Federal Reserve Bank of Richmond

Technology Diffusion and Productivity

Richmond, Virginia • June 7, 2019

Richmond Fed Research Conferences in 2019

April 19

University of Virginia-Richmond Fed Research Workshop (Spring)

May 16–17

Market Structure and the Macroeconomy

May 22–23

Technology-Enabled Disruption: Implications for Business, Labor Markets, and Monetary Policy

June 7

Technology Diffusion and Productivity Workshop

September 27

Regional Economics Workshop

October 2

Investing in Rural America

October 11

Richmond Fed-University of Virginia Research Workshop (Fall)

For more information, please visit: www.richmondfed.org/conferences_and_events

FEDERAL RESERVE BANK OF RICHMOND®

Richmond ▪ Baltimore ▪ Charlotte

Technology Diffusion and Productivity Workshop Agenda

June 7, 2019

9:00 AM	**Technology Diffusion: Evidence and Theory** **Nancy Stokey** University of Chicago
10:15 AM	**Idea Flows and Economic Growth** **Francisco Buera** Washington University in St. Louis **Robert Lucas** University of Chicago
11:30 AM	**Declining Search Frictions, Unemployment, and Growth** **Guido Menzio** New York University
12:30 PM	**Lunch Speaker** **Tom Barkin** Federal Reserve Bank of Richmond
2:00 PM	**Assessing the Gains from E-Commerce** **Pete Klenow** Stanford University
3:15 PM	**Two-Sided Market, R&D, and Payments System Evolution** **Zhu Wang** Federal Reserve Bank of Richmond
4:30 PM	**Innovation, Knowledge Diffusion, and Selection** **Danial Lashkari** Boston College

Conference Summary

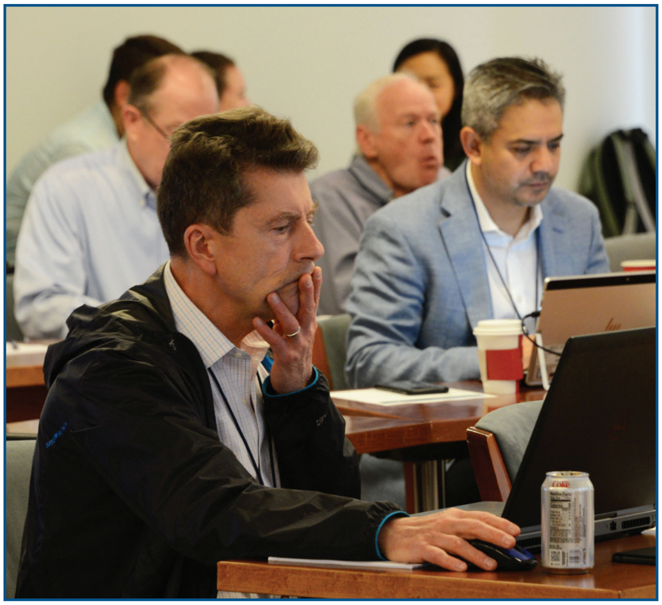
On June 7, economists gathered at the Richmond Fed for a workshop on "Technology Diffusion and Productivity." Nancy Stokey of the University of Chicago kicked off the event by addressing the big-picture question: Why should we be interested in technology diffusion? One answer, she proposed, is that a great deal of historical evidence shows that long-run economic growth comes from technical change — and this has been the case not only in the United States, but also in every other developed country. Moreover, the evidence suggests that differences in technology are what distinguish higher-income countries from lower-income countries. The story of technology, according to Stokey, involves more than just the invention of new machines and processes. A crucial element is how quickly and widely the new tools and ideas get adopted. This, in turn, will depend on the trade-off between how much the technology reduces ongoing costs versus how much it costs up front. She discussed many of the seminal empirical studies on technological diffusion, starting with Zvi Griliches's analysis of the diffusion of hybrid corn in the Midwest during the 1930s. (See *Econometrica*, October 1957.) A key takeaway of this work and the subsequent work it inspired is that good ideas that make economic sense tend to diffuse rapidly. Stokey discussed how this has been the case in many diverse industries, including brewing, bituminous coal, iron, steel, and railroads. To illustrate the importance of cost-benefit trade-offs for technology diffusion, she pointed to the example of tractors in agriculture. Although tractors had been around since about 1910, their adoption had been slow initially. Then they really took off during World War II due to labor shortages caused by the war effort.

Francisco Buera of Washington University in St. Louis and Robert Lucas of the University of Chicago discussed their paper, "Idea Flows and Economic Growth." They presented a variety of theoretical models that explore the "common idea that people learn from interactions with other people." One of the motivations for this research is to explain the observation that output per worker in the United States has grown at an average rate of about 2 percent per year for the past 150 years. Economists have attributed this labor productivity growth to better technology, better institutions, and total factor productivity growth. But, in some sense, these are just labels. Buera and Lucas argued that models of the "collective learning process" have the potential to provide a deeper explanation for long-run labor productivity growth. They began their discussion with a "bare-bones" model of an economy in which the evolution of individuals' productivity levels is determined by their random meetings with other individuals. In the basic model, each individual's productivity improves when he or she is fortunate enough to meet a more productive person. From this starting point, Buera and Lucas explored models with additional features and levels of complexity. In some of the models, the simple process of diffusion through chance meeting is sufficient to generate sustained growth. In others, this growth is augmented by additional forces. The authors expressed their hope that theoretical models like the ones they presented will become the basis for empirical studies that can relate worker earnings to education, work experience, and other factors.

Guido Menzio of New York University discussed his paper, "Declining Search Frictions, Unemployment, and Growth." The paper focuses on the surprising long-run stability of the Beveridge curve in the face of technological change. The Beveridge curve describes the historical relationship between unemployment and job vacancy rates. At the top of a business cycle, low unemployment is associated with a

high vacancy rate; at the bottom of the cycle, high unemployment is associated with a low vacancy rate. The long-run stability of the relationship has stood in contrast to the instability of many other macro relationships, such as the Philips curve and the money demand function. Menzio observed that economists have invoked technological change to explain the instability of some of these relationships. In this light, the long-term stability of the Beveridge curve is even more surprising because the theoretical models that have been built to explain it are based on search costs that arise due to incomplete information. These search costs allow for unemployment to exist side by side with job vacancies. But why has the Beveridge curve remained stable over the past 100 years when search technology has improved so dramatically? Menzio addresses this question by building a model that incorporates a worker's opportunity cost of taking a job (as well as an employer's opportunity cost of hiring a worker). In his model, these opportunity costs increase as information technology becomes more efficient. Improved search efficiency increases the speed of finding potential matches between workers and employees, but it also increases the opportunity cost of committing to a match. Workers encounter more job openings, so they become choosier. The Beveridge curve remains stable when these two effects counterbalance each other.

Pete Klenow of Stanford University discussed his recent paper, "Assessing the Gains from E-Commerce." The paper analyzes a large database of Visa credit card and debit card transactions and yields a number of interesting results. For instance, Klenow used the data to estimate that e-commerce has become about 8 percent of U.S. consumption, which is larger than a 6 percent estimate produced by the U.S. Bureau of Economic Analysis. He found a great deal of heterogeneity in e-commerce usage, with much higher usage rates in cities than in rural areas. This runs counter to the idea that rural areas might have had greater usage due to the convenience of online shopping as a substitute for stores that are not so close physically. Another finding is that higher-income households use cards more intensively than lower-income households. Klenow's analysis divides the consumer gains from e-commerce into two categories: first, the gains from convenience that are created by being able to buy an item online versus having to drive to a store, and second, the availability of a wider array of items online versus at a physical store. The paper estimates the division of gains by using data on the distances between cardholders and the brick-and-mortar stores closest to them, which provides information about the transportation cost savings of buying at, say, the Gap's online store instead of its physical store. Klenow estimates that the overall welfare gains from e-commerce amount to about 1 percent of U.S. consumption. Of this total, about 0.4 percent came from convenience gains and about 0.6 percent came from increased availability.



Zhu Wang of the Richmond Fed discussed his paper, "Two-Sided Market, R&D, and Payments System Evolution." Wang observed that the payment-card industry provides platforms that facilitate transactions between buyers and merchants. An interesting feature of the industry is that new technologies have to be adopted by different groups of agents and their decisions interact with each other and create network externalities. Wang's paper attempts to explain two puzzles: one is the arguably slow historical pace of payment-card adoption, and the other is the perception that merchant fees are "too high" in light of technological progress that might have been expected to reduce costs. The paper makes a new contribution to the literature by building a model in which a card network's R&D decisions are endogenous. In the model's competitive equilibrium, the card network charges a substantial markup that slows down the diffusion of electronic payment technology. Wang contrasts this case with the socially optimal outcome, which is characterized by more rapid diffusion. One of the paper's interesting results has to do with the effects of different government policies. He finds that policies that impose marginal cost pricing on card fees negatively affect R&D because the policies remove the incentive for networks to invest to reduce marginal cost. The analysis finds that the type of regulation that is currently implemented by the United States, Australia, and many countries in Europe may deliver better outcomes. Regulations in these countries often have taken the form of simple caps on card fees charged to merchants — a restriction that provides an incentive for networks to invest and lower costs.

Danial Lashkari of Boston College discussed his paper, "Innovation, Knowledge Diffusion, and Selection." The paper explores Schumpeter's idea of creative destruction, which can be illustrated using the example of Walmart, according to Lashkari. The firm built and improved on many of the ideas that had existed in the retail industry and eventually became so efficient that it drove many smaller retailers out of business. In the paper's model, such a process involves both positive and negative externalities. The positive externality occurs when the benefits of Walmart's new ideas spill over to other firms in the industry. The negative externality is the loss of potential new ideas and innovations due to the exit of Walmart's competitors from the industry. The model is based on a monopolistically competitive industry in which firms face fixed costs. Startups and incumbents both attempt to innovate and gain market share. They face positive and negative productivity shocks, and firms are forced to exit the industry when their productivity declines below a certain level. The paper focuses on the model's policy implications, which depend on firm size. As firms grow, the extent to which they stifle innovation by driving other firms out of business starts to overtake the positive spillover effects of their own innovations. Lashkari argued that optimal tax policy should take this trade-off into account by conditioning a firm's R&D tax credit on its market share. It turns out, he noted, that the current U.S. tax regime mimics certain elements of the model's optimal tax policy because the U.S. tax regime includes a binding cap on the amount of investment expenditure that a firm can use to gain tax benefits. ■

For more information on the research discussed at the conference, please see the interviews that follow. Interviews have been edited for length and clarity.

Technology Diffusion: Evidence and Theory

by Nancy Stokey

Economists have long recognized the importance of technological advancement for economic development. And while the invention of new machines and processes is a key part of the story, the pace of diffusion and widespread adoption of innovations is just as crucial. In her keynote presentation, Nancy Stokey of the University of Chicago gave a "bird's-eye view" of the economics literature on technology diffusion.

In your paper/presentation, you addressed the fundamental questions: What is technology diffusion, and why do we care about it?

Stokey: There's a wide variety of evidence that pretty clearly says long-run economic growth comes from technological change. That's true in the United States, it's true in every other developed country, and if you look at what distinguishes higher-income countries from lower-income countries, there again the evidence seems very compelling that it's differences in technology. So if we want to think about where growth comes from or how we could speed up development in the less-developed world, technology is at the heart of it.

Technological change has two parts. One is innovation or invention. The other part is, once a new product or a new process is invented, how quickly will it be adopted? And we see that the speeds vary a lot, just looking even casually at the historical evidence. Some innovations get adopted very quickly, while others take a long, long time to become widespread.

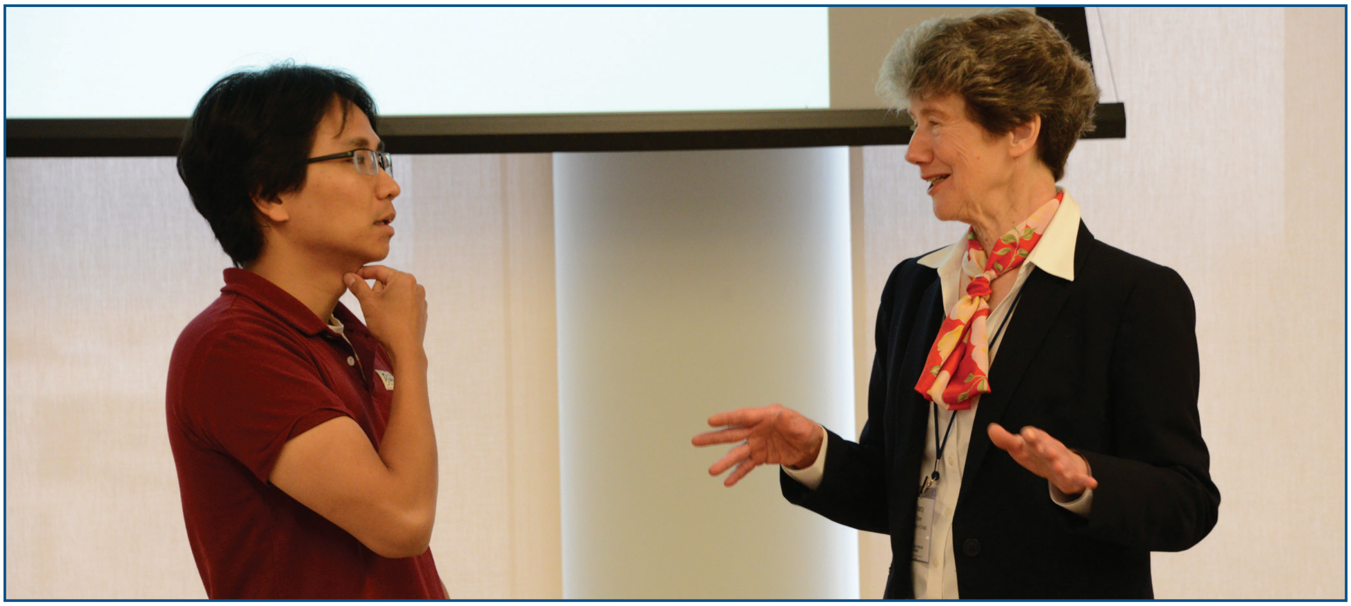
Could you please discuss some of the factors that affect the speed of technology diffusion?

Stokey: For the diffusion of producer technologies, used in the production of final goods and services, we expect that a key part of the story is how significantly the technology reduces costs. Another important part of the story is the up-front cost of the new technology itself. Many new technologies involve substantial investments in new equipment. So, on a very basic level, the speed of adoption depends on how large the cost savings are relative to the up-front costs.

You reviewed a number of case studies in your presentation. Maybe you could elaborate on Zvi Griliches's seminal study of hybrid corn (which appeared in* Econometrica *in October 1957).

Stokey: Griliches's study of hybrid corn is really the earliest one where anyone looked carefully at the adoption of a particular new technology. I like Griliches's study especially because it has two parts. One is just looking at the raw data, and the U.S. Department of Agriculture collected data that are very good and detailed for this technology. Griliches first distilled that raw evidence down to a couple of key parameters. He then used those parameters to describe how soon the technology was introduced, how quickly it spread, and how completely it swept across different geographic areas.

The hybrids were developed by the Department of Agriculture starting in the 1930s, and the hybrids had to be developed for local geographic areas. It wasn't just one hybrid; they were all a little bit different and tailored to local areas. Not surprisingly, they were first developed for the regions that were the most densely planted in corn, in what Griliches called the "corn belt" states, such as Iowa, Illinois, and Indiana.



The Department of Agriculture really did the innovation, and then they later passed the hybrids off to the private seed companies. Griliches says that the farming density mattered for how the marketing was done by the seed companies. The density affected how many doors they could knock on, or whether they could just supply it through a local store. If the area was very intensely planted in corn, it would also be a bigger market, and so more profitable for the seed growers. From the Department of Agriculture's point of view as well — as a matter of public policy, it makes sense to target the biggest markets first, areas with a lot of acreage planted in corn.

Your presentation also covered the work of another seminal contributor to the field, Edwin Mansfield, who looked at twelve major innovations in four industries.

Stokey: Mansfield's study is very much modeled on the study by Griliches, but it has some important differences. Mansfield looked at innovations in industry rather than agriculture and in industries with large firms. In addition, the innovations he looked at are different because they involve major investments in new equipment rather than just buying slightly more material inputs. Nevertheless, he finds that Griliches's main result is confirmed.

Griliches found that you can explain a lot on the basis of profitability — that the most profitable regions got the innovation the earliest and adopted it the fastest and most completely. Mansfield finds exactly the same pattern. And he finds that pattern for twelve innovations in four different industries. So Mansfield's evidence provides a lot of confirmation that what Griliches found was not particular to the hybrid-corn case.

It's interesting that he studied a lot of different industries, including bituminous coal, iron, steel, brewing, and the railroads. But the gist of it is if there's a good idea, it tends to be taken up pretty rapidly.

Stokey: One of the quickest was metal cans for beer. Those were a big hit: within a year or so, all the big sellers wanted to offer beer in cans. Evidently the cans were more convenient for a lot of customers.

They were definitely easier to crunch up. Your presentation also touched on the adoption of farm tractors during the 1940s.

Stokey: Cost reduction is also Manuelli and Seshadri's explanation for the adoption of tractors (from their 2014 article in the *American Economic Review*). The pace of adoption really took off during World War II, when the manpower shortage led to large wage increases. So their explanation makes perfect sense: a man with a tractor can work many more acres per day than a man with a team of horses, so tractors were a way to economize on manpower.

You spoke about agriculture and pointed to some really striking cross-country patterns. Maybe you could discuss some of the stylized facts and how they relate to the U.S. experience.

Stokey: Two patterns are very obvious in comparing productivity and employment in developing and developed countries. First, developing countries have lower labor productivity in both agricultural and nonagriculture sectors. Second, developing countries have much larger employment shares in agriculture.

There is nothing surprising about these two facts. But there is a third fact that is more subtle and is very surprising: the labor productivity gap in developing countries is even larger in agriculture than in nonagriculture sectors. Thus, developing countries seem to be specializing in the sector where they have a comparative disadvantage.

All three of the broad patterns one sees across countries in modern data appear as well if one looks at time series for countries that are already developed. Thus, they seem to be common to all economies over time and across geographic regions.

Are there any broad points you would like to make about technology and policy?

Stokey: One important question is why poor countries can't — or don't — develop faster than they do. We want to know how they can increase their wage rates and incomes faster. If you look at the history of the World Bank and how many different ideas they have tried over the decades, I would say they have not had such great success on the whole. They've spent lots of money, but the question is how much it has contributed to growth in the less-developed world. Unfortunately, I would say it's not clear that it's very much.

On the other hand, you can look at countries that have grown very rapidly in the modern era, such as Japan, Taiwan, and South Korea in the second half of the twentieth century, then in mainland China and now Southeast Asia. The lesson from these countries, I would say, is that what spurs growth is adopting new technologies and investing in new kinds of production.

Africa is one region that has still not done well economically. What would speed up growth there? One hypothesis is that they would benefit greatly from bringing in new technologies — new to them. Investing in manufacturing plants is an obvious choice. For example, textile production is an industry that countries often get into at an early stage of development. A second point — and this concerns developed countries — is that people are sometimes too pessimistic about future growth. They think that the days of income growth in the U.S. are over, that growth is slowing down or it's going to stop. It's this notion that all the good stuff has already been invented and there's not much left to do. This argument has come around every once in a while for centuries. It hasn't proven to be true yet, and I doubt if it's true now.

Finally, an important question my paper doesn't address is whether things are somewhat different in the modern world because industries are less competitive. They have fewer firms, they do a lot of innovation in-house, and they aren't eager to let their rivals have access to their ideas and inventions.

What would be a good example of this? The information technology sector?

Stokey: Yes. For many tech firms, legal cases are such a big issue now. Defending patents and fighting patent infringement cases seem as important as innovation. Some monopoly power over a new innovation is needed to provide an incentive to invest in R&D. But if firms are allowed to exercise monopoly power too vigorously, that's not good for society either. The same issue comes up in pharmaceuticals. If a company has developed a new drug and it's lifesaving for a group of people, there's almost no limit on the price they can charge. It's an old problem, but it seems to be more acute now than it was in the past. ■

Declining Search Frictions, Unemployment, and Growth

by Paolo Martellini and Guido Menzio

The Beveridge curve describes the historical relationship between unemployment and job vacancy rates — a relationship that has remained remarkably stable over the past 100 years. How can this be when job-search technology has changed so substantially? Guido Menzio of New York University and Paolo Martellini of the University of Pennsylvania address this question by building a model that incorporates a worker's opportunity cost of taking a job as well as an employer's opportunity cost of hiring a worker. Menzio discussed their model at the workshop.

Your paper deals with the Beveridge curve. Could you explain what that is?

Menzio: The Beveridge curve was an empirical finding from the late 1950s and early 1960s based partly on earlier observations made by British economist William Beveridge. Economists plotted unemployment rates and vacancy rates for the U.K. and found a very clear downward-sloping relationship, which said that when the vacancy rate in the U.K. was high, the unemployment rate was low, and when the vacancy rate in the U.K. was low, the unemployment rate was high. After that, there was a ton of empirical work about the Beveridge curve, mostly focused on cyclical fluctuations in the Beveridge curve. This work discovered the now well-known fact that, over the business cycle, the Beveridge curve tends to shift.

But the focus of this paper is instead on the long-run stability of the Beveridge curve, which is quite striking in light of how unstable most macroeconomic relationships are. If you think about the Phillips curve, the Phillips curve is all over the place. It's very hard to actually see the curve in the raw data. The demand curve for money is also very unstable.

What role does technological change play in these shifting relationships?

Menzio: A lot of economists think these relationships are unstable because of technological progress. So seeing the stability of the Beveridge curve was quite shocking in light of the leading theory of the Beveridge curve, which was developed mainly by the work of Dale Mortensen and Chris Pissarides in the early 1980s. Their theoretical idea is that unemployment and vacancies coexist because there are so-called search frictions in the labor market. That is, firms and workers don't find each other in some frictionless centralized exchange; instead they have to find each other in the wild, and this process takes time. And if it takes time for an unemployed worker to find a job, it also takes time for a firm to find a qualified worker for a position. So you're going to see the coexistence of unemployment and vacancies.

The entire theoretical explanation of the Beveridge curve rests on the premise that there are informational frictions. And so you would expect declining informational frictions to shift the Beveridge curve. For any given vacancy rates, you would expect the unemployment rate to be lower as technology allows information to be transmitted from firms to workers about their openings more quickly and efficiently. But we don't see that in the long-run data, and this is the starting point of the paper.

So the Beveridge curve has been relatively stable despite major technological changes. What does that tell you?

Menzio: This is a puzzle from the perspective of the theory. There are two ways to look at these data if we accept that information flows more efficiently now than a hundred years ago. Either the theory is completely wrong and the Beveridge curve has nothing to do with informational frictions, or we need to extend the theory to reconcile reduced search frictions with the stability of the Beveridge curve.

Could you give us a broad, intuitive explanation of how your theoretical work reconciles the stable Beveridge curve with improved search efficiency?

Menzio: Yes. So the point of the theory is very simple. It's the idea of thinking about how quickly an unemployed worker may find a job or how quickly a firm may find a worker to fill a job. And think of that speed as the outcome, as the product of two things. One is how easily the worker becomes aware of a job. And presumably this velocity has increased through time. On the other hand, there is a second factor: What are the odds that the worker is a sufficiently good fit for the job to be hired?

The idea is that, if firms become more selective about the workers they are willing to hire, this increased selectivity may offset the increased velocity of information diffusion. And if these two things just happened to balance each other exactly, we might see a stable Beveridge curve in the face of a constantly improving speed of communication.

When a firm takes a worker on board, it is giving up the option of finding someone better for the job. Likewise, from the perspective of the worker, when you take a job, you are giving up the option of fully staying in the market and finding something better. So how picky the firm will be and how picky the worker will be critically depend on the value of these options, and the value of your option depends on how quickly you become aware of alternatives. And so there is this balancing act taking place.

So improved communications increase the efficiency of search and also simultaneously increase choosiness.

Menzio: Yes. It's just one single force that drives up both the speed of contact and the pickiness, balancing things. One force counterbalances the other.

Does that mean your model is based on a bit of a knife-edged balancing act?

Menzio: Yes. This is still somewhat a theoretical exercise in the sense that we do not really believe that this condition exactly holds. And, in fact, the Beveridge curve is not perfectly stable. But it's useful and very much in the spirit of the growth literature. We are interested in finding that exact knife edge under which you would see no effect whatsoever of changes in information technology on the Beveridge curve. We don't believe it's literally true; we just believe it's a useful benchmark, especially looking at the relative stability of the unemployment-vacancy relationship.

There was an earlier literature on conditions for balanced growth in a neoclassical model. One of the stylized facts that has existed in the data is that the labor share remains constant, the interest rate remains constant, while wages go up. So the question was: Why is it the case that as GDP grows, GDP per capita grows and capital grows, the labor share remains constant, wages are growing, and hours worked per worker remain constant? And the solution, or knife edge, was that technological change must be labor-augmenting, and the preferences of the individuals must be such that income and substitution effects exactly cancel out to explain why, as wages rise constantly over time, people decide to work neither more nor less than before. And again, this is not a case where these facts are exactly true; they are merely stylized facts. They are approximately true.

And these knife-edge examples are very useful exercises. From a theoretical point of view, these results are somewhat troubling because they are fragile. But from an econometric point of view, these results are great because they give a very sharp identification of what we need to look for. Whenever a theory requires very special parameters to explain a particular phenomenon, that means that whatever you see in the world has very strong implications for what that means about the structure of the economy.

What are the implications of your model for search in cities versus rural areas?

Menzio: If I look at New York City and if I look at a small town, one market has 10 million workers and the other market has perhaps 100,000 workers. But these markets have roughly the same unemployment rates, the same speed at which unemployed workers find jobs, and the same vacancy-to-unemployment ratios. So people say, well this must mean that being in New York there is no advantage in the search process compared with being in a small town. However, this is no longer true when we live in the knife-edge world that I derived in the first part of the paper. So under the same conditions for which the Beveridge curve remains stable over time, you wouldn't be able to see a major difference in unemployment rates between New York and a small town. This is true even if there were increasing returns to scale in the search process. In New York City, workers find firms more quickly and firms find workers more quickly than in a small town. But this higher speed in New York City is going to be exactly offset by the higher pickiness of workers and firms about the quality of relationships, and you're going to see the same unemployment rate in the big city as in the small town.

What does your model say about our ability to measure search efficiency?

Menzio: If you believe in our theory, you cannot compare New York City to a small town and conclude anything about the returns to scale in the search process by simply looking at unemployment rates. It leaves a bit of an unsatisfactory feeling in your mouth, because you would like to measure these things.

The theory would say that if you want to figure out how much search frictions are going down over time or how much smaller they are in big cities than in small cities, you should look at how more selective firms become over time. We don't have a time series for that, which would be very nice. But we do have two points in time. There was this employment opportunity pilot project in the 1980s where they collected data about the hiring process of firms. And they found that, on average, a firm would look at about twenty applicants for each vacancy. More recently, in the 2010s, we have lots of data coming from online job platforms, where we can actually see how many people applied for particular jobs. And in the 2010s, we find that the number of applications received for each vacancy averaged about forty, which is roughly twice as many as in the 1980s.

Could you give us some of your observations about how your research fits into the broad literature on economics?

Menzio: Adam Smith discussed the idea that the extent of the market matters for how much people can specialize and do the things they are best at. And informational frictions are a constraint on the extent of the market not because they make the market big or small, but because they create delays and informational loss. I think it's a very compelling idea that, as you reduce this informational friction in the market, the market becomes in a sense thicker, and people can find the jobs they are best suited for.

This makes me think about one of the first important papers on search in the labor market. It is a paper by George Stigler from 1961 called "Information in the Labor Market." And it says something very interesting. It gives an example of what would happen if informational frictions were enormous. He said that, in a regime of ignorance, Enrico Fermi would have been a gardener and Von Neumann a checkout clerk at the drugstore. I think that is exactly what my paper is about. ■

Assessing the Gains from E-Commerce

by Pete Klenow et al.

Online shopping can be seen as offering consumers two advantages: convenience and breadth of choice. In his paper, Pete Klenow of Stanford University estimates that the overall welfare gains accruing to consumers from e-commerce are about 1 percent of U.S. consumption. Klenow's analysis finds that about 40 percent of these gains come from convenience and about 60 percent of the gains come from increased breadth of choice.

Your paper draws on a large database of credit and debit card transactions to analyze e-commerce in the U.S. over the past ten years. What are some of the more interesting stylized facts that come out of your work?

Klenow: One surprise is the extent of e-commerce. So, for example, the Bureau of Economic Analysis (BEA) has an estimate that e-commerce was equivalent to something like 5 percent of total U.S. consumption in 2017. But they're focusing on retail categories like clothing and electronics. We have a slightly broader view that incorporates transportation categories, including things like buying airline tickets online. We think that this kind of e-commerce has some of the same benefits as online shopping for electronics and clothing: you get to see a lot more varieties when you book a plane or, for that matter, get an Uber or a Lyft, and you have some more control over which type of car picks you up and exactly where it picks you up.

Including transportation turned out to be nontrivial and got us a significantly bigger number than the BEA's narrower estimate for retail e-commerce. I was surprised to see that, based on our estimates, e-commerce has gotten to be about 8 percent of all U.S. consumption.

We noticed that you found some interesting demographic patterns having to do with people's income and whether they live in rural or urban areas.

Klenow: Right. There's a lot of heterogeneity in e-commerce usage, and a bunch of it is geographic. Cities had much higher usage rates, and rural areas had much smaller percentages. So that kind of jumped out at us because a priori we might have thought rural areas would have higher e-commerce use because of the convenience of being able to access a lot of things online that aren't very close physically. Separately, income is a big predictor of whether you have a credit or debit card at all. If you look at households with incomes of $50,000 or above, the e-commerce share is almost 10 percent. If you look at households with incomes of $50,000 or below, the e-commerce share is below 4 percent.

Could you describe the Visa database that you analyzed? What are some of its advantages and potential limitations?

Klenow: First, we need to recognize that Visa is a transaction network. It's not the bank that issues credit cards or lends to people — it's really just recording the transactions as they happen and communicating that information to relevant parties. So what Visa can see is basically the amount you swipe at your transaction or the amount you enter into an e-commerce transaction. It doesn't see all the items you bought; it just sees the total amount. So if you spend $47 at Amazon, for example, that would just show up as Amazon Marketplace.

So the advantage of this kind of data is that it's got tremendous breadth of coverage, since Visa accounts for more than half of all debit and credit card transactions in the U.S. by volume. The things they don't get, of course, are things like housing and a lot of services that people don't pay for with credit cards.

And the data cover not only e-commerce, but also brick-and-mortar transactions, which shows us something about how consumers substitute between online and offline spending. That allows us to estimate how much consumers value the online option.

Your study makes a distinction between the gains in e-commerce that derive from increased convenience versus gains from increased availability. Could you explain the distinction?

Klenow: With convenience, the idea is that you know you're going to buy a particular, say, pair of jeans at the Gap, and the question is do you go into a brick-and-mortar Gap or do you buy it from Gap.com? So convenience is really not changing what you buy, it's just changing how easy it is to buy it and how convenient it is to buy it.

With availability, the idea is that you prefer a wider array of options. You may shop at Amazon instead of Barnes and Noble, for example, because Amazon has a greater number of titles available than an individual brick-and-mortar bookstore like Barnes and Noble.

What was a relationship you found between a cardholder's proximity to a retailer's brick-and-mortar store and the cardholder's propensity to buy at the same retailer's online site?

Klenow: When you decide to buy at a Gap physical store instead of Gap.com, how much is that a function of how far you are from the nearest Gap? So it's a question about choosing between the same merchant online and offline.

We found a very strong relationship. If you're right on top of a brick-and-mortar store, there is an 88 percent chance of carrying out the transaction in that physical store and only 12 percent of going online. But if you're, say, twenty miles from the nearest physical Gap store, then your probability of buying at that store falls to something like 66 percent and the probability of going online rises from 12 percent to 34 percent. So you see this kind of gravity at work in the retail context. The farther away you are from a physical store, the more likely you are to go online for the same merchant.

So how do you use this information to estimate the convenience gains from online buying?

Klenow: The key thing we do is take advantage of the fact that some households are close to a brick-and-mortar store and some are farther away. We say that's like effective price variation. If you're far away from a physical Gap store, it's as if the prices at that store are effectively higher because you've got to take into account your travel costs, both the direct cost of using a car or public transit, gasoline, depreciation, plus the time cost.

If online and offline shopping were literally perfect substitutes, then everybody who wasn't within a certain proximity of a brick-and-mortar store would go online. You don't see that; you don't see huge substitutability. That suggests they're not really the same product and there are some benefits to being able to shift between e-commerce and brick and mortar.

Your data show that a lot of online buying is from merchants that people don't even visit offline.

Klenow: We estimate that 88 percent of e-commerce spending is at merchants that a given cardholder never buys from offline. Some of that is kind of predictable. Amazon, for example, doesn't have any brick-and-mortar outlets, and a lot of the travel websites don't have brick-and-mortar options. But it also means a bunch of people go to Gap.com and just never go to the physical Gap store.

Maybe you could talk a little bit about your methodology and how you go about measuring the gains of e-commerce that are due to increased variety.

Klenow: Yes. Again we exploit distance because that is like effective price variation. But here we are not so much interested in the choice of Gap versus

Gap.com. We are more focused on the choice between a physical Gap store and, say, Amazon. Those might be less substitutable because you might have a bunch of options at Amazon that you don't see at an individual Gap store.

So when comparison is within a category like clothing, we likewise find this really striking gravity effect. When you're farther away from any brick-and-mortar clothing store, you're much more likely to go online. So that's a very strong pattern we can use. We again monetize the cost of traveling and say, here's how willing people are to substitute not between Gap and Gap.com, but between any brick-and-mortar clothing store that they can access versus any online clothing option.

So there we find less substitutability, which is kind of what you'd think, because these are different merchants they're accessing. And that's potentially why we find the majority of our gains, something like two-thirds, are coming from accessing different merchants and different types of products online.

What was your study's most unexpected finding?

Klenow: I think I would go back to population density. I'll give you an anecdote. My family went to Kauai not too long ago, and we saw Amazon packages everywhere, which made a lot of sense to us. Kauai is remotely located, has a low population density, and doesn't have many brick-and-mortar stores, so it makes sense that e-commerce would be really important for accessing products in a remote location like that.

We were stunned to find that that wasn't true in the Visa data, which show that the more remote you go, the less likely people are to even have credit and debit cards to transact online. That's more of a mechanical explanation than a deeper explanation, but it's what we see in the data.

What did this study tell you about possibilities for future research?

Klenow: Well, I think the question of how retailers grow is really interesting, and so a bunch of us on this team want to study how e-commerce firms acquire customers. One of the strengths of the Visa dataset is that it's not just the breadth of coverage of online and offline spending, but also being able to observe households, the cardholders.

Researchers who have examined online data, even the ones who see transactions and prices and quantities, can't trace that back to households and where they live and demographics in a very granular way. The Visa data allow you to do that with these anonymized cardholder identifiers.

So I think we can see how important acquiring customers is to growth, which normally we economists can't see very well because we might have household data that are pretty aggregated across consumption categories or retailer data that don't detail sales by different types of customers.

I think this dataset is well-suited to be able to help us understand how retailers grow, how important acquiring customers is for that growth, and how maybe e-commerce has changed those dynamics. People tend to focus on Amazon, this very large player, but at the same time you see many smaller online retailers growing faster.

I think there's a lot of fascination about what makes some firms grow rapidly and become really successful. Walmart took decades to build out to many different locations across the country, but maybe e-commerce is changing that dynamic so that the most efficient retailers can scale up more quickly. And we might have some benefits from that in terms of the growth we see in the United States. ■

Two-Sided Market, R&D, and Payments System Evolution

by Bin Grace Li, James McAndrews, and Zhu Wang

In the credit and debit card industry, new technologies affect multiple parties, including merchants, consumers, and the card networks. The decisions of these different groups interact with each other and create spillover effects called externalities. In his paper, Zhu Wang of the Richmond Fed attempts to explain two puzzles: first, the arguably slow historical pace of payment-card adoption, and second, the perception that merchant fees are too high in light of technological progress that should have reduced costs.

How does your paper contribute to the literature on technology diffusion?

Wang: The payment industry is a particular example of a platform industry that matches two sides of a market. In this kind of industry, the interesting thing about technology diffusion is that new technologies have to be adopted by different groups of agents, their adoption decisions interact with each other, and this process typically features some network externalities. So the adoption process is different from the standard way it has been treated in the literature. We have come up with a new theoretical framework to understand that, and based on that, we can give a realistic description of the industry and we can match theory with data.

You talked about the slow adoption of electronic-payment systems. Could you explain why that has been the case?

Wang: This is a puzzle because most electronic payments we are familiar with today actually were introduced several decades ago. Credit cards were introduced in the 1950s, and debit cards were introduced in the early 1980s. None of these things are really new, but several decades later we are still talking about why everyone is not yet using electronic payments. We still struggle with that because we would have thought that we would be further along because electronic payments are much more efficient than handling cash or checks.

There have been a number of claims by merchants that the card companies have acted in an oligopolistic manner and have used market power to drive up fees. Merchants have complained that it's a contradiction that IT costs are diminishing yet the fees look like they're not declining.

Wang: They are not. So a typical answer that the card networks would provide is: "We increase the fee on the merchant side, but we give more reward on the consumer side. Because this is a two-sided market, we need to balance the demand on both sides so we can boost the use of electronic payments." Without a model, it's hard to say which side is right or wrong. So at the beginning of the paper we pose two puzzles: one is slow adoption, and the other is the complaints of merchants about high fees in spite of technological progress.

In our paper, we address both puzzles with just one hypothesis, one line of reasoning. Our argument is that, given the market power of the card network, the network's profit-maximizing strategy will be to raise the fee on the merchant side and reduce the fee on the consumer side. We attempt to capture this behavior by building a diffusion model about a two-sided market. In this model, the card network charges a significant markup that holds back the diffusion of electronic-payment technology.

The contrasting case we provide is a social planner who does not want to charge a markup but just wants to maximize consumer welfare. This would give a different pricing pattern, also taking into account the two-sided market, and R&D will be different. So the pace of cost reduction will be different, and the diffusion pattern will be different. With a social planner, we will see much higher adoption and much quicker diffusion.

The payment-card network in your model makes investments in R&D. How does this affect its pricing decisions and the overall dynamics of the model?

Wang: The R&D component of the paper is really a new contribution to the literature because in the payment literature so far, people have typically examined static models — they don't usually focus on the dynamics. In the policy debate, one argument against regulation in this market is that price regulation would discourage card networks from doing R&D, so regulation may slow down technological progress, which may hurt consumers and merchants in the long run. But those concerns cannot be addressed using a static model.

To our knowledge, this paper is the first one in the literature to formally model the dynamic decision. And the result is very interesting. Actually, it was not what we expected. Before we wrote the paper, we thought that regulation, especially price regulation, must hold back R&D and thereby hurt efficiency in the longer run. So we built a model that incorporates R&D decisions and a two-sided market adoption externality. Indeed, we found that if regulators impose marginal-cost pricing for card fees, it negatively affects R&D, and we have a dynamically inefficient outcome.

That's true, but in reality, what the regulations typically do, like in the U.S. or in Europe or in Australia, is not really regulate two sides of the market. They typically only cap the merchant fee because merchants complain that the fees are too high. This type

of regulation has been implemented based on some naive argument about a one-sided market. But surprisingly, based on our analysis of a two-sided market, it turns out that a simple cap on merchant fees has a pretty good outcome compared with marginal-cost regulations that ensure the card companies have zero markups. The simple merchant fee cap actually creates a better outcome, which is a result we didn't expect when we started writing the paper.

How does the R&D investment improve the efficiency of card networks? And does R&D make consumers and merchants more willing to participate in a network?

Wang: Yes. In the model, R&D will push down the operational cost for the card network, and with this cost reduction, the card network will reevaluate its pricing. So, according to our paper, you will see the card network raising fees on the merchant side but reducing fees on the consumer side. The overall effect is that adoption by consumers and merchants will increase.

So you invest in R&D, which drives down costs, and that has implications for pricing and adoption by consumers and merchants.

Wang: Yes. So in a one-sided market, this intuition is very clear. If a monopoly sees its costs decline, it will charge a lower price, and the consumer will adopt more.

In your model, what are the roles played by fixed versus variable costs?

Wang: They play very important roles. In fact, the fixed-cost and variable-cost argument is rooted in some of the classic literature on technology diffusion, which argues that larger firms tend to adopt technology earlier than smaller firms. Typically, a new technology will require a significant fixed cost to adopt but will lower the marginal cost of operation. So a larger-sized firm will have a greater incentive to adopt because it pays the same fixed cost to adopt, but it reaps greater benefits from the lower marginal costs due to its higher production volume. In our model, this argument explains well how merchants make their card acceptance decisions.

But in our model of card networks, things are a little bit more complicated because we also talk about consumer adoption. In our model, consumers are heterogeneous due to their different levels of income and spending. When consumers adopt a new electronic payments network, they pay a fixed cost for adoption, but they lower their transaction cost on every payment they make using the network. So if a household has high income and spending, it may find it worthwhile to pay the cost to adopt. In this way, the consumer adoption pattern can be influenced by the income distribution and also can interact with the adoption pattern on the merchant side.

So the payment card system in general increases welfare by offering a way to conduct transactions more efficiently.

Wang: Yes. Our intuition here is that the cost of doing things electronically is much cheaper than handling physical currency and checks. So the resource cost to society of handling things electronically should be cheaper than handling things physically. But consumers and merchants do not face that resource cost directly. What they face is a fee that they each pay the card network. The difference between the fee they pay and the resource cost that society pays equals the markup the card network extracts. And if that markup is too high, it will hurt social welfare to some degree.

So what are the policy implications of your analysis?

Wang: As a benchmark, we examined the R&D decisions that a social planner would make to maximize welfare with full knowledge about how the economy works. We show that this social-welfare benchmark is very different from the model's solution when the card network operates with market power.

We then examine different regulatory approaches and compare their outcomes with the social optimum. One approach is marginal-cost regulation, which may be optimal in a static model but shuts off a network's incentive to do R&D in a dynamic world.

We also examine a different regulatory regime called merchant-fee-cap regulation, which is what most countries have adopted in practice. Based on our analysis, this regulation causes little dynamic inefficiency. In fact, this is a result we didn't expect when we wrote the paper. We thought if you put a regulation in, you will always slow down R&D no matter what. But it's not necessarily the case when you put a cap in only on the merchant side. A merchant fee cap can actually give more incentive for the network to do R&D. The reason is they want to get away from the cap as much as possible to keep some profit.

How do you think your research can be extended?

Wang: The payment industry is complicated because it involves two different distinctive groups of adopters: merchants and cardholders. But we could explore models that are even more complicated than this. In our paper, we only have a single, monopoly platform making dynamic pricing and R&D decisions. But in reality, there are multiple networks. So the competition or cooperation among those networks will make the whole dynamic even more complicated. ■

Innovation, Knowledge Diffusion, and Selection

by Danial Lashkari

Danial Lashkari of Boston College bases his model on Joseph Schumpeter's concept of creative destruction. In Lashkari's model, technology diffusion involves both positive and negative externalities. The positive externalities occur when a company such as Walmart innovates and its new ideas spill over to other firms in the industry. The negative externalities arise due to the loss of potential new ideas and innovations when Walmart's competitors exit the industry. The model analyzes a monopolistically competitive industry in which new entrants and incumbents alike attempt to innovate and gain market share.

Your paper explores Schumpeter's idea of creative destruction. Can you give some examples of how that process has played out in recent years?

Lashkari: I think in the past few decades, we've had many salient examples of the process of creative destruction. The rise of Walmart is a good example of how a firm can come in and build on some of the ideas that are currently used in an industry, in this case retail, and improve on them and provide services to consumers at a cheaper price or higher quality. But the rise of Walmart was obviously accompanied by the exit of lots of smaller retailers all across the U.S. So, on the one hand, you have this creative process of innovation and improvement; but on the other hand, you have the displacement of the very same firms that provided the original ideas in the industry. That's basically the process of creative destruction.

Interestingly enough, if you want to really push this idea, it's interesting to look at what's happened to Walmart with the arrival of Amazon. Amazon, perhaps inspired by the way Walmart was able to use IT, has now made serious inroads against Walmart on e-commerce.

In the transportation industry, the arrival of Uber and the consequences for the taxi industry is another really good example. A dynamic economy like the U.S. really thrives on this core engine of creative destruction, the constant arrival of new ideas that displace all these old ideas and push the frontiers of productivity.

Economists use the word "externality" to describe the effects of a firm's actions on other firms and individuals. Could you describe the distinction in your work between positive externalities caused by technological spillovers versus negative externalities that arise from the displacement?

Lashkari: When a firm like Walmart comes in and innovates, the ideas that it creates are not going to be confined within the boundaries of Walmart. Other firms, in this case Kmart and Target, will try to imitate some of the best ideas. This gives us a positive externality in the sense that when Walmart is investing in experimenting with these new ideas, it's not going to internalize the spillover benefits that its ideas may generate for other firms in the industry.

When Walmart is able to provide services at slightly lower prices than a current firm, obviously customers are going to shift toward Walmart. But Walmart does not internalize the negative effects of its innovation on its competitors. Now, what I'm emphasizing in my paper is a little bit more nuanced in the sense that I'm also focusing on the fact that when another incumbent firm goes out of business due to extra competition, in addition to the services it was providing, that firm was also providing its own ideas. So we're not only losing the business of that firm, but also the potential ideas and innovations the firm could have created.

Basically I am focusing on the positive externalities that Walmart's ideas could create in the industry, and I am comparing those to the negative externalities caused by the potential loss of new ideas and innovations that could have been provided by firms that are now being pushed out of the market.

Could you describe some of the key features of the economic growth literature that your paper builds on?

Lashkari: There has been, I would say, a sort of rebirth of the work on economic growth. We have now started to heavily leverage firm-level data, whereas an earlier wave of work had been more reliant on macro-level data. What we've done as researchers is to look at exactly where the innovative activity happens, which is at the level of firms. And we've realized there is a whole host of firm-level data that allows us to build our understanding of economic growth from the level of firms.

Could you discuss how your model works — the basic setup and some of the key assumptions and results?

Lashkari: The model is based on a monopolistically competitive industry in which firms are uncertain about their relative productivity. This is important because firms have some fixed costs of operation that make them unprofitable if their productivity declines below a certain level. So, the destructive part of creative destruction here works through an endogenous selection process whereby unprofitable firms exit the market due to the competition caused by new ideas. In this uncertain environment, firms invest in improving their productivity.

The model attempts to capture an environment in which startups and incumbent firms alike are trying to innovate and gain market share. You have the innovation efforts of incumbent firms, like Walmart, that are constantly trying to improve their productivity, and there are entrants coming in with newer innovations. In the model, all these firms operate in a stochastic environment in which they face positive and negative productivity shocks that could cause them to gain or lose market share.

The paper mostly focuses on the model's normative implications. I focus on how the model's market equilibrium deviates from the optimal allocation, and I find that the optimal market equilibrium is different from the optimal allocation in ways that are very tightly tied to the heterogeneity of firms. I basically calibrate the model using firm-level data, and I find that, based on a reasonable set of parameters, the largest and most productive firms show greater than optimal innovation activity. It turns out that the environment that produces the maximum rate of economy-wide growth is one where there are a lot of firms with intermediate-level size and productivity.

This result is important if you are interested in subsidizing innovation and R&D, as we do in the United States through policies such as R&D tax subsidies. One implication of the model is that you would want to focus your subsidies on firms that are in the intermediate level of productivity rather than subsidizing the largest and most productive firms.

What is it about larger firms that makes them tend to overinvest relative to the socially optimal level of investment?

Lashkari: There are two different sides to the calculus of externalities of innovation. For larger firms, the positive externality is that the ideas they create are more productive and more useful for all other firms.

The negative externality is that, when they are creating these ideas, these large firms basically move a lot of other firms out of business.

So do they effectively stifle innovation by other firms by eliminating them?

Lashkari: That's basically the idea. In the model, the degree to which large firms stifle the innovation of other firms exceeds the degree to which they are contributing to growth through their own innovations.

Is it also true that smaller-than-average firms are innovating too much?

Lashkari: They are, but it's not as significant for aggregate growth. In the model, the smaller, less-productive firms do not create very valuable innovations, so their positive and negative externalities are not sizeable. Subsidizing them is not going to be very useful.

In the paper, you talked about optimal tax/subsidy policies that might be able to move an economy of the sort described in your model closer to an optimal outcome. Could you talk about those policies and how they relate to current U.S. tax policy?

Lashkari: If you go back a few decades, most government support for R&D innovation was provided through grants. But there has been a shift over the past two or so decades in the U.S. The government has adopted a tax/subsidy scheme in which firms investing in R&D activity receive tax credits. With this type of policy, the government is not trying to pick the winning ideas, which is often the case with directed grants. Instead, generally speaking, the government provides all firms with some constant subsidy rate.

But my paper argues that the government should perhaps condition the subsidy rates on the productivity or market share of the firm. Now, there is an interesting twist to the way tax subsidies actually work in the U.S., which is that there is a cap to the amount of expenditures that firms can use to gain tax benefits. Since that cap is binding for the largest U.S. firms, it so happens that the current U.S. policy is approximating our model's optimal policy.

Should the optimal cap be set differently for different sectors, say, manufacturing versus retail?

Lashkari: In the context of the model, again, we go back to those two forces that we highlighted, the positive and the negative externalities. We would want to think about what are the drivers of those positive and negative externalities. What the calibration suggests is that in retail, as firms are becoming larger, the positive knowledge spillovers they create for other firms decline faster than in manufacturing. And we see that by observing that smaller firms have a much harder time trying to catch up with larger firms in retail compared with manufacturing. In retail, huge numbers of small, young firms die without ever getting even close to making it. In some sense, larger firms are safer from competition in retail than in manufacturing. That suggests to us that the ideas these larger firms are creating are not as easy for the rest of the industry to copy.

How do you plan to extend this research?

Lashkari: I am now working on some projects that try to build upon these ideas using firm-level data but in the context of France. And here the nice thing is the richness of the data that the French government is collecting from firms at the product level. So what we are able to observe here is not only the dynamics of firm sales and employment, but also the dynamics of the prices the firms are charging. And my hope is to build on some of the previous work in the literature to bring in a little bit more richness and a closer link to the data. I'm very much looking forward to what we are going to learn about these questions in the near future. A lot of other people are thinking about these questions, so research on innovation and growth at the firm level is an exciting area these days. ■

FEDERAL RESERVE BANK OF RICHMOND®

Richmond ▪ Baltimore ▪ Charlotte