Published Quarterly by the Federal Reserve Bank of Minneapolis

December 2011

The Region

Esther Duflo

The Value of Outreach

Implications of Innovation Policy

The Measure of Inflation

Social Insurance—A Forum

Research Digest: International Recession

Sargent & Sims Nobel Laureates

Volume 25 Number 4
December 2011
ISSN 1045-3369

The Region

2 Top of the Ninth
The Value of Outreach
Narayana Kocherlakota

4 Policies to Stimulate Innovation
What impact do innovation policies have on the macroeconomy?
Andrew Atkeson and Ariel T. Burstein

12 Interview with Esther Duflo
Douglas Clement

23 Nothing but Net
At Heller-Hurwicz Forum on Social Insurance, a synergy of research and policy
Douglas Clement

28 Taking the Measure of Prices and Inflation
A century of evolution has greatly improved price indexes. But work continues
Phil Davies

39 Research Digest
Not-so-great expectations
A global recession spurred by lost confidence in asset values
Phil Davies

46 Thomas Sargent and Christopher Sims
2011 Nobel Laureates

The Region
Federal Reserve Bank of Minneapolis
P.O. Box 291
Minneapolis, MN 55480-0291

Email: letters@mpls.frb.org
Web: minneapolisfed.org

Cover photograph by Peter Tenzer

The Value of Outreach

Lessons from a trip to the Williston oil basin

Narayana Kocherlakota
President
Federal Reserve Bank of Minneapolis

One thing that a Federal Reserve Bank president quickly learns to appreciate is the fundamental diversity of the U.S. economy by types of industry and by geography. Certainly every new president understands that our nation's $14 trillion economy is heterogeneous across sectors and regions, but that diversity is quickly brought home when one hears reports from members of the Bank's board of directors, advisory councils and others around the district. While general trends impact all sectors and regions in a national economy, these trends are shaped by ups and downs in a myriad of businesses located everywhere from rural areas to metro industrial parks to inner-city neighborhoods. When it comes to economic performance, there are always outliers on the upside and downside. Even so, sometimes there are cases that lie so far outside the norm that they attract extraordinary attention. North Dakota is one such case.

By now, most everyone who pays attention to business news has heard about the economic boom that is occurring primarily in western North Dakota, where most of the Williston oil basin resides, and also in extreme eastern Montana. Media outlets from cable news channels to national newspapers to web-based news services have carried breathless reports about the demand for workers, the high salaries, the traffic jams and, yes, the so-called man camps. As president of the Federal Reserve Bank of Minneapolis, I have been hearing these reports since I took office in late 2009. Members of our board of directors and our advisory councils have regaled us with tales of unprecedented economic growth while much of the rest of the country was still mired in the aftereffects of the worst economic slump since the Great Depression.

However, hearing or reading about this phenomenon is one thing; experiencing it is another. I recently had the opportunity to visit the "oil patch" region as part of a special meeting of our Helena Branch's board of directors in Sidney, Mont., which lies just west of the North Dakota border. Sidney is just a short drive—about 50 miles—from Williston, N.D. But one of the first things you notice on entering the oil patch is that you have to change your definition of "short drive."

Just like metropolitan areas, where commuting time is determined by congestion and not distance, the same is true of the highway connecting Williston and Sidney. The traffic jams begin in the city of Williston and extend along the rural stretches of

highway, as numerous large trucks and pickups that arrived to supply the oil industry compete with existing agricultural and local traffic. Road repair is a constant, as the heavy trucks take their toll. The rolling plains have a new visual landmark—natural gas flares that burn like eternal flames at the tops of oil wells, visible for miles at night.

Within the communities of Sidney and Williston, as well as other small towns in the region, the most common sign on a business is "Now Hiring." Billboards offer signing bonuses and good benefits if workers sign up, and radio stations carry a stream of ads hoping to lure workers. Business owners, especially service industries, tell stories of having to reduce hours or even eliminate shifts for lack of staff. And while much of the rest of the country continues to struggle with a decrease in residential housing prices, this is certainly not the case in the oil patch, where home prices are rising in an area that extends all the way to Minot and even Bismarck, as well as in the many small towns along the way.

I could go on, but as I indicated, there are plenty of stories out there about this extraordinary growth and the impact on the communities and the state, and we have provided a number of them—along with a lot of data—in our regional business publication, the *fedgazette*, and on our website, minneapolisfed.org. We will continue to monitor this regional phenomenon as it develops over the years.

Earlier I described the sudden growth in business activity in the Williston oil basin as a "boom," and that is a term that many local officials are trying to retire; some believe this remarkable growth as an industry will be around for decades to come. Advances in oil extraction technology mean that the billions of barrels of oil (experts disagree on how many billions, but all say the numbers are large) lodged within a shale deposit known as the Bakken can now be removed if oil prices remain above a certain number, roughly in the range of $50 to $60 per barrel. Having taken a tour of an oil rig near Williston, with its high-tech engineering applications, I find it easy to imagine that ongoing technological improvements will continue to reduce that threshold price.

Does this mean that North Dakota's oil boom will sustain its current pace and become a long-term growth industry? Maybe so, but many of the locals are circumspect. They have been down this road before, and they have long memories. An oil boom in the late 1970s and early 1980s ended abruptly, and many local businesses went bust. In scenes that foreshadowed the real estate splurge and collapse 20 years later, some oil patch towns were left with unfinished subdivisions and vacant commercial spaces that sat idle for years. With this in mind, many local businesses, bankers and community leaders are leery of a repeat performance, and so communities are struggling with how to manage this growth. Of course, problems associated with growth would be considered a luxury in many parts of the country, and North Dakotans know that they are living a charmed existence. In the many conversations I have had with North Dakotans on their economic good fortune, there is one common refrain: The state knows that it is the beneficiary of lucky circumstances—including the recent strength in agricultural prices—and luck doesn't last forever. They have no desire to repeat past mistakes.

I have made a number of trips to North Dakota and Montana since becoming president of the Federal Reserve Bank of Minneapolis and have discussed the impact of the oil boom on the state's economy with a number of people, but experiencing this phenomenon in person has helped me to better understand these reports and certainly has helped me wrap a story around all of the data. I have written in this space before about the importance of public outreach for the Federal Reserve, and this is especially true of Federal Reserve Bank presidents, who manage institutions that are not well understood by much of the American public. As one of those presidents, I have a responsibility to communicate with members of the public, especially people in the Ninth Federal Reserve District, and I take this responsibility seriously. My trips around the Ninth District entail public speeches and meetings with business owners, community leaders, bankers and local media, as well as informational tours. In the coming year, I plan to introduce more public meetings into my trips and to add even more discussion about the work of the Federal Reserve. I very much look forward to these visits, not only to inform people about the Federal Reserve, but also—as my recent trip to the Williston oil patch attests—to learn firsthand about the Ninth District's economy. R

Policies to Stimulate Innovation

How effective are policies to encourage investment in innovation by firms, and what impact do they have on the macroeconomy?

Andrew Atkeson
UCLA and Federal Reserve Bank of Minneapolis

Ariel T. Burstein
UCLA

Introduction[1]

National policymakers have long been interested in technological innovation and its potential contribution to economic growth and improved well-being. The Obama administration has embraced innovation as "the foundation of American economic growth and national competitiveness." In launching the "Strategy for American Innovation" in November 2010, the president remarked, "[T]he key to our prosperity … as it has always been—will be to compete by developing new products, by generating new industries, by maintaining our role as the world's engine of scientific discovery and technological innovation."[2]

Policies to encourage innovation by firms include government funding for research and development (R&D), direct and indirect subsidies, tax credits and other tax benefits, such as deductibility of research expenses. Other policies not typically thought of as aimed at stimulating innovation, such as the corporate profits tax, also impact firms' decisions to innovate. But to channel support effectively, policymakers need to know which policies are most successful in spurring innovation at companies, given their fiscal cost to taxpayers, and to what extent the firm-level innovation induced by these policies truly generates broader economic growth. Also important to policy: What factors influence the effectiveness of innovation subsidies in promoting economic well-being over the long term?

In particular, the idea that innovative activity by firms has "spillovers" that promote the wider diffusion

Economic Policy Papers *are based on policy-oriented research by Minneapolis Fed staff and consultants. The papers are an occasional series for a general audience. Views expressed are those of the authors, not necessarily of others in the Federal Reserve System.*

ABSTRACT

National policymakers have long been interested in technological innovation by firms and its potential contribution to economic growth and improved well-being. Policies to encourage innovation by firms include government funding for research and development, direct and indirect subsidies, tax credits and other tax benefits such as deductibility of research expenses. Other policies such as the corporate profits tax also impact firms' decisions to innovate. Which policies are most successful in spurring innovation at companies, given their fiscal cost to taxpayers? To what extent does the firm-level innovation induced by these policies truly generate broader economic growth?

This policy paper seeks to provide insight into key considerations in innovation policy. The overarching issue is: How do policies that affect firms' innovation costs and benefits impact aggregate innovation activity, output, productivity and welfare?

We establish a benchmark model of innovation that provides a straightforward procedure for estimating relative magnitudes of long-run macroeconomic impact of a range of innovation policy options. The procedure gauges approximate impact of two innovation policies on macroeconomic outcomes quite simply, through computing and comparing the government's fiscal expenditure on these two policies. Two innovation policies have approximately the same impact on aggregate innovation, output and productivity in the long run if they have the same fiscal impact on taxpayers.

The response of economic welfare and GDP over the long run to changes in innovation policy is highly sensitive to the size of innovation spillovers; welfare gains

Continued on page 6

Innovation

How do policies that affect firms' innovation costs and benefits impact aggregate output, productivity and welfare?

Andrew Atkeson

Ariel T. Burstein

ABSTRACT from page 4

could vary between virtually no change and a 50 percent increase in equivalent consumption, depending on spillover size.

Unfortunately, we cannot accurately measure these long-run effects without accurate estimates as to the magnitude of innovation spillovers. Results from our model indicate, however, that even under ideal conditions, it should be very difficult to measure spillovers using data on medium-term response of the macroeconomy from changes in innovation policy. That is, evidence from the medium term is not likely to help differentiate long-run effectiveness because all policies have similar medium-term outcomes regardless of the size of spillovers.

The clearest policy implication of our research is that to the extent that policymakers choose to subsidize innovative activity by firms, they should consider the full set of tax and regulatory policies that impact aggregate innovation through firm profitability. Taxing corporate profits or enacting regulations that make it more costly for firms to start up or operate has a significantly negative influence on innovation, undercutting the stimulative impact of R&D subsidization. The net effect may be to depress, rather than encourage, innovation by firms.

(intentional or inadvertent) of new innovations and knowledge created at just one location, firm or industry is central in justifying government subsidies for innovation. As a result, we want to know how important these spillovers are for the economy as a whole.

Economists who study these issues have generally followed two paths. The first is an effort to understand the impact of policy on innovation decisions taken *by individual firms*—the companies that develop new products and services or improve methods of production or service delivery. More relevant for overall economic well-being, the second looks at the *macroeconomy*, seeking to measure policy impact on a national level: To what extent do policies to encourage innovation generate broad economic growth? For the most part, these approaches have rarely intersected, leaving a significant gap in our knowledge of the mechanisms through which policy initiatives may or may not improve the economic well-being (or welfare) of Americans.

This policy paper reports on our effort to fill that gap by combining these two perspectives, the micro and the macro, thereby providing greater insight into several key considerations in innovation policy. (See "Aggregate Implications of Innovation Policy," Minneapolis Fed Staff Report 459, June 2011, online at minneapolisfed.org.) The overarching issue is: How do policies that affect firms' innovation costs and benefits impact aggregate output, productivity and welfare?

To answer this question, we have designed an economic model that is sufficiently detailed to capture the dynamic decisions of individual firms in response to innovation policy changes, yet is still mathematically manageable, allowing us to aggregate these many firm-level decisions and thereby gauge overall policy impact at the economywide (or "aggregate") level on output, productivity and economic well-being.

We break our analysis into two parts. In the first, we study which policies are most efficient in the long run in balancing their fiscal cost to taxpayers against their benefits in stimulating overall innovative activity by firms, including both those firms that are already operating when the policy is put in place and those that will enter under the new policy regime.

In the second part, we study how a policy-induced increase in innovative activity by firms impacts aggregate output, productivity and welfare (taking into account the fiscal cost to taxpayers of the policies used to stimulate that innovation) over both the long run and a medium-term horizon of 15 years.

Our research gives new answers to both of these questions.

Consider first the balance between the fiscal cost of various innovation policies and their effectiveness in stimulating innovative activities by firms. The standard analyses of fiscal efficiency of innovation policies attempt to fathom the many intricate details of the effects of the new policies on individual firms' decisions about changing investments.[3] By contrast, in our research, we embed a model of firms' innovation decisions in an overall model of the macroeconomy and show that with such a model, we can estimate the policy effectiveness of stimulating innovative activity simply by calculating the approximate impact of that policy on the profitability of new firms that might enter under this new policy.

Moreover, our results imply that, under some conditions, a broad set of innovation policies are all

equally efficient: Two policies have the same impact on aggregate innovation, output and productivity in the long run if they have the same fiscal impact on taxpayers. These results provide a simpler procedure for evaluating the effectiveness of innovation policy and other government efforts to stimulate innovative activity and suggest that policies currently in place to stimulate innovation might also be dramatically simplified.

To gauge policy strength in encouraging innovative investments by firms, we focus on the size of policy impact on the potential profitability of new firms. By doing so, our research implies that innovation subsidies and tax preferences are only part of a much broader set of government policies with both positive and negative effects on firms' incentives to innovate. In particular, the negative impact of the corporate profits tax on incentives to innovate through its impact on potential profitability of new firms[4] may very well undo, at the aggregate level, much of the benefit of current direct federal support for R&D. In this respect, our research indicates that in terms of their effect on innovation investment, the current mix of federal subsidy and tax policies may negate with one policy the impact of others.

On the second question of how a policy-induced increase in innovative activity by firms affects long-run aggregate output, productivity and welfare (taking into account the transition from the status quo to the long run), our research indicates that it may be very difficult to reach definitive conclusions, given available data. Our model predicts on the one hand that a policy-induced change in innovative activity by firms may have a very large impact on output and productivity in the long run and on welfare, particularly if the spillovers from innovative activity are large. On the other hand, if spillovers are small, this may not be the case (and policies to stimulate innovative activity may not raise economic well-being).

Our research indicates that analysts may not be able to distinguish between those divergent long-run outcomes because our model's predictions for the macroeconomic response to an innovation policy over a reasonable time horizon, such as 15 years, look quite similar whether spillovers are very large or very small. Our model's simulations of the economy's medium-term response to a significant increase in innovation subsidies suggest that analysts working with real-world data would have difficulty obtaining reliable estimates of the magnitude of innovation spillovers for the economy as a whole and hence the implications of actual policy changes for welfare.[5]

Our research approach

To analyze the micro- and macroeconomic dynamics of innovation policy, we've built an economic model that is rich, yet tractable. By this, we mean that it combines the fundamental and detailed elements of innovation processes at a company level, but nonetheless allows us to generate estimates of the overall national economic impact of these firm-by-firm decisions as influenced by changes in government innovation policy. This policy paper describes our model and research in broad terms, shares the analytical and quantitative insights we've gained and then discusses implications of these findings for both research and policy.

Our research approach and this paper consist of several steps.

- First, we build a model that enables us, with a two-stage procedure, to assess the impact of changes in firm-directed innovation policies on macroeconomic output, productivity and economic well-being.
- We then use this procedure to establish several analytical results about the long-run response of a macroeconomy, through its microeconomic units, to innovation policy change.
- These results allow us, in a third step, to assess the relative and absolute size of the medium-run and long-run macroeconomic impact generated by several distinct real-world innovation policy options.
- These findings imply several directions for future policy, discussed briefly in a final section.

Our model

We use a dynamic general equilibrium model common to macroeconomic research. It includes house-

holds that work and consume according to their preferences and budget constraints, firms that invest and produce with specified technologies and objectives, and a government that has objectives, revenues and expenditures.

To this standard framework, we introduce a number of special features that allow us to analyze macroeconomic (also referred to as "aggregate") implications of innovation policy.

We build a model of monopolistically competitive[6] firms that engage in either *process* innovation, which will increase their productivity (a more efficient assembly line, as a mundane example), or *product* innovation, which enables them to create a new type of product (an iPad, if they're very lucky and skilled). More simply put, firms can invest in R&D to become more productive or expand the range of goods available to consumers.

To consider the impact of policies on the cost of innovative activity, our model includes a *research good* that firms use as an input for innovation. Research goods aren't unlimited. Producing them requires a combination of goods and labor (lab equipment and scientists) along with scientific knowledge that is freely available. Also, most crucially for purposes of our analysis, the production of the research good benefits from *innovation spillovers*—the knowledge and experience that researchers gain through innovation activities that neither they nor their company directly capture (at least financially). Through these spillovers, current innovative investments by firms have an added benefit to society of increasing the productivity of future R&D workers and thus reducing the cost to firms of future innovation.

In our model, we consider the i*mpact of a range of subsidies* financed by taxes collected from households and equaling aggregate fiscal expenditures by government. These subsidies—fairly abstract when we derive analytical results and later made concrete in our quantitative estimates—include a subsidy to variable profits from production, a subsidy to process innovation and a subsidy to product innovation. In addition, firms are taxed on their use of physical capital, essentially a negative subsidy.

While this brief description hardly does justice to a relatively complex model, it provides a sense of the key features that allow us to analyze the impact of innovation policy on both individual firms and the macroeconomy.

We discovered that a relatively straightforward procedure—a two-step algorithm, or sequence of formulas in which results from the first are inputs to the second—would provide approximate estimates of the long-run impact on macroeconomic outcomes of changes in innovation policy and thereby enable us to compare the relative and absolute magnitudes of the impact of various policy alternatives.

A two-step procedure

Detailed examination of the interaction of these features and the more standard variables in our model yields insight into what is (and isn't) fundamental to analysis of the macroeconomic impact of changes in innovation policy. We discovered that a relatively straightforward procedure—a two-step algorithm, or sequence of formulas in which results from the first are inputs to the second—would provide approximate estimates of the long-run impact on macroeconomic outcomes of changes in innovation policy and thereby enable us to compare the relative and absolute magnitudes of the impact of various policy alternatives.

The first step in this procedure is using a basic formula to measure the impact of policy changes on the profits an entrepreneur might expect from starting a new firm.

The second step is to then use the model's macroeconomic structure to infer long-run changes in aggregate output and wages that must result, in general equilibrium, to restore the incentives of entre-

preneurs to create new firms or products in the face of the estimated change in expected firm profitability calculated in the first step.

In other words, the procedure gives us estimates of the new long-term level of macroeconomic outcomes that corresponds to whatever change in firm profits is generated by a new government innovation policy. And it does so without having to fathom the many intricate details of the new policy's effect on firms' decisions about changing investments, hiring, corporate structure and the like. We need only compute how the policy changes firm profitability—a far easier task.

This straightforward procedure (and the reasoning behind it) allows us to analyze more fully the implications of innovation policy changes. We do so in the next section, followed by an examination of the *quantitative* application of the procedure.

Analytical results

A central insight offered by our model and the algorithm just described is that a subsidy to all types of innovative activity has the same impact on macroeconomic outcomes as a direct subsidy to firm profitability. The reasoning is quite intuitive. Subsidizing a firm's innovative activities—in this case, by changing the price of the research good with a uniform subsidy to process and product innovation—lowers its costs, or equivalently, raises profits. Since profits here are the returns to innovation, supporting firm innovation through a subsidy has an identical impact on firm behavior and aggregates as a direct subsidy to firm profits.

We also find that, under some conditions, whether the subsidy is directed toward *process* or *product* innovation makes little difference in computing the effect on the *macroeconomy* as long as the impact on firm profitability is the same; this is because of dynamics that ensure that in macroeconomic equilibrium, with free entry of firms, companies will start up in an industry until doing so would no longer offer profits to entrepreneurs. (A policy directed specifically at either process or product innovation may have a dramatic impact on *firm*-level behavior, however, particularly on the innovative investments of existing firms.)

The zero-profit condition for entrepreneurs considering starting firms in a given industry limits the aggregate response of innovative investments by both existing firms and entrepreneurs contemplating a startup venture. This analytic insight is what is distinctive about our method for measuring the response of firms' innovative investments to a change in policy.

Previous research has often focused on the innovative response of existing firms only and neglected to consider that—in the long run, in general equilibrium—*the zero-profit condition for entrepreneurs creating new products is key* to assessing the overall response of the economy to the policy change. With this analytical insight, we argue that *regardless of how existing firm investments react to specific subsidies, the response of the macroeconomy will be the same.*[7]

In terms of policy, this implies further that, as we alluded to earlier, the details of firms' responses to changes in innovation policy are not of great importance for aggregate outcomes; beyond pure subsidization of profits, there is no special role for innovation policies. An example clarifies the implications of this argument. Consider the current design of the Research & Experimentation Tax Credit. This innovation policy sets out a complex set of rules by which a firm can gain a corporate tax credit for "qualifying research and experimentation expenditures" over and above a defined "baseline amount." The underlying idea is to reward existing firms only for new or incremental investments in innovation and to avoid subsidizing firms for innovation they would have done anyway.

Our research indicates that this policy focus on incremental expenditures at the firm level is misguided, since the impact on existing firms' investment is not the factor that determines the impact of the policy in the long run. Instead, it is the impact of the tax credit on the incentives of entrepreneurs to start new firms or introduce new products. Our results imply that the Research & Experimentation Tax Credit is, therefore, an administratively expensive way of offering a small reward to entrepreneurs who consider starting a new firm and spending money on R&D that qualifies for the credit somewhere down the line as their new firm grows. It would be more straightforward (and more efficient in terms of administrative costs) to subsidize firms in the relevant industry directly.

Quantitative results

These analytical results lead to the question of magnitude. How can we measure the effectiveness of various innovation policies in stimulating innovative investments given their fiscal cost to taxpayers? And how can we measure the impact of this induced innovation on aggregate productivity, output and welfare? We conducted two sorts of quantitative analyses. The first measured the *relative* impact of several innovation policy options. The second calculated the *absolute* size of the economic effect of parallel policy options.

Comparison of relative policy impact

To understand the effect of innovation policy on broad economic growth and welfare and to evaluate the relative efficacy of different policy options, a means of quantifying and comparing financial cause and effect—that is, cost and benefit—is essential. With our model, we show that, to a first-order (or ballpark) approximation, the relative impact of a policy change on firm profitability and on macroeconomic aggregates in the long run is proportional to the impact of the policy change on government fiscal expenditure.

In other words, to compare, roughly, how large an impact alternative innovation policy options will have both at the level of firm profits and on broad economic outcomes in the long run—GDP and productivity—we need only calculate how much that policy costs. The two figures aren't equal, just proportional, and the calculation is only a rough estimate, not a precise figure. But it means that to evaluate the relative merits of alternative policy options, we need only know their fiscal impact; the difficult task of gauging how millions of firms will respond to the policy isn't necessary.

To apply our results to actual policies in the United States, we looked at (1) the Research and Experimentation Tax Credit program, (2) federal spending on research and development and (3) the corporate profits tax. (Beyond the well-understood effects of the corporate profits tax on investments in physical capital, the tax influences innovation decisions in two ways: It affects variable after-tax profits generated from improved products or process, and firms may expense a portion of the cost of innovative activity and thus deduct these expenses from taxable profits. To the extent that firms are not able to fully deduct all of their expenses for innovation or are not able to carry forward all of the loss when attempts at innovation are unsuccessful, the net effect of the corporate profits tax is to reduce the profitability of starting a new firm or introducing a new product.)

Data from 2007 indicate that fiscal expenditure on the Research and Experimentation Tax Credit was $10 billion. In the same year, federal spending on the five categories grouped into R&D by the Office of Management and Budget—basic research, applied research, development, R&D equipment and R&D facilities—totaled $139 billion. (In contrast, business R&D spending in 2007 was far higher, about $260 billion.) Comparing these two figures (and applying the appropriate discount factor since subsidies to product innovation are paid upfront while variable profits are received in the future), we can clearly see that the *long-term impact on aggregate output of federal R&D spending is far larger than the impact of the Research & Experimentation Tax Credit.*

Calculating the impact of the corporate profits tax—which raised $445 billion in federal revenue in 2007—is more complicated because it depends on parameter values in a quantitative model that affect the physical capital-to-output ratio. But once parameters are chosen, we find that the long-run impact of the corporate profits tax (per dollar of revenue raised) exceeds that of innovation policies (per dollar spent) unless innovation spillovers are very large.

Hence, in our calibrated model, described below, reducing the corporate profits tax to collect $100 billion less in revenue would have a comparable or even larger impact on innovation spending and aggregate output in the long run than increasing either the Research & Experimentation Tax Credit or federal R&D spending by $100 billion, unless spillovers are very high. We thus conclude that *the corporate profits tax may very well be a relatively potent, counterproductive policy in terms of discouraging the long-run accumulation by firms of both physical, tangible capital and intangible capital (that is, patents, trademarks, intellectual property and the like).*

Comparison of absolute magnitude of policy impact

In a second quantitative exercise, we evaluated the absolute magnitude of both the long-run and medium-term impact on the macroeconomy of innova-

We found that none of the subsidies has significant impact on economic welfare if innovation spillovers are small. Output and productivity rise in the long run (and perhaps by a lot), but this increase comes at the cost of inefficiently high investments in innovation and low consumption by households in the transition from the present to the long run.

tion policies after putting some concrete figures into our model, giving it further realism by providing reasonable values for parameters such as the GDP growth rate, interest rate and capital depreciation. With this calibrated model, we measured the absolute magnitude of impact on GDP, welfare, productivity, research intensity and other economywide outcomes of two policies:

(1) A uniform subsidy to innovative activities (meaning that both process and product innovation would receive support).

(2) A subsidy to process innovation only.

In each case, the subsidy represented a fiscal expenditure of 3 percent of GDP, or about $420 billion in 2007 (similar to the revenue raised from corporate profits taxes that year). These are two typical policies aimed at stimulating innovation.[8]

Long-run response

In the long run, we find, innovation policies have an impact on the scale of firms' investments in innovation similar in magnitude to their fiscal impact, both relative to the level of GDP. Specifically, the *research intensity of the economy* (defined as the ratio of firms' spending on innovative activities to GDP) increases by roughly 3 percentage points of GDP in response to a subsidy of 3 percent of GDP. Moreover, this response of firms' innovative activity to innovation policy change is the same in the long run and roughly the same in the medium term regardless of the level of spillovers from innovative activity.[9]

Do these policies aimed at stimulating innovation increase consumers' welfare? The answer to this question is not obvious. At first glance, it appears that such a policy might not increase economic well-being—the taxes that consumers must pay to finance these innovation subsidies are roughly the same as the increase in firms' investments in innovation that result. In the absence of spillovers from firms' innovative activity, a policy of taxing households to pay for firms' investments in intangible capital is not likely to improve households' well-being. In the presence of spillovers, however, such a policy might bring substantial welfare benefits.

Our model confirms this logic. We found that none of the subsidies has significant impact on economic welfare *if* innovation spillovers are *small*. Output and productivity rise in the long run (and perhaps by a lot), but this increase comes at the cost of inefficiently high investments in innovation and low consumption by households in the transition from the present to the long run.

If spillovers are *large*, however, the subsidies have far greater impact on economic well-being. In fact, in this case, innovation subsidies of 3 percent of GDP can bring huge gains for households. The numbers from our model simulations below illustrate this point. We measure improvements in household economic welfare from policy changes by the amount that household consumption would have to be increased each and every year under the old policy to make households as happy as they would be with the consumption they attain under the new policy.

When we set our parameter for innovation spillovers at zero, the impact of the innovation subsidies on welfare is very close to zero—consumers would be just as happy with or without the innovation policy. In contrast, when we set our parameter for innovation spillovers close to its maximum possible value consistent with balanced growth, the impact of innovation policies on welfare is very large. Consumption under the old policy would need to rise by roughly 50 percent every year to attain the same level of household welfare as

Continued on page 43

Innovation from page 11

achieved in the equilibrium with innovation subsidies. Welfare gains like these are why Nobel Prize-winning economist Robert E. Lucas Jr. wrote that once one starts thinking about long-run growth and economic development, "it is hard to think about anything else" (p. 5).[10]

Our results on the long-run impact of innovation policies on aggregate output and productivity are also highly sensitive to our assumption for the parameter governing spillovers from innovative activity. When we set our parameter for innovation spillovers to zero, GDP is estimated to increase by a factor of only 1.03 (that is, by 3 percent) in the long run. In other words, in this case, the subsidies have little impact on either output or welfare.

But when we set the spillover parameter close to the maximum value consistent with balanced growth, the impact on GDP is much larger: It increases by a factor of 9.88 for policy 1 and 8.25 for policy 2. These nearly tenfold changes in GDP are comparable to the growth that the United States experienced from the beginning to the end of the 20th century and are brought about by a substantial, but perfectly feasible, level of innovation subsidies.

Clearly, *our model's implications for the long-run impact of a given change in policies vary tremendously depending on the assumed spillover parameter. If spillovers are large, there is a lot at stake for consumers in getting innovation policy right.*

Medium-term response

Our results on the impact of innovation policies on welfare and on output and productivity in the long run prompt the question: Can we use data on the response of the macroeconomy to changes in innovation policy over the medium term (say, 15 years) to figure out if spillovers from firms' innovative activities are small or large?

There is a large literature that attempts to answer this question, but, as Griliches (1988) and CBO (2005) discuss, the changes in the innovation intensity of the U.S. economy seen in the historical data are relatively small. It is therefore difficult to distinguish the effects on the macroeconomy of such small changes in R&D spending from the effects of all the other major factors at play—education, population growth and international trade, to name a few.

To shed light on the question of whether we might be able to measure economywide spillovers from innovative activity using available data even if we were to observe a large change in the innovation intensity of the economy arising from a change in innovation policy, we examined how the model performed over a shorter time frame, a 15-year medium-term period. The idea here is to understand transition dynamics—between now and the long-term equilibrium, *how* does the economy evolve, and *what factors are important* in that evolution? Again, and for all policies, we use the same subsidy size: a fiscal expenditure of about 3 percent of GDP.

Surprisingly, perhaps, we found that over this time frame, the two innovation policy options have a *similar impact* on economic growth *regardless of innovation spillover size*. In all cases, the cumulative factor increase in GDP in the 15th year is between 1.01 (or 1 percent) with no spillovers and 1.05 (5 percent) with high spillovers. Such small differences in GDP over a 15-year horizon would likely be difficult to discern in real-world data. Therefore, our results indicate that data on the response of GDP to innovation policy changes over the *medium* term will not shed much light on the size of such spillovers, suggesting that estimating policy outcomes over the *long* term will remain difficult, since an accurate measure of spillovers can't be obtained from shorter-term data.[11]

What explains the significance of spillovers for welfare?

The contrast in findings between long- and medium-term significance of innovation spillovers raises the question of why spillovers would have importance on innovation's macroeconomic impact *only* in the long run.

The intuition for this result is simply the idea of compound interest. Over the medium term, innovation policies have a similar impact on GDP growth regardless of the level of innovation spillovers. The real impact of spillovers comes only at longer time horizons. In the absence of spillovers, the boost to growth from innovation subsidies peters out relatively quickly and households are left paying roughly the same amount in taxes as the gain to innovation spending and the

increase in GDP achieved. In contrast, if spillovers are large, the boost to the growth of GDP from increased investments in innovation lasts for a long time, well beyond the medium-term horizon, and innovation spillovers compound over time, bringing large benefits associated with a moderate boost to growth that lasts over 100 years.

Summary and implications for policy

We've established a benchmark model of innovation that provides a straightforward procedure for estimating relative magnitudes of long-run macroeconomic impact of a range of innovation policy options. The procedure gauges approximate policy impact on macroeconomic outcomes quite simply, through computing the government's fiscal expenditure on innovation policies.

The response of economic welfare and GDP over the long run to changes in innovation policy is highly sensitive to the size of innovation spillovers; welfare gains could vary between virtually no change and a 50 percent increase in equivalent consumption, depending on spillover size.

Unfortunately, we cannot accurately measure these long-run effects without accurate estimates as to the magnitude of innovation spillovers. Results from our model indicate, however, that even under ideal conditions, it should be very difficult to measure spillovers using data on medium-term response of the macroeconomy from changes in innovation policy. That is, evidence from the medium term is not likely to help differentiate long-run effectiveness because all policies have similar medium-term outcomes regardless of the size of spillovers.

What does this imply for policy?

The clearest implication of our research is that to the extent that policymakers choose to subsidize innovative activity by firms, they should consider the full set of tax and regulatory policies that impact aggregate innovation through firm profitability. Taxing corporate profits or enacting regulations that make it more costly for firms to start up or operate has a significantly negative influence on innovation, undercutting the stimulative impact of R&D subsidization. The net effect may be to depress, rather than encourage, innovation by firms. R

Endnotes

[1] This paper is based on: "Aggregate Implications of Innovation Policy," Minneapolis Fed Staff Report 459, June 2011. The authors thank Doug Clement for assistance in preparing this text.

[2] See "Strategy for American Innovation: Introduction." Also see Chairman Ben Bernanke's May 16, 2011, speech, "Promoting Research and Development: The Government's Role," for a discussion of the importance of innovation by firms to long-run growth and a summary of the questions regarding the rationale for, the effectiveness of and the impact of federal support for research and development that we address in this policy paper.

[3] Specifically, following the methodology developed by Hall and Jorgenson (1976) for physical capital, a standard approach is to first estimate the impact of a policy change on the "user cost of R&D" and then estimate the elasticity of firms' demand for R&D in response to such a policy-induced change in the user cost of R&D. See Hall and Van Reenen (2000) and CBO (2007) for examples of such analysis.

[4] See Gentry and Hubbard (2000) and Cullen and Gordon (2007) for a discussion of the mechanisms through which the U.S. tax structure reduces the incentives of entrepreneurs to start new firms.

[5] In this sense, our research casts doubt on the methods economists have previously used to measure the relationship between innovative activity by firms and aggregate productivity in the long run. See, for example, CBO (2005) and Hall, Mairesse and Mohnen (2009) for summaries of this research.

[6] A monopolistically competitive market combines characteristics of competition and monopoly. There are many buyers and many firms, with free exit and entry into industries, as under perfect competition. But consumers perceive sufficiently great nonprice differences (branding, for example) among similar products that producers can exercise a degree of control over pricing, as in a monopoly. Brand-name cereals and restaurants are textbook cases; laptop computers might be another example.

[7] Innovation policies in our model do impact the user cost of R&D and do have an impact on the innovative investments by incumbent firms that does depend on the responsiveness of these incumbent firms' innovative investments to changes in the user cost of R&D. This responsiveness, or elasticity, of R&D investments is not of first-order importance, however, in the calculation of how a change in inno-

vation policy affects the expected profitability of a new firm. For example, in calculating the impact on the expected profitability of a new firm from a change in a tax credit for R&D, what is of first-order importance is the change in taxes that a new firm can expect to pay given the investments in R&D that it had planned to undertake before the policy change was proposed. For small changes in policy, the additional accuracy gained by considering the impact on the expected profitability of new firms that arises from considering changes in policy and firms' investments simultaneously is necessarily very small.

[8] In the full paper, we also consider a third policy, a subsidy to physical capital to compare the impact of policies aimed at promoting firms' investment in intangible capital and those promoting investment in physical or tangible capital. We make this comparison to analyze the impact of the corporate profits tax, which is a combination of taxes on firms' profits from intangible and tangible capital.

[9] Our findings here are consistent with those summarized by Hall and Van Reenen (2000) on the effectiveness of fiscal incentives for R&D.

[10] See Lucas (1988).

[11] We note that substantial research has sought to establish a link between research intensity and output or productivity. This research has generally used regression analysis of disaggregated data at the firm or industry level. Unfortunately, this evidence is less than conclusive for answering the questions addressed by this paper.

First, many of these results are driven by long-term differences across firms or industries: Firms and industries that invest more in R&D also appear to have higher levels of productivity. It is not clear, however, how to interpret this observation. Klette and Kortum (2004), for example, argue that it should be accounted for by models with intrinsic factors that vary across firms and industries and that it does not necessarily indicate that a policy of stimulating further R&D would have a substantial impact on the aggregate economy.

Second, even in our model, a policy aimed at stimulating innovative investments by a *select group* of firms or industries can have a large impact in the short term on output and productivity that suggests spillovers are high even if *aggregate* spillovers are absent. It is a simple matter for a subset of firms or industries to invest in innovation and grow at the expense of the other firms or industries in the economy. Therefore, evidence of specific firm or industry responses to policy changes does not necessarily shed light on the central question of the *macroeconomic* response.

References

Bernanke, Ben S. 2011. "Promoting Research and Development: The Government's Role." Speech at the Conference on "New Building Blocks for Jobs and Economic Growth." Washington, D.C., May 16. Online at federalreserve.gov.

Congressional Budget Office (CBO). 2005. "R&D and Productivity Growth." Background Paper, June.

Congressional Budget Office (CBO). 2007. "Federal Support for Research and Development." A CBO Study, June.

Cullen, Julie Berry, and Roger H. Gordon. 2007. "Taxes and Entrepreneurial Risk-Taking: Theory and Evidence for the U.S." *Journal of Public Economics* 91(7/8): 1479-1505.

Gentry, William M., and R. Glenn Hubbard. 2000. "Tax Policy and Entrepreneurial Entry." *American Economic Review* 90 (2): 283-87.

Griliches, Zvi. 1988. "Productivity Puzzles and R&D: Another Nonexplanation." *Journal of Economic Perspectives* 2 (4): 9-21.

Hall, Bronwyn H., Jacques Mairesse and Pierre Mohnen. 2009. "Measuring the Returns to R&D." *National Bureau of Economic Research Working Paper* 15622.

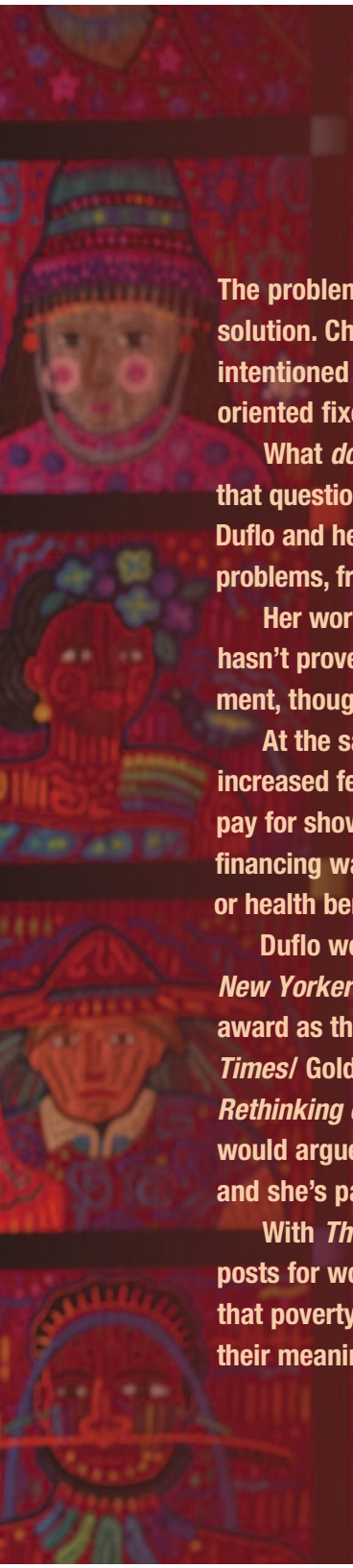
Hall, Bronwyn H., and John Van Reenen. 2000. "How Effective Are Fiscal Incentives for R&D? A Review of the Evidence." *Research Policy* 29 (4/5): 449-69.

Hall, Robert E., and Dale W. Jorgenson. 1976. "Tax Policy and Investment Behavior." *American Economic Review* 57: 391-414.

Klette, Tor Jakob, and Samuel Kortum. 2004. "Innovating Firms and Aggregate Innovation." *Journal of Political Economy* 112 (5): 986-1018.

Lucas, Robert E. Jr. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (1): 3-42.

The tapestry behind Esther Duflo, “Peoples of the World,” was handcrafted by Japanese artist Fumiko Nakayama. It was donated by MIT alumnus Mohammed Abdul Latif Jameel, a major J-PAL funder.



Esther Duflo

The problems of poverty in the developing world are extreme, extensive and seemingly immune to solution. Charitable handouts, massive foreign aid, large construction projects and countless other well-intentioned efforts have failed to alleviate poverty for many in Asia, Africa and Latin America. Market-oriented fixes—improved regulatory efficiency and lower trade barriers—also have had limited effect.

What *does* work? MIT economist Esther Duflo has spent the past 20 years intensely pursuing answers to that question. With randomized control experiments—a technique commonly used to test pharmaceuticals—Duflo and her colleagues investigate potential solutions to a wide variety of health, education and agricultural problems, from sexually transmitted diseases to teacher absenteeism to insufficient fertilizer use.

Her work often reveals weaknesses in popular fixes and conventional wisdom. Microlending, for example, hasn't proven the miracle its advocates espouse, but it can be useful in the right setting. Women's empowerment, though essential, isn't a magic bullet.

At the same time, she's discovered truths that hold great promise. A slight financial nudge dramatically increased fertilizer usage in a western Kenya trial. Monitoring teacher attendance, combined with additional pay for showing up, decreased teacher absenteeism by half in northwest India. Better access to credit for financing water connections in urban Morocco significantly improved family well-being, even without income or health benefits.

Duflo would resist the oxymoronic label, but she is something of a rock star economist: profiled by the *New Yorker* last year, honored as a MacArthur "genius" in 2009, recipient of the 2010 John Bates Clark award as the best economist under 40, and winner—with coauthor Abhijit Banerjee—of the *Financial Times*/ Goldman Sachs "Business Book of the Year" award in November 2011 for *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty.* She wears celebrity awkwardly. The work is important, she would argue, not the individual. But she knows well that her fame, such as it is, helps promotes her cause, and she's passionately devoted to improving the welfare of the poor.

With *The Region*, she discusses strengths and limits of experimental methods, why reserving leadership posts for women makes sense and a future agenda for development economics. Above all, she emphasizes that poverty and its solutions are multidimensional. "If we think of them each as an isolated data point, their meaning is limited," she said of her research results. "But together, they start painting a picture."

Photographs by Peter Tenzer

A DEEPER CONCEPT OF POVERTY

Region: In *Poor Economics* with Abhijit Banerjee, in your 2007 *Journal of Economic Perspectives* piece together and elsewhere, you describe a richer concept of poverty—more nuanced than the traditional concepts of starving masses or Schultz's phrase: "poor but efficient."[1]

Would you start by explaining that deeper concept and then discuss what it means in terms of how to approach solutions to the problems of poverty?

Duflo: The short answer is that there is not one thing. You can't replace any of the clichés by yet another one. It's a very natural thing to do, to try to reduce a problem to a single dimension, and I think that's what people have done. And in a sense, in each of these clichés that has coexisted or existed in succession, there is a certain amount of truth. It is just that you need some combination of them.

For example, I think "poor but efficient" is actually a very useful step in starting to think about the world of the poor—much better than the ideas that existed before. It was a very foundational step to think, "Well, we are just going to consider the poor in the way that we considered anybody else at the time in economic models," which is people who are making rational decisions.

So there is a lot of use to that, but now we can also include what we've incorporated in the last few years in economics as well, which is that people are not always acting fully rationally. Or they are acting rationally, but they don't have full information. Or they have constraints on what they can do because other people lack information about them.

And so, you add this complexity, and then an added layer is the psychological limitations. You recognize that the poor, like any of us, are social beings who exist within a social context. They have friends and values, and things like that, and all of these things have their weight into the way people make decisions. And so sadly, there's no replacing what is there by something else, but it's keeping in mind the whole picture in thinking about how people make decisions.

"Poor but efficient" is a very useful step in starting to think about the world of the poor. … But now we can also include what we've incorporated in the last few years in economics as well, that people are not always acting fully rationally. Or they are, but they don't have full information. Or they have constraints on what they can do. … And so then you recognize that the poor, like any of us, exist within a social context.

Region: That's a lot of heterogeneity, no?

Duflo: Yes, it is a certain amount of heterogeneity, but that doesn't mean it's unpredictable, because there is some logic to this heterogeneity. You can predict on the basis of other things you have seen before. For example, you can predict what's likely to be a constraint, or under what conditions something is likely to be a constraint.

And therefore, you can foresee where policy action might be effective. You might be wrong, of course. You don't want to replace what was there before, either, by saying, "Oh, we do not know." I think we are very much in favor of an analytical approach where you can predict and understand how people behave and why they behave the way they do. It's just that it requires that we keep a lot of threads together.

THE EXPERIMENTAL APPROACH: ADVANTAGES AND LIMITATIONS

Region: You've been a pioneer of experimental research in developing countries. Your work has inspired many others and helped revitalize development economics generally. But as you're well aware, the experimental approach has also been criticized by some economists, for a variety of reasons, including lack of generalizability, ethical considerations, compliance issues and other issues—points James Heckman made 20 years ago.[2]

What are the central advantages of the experimental approach over other methods—observational studies, for instance, or structural estimation techniques? Or is it complementary to those? And the follow-up question is, Which of

the criticisms do you consider well founded, and how do you address them?

Duflo: There are a number of distinct advantages over other methods. One, of course, is the obvious one, which is that running an experiment gives a handle on causality, at least in the particular context in which you run your experiment. When you run an experiment, you modify the conditions in one group and not in another group. Assuming the experiment was well run, you know that whatever different outcomes or behavior you measure are due to the modification of conditions.

One criticism that I don't find useful is, "Oh, but if the experiment is not well run, then that's not true." Well, that's obvious. I don't know how that helps us. That's always the case—true in the lab, true everywhere.

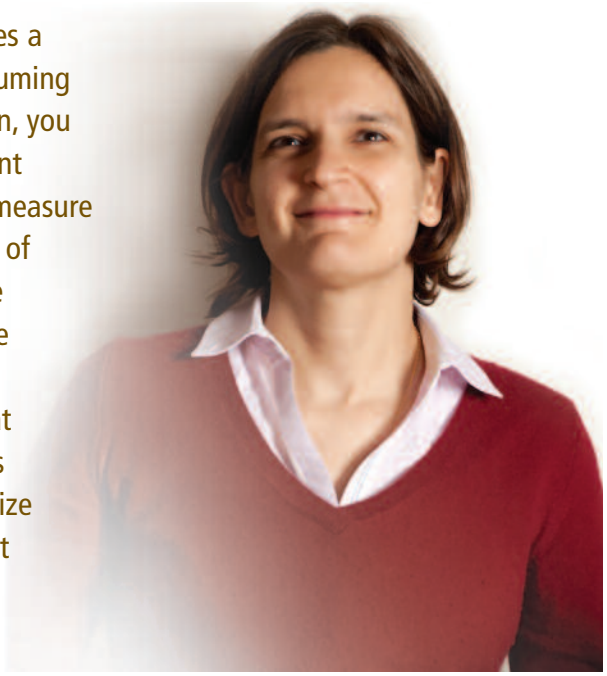
So getting a handle on causality is the first advantage, and it's the one that is easiest to explain to policymakers. If you want to know whether your policy works, that is a very transparent way to do it. And you know it with much more certainty than you would have without an experiment, because usually in the real world when things differ, there are reasons for it, and that possibility prevents you from estimating the causal effects of the policy or intervention.

But there are other advantages, which are more subtle. One is that you can sometimes with experiments estimate things that you just could not estimate in any other way. It's not that you can do it *better* with an experiment; it's that there is *no other way* to get at it.

Region: For example?

Duflo: Suppose you are interested in a range of price elasticity. You might be able to experiment with prices that are just not observed in nature. For example, when people sell things in the market, they don't sell things at the price of zero. [Laughter] You might need an experiment in order to test zero. In fact, that's a point Heckman made a long time ago, saying that that's what experiments should be used for.

Running an experiment gives a handle on causality. … Assuming the experiment was well run, you know that whatever different outcomes or behavior you measure are due to the modification of conditions. … But there are other advantages, which are more subtle. … The only criticism of experiments that I think is really useful is this question of, Does it generalize or not? With the caveat that that question applies pretty much to any method.



Say you are interested in estimating people's response to increasing or decreasing their wage. That response is a combination of an income effect and a substitution effect. In the real world, you can't really distinguish the two, because whenever wages increase, the two things happen.

But in an experimental context, you *can* separate the two. You can give people a bunch of income that doesn't correspond to a wage, which you wouldn't be able to do in a real world setting, so you can estimate the income effects separately from the substitution effect. This was what the negative income tax experiment set out to do, and Heckman was actually quite in favor of this particular use of experiments (in the 1991 article you cite).

A recent example of this is an experiment by Rob Jensen and Nolan Miller, where they look at the effect on consumption of changes in the price of rice.[3] If you decrease the price of rice, will people consume more rice or less rice? In the real world, it's very difficult to know that because whenever the price of rice decreases, that's the result of a combination of supply and demand factors, and isolating variation in the price of rice as purely exogenous is essentially impossible.

So you need an experiment to know, and in fact they found something very interesting when they did this experiment in one place in China where rice is a very important part of the food basket for the poor. And they found that when the price decreased, people ate less rice, not more rice, which means rice is a Giffen good [a product that consumers demand more of as its price rises because the income effect dominates the substitution effect].

Region: I didn't know they existed.

Duflo: Well, exactly. Whether an actual Giffen good exists has been a question since … since [pause]

Region: Giffen himself, I suppose.

Duflo: [Laughter] [Alfred] Marshall brought it up, but he attributed the observation to one Dr. Giffen. And I think this experiment is very fascinating, because you can't investigate it any other way. I

think you can't dispute the fact that rice, in this particular place in China, is a Giffen good.

But then it comes to one of the criticisms: "It doesn't generalize." Yes, it doesn't mean that rice is a Giffen good here in the United States. I'm not interested in that question. But the fact that there is one Giffen good somewhere I think makes this interesting. It is incremental knowledge for how we think about the world and is very, very, very important for what we think about the poor and food. And in particular, in the policy domain, it shows that policies that subsidize the price of staples—which is quite common—might be counterproductive from the view of getting people to eat more. It still might be good for the poor, because they consume a lot of staples, and subsidizing a staple improves their income. But if the objective was to make people eat more, that's not necessarily the way.

That does not mean that it would be true in India, but the very fact that there is this possibility means that we want to investigate this question more. And we can try a similar experiment elsewhere to see in what conditions this will reproduce. With a Giffen good, the advantage is that we have a very established theory that helps us think what's likely to be a Giffen good. It has to be something that is a very big part of the budget so that the income effect is large. And it must be an inferior good.

That gives us a sense of, in another place, how would we go about looking for a good that's likely to have the same characteristics? Maybe there are no Giffen goods here because no goods have those characteristics. But maybe if we went to Ethiopia, it would be whatever is the staple food there. We can see what's the share of this staple in people's budgets and get some idea of what we are looking for.

The only criticism of experiments that I think is really useful is this question of, Does it generalize or not? With the caveat that that question applies pretty much to any method. The only way in which experiments are different is, because there are cases—this is a point Heckman made a long time ago—where the experiment modifies the population that you study because not everybody would even agree to be in an experiment.

For example, not every city accepts to be the site of a job market experiment or not every nongovernmental organization accepts to work with you, so any result you find is specific to the context of people who agree to work with you, and the people who agree to work with you might be different. So I think that's a very relevant point, which is specific to experiments.

But that is different from the point generally made, which is that if I have a result somewhere, it may not apply elsewhere. That applies to just about any result from any research approach. Science makes progress thanks to the interplay of theory and experiment that helps us draw generalized knowledge from individual observations.

The only reason we discuss it with experiments is that we have solved the other problems, so there is more time to discuss that. Until now, there has not really been a problem of worrying whether things were generalizing. Sometimes people use quite subtle sources of variation to identify things they are interested in, so they're looking at very few people. And these very few people are very particular, so we cannot generalize from estimates that are identified from variation that is affecting just 5 percent or 2 percent of the population.

But that being said, it's still the case that the question of generalization applies to any study, and thinking about this question is useful and important. I think the answer is that you'll never be able to interpret a single experimental result, except if it's something like the Jensen result, which is kind of a counterexample, it's showing something quite [pause]

Region: Counterintuitive.

Duflo: Yes, counterintuitive. But if you had done 10 experiments where you had found that these 10 goods are Giffen goods, it doesn't tell you that the 11th is, or isn't, a Giffen good. So in most cases, one single result is not sufficient to reach broad conclusions.

But on the other hand, typically these experiments are informed by a broad theory. Either it's implicit or explicit—but very often explicit—and an experiment is set up in part to test that theory. Or even if it's set up to evaluate a policy that someone is running and wants to evaluate, researchers usually get interested because they can put that in a framework. It helps them test some hypothesis.

That's the theoretical framework which helps make sense of the result of the experiments. In a sense, that's why we wrote this book, *Poor Economics*. That's a little bit of, if you take each of these experiments—and not only experiments; we also have lots of nonexperimental research, there are descriptive results, et cetera—fitting together. If we think of them each as an isolated data point, their meaning is limited. But together they start painting a picture.

Region: Like the Fumiko Nakayama tapestry on P-LAB's wall. [See photo description on page 12.]

Duflo: Yes. They all fit into a greater picture.

Sometimes, maybe, an isolated result is very puzzling. So we just have to set it aside and wait to see how other results will fit with it. Maybe it was a fluke. Maybe things actually continue going in that direction, and then it will push theory to develop an answer to this. The theory changes and then that generates a new wave of experiments. Research moves like that.

If you look at the developments of the last 15 years, I would say it's that process you see happening. As I wrote

in a simple paper, "Poor but Rational?"[4] in some sense, we had done the first phase at that point; we had a lot of experiments and nonexperiments whose results were a bit odd, and we didn't have a framework to think about them. Since then, people have developed much more of a theoretical framework to think of behavioral economics of poverty. And the new wave of experiments helps fill in that framework.

WOMEN'S EMPOWERMENT AND MARKET FAILURE

Region: At the Fed's Jackson Hole symposium in late August, you suggested that reducing a variety of market failures could better ensure that the well-being of the poor improves as nations grow economically—that there would not be a growth/equity trade-off.[5]

You've also written that while development and women's empowerment are reciprocally intertwined, neither ensures the other.[6] In other words, that growth doesn't guarantee gender equity, and empowerment won't improve all aspects of life.

I read both pieces and wondered whether the market failure argument that you apply to income equality and growth is also relevant to gender equality. That is, can policies to reduce market failures better ensure women's empowerment as nations develop?

Duflo: To a point, yes, there are cases where we see that, I think. For example, very few women are elected as policymakers; it could just be that people don't like to be led by a woman. So then there's no market failure. As women, we may not like it, but that's the market equilibrium.

But it could be that it's because people think that women are not going to be good. Or even not that they think that they won't be good, but maybe they are just worried because they have never seen a woman lead, so their priors—that is, their prior beliefs—are very diffused. It seems to them that it is a very risky proposition to elect a woman, because they don't know whether women are good in general, or not so good in general, so there is much more variance in their estimate of how good a woman will be, compared to a man. So they go for a man always just because they have gone for men always, and it's the safe thing to do.

Region: The known quantity is the default.

Duflo: They know how men are, typically. They don't know about women in that position, and you don't want to take risk in politics. That could be very inefficient because it means they never elect women and never find out that women maybe have the same average quality. In that case, they are depriving themselves of half of the pool of capable leaders. And there it's kind of a market failure.

So if you force people to experiment with women and they discover that women are fine, then they start to elect them themselves.

Region: And the forced experiment with women might be through a reservation policy, such as the one you've studied in India.[7]

Duflo: Yes, for example, through reservations. And after reservations go away, they may continue to elect women.

Region: If the results are good—if women are seen to be effective.

Duflo: Yes. If they are not, then they shouldn't be in office. So that's an example where it's a market failure and one that can be addressed by forcing people to experiment. I think that's a much better rationale for reservations than one that is typically made, in terms of outcomes. I think changing perception is a better argument. There is no downside for people to acquire information that they didn't have. Maybe taste discrimi-

It was actually an efficient thing to reserve policy positions for women for five years or 10 years [in India] … just so that people experience the fact of having a woman lead and realize that women are not what they thought. … But I just think that the business case argument should be used when we have evidence for it, and it shouldn't be used when we don't.

nation against women per se had gone away a long time ago, but people still won't elect them just because of this statistical discrimination. I don't think that's true—in our survey in India, people were not shy to admit that they don't like female leaders, but that's possible.

Region: You said, I think, in the women's empowerment paper that to bring about equality might require that policies favoring women should be in place for a long time to come. It's a provocative statement. How have colleagues and policymakers reacted to it?

Duflo: Colleagues are fine with it. Policymakers usually are a bit sad, particularly those who are advocates of women, because the way the case has traditionally been made for empowering women is a business case. To say that you should do it because it is good for everybody.

It's really a whole bunch of arguments, like women will be less likely to be corrupt, they will invest more in girls—you name it. They have better investment opportunities because no one has given them money before. There is a range of claims that people make to say that discriminating in favor of women is the policy efficient thing to do.

I think that's a slightly dangerous case to make, because if you find out eventually that that's not true, it's going to be apparent, and then once the business case disappears—that is, you have problems—people will say, "You fooled us on the business case," and you'll get this backlash.

So I think it's better to call a spade a spade and to say, "Well, if you look at the rich countries, there is still plenty of discrimination against women." So if you care about equality for its own sake, then you might have to continue to help out for a while. We don't know; eventually it might disappear, but it might take some time. We don't know how long because we see we still have discrimination here in the United States, in some domains.

When I started working in this field, there was a range of questions that were all fitting together in a theoretical framework, but empirically, the questions were quite open. … The idea of carefully looking at data while being inspired by a model is actually quite a tradition in development economics.

That's *not* to say there is *never* a business case. For example, I just made one for the political reservations, that it was actually an efficient thing to do to reserve policy positions for women for five years or 10 years. Have them in place just so that people experience the fact of having a woman lead and realize that women are not what they thought.

But I just think that the business case argument should be used when we have evidence for it, and it shouldn't be used when we don't. Policymakers always want to go back to the business case because it goes well, you know? It's win-win, and win-win arguments are very popular.

Region: Especially in politics, so the politician needn't worry about alienating those who might lose.

Duflo: In politics, exactly, win-win arguments really have this attraction. People want something for nothing, and they can't always have it.

Region: Spoken as a true economist: no free lunches.

Duflo: Actually, as an applied economist, I'm one to think that often you *can* get something for *next* to nothing. There are many cases where there is a lot of inefficiency, and they could be improved. That would increase the size of the cake as well as redistributing it. So in general I'm more on the side of thinking that sometimes it's possible. I just think that we should make these claims only when we are able to make them.

THE COMMON THREAD

Region: You've done such a broad range of empirical work, from microfinance to fertilizer use, teacher absenteeism, school construction, water supplies. And you've conducted these studies from Indonesia to Cote d'Ivoire.

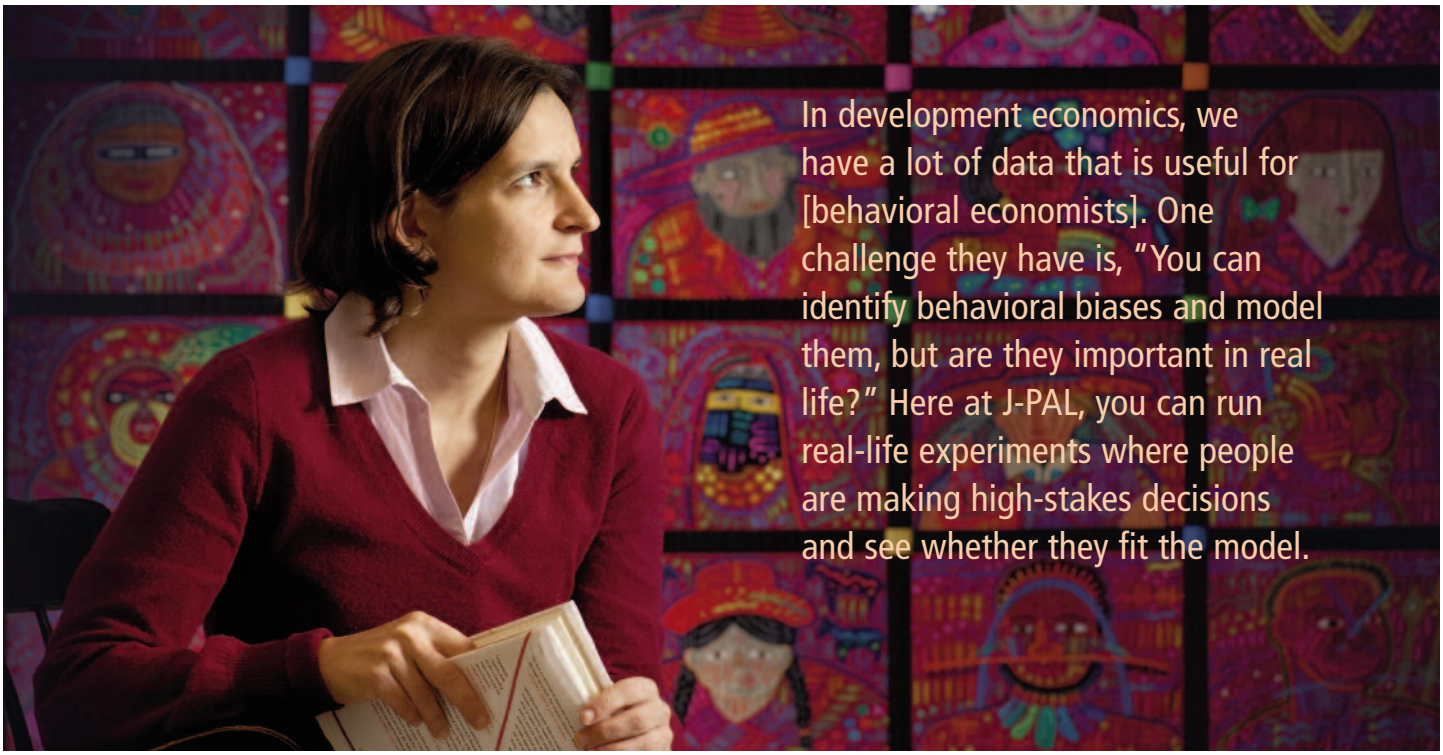
What is the common thread? What are the fundamentals that motivate and give coherence to this wide range of work?

Duflo: So [long pause]. Early on in my career, I guess, that's a question which I wasn't asking myself, and I wasn't, in the sense, particularly interested in that. There are so many questions that are important in development that we know little about. So if I can get an opportunity to answer these questions, I should go for it. I guess that's why there's such a range of things I've studied.

In that range, there was one common thread, which is methodological: If I've made causal statements, they are accurate; these are *true* natural experiments or *true* randomized experiments. So that's always been there.

But in terms of the topics, I guess I've looked at the common core of the type of questions that we as a development community think are important: education, infrastructure, et cetera. And on those questions, to just do what I could do.

I didn't feel that more focus than that was particularly needed because, you know, it was such an open field. It was an excellent field when I entered it. I started at MIT in 1995. It was an ideal time to start working on economic development because the early 1990s had seen a lot of really fundamental theoretical work, particularly by Abhijit Banerjee, Andy Newman, Debraj Ray, Tim Besley and other applied theorists and building on Joseph Stiglitz. So, it's a field that had first been reborn through



applied theory. When I started working in this field, there was a range of questions that were all fitting together in a theoretical framework, but empirically, the questions were quite open.

Not very many people were working in empirical research then. I mean to say, of course, that the field has a long tradition of both empirical and theoretical work, but it was just not a very thick field. There's a tradition of very, very good people. But with all of the theoretical work laid out in the early 1990s, it left many empirical questions wide open.

And then more or less at the same time, there was all of this work in labor economics in the U.S., and later in public finance, that was improving methodology in terms of demonstrating causality. So we had the whole field of development economics in which to make the two play together. The idea of carefully looking at data while being inspired by a model is actually quite a tradition in development economics. For example, Schultz's concept of "poor but rational" was very grounded in the theory of the time, but it also looked at the data that then existed.

That continued constantly over time, and we're doing the same thing—leaning on the theory and using the tools that had been developed more recently. It was a great time, given that these two things were available.

And now you can judge yourself whether or not *Poor Economics* has some common thread. There is no grand unifying theory of everything, which comes back to this idea that there is not one vision of the poor that can explain everything, but there are a number of insights that run through it all. And my work is maybe always, or often, trying to push one of these insights or show evidence—either very traditional things, like how people respond to financial incentives, or slightly more recent things like the fertilizer work, which links more to behavioral economics.

BEHAVIORAL BIASES

Region: Let's jump to that. You mentioned it earlier as well, the behavioral biases and blocks that limit what might be considered rational behavior. You've done research on this with Kremer and Robinson on fertilizer and with Banerjee and others on immunization.[8] Utility-maximizing agents are a bedrock assumption in traditional economic models, so how do you change theory to account for such biases, and does doing so limit the explanatory power and generalizability of theory?

Duflo: Well, the theoretical advances come from people who have done the behavioral work. I tend to take models and apply them to the circumstances that I have. But when you take the workhorse model and add hyperbolic discounting, people are still maximizing utility; they're just maximizing a utility function that's different from the one we used to work with, and probably more realistic.

It's like when information economics came in the 1980s, or the late '70s. Before that, the models assumed that people were perfectly informed when of course they are not. And people like Stiglitz showed that we can incorporate imperfect, asymmetric information and still work with that. And again, you don't have one universal model of everything.

Region: But eventually do they become analytically intractable?

Duflo: There are more things to deal with, but no. Maybe they are harder to put in a big machine to explain the entire economy, like a "Minnesota school" economist would like to do, but they can still be worked with. And if it hasn't already happened, it won't take very long for it to be incorporated in macro models as well.

Region: Who are the behavioral economists that you take your lead from?

Duflo: David Laibson, Matthew Rabin, Sendhil Mullainathan, Dick Thaler.

In development economics, we have a lot of data that is useful for them. One challenge they have is, "You can identify behavioral biases and model them, but are they important in real life?" Here at J-PAL, you can run real-life experiments where people are making high-stakes decisions and see whether they fit the model.

MACRO & MICRO

Region: Most of your work—perhaps all—is at the microeconomic level. How do you integrate your findings with macroeconomic considerations that affect development—aggregate growth, international trade, and fiscal and monetary policy?

Macro models should be micro-founded with the right micro assumptions. "Right" both in terms of incorporating the important dimensions that really need to be there, like credit constraints or some other reason why resources don't flow to their most efficient use, and right quantitatively in terms of micro parameters. … The agenda is very young, but it's being moved.

More About Esther Duflo

Current Positions

Abdul Latif Jameel Professor of Poverty Alleviation and Development Economics, Massachusetts Institute of Technology, since 2005; founder and co-director, Abdul Latif Jameel Poverty Action Lab (J-PAL), since 2005; on MIT faculty since 1999

Previous Positions

Director, Development Program, Center for Economic Policy Research

Research Associate, National Bureau of Economic Research

Board Member, Bureau for Research and Economic Analysis of Development

Fellow, MacArthur Foundation

Fellow, American Academy of Arts and Sciences

Member, Board of Editors, *Annual Review of Economics*

Founding Editor, *American Economic Journal: Applied Economics*

Honors and Awards

David N. Kershaw Award, Association for Public Policy Analysis and Management, 2011

Médaille de l'Innovation, Centre National de la Recherche Scientifique, 2011

Thomas C. Schelling Award, Harvard Kennedy School, 2011

John Bates Clark Medal, American Economic Association, 2010

Calvó-Armengol International Prize, Barcelona Graduate School of Economics, 2009

Best Young French Economist Prize, Le Monde/Cercle des Économistes, 2005

Elaine Bennett Prize for Research, American Economic Association, 2003

Publications

Author of numerous journal articles, particularly on microeconomic issues in developing countries, including household behavior, education, access to finance, health and policy evaluation through use of randomized control trials. Co-author (with Abhijit Banerjee) of *Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty*, selected as the *Financial Times*/Goldman Sachs "Business Book of the Year" in 2011

Education

Massachusetts Institute of Technology, Ph.D. in economics, 1999

DELTA, Paris, master's in economics, 1995

L'Ecole Normale Supérieure, Paris, 1994

Developing theory is not really my role. I'm an empirical person. That's what I'm good at. I'm not going to start writing theory. It's certainly not J-PAL's role. We run experiments; that's what we do.

Duflo: Banerjee and I have a chapter in the *Handbook of Economic Growth* called "Growth Theory from the Lens of Development Economics" that tries to get at that.[9] And the point we are making is that these macro models should be micro-founded with the right micro assumptions. "Right" in terms of incorporating the important dimensions that really need to be there, like credit constraints or some other reason why resources don't flow to their most efficient use, and right quantitatively in terms of micro parameters.

We do it in a primary school sort of way at the end of the paper. I think since then there has been much, much more involved work to do this well. That's something that Rob Townsend here at MIT has been doing for a while. A lot of people trained more in the [Universities of] Minnesota- and Chicago-type of traditions are good at it and have been doing it since then. So, Townsend, Paco (Francisco) Buera, Pete Klenow. Other younger economists are also going in this direction.

I'm not saying we couldn't do it here, and to some extent it is being done here under Rob Townsend. But in any case, I think that is the way to integrate micro and macro development economics. That is, use the micro insights to estimate parameters and also all the important things that it tells you about the way life is, I guess. What needs to be taken into account, like the shape of the production function, whether people need collateral to borrow—incorporate those constraints in models and then do the same calibration exercises.

I think the agenda is very young, but it's being moved. That's one thing that I'm not going to do. It's not my comparative advantage. But I think someone should be doing it, and in fact, they *are*, in a very fruitful way.

RESEARCH AGENDA

Region: In September 2010, you released a paper, "A Research Agenda for Development Economics."[10] It suggests (a) revitalizing applied theory to address limits exposed in earlier theory by recent empirical work, (b) expanding empirical research and (c) expanding both theory and empirical research on aggregate consequences of micro distortions.

Duflo: Yes, the third is what I've just talked about. And the first point we've discussed a bit. So, point two, obviously I have to preach a bit for my own parish along the way.

Region: Yes, please, let's focus on the first two. How can theory be revitalized, and how will empirical work expand? And how do you and J-PAL intend to allocate your time, your resources?

Duflo: This is already happening in a sense. You know, we know more than we did five years ago; there are more applied theory papers that have come out. But it's true that, as I was saying, the really big growth spurt of developments in the late '80s, early '90s was applied theory, and it gave the framework that all of the empirical work built on.

But now sometimes some of the limits of those models have been shown, and it would be nice to have other things as well. Now people should go back to doing that, and I think they will, naturally.

And if they don't, it's not because of development economics per se; it's because economics as a field generally is not very sympathetic to applied theory at the moment. I think theoretical work needs to be considered "hard core" to be interesting. That's not specific to *development* economics; it's a general issue in economics as a field.

But you know, this type of thing comes and goes, so applied work will come back. Developing theory is not really my role. I'm an empirical person.

That's what I'm good at. I'm not going to start writing theory. It's certainly not J-PAL's role. We run experiments; that's what we do.

But to the extent that we train students, we make sure that training in development economics has a good balance of theory and empirical work. We are lucky to have Abhijit Banerjee and Rob Townsend at MIT, who can make sure that this happens!

Region: Thank you very much for your time today. R

—Douglas Clement
Oct. 25, 2011

Endnotes

[1] See Schultz (1964).

[2] See Heckman (1991).

[3] See Jensen and Miller (2008).

[4] See Duflo (2006).

[5] See Duflo (2011b).

[6] See Duflo (2005a).

[7] See Duflo et al. (2009) and Duflo (2005b).

[8] See Duflo, Kremer and Robinson (2009) and Banerjee et al. (2010).

[9] See Banerjee and Duflo (2005).

[10] See Duflo (2011a).

References

Banerjee, Abhijit, and Esther Duflo. 2005. "Growth Theory through the Lens of Development Economics." In *Handbook of Economic Growth*, Steve Durlauf and Philippe Aghion (eds.) North Holland: Elsevier Science, pp. 473-552.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster and Dhruva Kothari. 2010. "Improving Immunization Coverage in Rural India: A Clustered Randomized Controlled Evaluation of Immunization Campaigns with and without Incentives." *British Medical Journal* 340:c2220.

Duflo, Esther. 2005a. "Gender Equality in Development." Policy Paper 001. Bureau for Research and Economic Analysis of Development.

Duflo, Esther. 2005b. "Why Political Reservations?" *Journal of the European Economic Association* 3 (2-3): 668-78.

Duflo, Esther. 2006. "Poor but Rational?" In *Understanding Poverty*, Abhijit Banerjee, Roland Benabou and Dilip Mookherjee (eds.) New York: Oxford University Press, pp. 367-78.

Duflo, Esther, Lori Beaman, Raghabendra Chattopadhyay, Rohini Pande and Petia Topalova. 2009. "Powerful Women: Does Exposure Reduce Bias?" *Quarterly Journal of Economics* 124 (4): 1497-1540.

Duflo, Esther, Michael Kremer and Jonathan Robinson. 2009. "Nudging Farmers to Use Fertilizer: Evidence from Kenya." Working Paper 15131. National Bureau of Economic Research. Forthcoming in *American Economic Review.*

Duflo, Esther. 2011a. "A Research Agenda for Development Economics." American Economic Association, Ten Years and Beyond: Economists Answer NSF's Call for Long-Term Research Agendas. Available at Social Science Research Network: http://ssrn.com/abstract=1888605.

Duflo, Esther. 2011b. "Balancing Growth with Equity: The View from Development." Prepared for 2011 Jackson Hole Symposium. Available at http://www frbkc.org/publicat/sympos/2011/2011.Duflo.Paper.pdf.

Heckman, James J. 1991. "Randomization and Social Policy Evaluation." Technical Working Paper 107. National Bureau of Economic Research.

Jensen, Robert, and Nolan Miller. 2008. "Giffen Behavior and Subsistence Consumption." *American Economic Review* 98 (4): 1553-77.

Schultz, Theodore W. 1964. *Transforming Traditional Agriculture.* New Haven, Conn.: Yale University Press.

Nothing but Net

At Heller-Hurwicz Forum on Social Insurance, a synergy of research and policy

We've learned a great deal about both the frontiers of research and the challenges of policy and its implementation. Fortunately, we have also seen that promising new economic research holds great potential for the design of better, more effective social insurance.

—V. V. Chari

All branches of government, including the Supreme Court, are currently debating social insurance—protection of the people, by the people and for the people against uncertainty in life. This so-called safety net—including Social Security, unemployment and disability insurance, Medicare and Medicaid—is often taken for granted. But, of course, these forms of insurance are neither guaranteed nor inexpensive.

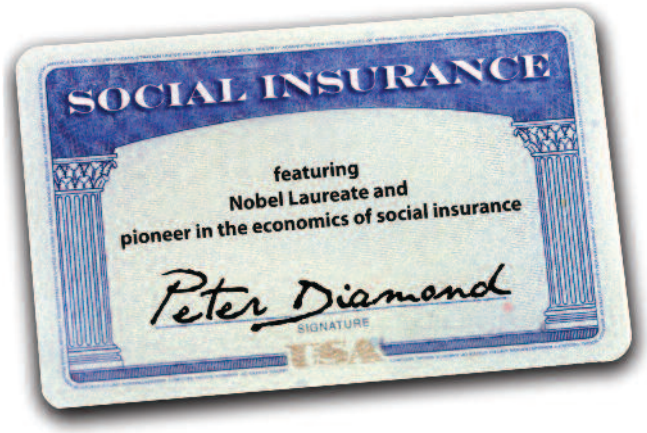
Moreover, while citizens value such programs highly, providing them is costly both in their obvious fiscal impact and through their subtler incentive effects. Economists point out that unemployment insurance, disability payments and retirement pensions may affect labor supply. Health services are likely used more when their cost is subsidized. Taxing wages and capital can discourage work and investment.

Thus, the structure and dynamics of social insurance programs have tremendous economic consequences, and for decades, economists have studied how to design an effective and efficient safety net and how to generate tax revenue to pay for it. At the University of Minnesota and the Minneapolis Fed, in particular, researchers have pioneered optimal design of insurance and taxation policies and conducted innovative research into health and economic risk over the life cycle.

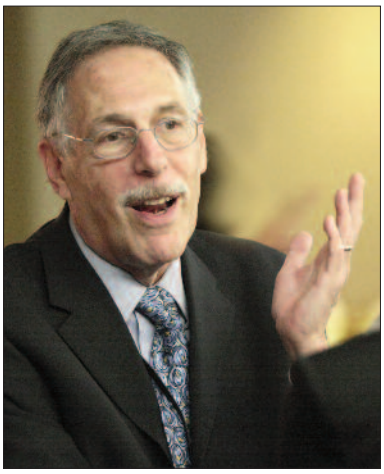
HHEI's first forum

Social insurance was therefore both a fitting and timely topic for the Heller-Hurwicz Economic Institute's first annual policy forum, held Nov. 16-17, 2011, at the University of Minnesota. HHEI was launched in 2010 to help shape public policy with insights from cutting-edge economic research, and the design of social insurance programs builds solidly on the legacies of the institute's guiding lights, Walter Heller and Leo Hurwicz.

The "Inaugural Forum on Social Insurance" immersed its roughly 200 registered participants in



theory, policy and practice. Both days began with presentations from economists on taxation, social insurance and government spending. The afternoons consisted of panel discussions—on tax policy, pension programs and health care—led by policymakers, policy advisers and practitioners. The keynote address was delivered by 2010 Nobel Laureate Peter Diamond, the former MIT economist renowned for his analysis of optimal taxation, labor markets and social insurance.



When Chari called me and asked would I be the keynote speaker at the first Heller-Hurwicz event, I jumped at the chance.

—Peter Diamond

Taxes, expenditures and the size of government

Day one began with MIT economist Iván Werning's theoretical exploration of social insurance and optimal taxation policy. Donald Marron from the Urban-Brookings Tax Policy Center then presented an analysis of federal government expenditure. Laurence Kotlikoff of Boston University followed, asserting that the United States is already "bankrupt and we don't even know it."

Health care panel:
Douglas Holtz-Eakin
Larry Jacobs
John Marty

Photos by Everett Ayoubzadeh and Douglas Clement

Alan Viard

Following lunch, a tax policy panel discussion, moderated by V. V. Chari, founding director of the HHEI and a Minneapolis Fed consultant, with Kotlikoff and the American Enterprise Institute's Alan Viard, focused on using economic principles to design a better tax system. A second panel discussion followed: using mechanism design theory to build viable pension programs. This discussion, moderated by the university's Art Rolnick, former research director at the Minneapolis Fed, included Leo de Bever of the Alberta Investment Management Corp., Martin Skancke, formerly of the Norwegian Ministry of Finance, and Kurt Winkelmann of MSCI.

Stefania Albanesi

The evening event began with Peter Diamond's recollections of both Heller and Hurwicz. "When Chari called me and asked would I be the keynote speaker at the first Heller-Hurwicz event, I jumped at the chance," he said, "because I have extremely positive feelings about the accomplishments of both of them [and] brief but warm personal connections." Diamond then discussed recent research on the forum's policy focus: "Resource allocation and economic stabilization: Taxes, spending, regulation and social insurance."

Theory and practice

The forum's final day started with three research papers on social insurance and taxation. First, Columbia University's Stefania Albanesi presented on the optimality of tax policies that front-load distortions (raising labor taxes now to finance tax cuts in the future, for example); then Emmanuel Farhi of Harvard on insurance and taxation over the life cycle and, finally, Mikhail Golosov of Princeton on optimal dynamic taxation.

After lunch, Stephen Parente of the University of Minnesota explored health care finances, stressing the strain on future revenue streams of projected program expenses. The forum concluded with an afternoon health care panel moderated by the University's Larry Jacobs, at which state Sen. John Marty and Douglas Holtz-Eakin, former director of the Congressional Budget Office, debated the future of health care entitlement systems.

Mikhail Golosov and Emmanuel Farhi

New possibilities

HHEI annual forums are "designed to advance emerging theories, push the boundaries of economic theory and open new possibilities in the face of pressing problems," according to the institute's website. Through the depth of research and discussion at the inaugural forum, observed Chari, "we've learned a great deal about both the frontiers of research and the challenges of policy and its implementation. Fortunately, we have also seen that promising new economic research holds great potential for the design of better, more effective social insurance." Referring in part to a post-forum celebration with 2011 Nobel laureates and former University of Minnesota economists Thomas Sargent and Christopher Sims (see page 46), he added, "And moreover, we found that, as expected, the future of Minnesota economics is as bright as its past."

—Douglas Clement

While citizens value such programs highly, providing them is costly both in their obvious fiscal impact and through their subtler incentive effects. The structure and dynamics of social insurance programs have tremendous economic consequences.

To learn more about HHEI and the policy forum on social insurance, visit http://hhei.umn.edu/policyForum2011/.

Kurt Winkelmann

Peter Diamond and Robert Lucas

Peter Diamond and Larry Jones

Douglas Holtz-Eakin

Taking the Measure of Prices and Inflation

A century of evolution—and near-constant criticism—has greatly improved price indexes. But work continues to perfect these closely watched economic indicators

Phil Davies
Senior Writer

Price fluctuations have been an economic phenomenon since ancient times. Escalating prices led Roman Emperor Diocletian to enact price controls in A.D. 301, but his "Edict on Maximum Prices" failed to restore financial order. In England during the 16th and early 17th centuries, an inflow of plundered treasure from the New World contributed to 4 percent annual price inflation. The United States experienced sharp price increases during the Revolutionary War—when Congress' enthusiastic printing of money to finance the war triggered rapid depreciation of the Continental dollar—and during the Civil War. The Great Depression and many other periods in U.S. history have seen the opposite phenomenon—deflation, or declining prices.

Attempts to measure inflation rates—how much prices rise or fall over time—also go back a long way. Economists experimented with index formulas for gauging the average price of various commodities in the 19th century, and the first official indexes measuring broad prices and their changes were developed by government agencies in the early 1900s as a means of settling wage disputes.

Over the past century, government entities such as the U.S. Bureau of Labor Statistics (BLS)—keeper of the Consumer Price Index—have strived to develop more accurate measures of changes in consumer prices, often under intense scrutiny from stakeholders in price fluctuations. Few government statistics have generated as much debate over the decades as price indexes, which have assumed an increasingly important role in the economy.

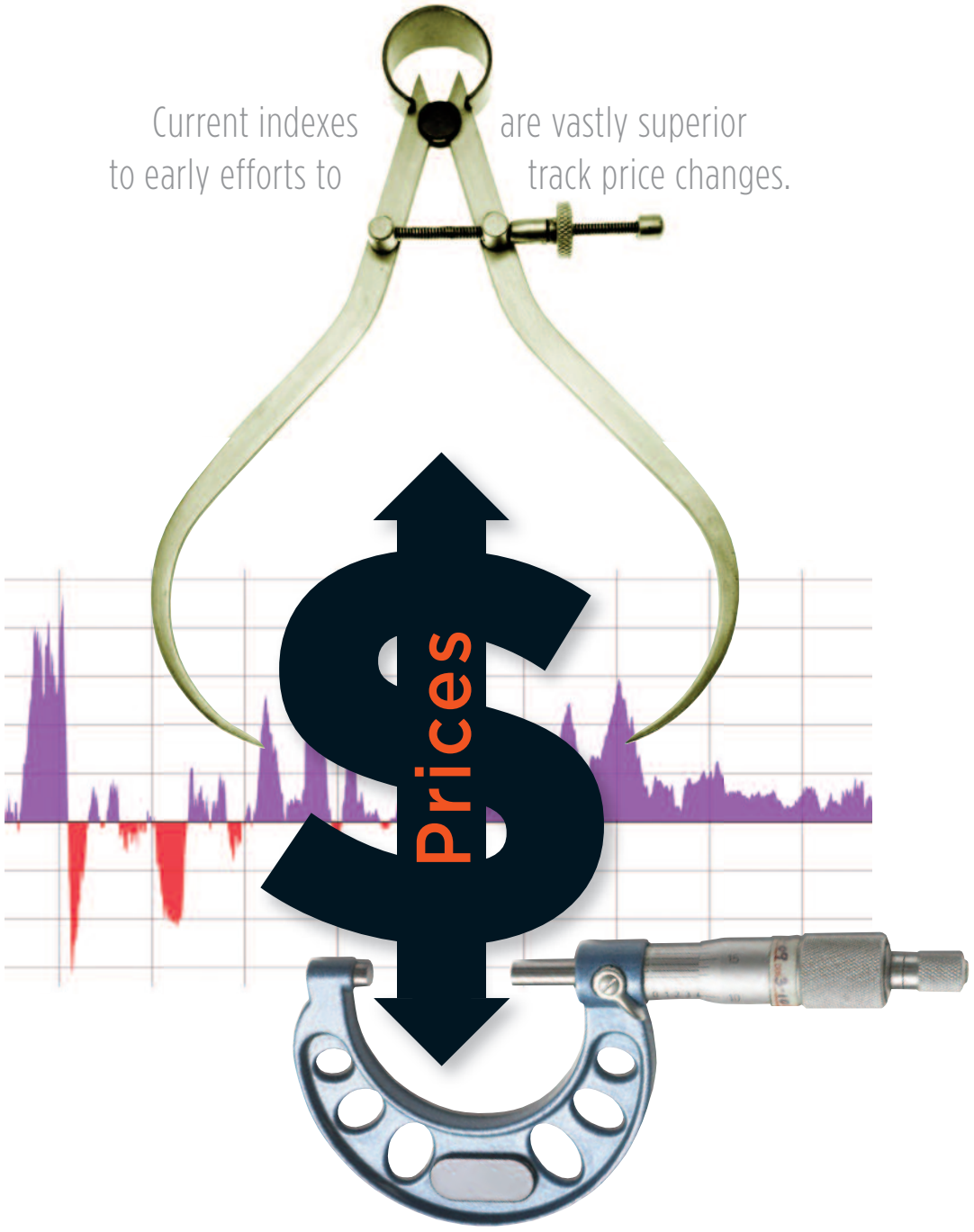
Conceived as tools for adjusting wages and business contracts to current prices, over time price indexes developed into cost-of-living escalators for government programs and crucial indicators of economic performance. Economists began to watch closely the rate of change of price indexes—the inflation rate. A marked and persistent rise in inflation usually presages higher interest rates, which could in turn reduce investment by businesses, slowing economic growth for a time. When inflation approaches zero or becomes negative, it raises the specter of sustained deflation.

No wonder, then, that price indexes are closely watched and often questioned. Economists in particular frequently tussle over the finer points of computing price indexes, said Jack Triplett, an economist with the Brookings Institution who has done extensive research on price measures. "There are great demands that economists make on the accuracy of the CPI, demands that they don't make on other data they use," he said. The same could be said of the Personal Consumption Expenditures (PCE) price index, an alternative inflation measure published by the Bureau of Economic Analysis (BEA).

But getting an accurate reading of the overall change in consumer prices isn't as easy as it would first appear; all kinds of methodological problems bedevil designers and overseers of price indexes. Among the challenges: ensuring that the index rep-

Current indexes
to early efforts to

are vastly superior
track price changes.



And the fundamentals of inflation
measurement are no longer in dispute.

resents the consumption patterns of the U.S. population, accounting for the inclination of consumers to substitute items (say, ground turkey for hamburger) in response to relative price changes and adjusting prices to reflect quality improvements and the introduction of new products.

Over the decades, government agencies have been prodded to address these issues and improve price indexes. The nearly constant criticism came initially from trade unions and politicians, and later from economists. During World War II, trade unions insisted that the BLS's cost-of-living index (predecessor of the CPI) grossly underestimated the rate of wartime inflation. In the 1990s, a congressional commission appointed to review the CPI took the opposite tack, maintaining that the index overestimated the inflation rate.

The science of gauging consumer inflation has made great strides over the past century; current measures are vastly superior to early efforts to track price changes, and the fundamentals of inflation measurement are no longer in dispute. But price indexes remain works in progress; economists continue to sweat the technical details in an effort to improve them further.

Can you believe the prices these days?

Statisticians developed the first gauges of changing price levels over 200 years ago; during the American Revolution, the Massachusetts Legislature developed a rudimentary price index used to adjust soldiers' wages as the dollar declined in value. The index comprised an average of the prices of four staple commodities: Indian corn, beef, sheep wool and finished leather.[1]

But the science of measuring price trends wasn't developed until the late 19th century, when economists and statisticians came to grips with issues such as index formulas, weights (relative expenditures on different items) and sampling error. In the 1870s, the German econometrician Étienne Laspeyres invented the index formula for determining price changes that is at the heart of the CPI. About the same time, British economist William Stanley Jevons advocated the creation of a government-authorized "tabular standard of value" that could be used to update contract prices by measuring average variations in the purchasing power of gold.

In the United States, the early 1900s were a period of massive labor unrest. Prices for most goods were on the rise after a long period of stable or slowly declining prices, and workers in a range of industries were agitating for higher pay.[2] Public and private employers wanted a reliable gauge of living expenses for use in wage negotiations, but none existed at the time. The only official price indexes were for selected wholesale commodities sold in a limited number of cities.

In 1904, the federal Bureau of Labor (forerunner of the BLS) published a monthly index of retail food prices gleaned from 800 merchants in large industrial centers. The index, covering the past 13 years, priced 30 principal food items and weighted them according to average consumption. Within a few years, the food price index reflected data gathered from over 1,000 retail establishments in 40 states. The bureau also collected wage data in the surveyed cities.

For the first time, changes in workers' pay could be compared with changes in the price of food. But labor groups and some politicians blasted the index, charging that it was politically motivated and that it failed to reflect the diminished purchasing power of workers. "It will take more than [the food index] to convince the housewives of the nation that wages have increased in proportion to the increase in prices," declared the International Association of Machinists.[3]

World War I provided the impetus for a much more ambitious indexing project—a set of new

In Brief

Price trackers

- Few government statistics have generated as much debate over the past century as price indexes. Conceived as tools for settling wage disputes, price indexes developed into cost-of-living escalators and crucial economic indicators.
- Nearly constant criticism, initially from trade unions and later from economists, prodded government agencies to improve price indexes by accounting for factors such as consumer substitution, quality change and new products.
- Today's price indexes are far more accurate measures of consumer inflation than early efforts to track price changes. But technical issues still dog economists and statisticians striving to further improve price indexes.

Counting prices at the BLS: tabulating room, about 1935; UNIVAC operator's console, 1965

retail indexes covering an array of domestic items in addition to food. As in previous conflicts, price increases accelerated during the war, particularly in shipbuilding centers. Intent on setting equitable wages for factory workers vital to the war effort, the National War Labor Board in 1918 called upon the BLS to produce nationwide data on the "cost of living"; changes over time in this index would indicate how much household income would have to change to maintain roughly the same standard of living. The stated goal of this exercise: "[I]nsure the subsistence of the worker and his family in health and reasonable comfort."[4]

President Woodrow Wilson allocated the bureau $300,000 ($4.4 million in today's dollars) to conduct a national study of prices and household expenditures. Over the next two years, BLS agents fanned out across the country to collect prices for about 145 consumer products and services. Price takers carefully specified items to make pricing of identical or similar items easier in future surveys and surveyed about 12,000 working-class families in 42 states to gather information about income and consumption patterns.

In 1919, the bureau released the first comprehensive set of cost-of-living indexes for 31 major industrial and shipbuilding centers. Thereafter, updated indexes were issued semiannually for individual cities (Washington, D.C., was added in 1921) and the nation as a whole. In the Roaring '20s, an era of rising incomes and economic growth, the indexes quickly became standard tools for negotiating wage increases.

"Hidden" price increases

In the 1930s, widespread financial hardship prompted revisions to the BLS's cost-of-living index, which was still based on consumer purchases during the First World War. The BLS updated and enlarged the index's basket of goods to better reflect spending by wage-earning households and conducted a new family expenditure survey from 1934 to 1936. The new survey showed a marked rise in the U.S. standard of living since the war, despite the privations of the Great Depression. Workers and their families were buying more ready-made clothing, heating oil and refrigerators, eating more fruits

and vegetables, and spending more money at the beauty shop and on driving vacations.

Price controls and rationing during World War II brought more changes to the cost-of-living index—scarce items such as refrigerators, automobiles and tires were removed, for example—and provoked heated opposition by trade union leaders who distrusted the BLS's methodology.

In spite of federal price controls, prices started rising in 1942 as factories switched from production of goods demanded by consumers to the manufacture of ships, tanks, munitions and other war materiel. After the War Labor Board tied wage increases to the cost-of-living index in what became known as the "Little Steel Agreement," labor unions attacked the index, charging that it failed to capture the full rise in living costs for industrial workers. In an effort to refute the BLS's figures, a number of unions collected their own retail price data.

Controversy over the index contributed to labor unrest during the war; ignoring no-strike pledges by union leaders, workers in a number of industries staged a series of wildcat strikes in 1943 and 1944. Toward the end of the war, the War Labor Board created a presidential committee to investigate the unions' charges. Before it could rule, labor representatives on the committee, including George Meany of the American Federation of Labor (AFL), issued their own blistering critique of the BLS's price index. In their 1944 report, the labor leaders estimated that the actual rise in the cost of living from 1941 to 1943 was almost twice the 23 percent reported by the BLS (see chart at right).

As Triplett and Marshall Reinsdorf of the BEA note in a 2006 paper, the unions ascribed the "hidden" price increases to skewed BLS sampling that omitted items not subject to price controls and a decline in the quality of items such as shoes due to wartime shortages of material and labor.[5] In a speech, Meany questioned the scientific validity of the cost-of-living index and accused the bureau of "obsequiously" going along with an alleged Roosevelt administration plan to freeze wages.[6]

The BLS vigorously defended the index, supported by experts involved in the review process who estimated that the index underestimated wartime price increases by only 3 percent to 4 percent. But in a nod to its critics, the bureau changed the name of the index in 1945. Henceforth it would be known as the Consumer Price Index—an acknowledgment that the measure was not a true cost-of-living index because it didn't fully capture changes in product quality or consumer substitution of items when one becomes cheaper (or its price rises less) than another.

Unions remained suspicious of the CPI into the early 1950s. The AFL and the Congress of Industrial Organizations continued to push for improvements in the index to allow it to serve as a broad measure of living costs, while more radical labor groups damned it as a flawed, politicized instrument of wage suppression.

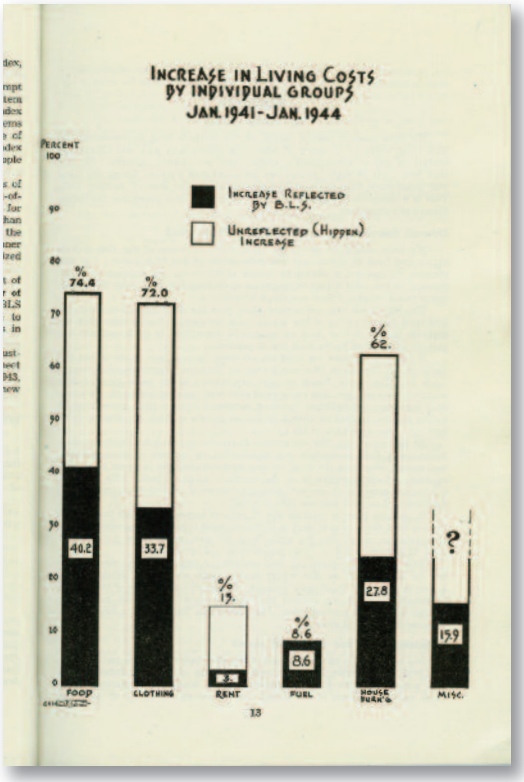
Finger on the pulse

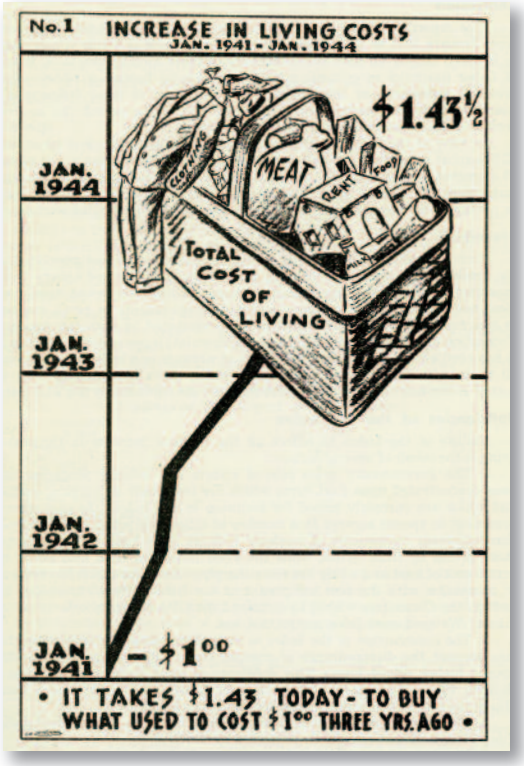
Before World War II, price indexes were used chiefly as yardsticks for adjusting wages, rents, royalties and other contracts to the cost of living. In the postwar economy, they took on an additional function as a barometer of the general level of inflation, a key indicator of economic performance.

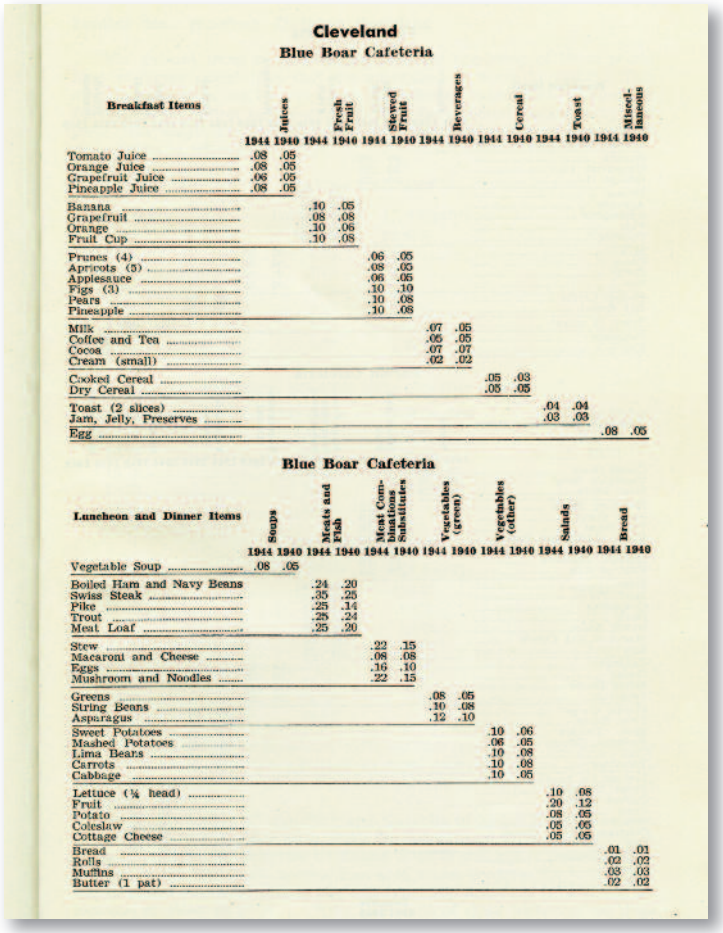
After price controls were phased out in the late 1940s, the inflation rate accelerated; the federal government became increasingly anxious about rising prices and turned to the CPI and BLS wholesale price indexes as a guide to monetary policy. Economists at universities and private research institutions also wanted a reliable gauge of changes in the overall price level.

In 1951, the economics arm of the U.S. Department of Commerce introduced a new price index measure designed specifically as a macroeconomic indicator. The "implicit price deflator," based on data on personal consumption in the national accounts, was broader in scope than the CPI and used a different index formula to measure average change in the price of consumer goods and services. The deflator would evolve into today's PCE price index.

The growing importance of the CPI to economic analysis exposed the measure to renewed criticism—leveled not by labor groups, but by professional economists. A 1961 review of the CPI and other federal price indexes commissioned by the Eisenhower administration delved into statistical problems that had long simmered in academic circles. One issue was how to account for changes in product quality; if a washing machine costs the same as last year's model but performs better, its







Cleveland
Blue Boar Cafeteria

Breakfast Items	Juices 1944	Juices 1940	Fresh Fruit 1944	Fresh Fruit 1940	Stewed Fruit 1944	Stewed Fruit 1940	Beverages 1944	Beverages 1940	Cereal 1944	Cereal 1940	Toast 1944	Toast 1940	Miscellaneous 1944	Miscellaneous 1940
Tomato Juice	.08	.05												
Orange Juice	.08	.05												
Grapefruit Juice	.06	.05												
Pineapple Juice	.08	.05												
Banana			.10	.05										
Grapefruit			.08	.08										
Orange			.10	.06										
Fruit Cup			.10	.08										
Prunes (4)					.06	.05								
Apricots (3)					.08	.05								
Applesauce					.06	.05								
Figs (3)					.10	.10								
Pears					.10	.08								
Pineapple					.10	.08								
Milk							.07	.05						
Coffee and Tea							.05	.05						
Cocoa							.07	.07						
Cream (small)							.02	.02						
Cooked Cereal									.05	.03				
Dry Cereal									.05	.05				
Toast (2 slices)											.04	.04		
Jam, Jelly, Preserves											.03	.03		
Egg													.08	.05

Blue Boar Cafeteria

Luncheon and Dinner Items	Soups 1944	Soups 1940	Meats and Fish 1944	Meats and Fish 1940	Meat Combinations Substitutes 1944	Meat Combinations Substitutes 1940	Vegetables (green) 1944	Vegetables (green) 1940	Vegetables (other) 1944	Vegetables (other) 1940	Salads 1944	Salads 1940	Bread 1944	Bread 1940
Vegetable Soup	.08	.05												
Boiled Ham and Navy Beans			.24	.20										
Swiss Steak			.35	.25										
Pike			.25	.14										
Trout			.25	.24										
Meat Loaf			.25	.20										
Stew					.22	.15								
Macaroni and Cheese					.08	.08								
Eggs					.16	.10								
Mushroom and Noodles					.22	.15								
Greens							.08	.05						
String Beans							.10	.08						
Asparagus							.12	.10						
Sweet Potatoes									.10	.06				
Mashed Potatoes									.06	.05				
Lima Beans									.10	.08				
Carrots									.10	.08				
Cabbage									.10	.05				
Lettuce (¼ head)											.10	.08		
Fruit											.20	.12		
Potato											.08	.05		
Coleslaw											.05	.05		
Cottage Cheese											.05	.05		
Bread													.01	.01
Rolls													.02	.02
Muffins													.03	.03
Butter (1 pat)													.02	.02

Excerpts from labor unions' 1944 critique of the BLS's cost-of-living index

price has fallen in real terms. Another was how to treat the cost of shelter, which can be viewed as an investment as well as a living expense.

"It isn't that people way back then didn't understand what the issues were," Triplett said. "It's just that these were things that were hard to resolve. People knew that the quality change problem was serious; they just didn't know what to do about it."

A panel of experts chaired by University of Chicago economist George Stigler exhorted the BLS to research potential solutions to these problems, including using the rental equivalence method (estimating market rents for owner-occupied homes) to measure changes in shelter expenses and sampling items more frequently to reflect purchases of new or improved products.

The Stigler Commission's recommendations were barely noticed by the public, but they prompted intensive research on price indexes that over time greatly improved the CPI and other price indexes.

Just WIN, baby

Taking the measure of inflation took on fresh urgency in the late 1960s and 1970s, when inflation threatened to spiral out of control. Energy prices soared, contributing to annual inflation rates above 4 percent in the early 1970s. "Stagflation"—high inflation coupled with slow economic growth—gripped the nation, spurring the Nixon administration to impose price controls and President Gerald Ford to launch Whip Inflation Now (WIN), a much-lampooned initiative to foster energy conservation and cut consumer spending.

Fighting inflation was job one for policymakers, who sought more precise intelligence on price movements and their interplay with economic output and employment. The BLS and other agencies broadened the scope of their price indexes and developed new ones to obtain a closer reading of inflationary trends and their impact on consumers.

In 1978, the BLS split the CPI into two measures, each representing the buying habits of distinct populations. A new CPI for all urban consumers (CPI-U) expanded the index's geographic reach beyond large cities to smaller urban areas and added previously excluded groups of consumers such as salaried employees, part-time workers, the unem-

I say CPI, you say PCE

One measure of inflation is the popular one, the index that hogs the limelight in the media and around the water cooler, especially when prices are rising. The other gauge of price change is the shy one, the little-known measure that gets attention only from policy wonks and macroeconomists.

Both the Consumer Price Index for all Urban Consumers (CPI-U) and the Personal Consumption Expenditure Index (PCE) track changes in prices paid by consumers for goods and services. Both measures—the first published by the U.S. Bureau of Labor Statistics (BLS), the second by the Bureau of Economic Analysis—have "core" versions that exclude food and energy prices to help inflation watchdogs such as the Federal Reserve anticipate future movements in the headline, or overall, index.

Yet, like lenses in a pair of binoculars that view objects from divergent angles, the indexes show slightly different inflation rates. Although they usually move in parallel when prices rise or fall, the PCE has historically traced a lower path than the CPI. However, since the 2000s, the average gap between the two measures has narrowed.

These alternative measures of consumer inflation reflect fundamental differences in the way the two indexes are constructed. Each has its own underlying concept, data sources and formula for calculating price changes.

The CPI takes an in-the-trenches approach to measuring inflation, tracking the change in

price of a market basket of goods purchased by all urban households. The BLS collects prices from more than 25,000 retail and service outlets in 87 urban areas across the country to create the national CPI. In contrast, the PCE measures price changes for goods and services within the framework of the National Income and Product Accounts, a comprehensive set of figures for the total value of output and income in the U.S. economy.

Because of these different approaches, the PCE measures a broader swath of personal consumption than the CPI. For instance, the PCE captures expenditures by rural as well as urban consumers and includes spending by nonprofit institutions that serve households. And while the CPI records only out-of-pocket spending on health care by consumers, the PCE also tracks personal medical expenses paid by employers and federal programs such as Medicare. However, over 70 percent of the price data in the PCE is drawn from the CPI.

Weight for it

The weights (relative consumer expenditures) assigned to prices are crucial, and the CPI and the PCE derive their weights from different sources. The CPI reflects reported consumption in the Consumer Expenditure Survey, conducted for the BLS by the U.S. Census Bureau. To determine its expenditure shares, the PCE relies on business surveys such as the Census Bureau's annual and monthly retail trade surveys. Shelter accounts for the biggest difference in weighting between the two indexes; the share of personal spending devoted to housing is larger in the CPI because nonshelter expenditures in the CES are less than those estimated from business surveys.

Another key distinction between the indexes is the mathematical formula used to aggregate myriad prices and sub-indexes into a measure of overall inflation. The CPI's "fixed-weight" formula calculates price changes from a base period whose expenditure weights are updated roughly every two years. The PCE uses a formula (developed by U.S. economist Irving Fisher in the early 20th century) that takes the average of two fixed-weight measures of price change—one based on weights in the current period and the other based on weights in the preceding period.

An important benefit of the PCE's formula is that it automatically adjusts for consumer substitution among general categories of goods (such as from grapes to apples) as relative prices change. Studies have shown that this "formula effect" accounts for almost half of the gap between the CPI and PCE inflation rates.

Other, minor differences between the indexes include alternative ways of adjusting for seasonality and figuring changes in airfares and gasoline prices.

The Federal Reserve and many economists hew to the PCE as an inflation measure. The Fed switched from the CPI to the PCE in 2000. In addition to the PCE's broad scope and index formula, the Board of Governors has said that it prefers the measure's historical consistency, valuable for research. Unlike CPI figures—which once published cannot be changed because they are written into contracts—previously released PCE data are continuously revised to reflect updated information and refinements in measurement techniques.

—Phil Davies

ployed and retirees. This is the measure of "headline," or overall, inflation that is reported most widely each month.

The traditional CPI dating to World War I continued as the CPI for urban wage earners and clerical workers (CPI-W), with weights reflecting the consumption patterns (more gasoline purchased by commuters, for example) of wage earners.

Research by economist Robert Gordon of Northwestern University gave rise to the CPI excluding food and energy, a measure of "core" inflation that first appeared in the annual *Economic Report of the President* in 1980. Gordon and many other economists believed that fluctuating prices for energy and food in global markets obscured the elusive inflation "signal" that should inform fiscal and monetary policy. In particular, oil prices fixed by producers in the Middle East were viewed as artificial distortions of the inflation rate, says John Greenlees, chief of the BLS's Division of Price and Index Number Research.

"Part of the idea behind a core index was to get rid of volatility or noise in the index," he said, "and part of it was to try to limit the measure to price changes that really reflected [inflationary forces] within the U.S. as opposed to something that's just arbitrarily set by Arab states."

The inflationary 1970s also saw major revisions to the PCE implicit price deflator, which was primarily used for macroeconomic analysis and forecasting. In 1976, the BEA changed the formula of the deflator to focus on pure price change instead of measuring changes in both prices and quantities of items purchased by households. This modification made the PCE a full-fledged measure of consumer inflation akin to but distinct from the CPI. (For more on the PCE and how it differs from the CPI, see sidebar on page 34.)

Greenspan speaks; Boskin weighs in

Sharp increases in interest rates in the late 1970s reined in inflation, ushering in the Great Moderation—an extended period of modest business cycles and low inflation. Even though the stagflation beast had been tamed, price indexes remained key indicators of economic performance, carefully watched by the Fed and other policymakers. And the CPI became more and more embedded in government operations as a cost-of-living escalator. Social Security benefits had been indexed to the CPI since the early 1970s; beginning in 1985, the CPI-U was used to make adjustments to federal income tax brackets, exemptions and deductions.

Over the next 30 years, the BLS and the BEA significantly improved their indexes by gathering more timely price and expenditure data and by revising methods of calculating price changes to reduce—if not entirely eliminate—statistical biases.

The BLS had resisted for over 20 years the Stigler Commission's recommendation that it revamp its method of calculating changes in shelter costs. The Bureau had stuck with an asset-based approach to measuring homeowner costs—tracking house prices. But in the early 1980s, after a period of rapid home appreciation, the agency warmed to the idea—endorsed by most economists of the day—that including home prices in the CPI distorted the index.

In 1983, the BLS switched to the rental equivalence method of measuring homeowner costs: CPI price gatherers consider how much a house would rent for if the owner rented it to someone else.

A major impetus for further changes in the CPI came from the Boskin Commission, an advisory group appointed by the U.S. Senate Finance Committee in 1995 to study the CPI.

Federal Reserve Chairman Alan Greenspan had caused a furor by stating in testimony to the committee that aspects of the CPI's construction "point in the direction of an overstatement of increases in the cost of living."[7]

The report of the commission, chaired by Stanford University economist Michael Boskin, supported Greenspan's assertion. In a turnabout from past criticism of the CPI that it downplayed inflation, the Boskin Commission found that the current index overestimated the rise in living costs by just over 1 percentage point per year.

The implications of this were enormous, provoking intense interest in the Boskin Report from economists, politicians and journalists around the world. If the CPI overstated inflation, then economic output and productivity had grown more than previously believed in the 1990s. Real median income had risen more than the official CPI

indicated. And—an eye opener for seniors and taxpayers—Social Security cost-of-living adjustments were too high and should be trimmed. The Boskin Commission's finding also held great import for interest rate policy at the Federal Reserve.

The commission recommended major revisions to the CPI that would largely correct the upward bias and make the measure a more accurate cost-of-living index, that is, an index that accounted to a greater degree for factors such as quality change, consumer substitution and technological innovation.

Reconstruction projects

The Boskin Commission found that more than half of the CPI's upward bias was due to the failure of the index to keep up with new products coming into the market and quality improvements in existing products.

Because household expenditure surveys were conducted only once every 10 years, the CPI missed the introduction of new products such as the microwave, VCR and cellular phone. By the time they entered the index (cell phones finally made the cut in 1998), they were much less expensive, but this price drop was missed by the index. To catch new products earlier, the BLS switched to a continuous survey process in which the CPI's expenditure weights are updated roughly every two years.

To address the quality change issue, the BLS expanded its use of "hedonic regression modeling," a technique developed in the 1930s to measure the impact of quality improvement (or decline) on the price of a product. If the price of an item rose, the portion of the increase that reflected improved quality didn't count as inflation in the CPI. The BLS had been using hedonic models to track the prices of clothing since the late 1980s; in response to the Boskin Report, the Bureau began applying the method to appliances and electronics, including TVs and computers.

Other changes to the CPI tackled the vexing problem of consumer substitution, which the Boskin Report found also contributed to the index's overstatement of inflation. The BLS had never found a solution to this issue, which had been raised by the Stigler Commission as well. The CPI's method of figuring monthly price changes couldn't accommodate the readiness of shoppers to substitute away from items that increased in relative price.

In 1999, the BLS partially solved the problem by modifying the Laspeyres formula used to measure price change at the basic item level (apples in Milwaukee or steaks in New York) in the CPI. The new calculation assumed that consumers responded to relative price hikes by buying similar, cheaper items; as a result, overall prices rose less than they did under the old formula.

Nevertheless, the CPI was still subject to bias from "upper level" substitution—consumer tradeoffs among broader item categories such as rice and pasta. To address this shortcoming, the bureau in 2002 added yet another measure of price change to its family of indexes—the chained CPI for all urban consumers. Intended primarily a research tool, the C-CPI-U relies on a different price-change formula that better accounts for consumer substitution among all item classes.

The Boskin Commission didn't examine the PCE index. Since its inception, the BEA's price deflator has largely avoided the scrutiny directed at the CPI—perhaps because the PCE has never served as a cost-of-living index. The CPI "affects more people's lives," observed Clinton McCully, a lead researcher in the Consumption Branch of the BEA, which publishes the PCE. "A lot of money is riding on the CPI that's not riding on the PCE—things like Social Security adjustments [and commercial] contracts."

But like its more famous cousin, the PCE also was revamped in the mid-1990s. Seeking to hone the index as a macroeconomic indicator, the BEA reworked the PCE into an index whose formula, like the C-CPI-U, accounts for broad month-to-month changes in consumption patterns. In 1995, the BEA added a PCE index that excluded food and energy prices, analogous to the CPI "core" price index.

Chasing the dream

Tremendous progress has been made in measuring inflation; thanks to more than 100 years of research

and development, today's price indexes are far more accurate than the first food and commodity price indexes. But the work begun in the 1800s by pioneers such as Laspeyres and Jevons and carried forward into the 21st century by teams of researchers at the BLS and the BEA is not yet finished.

Economists and statisticians still struggle with technical problems that dogged researchers in the 1930s. For example, new products trickle quite slowly into the CPI, despite the BLS's efforts to continuously update its basket of goods. It can take up to four years for the latest tablet computer to be fully represented in the national index as price takers make their rounds of stores across the country. So, aggregate monthly price changes for that item are missed. (This type of omission also affects the PCE, because 75 percent to 80 percent of personal expenditures tracked by the PCE consist of CPI price data.)

Another challenge is fully accounting for quality change in certain markets such as computer software, health care and the airline industry. Medical technology, for instance, is in constant flux, with new treatments quickly supplanting the old. If a hospital charges more for hip replacement surgery, how much of the price increase is due to better quality—improved techniques and more advanced materials that reduce pain and speed healing? Difficult to tell, considering that complete recovery may take years.

Similarly, cinema admissions and airfares defy hedonic analysis; the BLS doesn't try to compare the quality of new movies to last year's or put a value on reduced leg space in jetliners. "There are just things we miss," Greenlees said. "It's not because we don't want to adjust for quality change … it's really that there's no known way to make these kinds of adjustments."

Researchers constantly tweak indexes and experiment with new ones in an ongoing effort to more precisely measure price change. In 2009, the BEA modified the "core" PCE index to include previously omitted meals and beverages purchased away from home. The Cleveland Federal Reserve computes median and trimmed mean versions of the CPI—alternate measures of consumer inflation that eliminate the high and low extremes of prices for all items.

The latest wrinkle in price indexes is the emergence of "web scraping," the practice of crunching price data gathered daily from the Internet. The Billion Prices Project, an index published by the Massachusetts Institute of Technology, tracks prices of over a half million products and services sold online in seven countries. The limitations of such web indexes are obvious—many products and services aren't sold online. But by harnessing the web's instantaneous access and expansive coverage, they have the potential to make inflation measures more timely and accurate.

In his 1887 essay, "Remedies for Fluctuations of General Prices," British economist Alfred Marshall wrote that "an absolutely perfect standard of purchasing power is not only unattainable but even unthinkable."[8] Marshall may be proven right, but that possibility doesn't deter economists and statisticians from chasing the dream. "The CPI is a little like cosmology; there's always a new problem to explore," Triplett said. "You're always looking to make changes, and in doing research, that opens up another question sometimes." R

Endnotes

1 Thomas M. Humphrey. 1979. "The Concept of Indexation in the History of Economic Thought." In *Federal Reserve Readings on Inflation,* Federal Reserve Bank of New York, 241.

2 Hayes, Matthew. 2011. "The Social History of Quantifying Inflation: A Sociological Critique." *Journal of Economic Issues* 45 (1), 101.

3 Goldberg, Joseph, and William T. Moye. 1985. *The First Hundred Years of the Bureau of Labor Statistics.* Bulletin 2235, Bureau of Labor Statistics, 138.

4 Ibid., 103-04.

5 Reinsdorf, Marshall, and Jack E. Triplett. 2006. "A Review of Reviews: Ninety Years of Professional Thinking About the Consumer Price Index." In *Price Index Concepts and Measurement,* National Bureau of Economic Research, 12.

6 Goldberg, 156.

7 Greenspan, Alan. 1995. Testimony before the Committee on Finance, U.S. Senate, March 13.

8 Humphrey, 246.

Research Digest

In this issue, Research Digest *summarizes recent work by* **Fabrizio Perri** *and* **Vincenzo Quadrini**.

Fabrizio Perri

Not-So-Great Expectations

Research suggests that a loss of confidence in asset values contributed to the global recession.

The Great Recession was remarkable for both its depth—slumping economic output, plummeting asset prices, heavy job losses—and the similar way it played out across the industrialized world. Output, spending, investment and employment were all hit hard in 2008, both in the United States and in other industrialized countries such as Canada, Japan, Germany and France. It was as if many of the world's leading economies had jumped off a cliff together.

Recent research by University of Minnesota economist Fabrizio Perri, a consultant to the Minneapolis Fed, and Vincenzo Quadrini, an economist at the University of Southern California, offers an explanation for why nations marched in lock step into the downturn.

In "International Recessions" (Minneapolis Fed Staff Report 463, online at minneapolisfed.org), the authors describe a self-reinforcing process in which tighter credit conditions arise from expected low values for company assets. Tighter credit in turn stunts economic growth. Cross-border financial ties ensure that pessimism takes hold globally, triggering widespread economic woe.

"In a financially integrated world, these expectations are coordinated across countries," Perri said in an interview. "If a crisis of this type happens, it's necessarily a global crisis."

All together now

Not since the Great Depression of the 1930s have so many countries moved in such macroeconomic unison as they did during the 2007-09 recession. Perri and Quadrini show, for example,

Research Digest

In this issue, Research Digest *summarizes recent work by* **Fabrizio Perri** *and* **Vincenzo Quadrini**.

Fabrizio Perri

Not-So-Great Expectations

Research suggests that a loss of confidence in asset values contributed to the global recession.

The Great Recession was remarkable for both its depth—slumping economic output, plummeting asset prices, heavy job losses—and the similar way it played out across the industrialized world. Output, spending, investment and employment were all hit hard in 2008, both in the United States and in other industrialized countries such as Canada, Japan, Germany and France. It was as if many of the world's leading economies had jumped off a cliff together.

Recent research by University of Minnesota economist Fabrizio Perri, a consultant to the Minneapolis Fed, and Vincenzo Quadrini, an economist at the University of Southern California, offers an explanation for why nations marched in lock step into the downturn.

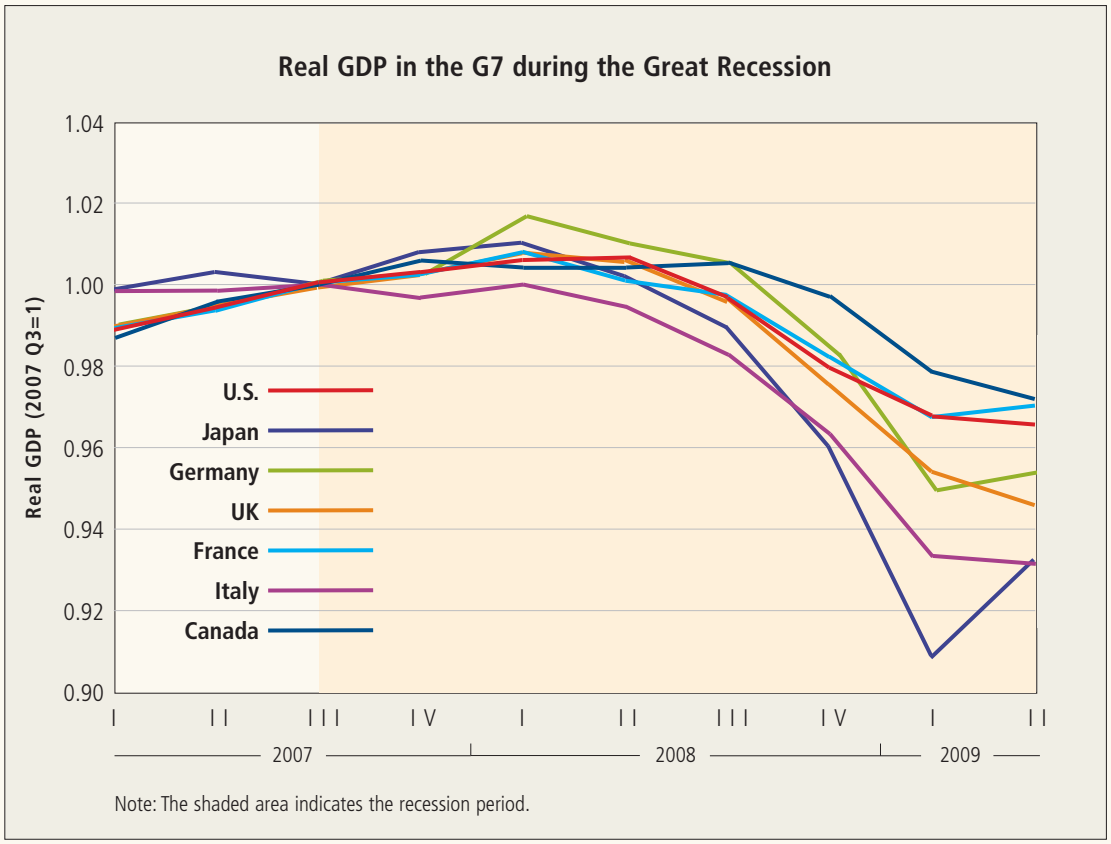
In "International Recessions" (Minneapolis Fed Staff Report 463, online at minneapolisfed.org), the authors describe a self-reinforcing process in which tighter credit conditions arise from expected low values for company assets. Tighter credit in turn stunts economic growth. Cross-border financial ties ensure that pessimism takes hold globally, triggering widespread economic woe.

"In a financially integrated world, these expectations are coordinated across countries," Perri said in an interview. "If a crisis of this type happens, it's necessarily a global crisis."

All together now

Not since the Great Depression of the 1930s have so many countries moved in such macroeconomic unison as they did during the 2007-09 recession. Perri and Quadrini show, for example,

Research Digest



Note: The shaded area indicates the recession period.

that economic output in the seven largest industrialized countries followed a nearly identical path (see accompanying chart); gross domestic product in the United States and in other Group-of-Seven nations peaked in the first quarter of 2008, then nosedived that fall. Financial markets on three continents also moved as one; in both the United States and the G7, stock prices tumbled and commercial lending fell sharply in the midst of the recession.

Why did the Great Recession affect so many countries at the same time, with similar consequences for firms, workers and investors? The economists attribute this striking global synchrony—and the severity of the recession—to a broad and deep contraction in lending.

Other investigators have studied credit shocks—unexpected changes in the liquidity of capital—as the impetus for international business cycles. But most studies have treated such shocks as originating outside the economy of a particular country. An "exogenous" shock may entail a change in credit conditions—tighter credit access, for instance—in one country that affects the economic performance of other countries.

Research Digest

The authors describe a self-reinforcing process in which tighter credit conditions arise from expected low values for company assets. Tighter credit in turn stunts economic growth. Cross-border financial ties ensure that pessimism takes hold globally, triggering widespread economic woe.

But such a scenario, in which shocks are transmitted from one country to the next, doesn't fit the pattern of deteriorating credit conditions during the downturn. If credit had tightened only in one country, domestic firms could have borrowed instead from unconstrained lenders in other countries. But during the recent recession, access to credit diminished in London, Tokyo, Frankfurt and New York, roughly at the same time and in equal measure.

Perri and Quadrini attempt to explain this worldwide credit contraction by proposing an additional "endogenous" mechanism for multinational credit shocks—a seismic shift in the global credit environment made possible by financial bonds among countries.

Altered states

To make their case, the researchers develop a two-country economic model representing the United States and the rest of the G7 in which companies rely on credit to finance hiring and to pay dividends to shareholders. Access to credit in the model depends on how lenders perceive the value of firms' physical assets—facilities, inventory and other holdings that would be liquidated in the event of a loan default.

When lenders anticipate low resale values for these assets, they withhold credit, fearing that they won't be able to recover the full amount of loans made to defaulting firms. Constrained credit deprives healthy firms of capital they need to buy the assets of defaulting firms, depressing market prices—and confirming the expectations of lenders. Credit remains tight, forcing firms to lay off workers and causing declines in consumer spending and investment.

Conversely, when lenders expect high resale values for the assets of defaulting firms, they lend freely, keeping market prices high and stimulating hiring and economic growth.

Thus, rational expectations for asset values become self-perpetuating, locking the economy into one of two possible stable states, or equilibria—one with loose credit and the other with tight credit. Crucially, in the model, altered price expectations in one country instantly change expectations and credit conditions in the other, because of the web of cross-border ownership, investment and banking arrangements linking financial markets today.

Perri and Quadrini don't address *why* expectations change, but point to the bankruptcy of the Lehman Brothers investment house in September 2008 as a watershed event in the Great Recession. "The Lehman default could be interpreted as the trigger that switched the world economy from an equilibrium with globally loose credit to an equilibrium with tight credit and shortage of liquidity, causing widespread contraction in economic and financial activities," they write.

Macroeconomic impacts in the model match up well with the behavior of real-world economies wracked by financial crises. Increasing use of credit before the crisis gradually increases employment, helping to lift the economy. But when credit suddenly tightens, firms are forced to quickly shed workers, causing a marked downturn in economic activity. The longer the period of credit expan-

Research Digest

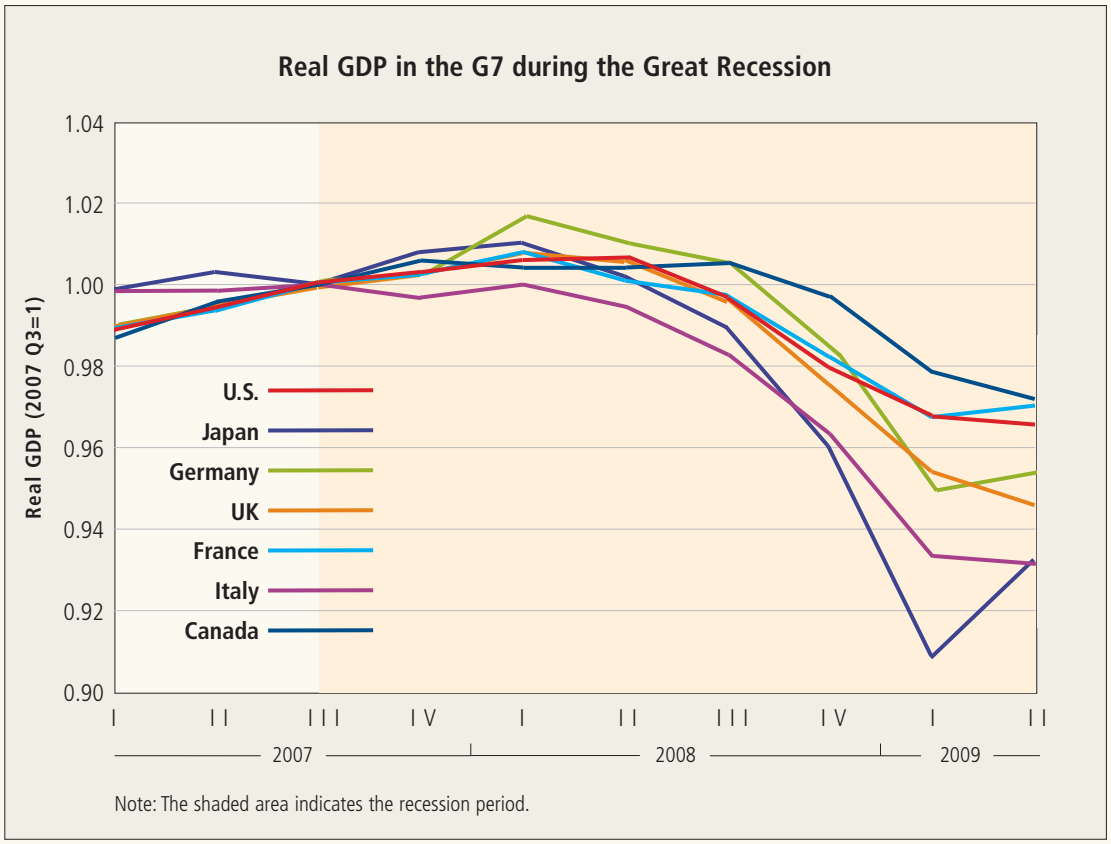
When credit suddenly tightens, firms are forced to quickly shed workers, causing a marked downturn in economic activity. The longer the period of credit expansion, the harder the fall when lenders' expectations change.

sion, the harder the fall when lenders' expectations change. Hence, the sharp drop in employment and GDP in the G7 countries during the last recession, after several years of strong—but perhaps unsustainable—credit growth.

The idea of global economic upheaval resulting from a worldwide shift in expectations has important policy implications, the economists note. If the greatest danger in a financially borderless world is fear of economic distress, then governments can take coordinated action to restore calm by, for example, providing financial markets with liquidity to support asset prices.

—Phil Davies

Research Digest



Note: The shaded area indicates the recession period.

that economic output in the seven largest industrialized countries followed a nearly identical path (see accompanying chart); gross domestic product in the United States and in other Group-of-Seven nations peaked in the first quarter of 2008, then nosedived that fall. Financial markets on three continents also moved as one; in both the United States and the G7, stock prices tumbled and commercial lending fell sharply in the midst of the recession.

Why did the Great Recession affect so many countries at the same time, with similar consequences for firms, workers and investors? The economists attribute this striking global synchrony—and the severity of the recession—to a broad and deep contraction in lending.

Other investigators have studied credit shocks—unexpected changes in the liquidity of capital—as the impetus for international business cycles. But most studies have treated such shocks as originating outside the economy of a particular country. An "exogenous" shock may entail a change in credit conditions—tighter credit access, for instance—in one country that affects the economic performance of other countries.

Research Digest

The authors describe a self-reinforcing process in which tighter credit conditions arise from expected low values for company assets. Tighter credit in turn stunts economic growth. Cross-border financial ties ensure that pessimism takes hold globally, triggering widespread economic woe.

But such a scenario, in which shocks are transmitted from one country to the next, doesn't fit the pattern of deteriorating credit conditions during the downturn. If credit had tightened only in one country, domestic firms could have borrowed instead from unconstrained lenders in other countries. But during the recent recession, access to credit diminished in London, Tokyo, Frankfurt and New York, roughly at the same time and in equal measure.

Perri and Quadrini attempt to explain this worldwide credit contraction by proposing an additional "endogenous" mechanism for multinational credit shocks—a seismic shift in the global credit environment made possible by financial bonds among countries.

Altered states

To make their case, the researchers develop a two-country economic model representing the United States and the rest of the G7 in which companies rely on credit to finance hiring and to pay dividends to shareholders. Access to credit in the model depends on how lenders perceive the value of firms' physical assets—facilities, inventory and other holdings that would be liquidated in the event of a loan default.

When lenders anticipate low resale values for these assets, they withhold credit, fearing that they won't be able to recover the full amount of loans made to defaulting firms. Constrained credit deprives healthy firms of capital they need to buy the assets of defaulting firms, depressing market prices—and confirming the expectations of lenders. Credit remains tight, forcing firms to lay off workers and causing declines in consumer spending and investment.

Conversely, when lenders expect high resale values for the assets of defaulting firms, they lend freely, keeping market prices high and stimulating hiring and economic growth.

Thus, rational expectations for asset values become self-perpetuating, locking the economy into one of two possible stable states, or equilibria—one with loose credit and the other with tight credit. Crucially, in the model, altered price expectations in one country instantly change expectations and credit conditions in the other, because of the web of cross-border ownership, investment and banking arrangements linking financial markets today.

Perri and Quadrini don't address *why* expectations change, but point to the bankruptcy of the Lehman Brothers investment house in September 2008 as a watershed event in the Great Recession. "The Lehman default could be interpreted as the trigger that switched the world economy from an equilibrium with globally loose credit to an equilibrium with tight credit and shortage of liquidity, causing widespread contraction in economic and financial activities," they write.

Macroeconomic impacts in the model match up well with the behavior of real-world economies wracked by financial crises. Increasing use of credit before the crisis gradually increases employment, helping to lift the economy. But when credit suddenly tightens, firms are forced to quickly shed workers, causing a marked downturn in economic activity. The longer the period of credit expan-

Research Digest

When credit suddenly tightens, firms are forced to quickly shed workers, causing a marked downturn in economic activity. The longer the period of credit expansion, the harder the fall when lenders' expectations change.

sion, the harder the fall when lenders' expectations change. Hence, the sharp drop in employment and GDP in the G7 countries during the last recession, after several years of strong—but perhaps unsustainable—credit growth.

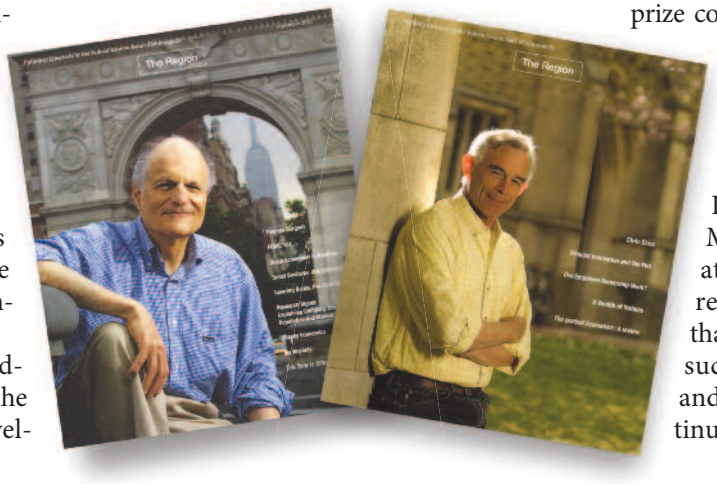
The idea of global economic upheaval resulting from a worldwide shift in expectations has important policy implications, the economists note. If the greatest danger in a financially borderless world is fear of economic distress, then governments can take coordinated action to restore calm by, for example, providing financial markets with liquidity to support asset prices.

—Phil Davies

Thomas Sargent and Christopher Sims

2011 Nobel Laureates

In the early morning of Oct. 10, Thomas Sargent and Christopher Sims each received a phone call from Sweden informing them that they had been selected as joint winners of the 2011 Nobel Memorial Prize in Economic Sciences. They were undoubtedly the only economists who were surprised to hear that news. Their pioneering scholarship has long been recognized as fundamental to the theory and practice of macroeconomics. As the Nobel Committee noted, "Their combined work constitutes a solid foundation for modern macroeconomic analysis. It is hard to envisage today's research without this foundation."



Working independently in the 1970s, the two economists developed methods for answering central questions about the relationship between economic policy and macroeconomic variables such as GDP, inflation, employment and investment. Because policy and the macroeconomy affect one another, it can often be hard to distinguish between cause and effect. Sargent and Sims developed distinct techniques, each with advantages in particular circumstances, for teasing out the direction of causality.

"Almost 40 years later, their thinking informs the making of macroeconomic policy around the world," said Narayana Kocherlakota, president of the Minneapolis Fed and formerly on the faculty of the University of Minnesota, where Sargent and Sims worked in the 1970s and 1980s. "I'm especially proud that much of the research recognized by the prize committee was done at the Federal Reserve Bank of Minneapolis and the University of Minnesota. My predecessors at the Federal Reserve Bank of Minneapolis deliberately fostered a research environment that could give rise to such important work, and this tradition continues today." R

Patrick Kehoe, Minneapolis Fed, Princeton, University of Minnesota

"Their combined work constitutes a solid foundation for modern macroeconomic analysis. It is hard to envisage today's research without this foundation."

—Nobel Committee

Kei-Mu Yi, Minneapolis Fed

Ellen McGrattan, Minneapolis Fed and University of Minnesota

Nancy Stokey, University of Chicago

Robert Litterman, (1980 Ph. D., U of Minn.), Kepos and HHEI Advisory Boards

Photos by Everett Ayoubzadeh

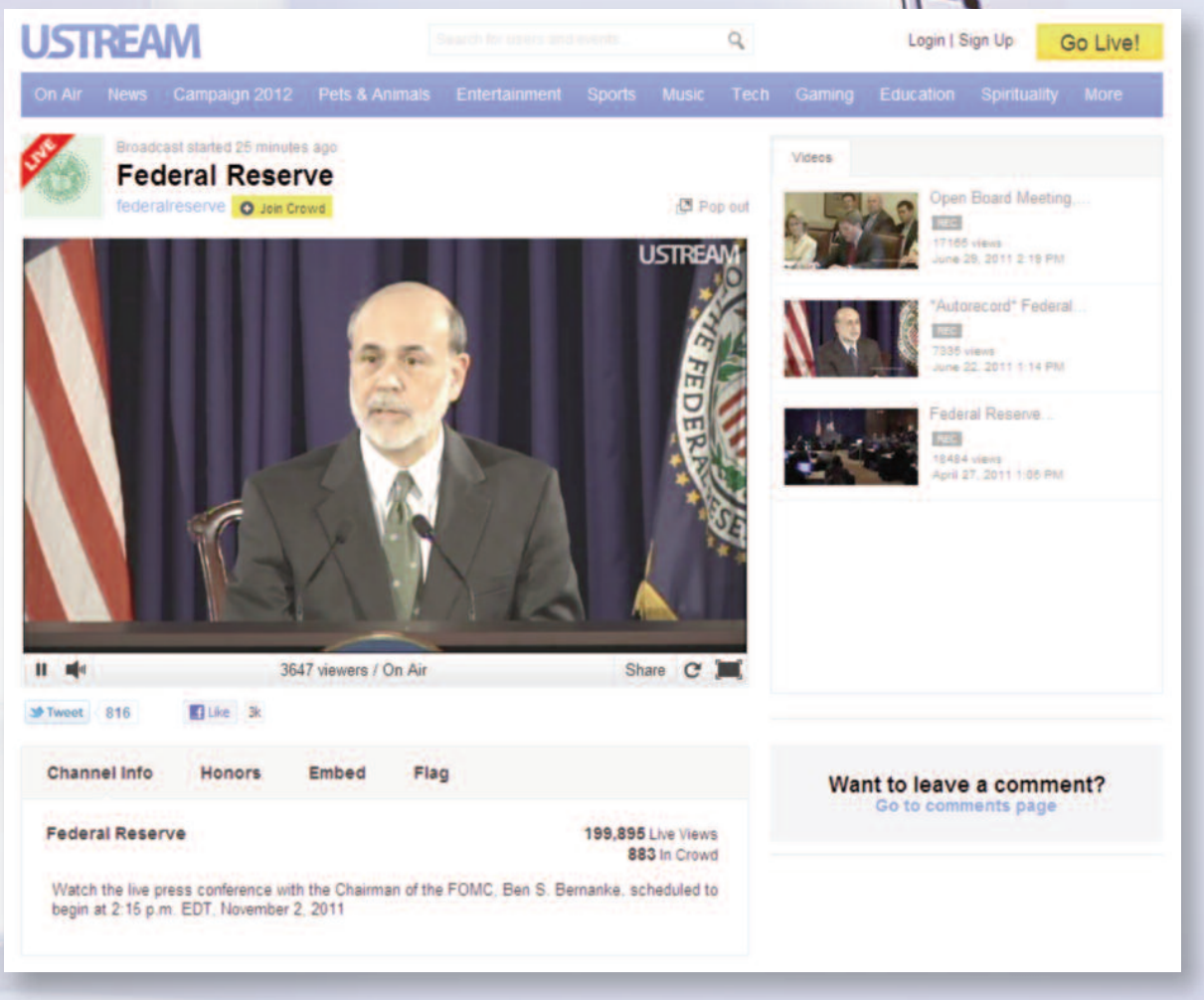
For more about Sargent, read the September 2010 *Region* interview
http://www.minneapolisfed.org/publications_papers/pub_display.cfm?id=4526

For more on Sims, go to the June 2007 *Region*
http://www.minneapolisfed.org/publications_papers/pub_display.cfm?id=3168

For news and background on their 2011 Nobel Prizes, see
http://www.nobelprize.org/nobel_prizes/economics/laureates/2011/#

Photos above are from the Nov. 17 reception in honor of Sargent and Sims, sponsored by the University of Minnesota and the Minneapolis Fed.

Virtual Fed



Transparency.tv

There are good reasons to believe that monetary policy should be made with a high degree of transparency, rather than behind closed doors. Doing so gives the public a good sense of policymakers' goals and minimizes destabilizing surprises when policies change.

As part of its ongoing efforts to bring more transparency to monetary policy, the Federal Reserve Board of Governors announced last March that Chairman Ben Bernanke would hold press conferences four times a year. These sessions are held immediately after those quarterly meetings of the Federal Open Market Committee when it releases its economic outlook. After a brief statement reviewing the FOMC's policy decision, in the context of its economic projections and policy strategy, the chairman answers reporters' questions about the outlook and the thinking behind the Committee's actions. The most recent post-FOMC press conference was Nov. 2, 2011.

But you don't have to be a reporter to listen in. The Board is live streaming—and archiving—the conferences as well as other events, so anyone with an Internet connection can be a fly on the wall. Buzz in, virtually, at: http://www.ustream.tv/federalreserve. Archives are available at http://www.federalreserve.gov/mediacenter/media.htm.

—Joe Mahon