment. Why might districts have used these funds for capital improvement? Since this aid was temporary, school districts may have smoothed the benefits of the aid over time by making long-lived physical investments.

Fourth, we build and calibrate a model of dynamic decisionmaking by a forward-looking school district. We show that the small employment effect and relatively large investment effect fall out of a fully specified and realistic dynamic programming problem.

We also use our theoretical model as a laboratory to understand the effects of different types of policy. Our main finding is that forcing school districts to use all the stimulus money on labor has no additional effect on the employment outcome. School districts that are forced to use stimulus money only on employment reduce their labor spending from state and local funding sources and substitute this shortfall with stimulus money, leaving the net employment outcome unchanged. We show that an alternative policy requiring districts to spend most of their revenue (from both stimulus and state and local sources) has a more significant effect on employment.

With respect to existing work, there is little economic research on the act's education component. Two exceptions are Dinerstein et al. (2013), who study the impact of the act on universities, and Chakrabart and Setren (2011), who examine the impact of the recession and the early part of the Recovery Act on school districts in the state of New York. More generally, other studies using microeconomic evidence that study the overall impact of the Recovery Act have focused mainly on economy-wide labor market outcomes. These include Chodorow-

Reich et al. (2012); Conley and Dupor (2013); Dupor and McCrory (2018); Feyrer and Sacerdote (2012); and Wilson (2012).

Another line of research studies how federal grants to schools influence school spending. Gordon (2004) studies the impact of additional federal grants to school districts serving economically disadvantaged children through the No Child Left Behind Act of 2001. She finds that, although the additional federal grants initially caused a dollar-for-dollar increase in school spending, over time school districts offset those increases with reductions in their own contributions to education funding.

Lundqvist, Dahlberg, and Mork (2014) study the impact of intergovernmental grants to local governments in Sweden and find that the grants do not stimulate local public employment. Evans and Owens (2007) study the extent to which federal grants to fund new police hires increased the size of local police forces versus simply supplanting local funding. They found that for every four officers payrolled by a grant, in an accounting sense, a police force actually increased by only a little over two officers.⁶

2 EMPIRICAL ANALYSIS

2.1 The Data

The Sample. Our unit of observation is a public school district.⁷ During the 2010 school year (SY), there were 16,117 such districts in the United States. We restrict our attention to districts with more than 500 students during that year. We also exclude districts missing requisite data, which leaves 6,786 districts.⁸

Outcome Variables (Δ Job-Years, Δ Expenditures, and Debt accumulation). Our first outcome variable measures school district employment. It is the change in employment from a base of the 2007 SY over the first two school years in which the act was fully in effect: that is, the 2009 SY and the 2010 SY.⁹ Employed persons include teachers, aides, guidance counselors, librarians, district administrators, and other support staff. The data are self-reported by school districts in the annual Common Core of Data Local Education Agency Universe Survey.

Let $Y_{j,k}$ denote employment by district j during school year k . Then,

$$\Delta\text{Job-years}_j = \frac{1}{\text{Pop}_j} \sum_{k=2009}^{2010} (Y_{j,k} - Y_{j,2007}),$$

where Pop_j is district- j enrollment in the 2007 SY.

Our data on total expenditures $S_{j,t}$ and debt are from the annual Local Education Agency Finance Survey. From these variables, we calculate our next two outcome variables. We measure expenditures as the per student cumulative spending in the 2009 SY and 2010 SY relative to a pre-act baseline:

$$\Delta\text{Expenditures}_j = \frac{1}{\text{Pop}_j} \sum_{k=2009}^{2010} (S_{j,k} - S_{j,2007}).$$

Debt accumulation is the change in the per student debt of a district over the two school years following the act's passage:

$$\text{Debt accum}_j = \frac{1}{\text{Pop}_j} (\text{End of 2010SY Debt}_j - \text{End of 2008SY Debt}_j).$$

Treatment Variable (V). First, let $\tilde{V}_j$ be the Recovery Act dollars outlaid to school district j from the time of enactment through 2011:Q2.¹⁰ Outlaid dollars are defined as dollars paid by the federal government to a recipient organization. These amounts are constructed using quarterly reports filed by recipients on the website FederalReporting.gov.¹¹ Finally, we scale by the district enrollment and report values in millions of dollars:

$$V_j = \frac{\tilde{V}_j}{10^6 \times \text{Pop}_j}.$$

Nearly all of the education dollars authorized by the act were outlaid by the end of 2011:Q2.

Instrument Variables (V^{SN} and V^{SEF}). Since the allocation of the act's school funding was perhaps in part endogenous, we use instrumental variables (IVs). We have two IVs. Our first IV is the ratio of the number of special-needs students within a district to the overall student enrollment in that district in 2007.¹² Denote this variable as V_j^{SN} . While the fraction of special-needs students in a school district is likely to affect the Recovery Act funding that a district receives, it is plausibly uncorrelated with the business cycle conditions and tax revenue stress that the district faced.

Our second IV is the per student value of special-education funding outlaid as part of the Recovery Act, defined as V_j^{SEF} , through 2011:Q2.

The main channel through which the federal government supports special education is through the Individual's with Disabilities Act (IDEA). Most of the Recovery Act special-education money was tied to the IDEA program. While there are several subprograms within IDEA, the lion's share of funding comes through Part B of IDEA. The Recovery Act funding formula follows the IDEA Part B formula.¹³

Recovery Act IDEA Part B grants were add-ons to regular annual IDEA Part B grants to states. The national federal fiscal year (FFY) 2009 regular grant amount was \$11.5 billion. The first \$3.1 billion (from both regular funding and the Recovery Act add-on) was divided among states so they were guaranteed to receive their FFY 1999 awards. Once this requirement was met, the remaining part of the national award was allocated among the states according to the following rule: "85% are allocated to States on the basis of their relative populations of children aged 3 through 21 who are the same age as children with disabilities for whom the State ensures the availability of a free appropriate public education and 15% on the relative populations of children of those ages who are living in poverty."¹⁴ The Recovery Act add-on totaled \$11.3 billion. Since, at the margin, the FFY 1999 requirements had already been met by the regular awards, every Recovery Act dollar was in effect assigned according to the 85/15 percent rule.

We now address how funds were assigned from state education agencies to local education agencies (LEAs). These initial allocations too were made at the federal level. Each LEA was

first allocated a minimum of its FFY 1999 award.¹⁵ Beyond these minimums, which were already met by the regular annual award amounts, a slightly different 85/15 rule was used. Within each state, 85 percent of the dollars were allocated according to the share of school-age children in the LEA and 15 percent were allocated according to the share of those children living in poverty. After meeting these stipulations, states were allowed to reallocate funds as explained below. Before we explain how reallocations worked, we ask whether the observed spending data at the within-state level are explained by the simple formulary rule.

Let $P_{j,s}$ and $\tilde{P}_{j,s}$ be the enrollment of students and students in poverty, respectively, in district j and state s . Let $IDEA_{j,s}$ denote the total Recovery Act special-needs funding in district j in state s .¹⁶

Thus, within each state, the district-level per pupil IDEA amount would be perfectly predicted by the ratio of the enrollment of low-income students to the overall enrollment in the district if the simple formula were used. Next, we run state-level regressions to check this conjecture for the 46 states for which we have fully reported IDEA amounts. The R^2 values from these regressions are generally very low: 25 values are less than 0.01. Only six of the R^2 values are greater than 0.1 and only one is greater than 0.3.¹⁷ This tells us that factors other than the poverty rate in each district are influencing the allocation of IDEA funds.

This brings us to the rules for reallocation of dollars within states across LEAs, given by Code of Federal Regulation 300.707(c)(1). It states,

If an SEA [state education agency] determines that an LEA is adequately providing FAPE [free appropriate public education] to all children with disabilities residing in the area served by that agency with State and local funds, the SEA may reallocate any portion of the funds under this part...to other LEAs in the State that [are] not adequately providing special-education and related services to all children with disabilities residing in the area served by those LEAs.

Based on the legislation and given the low set of R^2 values above, we conclude that the primary reason that IDEA money was allocated differently from the formulary rule is that some states were able to meet their funding requirements for special-needs students in some districts without drawing on Recovery Act IDEA funds. These funds were then reallocated to districts with additional funding requirements for special-needs students. Differences in funding requirements across districts were likely due to factors such as the number of special-needs students, the types of disabilities and their associated costs, and the districts' own funding contributions for providing services to these special-needs students. Our exogeneity assumption is that this set of factors driving redistribution of IDEA funds is orthogonal to the error term in the second-stage equation.

Conditioning Variables (X). We include the following conditioning variables, which we partition into three types:

- *pre-recession education variables*: the 2007 SY values of the teacher-student ratio, staff-student ratio, and expenditures per pupil and the change in debt per pupil over the 2007 SY;
- *nonfinancial variables*: the ratio of African American plus Hispanic enrollment to total enrollment, the natural log of enrollment, seven regional dummy variables, and a constant; and

Table 1**Summary Statistics**

Variable	Mean	SD	10th Percentile	90th Percentile
Change in total revenue (\$pp)	838.85	3,186.31	−1,778.23	3,635.67
Change in expenditures (\$pp)	689.81	5,140.74	−3,492.91	4,976.10
Recovery Act education spending (\$pp)	1,013.20	766.98	446.04	1,569.25
Recovery Act IDEA spending (\$pp)	178.48	480.00	0.00	288.82
Change in the wage bill (\$pp)	642.68	1,397.71	−926.05	2,256.92
Change in the number of job-years (\$pp)	0.00	0.03	−0.03	0.02
Debt accumulation (\$pp)	59.69	7,443.66	−2,381.30	2,984.66
Log of enrollment	7.83	1.09	6.55	9.32
2007 SY values of				
Number of teachers (pp)	0.06	0.01	0.05	0.08
Number of staff (pp)	0.12	0.03	0.08	0.16
End of school year debt (\$pp) [†]	10.88	2.99	8.23	14.40
One-year debt change (\$pp)	3,653.55	3,0046.28	−3,000.00	9,662.00
Minority rate	0.24	0.27	0.02	0.69
Poverty rate	0.03	0.02	0.01	0.05
Self-sufficiency ratio	0.41	0.20	0.19	0.71
Total Recovery Act education spending = \$36 billion				
Total Recovery Act IDEA spending = \$7 billion				
Number of observations = 6,786				
NOTE: The unit of observation is a U.S. school district. The above sample excludes districts with enrollments less than 500 in the 2010 SY. [†] denotes a variable has been divided by 1,000. SD, standard deviation; pp, per pupil.				

- *school district financials*: the poverty rate, the fraction of revenue from local sources, and the cumulative change in revenue from nonfederal sources.

Details regarding a few of these variables are in order. The poverty rate is the number of young persons living in poverty relative to the total population of persons living within each school district's borders. The change in revenue from the nonfederal sources variable is given by

$$\frac{1}{\text{Pop}_j} \sum_{k=2008}^{2010} \left(R_{j,k}^{\text{non fed}} - R_{j,2007}^{\text{non fed}} \right),$$

where $R_{j,k}^{\text{non fed}}$ is the district- j revenue from nonfederal sources in school year k . The primary nonfederal sources are from within the district and the state government.

Table 1 provides summary statistics for the variables in our analysis.

2.2 The Econometric Model

We use two-stage least-squares regression for our estimations. The statistical model for the $\Delta\text{Job-years}$ equation is

$$(2.1) \quad \begin{aligned} V_j &= \theta_1 V_j^{SEF} + \theta_2 V_j^{SN} + \psi X_j + v_j \\ \Delta\text{Job-years}_j &= \beta_{JY} \hat{V}_j + \gamma X_j + \varepsilon_j, \end{aligned}$$

where $\hat{V}_j$ are the fitted values from the first-stage regression. The parameter of interest is β_{JY} . The statistical model for the other two outcome variables simply replaces $\Delta\text{Job-years}_j$ with $\Delta\text{Expenditures}_j$ or Debt Accum_j . Our estimates are weighted by district enrollment, and we report robust standard errors (SEs).

3 RESULTS

3.1 Benchmark Results

The Employment Effect. Table 2 contains our benchmark estimates. Column (i) shows the job-years response to grants. The coefficient on education spending equals 1.47 (SE = 1.32): Every \$1 million in grants increased district employment by 1.47 relative to a no-Recovery Act baseline. Note that our construction of the outcome variable is such that one job should be interpreted as lasting one year. This estimate is not statistically different from zero, but is estimated sufficiently precisely to conclude that the jobs effect was small at best. At the upper end of the 95 percent confidence interval, the employment effect was 4.05 jobs per \$1 million spent. We view this as quantitatively small, bearing in mind that the average education industry wage was roughly \$50,000 during this period.¹⁸ The estimates for other outcome variables (presented below), elucidate two reasons why there was a small effect, if any, on education jobs. First, a large portion of the grants did not translate into greater district-level expenditures. Second, district-level expenditures that did arise from the grants were used mainly for capital expenditures.

Next, using the job-years response estimate, we calculate the implied total number of education job-years resulting from the act's education component. Taken at the upper end of its 95 percent confidence interval, our estimate is 260,000 jobs.¹⁹ As explained in the introduction, this is substantially lower than the corresponding number based on the payroll count data reported at Recovery.gov.²⁰

The bottom rows of Table 2 report key statistics from the first-stage regressions. The first-stage results indicate that we have two strong instruments. The partial F -statistic is 2,589.38, with a pointwise t -statistic of 11.72 for the special-education student-ratio instrument and 70.54 for the Special Education Funds instrument.

Our findings related to the jobs effect raise this question: Why were so few, if any, education jobs created as a result of the act? One possibility is that district administrators viewed their staff, particularly teachers, as so important to their mission that districts receiving relatively little aid found ways to close budget gaps without firing many staff. Also, districts that received relatively generous Recovery Act grants may have been less willing to hire new staff

Table 2

Estimates of the Impact on Staff Employment, Expenditures, and Debt Accumulation in \$1 Million of Recovery Act Education Grants, Benchmark Results

	Δ Staff jobs (i)	Δ Total expenditures (ii)	Debt accumulation (iii)
Recovery Act education	3.09*** (0.66)	516.35*** (103.66)	–254.38 (162.75)
<i>Ln</i> (population)	0.07 (0.22)	166.48*** (33.86)	81.29 (53.16)
Minority ratio	–0.02*** (0.00)	1.54*** (0.21)	–0.93*** (0.33)
Poverty rate	–0.02 (0.02)	–16.84*** (3.50)	–1.31 (5.49)
Nonfederal spending change per pupil	1.64*** (0.10)	642.90*** (15.60)	–114.60*** (24.49)
Self-supporting school district	–0.01 (0.00)	–1.65*** (0.28)	1.01** (0.44)
Teachers per pupil, lag	–0.30*** (0.04)	57.13*** (6.03)	–30.30*** (9.46)
Staff per pupil, lag	–0.35*** (0.01)	7.76*** (2.22)	10.57*** (3.49)
Total expenditures per pupil, lag	0.00*** (0.00)	0.00** (0.00)	0.00 (0.00)
Debt change per pupil, lag	0.00*** (0.00)	0.00** (0.00)	0.00** (0.00)
Region dummies	Yes	Yes	Yes
No. of observations	7,519	7,519	0.03
R^2	0.30	0.37	0.69
First-stage results			
Special ed. ratio (<i>t</i> -statistic)	11.72	11.72	11.72
IDEA Recovery Act aid (<i>t</i> -statistic)	70.54	70.54	70.54
Partial <i>F</i> -stat	2,589.38	2,589.38	2,589.38

NOTE: Each estimation also includes additional conditioning variables described in the text. The regressions are enrollment weighted. SEs are in parentheses. *** denotes 1 percent and ** 5 percent statistical significance. The expenditure and debt accumulation variables are in units of thousands of dollars.

for risk that, once the short-lived grants were spent, the new staff would need to be laid off. Adjusting the capital outlays was an alternative way to spend grant dollars. We provide empirical evidence of and theoretical justification for a capital outlay response later in the article.

If neither districts receiving large grants nor those receiving small grants significantly adjusted their staff levels in response to the shock, then we should expect our IV estimates to reflect a small jobs effect. The absence of significant changes in staffing levels is consistent with narrative descriptions of districts' responses to the most recent recession. Cavanaugh (2011) explains that school officials initially responded to budget stress caused by the recession "at the periphery," for example, by cutting travel, delaying equipment upgrades, or scaling back extracurricular activities and art and music programs. Other evidence, based on surveys of school administrators, comes from the American Association of School Administrators (2012), which lists many ways that school administrators filled budget gaps during the period without firing employees. These include furloughing personnel, eliminating or delaying instructional improvement initiatives, deferring textbook purchases, and reducing high-cost course

offerings. While each of these strategies may have marginally reduced the quality of education services provided by the schools, the changes did not directly affect the total number of district employees.

Note that if jobs were created outside the district, perhaps because of a “Keynesian multiplier” effect, this is not reflected in our estimates, because we examine only school district employment.

The Expenditure Effect. Column (ii) of Table 2 reports estimates for the Δ Expenditure specification. The point estimate on Recovery Act education spending equals \$516 (SE = \$103). This implies that \$1 million in education grants resulted in an increase in expenditures of approximately \$516,000 over the first two full school years following the act’s passage. Thus, only about half of the aid to a district actually translated into more expenditures in that district. One explanation for this result may be that there was substantial “crowding out” of contributions by local and state governments to public education when school districts received Recovery Act dollars.

This finding relates to previous research on whether federal grants crowd out state and local spending. In a simple political economy model, Bradford and Oates (1971) show conditions under which crowding out occurs.

The Debt Accumulation Effect. Column (iii) of Table 2 presents the results with debt accumulation per pupil over the two years following the act’s passage as the outcome variable. The point estimate on the Recovery Act spending variable is –\$254 (SE = \$162). Based on the point estimate, districts that received relatively more aid tended to increase their debt positions.

3.2 Additional Results

Table 3 gives the responses of the outcome variables for several variations on the benchmark specification. Panels A and B provide the weighted and unweighted specifications, respectively. The first rows contain the benchmark estimates.

Column (i) of Table 3 presents the job-years estimates for all of the alternative specifications. The majority of estimates are close to the benchmarks. There are three things worth noting. First, not weighting by enrollment has very little effect on the estimate. Second, the “Ordinary least squares” (OLS) row is identical to the benchmark specification except we estimate via OLS rather than IVs. We expect that spending would have an expansionary effect on our outcome variables and money would be targeted toward more severely affected districts. Thus, the direction of endogeneity would suggest the OLS estimates are downwardly biased relative to our IV results. For job-years, the relative value of the point estimates are in the opposite direction; however, they are both quantitatively small and not statistically different from each other. Third, instrumenting with only the special-education ratio generates substantial increases in the jobs and expenditure effects relative to the benchmark specification. The job-years estimate increases to 8.04 (SE = 7.36). Note that we are unable to reject a zero jobs effect for this specification. This specification results in the strongest jobs and expenditure effects of all of the alternative estimated models. Interestingly, the large jobs and expenditure effects are diminished substantially in the corresponding unweighted estimates (see Panel B).

Table 3

Estimates of the Impact On Job-Years, Total Expenditures, and Debt Accumulation of \$1 Million in Recovery Act Education Funding, Alternative Specifications

	Δ Job years (i)	Δ Expenditures (ii)	Debt accumulation (iii)	First-stage partial <i>F</i> -statistics (iv)
A: Weighted by enrollment				
Benchmark	1.47 (1.32)	570.10*** (196.60)	340.63* (185.12)	79
Ordinary least squares	2.11** (1.00)	165.46 (116.72)	30.25 (242.32)	N/A
IDEA instrument only	1.39 (1.32)	524.99*** (190.14)	229.73 (181.77)	100
Special-education ratio instrument only	8.04 (7.36)	2,339.31*** (815.61)	3,977.68*** (1,352.86)	30
Drop region dummies	1.28 (1.21)	621.54*** (209.81)	232.14 (195.77)	89
Drop all lagged variables	0.97 (1.07)	216.97** (105.19)	388.78** (186.16)	78
B: Unweighted results				
Benchmark	0.12 (0.63)	346.08*** (125.01)	461.15*** (164.84)	540
Ordinary least squares	-0.32 (0.37)	60.45 (97.03)	199.82 (132.80)	N/A
IDEA instrument only	0.18 (0.68)	349.87*** (127.50)	401.69 (159.89)	1,027
Special-education ratio instrument only	3.29 (0.24)	772.27 (994.30)	8,515.06*** (2,437.75)	31
Drop region dummies	0.09 (0.53)	323.39*** (125.15)	588.84*** (195.27)	572
Drop all lagged variables	-0.19 (0.24)	145.13** (66.51)	463.07*** (163.62)	528

NOTE: Each estimation includes the conditioning variables described in the text. SEs are in parentheses. *** denotes 1 percent, ** 5 percent, and * 10 percent statistical significance.

Column (ii) of Table 3 presents the total expenditure estimates. (Recall that the coefficient is interpreted as the thousands of dollars by which expenditures increase for a \$1 million Recovery Act education grant to the district.) Thus, if the value is less than \$1,000, then there is some crowding out of grants because part of the aid is not passing through to expenditures. The majority of estimates are close to the benchmark estimate and exhibit substantial crowding out. Comparing the benchmark and OLS estimates, the direction of endogeneity bias is what one would expect.

Column (iii) of Table 3 presents the debt accumulation estimates. The benchmark estimate shows a statistically significant positive effect. All of the alternative specifications have a positive point estimate, with roughly one-half being statistically different from zero. The only outliers are the “Special-education ratio instrument only” cases, both weighted and unweighted. The point estimates for these specifications jump to about \$4.0 million and \$8.5 million, respectively. We view these values as implausibly large. For debt accumulation, the OLS estimate is lower than the IV estimate, which may be somewhat surprising.

Column (iv) of Table 3 contains the partial *F*-statistic for each specification. None of the values indicate a weak instrument problem, although the statistic is dramatically lower for the “Special-education ratio instrument only” specifications.

Table 4**Estimates of the Impact of Staff Employment of \$1 Million Recovery Act Education Grants, By Job Type**

	Δ Teacher JY (i)	Δ Nonteacher JY (ii)
Recovery Act education spending per pupil	-0.03 (0.51)	1.50 (0.99)
Full controls	Yes	Yes
No. of observations	6,786	6,786
Partial <i>F</i> -statistic	78.93	78.93

NOTE: Each estimation includes the conditioning variables described in the text. SEs are in parentheses. JY, job-years.

Table 5**Estimates of the Impact on Expenditures of \$1 Million of Recovery Act Education Funding, by Major Expenditure Categories**

	Δ Expenditures (i)	Δ Capital (ii)	Δ Salaries (iii)	Δ Benefits (iv)
Recovery Act education spending per pupil (\$)	570.10*** (196.60)	390.82*** (149.34)	8.92 (41.23)	79.38* (48.00)
Full controls	Yes	Yes	Yes	Yes
No. of observations	6,786	6,786	6,786	6,786
Partial <i>F</i> -statistic	78.93	78.93	78.93	78.93

NOTE: Each estimation includes the conditioning variables described in the text. The regressions are enrollment weighted. SEs are in parentheses. *** denotes 1 percent and * 10 percent significance. Expenditures and debt accumulation variables are in units of thousands of dollars. The benefits variable is in units of millions of dollars.

Next, we consider the type of education jobs affected. Did the grants create and save teachers' jobs or those of other employees? Table 4 presents the estimates for the benchmark specification except we estimate the equation separately for the changes in the numbers of teaching and nonteaching employees.

Column (i) of Table 4 shows that there was no statistically significant effect on the number of teacher jobs created or saved. The point estimate equals -0.03 ($SE = 0.51$). District administrators may have sought, as a top priority, to maintain class sizes at their pre-recession levels. This constancy may have been achieved by neither hiring nor firing teachers on net.

The employment effect came through nonteaching jobs. As seen in column (ii), each \$1 million resulted in 1.50 ($SE = 0.99$) additional job-years of nonteaching employment, although this too is not statistically different from zero.

Next, Table 5 examines the categories of spending that account for most of the effect on total expenditures. In columns (ii) through (iv), we estimate the benchmark model except we

in turn replace the change in total expenditures with the change in a component of total expenditures.

Column (ii) shows that Recovery Act aid had a substantial effect on capital outlays. Roughly 70 percent of all expenditures came in the form of capital outlays.²¹ Why might districts have used so much of their grant money for investments? First, suppose a district seeks to keep its provision of education services as well as keep those provided services relatively smooth over time, in a manner similar to the permanent income model of consumption smoothing. Second, suppose education services are a function of labor (i.e., the number of staff) and capital. In this case, a district that receives a one-time grant may seek to spread the benefits of this grant over many periods by using part of the grant to increase its capital stock.

Likewise, a district that received a relatively small amount of aid may have found that the best way to close budget gaps was to temporarily cut back on capital investment rather than lay off staff. Because the capital stock depreciates slowly, a temporary interruption in investment would likely have only a small effect on the quality of education services that the school could provide.

Recall that earlier, we document that Recovery Act aid tended to increase debt accumulation. This effect may be related to the positive effect of aid on capital expenditures shown in Table 5. Suppose that, upon receipt of Recovery Act funds, a district decided to spend part of its funds on capital, such as construction. The district may have chosen to boost the dollars available for construction by leveraging up the grant aid by borrowing. Under this scenario, had the district attempted to finance the entire capital project only with debt, it may have been unable to secure the funds or get a reasonable financing rate. Thus, it is possible that grants may have led to borrowing rather than saving by some districts.

Note that the construction spending itself is likely to have a positive jobs effect because of building contractors the district might hire. These numbers are not reflected in our employment estimate, because we restrict attention to school district employees.²²

Column (iii) of Table 5 reports the impact of aid on salaries, which was small and not statistically different from zero.²³ Since the employment effect was so small, it is not surprising that we do not recover a substantial wage effect. Column (iv) of Table 5 implies that \$1 million in aid increased benefits paid by the school district by about \$79 million.

4 A MODEL OF SCHOOL DISTRICT HIRING AND CAPITAL DECISIONS

In this section, we study the dynamic optimization problem of a school district facing stochastic revenue shocks.

In the previous section, we found that the ratio of stimulus spending for paying education workers relative to capital investment was 0.25. This may be puzzling since, as we explain below, the long-run average of this ratio equals 8. Second, there was a small effect on non-teacher staffing and no effect on the number of teachers employed. Our model simulations roughly match both of these findings.

Moreover, our model allows us to estimate the medium- and long-run effects of these grants and provides a laboratory to study the effects of alternative hypothetical stimulus programs aimed at schools.

Table 6**Volatility of the Teacher-Student and Nonteacher-Student Ratios**

	Time-series variance of log deviations from the aggregate trend of the per student ratios			
	Teacher (i)	Nonteacher (ii)	Teacher (iii)	Nonteacher (iv)
10th Percentile	0.0012	0.0036	0.0004	0.0008
25th Percentile	0.0018	0.0062	0.0007	0.0017
Median	0.0033	0.0122	0.0016	0.0038
75th Percentile	0.0062	0.0245	0.0039	0.0093
90th Percentile	0.0104	0.0555	0.0092	0.0232
Years	1994-2011		2006-11	
Number of districts	1,901		4,291	

4.1 The Stylized Facts

We begin by documenting two stylized facts about education spending by analyzing a 17-year panel of district-level data ending with the 2011 SY.²⁴ The facts provide guidance for building and then calibrating our economic model.

Our panel covers a long time span, and some of our series contain time trends. As such, we detrend every variable x_t by its aggregate (over districts) gross growth rate between period t and Q , the final period in our sample. The cumulative growth rate is

$$cg_{x,t} = \frac{\sum_{i \in I} x_{i,t}}{\sum_{i \in I} x_{i,Q}},$$

where I is the set of all districts. The detrended district-level variable is then $\tilde{x}_t$, thus

$$\tilde{x}_{i,t} = \frac{x_{i,t}}{cg_{x,t}}.$$

Unless otherwise noted, each variable is scaled by its district enrollment.

Stylized Fact 1: *The teacher-student ratio is less volatile than the nonteacher-student ratio.*

For each district i , we compute the time-series variance of the log deviation of the employment levels of teachers, T , and nonteachers²⁵, N :

$$v_{x,t} = \text{variance across } t \text{ of } \log \left[\frac{\tilde{x}_{i,t}}{\frac{1}{Q} \sum_t \tilde{x}_{i,t}} \right]$$

for $x \in (T, N)$.

Table 7**Volatility of the Pay-to-Student and Capital-to-Student Ratios**

	Time-series variance of log deviations from the aggregate trend of the per student ratio	
	Salary + benefits (i)	All capital outlays (ii)
10th Percentile	0.0013	0.3339
25th Percentile	0.0021	0.5860
Median	0.0036	0.9580
75th Percentile	0.0064	1.4720
90th Percentile	0.0111	2.0509
Years	1994-2010	
Number of districts	6,092	

Columns (i) and (ii) of Table 6 contain the across-district median values (along with the 10th, 25th, 75th, and 90th percentile values) of $v_{T,i}$ and $v_{N,i}$. Observe that the nonteacher-student ratio is more variable than the teacher-student ratio. The difference in variability ranges from three times as high for the 10th percentile, four times as high for the median, and over five times as high for the 90th percentile.

As further robustness, columns (iii) and (iv) contain the statistics for a smaller subsample that includes data from the most recent six years. Whereas the shorter time horizon results in a reduced value of the magnitude of the variance, as in the full sample, in this subsample, the teacher-student ratio remains less variable than the nonteacher-student ratio.

Stylized Fact 2: *Capital spending is more volatile than labor spending.*

Next, we consider the behavior of two categories of spending: capital expenditures and labor expenditures. Capital expenditures are the sum of spending on CLS and equipment with an expected life of five or more years. Labor expenditures include salaries and benefits of district employees.²⁶ We convert each variable into real terms using the gross domestic product deflator with a base year of 2011.

Table 7 reports the across-district median values (along with the 10th, 25th, 75th, and 90th percentile values) of the time-series volatility of expenditures on total real salary plus benefits and real capital outlays, where each volatility is calculated as the time-series variance of the log deviations of the variable from its aggregate trend using equation (4.1). Note that capital expenditures are significantly more variable than labor expenditures. At the median level of variability, capital expenditures are over 580 times more variable than labor expenditures.

We then divide the capital category into spending of two types: CLS and equipment. Table 8 reports the volatility of these variables. Even though the equipment component itself is volatile, most of the volatility in capital is driven by CLS. This fact, coupled with the facts

Table 8**Volatility of the (Salary + Benefits)-to-Student and the Investment-to-Student Ratios**

	Time-series variance of log deviations from the aggregate trend of the per student ratio		
	CLS (i)	Equipment (ii)	All capital outlays (iii)
10th Percentile	0.5657	0.1141	0.3339
25th Percentile	1.0089	0.1861	0.5860
Median	1.7008	0.3255	0.9580
75th Percentile	2.6626	0.5898	1.4720
90th Percentile	3.9451	1.1045	2.0509
Years	1994-2010		
Number of districts	6,092		

that CLS makes up roughly 80 percent of all capital investment and that labor expenditures are not highly volatile, pushes us toward a theory in which districts (i) tend to use large revenue gains and (ii) make up for revenue shortfalls largely by either investing in, or delaying expenditure on, long-lived capital goods.

4.2 The Economic Model

Consider a school district that uses an exogenous stream of revenue, R , to hire workers and buy capital to provide education services to its students. Its revenue process is given by the following AR(1) process:

$$(4.1) \quad R' = \rho R + (1 - \rho)\bar{R} + \epsilon_R \text{ with } \epsilon_R \sim \mathcal{N}(0, \sigma_R),$$

where $\rho \in (0, 1)$ and $\bar{R}$ is fixed. Revenue, as well as other variables in the model, are per pupil.

A district's one-period welfare function is

$$(4.2) \quad W(T, N, K) = \alpha U(T; \xi_T) + \gamma U(N; \xi_N) + \eta U(K; \xi_K),$$

where T , N , and K are the number of teachers, number of nonteachers, and quantity of capital, respectively. Moreover, let $U(X; \xi) = X^{1-\xi}/(1-\xi)$. The function W depends on inputs that improve "student outcomes," broadly defined.

The district's dynamic optimization problem is given by the following recursive functional equation:

$$V(K; R) = \max_{T, N, I} \{ W(T, N, K) + \beta E[V(K'; R') | R] \}$$

subject to

$$(4.3) \quad R = w_T T + w_N N + I$$

$$(4.4) \quad K' = (1 - \delta)K + I$$

and nonnegativity constraints on T , N , and K . Also, I represents investment in the capital good, and values with a superscript prime give the next-period realization of that variable. For example, K' gives the next-period realization of capital, K .

Next, equation (4.3) is the district budget constraint, with w_T and w_N representing the teacher wage and nonteacher wage, respectively. Also, equation (4.4) is the capital law of motion and δ is the capital depreciation rate.

Every period the school district receives revenue that it optimally allocates to (i) the hiring of teachers and nonteachers and (ii) capital acquisition. Whereas the number of teachers and nonteachers hired affects only the current period's welfare, the durable nature of capital results in a multiperiod effect. As we discuss in the next section, the dynamics that result from allowing the district to choose a durable input are important for understanding why the 2009 Recovery Act had a small effect on hiring but a large effect on capital outlays.

4.3 Calibration and Simulations

Table 9 provides the parameter values for the model. The model period is one year. We begin our calibration by setting the discount factor $\beta = 0.96$ to match a 4 percent annual real interest rate.

Next, in the data, the capital stock is composed of two different basic types: equipment with more than a five-year lifespan and CLS. CLS account for roughly 75 percent of the capital outlays and depreciate at a 1.88 percent annual rate, while equipment accounts for roughly 25 percent of capital outlays and depreciates at a 15 percent annual rate.²⁷ As such, we set $\delta = 0.0516 (= 0.75 \times 0.0133 + 0.25 \times 0.16)$.

Across districts, the median wage bill per student is \$8,128, for which 48 percent goes toward teacher pay and 52 percent goes toward nonteaching staff pay. Thus, teacher compensation is \$3,901 and nonteacher compensation is \$4,227 per pupil. The median teacher-student ratio is 1:15.5, and the median nonteacher-student ratio is 1:16. As a result, teacher and nonteacher wages are set at $w_T = \$60,472 (= 15.5 \times \$3,901)$ and $w_N = \$67,625 (= 16 \times \$4,227)$, respectively.

The persistence of the AR(1) revenue process is directly estimated from the data. The median autocorrelation of expenditures is 0.47. The average revenue is set at $\bar{R} = \$8,128 + \$988 = \$9,116$.

Six parameters remain: The welfare elasticities, ξ_T , ξ_N , and ξ_K ; the relative shares of teachers, α , and nonteachers, γ ; and the standard deviation of the revenue process, σ_R .

First, we set $\xi_K = 1.0$ and then jointly calibrate the remaining five parameters to match the following five targets: The average teacher-student ratio is ; the average nonteacher-student ratio is 0.062; the nonteacher-student ratio is four times as volatile as the teacher-student ratio; the average salary volatility is 0.0036; and the average investment volatility is 0.95.

Table 9**Parameter Values**

Parameter	Value	Description	Explanation
β	0.96	Discount factor	Standard value for annual discount factor
$^*\delta$	0.0516	Depreciation rate	Calculated using the BEA data on the depreciation of buildings and equipment
w_T	\$60,472	Wage rate for teachers	Set equal to avg. teacher wage in the data
w_N	\$67,625	Wage rate for nonteachers	Set equal to avg. non-teacher wage in the data
$\bar{R}$	\$9,116	Average revenue per pupil	Sum of average revenue on labor + capital
$^*\rho$	0.47	Persistence of revenue process	Set equal to the autocorrelation of district-level expenditures in the data
ξ_K	1.0	Welfare elasticity of capital	Normalized to 1
ξ_T	1.52	Welfare elasticity of teachers	Jointly calibrated to match (i) avg. teacher-student ratio = 0.064, (ii) avg. nonteacher-student ratio = 0.062, (iii) teacher-student ratio 4x more volatile than nonteacher-student ratio, (iv) volatility of total salary = 0.0036, and (v) avg. volatility of investment = 0.95
ξ_N	0.76	Welfare elasticity of nonteachers	
α	0.086	Welfare share for teachers	
γ	0.751	Welfare share for nonteachers	
σ_R	1,150	SD of shock to revenue	

NOTE: SD, standard deviation; BEA, Bureau of Economic Analysis; avg., average.

4.4 The Effect of a Recovery Act-Sized Shock

To simulate the effects of the Recovery Act, we alter equation (4.3) to

$$(4.5) \quad R + A = w_T T + w_N N + I,$$

where A denotes the net magnitude of the Recovery Act shock to revenue after accounting for any loss in revenue at the district level. From our benchmark regression analysis (see Table 2), we estimate the size of this shock to be \$570 per student. As a result, we set $A = \$570$ in the period of the shock and $A = \$0$ otherwise. For a transparent comparison with our regression results, all results below are for a \$1 million shock. In the data, the gross magnitude of the Recovery Act shock before accounting for any loss in revenue at the district level was approximately \$1,000 per student. As a result, to find the \$1 million response, we multiply the per student values by 1,000.

Figure 2 plots the effect of the spending shock. The left panels show the per period impulse responses, and the right panels show the cumulative responses. As seen in the figure, over the first two years, the additional revenue creates 1.4 nonteaching staff jobs and 0.7 teaching jobs and increases investment by \$435,000 for each \$1 million spent. Note that other than the size of the shock, the model was calibrated independently of the regression results. Consequently, the consistency between our regression results and the dynamic model provides further evidence for a small effect of the Recovery Act on employment.

Figure 2**Impulse Responses to a Recovery Act-Sized Shock**

The large effect on investment is driven by a motive to smooth the value of education inputs over time. For the purpose of intuition, suppose a school district had two mutually exclusive uses of new funds: (i) increasing the number of staff for one year or (ii) engaging in additional investment for one year. The latter option leads to more capital in both the short and the intermediate runs, which increases education services. Also, since the capital is now higher, the district can cut back marginally on investment in periods after the shock and use the funds saved to increase its staffing levels. The latter option leads to an increased and smoother path of inputs over time—as well as higher welfare.

To illustrate this effect, Figure 2 graphs the responses of the district in a calibration where $\delta = 1.0$; that is, capital depreciates fully after one period. As the figure shows, once the district loses access to interperiod savings, the employment effect increases.

Note that our environment does not permit the district to smooth the benefit of the revenue shock over time using savings or a similar means to deficit reduction. If we were to extend the model to permit these options, districts would use these financial instruments as well as capital accumulation in an optimal policy. Note, however, that our regression results instead find that deficits increased upon receipt of Recovery Act grants. As explained earlier, the increased deficits may be a result of districts pairing new capital spending with increased leverage through higher debt levels.

Next, one of our stylized facts was that the volatility of the number of teachers is significantly lower than that of nonteachers. We conjecture that this occurs because there may be little flexibility in hiring or laying off teachers. Consider a school that teaches five subjects—math, English, Spanish, social studies, and science—to 80 students and currently hires one teacher for each subject. This school may be unable to lay off a teacher because doing so would lead to one fewer subject being taught. If the school wanted to add one teacher, the additional teacher could not teach a bit of all five subjects. Thus, the marginal benefit of hiring, say, one extra math teacher, is very low. On the other hand, hiring nonteaching staff across the district likely would not face classroom indivisibility constraints. The relatively low volatility of teacher employment can be achieved in the model with a high value of ξ_T relative to ξ_N . Thus, $\xi_T > \xi_N$ proxies for relatively low flexibility in changing the teaching-staff level. Figure 2 plots impulse responses if $\xi_T = \xi_N$.²⁸ When the elasticities for teachers and nonteachers are identical, the response between them is more closely aligned.

Our model also permits us to estimate the shock's long-run effects. As discussed earlier, the initial effect of the shock is driven largely by an education-services-smoothing motive that results in accumulating capital initially. This, in turn, frees up future resources for hiring teachers and nonteachers. As shown in Figure 2, the cumulative 10-year effect is approximately two teachers and four nonteachers per \$1 million spent. Note that these effects are larger than the two-year effect. The long-run effect still dwarfs the Council of Economic Adviser's estimate that over 750,000 education jobs were created or saved by the act. At 2.25 jobs per \$1 million in two years and six jobs per \$1 million in 10 years, the \$64.7 billion spent by the Department of Education creates 146,000 jobs in the first two years and 388,000 jobs in the first 10 years following the act's passage.

Policy Analysis. Our model provides a laboratory to study the effects of alternative ways to implement a stimulus program. First, a simple—and it turns out simplistic—policy would require all districts to use stimulus money only on employment; that is,

$$(4.6) \quad A \leq w_T T + w_N N.$$

Figure 3 plots the response to this policy. The policy has no effect relative to the “no constraint” case presented above. This is because a district's existing revenue and the stimulus money are fungible. In response to a stimulus shock, a district can cut back on using its existing revenue to pay labor and instead use the stimulus money to hire workers. The district would meet the requirement of using stimulus money to hire workers and maintain the no-constraint outcome.

Figure 3**Impulse Responses to a Recovery Act-Sized Shock: All Funding Used on Labor**

Consider an alternative policy where, instead, the federal government requires that in the period of the shock at least a fraction ϕ of all revenue must be used to pay workers:

$$(4.7) \quad \phi(R + A) \leq w_T T + w_N N.$$

We simulate the model under this policy, setting $\phi = 0.875$, which we find achieves the maximum employment effect (while keeping investment constant). Figure 4 gives the results of this exercise: There is a significantly larger response, with nine new jobs (three teaching plus six nonteaching) in the year of the shock.

Our model also allows us to consider much richer policy alternatives where the percentage of revenue depends on the amount of revenue and capital at the district level. In Figure 5, we first calculate the pre-stimulus response of the district and then require the district to use all

Figure 4

Impulse Responses to a Recovery Act-Sized Shock: Part of Funding Used on Labor

its stimulus revenue plus all revenue it would have used toward hiring labor had it not received the stimulus revenue. The figure then plots what percentage of the total post-stimulus revenue this amount would have been.

As Figure 5 shows, optimal policy is for the government to impose that a larger percentage of revenue be used on labor in districts with lower revenues and high levels of capital. Districts with lower levels of revenue in particular would be motivated to use the additional stimulus revenue received from the government on capital.

Figure 5**Optimal Policy: Heterogeneous Policy Analysis to Generate Maximum Employment Effects**

5 CONCLUSION

This article explores the impact of countercyclical government spending on the education sector. Empirically, we find that the Recovery Act's education component had a small effect on nonteacher staff levels, no effect on teacher staff levels, and a substantially less than one-for-one response of district-level expenditures. To the extent that government grants increased district expenditures, the increases largely took the form of capital outlays. The grants also stimulated district debt accumulation.

These findings should not be entirely surprising given the decentralized nature of the act's implementation plan. The allocation process was multitiered, with local and state governments allowed latitude as to how Recovery Act dollars were spent. First, state governments maintained substantial control over how they spent their own revenue. This created an environment where stimulus dollars might be used to replace state contributions.²⁹

After passing through the state level, the Recovery Act dollars were spent by individual districts largely at their own discretion. Given that the stimulus dollars were temporary, districts had an incentive to smooth out the spike in additional education services that they could provide by investing in equipment and structures. This objective is one potential explanation for the small employment effect estimated in this article. ■

APPENDIX

Table A.1

Number of Jobs Directly Created and Saved Through Grants, Contracts, and Loans Administered by the U.S. Department Of Education, First Two School Years Following Enactment of the Recovery Act

Quarter	Education jobs
2009:Q3	397,982.43
2009:Q4	423,616.33
2010:Q1	470,197.34
2010:Q2	454,281.08
2010:Q3	344,308.14
2010:Q4	309,187.21
2011:Q1	319,494.26
2011:Q2	307,901.15
Total (annualized)	756,741.99

NOTE: Jobs are measured units of full-time equivalents.

SOURCE: Recovery.gov.

NOTES

- ¹ This includes the Office of Special Education and Rehabilitative Services Special Education Fund (SEF) (\$12.2 billion) and the following Office of Elementary and Secondary Education programs: Education Stabilization funds (\$42.0 billion), Compensatory Education for the Disadvantaged (\$12.4 billion) and the School Improvement Program (\$0.7 billion).
- ² The federal government's objectives for each of the programs were explicit and usually involved, in part, an attempt to stimulate economic activity. For example, recipients were advised that "Among other things, the Education Stabilization funds may be used for activities such as: paying the salaries of administrators, teachers, and support staff; purchasing textbooks, computers, and other equipment," according to U.S. Department of Education (2009a) implementation guidance.
- ³ See Biden (2011).
- ⁴ See also Congressional Budget Office (various quarterly reports).
- ⁵ See Table A.1 in the appendix for a quarterly breakdown of the payroll count data extracted from Recovery.gov. Here, a job is measured as lasting one year and as a "full-time equivalent" of one respective position.
- ⁶ See also Knight (2002) for another example of crowding out of federal grants on state government spending.
- ⁷ Our usage of the term "school district" is synonymous with the term "local education agency" (LEA) used in the education policy area. In the education policy jargon, our sample is made up of school districts and a small number of regional educational service agencies.
- ⁸ For example, we were forced to exclude data from all districts in Iowa, Montana, New Hampshire, Pennsylvania, and Vermont because the Recovery Act spending information was reported in a manner that did not allow us to match the Recovery Act grants to school district spending and employment variables. We also excluded Hawaii because the entire state is a single school district.

⁹ We exclude the 2008 SY because it includes only a few months during which the Recovery Act was in effect.

¹⁰ We use outlays through 2011:Q2 because this aligns our Recovery Act data sample with the end of the 2010 school year.

¹¹ After processing and data verification by the Recovery Accountability and Transparency Board, these data were posted on the website Recovery.gov. A user's guide for these data is contained in Recovery Accountability and Transparency Board (2009).

¹² These data are also from the Common Core of Data Universe Survey. As the data documentation explains, special-needs students are defined as "all students having a written Individualized Education Program (IEP) under the Individuals with Disabilities Act (IDEA), Part B." The IDEA is a comprehensive statute originally passed in 1990 to ensure all students with disabilities are entitled to a free appropriate public education.

¹³ See U.S. Department of Education (2009b) and New America Foundation (2014).

¹⁴ See Enclosure B of U.S. Department of Education (2009b), which contains a description of the allocation of Recovery Act funds across states.

¹⁵ Federal code also describes how minimum awards are determined for LEAs created after 1999.

¹⁶ Based on the above formula, the distribution of Recovery Act IDEA funds would be

$$IDEA_{j,s} = \left(0.85 \times \frac{P_{j,s}}{\sum_{i=1}^{N_s} P_{i,s}} + 0.15 \times \frac{\tilde{P}_{j,s}}{\sum_{i=1}^{N_s} \tilde{P}_{i,s}} \right) IDEA_s.$$

Letting P_s and $\tilde{P}_s$ denote the sum within state of the two district-level enrollment variables, we can rewrite the above equation as

$$\frac{IDEA_{j,s}}{P_{j,s}} = \left[0.85 \times \frac{1}{P_s} + 0.15 \times \frac{1}{\tilde{P}_s} \left(\frac{\tilde{P}_{j,s}}{P_{j,s}} \right) \right] IDEA_s.$$

¹⁷ As an additional measure, we include the poverty rate as an additional control in our estimation.

¹⁸ The mean annual wage for U.S. workers in the "Education, Training and Library" occupation was \$49,530 in 2009.

¹⁹ We calculate this number by multiplying the 95 percent upper bound of the job-years coefficient confidence interval by the cumulative total Recovery Act education spending through the 2010 SY. This calculation assumes that the treatment effect is the same for districts within our sample as for those excluded from the sample.

²⁰ See Table A.1 in the appendix for a tabulation of the Council of Economic Advisers payroll count data.

²¹ Capital outlays include spending on CLS and equipment.

²² Dupor and McCrory (2018) conduct a cross-regional analysis of the act in a broader context than only education. That paper examines employment from all sectors and the act's entire spending component, in contrast to that solely from education. They find a larger jobs effect than that estimated in the current article.

²³ The salary and benefits variables are constructed in the equivalent manner as the variable for total expenditures was constructed.

²⁴ We use the merged Universe and Finance surveys of the Common Core School District dataset. The 1994 SY is the first year for which the entire dataset is available. As in the article's previous section, we drop districts that report fewer than 500 students.

²⁵ Nonteaching staff includes instructional aides, guidance counselors, library/media staff, administrative support staff, and so on.

²⁶ We exclude services and nondurable goods expenditures in our descriptions here. In regression results not provided in the article (but available on request), we establish that there was a negligible effect of grants on these types of spending. We also exclude debt service payments, payments to other districts, and expenditures on nonelementary/secondary programs because they make up only 10 percent of the average district's spending and are outside of our model.

²⁷ See "BEA Depreciation Estimates" at https://apps.bea.gov/iTable/index_FA.cfm.

²⁸ The values of α and γ are jointly determined with ξ_T and ξ_N .

²⁹ As Inman (2010, Abstract) writes, “States are important ‘agents’ for federal macro-policy, but agents with their own needs and objectives.”

REFERENCES

- American Association of School Administrators. “Weathering the Storm: How the Economic Recession Continues to Impact School Districts.” March 2012; https://www.aasa.org/uploadedFiles/Policy_and_Advocacy/files/Weathering_the_Storm_Mar_2012_FINAL.pdf.
- Biden, J. “Roadmap to Recovery.” 2011; <http://www.whitehouse.gov>.
- Bradford, D. and Oates, W. “The Analysis of Revenue Sharing in a New Approach to Collective Fiscal Decisions.” *Quarterly Journal of Economics*, 1971, 85, pp. 416-39; <https://doi.org/10.2307/1885931>.
- Cavanaugh, S. “Education Regroups in Recession’s Aftermath.” *Education Week*, 2011, 30, p. 6.
- Chakrabarti, R. and Setren, E. “The Impact of the Great Recession on School District Finances: Evidence from New York.” Staff Report 534, Federal Reserve Bank of New York, 2011; <https://doi.org/10.2139/ssrn.1979694>.
- Chodorow-Reich, G.; Feiveson, L.; Liscow, Z. and Woolston, W. “Does State Fiscal Relief During Recessions Increase Employment? Evidence from the American Recovery and Reinvestment Act.” *American Economic Journal: Economic Policy*, 2012, 4, pp. 118-45; <https://doi.org/10.1257/pol.4.3.118>.
- Congressional Budget Office. “Estimated Impact of the American Recovery and Reinvestment Act on Employment and Economic Output.” Various quarterly reports; http://www.cbo.gov/sites/default/files/05-25-arr_0.pdf. (Note that each report has a different URL. The URL to one of the reports is provided.)
- Conley, T. and Dupor, B. “The American Recovery and Reinvestment Act: Solely a Government Jobs Program?” *Journal of Monetary Economics*, 2013, 60, pp. 535-49; <https://doi.org/10.1016/j.jmoneco.2013.04.011>.
- Council of Economic Advisers. “The Economic Impact of the American Recovery and Reinvestment Act of 2009.” Various quarterly reports; http://www.whitehouse.gov/sites/default/files/cea_8th_arra_report_final_draft.pdf. (Note that each report has a different URL. The URL to one of the reports is provided.)
- Dinerstein, M.; Hoxby, C.; Meyer, J. and Villanueva, P. “Did the Fiscal Stimulus Work for Universities,” in *How the Financial Crisis and Great Recession Affected Higher Education*. University of Chicago Press, 2013, pp. 263-320.
- Dupor, B. “The 2009 Recovery Act: Directly Created and Saved Jobs Were Primarily in Government.” *Federal Reserve Bank of St. Louis Review*, Second Quarter, 2014, 96(2), pp. 123-45; <https://doi.org/10.20955/r.96.123-146>.
- Dupor, B. and McCrory, P. “A Cup Runneth Over: Fiscal Policy Spillovers from the 2009 Recovery Act.” *Economic Journal*, 2018, 128(611), pp. 1476-508; <https://doi.org/10.1111/eoj.12475>.
- Evans, W. and Owens, E. “COPS and Crime.” *Journal of Public Economics*, 2007, 91, pp. 181-201; <https://doi.org/10.1016/j.jpubeco.2006.05.014>.
- Executive Office of the President of the United States. “Educational Impact of the American Recovery and Reinvestment Act.” Domestic Policy Council in Cooperation with the U.S. Department of Education, October 2009; <https://www.hsdl.org/?abstract&did=32124>.
- Feyrer, J. and Sacerdote, B. “Did the Stimulus Stimulate? Effects of the American Recovery and Reinvestment Act.” Dartmouth College, 2012; <https://doi.org/10.3386/w16759>.
- Gordon, N. “Do Federal Grants Boost School Spending? Evidence from Title I.” *Journal of Public Economics*, 2004, 88, pp. 1771-792; <https://doi.org/10.1016/j.jpubeco.2003.09.002>.
- Gramlich, E. “Stimulating the Macroeconomy through State and Local Governments.” *American Economic Review*, 1979, 68, pp. 180-85.
- Inman, R. “States in Fiscal Distress.” NBER Working Paper 16086, National Bureau of Economic Research, 2010; <https://doi.org/10.3386/w16086>.

- Knight, B. "Endogenous Federal Grants and Crowd-out of State Government Spending: Theory and Evidence from the Federal Highway Aid Program." *American Economic Review*, 2002, 92, pp. 71-91; <https://doi.org/10.1257/000282802760015612>.
- Lundqvist, H.; Dahlberg, M. and Mörk, E. "Stimulating Local Government Employment: Do General Grants Work?" *American Economic Journal: Economic Policy*, 2014, 6, pp. 167-92; <https://doi.org/10.1257/pol.6.1.167>.
- New America Foundation. "Individuals with Disabilities Education Act-Funding Distribution." Federal Education Budget Project, April 2014; <http://febp.newamerica.net/background-analysis/individuals-disabilities-education-act-funding-distribution>.
- Recovery Accountability and Transparency Board. "Recovery.gov: Download Center Users Guide." 2009; <http://www.recovery.gov>.
- U.S. Department of Education. "Guidance on the State Fiscal Stabilization Fund Program." April 2009a; <http://www2.ed.gov/programs/statestabilization/guidance.pdf>.
- U.S. Department of Education. "Guidance: Funds for Part B of the Individuals with Disabilities Act Made Available under the American Recovery and Reinvestment Act of 2009." July 1, 2009b, revise; <http://www2.ed.gov/policy/gen/leg/recovery/guidance/idea-b.pdf>.
- U.S. Department of Education. "Memorandum: The American Recovery and Reinvestment Act of 2009 Individuals with Disabilities Education Act Part B Grants to States and Preschool Grants." Office of Special Education and Rehabilitative Services, April 1, 2009c (including Enclosures A, B and C).
- Wilson, D. "Fiscal Spending Jobs Multipliers: Evidence from the 2009 American Recovery and Reinvestment Act." *American Economic Journal: Economic Policy*, 2012, 4, pp. 251-82; <https://doi.org/10.1257/pol.4.3.251>.

Taking Stock of the Evidence on Microfinancial Interventions

Francisco J. Buera, Joseph P. Kaboski, and [Yongseok Shin](#)

We review the empirical evidence on microfinance and asset grants to the ultra poor or microentrepreneurs and use quantitative economic theory to account for this evidence. Properly executed, these interventions can help segments of the population increase their income and consumption, but neither literature gives much reason to believe that such interventions can lead to wide-scale, transformative impacts akin to escaping aggregate poverty traps. (JEL O16)

Federal Reserve Bank of St. Louis *Review*, Second Quarter 2020, 102(2), pp. 173-202.
<https://doi.org/10.20955/r.102.173-202>

1 INTRODUCTION

Microfinancial interventions are often designed as responses to poverty traps, where the poor cannot invest because they lack wealth and without investment poverty persists. The past decade of empirical development research has produced a host of highly insightful, well-identified evaluations of the impacts of microfinancial interventions. These interventions include microcredit programs; asset grants to microentrepreneurs; and small asset transfers to the very poor, regardless of their entrepreneurial status. The aim of this article is to take stock of the state of our knowledge.

We approach the topic in two steps. First, we review the findings to crystallize the salient patterns. Second, and of equal importance, we assess our understanding of these empirical patterns through the lens of economic theory. Reflecting on the policy lessons of the East Asian miracles, Robert E. Lucas, Jr. once observed,

If we understand the process of economic growth—or of anything else—we ought to be capable of demonstrating this knowledge by creating it in these pen and paper (and computer-equipped) laboratories of ours. If we know what an economic miracle is, we ought to be able to make one. (Lucas, 1993, p. 271)

Francisco J. Buera is a professor of economics at Washington University in St. Louis. Joseph P. Kaboski is a professor of economics at the University of Notre Dame. Yongseok Shin is a professor of economics at Washington University in St. Louis and a research fellow at the Federal Reserve Bank of St. Louis. This article is based on the authors' book chapter of the same name (Buera, Kaboski, and Shin, 2019) in *The Economics of Poverty Traps*. The authors thank the *Review* editor and an anonymous referee for their helpful editorial suggestions.

© 2020, Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

The same is true for poverty traps and financial interventions. If we truly understand why an intervention works, we ought to be able to recreate the empirical patterns in our theories. Such an understanding is necessary to design our policy interventions, apply them with confidence in new contexts, and make projections for larger-scale programs that will have macroeconomic consequences.

Toward the first step, this article reviews the lessons from the empirical literature on microfinancial interventions. At least three general lessons arise consistently. First, no policies produce large-scale miracle escapes from poverty traps. That is, although some of the policies have led to sustained gains, none has been shown to lead to permanent increases in income or consumption well beyond poverty levels or to extended and sizable increases in *the rate of growth of* income, consumption, and capital that predict such escapes. Second, take-up rates for microcredit are typically low, while those of asset transfer programs are understandably much higher. Third, heterogeneous responses to policies are evident in almost all studies, where impacts vary by initial wealth, size of the intervention, gender, ability, entrepreneurial status, financial access, and time frame. Variation in measurement and context (e.g., rural vs. urban settings and the degree of preexisting financial development) may also play a role.

The most interesting patterns emerge from a comparison across interventions. Although individual-level microcredit interventions can lead to increases in credit, entrepreneurial activity, and investments, they have been much less successful in leading to higher incomes or consumption. Among these interventions, only larger-scale village fund programs are shown to raise income and possibly consumption. Microcredit interventions often show relatively larger impacts on existing and marginal entrepreneurs. Small asset grants of less than \$200—at purchasing power parity (PPP)—to entrepreneurs often lead to stronger increases in capital and profits, with typically high returns on assets. Grants to “ultra poor” households often have led to changes in income-generating activities, higher asset levels and capital, and increases in consumption of up to 30 percent.

Natural questions are, what leads to such very different outcomes, and what do these outcomes say about the relevant economic mechanisms at play? Even to replicate the outcomes of these different policies in varying contexts, we need an understanding of these mechanisms. Lucas (1993, p. 252) is again much more eloquent:

[S]imply advising a society to “follow the Korean model” is a little like advising an aspiring basketball player to “follow the Michael Jordan model.” To make use of someone else’s successful performance at any task, one needs to be able to break this performance down into its component parts so that one can see what each part contributes to the whole, which aspects of this performance are imitable and, of these, which are worth imitating. One needs, in short, a theory.

A purely qualitative theory is useful in terms of organizing ideas and checking the internal consistency of one’s reasoning, but we also want to know how well such a theory can *quantitatively* explain our observations, which is undoubtedly a higher hurdle.

Toward the second step, we review existing quantitative theory on financially constrained entrepreneurial decisions. A representative model in this literature incorporates much of

what seems a priori essential in the economics involved: ex ante heterogeneity in wealth and ability, entrepreneurial decisions on both the extensive (entry) and intensive margins (scale), stochastic shocks, “necessity” entrepreneurs, and financial constraints that interact with wealth and ability. The combination of heterogeneity, intensive margins, and stochastic shocks provide enough smoothness and mixing so that poverty traps at the individual level (where investment decisions and asset and income paths depend critically on initial wealth levels) become irrelevant at the economy level (where a unique stationary equilibrium exists). Using this model, we simulate analogs of microcredit programs and asset grants targeted toward the poor and small entrepreneurs. Within our microcredit programs, we further vary the interest rates faced by borrowers. Some of these simulations reproduce results from our earlier work (Buera, Kaboski, and Shin, 2012, 2014), while others are unique to this article.

We show that the model captures many of the qualitative and quantitative patterns observed empirically in the interventions, but we also learn lessons from where it fails. For asset grants, the model shows that marginal entrepreneurs enter and that capital, income, and consumption increase, while assets tend to decline over time. However, the model does not generate the large increases in income, and we conjecture that the model fails to account for increases in labor supply in certain economic situations (e.g., where the market labor is limited for women). Moreover, the training components of such interventions might increase the effective ability of livestock entrepreneurs, or the real-world projects may somehow target people with higher ability (i.e., marginal entrepreneurs).¹ Indeed, we show that marginal products of capital for poor existing entrepreneurs are quite high in the model. For microcredit, the simulations capture low take-up rates, borrowing, and impacts concentrated in the higher end of the ability distribution, and small increases in entrepreneurship mostly due to the entry of marginal entrepreneurs. The baseline model somewhat overpredicts the increases in investment. However, with realistically higher interest rates on microloans, the model limits microcredit along the extensive and intensive margins and dampens the impacts of microcredit.

However, the simulations illuminate some long-run and general-equilibrium implications: First, microfinancial interventions can have substantial steady-state and transitional impacts on development measures (income, consumption, productivity, etc.). But the simulations show no escape from aggregate poverty traps that operate through wealth distributions and general-equilibrium effects, since these traps do not exist. In this sense, we are unable to “make a miracle.” Second, the simulations show that one-time redistributions in the form of asset grants alone tend to have only short-run aggregate and distributional impacts, as infused assets are eventually depleted over time. In contrast, microcredit—at least subsidized low-interest microcredit—has potentially longer-run impacts because of its permanent availability and general-equilibrium impact through wages. The cost effectiveness of smaller but sustained subsidies to microcredit versus one-time asset grants is therefore of interest. This result also suggests the importance of proper targeting and technical training for asset grant programs to have persistent effects.²

2 MICROEMPIRICAL ESTIMATES

In this section, we review the evidence on asset grants to microentrepreneurs and the ultra poor and on microcredit interventions. We then hypothesize about potential explanations for the patterns that emerge.

2.1 Asset Grants to Microentrepreneurs

Field experiments involving asset grants to microentrepreneurs have been undertaken in multiple countries: for example, Sri Lanka, Mexico, Ghana, and Nigeria. With one exception, all studies found significant profit increases from these asset grants. These findings are important experimental evidence for the long-held conjecture that at least some microentrepreneurs can generate above-market returns to capital, which in turn is evidence of the existence of financial constraints. We summarize these studies in Table 1.

The Sri Lanka study (de Mel, McKenzie, and Woodruff, 2008) identified about 400 non-employer entrepreneurs in urban areas of Sri Lanka and gave them small one-time grants either in kind (inventories or equipment) or in cash. They randomized between small and large grants equaling 460 or 920 PPP dollars, or roughly three to six months of average profits for these entrepreneurs. The impacts on investment were sizable: Capital increased by 70 to 130 percent of the grant at 24 months (i.e., roughly the size of the grant), and monthly profits increased by 4 to 6 percent of the original grant. The implied monthly return on the grant was substantially above market interest rates, implying recovery of the original amount after one and a half to two years—if it had been a loan. Moreover, the timing of the growth could be characterized as immediate and stable. Indeed, the point estimates of the follow-up work in de Mel, McKenzie, and Woodruff (2012) show stability of higher profits even after five years, and the results are statistically significant. The sizable returns are evidence of potential financial frictions limiting profitable investments. The fact that these impacts are stable over time but do not lead to virtuous cycles of continual reinvestment and growth, however, indicates that the gains to relaxing these constraints may be limited.

The Mexico study by McKenzie and Woodruff (2008) is a similar study, lending further support to these findings. Their study is smaller than the Sri Lanka study both in sample size (with about 200 entrepreneurs) and grant size (210 PPP dollars). They found extremely large returns to these small grants, between 20 to 33 percent *per month* at one year, but acknowledged that the 35 percent sample attrition rate was potentially problematic.

Although returns may be high, the original Sri Lanka study also emphasized the strong heterogeneity in returns to capital, however. In particular, the authors found that the impacts were driven by those with disproportionately low levels of wealth and those with higher ability (measured by education attainment or through digit recall tests). They found smaller returns on larger grants, and the returns were driven overwhelmingly by grants to men rather than to women. In their study on Ghana, Fafchamps et al. (2013) examine the impact of grants to women. They granted about 280 PPP dollars to about 800 microentrepreneurs and also found large increases in monthly profits, about 15 percent of the original grant. These results again imply high rates of return, but in contrast to the Sri Lanka study, they found that in-kind

Table 1
Studies of Grants to Microentrepreneurs

Study	de Mel, McKenzie, and Woodruff (2008)	McKenzie and Woodruff (2008)	Fafchamps et al. (2013)	Karlan, Knight, and Udry (2015)	McKenzie (2016)
Country	Sri Lanka	Mexico	Ghana	Ghana	Nigeria
Sample	408, Non-employers	198, Self-employed	793, Self-employed	160, Tailors employing 3 or fewer	1,831, Young applicants, “ordinary merit” winners
Intervention	\$460 to \$920 PPP (cash or in kind)	\$210 PPP (Cash or in kind)	\$280 PPP (Cash or in kind)	\$370 PPP (Cash), plus consulting	\$98,200 PPP (Cash), plus business training
Time horizon	24 Months	12 Months	12 Months	14 Months	12 Months
Profit (change relative to grant)	4-6% Per month	20-33% Per month	15% Per month	-67% over 14 months	23% over 12 months
Capital (change relative to grant)	70-130%	N/A	20-105%	-250%	N/A

grants yielded larger impacts than cash grants. Moreover, the in-kind grants increased profits for female entrepreneurs, whereas the Sri Lanka grants did not.

The Karlan, Knight, and Udry (2015) study, also on Ghana, provides a reminder that high returns to microentrepreneurs are not realized always and everywhere, however. They found that grants significantly decreased profits, by as much as 67 percent of the size of the initial grant. Their study experimented with a program of grants and consulting, and neither intervention proved effective. They found some positive short-run changes that quickly reversed course. A few ways in which this study differs from the Fafchamps et al. (2013) study should be noted, however. First, the sample size of 160 entrepreneurs was much smaller, about one-fifth of the other study. Given the multiple branches of the sample, it may simply be that the control group was a statistical anomaly. Second, the Karlan, Knight, and Udry (2015) study provided cash grants, while the impacts in the Fafchamps et al. (2013) study, which provided either cash or in-kind grants, were larger for in-kind grants. Third, this study focused on a particular occupation, tailors, and perhaps the industry differs from the typical microentrepreneur industry. Finally, their targeting rule allowed for entrepreneurs running slightly larger businesses, with up to 3 employees. In practice, the differences were not large, as the entrepreneurs averaged 0.35 employees and 0.86 apprentices. In addition, baseline profits were larger, so that their larger grants of 370 PPP dollars amounted to about 6 weeks of profit, comparable to the grant size in the other study.

Only one study, McKenzie (2016), has looked at the impacts of larger grants on larger firms. That study examines the impacts of average grants of nearly 100,000 PPP given to young, aspiring entrepreneurs in Nigeria. The experiment stems from an entrepreneurship competition in which applicants submitted business plans and received business training. The randomization was among a middle group of 1,200 applicants deemed of “ordinary merit”—a selected group of applicants, but not the most promising. Profits increased by 23 percent, implying a monthly rate of return of 1 to 2 percent, somewhat lower than in other studies but comparable to market rates for small-to-medium enterprises in Nigeria. Thus, with more financial access, the control group should have been able to invest.

In summary, the bulk of the evidence shows sizable returns to capital grants of modest sizes, equaling up to six months of profits of existing microentrepreneurs. On average, these grants lead to higher investment and profits, though the impacts are heterogeneous. The returns are somewhat lower for the wealthy, the less able, and female entrepreneurs.

2.2 Asset Grants to the Ultra Poor

Microentrepreneurs are often not the poorest of the poor, those living on only a few PPP dollars per day. A natural question for poverty alleviation is whether asset grants could have substantial impacts on this population. Many of the ultra poor are only involved in subsistence agriculture, where the results for existing microentrepreneurs are less relevant. Although the wealth results from the entrepreneur studies might make us expect high returns, the results for low ability and female entrepreneurs suggest otherwise. In any case, a wide set of recent studies has given us strong evidence on the impact of asset grants to rural, ultra-poor households with female heads. We summarize them in Table 2.

Table 2**Studies of Grants to the Ultra Poor**

Study	Bandiera et al. (2016)	Banerjee et al. (2015b), ex. India [†]	Banerjee et al. (2011)	Morduch, Ravi, and Bauchet (2012)	Blattman, Fiala, and Martineez (2014)	Blattman et al. (2016)	Haushofer and Shapiro (2013)
Country	Bangladesh	Five countries	India (WB)	India (AP)	Uganda	Uganda	Kenya
Sample	6,700 Women	9,500 (900 to 2,600 Per country), women	800 Women	3,500 Women	1,900 Younger adults	1,800, younger women	1,380, Men and women
Randomization level	Village	Village and individual	Individual	Village	Groups of 10-40	Village	Village and individual
Intervention	\$520 PPP or 2 cows, plus technical training	\$450-\$1,280 PPP plus consumption support	\$330 PPP plus consumption support, technical training, forced saving	\$510 PPP plus technical training, forced saving, health service, group building	\$1,310 PPP plus artisan training	\$380 PPP plus business training, group building	\$404-\$1,520 PPP plus mobile money access
Horizon	48 months	36 months	18 months	18 months	47 months	16 months	4 months
Income change	+44%	Sig. positive	+39%	Insignificant	+43%	+70-150%	+34%
Income activity	Specialized self-emp. +15 pp, self-emp. hours +106%	Sig. increase in productive asset in ETH, GHA, PAK; sig. increase in hours in ETH (17%)	48% Increase in hours worked, income from business labor	Increase in livestock income	Nonag. hours +56%, overall labor supply +19%	Hours +60%, nonag. hours +100%	Business, ag. expenses rise
Increase in assets	137% of grant	Sig. in ETH (83% of grant), PAK (14%)	Sig. positive	No impact, except the probability of owning livestock	34% of transfer, 68% of original investment	Sig. positive	35% of grant
Consumption change	10%	Sig. in ETH (18%), GHA (10%)	29%	Insignificant	Sig. positive	30%	23%

NOTE: [†]Banerjee et al. (2015b) consider programs in six countries, one of which (India) is analyzed in depth in Banerjee et al. (2011). Since we report on Banerjee et al. (2011) separately in this table, we show the results on the other five countries in Banerjee et al. (2015b), pp. percentage points; sig., significant(y); self-emp., self-employed; ETH, Ethiopia; GHA, Ghana; PAK, Pakistan; nonag., nonagricultural; ag., agricultural.

Several of the studies focus on a standardized program developed in Bangladesh by BRAC and exported to other countries. The studies focused on households headed by a female and experimented with in-kind transfers of livestock, amounting to roughly four to eight goats or one to two cattle/buffaloes. In PPP terms, the value of these assets are in the ballpark of the microentrepreneur grants described above, but they are somewhat larger and certainly larger as a fraction of the recipients' income. More important, the program is not a simple asset grant but is instead the lead part of a set of services offered to the participant households that together are designed as a micro-level "big push." These other services can include required or encouraged savings; technical assistance, often in the area of livestock rearing; and a consumption supplement. One key purpose of these other services is to lower the chances that the household would need to liquidate the livestock assets for short-term needs.

Bandiera et al. (2016) evaluate the ultra-poor program in the setting where it was developed, Bangladesh.³ Their results are the most impressive of these programs. Randomizing at the village level, they report experimental results up to four years after the livestock grants for a sample of 6,700 female-headed households. Four years out, the treatment group had higher assets, which exceeded the original value of the asset grant by 40 percent. The fraction of women specializing in self-employment increased by 15 percentage points, and labor hours in self-employment doubled. Income was 44 percent higher as well. Putting this into perspective, this amounts to extra income equivalent to 22 percent of the initial asset grant *per month*, comparable to the very high returns of microentrepreneurs in Mexico. However, the program also involved technical-assistance costs.⁴ Moreover, they find that consumption was 10 percent higher. Looking at the dynamics between two and four years, they find growth in assets, income, and consumption, but labor supply remained stable.

The largest and broadest study is Banerjee et al. (2015b), which presents experimental results for Ethiopia, Ghana, Honduras, Pakistan, Peru (and West Bengal, which we discuss below). The samples in these countries range from 900 to 2,600, and 9,500 female-headed households are involved in the analysis combined. They evaluate the impacts three years out and find that assets are higher, but by less than the initial asset transfer. There is a great deal of variation across countries. Assets are significantly higher in Pakistan and Ethiopia, but the point estimates constitute 14 percent and 83 percent of the initial transfer, respectively. The study combines multiple measures into indices, which allows for more statistical power in terms of finding significant tendencies but makes it difficult to compare their reported magnitudes with other studies or theory. Nonetheless, they find statistically significant increases in their income index and in productive assets in Ethiopia, Ghana, and Pakistan and a significant 17 percent increase in hours in Ethiopia. They find significant consumption increases of 18 percent in Ethiopia and 10 percent in Ghana.

The results in Banerjee et al. (2015b) are reported in more depth in Banerjee et al. (2011). In a sample of 800 women, where individual rather than village randomization was used, they find a substantial increase in assets, income, and consumption at 18 months. The measured increase in income of 39 percent amounts to a monthly return of 12 percent on the value of the asset. Here the cost of the program involves not only the grant and technical assistance, but also up to nine months of food supplements (tantamount to per capita consumption) and a

saving requirement of three dollars per month. Nevertheless, the returns are sizable. Moreover, the program led to an increase in consumption of 29 percent. Because measured consumption exceeds measured income, as is typical in survey data from developing countries, the absolute increase in consumption exceeds the increase in income. The consumption increase is thus financed not only by increased income-generating activities but also by sales of assets.

A larger study in another Indian state (Andhra Pradesh) finds less-promising results, however. In a sample of 3,500 female-headed households, Morduch, Ravi, and Bauchet (2012) find no significant effects on income or consumption. They find increases in livestock and livestock income, but these are offset by lower levels of labor income. Like the program in Bangladesh, this program incorporated technical assistance and mandatory savings, but it also differed in that it had a health component but no food supplement.

Looking across these studies, there is a pattern of sizable increases in income and consumption along with increases in hours. Bangladesh, West Bengal, and Ethiopia are the countries with the most promising results, and all showed increases in labor supply. In Andhra Pradesh, where jobs were widely available because of the presence of the National Rural Employment Guarantee Act, there was no increase in income or consumption, only movements from labor income to self-employed agricultural income.

Bandiera et al. (2016) build a model where labor supply plays a key role, and it may be that such a model is only relevant in particular economic environments. Indeed, it is interesting to note that the two places with the largest observed gains, Bangladesh and West Bengal, have strong affinities, both cultural and socio-economic.

Other asset grant programs in east Africa exhibit positive yet relatively modest impacts. Blattman, Fiala, and Martinez (2014) examine transfers targeted toward young adults (aged 18-35) rather than women. The grants were cash and sizable on average (1,310 PPP dollars), especially relative to the recipients' income. The grants were made at the group level, and in part they were used to finance artisanal training. About four years after the grant, the grantees had higher assets, with the difference being 34 percent of the original transfer or 68 percent of the original asset investment. Nonetheless, income was 43 percent higher, and this additional income constituted a monthly increase of about 5 percent, comparable to the returns to Sri Lankan entrepreneurs. On average, the grantees had 19 percent higher labor supply and 56 percent higher labor supply in nonagricultural/skilled-labor activities.

As mentioned, the additional assets four years out are only a fraction of the original transfer. Indeed, the program had larger effects two years out. After four years, nearly half of the recipients no longer practiced their trade. Although the program did not have a gender focus, the decline between years two and four is driven overwhelmingly by men. Nonetheless, the program is estimated to have a positive net present value.

Blattman et al. (2016) examine another program in Uganda, but one that targets women in war-torn areas of the country. The cash grants were considerably smaller (380 PPP dollars) and constituted just 17 percent of the total costs of the program, which included business-skills training, follow-up supervision, and group-building activities. The program was evaluated at 16 months: The recipients had 60 percent higher labor supply and nearly twice as many hours in nonagriculture as those in the control, and their consumption was 30 percent higher. The

increase in monthly income amounted to 7 percent of the initial transfer, again comparable to the Sri Lankan returns.

The final study reviewed is Haushofer and Shapiro (2013), which examines a program in Kenya offering grants averaging about 800 PPP dollars. The study had multiple levels of randomization including grant size, the gender of recipients, and the timing of payments. Smaller grants were made over nine months, while larger grants were made over 16 months. In principle, the drawn-out payments might help households that struggle with inconsistent intertemporal preferences unless a lump sum is needed for an indivisible, illiquid investment. The overall time horizon is much shorter, however, averaging about four months, which overlaps with the payment schedule. Over this short run, the program led to increases in income and consumption, but the monthly increase in income constituted just 2 percent of the average total transfer, somewhat lower than in the other studies. Using both a village- and an individual-level design, they find no evidence of spillovers to nonparticipants, which is in harmony with the other studies.

In sum, the asset grant programs to poor rural, usually female-headed households lead to substantial increases in assets, income, and consumption. With the exception of the Bangladesh study, the existing evidence shows the initial increase in assets dissipating over time, however.⁵

2.3 Microcredit Evaluations

The high apparent marginal returns on assets among portions of microentrepreneurs and the ultra poor suggest that financial frictions may be prohibiting these groups and could motivate the use of microcredit as an alternative program for these populations, one that could potentially improve on asset grant programs in terms of both cost-effectiveness and the ability to identify high returns. Indeed, this is the original, anecdote-based motivation for microcredit as a transformative financial intervention. A host of recent research has given a more nuanced and sober assessment of its impacts, however.

Banerjee, Karlan, and Zinman (2015c) report the results of six recent randomized evaluations of microcredit programs in Bosnia-Herzegovina, Ethiopia, India, Mexico, Mongolia, and Morocco. These are summarized in Table 3.⁶ In PPP terms, the average loans have magnitudes similar to those of the asset grants, although somewhat larger. The studies tend to find (i) relatively low take-up rates; (ii) increases in credit overall; (iii) increases in business activity, but (iv) little impact on overall measures of profits, income, or consumption. Together with these studies, Table 3 also includes two evaluations of village fund programs in China (Cai, Park, and Wang, 2016) and Thailand (Kaboski and Townsend, 2005), which show more positive results. There are some common findings, but also remarkable differences in both the programs and findings.

The first study, Attanasio et al. (2015), evaluates an expansion of microcredit within villages in Mongolia. Although generally Mongolia has a strong microcredit presence, the villages studied have relatively low baseline usage. The unique aspect of this study is the variation between joint liability and individual liability loans. The loans are relatively short-term (six months), and the study finds that after 19 months, roughly half of those surveyed have taken

Table 3

Studies of Microcredit

Study	Attanasio et al. (2015)	Crépon et al. (2015)	Tarozzi, Desai, and Johnson (2015)	Banerjee et al. (2015a)	Angelucci, Karlan, and Zinman (2015)	Augsburg et al. (2015)	Kaboski and Townsend (2011, 2012)	Cai, Park, and Wang (2016)
Country	Mongolia	Morocco	Ethiopia	India	Mexico	Bosnia and Herzegovina	Thailand	China
Sample	600 Rural, women micro-entrepreneurs	5,600 Rural, at least partly self-employed	6,300 Rural, poor, potential entrepreneurs	6,900 Urban, women	16,600, Women	1,000 Marginal borrowers	1,000 Rural, no targeting	1,200 Rural, no targeting
Randomization level	Village	Village	Peasant association	Neighborhood	Village, neighborhood	Individual	Village	Village
Average loan size	\$700 PPP	\$1,080	\$500	\$600	\$450	\$1,820	16,700 Thai baht	5,000 Chinese yuan
Nominal APR	27%	15%	12%	24%	110%	22%	7%	8%
Average loan term	6 months	16 months	12 months	12 months	4 months	14 months	12 months	12 months
Horizon	19 months	24 months	36 months	39-42 months	27 months	14 months	24 months	24 months
Take-up	50-57%	13pp	25pp	17%	19%	99%, by design	54%	29%
Overall credit change	+67%	+64%	+195%	+63%	+18pp (Fraction with loan)	+19pp (Fraction with loan)	+50%	+23pp (Fraction with loan)
Change in entrepreneurs	Fraction of entrepreneurs +8pp	Insignificant, as expected	Livestock revenue and crop exp. rise	Fraction of entrepreneurs +2pp	Revenue and crop exp. rise	Insignificant	Insignificant	Cash crop land +63%
Change in capital	Insignificant	+29%	Insignificant	+25%	-18%	Insignificant	Insignificant	+47% (in Husbandry)
Labor supply change	+57%	Decreased non-self-employed hours	Insignificant	Insignificant	N/A	Insignificant	N/A	+8%, Driven by migrant labor
Profit change	Insignificant	+40%	+68%	+57%	Insignificant	+34%	Income +35%	Income +50% (husbandry income +53%)
Consumption change	+11%	Insignificant	N/A	Insignificant	Insignificant	-16%	+10%	+8% Insignificant point est.

NOTE: pp, percentage points; exp, expenses; est., estimate.

up loans, which is higher than in the other studies. The intervention increases the fraction with loans by 26 percentage points and the level of credit overall by 67 percent. They also find an 8-percentage-point increase in the fraction of self-employed and a 57 percent increase in labor supply. This is the lone study of traditional microcredit that finds any evidence of an increase in consumption, an 11 percent increase that seems to be driven by a significant increase in food consumption.

Crépon et al. (2015) and Tarozzi, Desai, and Johnson (2015) study expansions of microcredit programs into rural areas, Morocco and Ethiopia, respectively. In Morocco, the program targeted those already involved in activities other than crops. Thus, it is unsurprising to not see an increase in the fraction of people involved in self-employment activities. After two years, the program still had low take-up, with just 13 percentage points more having borrowed, but that led to a 64 percent increase in credit overall. Capital increased by 29 percent, and there was a decrease in labor supplied to non-self-employment activities. This yielded an increase in profits of 40 percent, which was marginally significant, but no significant impact on consumption.

The Tarozzi, Desai, and Johnson (2015) study involves repeated cross sections of households, but effectively panels of villages and “peasant associations,” which are the unit of randomization. The microcredit program was joint with a family planning intervention that was ex post ineffectual. After three years, the fraction with loans was 25 percentage points higher in treatment villages and credit had increased by 195 percent. Still, they found no impacts on businesses, capital, or profits, despite the program targeting potential entrepreneurs. The survey did not measure consumption.

Banerjee et al. (2015a) evaluate an urban expansion of microcredit in India, while Angelucci, Karlan, and Zinman (2015) combine both rural and urban expansions in Mexico. They find take-up rates below 20 percent. Both programs find substantial increases in credit and its prevalence, and different measures of business activity, but neither finds a significant effect on profits (although the point estimate for India is sizable) or consumption. India shows an increase in assets, while Mexico shows a substantial decline. The Mexico program is unique in that the loans were very short term (averaging four months).

The Bosnia-Herzegovina study, Augsburg et al. (2015), stands apart in several ways. First, it randomized at the individual level, targeting marginal borrowers who otherwise would not have qualified for loans.⁷ Second, the loan amounts were substantially higher, averaging 1,820 PPP dollars. By design, the take-up rate approaches 100 percent. Still, the study finds significant impacts only on credit and not on entrepreneurship, profits, or consumption. Naturally, marginal borrowers make a unique sample, which may partially explain the latter result.

The two remaining studies examine village fund interventions and yield somewhat more positive results. Village funds differ in that they are largely independent of existing microfinance institutions and instead involve a transfer of funds to a village in order to set up its own quasi-formal institution. Kaboski and Townsend (2011, 2012) study the introduction of village funds in Thailand. Although they lack a randomized control, the fact that the government gave the same amount of funds to all villages, regardless of their size, makes village size an effective instrument for the intensity of the treatment. In the first two years, they find a

near doubling of the level of short-term credit in the villages, a 35 percent increase in income, and a 10 percent increase in consumption. Followed over six years, the increase in credit is stable, but the increases in consumption and income are concentrated in the early years.

Cai, Park, and Wang (2016) examine a similar village fund program in China, but one that had a randomized introduction at the village level. After two years, there is a 23-percentage-point increase in the probability of having a loan, substantial increases in resources going to cash crops and animal husbandry, and a 50 percent increase in income per capita. Interestingly, total working days increase, but this result is driven by migrant labor outside of the village (and province) rather than self-employed labor or labor within the village.

The setup of the Cai, Park, and Wang (2016) study allows us to compare the results using the experimental variation with the results using the quasi-experimental variation in village size of Kaboski and Townsend (2011, 2012). The results largely validate the village-size approach, although the standard errors rise, highlighting the improved identification with field experiments.

A few other nuanced findings from the empirical work deserve discussion. First, impacts tend to be heterogeneous. Kaboski and Townsend (2011) showed that households who are at the margin of large indivisible investments benefited the most. Quantile regressions in the articles in the same *American Economic Journal: Applied Economics* issue show that impacts are often concentrated among the very highest percentiles. Banerjee et al. (2015a) provide further evidence that positive impacts are concentrated among existing entrepreneurs.

Second, Angelucci, Karlan, and Zinman (2015), Crépon et al. (2015), and Cai, Park, and Wang (2016) examine impacts on (expected) nonparticipants and find no spillovers. In contrast, Kaboski and Townsend (2012) find impacts of the Thailand village fund program on local wages. Using the government-driven microfinance crisis and subsequent collapse of microfinance in Andhra Pradesh as a source of quasi-experimental variation, Breza and Kinnan (2018) find that day wages declined in areas where microcredit contracted more severely. Whether general-equilibrium spillovers are important may depend greatly on the structure of the labor market and the relative importance of microfinance.

Third, the impact of the introduction of a program on the use of other credit products varies by study. Some find that the new intervention has no effect (Attanasio et al., 2015, and Tarozzi, Desai, and Johnson, 2015), others find that it crowds out other sources (Augsburg et al., 2015; Banerjee et al., 2015a; and Cai, Park, and Wang, 2016), while still others actually find crowding in (Kaboski and Townsend, 2011, 2012; Angelucci, Karlan, and Zinman, 2015; Crépon et al., 2015; and Greaney, Kaboski, and Leemput, 2016). Even at a more disaggregate level (e.g., bank loans, informal loans), the impacts vary from crowding out to crowding in.

Fourth, the long-term impacts have been examined in two papers with different results. In Thailand, Kaboski and Townsend (2011) find that impacts fluctuate over six years but are concentrated in the early years. Banerjee et al. (2014) find fluctuations in treatment effects over time but also find some contrasting results, at least for existing entrepreneurs. They examine the impact of the collapse of microcredit in Andhra Pradesh again, looking at whether the benefits persist even after microcredit has exogenously declined. They find that the existing entrepreneurs were more profitable six years later, but that the more reluctant entrepreneurs' profitability had declined.

Finally, impacts tend to vary substantially based on program details.⁸ Attanasio et al. (2015) found that only joint liability loans led to positive impacts. Although Field and Pande (2008) found no impact of moving from weekly to biweekly payment frequency, Field et al. (2013) show that a two-month delay before the onset of required repayment leads to higher levels of entrepreneurial investments. Finally, Greaney, Kaboski, and Leemput (2016) show that the contractual structure of the administrative agents in self-help groups impacts both entrepreneurial activities and group membership.

2.4 Taking Stock Across Interventions

The evaluations uncover some commonalities but also strong differences across the interventions.

Among the commonalities, one important theme is that of individual heterogeneity. The entrepreneur grants focused on initial assets, ability, and gender. In many countries, the ultra-poor programs showed broad-based impacts (Banerjee et al., 2015b), but even they exhibit a factor of 20 difference in the impacts on income between the 90th and the 10th percentiles. Moreover, while those specializing in wage labor shifted activities toward self-employment, the impacts on earnings were much larger for those already specialized in self-employment (Bandiera et al., 2016). The microcredit work highlighted the low take-up and concentration of the largest impacts near the very top of the distribution.

A second related generalization is that even among existing entrepreneurs, interventions can increase profitability, indicating constraints along the intensive margin. The intensive-margin impacts of the entrepreneur grants are obvious, but we also find impacts of microcredit and ultra-poor grants among the existing self-employed. On the other hand, the ultra-poor grants also show impacts along the extensive margin of entrepreneurship, perhaps only for the severely constrained, however.

A third common finding was a general lack of sustained growth patterns, at least among the bulk of the population. Among those studies with multiple endlines, impacts were generally realized fairly rapidly and either remained steady or fell over time. Across the ultra-poor programs, the additional assets at endline were generally smaller than the initial grants. The one exception is the Bangladesh ultra-poor program, which led to increases in assets, income, and consumption even between years two and four.

The key difference across the interventions is the smaller impacts of microcredit on income and consumption relative to those from grants to entrepreneurs (which impacted profits positively) and to the ultra poor. We hypothesize several possible reasons for this difference, along with providing some supportive evidence.

The most obvious explanation is that the burden of repayment limits the impact of microcredit relative to grants. Take-up tends to be low, and so—in the absence of strong spillovers—much of the population is simply not affected. The need to repay could also lower the impact on consumption, even among those who borrow. However, we also see small impacts on income, indicating that this is unlikely to be the only factor. Repayment can impact the income-generating investments themselves. First, by definition, relatively high interest rates make investments less profitable. Second, the need to make immediate repayments may limit

investments with longer horizons, even if they are otherwise profitable. Interest rates of 2.5 percent per month are high, certainly higher than the returns exhibited by some in the grants studies.

In considering the burden of repayment, the village fund programs in China and Thailand are of particular interest, since they fall somewhere in between grants and pure microfinance. The fund itself is a grant to the village, but it is channeled to the villagers in the form of loans that need to be repaid. They had lower interest rates (8 and 7 percent, respectively) and longer payment schedules (a single repayment at the end of the loan). They showed relatively high take-up (54 percent and 29 percent, respectively) and resulted in strong increases in income (in both countries) and consumption (significant in Thailand). In addition, the microcredit study in Morocco allowed for a two-month grace period for animal husbandry investments, and it was the only pure microcredit study to find any evidence of higher profits. This is again consistent with Field et al. (2013), which documents the impact of a two-month grace period in India.

Another explanation is the difference in the targeted population of microcredit relative to the grant programs. The programs to the poor show that grants can have large impacts on very poor populations (at least in the short run) and that entrepreneurship grants also can have larger impacts on those with fewer assets. Microfinance programs, however, often do not lend to the poorest populations. Related, microcredit may be a “small” intervention in the sense that, in many places, those with the most to gain by borrowing may already have access to other forms of credit, and so the interventions are only changing the terms. Those whose investments respond to small changes in terms are likely those with the most marginal returns. This would be a good description of those areas in which significant crowding out was observed.

Another difference is that microcredit programs have often targeted women, and with cash. The entrepreneurship grants found that it was difficult to increase the profitability of women entrepreneurs, at least with cash. They also found that more-educated entrepreneurs exhibited bigger impacts, but women tend to be less educated than men in many developing countries. Here the village fund programs are again an interesting comparison, since they did not target women specifically and the gender education gap is small in both China and Thailand.

3 TAKING STOCK OF THEORY

We now turn to evaluating our understanding of these empirical patterns through the lens of quantitative theory. We present the basic model as developed in a series of papers: Buera and Shin (2011, 2013) and Buera, Kaboski, and Shin (2011, 2012, 2014). The model captures many common elements in the theoretical and quantitative literature, and financial frictions have quantitative bite both in the steady state (Buera, Kaboski, and Shin, 2011) and transitionally (Buera and Shin, 2013) (for a more comprehensive review of this literature, see Buera, Kaboski, and Shin, 2015). We evaluate the theory a priori based on its consistency with many of the common patterns above, discuss its implications on poverty traps, and then assess its ability to predict the variety of interventions.

3.1 Basic Model

Consistent with the commonalities discussed above, the quantitative theory has focused on models with (i) extensive entrepreneurship decisions; (ii) intensive investments; (iii) individual heterogeneity in wealth, productivity/ability, and whether or not entrepreneurship is simply a matter of necessity; and (iv) forward-looking decisions regarding entrepreneurship, investment, and saving. We reproduce this basic model below.

Individuals differ in terms of their productivity as workers x and entrepreneurs z . As entrepreneurs, they produce output using capital k , labor l , and a diminishing-returns-to-scale production function $zk^\alpha l^\theta$. Worker productivity and entrepreneurial productivity follow Markov processes that are independent of each other. Specifically, with probability γ , the value of entrepreneurial productivity remains constant from one period to the next, $z' = z$, and, with probability $1 - \gamma$, it is a random draw from a Pareto distribution, $z' = \zeta \sim -\eta\zeta^{-\eta-1}$. The stochastic nature captures the possibility of both positive and negative shocks to business profitability, which we observe in the data.⁹ A worker's productivity, or efficiency units of labor, is assumed to follow a two-state symmetric Markov chain, $x \in \{x_l, x_h\}$, with $x_l < x_h$. The probability of a shock remaining in its current value is π and $\mathbb{E}[x]$ is normalized to 1.

The financial frictions in the model follow a simple yet useful form and stem from limited enforceability of contracts. It should be noted that credit in this model is capital rental within a period. Between-period consumption loans are not allowed, which implies that individual financial wealth a must be nonnegative. By defaulting on their credit contracts, entrepreneurs can keep a fraction $1 - \phi$ of the period's output net of labor costs and the same fraction of the undepreciated capital, but they lose their financial wealth a . Defaulting individuals regain full access to credit markets in the following period, and hence the limited commitment constraint has a simple static representation.

Given the interest rate r and the wage per efficiency units of labor at time t , w_t , the problem of an individual with wealth $a \geq 0$ and worker/entrepreneurial productivity x and z at time t is recursively formulated as

$$\begin{aligned} v_t(a, x, z) = & \max_{c, a', k, l \geq 0, e \in \{0, 1\}} \left\{ \frac{c^{1-\sigma}}{1-\sigma} + \beta \mathbb{E}_{x', z'} [v_{t+1}(a', x', z') | x, z] \right\} \\ \text{s.t. } & c + a' + T_t(a) - S_t(a) \leq e [zk^\alpha l^\theta - (r + \delta)k - w_t l] + (1 - e)xw_t + (1 + r)a \\ \text{and } & zk^\alpha l^\theta - w_t l - (r + \delta)k + (1 + r)a \geq (1 - \phi) [zk^\alpha l^\theta - w_t l + (1 - \delta)k] \text{ when } e = 1, \end{aligned}$$

where c is consumption and e is the discrete occupational choice ($e = 1$ for an entrepreneur and $e = 0$ for a wage worker). The second inequality captures the financial friction for an entrepreneur, which places an upper bound on available capital. Buera, Kaboski, and Shin (2012) show this reduces to $k \leq \bar{k}(a, z; \phi)$, where $\bar{k}$ is increasing in wealth a , ability z , and ϕ . In our modeling of financial frictions, ϕ is the unique parameter indexing the enforceability of contracts across countries, and so it captures financial development and the availability of credit.¹⁰ As ϕ varies from zero to 1, the model spans the spectrum of cases from financial autarky to perfect credit markets.

The basic components of the model can be calibrated quantitatively to key measurables, including the firm-size distribution (which identifies thick-tailed ability distributions), the income distribution (which, given the thick tails, identifies the return-to-scale parameters), and larger firms' exit rates (which identify the frequency of shocks to productivity). The parameters of the labor income process can be calibrated to the autocorrelation and standard deviation of income in rural areas of developing countries, which reflect the dearth of labor market opportunities. Given the distribution of heterogeneous productivities in the population, this model can be aggregated to solve for endogenous levels of financial intermediation, productivity, aggregate capital, etc. One can do this within the framework of a partial-equilibrium model (where wages and interest rates are taken as given), a small open economy (where wages are endogenous but the interest rate is given), or a fully general equilibrium. Both steady-state and transitional analyses are computationally tractable.

We now offer a few words on some implicit modeling assumptions vis-à-vis real world empirics. First, we follow the quantitative literature in fixing total labor supply (hours worked in business or the labor market) exogenously. Although labor supply was often impacted in the experimental work cited above, this is a reasonable benchmark for looking across programs because the sign of impacts varied across studies. Nonetheless, labor supply seemed to play an important role in the more successful grants to ultra-poor woman. Second, occupational choice is binary. Empirically, we often observe households and even individuals whose income and hours are attributed to both labor and business/self-employment. Nonetheless, we view this as preferable to ignoring the natural indivisibility that comes from the minimum efficient scale or fixed costs. Buera, Kaboski, and Shin (2011) model these fixed costs explicitly and emphasize how they vary across sectors, and Buera, Kaboski, and Shin (2014) argue that such fixed costs may be necessary to explain the persistent effects on the right tail of the wealth distribution from the land distribution in Bleakley and Ferrie (2013). Similarly, Banerjee et al. (2014) argue that their microfinance results are consistent with a model with fixed costs and technology choice within industries.

One can easily consider the decisions of a single individual taking prices as given, or a full general equilibrium. One can consider either a stationary equilibrium where aggregates and prices are stable or a dynamic equilibrium where these aggregates and prices transition over time. The model therefore holds a theory for both household and aggregate behavior, and the latter allows us to have insight into potential impacts of both scaled up microfinancial interventions and macro policies.

3.2 Financial Frictions and Poverty Traps

When considering the role of poverty traps in the model, it is important to distinguish between individual and aggregate poverty traps. Within the model, we define poverty traps as self-reinforcing differences in steady-state income that result from differences in initial wealth conditions. Without financial frictions, agents with identical productivities would have identical occupational and productive choices regardless of their wealth. Since all individual decisions coincide, aggregate productive behavior (and ultimately aggregate savings behavior) is unaffected by the distribution of wealth.

Figure 1**Occupation and Saving Decision Map**

SOURCE: Authors' calculations.

However, when financial frictions are present, the model can lead to individual-level poverty traps in which agents with identical productivities but different initial wealth levels behave differently and their wealth levels diverge. Buera (2008) and later Banerjee and Moll (2010) show the importance of the lack of self-financing in driving these poverty traps. Initial wealth determines how quickly saving to self-finance would pay off, and agents with low initial wealth do not find it optimal to save for so long. At the macro level, financial frictions lower the demand for capital, while self-financing motives increase the supply. Both of these in turn lower equilibrium interest rates, leading those with no intention of becoming entrepreneurs to dissave instead.

We visually demonstrate this in Figure 1, which plots normalized net worth (a in the model) against unconstrained profits (a linear function of z in the model). The solid blue curve illustrates the occupational choice decision as a function of individual wealth and productivity. Under financial frictions it is not only high productivity that leads people to become entrepreneurs but also high wealth. The red-dashed curve represents the thresholds above which agents save and below which they dissave. For agents with high productivity, the wealthiest agents save, while the poorest agents dissave. The intersection of the occupational choice and poverty trap curves indicates that there are workers who are saving to eventually escape poverty, while there are rich entrepreneurs who will eventually fall into poverty. Of course, shocks to ability can alter these dynamics, so that these poverty traps are not absorbing states in the long run.

Beyond individual-level poverty traps, however, many stylized theories of entrepreneurial choice predict the possibility that financial frictions can lead to poverty traps for entire economies by distorting entrepreneurship (e.g., Banerjee and Newman, 1993; Ghatak and Jiang, 2002; Aghion and Bolton, 1997; and Piketty, 1997).¹¹ Poverty traps arise in these models because initial distributions affect general-equilibrium wages or interest rates, and in turn aggregate dynamics. If few people have the required initial assets to become entrepreneurs, wages and interest rates will be low, which leads to the persistence of a wealth distribution in which few have the resources to become entrepreneurs. The models typically assume a small role for forward-looking self-financing motives and intensive margins in the scale of establishments that respond to the low cost of labor and capital. For example, Banerjee and Newman (1993) and Ghatak and Jiang (2002) lack an intensive margin in the demand for labor that would make equilibrium wages respond continuously rather than discretely. Piketty (1997) lacks any labor market and, like Aghion and Bolton (1997), also abstracts from an intensive margin for capital that could respond to the interest rate. Moreover, all of these models have warm-glow savings behavior.

Qualitatively, the mechanisms emphasized in the poverty trap literature (lower interest rates and wages due to constrained entrepreneurial borrowing) are present in our benchmark model; and indeed, with the self-financing motive, the impact on interest rates can be exacerbated. The benchmark model also contains nonconvexities in production, which could generate multiple equilibria. Nevertheless, quantitative versions of these models, when properly mapped to the data, do not lead to aggregate poverty traps (e.g., Giné and Townsend, 2004, and Buera, Kaboski, and Shin, 2011)—but only slower convergence to a unique stationary distribution, the main point of Buera and Shin (2013).¹² As explained in Buera, Kaboski, and Shin (2014), aggregate poverty traps disappear once one relaxes the above mentioned simplifying assumptions needed for analytical tractability. In addition to the intensive margins, the productivity shocks ensure churning in the distribution of wealth and ability that leads to uniqueness.

3.3 Assessing Poverty Interventions

Variants of the above model have been simulated to assess asset grants (Buera, Kaboski, and Shin, 2014) and microcredit (Buera, Kaboski, and Shin, 2012). We follow those calibrations, which map the model to the moments for India's firm-size distribution and dynamics (which helps identify the z distribution and the γ shocks to z), wealth concentration (which, together with the z distribution, captures the returns to scale/share of entrepreneurial profits), labor income dynamics (which identify the x shocks), and the level of external finance relative to income (which identifies the ϕ parameter). In these papers, short-run partial-equilibrium simulations are compared with some of the above empirical results, and the long-run general-equilibrium and macroeconomic implications of the scaling up of these programs are assessed. We review these results before extending them to evaluate the role of high lending rates on the impacts of microcredit.

3.3.1 Returns to Capital Among Entrepreneurs. Before using the quantitative model to assess the aggregate and distributional impacts of poverty interventions, we first illustrate the

Figure 2**Average Marginal Product of Capital, by Entrepreneurial Wealth and Size (Employment)**

distribution of the marginal product of capital in the model economy. These returns provide a natural benchmark to compare with the estimates of the return to capital from the asset grants to microentrepreneurs in Section 2.1. The microestimates on the return to capital provide a natural test of the quantitative theory.

In Figure 2 we present the marginal product of capital among entrepreneurs. In the top panel, we show the average marginal product of capital among entrepreneurs in a given wealth percentile. For entrepreneurs in the bottom 10 percentiles, the (annual) return to capital net of depreciation is between 25 and 75 percent. The measured returns to capital to large interventions in the Nigeria study (McKenzie, 2016) are close to the lower end of this range, while the returns in the Sri Lanka study (de Mel, McKenzie, and Woodruff, 2008) are slightly above the upper end of this range. The large returns found in the Mexico study (McKenzie and Woodruff, 2008) or the Ghana study (Fafchamps et al., 2013) are only rationalized if they are interpreted as capturing an average return to capital for entrepreneurs in the lower end of the asset distribution. In the bottom panel, we plot the average marginal product of capital among entrepreneurs with a given establishment size, measured by the number of entrepreneurs. Since, with few exceptions, the real-world programs relate to the self-employed without employees, the relevant group in the model is the nonemployers in the far-left end, whose average annual return exceeds 15 percent. Overall, this figure shows that the returns to capital are very heterogeneous among entrepreneurs in the calibrated model.

3.3.2 Asset Grant Programs. Buera, Kaboski, and Shin (2014) assesses the role of asset grant programs in the context of a small open economy with a fixed interest rate.¹³ In particular, that study analyzes the transitional dynamics following an unexpected redistribution of wealth from the wealthiest to the poorest. The redistribution establishes at that point minimum wealth in the economy equal to double the average annual wage in the initial stationary equilibrium and funded in an extreme fashion by instituting a one-time, 100 percent tax on wealth above a particular threshold, $\bar{a}$.¹⁴ The size of the redistribution is fairly comparable to the asset grants to the poor summarized in Table 2, which is estimated to range from about five months to three years of income.¹⁵ The cash grants to entrepreneurs in Table 1 are a bit smaller in absolute terms, but much smaller as a fraction of reported income of the entrepreneurs (two weeks to six months of baseline profits). The exercise in Buera, Kaboski, and Shin (2014) is less comparable with the entrepreneurial grants, however, since they target the poor rather than small-scale entrepreneurs.

In medium-run projections in Buera, Kaboski, and Shin (2014), the impacts dissipate over time but largely remain after four years, but they are substantially smaller than those found in the empirical study. The program matches the empirics in that the fraction of people that pass over the poverty trap thresholds illustrated in Figure 1 is relatively small. Initially, 17 percent of the treated population switches to entrepreneurship. This compares well with the 15-percentage-point increase in Bandiera et al. (2016). In terms of labor hours, it constitutes a roughly 200 percent increase in entrepreneurship hours for this population, which is greater than the results in Table 2, which vary between about 50 and 110 percent. In this sense, the strict occupational choice may lead to too strong a result for hours.

Qualitatively, the model also predicts an increase in earnings, but quantitatively the effects on earnings are just 4 percent. This is in line with the negligible impact on income reported in Morduch, Ravi, and Bauchet (2012) but substantially smaller than the promising results found in the other studies in Table 2. Recall that where increases in income and consumption were substantial, they were accompanied by increases in labor supply, and the labor-leisure decision is not considered in the model. Alternatively, perhaps we can interpret these increases in income as resulting from an increase in productivity z as a result of technical training, which is not in the benchmark model. Thus, the large and sustained increase in earnings that some of the empirical work reports is not really a puzzle for theory.

The model matches the studies, with the exception of Bandiera et al. (2016), in that the impacts fall over time, however. After four years, the entrepreneurship rate is just 8 percentage points higher and earnings are just 3 percent higher. Again, in line with the empirics, the impacts in the model are very heterogeneous across individuals, with the earnings of the treated individuals in the 90th (95th) percentile of the entrepreneurial productivity increasing by 11 (15) percent in the second year. Thus, the model has the potential to have somewhat larger impacts for marginal entrepreneurs, but not the large gains reported in Table 1. In any case, the model is certainly consistent with a lack of a virtuous cycle of growth for the average recipient.

Although the model cannot match the magnitude of the observed income increase, the long-run macroeconomic impacts reported in Buera, Kaboski, and Shin (2014) are still of

interest. These aggregate effects include those on both the recipients and nonrecipients and those taxed by the redistribution. The wealth grants have a positive effect on aggregate total factor productivity (TFP) but a relatively larger negative impact on aggregate capital. The increase in TFP is due to the net entry of productive entrepreneurs and the capitalization of poor entrepreneurs with relatively high marginal products of capital. On impact, the decline in capital arises for the following reasons. The funds for the wealth grants come from rich active entrepreneurs who decrease their capital input by more than the drop in their wealth, since the acquisition of capital is based on leveraging wealth as collateral. This decrease is not completely offset by the grant recipients, because not all of them choose to become active entrepreneurs. In a small open economy, the redistribution of wealth therefore leads to a drop in the capital used in production and a capital outflow.

The net effect of the increase in productivity but decrease in capital on aggregate per capita income is negative but small. Although these mechanisms may be offset by the larger gains in income experienced by recipients in the empirical work, the capital decline mechanism in the model may still be an important consideration at the macro level.

All of these impacts, however, are transitional. Since the overall distribution of ability remains constant, wealth levels gradually return to their stationary distribution. There is no aggregate poverty trap to begin with, and so in the long run, the one-time wealth redistribution can cause no aggregate escape and the economy returns to its original state.

3.3.3 Microcredit Programs. Buera, Kaboski, and Shin (2012) report parallel results for the impacts of microcredit. They model microcredit as a new alternative intermediation technology that allows anyone access to a small level of credit for capital, b^{MF} , regardless of wealth or ability. The financial constraint on capital is thus relaxed to a choice between formal credit and microcredit:

$$(1) \quad k \leq \max \left\{ \bar{k}(a, z; \phi), a + b^{MF} \right\},$$

where the second element of the maximum captures the microcredit option. In the literature, microcredit is viewed as an innovation that substantially raises the probability that small uncollateralized loans are repaid. Many researchers have hypothesized about the exact nature of this technology, including joint liability, the high frequency of repayment, and dynamic incentives. Empirical tests have not clearly validated these hypotheses. As in Buera, Kaboski, and Shin (2012), we remain agnostic about the reason for the high repayment rates and simply model microfinance as a technology that guarantees the full repayment of uncollateralized loans of small sizes. While Buera, Kaboski, and Shin (2012) consider several variants of models, we report the results without the labor shock, which provides the simplest benchmark.¹⁶

Their benchmark analysis sets b^{MF} at 150 percent of annual wages, which implies a maximum microloan size relative to per capita expenditures of 1 and an average microloan size relative to per capita expenditures of 0.1. This average is comparable to the levels of average loan size to income of 6 to 43 percent reported by Banerjee, Karlan, and Zinman (2015c) in all but one of the empirical evaluations in the noted *American Economic Journal: Applied Economics* issue. In the aggregate, this value of b^{MF} leads to total microcredit constituting 30

percent of overall credit, somewhat smaller than the 33 percent in Thailand (Kaboski and Townsend, 2011, 2012) or the 44 percent in India (Banerjee, Karlan, and Zinman, 2015c).

The model does well in generating small take-up rates, 11 percent in the population overall, somewhat lower than the empirical estimates in Table 3, but those microcredit programs targeted marginal populations. The model also does well in predicting heterogeneous impacts, where both take-up and impacts are concentrated in the top decile of the entrepreneurial ability distribution.

In the short run, partial-equilibrium (i.e., small-scale) simulations, the model predicts significant impacts on entrepreneurship (a 4-percentage-point increase) overall and investment by borrowers (a 46 percent increase), but small effects on overall consumption (a 1 percent increase). The increase in entrepreneurship is on the high end of the empirical studies, with no impact in several countries but ranging from 2 percentage points in India to 8 percentage points in Mongolia, and the increase in investment is also larger than in most. The China study shows a 48 percent increase in investment, however, which is comparable to the simulation. The small increase in consumption is in line with the majority of the studies. Thailand and Mongolia show significant increases of roughly 10 percent, while the others exhibit negligible increases. In sum, the model does well in predicting the impact on entrepreneurship and consumption, but somewhat overpredicts the impact on investment.¹⁷

The aggregate impacts of microcredit are similar to those of the asset grants in the model. Capital decreases as income and resources are redistributed toward the poor, who have lower saving rates. The impacts on TFP is positive, but on net, the impacts of microcredit on per capita income are small.¹⁸

The main long-run impact of microcredit is that it is highly redistributive. Indeed, in contrast to one-time asset grants, the permanent availability of microcredit to poor households has long-run impacts. Despite low take-up, the option to finance entrepreneurship leads to a general-equilibrium increase in wages, which is consistent with the findings of Kaboski and Townsend (2012) and Breza and Kinnan (2018). When widely available, microcredit can therefore be highly redistributive, even if the take-up rates are low. Of course, these impacts are within a model where microcredit has substantial impacts on entrepreneurship and investment.

3.4 New Analysis with Interest Rate Spreads

In Section 2.4, we conjectured that the interest rate charged on microloans may contribute to the varying impacts both within microfinance interventions and across microcredit and asset grant interventions. We pursue this formally here within the context of our model. In the model results of Section 3.3.3, the interest rate on microcredit was the same as the low rate available to savers and borrowers from formal finance. Here we add microcredit-specific intermediation costs that lead to higher interest rates on microfinance loans. Microfinance loans that are not heavily subsidized typically carry high interest rate spreads, which reflect the high intermediation cost per loan of extending and (frequently) collecting repayments of a small-sized loan. In principle, the variation in these spreads may reflect different rates of subsidies toward microfinance. We simulate each lending rate as its own unique scenario.

Table 4**Simulation with Different Interest Rates on Microloans**

	Short-run partial equilibrium			Long-run general equilibrium		
Microcredit lending rate	–4%	6%	36%	–4%	6%	36%
Wage	1 by definition			1.05	1.04	1.01
Output	1.07	1.04	1.02	1.02	1.02	1.01
Capital	1.03	1.01	1.01	0.94	0.96	1.00
TFP	1.00	1.03	1.02	1.04	1.03	1.01
Consumption	1.01	1.01	1.01	1.03	1.03	1.00
Average z (active entrepreneurs)	1.01	1.02	1.01	1.03	1.04	1.02
Fraction of entrepreneurs [†]	+0.04pp	+0.01pp	+0.00pp	+0.03pp	+0.02pp	+0.00pp

NOTE: [†]Deviations from the no-microcredit economy. All other quantities are divided by their respective values in the no-microcredit economy. pp, percentage point.

In the simulations below, the interest rate on savings is –4 percent. Table 4 reports results for interest rates charged on microcredit of –4 percent, 6 percent (comparable to the low interest village funds in Table 3), and 36 percent (toward the higher end of the interest rates reported in Table 3), where all quantities are normalized by the respective levels in the no-microcredit economy. Focusing on the short-run partial-equilibrium results in the first three columns, we see that the interest rates matter considerably. At the market interest rate, in the short run, the model predicts a 4-percentage-point increase in the fraction of entrepreneurs in the population, a 3 percent increase in capital, and 7 percent increase in output. At the intermediate interest rate, comparable to the village funds, these impacts are smaller, but there is still a 4 percent increase in output, and a 1-percentage-point increase in the fraction of entrepreneurs. At the high interest rate, the impact on output is just 2 percent and there is no impact on entrepreneurship. In the short run, the increase in consumption—in this model without necessity entrepreneurs—is small across the board, just 1 percent, however.

Focusing on the long-run general equilibrium impacts, we see the patterns discussed above for the low and moderate interest rates: increases in wages, reductions in capital accumulation, and increases in TFP leading to small changes in output, but a somewhat larger increase in consumption. At the high interest rate, even these modest impacts all but disappear.

Figure 3 gives more insights into these patterns. It plots take-up rates, microcredit as a fraction of total credit, and impacts on income and consumption at different percentiles of the entrepreneurial ability (z) distribution for microcredit with different lending rates. The figure crystallizes the heterogeneous impacts in the model. At all interest rates, borrowing and impacts are concentrated near the higher end of the ability distribution, but as interest rates increase, take-up falls and becomes even more concentrated near the top (top-left panel). Microcredit relative to total credit shows even sharper declines as interest rates increase, showing that the intensive margin also responds negatively (top-right panel). The lower panels show that the impacts on income and consumption are even more muted than those on credit.

Figure 3**Short-Run Effect of Microcredit, with Different Lending Rates**

In sum, higher interest rates, as expected, dampen the effects of microfinance, bringing the short-run investment and entrepreneurship predictions closer to the existing empirical evidence on the effect of microfinance reviewed in Table 3. Another implication is the stronger positive selection of whom uses microloans. As shown in Figure 3, the higher the lending rates, the more concentrated the effect of microfinance at the top of the entrepreneurial ability distribution.

4 CONCLUDING REMARKS

We first reviewed the empirical evidence on the effect of asset grants and microcredit programs and then showed how these findings, to the extent that a pattern exists, can be explained by a model in which agents make optimal decisions subject to financial frictions. The simplest model underestimated the impact of asset grants in the short run and overestimated that of microfinance, although the microfinance results were more comparable with the estimates from village fund programs. We first conjecture that a modified version of our model with technical training, a common element of real-world asset grant programs, can

replicate the empirical evidence on asset grants. We then show that introducing realistic levels of spreads between lending and deposit rates makes the model's short-run results align with the empirical evaluations of microfinance. A central finding from both empirical and quantitative research is the lack of dramatic escapes from poverty traps.

Because the empirical studies are small-scale relative to the overall economy, and the followups are performed at most a few years after program implementation, we learn only short-run partial-equilibrium outcomes from them. Having a fully specified equilibrium model allows us to consider the macro-level effects of scaled-up programs and also programs over longer horizons. We find that it would be erroneous to simply extrapolate the short-run partial-equilibrium empirical results to predict long-run general-equilibrium effects. While one-time asset grant programs hold a lot of promise based on the empirical evidence, the model shows that they have negligible longer-run effects, since the economy, absent any other permanent change, will revert eventually to its unique invariant distribution. On the other hand, while microcredit programs in the real world have low take-up rates and small overall impact, the model shows that, once they are scaled up, through the increase in wages, even non-borrowers will be positively affected by the programs, consistent with village fund programs that showed local labor-market equilibrium effects. Again, neither intervention leads to escapes from poverty traps, even when scaled up to the full economy.

More broadly, we see large gains from trade between micro and macro developments. The well-established micro-experimental evidence helps us enhance theoretical models, while quantitative theory is a natural guide to interpreting the micro evidence and making predictions on what can be expected when existing programs are scaled up over time. ■

NOTES

- ¹ Business training interventions have not proven particularly effective, but the training in these programs involves technical training regarding livestock rearing.
- ² Kaboski and Townsend (2011) compare asset grants to village funds and find that the latter are more cost-effective overall. This model has indivisibilities but only an intensive investment margin and is a partial equilibrium model.
- ³ Emran, Robano, and Smith (2014) find very similar results for the BRAC program in Bangladesh based on earlier data but using nonexperimental methods to account for selection bias.
- ⁴ In conversations, the authors also reported the presence of food supplements and savings encouragement.
- ⁵ This is difficult to assess for papers such as Blattman et al. (2016), who only provide a normalized asset index. Asset transfers of land as in Bleakley and Ferrie (2013) and Keswell and Carter (2014) show long persistence, although Bleakley and Ferrie (2013), perhaps surprisingly, only find it on the right tail of the distribution.
- ⁶ Some of the information reported comes from the individual papers, and some comes from the Banerjee, Karlan, and Zinman (2015c) overview article.
- ⁷ Karlan and Zinman (2010, 2011) follow a similar approach.
- ⁸ Kaboski and Townsend (2005) is an early paper showing the importance of program policies for impacts in a non-experimental setting.
- ⁹ Note that we model shocks to productivity rather than assets as in Ikegami et al. (2016).
- ¹⁰ Buera, Kaboski, and Shin (2015) review alternatives to this form in the literature.
- ¹¹ Matsuyama (2011) provides an excellent recent review of these and related results.
- ¹² See also Moll (2014) for a theoretical analysis of this point.
- ¹³ In order to capture the poor saving opportunities in developing countries, we set this interest rate to zero, which is 2 percent lower than the historical average in developed economies.
- ¹⁴ Specifically, we implement in the initial period the wealth grant $S_0(a) = \max\{2w\mathbb{E}[x] - a, 0\}$, which is financed by a one-time tax over the wealthiest individuals, $T_0(a) = \max\{a - \bar{a}, 0\}$, where $\bar{a}$ is chosen to satisfy the static government budget constraint.
- ¹⁵ These calculations are complicated because income may be underestimated (e.g., people underreport noncash income or income of other household members) and also because many studies report household income, which we need to convert into income per working-age household member.
- ¹⁶ In an extension in the paper, the authors consider a stark calibration of a labor shock where $x_i = 0$ to capture an individual who is forced into entrepreneurship because the individual has no labor market option. They choose the process of the labor shock to match the high rates of entrepreneurship that are typical in developing countries.
- ¹⁷ In the version of the model in which a sizable fraction of the population faces a lack of labor market opportunities and therefore become necessity entrepreneurs, microcredit has a bigger impact on consumption (a 20 percent increase).
- ¹⁸ In the case with necessity entrepreneurs, the effects on per capita income can be even negative, although consumption increases.

REFERENCES

- Aghion, P. and Bolton P. "A Theory of Trickle-Down Growth and Development." *Review of Economic Studies*, 1997, 64, pp. 151-72; <https://doi.org/10.2307/2971707>.
- Angelucci, M.; Karlan, D. and Zinman, J. "Microcredit Impacts: Evidence from a Randomized Microcredit Program Placement Experiment by Compartamos Banco." *American Economic Journal: Applied Economics*, 2015, 7, pp. 151-82; <https://doi.org/10.1257/app.20130537>.

- Attanasio, O.; Augsburg, B.; De Haas, R.; Fitzsimons, E. and Harmgart, H. "The Impacts of Microfinance: Evidence from Joint-Liability Lending in Mongolia." *American Economic Journal: Applied Economics*, 2015, 7, pp. 90 -122; <https://doi.org/10.1257/app.20130489>.
- Augsburg, B.; De Haas, R.; Harmgart, H. and Meghir, C. "The Impacts of Microcredit: Evidence from Bosnia and Herzegovina." *American Economic Journal: Applied Economics*, 2015, 7, pp. 183-203; <https://doi.org/10.1257/app.20130272>.
- Bandiera, O.; Burgess, R.; Das, N.; Gulesci, S.; Rasul, I. and Sulaiman, M. "Labor Markets and Poverty in Village Economies." Unpublished manuscript, 2016; <https://doi.org/10.1257/rct.2231>.
- Banerjee, A.; Duflo, E.; Chattopadhyay, R. and Shapiro, J. "Targeting the Hardcore Poor: An Impact Assessment." Unpublished manuscript, MIT, 2011.
- Banerjee, A.; Breza, E.; Duflo, E.; Glennerster, R. and Kinnan, C.G. "Does Microfinance Foster Business Growth? The Importance of Entrepreneurial Heterogeneity." MIT, 2014.
- Banerjee, A.; Duflo, E.; Glennerster, R. and Kinnan, C. "The Miracle of Microfinance? Evidence from a Randomized Evaluation." *American Economic Journal: Applied Economics*, 2015a, 7, pp. 22 -53; <https://doi.org/10.1257/app.20130533>.
- Banerjee, A.; Duflo, E.; Goldberg, N.; Karlan, D.; Osei, R.; Pariente, W.; Shapiro, J.; Thuysbaert, B. and Udry, C. "A Multifaceted Program Causes Lasting Progress for the Very Poor: Evidence From Six Countries." *Science*, 2015b, 348; <https://doi.org/10.1126/science.1260799>.
- Banerjee, A.; Karlan, D. and Zinman, J. "Six Randomized Evaluations of Microcredit: Introduction and Further Steps." *American Economic Journal: Applied Economics*, 2015c, 7, pp. 1-21; <https://doi.org/10.1257/app.20140287>.
- Banerjee, A. and Moll, B. "Why Does Misallocation Persist?" *American Economic Journal: Macroeconomics*, 2010, 2, pp. 189-206; <https://doi.org/10.1257/mac.2.1.189>.
- Banerjee, A. and Newman, A.F. "Occupational Choice and the Process of Development." *Journal of Political Economy*, 1993, 101, pp. 274-98; <https://doi.org/10.1086/261876>.
- Blattman, C.; Fiala, N. and Martinez, S. "Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda." *Quarterly Journal of Economics*, 2014, 129, pp. 697-752; <https://doi.org/10.1093/qje/qjt057>.
- Blattman, C.; Green, E.P.; Jamison, J.; Lehmann, M.C. and Annan, J. "The Returns to Microenterprise Support among the Ultrapoor: A Field Experiment in Postwar Uganda." *American Economic Journal: Applied Economics*, 2016, 8, pp. 35-64; <https://doi.org/10.1257/app.20150023>.
- Bleakley, H. and Ferrie, J.P. "Up from Poverty? The 1832 Cherokee Land Lottery and the Long-run Distribution of Wealth." NBER Working Papers 19175, National Bureau of Economic Research, 2013; <https://doi.org/10.3386/w19175>.
- Breza, E. and Kinnan, C. "Measuring the Equilibrium Impacts of Credit: Evidence from the Indian Microfinance Crisis." NBER Working Paper 24329, National Bureau of Economic Research, 2018.
- Buera, F.J. "Persistency of Poverty, Financial Frictions, and Entrepreneurship." Unpublished manuscript, Northwestern University, 2008.
- Buera, F.J.; Kaboski, J.P. and Shin, Y. "Finance and Development: A Tale of Two Sectors." *American Economic Review*, 2011, 101, pp. 1964-2002; <https://doi.org/10.1257/aer.101.5.1964>.
- Buera, F.J.; Kaboski, J.P. and Shin, Y. "The Macroeconomics of Microfinance." NBER Working Papers 17905, National Bureau of Economic Research, 2012; <https://doi.org/10.3386/w17905>.
- Buera, F.J.; Kaboski, J.P. and Shin, Y. "Macro-Perspective on Asset Grants Programs: Occupational and Wealth Mobility." *American Economic Review*, 2014, 104, pp. 159-64; <https://doi.org/10.1257/aer.104.5.159>.
- Buera, F.J.; Kaboski, J.P. and Shin, Y. "Entrepreneurship and Financial Frictions: A Macroeconomic Perspective." *Annual Review of Economics*, 2015, 7, pp. 409-36; <https://doi.org/10.1146/annurev-economics-080614-115348>.
- Buera, F.J.; Kaboski, J.P. and Shin, Y. "Taking Stock of the Evidence on Micro-Financial Interventions," in C.B. Barrett, M.R. Carter, and J-P Chavas, eds., *The Economics of Poverty Traps*. National Bureau of Economic Research and University of Chicago Press, 2019, pp. 189-221.

- Buera, F.J. and Shin Y. "Self-insurance vs. Self-financing: A Welfare Analysis of the Persistence of Shocks." *Journal of Economic Theory*, 2011, 146, pp. 845-62; <https://doi.org/10.1016/j.jet.2011.01.003>.
- Buera, F.J. and Shin Y. "Financial Frictions and the Persistence of History: A Quantitative Exploration." *Journal of Political Economy*, 2013, 121, pp. 221-72; <https://doi.org/10.1086/670271>.
- Cai, S.; Park, A. and Wang, S. "Microfinance Can Raise Incomes: Evidence from a Randomized Control Trial in China." Technical report, Hong Kong University of Science and Technology, 2016.
- Crépon, B.; Devoto, F.; Dufló, E. and Pariente, W. "Estimating the Impact of Microcredit on Those Who Take It Up: Evidence from a Randomized Experiment in Morocco." *American Economic Journal: Applied Economics*, 2015, 7, pp. 123-50; <https://doi.org/10.1257/app.20130535>.
- de Mel, S.; McKenzie, D. and Woodruff, C. "Returns to Capital in Microenterprises: Evidence from a Field Experiment." *Quarterly Journal of Economics*, 2008, 123, pp. 1329-72; <https://doi.org/10.1126/science.1212973>.
- de Mel, S.; McKenzie, D. and Woodruff, C. "One-Time Transfers of Cash or Capital Have Long-Lasting Effects on Microenterprises in Sri Lanka." *Science*, 2012, 335, pp. 962-66; <https://doi.org/10.1126/science.1212973>.
- Emran, M.S.; Robano, V. and Smith, S.C. "Assessing the Frontiers of Ultrapoverty Reduction: Evidence from Challenging the Frontiers of Poverty Reduction/Targeting the Ultra-poor, an Innovative Program in Bangladesh." *Economic Development and Cultural Change*, 2014, 62, pp. 339-80; <https://doi.org/10.1086/674110>.
- Fafchamps, M.; McKenzie, D.; Quinn, S. and Woodruff, C. "Microenterprise Growth and the Flypaper Effect: Evidence from a Randomized Experiment in Ghana." Technical report, University of Oxford, 2013; <https://doi.org/10.1016/j.jdevco.2013.09.010>.
- Field, E. and Pande, R. "Repayment Frequency and Default in Microfinance: Evidence from India." *Journal of the European Economic Association*, 2008, 6, pp. 501-09; <https://doi.org/10.1162/JEEA.2008.6.2-3.501>.
- Field, E.; Pande, R.; Papp, J. and Rigol, N. "Does the Classic Microfinance Model Discourage Entrepreneurship among the Poor? Experimental Evidence from India." *American Economic Review*, 2013, 103, pp. 2196-226; <https://doi.org/10.1257/aer.103.6.2196>.
- Ghatak, M. and Jiang, N.-H. "A Simple Model of Inequality, Occupational Choice, and Development." *Journal of Development Economics*, 2002, 69, pp. 205-26; [https://doi.org/10.1016/S0304-3878\(02\)00059-7](https://doi.org/10.1016/S0304-3878(02)00059-7).
- Giné, X. and Townsend, R.M. "Evaluation of Financial Liberalization: A General Equilibrium Model with Constrained Occupation Choice." *Journal of Development Economics*, 2004, 74, pp. 269-307; <https://doi.org/10.1016/j.jdevco.2003.03.005>.
- Greaney, B.; Kaboski, J.P. and Leemput, E.V. "Can Self-Help Groups Really Be 'Self-Help'?" *Review of Economic Studies*, 2016, 83, pp. 1614-44; <https://doi.org/10.1093/restud/rdw004>.
- Haushofer, J. and Shapiro, J. "Household Response to Income Changes: Evidence from an Unconditional Cash Transfer Program in Kenya." Technical report, Abdul Latif Jameel Poverty Action Lab, 2013.
- Ikegami, M.; Carter, M.R.; Barrett, C.B. and Janzen, S.A. "Poverty Traps and the Social Protection Paradox." NBER Working Papers w22714, National Bureau of Economic Research, 2016; <https://doi.org/10.3386/w22714>.
- Kaboski, J. and Townsend, R. "Policies and Impact: An Analysis of Village-Level Microfinance Institutions." *Journal of the European Economic Association*, 2005, 3, pp. 1-50; <https://doi.org/10.1162/1542476053295331>.
- Kaboski, J. and Townsend, R. "A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative." *Econometrica*, 2011, 79, pp. 1357-406; <https://doi.org/10.3982/ECTA7079>.
- Kaboski, J.P. and Townsend, R.M. "The Impact of Credit on Village Economies." *American Economic Journal: Applied Economics*, 2012, 4, pp. 98-133; <https://doi.org/10.1257/app.4.2.98>.
- Karlan, D.; Knight, R. and Udry, C. "Consulting and Capital Experiments with Microenterprise Tailors in Ghana." *Journal of Economic Behavior and Organization*, 2015, 118, pp. 281-302; <https://doi.org/10.1016/j.jebo.2015.04.005>.
- Karlan, D. and Zinman, J. "Expanding Credit Access: Using Randomized Supply Decisions to Estimate the Impacts." *Review of Financial Studies*, 2010, 23, pp. 433-64; <https://doi.org/10.1093/rfs/hhp092>.
- Karlan, D. and Zinman, J. "Microcredit in Theory and Practice: Using Randomized Credit Scoring for Impact Evaluation." *Science*, 2011, 332, pp. 1278-84; <https://doi.org/10.1126/science.1200138>.

- Keswell, M. and Carter, M.L. "Poverty and Land Distribution." *Journal of Development Economics*, 2014, 110, pp. 250-61; <https://doi.org/10.1016/j.jdeveco.2013.10.003>.
- Lucas, Robert E.J. "Making a Miracle." *Econometrica*, 1993, 61, pp. 251-72; <https://doi.org/10.2307/2951551>.
- Matsuyama, K. "Imperfect Credit Markets, Household Wealth Distribution, and Development." *Annual Review of Economics*, 2011, 3, pp. 339-62; <https://doi.org/10.1146/annurev-economics-111809-125054>.
- McKenzie, D. and Woodruff, C. "Experimental Evidence on Returns to Capital and Access to Finance in Mexico." *World Bank Economic Review*, 2008, 22, pp. 457-82; <https://doi.org/10.1093/wber/lhn017>.
- McKenzie, D.J. "Identifying and Spurring High-Growth Entrepreneurship: Experimental Evidence from a Business Plan Competition." Unpublished manuscript, World Bank, 2016; <https://doi.org/10.1596/1813-9450-7391>.
- Moll, B. "Productivity Losses from Financial Frictions: Can Self-financing Undo Capital Misallocation?" *American Economic Review*, 2014, 104, pp. 3186-221; <https://doi.org/10.1257/aer.104.10.3186>.
- Morduch, J.; Ravi, S. and Bauchet, J. "Failure vs. Displacement: Why An Innovative Anti-Poverty Program Showed No Net Impact." Discussion Paper Series, 32, 2012, PRIMCED.
- Piketty, T. "The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing." *Review of Economic Studies*, 1997, 64, pp. 173-89; <https://doi.org/10.2307/2971708>.
- Tarozzi, A.; Desai, J. and Johnson, K. "The Impacts of Microcredit: Evidence from Ethiopia." *American Economic Journal: Applied Economics*, 2015, 7, pp. 54-89; <https://doi.org/10.1257/app.20130475>.

On the Aggregate Implications of Removing Barriers to Formality

Catalina Granda and Franz Hamann

This article examines the aggregate implications of several policies aimed at removing barriers to formality. To this end, we build a dynamic equilibrium model in which heterogeneous agents choose to work for a wage or operate a technology in the formal or informal sector, based on the costs and benefits associated with these occupational choices. Formality entails compliance with taxes, a minimum wage scheme, and firm operation costs but has a productivity advantage stemming from access to external finance and legal enforcement mechanisms. Informal activities avoid taxes and regulations without detection or punishment. The simulation results suggest that eliminating formal operation fees leads to firm formalization, earnings redistribution, and increases in total factor productivity and welfare. In addition, eliminating the income tax reduces labor informality. These two policies, taken together, generate full formalization and gains in redistribution, efficiency, and welfare that are even greater than when all the barriers to formality are jointly removed. In contrast, eliminating the minimum wage has strong adverse effects on labor formalization and little impact on productivity. Eliminating the payroll tax leaves the occupational composition nearly unchanged and productivity and welfare as well. (JEL E26, H20, J30, L26, O17)

Federal Reserve Bank of St. Louis *Review*, Second Quarter 2020, 102(2), pp. 203-20.
<https://doi.org/10.20955/r.102.203-20>

1 INTRODUCTION

An extensive phenomenon worldwide, but especially in developing economies, is the existence of large informal sectors. Indeed, Schneider, Buehn, and Montenegro (2010) report average sizes of over 40 percent of gross domestic product (GDP) for Sub-Saharan Africa, Latin America, and the Caribbean, which contrast with about 17 percent in high-income Organisation for Economic Co-operation and Development (OECD) countries. While both

Catalina Granda is an associate professor of economics at the Universidad de Antioquia. Franz Hamann is director of the department of macro-economic modeling at Banco de la República. The authors thank one anonymous referee for very detailed suggestions. Catalina Granda gratefully acknowledges financial support from the program “*Inclusión productiva y social: programas y políticas para la promoción de una economía formal*,” code 60185, which conforms to the Alianza EFI – Economía Formal Inclusiva, under the contingent recovery grant No. FP44842-220-2018. All errors and omissions are our own. Amalia Rodríguez provided valuable assistance in microdata processing.

© 2020, Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Banco de la República or its board of directors, the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

Figure 1**Saving Motives by Sector (percent)**

NOTE: See note 2 for the categorization of motives.

SOURCE: Authors' calculations based on the 2013 wave of the ELCA survey.

firms and workers often operate in the informal sector to avoid cumbersome regulations and taxation, there are costs—such as lack of access to the judicial and legal system and to financial markets (including insurance against retirement income risk and other risks)—that are likely to exert a negative influence on productivity, growth, and welfare.

In view of the potential advantages of overcoming the informal sector, this article develops a framework to quantitatively examine the macroeconomic implications of several policies aimed at removing barriers to formality. To this end, we build a dynamic equilibrium model wherein heterogeneous agents choose whether to operate in the formal sector or the informal sector. These decisions result from analyzing the costs and benefits associated with such occupational choices and take into account a variety of policies. Thus, individuals in the formal sector must comply with taxes, a minimum wage, and firm operation costs but also have a productivity advantage stemming from access to external finance and legal contract enforcement mechanisms. Individuals in the informal sector avoid taxes and regulations without being audited and punished.

In the model, there are financial market imperfections that result in agents' inability to insure against idiosyncratic uncertainty, thus inducing them to save for precautionary reasons. This behavior is consistent with evidence from developing countries suggesting that buffering unexpected events seems to be the main motivation for saving among low- and middle-income individuals. In particular, using data from the Colombian Longitudinal Survey (henceforth ELCA),¹ we find this to be the most important motive irrespective of the occupation and/or sector of operation.² Moreover, according to Figure 1, individuals employed in the informal sector save more for precautionary reasons than those employed in the formal sector.³

The model is calibrated to match important aspects of the Colombian microdata, as this developing country is highly regulated (considering both labor and entry regulations), with extensive informality at the firm level and in the labor market.⁴ This procedure allows us to assess the impact of a broad array of highly debated formalization policies on the economy's extent of informality, productivity, and welfare. Specifically, the policies considered involve eliminating labor taxation (i.e., payroll and income), fixed costs of formal operation, and the minimum wage.

The simulation results suggest that eliminating both the income tax and the fixed costs of formal operation substantially improves the occupational composition. Indeed, eliminating the fixed costs of formal operation leads to firm formalization, earnings redistribution, and increases in aggregate efficiency and welfare. In turn, reducing the income tax rate to zero is an extremely effective policy for promoting labor formality, but it is inconsequential for earnings concentration and productivity. These two policies, taken together, generate full formalization and gains in redistribution and efficiency that are even greater than when all the barriers to formality are jointly removed.

In contrast, eliminating the payroll tax and the minimum wage is not beneficial for overcoming the informal sector and improving aggregate efficiency. With no payroll tax, both labor and firm informality remain nearly unchanged, as do productivity and welfare. Further, eliminating the minimum wage has strong adverse effects on the formal sector and little impact on efficiency. These results can be rationalized by the insurance role that a nonbinding minimum wage plays in our model, which mitigates financial incompleteness for low-productivity individuals who work in the formal sector.

Some of the policies considered in this article have been widely analyzed at an empirical level. Indeed, a number of studies have shown that high nonwage costs and wage inflexibility associated with the minimum wage discourage formal employment and lead to high informal employment rates (see, for instance, Bell, 1997; Kugler and Kugler, 2009; and Mondragón-Vélez, Peña, and Wills, 2010). Further, the literature on entry regulation and formalization of microenterprises in developing countries suggests that reforms intended to simplify business registration in the past decades have resulted in modest increases in the number of formal firms (see Bruhn and McKenzie, 2014, and the references therein).

Theoretically, this article is related to a strand of the literature that analyzes the aggregate effects of different formalization policies in developing countries. In a recent study, Ulyssea (2018) develops an equilibrium model wherein formal firms face fixed costs of registration and comply with revenue and labor taxes, yet they may avoid the latter by hiring informal workers. Informal firms in turn are able to evade all taxes and regulations, but they face a detection cost. The author uses the model to conduct counterfactual analyses of several policies toward informality, considering two experiments of interest: (i) reducing formal sector entry costs and (ii) a payroll tax cut. These experiments show that firm and labor informality need not move in the same direction as a result of policy changes. In particular, he finds that lowering registration costs is not effective in reducing labor informality and that lowering the payroll tax reduces the number of informal firms only slightly. Overall, his findings suggest that fewer informal firms and workers are not necessarily associated with higher output, total factor productivity (TFP), or welfare.

Furthermore, Ulyssea (2010) examines the role of labor market institutions and entry regulations in the size of the informal sector and in overall labor market performance. To that end, he develops a two-sector matching model that incorporates the main features of developing countries' labor and entry regulations. The simulation results indicate that reducing payroll taxes and increasing unemployment benefits are not effective policies for reducing informality and improving labor market indicators. In contrast, lowering the costs of entry into the formal sector significantly reduces the size of the informal sector and substantially improves employment composition. The author concludes that the best option for decreasing informality and improving labor market performance and welfare would be to reduce the formal sector's entry costs, instead of intensifying punishment and auditing informal activities.

Moreover, this article is related to a number of recent studies that quantitatively examine the impact on economic development of financial market imperfections and costs of creating and operating formal sector firms. D'Erasmus and Moscoso Boedo (2012) propose a general equilibrium model of firm dynamics, finding that countries with low degrees of debt enforcement and high costs of formality are characterized by low allocative efficiency and a large share of output produced by low-productivity informal firms. Somewhat similarly, Antunes and Cavalcanti (2007) construct a model with three occupational choices (worker, formal entrepreneur, or informal entrepreneur) and inequality in wealth and entrepreneurial ideas to assess how much of the cross-country variation in the size of the informal sector and per capita income can be attributed to entry barriers and limited enforcement of financial contracts. These authors find, among other results, that contract enforcement and regulation costs interact in nonlinear ways and cannot account for much of the output differences across countries.

In the same vein, Lopez-Martin (2019) builds a model of entrepreneurship with the same three occupational choices and financial frictions to evaluate the impact of several formalization policies. His findings suggest that eliminating registration costs has modest effects but that improving access to credit is key to reducing the informal sector size and increasing aggregate TFP and output per worker. Lastly, Araujo and Rodrigues (2016) analyze the role of taxation and credit constraints on formalization, aggregate efficiency, and income distribution. These authors find that, taken together, the distortions included in their model are able to generate substantial inefficiency and inequality. Further, while the efficiency implications of removing these distortions largely come from borrowing constraints, they find that tax rates are the main driver behind inequality reductions.

The article is organized as follows. The main aspects of the model economy are described in Section 2, whereas an assessment of its quantitative performance is left to Section 3. In Section 4, we examine how the informal sector and several macroeconomic aggregates behave in the face of changes in the policy parameters mentioned above. Section 5 concludes.

2 THE MODEL

We propose a model of occupational choice featuring taxation, wage rigidities, and fixed costs of formal operation. The model economy is small and open and financial markets are incomplete, but individuals use a risk-free asset to insure against idiosyncratic uncertainty. Our analysis focuses on the stationary equilibrium.

The economy is populated by a continuum of infinitely lived individuals who are heterogeneous with respect to their productivity z , their occupation o , and their wealth a . While the latter two are chosen endogenously by forward-looking decisions, the former is an exogenous stochastic process that evolves over time according to the Markov transition probabilities $p(z', z) = \Pr(z_{t+1} | z_t)$. Depending on an individual's occupation, z can be regarded as entrepreneurial ability or units of labor efficiency.

In each period, individuals choose whether to work for a wage or to operate a business in one of two sectors: a formal sector encompassing all production produced strictly following all regulations in place and an informal sector encompassing all production unreported to the tax authorities and conducted with workers not hired under a regular contract. Their choices are based on the comparative advantages associated with each occupational status as described below.

Specifically, at the beginning of each period, individuals are either entrepreneurs or workers in either the formal or informal sector,

$$o \in \mathcal{O} = \{\text{formal entrepreneur, informal entrepreneur, formal worker, informal worker}\},$$

and hold some level of a risk-free asset, $a \in \mathcal{A} = \{a_0 < a_1 < \dots < a_{n_a}\}$, such that their asset choices are bounded by the no-borrowing constraint $a_0 \geq 0$.

2.1 Technology and Preferences

Individuals discount their future utility using the discount factor $\beta \in (0, 1)$. Their problem is to maximize lifetime utility as described by

$$U = E_0 \sum_{t=0}^{\infty} \beta^t u(ct),$$

where per-period utility exhibits constant relative risk aversion:

$$u(c_t) = \frac{c_t^{1-\sigma} - 1}{1-\sigma}.$$

To this end, individuals face two types of decisions: a *static* one, whereby entrepreneurs choose the optimal factor demands, and a *dynamic* one, which involves an agent's occupation o' and asset holdings a' next period.

As for the static decision, an entrepreneur with ability z uses labor to produce a homogeneous consumption good. At the beginning of each period, the entrepreneur operates the technology $f^s(\cdot)$ in either of the two sectors, $s = \{f, i\}$. The entrepreneur's production technology is of the form

$$f^s(z_t, n_t^s) = z_t Z^s (n_t^s)^\mu,$$

where n_t^s is the level of labor used in period t and Z^s is a sector-specific productivity parameter such that $Z^f > 1$ and $Z^i = 1$. Further, it is assumed that $0 < \mu < 1$, meaning that production exhibits decreasing returns to scale. Total production is defined as $y_t^{\text{tot}} = \sum_s y_t^s$.⁵

This specification of production technologies implies that formal sector firms enjoy a productivity advantage relative to informal sector firms. Such an advantage constitutes an incentive for entrepreneurs to operate formally and can be thought of as stemming from the ability to engage in legal contracts and access infrastructure facilities and financial markets, among others (Perry et al., 2007; Dabla-Norris, Gradstein, and Inchauste, 2008).⁶

2.2 Individual's Problem

The incentives individuals face depend on whether they decide to remain as workers in either of the two sectors or switch and become a formal entrepreneur. In particular, if an individual chooses to become a formal entrepreneur, he must pay a creation cost for his business η , which is denominated in units of the consumption good.

Also, a formal entrepreneur hires each labor unit at a wage w^f and pays constant tax rates on profits, $\tau^c \geq 0$, and payroll, $\tau_f^w \geq 0$. An entrepreneur operating in the informal sector, in contrast, must pay a wage per unit of labor of w^i and pays no taxes ($\tau^c = \tau_i^w = 0$).

Likewise, an individual who decides to be a formal worker earns a wage $w^f = \max\{w_{min}, wz\}$, the maximum wage between the minimum wage and a labor efficiency wage, and pays a fixed tax rate on labor income $\tau^y \geq 0$. If the individual instead chooses to work informally, the individual is paid the wage $w^i = wz$ and charged no taxes.

There are no switching costs on labor supply. Moreover, irrespective of their sector of operation, entrepreneurs choosing to liquidate their formal businesses and move to another occupation face no exit costs. The government cannot distinguish an individual's occupation, and hence informal activities are neither monitored nor punished. All revenue collected from taxes and operation costs is used for unproductive activities.

Accordingly, an individual's earnings depend on the individual's current occupation and can be summarized in the following reward function⁷:

$$r(z, o) = \begin{cases} (1 - \tau^c) \pi^f(z), & o = \text{formal entrepreneur} \\ \pi^i(z), & o = \text{informal entrepreneur} \\ (1 - \tau^y) w^f, & o = \text{formal worker} \\ w^i, & o = \text{informal worker} \end{cases},$$

where the indirect profit function from running a business in sector s , π^s , is defined as

$$\pi^s(z; w^s) = \max_{n^s} \left\{ f^s(z, n^s) - (1 + \tau_s^w) w^s n^s \right\}$$

and the associated labor demands are $n^s(z; w^s)$ with $n^s \geq 0$.

Thus, the recursive representation of an individual's problem is given by the following Bellman equation:

$$(1) \quad \begin{aligned} v(z, a, o) = & \max_{c \geq 0, a' \in \mathcal{A}, o' \in \mathcal{O}} \left\{ u(c) + \beta \sum_{z'} p(z', z) v(z', a', o') \right\} \\ \text{subject to} \quad & c + a' = r(z, o) + R^* a - \eta 1_{-o'} \\ & a \geq 0, \end{aligned}$$

where $1_{-o'}$ is an indicator variable that equals 1 if an individual starts or continues operating a formal enterprise and zero otherwise. Note that R^* is the gross rate of return on assets.

The value function v is defined on three state variables: productivity, $z \in \mathcal{Z}$; net wealth, $a \in \mathcal{A}$; and occupation, $o \in \mathcal{O}$. The decision variables are current consumption, c , and net asset holdings and occupation next period, $a' \in \mathcal{A}$ and $o' \in \mathcal{O}$, respectively. Letting ω be the vector of state variables, $\omega = (z, a, o)$, the optimal policy function for the decision variables can be denoted by $x(\omega) = \{c(\omega), a'(\omega), o'(\omega)\}$. Accordingly, the controlled-state process of the just-described dynamic program is a Markov chain with transition probability matrix $\mathcal{P}$ and ergodic distribution h .

This framework implies that individuals save to smooth consumption. Note that earnings from all occupations are stochastic, but likely more so for informal workers and (both formal and informal) entrepreneurs. In this sense, labor market rigidities secure flat income flows and determine how many workers join the formal sector, in turn affecting how many join the informal sector. Put differently, for workers in the formal sector, the minimum wage policy provides insurance from the downside risk of stochastic productivity associated with entrepreneurship and informality. Such is the main benefit for formal workers, even though informal workers may earn more.

2.3 Equilibrium

Notice that the solution to the discrete dynamic program described above depends on the equilibrium values of the formal and informal wages (w^f, w^i). In this regard, the minimum wage implies a lower bound on wages earned by workers in the formal sector. Thus, there is an individual productivity threshold below which lower-productivity workers earn a higher wage than they would have earned in the absence of the minimum wage. Since workers can move freely across sectors, however, formal and informal wages are determined endogenously by the mass of individuals willing to work in either sector at wages w^f and w^i , respectively, relative to the labor demand by formal and informal entrepreneurs at those wages. To sum up, whereas W_{min} is taken as a parameter, w^i and w^f constitute an equilibrium outcome.

Let $\mathcal{O}_f$ and $\mathcal{O}_i$ be the sets of agents who optimally choose to be in the formal or informal sector, respectively. A stationary equilibrium for this economy consists of an invariant distribution h of the state variables ω ; a set of policy functions $x(\omega)$; and labor decisions by entrepreneurs $\{n^s(z)\}$ such that, given wages (w^f, w^i), tax rates (τ^c, τ_s^w, τ^y), and the cost of creating a formal firm η ,

- individuals solve their optimization problem, the Bellman equation (1);
- the distribution h is stationary: $h = \mathcal{P}'h$; and
- both the formal and informal labor markets clear.

Note that the capital market does not necessarily clear. Excess capital supply is exported abroad, with no effect on the interest rate, in this small open economy.

3 QUANTITATIVE PERFORMANCE

3.1 Calibration

The model is calibrated to be consistent with a number of features of the Colombian economy. To this end, we divide the parameter vector into two groups. The first group includes preference parameters that are difficult to identify using our data and to which we assign values that are common in the dynamic general equilibrium literature. Accordingly, the period is set to one year so that the discount factor β is equal to 0.972; also, the risk aversion coefficient σ is set to 2. These two values are within the ranges reported in studies of emerging market economies.⁸

Furthermore, the first group takes into account the labor share, so we set μ to 0.06 for both the formal and informal sectors. This parameter value is also standard in the literature. Moreover, the real interest rate is set to 2.5 percent, a value close to the average ex-post consumer price index deflated yield on 3-month U.S. Treasury bills for the 1955-2014 period. This figure is 0.77 percent plus 158 basis points, the latter being the average Colombian emerging markets bond index (EMBI) spread for 2013.

Moving to the policy parameters, we assign both the labor income tax rate τ^y and the payroll tax rate τ_f^y in the formal sector a value of 0.15. Also, for our benchmark calibration, we set the tax rate on profits equal to zero. The first two parameter values are low relative to the actual burden of taxation in Colombia and other Latin American countries (see Granda and Hamann, 2015, and Ulyssea, 2018), and their choice is arbitrary. We perform experiments changing these tax rates in a subsequent section.

The second group of parameters comprises the per-period fixed costs of formal operation η , the minimum wage w_{min} , the formal sector productivity advantage Z^f , and two parameters that characterize the variance and persistence of individual productivity.⁹ These parameters are calibrated jointly to replicate the patterns of cross-sectoral occupations and earnings inequality computed for 2013 from two different data sources: the ELCA survey mentioned above and the National Household Survey (GEIH, for its initials in Spanish).¹⁰ Specifically, the moments targeted are as follows:

- (i) the share of entrepreneurs;
- (ii) the share of informal entrepreneurs;
- (iii) the share of informal workers;
- (iv) the ratio of the minimum wage to median earnings; and
- (v) the bottom 40 percent of the earnings distribution.

Note that by targeting the first three moments, we can also match their complements: the share of workers and the shares of both formal entrepreneurs and workers.

Table 1 summarizes the calibration strategy for all parameters. The calibrated value of the formal sector's fixed cost of operation η is 0.12, somewhat low compared with that found in cross-country studies. In particular, the World Bank's (various years) *Doing Business* project

Table 1
Summary of Calibrated Parameters

Parameter	Value	Description	Source
β	0.972	Discount factor	Dynamic general equilibrium literature
σ	2.000	Risk aversion coefficient	
μ	0.600	Labor share	
R	1.025	Real interest rate	3-month U.S. T-bills + EMBI 2013
τ_f^w	0.150	Payroll tax rate	Arbitrary
τ^y	0.150	Labor income tax rate	
η	0.120	Fixed costs of formal operation	Joint calibration
w_{min}	0.067	Minimum wage	
Z^f	1.300	Formal productivity advantage	
σ_z	0.500	Standard deviation of individual productivity	
ρ_z	0.900	Persistence of individual productivity	

reports an average cost of registering a business in Colombia during the 2004-14 period of 16.5 percent of per capita income.

In contrast, the values of the parameters that characterize the process for individual productivity are relatively higher than the ones used in quantitative models to evaluate the misallocation costs of financial constraints. More specifically, these values are well above those of Lopez-Martin (2019), who takes the median of the cross-country estimates for these two parameters reported in Asker, Collard-Wexler, and Loecker (2014) for a large set of emerging economies.

3.2 Model Assessment

To evaluate the performance of the model economy, Table 2 displays some distributional statistics in stationary equilibrium, allowing for comparison with the ones obtained from the data. For the computations, informality is defined by the absence of contributions to the social security system, be it a healthcare or pension scheme.¹¹

In the first panel, we report some statistics for occupations. Note that both the ELCA and the GEIH encompass the whole range of occupational statuses. Thus, to compute the occupational and sectoral compositions so as to make moments from the benchmark economy comparable with those of the data, we have taken from the surveys data pertaining only to workers and entrepreneurs. It can be seen that the model does a good job in replicating the observed distribution of these two occupations across the formal and informal sectors.

The other panels report computations involving the minimum wage as a fraction of both mean and median earnings (second panel) and statistics for earnings shares held by all individuals in the corresponding quintile (third panel). Comparing the model with the data, it can be confirmed that the targeted moments are well matched. Specifically, the benchmark

Table 2
Distributional Statistics

	Data		
	ELCA	GEIH	Model
Occupations (percent)			
Workers	60.47	55.70	55.25
Formal	82.60	62.97	71.58
Informal	17.40	37.03	28.42
Entrepreneurs	39.53	44.30	44.75
Formal	39.01	14.44	23.51
Informal	60.99	85.56	76.49
Minimum wage to earnings			
Minimum wage/mean earnings	0.566	0.614	0.494
Minimum wage/median earnings	0.172	0.274	0.289
Earnings distribution (percent)			
0 to 40 percent	13.47	17.66	14.34
40 to 60 percent	15.43	14.34	12.21
60 to 80 percent	19.11	19.34	20.86
80 to 100 percent	51.99	48.67	52.59
NOTE: Data correspond to statistics for 2013.			

economy replicates the ratio of the minimum wage to median earnings and the share held by people in the bottom 40 percent of the earnings distribution relatively closely.

Last but not least, the calibration yields parameter values that allow the benchmark economy to resemble a number of moments that were not targeted explicitly. In the third panel of Table 2, the model is shown to generate earnings shares held by individuals in the middle to top quintiles that are roughly consistent with the evidence obtained from both the ELCA and the GEIH surveys.

4 POLICY EXPERIMENTS

In this section, we use the calibrated model to analyze a number of alternative policy scenarios. These policies traditionally have been proposed to remove barriers to formalization and include reductions in labor taxes, the minimum wage, and fixed costs of formal operation. We focus our attention on the impact on occupational choices across the formal and informal sectors, on the earnings distribution, and on production efficiency (i.e., TFP).¹² Further, we study the implications for social welfare by measuring consumption equivalent variations.

Table 3**Policy Experiments: Removing Barriers to Formality**

	Benchmark	No fixed costs	No minimum wage	No payroll tax	No income tax	No interventions
Occupations (percent)						
Workers	55.25	46.35	55.65	63.67	63.67	50.58
Formal	71.58	78.38	0.00	57.14	100.00	44.69
Informal	28.42	21.62	100.00	42.86	0.00	55.31
Entrepreneurs	44.75	53.65	44.35	36.33	36.33	49.42
Formal	23.51	100.00	24.03	39.78	39.78	100.00
Informal	76.49	0.00	75.97	60.22	60.22	0.00
Minimum wage to earnings						
Minimum wage/mean earnings	0.494	0.438	0.000	0.482	0.476	0.000
Minimum wage/median earnings	0.289	0.305	0.000	0.275	0.268	0.000
Earnings distribution (percent)						
0 to 40 percent	14.34	15.58	12.14	14.05	14.49	14.32
40 to 60 percent	12.21	12.61	12.10	11.97	12.09	13.29
60 to 80 percent	20.86	21.27	21.46	20.96	20.67	21.85
80 to 100 percent	52.59	50.54	54.31	53.02	52.74	50.54
Production efficiency and welfare						
TFP	1.000	1.162	1.000	1.017	1.014	1.147
Welfare gain		0.047	-0.026	0.024	0.051	0.065

Thus, in what follows, we present the average welfare gains from each policy scenario, using the stationary distribution from the benchmark economy.

The first set of experiments we consider pertain to the elimination of each of the government interventions mentioned above (i.e., labor taxes, the minimum wage, and fixed costs of formal operation). These experiments are represented by reducing the corresponding policy parameters (τ^y , τ_s^w , w_{min} , η) from their calibrated values to zero. The results from these reductions are presented in Table 3, wherein, to facilitate comparison, we reproduce the performance of the benchmark economy (second column) and that in which none of the concerned interventions is in place (last column).

As the table shows, the policies that encourage formalization the most are the elimination of the labor income tax and fixed costs of formal operation. These policies increase the appeal of joining the labor force and entrepreneurship in the formal sector, respectively, and thereby make formality an attractive choice. While these effects are in line with intuition, those from eliminating the minimum wage and the payroll tax seem a bit puzzling. Indeed, eliminating the minimum wage, rather than promoting formalization, leads to the opposite effect, as all

Table 4**Policy Experiments: Minimum Wage and Fixed Costs of Formal Operation**

	$\eta = 0$			$w_{min} = 0$		
	$\tau^y = \tau_f^w = 0.15$	$\tau_f^w = 0$	$\tau^y = 0$	$\tau^y = \tau_f^w = 0.15$	$\tau_f^w = 0$	$\tau^y = 0$
Occupations (percent)						
Workers	46.35	50.57	47.18	55.65	63.67	55.65
Formal	78.38	58.40	100.00	0.00	0.00	40.83
Informal	21.62	41.60	0.00	100.00	100.00	59.17
Entrepreneurs	53.65	49.43	52.82	44.35	36.33	44.35
Formal	100.00	100.00	100.00	24.03	39.78	24.03
Informal	0.00	0.00	0.00	75.97	60.22	75.97
Minimum wage to earnings						
Minimum wage/mean earnings	0.438	0.405	0.408	0.000	0.000	0.000
Minimum wage/median earnings	0.305	0.859	0.776	0.000	0.000	0.000
Earnings distribution (percent)						
0 to 40 percent	15.58	15.54	16.09	12.14	12.04	12.41
40 to 60 percent	12.61	12.89	12.64	12.10	12.07	12.25
60 to 80 percent	21.27	21.44	21.23	21.46	21.42	21.51
80 to 100 percent	50.54	50.13	50.04	54.31	54.48	53.82
Production efficiency and welfare						
TFP	1.162	1.147	1.159	1.000	1.017	1.000
Welfare gain	0.047	0.075	0.071	-0.026	0.000	-0.026

workers become informal. As for eliminating the payroll tax, it does not imply qualitative changes in the sectoral and occupational compositions relative to the benchmark economy.

Furthermore, eliminating the fixed costs of formal operation turns out to be the policy that generates higher redistribution: It leads to the greatest decrease in the share of earnings held by the top 20 percent and to the greatest increase in the share held by the bottom 40 percent of income earners. In contrast, the policy that translates into higher earnings concentration is the elimination of the minimum wage, which generates the greatest increase in the share held by the top quintiles and the greatest decrease in the share held by the lowest ones. This result is to be expected given that removal of the minimum wage leaves wages entirely at the mercy of market forces and thereby ends the lower bound on earnings for low-productivity individuals.

Moreover, eliminating the fixed costs of formal operation is the most beneficial policy in terms of production efficiency, whereas eliminating the minimum wage is the least. Table 3 shows that eliminating the fixed costs of formal operation generates the greatest increases in

TFP, while eliminating the minimum wage results in little to no change. In the former case, the TFP gains reflect variations in the occupational composition: All firms choose to operate in the formal sector, which by assumption is more productive, thus leading to an increase in aggregate efficiency. Similarly for the latter case, lack of variation in the entrepreneurial shares when all workers join the informal sector turns out to be inconsequential for overall productivity.

Note that all these effects take place whether removal of the interventions is implemented in isolation or combined with other policies such as payroll and/or income tax cuts (Table 4). In this regard, Table 3 suggests that increases in production efficiency are higher when the fixed costs of formal operation are dropped solely (third column) than when none of the policy distortions is in place (last column). Further, removing all the barriers to formality considered in this article does not necessarily imply the highest TFP increases.

As for social welfare, Table 4 shows that eliminating for formal firms both the fixed costs of operation and payroll tax at the same time leads to the greatest welfare gains. This contrasts with eliminating the minimum wage, from which, either alone or combined with other tax cuts, the welfare gains are negative to null (see Tables 3 to 5). The reason for this paradoxical result has to do with the insurance nature of the minimum wage, which mitigates financial incompleteness especially for low-productivity individuals in the formal sector. Thus, absent this insurance, these individuals are no longer able to smooth consumption and hence there is a welfare loss.

The findings above seem to contradict those from related empirical studies. In particular, Bruhn and McKenzie (2014) show that reforms intended to simplify business registration in developing countries have led to modest results in terms of formalization of microenterprises. Note, however, that the formal sector's fixed costs featured in our model pertain not only to entering but also to staying in this sector. Also, Kugler and Kugler (2009) find that payroll taxes always decrease formal employment or increase informal employment in the presence of downward rigidities from government-mandated minimum wages. The divergence in this case has to do with our model assumption that the minimum wage is not binding, which makes it easier for firms to pass on payroll tax rate changes to workers via wages.

Nonetheless, our results are in line with the model simulations of Ulyssea (2010), who finds that lowering the costs of entry into the formal sector significantly reduces informality and increases average productivity and welfare. In contrast, payroll tax cuts have little impact on employment composition and hence productivity and welfare exhibit only a slight improvement. Somewhat similarly, Ulyssea (2018) shows that reducing formal sector entry costs leads to a substantial fall in the share of informal firms but essentially no increase in the share of labor in the informal sector. However, the general equilibrium effects of this intervention imply reductions in aggregate TFP and welfare, which are the opposite of the findings presented here.

Likewise, our results are certainly in line with those of Araujo and Rodrigues (2016). These authors find that eliminating all distortions in their model can improve aggregate efficiency and reduce inequality considerably, largely because this causes entrepreneurs to switch to the formal sector. While most of the efficiency improvement comes from removing credit frictions, lower tax rates are the main driver of reduced inequality. Yet, contrary to our findings, they find that eliminating only the labor tax reduces inequality but also makes the economy less

Table 5**Further Experiments: Both Minimum Wage and Income Tax**

	$w_{min} = 0, \tau^y = 0$			
	$\tau_f^w = 0.15$	$\tau_f^w = 0$	$\eta = 0.12$	$\eta = 0$
Occupations (percent)				
Workers	55.65	63.67	55.65	46.61
Formal	40.83	44.33	40.83	40.77
Informal	59.17	55.67	59.17	59.23
Entrepreneurs	44.35	36.33	44.35	53.39
Formal	24.03	39.78	24.03	100.00
Informal	75.97	60.22	75.97	0.00
Earnings distribution (percent)				
0 to 40 percent	12.41	12.31	12.41	13.91
40 to 60 percent	12.25	12.21	12.25	12.99
60 to 80 percent	21.51	21.44	21.51	21.86
80 to 100 percent	53.82	54.03	53.82	51.25
Production efficiency and welfare				
TFP	1.000	1.017	1.000	1.161
Welfare gain	-0.026	0.000	-0.026	0.033

efficient. This is because it lowers gross wages, thereby inducing some entrepreneurs to migrate to the informal sector, where technology is less productive.

Consistent with the analysis above, it is worth noting that the only policy that results in full formalization, meaning that all workers and entrepreneurs join the formal sector, is one that eliminates both the labor income tax and the fixed costs of formal operation (fourth column of Table 4). This policy scenario also generates the greatest redistribution of earnings of all the scenarios considered in this article. Moreover, it implies both production efficiency and social welfare gains that are even greater than in the case of no government intervention whatsoever, although these certainly are not the highest among all the policy experiments.

Now, in view of the formalization effects of the policy just considered, we look at the implications of removing government interventions related to the minimum wage and the labor income tax at the same time. The results are displayed in the last column of Table 4 and in Table 5, where it can be seen that the joint removal of these interventions does not lead to full formalization or incentivize all workers to choose informality either. As regards redistribution, efficiency, and welfare, the consequences of eliminating the minimum wage are shown to be greater than those from eliminating only the labor income tax.

Finally, we also consider in Table 5 the impact of eliminating both the minimum wage and the income tax in conjunction with eliminating the payroll tax and the fixed costs of

formal operation. Notice that of these experiments, only eliminating the fixed costs overcomes the impact of eliminating the minimum wage in all respects. This policy generates a pattern of earnings redistribution that, although not as remarkable as that observed in (columns 2 to 4 of) Table 4, is certainly the most prominent of all the scenarios considered therein. Further, this policy generates one of the highest increases in production efficiency as well as nonnegligible welfare gains.

5 CONCLUDING REMARKS

This article examines the aggregate implications of several policies aimed at removing barriers to formality. To this end, we develop a dynamic equilibrium model in which heterogeneous agents choose to work for a wage or operate a technology in the formal or informal sector, based on the costs and benefits associated with such occupational choices. Formality specifically entails compliance with taxes, a minimum wage scheme, and firm operation costs, but also has a productivity advantage from access to external finance and legal contract enforcement mechanisms. Informal activities avoid taxes and regulations without detection and punishment. Our model is calibrated to approximate some features of the Colombian economy in the early 2010s.

The simulation results suggest that a crucial determinant of informality is the magnitude of formal sector operation costs. High operation costs are associated with firm informality, increased earnings concentration, and aggregate efficiency and welfare losses. Further, eliminating the income tax is an extremely effective policy for reducing labor informality, but has no effect on the earnings distribution or total factor productivity. Jointly eliminating these two government interventions generates full formalization and leads to the greatest redistribution of earnings and even greater efficiency and welfare gains than eliminating all barriers to formality.

In stark contrast, eliminating only the minimum wage has strong adverse effects on labor formalization and welfare and is inconsequential for productivity. Also, reducing the payroll tax rate to zero has little impact on the occupational composition (i.e., informality remains nearly unchanged), efficiency, or welfare. When these two policies are combined, the implications on formalization and earnings distribution of eliminating the minimum wage are greater than those from eliminating only the payroll tax. These results are admittedly at odds with related empirical evidence (see, for instance, Bell, 1997, and Kugler and Kugler, 2009), but can be rationalized by the insurance role that the minimum wage plays in the model, which mitigates financial incompleteness for low-productivity workers in the formal sector.

Our approach, however, is not without limitations. Noticeable among these is how the presence of a binding minimum wage might affect the results presented here. This is especially important given the fact that the minimum wage is binding in some countries (see Neumark and Munguia Corella, 2019). Another interesting possibility would be the introduction of auditing and punishment of informal activities. Also, to improve our proposed model, a worthwhile specification would involve financial market incompleteness in the form of collateral constraints. These extensions would certainly make for a richer economic environment,

thus allowing us to shed further light on issues currently at the heart of academic and policy discussions. ■

APPENDIX: A SOLUTION METHOD

The model is solved by a combination of several numerical procedures.¹³ First, we proceed by applying state-space discretization. The process for individual productivity $z \in \mathcal{Z}$ thus is discretized into 15 states following the method proposed by Rouwenhorst (1995). In the same vein, we construct a grid of 200 points for the asset level $a \in \mathcal{A}$. Likewise, there are four occupations $o \in \mathcal{O}$ in the model. Hence, the state space comprises $\mathcal{Z} \times \mathcal{A} \times \mathcal{O}$.

The equilibrium solution involves finding out the informal wage by applying the bisection method to an algorithm solving the dynamic programming problem. Such an algorithm takes a given informal wage to solve the entrepreneurs' static problem, compute individual utility, solve Equation (1) through policy function iteration, and calculate the excess demand for labor in the formal sector under the resulting stationary occupations. The bisection technique permits ascertaining the wage rate that nullifies the excess demand.

Once equilibrium convergence is reached, the ergodic distribution of combined assets and occupations is obtained. Then we compute the stationary values of the endogenous variables.

NOTES

¹ Designed and carried out every three years by the Universidad de los Andes, this survey collects information from approximately 10,800 households representative of low- and middle-income socioeconomic groups at the national level as well as for five geographical regions in Colombia. For methodological details, see Bernal et al. (2014). Our analysis focuses on the second wave of the survey, conducted in 2013, and is based on responses by individuals from the urban module.

² The ELCA surveyors ask respondents their reasons for saving. We categorize the responses into six motives as follows: (i) retirement (for the future and old age); (ii) precaution (for unexpected events); (iii) asset purchases (to buy a house, a car, or other assets); (iv) to start a business (to start up or invest in a business); (v) education (to pay for their own education or that of their children); (vi) other planned expenses (for entertainment and recreational purposes); and (vii) other motives (for reasons different from the above).

³ To measure the extent of informality, we apply the legal definition according to which the informal sector encompasses those individuals who do not make contributions to the social security system (i.e., either healthcare or pension schemes). However, to check the robustness of our computations, we use several specific defining criteria for firm and labor informality as presented in Granda and Hamann (2015, Appendix B). In all cases, results are similar to the ones using the legal definition.

⁴ Some estimates suggest that the informal sector in Colombia comprised about 37 percent of GDP and 74 percent of the labor force in the early to mid-2000s (see, respectively, Schneider, Buehn, and Montenegro, 2010, and Bernal, 2009).

⁵ Note that since both formally and informally produced goods are identical, they must have the same price in equilibrium, $q_t^f \equiv q_t^i, \forall s$. For simplicity, this price is imposed along the solution and normalized to unity.

⁶ Also, this characterization might reflect the fact that the informal sector is typically unproductive compared with the formal sector (see La Porta and Shleifer, 2014).

⁷ Time- t subscripts are omitted hereafter, with the prime symbol (') denoting a variable next period.

- ⁸ Note that the chosen value for the discount factor also allows a well-defined ergodic distribution of assets for all four occupations.
- ⁹ The process for individual productivity is discretized into 15 states using the Rouwenhorst (1995, Appendix 6) method. We assume that the probability of staying in the lowest state and the probability of remaining in the highest state are both equal.
- ¹⁰ The GEIH survey collects demographic and socioeconomic data from households in 13 metropolitan areas on a monthly basis. Our analysis makes use of seven specific chapters of the survey: Chapters B and C, about the dwelling and the households living therein; Chapter E, about people's general characteristics; Chapter F, compiled social security (health) information; Chapter G, about education; Chapter I, data about people employed; and Chapter N, which includes information about other income sources. Also, following Mondragón-Vélez, Peña, and Wills (2010), we restrict the sample to those 15 to 70 years of age that report between 16 and 84 weekly hours worked on the main job. For further details, see Granda and Hamann (2015, Appendix B).
- ¹¹ See note 3.
- ¹² Note that TFP has been normalized to 1 in the benchmark model.
- ¹³ The computational implementation of the solution is based in part on some routines contained in the CompEcon toolbox (see Miranda and Fackler, 2002). The Matlab code files are available upon request.

REFERENCES

- Antunes, Antonio R. and Cavalcanti, Tiago V. de V. "Start Up Costs, Limited Enforcement, and the Hidden Economy." *European Economic Review*, 2007, 51(1), pp. 203-24; <https://doi.org/10.1016/j.eurocorev.2005.11.008>.
- Araujo, Julia P. and Rodrigues, Mauro. "Taxation, Credit Constraints and the Informal Economy." *EconomiA*, 2016, 17(1), pp. 43-55; <https://doi.org/10.1016/j.econ.2016.03.003>.
- Asker, John; Collard-Wexler, Allan and De Loecker, Jan. "Dynamic Inputs and Resource (Mis)Allocation." *Journal of Political Economy*, 2014, 122(5), pp. 1013-63; <https://doi.org/10.1086/677072>.
- Bell, Linda A. "The Impact of Minimum Wages in Mexico and Colombia." *Journal of Labor Economics*, 1997, 15(3), pp. S102-35; <https://doi.org/10.1086/209878>.
- Bernal, Raquel. "The Informal Labor Market in Colombia: Identification and Characterization." *Desarrollo y Sociedad*, 2009, 63(1), pp. 145-208; <https://doi.org/10.13043/dys.63.4>.
- Bernal, Raquel; Cadena, Ximena; Camacho, Adriana; Cárdenas, Juan Camilo; Fergusson, Leopoldo; Ibáñez, Ana María; Peña, Ximena and Rodríguez, Catherine. "Encuesta Longitudinal Colombiana de la Universidad de los Andes - ELCA 2013." Documento CEDE 2014-42, 2014.
- Bruhn, Miriam and McKenzie, David. "Entry Regulation and the Formalization of Microenterprises in Developing Countries." *World Bank Research Observer*, 2014, 29(2), pp. 186-201; <https://doi.org/10.1093/wbro/lku002>.
- Dabla-Norris, Era; Gradstein, Mark and Inchauste, Gabriela. "What Causes Firms to Hide Output? The Determinants of Informality." *Journal of Development Economics*, 2008, 85(1-2), pp. 1-27; <https://doi.org/10.1016/j.jdevco.2006.05.007>.
- D'Erasmus, Pablo N. and Moscoso Boedo, Hernan J. "Financial Structure, Informality and Development." *Journal of Monetary Economics*, 2012, 59(3), pp. 286-302; <https://doi.org/10.1016/j.jmoneco.2012.03.003>.
- Granda, Catalina and Hamann, Franz. "Informality, Saving and Wealth Inequality in Colombia." IDB Working Paper 575, Inter-American Development Bank, 2015; <https://doi.org/10.2139/ssrn.2581400>.
- Kugler, Adriana and Kugler, Maurice. "Labor Market Effects of Payroll Taxes in Developing Countries: Evidence from Colombia." *Economic Development and Cultural Change*, 2009, 57(2), pp. 335-58; <https://doi.org/10.1086/592839>.
- La Porta, Rafael and Shleifer, Andrei. "Informality and Development." *Journal of Economic Perspectives*, 2014, 28(3), pp. 109-126; <https://doi.org/10.1257/jep.28.3.109>.
- Lopez-Martin, Bernabe. "Informal Sector Misallocation." *Macroeconomic Dynamics*, 2019, 23(8), pp. 3065-098; <https://doi.org/10.1017/S1365100517001055>.

Granda and Hamann

- Miranda, Mario J. and Fackler, Paul L. *Applied Computational Economics and Finance*. The MIT Press, 2002.
- Mondragón-Vélez, Camilo; Peña, Ximena and Wills, Daniel. "Labor Market Rigidities and Informality In Colombia." *Economía*, 2010, 11(1), pp. 65-95; <https://doi.org/10.1353/eco.2010.0009>.
- Neumark, David and Munguia Corella, Luis Felipe. "Do Minimum Wages Reduce Employment in Developing Countries? A Survey and Exploration of Conflicting Evidence." NBER Working Paper 26462, National Bureau of Economic Research, 2019; <https://doi.org/10.3386/w26462>.
- Perry, Guillermo E.; Maloney, William F.; Arias, Omar S.; Fajnzylber, Pablo; Mason, Andrew D. and Saavedra-Chanduvi, Jaime. *Informality: Exit and Exclusion*. The World Bank, 2007; <https://doi.org/10.1596/978-0-8213-7092-6>.
- Rouwenhorst, K. Geert. "Asset Pricing Implications of Equilibrium Models of Business Cycles," in Thomas F. Cooley, ed., *Frontiers of Business Cycle Research*. Princeton University Press, 1995, pp. 294-330.
- Schneider, Friedrich; Buehn, Andreas and Montenegro, Claudio. "New Estimates for the Shadow Economies All Over the World." *International Economic Journal*, 2010, 24(4), pp. 443-61; <https://doi.org/10.1080/10168737.2010.525974>.
- Ulyssea, Gabriel. "Regulation of Entry, Labor Market Institutions and the Informal Sector." *Journal of Development Economics*, 2010, 91(1), pp. 87-99; <https://doi.org/10.1016/j.jdevco.2009.07.001>.
- Ulyssea, Gabriel. "Firms, Informality, and Development: Theory and Evidence from Brazil." *American Economic Review*, 2018, 108(8), pp. 2015-47; <https://doi.org/10.1257/aer.20141745>.
- World Bank. *Doing Business*. Various years; <http://www.doingbusiness.org/>.