



2020

Richmond Fed Research Digest

Summaries of work by economists in the Bank's Research Department
published externally from June 1, 2019, through May 31, 2020

The Federal Reserve Bank of Richmond produces several publications that feature the work of economists in its Research Department, but those economists also publish extensively in other venues. *Fed Digest* brings this research together in one place with brief summaries, full citations, and links to the original articles.

Underlining indicates links to bios of Richmond Fed economists.

Risk, the College Premium, and Aggregate Human Capital Investment

By [Kartik Athreya](#) and [Janice Eberly](#)

American Economic Journal: Macroeconomics, forthcoming

In an article forthcoming in the *American Economic Journal: Macroeconomics*, Kartik Athreya of the Richmond Fed and Janice Eberly of Northwestern University study an apparent puzzle. Data over the past several decades show that the earnings premium associated with college education has continued to increase. However, rates of college enrollment have not kept pace. This raises the question: Why has investment in college education stagnated when the monetary incentive to invest apparently has increased?

Athreya and Eberly suggest that the answer hinges on the riskiness of educational investment. In practice, students face a substantial risk of not completing college. (The noncompletion rate in the United States is currently running as high as 50 percent.) And, in the event of noncompletion, costs can be substantial when out-of-pocket expenses are added to the opportunity cost of foregone earnings. Given the uncertainty of completing a college education, it seems reasonable that some individuals may decide to forego enrollment despite the higher college premium.

The authors explore the implications of college noncompletion risk by using a standard model of savings and investment augmented to allow for risky college investment—that is, a college investment that may or may not bear fruit. Parameterizing the model so that it accounts for college enrollment and completion data, the researchers then use the model to examine the responsiveness of college enrollment and attainment rates to changes in the college earnings premium.

They emphasize two main findings. First, college enrollment is not likely to be sensitive to small changes in the college earnings premium. Second, college attainment rates will remain insensitive to even large increases in the college premium. These results are driven by a self-selection mechanism and the fact that college-completion risk varies greatly across the population. Partly due to self-selection, individuals who do not enroll tend to have the highest drop-out risk. And for those individuals, enrollment rates tend to be relatively unresponsive to increases in the college premium. Moreover, even if they are induced to enroll by a higher college premium, their relatively low completion rates imply that their enrollment in college is unlikely to substantially improve the societal college-attainment rate. This pattern—absent changes in college readiness—may portend

a continuing trend of increased earnings inequality. If investment in education is relatively insensitive to the college earnings premium and the supply of college-educated workers is therefore inelastic, the rewards to those who do attain college degrees likely will increase even more.

The authors' findings are consistent with slower future economic growth due to limited human capital accumulation, which inhibits the economy's ability to exploit skill-based technological advancement. They suggest that the significant growth in college attainment from 1979–97 is not likely to recur anytime soon, so the relatively inelastic supply of college-educated labor is likely to be a long-lasting phenomenon. But Athreya and Eberly conclude with some optimism. They note that their *ceteris paribus* analysis is built on the assumption that college readiness is fixed, an assumption that may not hold in practice. Indeed, they argue that their results highlight the need to devote more research to the economy-wide distribution of college preparedness, and they point to a large and growing body of work emphasizing early childhood interventions to underpin the effectiveness of later educational investments.

<https://www.aeaweb.org/articles?id=10.1257/mac.20160396>

Banker Compensation, Relative Performance, and Bank Risk

By *Arantxa Jarque and Edward Simpson Prescott*

Journal of the Japanese and International Economies, June 2020, vol. 56

Controlling bank risk via regulation of compensation arrangements is a new focus of bank regulation. Motivated by the belief that bank compensation practices contributed significantly to the financial crisis of 2007–08, several countries have imposed or proposed new guidelines. For example, the Board of Governors of the Federal Reserve System told U.S. banks in 2010 that their compensation arrangements should provide “incentives that appropriately balance risk and reward.” And in Europe, caps on variable pay relative to base pay have been implemented.

In an article forthcoming in the *Journal of the Japanese and International Economies*, Arantxa Jarque of the Richmond Fed and Edward Simpson Prescott of the Cleveland Fed model a bank as a large number of independent loan officers subject to a common bank-specific shock. They focus on

loan officers, traders, and other low-ranking employees—rather than chief executive officers or other top executives—for two reasons. First, since CEOs cannot directly control the actions of their subordinates, they have to rely on indirect methods, such as delegation of authority, internal controls, and compensation. In the end, a bank's risk profile is mostly determined by the actions of its loan officers and other employees. Second, despite the high level of CEO pay, the vast majority of labor compensation goes to other employees, so compensation regulations have the largest effect on them.

Jarque and Prescott show that the connection between compensation and bank risk is not straightforward and depends on the production technology.

The model highlights three connections between compensation and risk. First, because each loan officer has a miniscule effect on the performance of a large bank, risk is determined by the correlation of loan officers' returns, not the risk of an individual loan officer's projects. Second, compensation contracts rely heavily on relative performance because comparing loan officers' returns can be highly informative about effort. (Both of these implications are absent from the single-agent CEO model.) Third, loan officer effort is under-provided relative to the social optimum due to limited liability and the safety net; however, risk-averse loan officers need to be compensated for extra risk, which adds an additional cost to exploiting the safety net, potentially mitigating the risk-shifting problem. Furthermore, the mapping from bank risk to the optimal contract suggests that the commonly held perception that high bonuses create bank risk is not necessarily true. For example,

when loan officer returns are perfectly uncorrelated, there is no bank risk because the loan officer risk is entirely idiosyncratic and averages out. Consequently, compensation is *irrelevant* for bank risk, though it may matter for bank profits, and it certainly matters for the risk to the loan officers. In contrast, when loan officer returns are perfectly correlated, loan officer effort can be perfectly inferred from bank output, so there is no moral hazard problem, and the officer can be paid a wage. Here the correlation in returns means that there is a lot of risk for the bank, and the authors show that, under reasonable conditions, low wages could create more risk than high wages.

In summation, Jarque and Prescott demonstrate that the connection between compensation and bank risk is not straightforward and depends on the production technology. Evaluating bank risk requires a detailed understanding of the production technology to identify precise effects. Nevertheless, their analysis shows the importance of relative-performance schemes in compensation and implies that finding ways relative performance can increase correlation in returns is a productive strategy for identifying risky compensation practices.

<https://doi.org/10.1016/j.jjie.2020.101077>

Medical Spending, Bequests, and Asset Dynamics around the Time of Death

By [John Bailey Jones](#), [Mariacristina De Nardi](#), [Eric French](#), [Rory McGee](#), and [Rachel Rodgers](#)
National Bureau of Economic Research Working Paper No. 26879, March 2020

Almost 40 percent of U.S. wealth resides in households whose head is age sixty-five or older. In a basic life-cycle model in which the time of death of the head of household is known, the net worth of the household tends to decrease more slowly than the rate predicted by the model. An active research literature has arisen around this retirement-savings puzzle. One potential solution is that older households are holding on to their assets because they fear incurring high medical expenses at the ends of their lives. Another potential explanation is that older households are preserving assets to bequeath to their heirs. Determining which of these motives is more powerful has proven challenging, in large part because the two motives generate similar savings behavior.

One approach to differentiating between these two motives is to analyze dissaving decisions near the end of life. If older households are saving to cover end-of-life medical spending, we should observe at least some households incurring significant medical expenses that decrease their wealth in their final years. On the other hand, if medical expenses do not rise significantly at the end of life and if people are not concerned with leaving assets to their heirs, we should see households consuming their wealth faster as death becomes imminent.

In a National Bureau of Economic Research working paper, John Bailey Jones of the Richmond Fed, Mariacristina De Nardi of the University of Minnesota, Eric French of University College London, Rory McGee of University College London, and Rachel Rodgers of the Richmond Fed take a first step toward that type of analysis by documenting the dynamics of wealth and medical expenses around the time of a person's death. To do so, they employ an event-study approach using data from the oldest cohorts in the Health and Retirement Study. For each "treatment" household experiencing a death, they identify a similar "control" household that does not experience a death until six-to-ten years later.

Using fixed-effects regression, the authors estimate the asset trajectories of the treatment households as death approaches, along with the asset trajectories of their matched control households over the same time period. They find that a household's wealth declines significantly when one of its members dies. Their estimates suggest that in the six years preceding their deaths, the assets of dying singles decline by \$20,000 more than those of otherwise similar survivors. The assets of

couples who lose a spouse fall by \$90,000, a much larger amount. These decreases in net worth are concentrated among wealthier households. Households in the bottom half of the initial wealth distribution experience virtually no death-related declines whatsoever.

Having established that assets fall at the time of death, the next order of business is explaining why. From an accounting perspective, elevated medical spending can account for all of the decline observed for singles and around one-quarter of the decline observed for couples. Bequests from the dying spouse to nonspousal heirs, such as children, are enough to account for the remainder and then some. More empirical work on such bequests, along with similarly timed inter-vivos transfers, is in order. It is also important to examine the potential economic motivations behind nonspousal bequests.

<https://doi.org/10.3386/w26879>

The Wage Penalty for Married Women of Career Interruptions: Evidence from the 1970s and the 1990s

By [John Bailey Jones](#), [Minhee Kim](#), and [Byoung G. Park](#)
Oxford Bulletin of Economics and Statistics, forthcoming

In the latter part of the twentieth century, married women increased their labor force participation substantially. They also greatly reduced the number and duration of their career interruptions, suggesting that career interruptions had become more costly for them. To understand these trends, it is important to examine changes in the career-interruption penalties that married women faced during this period. While many researchers have studied career-interruption penalties generally, few have explored their evolution over time.

In an article forthcoming in the *Oxford Bulletin of Economics and Statistics*, John Bailey Jones of the Richmond Fed, Minhee Kim of the Korea Information Society Development Institute, and Byoung G. Park of the University at Albany (SUNY) use a work-history model and data from the Panel Study of Income Dynamics to compare career-interruption penalties for married women in the 1970s versus those in the 1990s. They find that in the 1970s (data from 1972–79), married women who were out of work in the preceding year earned 40.4 percent less than married women who remained employed. That wage penalty increased to 73.7 percent in the 1990s (data from 1990–97). For married men, comparing the same periods, the career-interruption penalty *decreased*, and the ratio of the wage penalty for married women to that of married men increased from 1.33 in the 1970s to 2.43 in the 1990s.

The results show that changes in occupation- and education-specific wage penalties are the main contributors to the rise in the overall penalty for career interruption.

The authors note that causation is difficult to determine, but during this timeframe, women's levels of educational and professional attainment increased dramatically. Human capital theory suggests and empirical studies have shown that women with college educations and/or jobs

in professional occupations face higher career-interruption wage penalties, which would tend to discourage breaks in employment. There are many potential reasons why women increased their levels of educational and professional attainment from the 1970s to the 1990s: one possibility is that they were responding to changes in their expectations about career interruptions. If married women had come to expect fewer and shorter career interruptions, then wage penalties for employment breaks would have become less of a perceived threat to their lifetime earnings. If this were the case, they would have become more willing to choose careers with higher interruption costs, especially if such careers also generated higher incomes. While this change in selection dynamics might appear as changes in occupation and education, it also might take the form of alternative career choices within the same occupation groups and education groups.

Jones, Kim, and Park consider how much the distribution of married women across education groups and occupation groups changed between the 1970s and the 1990s and how much the wage penalties for each of these groups changed over the same period. They use Blinder-Oaxaca decompositions to allocate the estimated increase in the aggregate wage penalty between the effects of changes in the education and occupation distributions and the effects of changes in the group-specific wage penalties. In the decomposition for occupation, changes in occupation-specific wage penalties account for 83.7 percent to 94.1 percent of the total increase in the wage penalty. Similarly, in the decomposition for education, changes in education-specific wage penalties are the main contributors to the increase in the overall wage penalty. Both results show that these changes account for most of the increase in the wage penalty, as opposed to composition effects, which the authors found to be small.

<https://doi.org/10.1111/obes.12356>

An Estimated Structural Model of Entrepreneurial Behavior

By [John Bailey Jones](#) and [Sangeeta Pratap](#)
American Economic Review, forthcoming

Owner-operators manage four out of five U.S. businesses. These entrepreneurs make decisions about how their businesses borrow, invest, produce, and pay dividends. They also maintain the option to liquidate their businesses and join the labor force as wage workers. What mechanisms explain their behavior? Researchers have approached this question from many angles, alternatively focusing on the roles of financial constraints, the nonpecuniary benefits of business ownership, and cross-firm differences in productivity and risk aversion.

John Bailey Jones of the Richmond Fed and Sangeeta Pratap of Hunter College shed new light on owner-operator behavior in an article forthcoming in the *American Economic Review*. The authors formulate a rich life-cycle model of entrepreneurial behavior and estimate the model using data from the Dairy Farm Business Summary (DFBS), an annual survey of New York dairy farmers conducted by the Dyson School of Applied Economics and Management at Cornell University. The data include detailed records of revenues, expenses, assets, and liabilities. Physical measures, such as farm acreages and herd sizes, are also collected. The data are reviewed by the DFBS staff, who also construct income statements, balance sheets, cash-flow statements, and a variety of productivity statistics and financial ratios.

Using this database, the researchers make two contributions to the literature. First, to the best of their knowledge, they are the first to formulate and estimate a model of entrepreneurial behavior that jointly accounts for financial constraints, nonpecuniary benefits, and risk-taking. Such a joint treatment is essential because none of the driving mechanisms can be estimated accurately in isolation. Second, by using data that focus on one industry in one region, the researchers are better able to identify dynamic mechanisms and minimize problems associated with heterogeneity. For example, within the parameters of the model, they are able to calculate revenue-productivity shocks for each firm, which they decompose into three parts: an aggregate shock due to the price of milk, a permanent farm-specific component, and an idiosyncratic farm-specific component. Such identification is not possible in the cross-sectional datasets typically used in the literature.

One of the authors' main findings is that the effects of financial constraints and nonpecuniary benefits are of first-order importance. Financial constraints appear to exert a strong influence on the intensive margin of operation—that is, on investment and output. Collateral constraints hinder capital accumulation (especially among highly productive firms seeking to expand), while liquidity

constraints force firms to hold cash and divert resources from investment, with similarly negative but smaller effects. Nonpecuniary benefits exert more of an influence on the extensive margin—that is, on the decision to continue operations. The authors find that the least productive farms in the DFBS database have very low financial returns, and many of them likely would exit the industry if it were not for nonpecuniary benefits.

In contrast, Jones and Pratap find that revenue risk appears to play a relatively minor role in decision-making and that eliminating uncertainty in revenue would encourage only a modest expansion in operations, especially compared with the effects of relaxing financial constraints. For example, their estimates suggest that dairy farmers would find only limited benefits from the 2014 Farm Bill's program to insure dairy business operating margins—a result that may explain the program's limited uptake.

<https://www.aeaweb.org/articles?id=10.1257/aer.20170370&&from=f>

Detecting and Analyzing the Effects of Time-Varying Parameters in DSGE Models

By Fabio Canova, Filippo Ferroni, and [Christian Matthes](#)

International Economic Review, February 2020, vol. 61, no. 1, pp. 105–125

In macroeconomics, it is standard to work with models that are structural—that is, models where the parameters characterizing the preferences and constraints of agents and technologies that produce goods and services are invariant to changes in the parameters describing government policies. Many economists have shown that the parameters of dynamic stochastic general equilibrium (DSGE) models are not time-invariant. These parameter variations do not necessarily imply that DSGE models are not structural, but they create concerns about the interpretation of results. To include time variations in DSGE models, researchers have followed the vector autoregression (VAR) literature in which parameter variations are assumed to be exogenous, drifting smoothly as independent random walks or switching between a finite number of states. Many economic

questions, however, hint at the possibility that parameter variations may instead be endogenous. Clearly, analyses conducted under the assumption of time-invariant models or exogenous rather than endogenous forms of time variation may lead to misleading conclusions regarding, for example, the welfare costs of business cycles.

Canova, Ferroni, and Matthes address the misspecification induced by neglected parameter variation and the consequences of assuming incorrect forms of time variation.

In an article in the *International Economic Review*, Fabio Canova of the Norwegian Business School, Filippo Ferroni of the Chicago Fed, and Christian Matthes of the Richmond Fed and Indiana University address the misspecification induced by neglected parameter variation and the consequences of assuming incorrect forms of time variation. They characterize the approximate decision

rules of a DSGE model when parameter variation is present, discuss whether constant-parameter models provide good approximations of the data generating process (DGP), and examine whether time-varying parameter DSGE models generate decision rules comparable to those of time-varying parameter VARs. They provide diagnostics to detect time-varying misspecification and study the consequences of using time-invariant models when the DGP features parameter variations.

When parameter variations are present, a constant-parameter model does not “reasonably” approximate the DGP. When linearized solutions are considered, there are two special cases when the dynamics in response to structural shocks are the same. When second-order solutions are considered, structural responses in time-varying and constant-parameter models are proportional only when parameter variations are exogenous. For higher-order solutions, the structural responses are highly distorted. Constant and time-varying parameter models produce dynamics that are differ-

ent because income and substitution effects are altered. Disturbances to the parameters add to the uncertainty of the environment, making agents prefer to consume more today relative to the future given the same transitory fluctuations in income.

Canova, Ferroni, and Matthes show that linear approximations do not produce time-varying decision rules and that higher-order approximations can do this only if parameter disturbances are interpreted as (reduced-form) decision-rule coefficients. Still, regardless of the order of approximation employed, structural responses will be time-invariant. The diagnostics the authors design are able to detect neglected parameter variations and distinguish exogenous and endogenous forms of time variations. They highlight that certain identification problems noted in the literature may be the result of neglected time variations. Their Monte Carlo study indicates that parameter and impulse response distortions may be large, even for modest parameter variations.

Finally, the authors show that the parameter regulating moral hazard in the Gertler and Karadi (2010) model is likely to be time-varying. When Canova, Ferroni, and Matthes allow variations to be a function of net worth, the fit of the model dramatically improves because there is an additional propagation channel that makes spread and output responses to capital-quality shocks stronger and more persistent.

<https://doi.org/10.1111/iere.12418>

Bubbly Recessions

By *Siddhartha Biswas, Andrew Hanson, and [Toan Phan](#)*

American Economic Journal: Macroeconomics, forthcoming

In recent decades, many countries—including Japan, the United States, and several European nations—have experienced episodes of rapid speculative booms and busts in asset prices followed by declines in economic activity and, in some cases, persistent recessions. More generally, throughout history, the collapse of large asset and credit booms have tended to precede financial crises and recessions. This experience has led policymakers to become increasingly aware of the potential risks of asset-price bubbles, leading to discussions of macroprudential regulation, such as “leaning-against-the-bubble.”

Despite the recent proliferation of economic literature on asset bubbles, relatively little theoretical analysis has been published regarding potential efficiency trade-offs between the booms and busts of bubbly episodes. In most rational bubble models—those commonly used to study the effects of bubbles in general equilibrium—private agents correctly perceive the risk of speculating in a bubbly asset, and the bubbles generally improve financial system efficiency. In an article forthcoming in the *American Economic Journal: Macroeconomics*, Siddhartha Biswas and Andrew Hanson of the University of North Carolina at Chapel Hill and Toan Phan of the Richmond Fed develop a tractable general equilibrium model to address the questions of when and how rational bubbles can lead to inefficiencies and to evaluate the welfare trade-offs. They focus on the combination of financial frictions and downward wage rigidities during bubbly episodes.

The authors model an economy in which entrepreneurial agents with heterogeneous productivity accumulate capital and face financial frictions that constrain their ability to borrow from each other. If borrowing in the credit and capital markets cannot satisfy the demand for savings, speculative bubbles may arise in which agents purchase overvalued assets because they expect to sell them later for even higher prices. These soaring valuations lead to a boom in entrepreneurial net worth, credit, investment, output, wages, and consumption. When the boom finally goes bust, the economy simply reverts to its prebubble state. Therefore, with financial frictions alone, the model implies only that speculative bubbles help to crowd in productive investment and improve the overall efficiency of the economy, as implied in most existing expansionary bubble frameworks.

The theory implies that a policy of leaning against bubbles by taxing speculative investment could reduce net losses from boom-bust trade-offs.

However, the implications change when the researchers add downward wage rigidities to their model. When an expansionary bubble collapses, the net worth of entrepreneurs also collapses, leading to contractions in credit, investment, and labor demand. In a flexible labor market, wages will fall to clear the labor market. However, if real wages are downwardly rigid, there will be rationing in the labor market, resulting in involuntary unemployment. This drop in employment reduces the return on capital investment, which then lowers entrepreneurs' net worth. This effect leads to further contraction in capital investment because entrepreneurs' ability to borrow and invest depends on their net worth. This contraction, in turn, causes further downward pressure on labor demand and wages, thus reducing future capital accumulation, and so on. This vicious cycle continues until the capital stock has fallen sufficiently, often undershooting the pre-bubble steady state.

In short, Biswas, Hanson, and Phan's theory shows how bubbly episodes could trigger deep and persistent recessions. They demonstrate that when a bubble is sufficiently risky and when the wages are sufficiently rigid, society would be better off without the bubble. In this way, their model provides a step toward bridging the views of policymakers and theoretical models of bubbles. In particular, the authors' theory implies that a "leaning-against-the-bubble" policy of taxing speculative investment could reduce net losses from boom-bust trade-offs.

<https://www.aeaweb.org/articles?id=10.1257/mac.20180083&&from=f>

Regressive Welfare Effects of Housing Bubbles

By Andrew Graczyk and [Toan Phan](#)
Macroeconomic Dynamics, forthcoming

Many countries have experienced episodes of bubble-like booms in asset prices. Examples include the real estate booms in Japan in the 1980s, Southeast Asia in the 1990s, and the United States in the 2000s. In general, when the demand for saving/investment outstrips the demand for borrowing in the credit market, investment returns are depressed and real estate investment can serve as a prominent alternative for savers and investors. Thus, a low interest rate environment, as seen in the 2000s, provides fertile ground for the emergence of asset bubbles, including housing bubbles. Given the prevalence of such episodes in recent decades, a central question has arisen for economists and policymakers: What are the welfare effects of asset bubbles?

In an article forthcoming in *Macroeconomic Dynamics*, Toan Phan of the Richmond Fed and Andrew Graczyk of Wake Forest University highlight the nuanced welfare effects of asset bubbles, especially those in housing or some other fundamentally useful asset. The authors develop a simple overlapping generations model of bubbles with income inequality and financial frictions. To establish income inequality, agents are born with endowments that are either high or low. The low-endowment agents borrow to purchase the amount of housing that maximizes their utility. In contrast, the high-endowment agents purchase enough housing to maximize their utility plus additional housing to serve as saving or investment vehicles. In the presence of financial frictions, such as imperfect contract enforcement, the low-endowment borrowers face a binding credit constraint, modeled as an exogenous limit on their debt capacity. In equilibrium, that constraint also effectively limits how much the high-endowment savers can invest in the credit market. This shortage of saving/investment opportunities in the credit market can lead to an equilibrium interest rate that is below the economy's growth rate. This low interest rate environment, in turn, facilitates the emergence of asset bubbles.

In a housing-bubble equilibrium, "speculative" investment demand for housing among savers is similar to demand for a bubble asset in a standard rational bubbles model. In other words,

an agent (the saver) purchases extra homes and/or larger homes because he or she expects to sell them to someone else for higher prices in the future. The authors show that as long as the housing bubble persists, it produces opposite welfare effects on the high-endowment savers and the low-endowment borrowers. The persistent bubble increases investment returns for the high-income savers (relative to the prebubble benchmark). In contrast, the persistent bubble reduces the welfare of the low-endowment borrowers because it raises the price of housing as the speculative demand of the savers crowds out the utility demand of the borrowers, who derive a relatively higher marginal utility from housing in equilibrium. This scenario also implies a feedback loop in which high income inequality depresses the interest rate, which facilitates the existence of housing bubbles, which produce regressive welfare effects, which create greater income inequality, and so on.

The regressive welfare effects are reduced if the model considers only pure bubbles—that is, bubbles based on assets with little or no fundamental value, such as the tulip bulbs that became exorbitantly overpriced in seventeenth century Holland. Pure bubbles provide an investment vehicle for savers, but unlike the housing-bubble equilibrium, credit-constrained borrowers have little or no demand for a pure-bubble asset. And, since investing in the pure-bubble asset provides a substitute for investing in housing, the crowding-out effect that was prevalent in the housing bubble is absent in the pure-bubble equilibrium.

<https://doi.org/10.1017/S1365100519000981>

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

By Bin Grace Li, James McAndrews, and [Zhu Wang](#)
Journal of Monetary Economics, forthcoming

The payments system is vital to the American economy. In 2011, U.S. consumers made an estimated 158 billion purchases worth about \$8.3 trillion, not counting purchases made by businesses or governments.

In the past several decades, the payments system has been migrating steadily from paper (currency and checks) to electronic transfers (including credit cards and debit cards), but many economists had expected this migration to occur more rapidly. Another puzzle is that card fees paid by merchants have gone up, while card fees paid by consumers have gone down. Merchants claim that card networks (Visa and MasterCard, for example) and the cards' issuing banks have used their market power to drive up the fees that merchants pay. These higher fees stand in sharp contrast to technology-driven cost reductions in the payment-card industry, a contrast that has produced new government regulation and a spectacular amount of antitrust litigation.

In an article forthcoming in the *Journal of Monetary Economics*, Bin Grace Li of the International Monetary Fund, James McAndrews of TNB USA Inc., and Zhu Wang of the Richmond Fed address these puzzles with a new theoretical framework for payments system evolution and the accompanying issues of efficiency and competition. They base their analysis on a two-sided market where merchants and consumers make adoption and usage decisions for payment cards. The economics of these choices include the fixed costs of adoption and the marginal benefits of usage intertwined with the heterogeneity of consumer income and merchant size. The authors embed their theory in a fully dynamic model in which a monopoly card-payment network sets usage fees and conducts research and development to lower costs. They calibrate the model with U.S. payment-card pricing, adoption, and usage data to conduct welfare and policy analyses.

Their findings suggest that the market power of payment-card networks (a monopoly in the model) plays an important role in explaining the slower-than-expected adoption and the asymmetric pricing of payment cards. In contrast, a Ramsey social planner who aims to maximize social welfare would set lower usage fees and conduct more R&D, thus achieving both higher levels of adop-

tion by merchants and higher levels of usage by consumers. There are reasons why the monopoly network and the Ramsey social planner would behave differently. For the monopoly network, increasing the merchant fee leads to higher retail prices, which allow the network to extract more profit. Moreover, because the network does not gain anything from cash usage, increasing the merchant fee reduces cross subsidies from card users to cash users among large merchants who serve both card and cash customers. In contrast, the Ramsey social planner cares about consumers'

The authors' findings suggest that the market power of payment-card networks plays an important role in explaining slower-than-expected adoption and asymmetric pricing.

real total purchases rather than their nominal card spending. The social planner also cares about the welfare of both card users and cash users. In terms of R&D decisions, the monopoly focuses on increasing its profit, which is only a subset of the social welfare that the social planner would value, so the monopoly invests less in R&D than the social planner does.

Li, McAndrews, and Wang also consider several regulatory interventions. They find that while regulating card fees based on marginal costs could maximize consumer welfare in a static setting, it would shut off the card network's R&D, which might reduce social welfare in a dynamic setting. In comparison, they find that capping merchant fees may improve consumer welfare without causing much dynamic inefficiency.

<https://doi.org/10.1016/j.jmoneco.2019.06.005>

Investigating Nonneutrality in a State-Dependent Pricing Model with Firm-Level Productivity Shocks

By Michael Dotsey and [Alexander L. Wolman](#)

International Economic Review, February 2020, vol. 61, no 1, pp. 159–188

Within the confines of a standard DSGE framework, Mike Dotsey of the Philadelphia Fed and Alexander L. Wolman of the Richmond Fed construct a state-dependent pricing model with idiosyncratic productivity variation among firms and small fixed costs of price adjustment. Two parameterizations of the resulting model can match several facts concerning firms' pricing behavior.

Unlike other models in the literature, their model matches the following six facts—drawn mainly from Klenow and Malin (2011): (1) there are both large positive and large negative price changes that are intermixed with many small price changes; (2) average price changes are an order of magnitude larger than needed to keep up with inflation; (3) many prices are set infrequently, with changes occurring less than once a year, whereas some prices appear to be completely flexible; (4) aggregate hazards are relatively flat; (5) the size of a price change is not related to the time since the last price change; and (6) price changes are positively autocorrelated.

The authors calibrate their model to match the distribution of price changes as well as the average duration of prices documented in Klenow and Kryvtsov (2008) and the positive autocorrelation of price changes discussed in Klenow and Willis (2016). Then the authors investigate the model's implications for responses to monetary shocks and find that two parameterizations can match the six facts while generating significant and persistent nonneutrality. The two parameterizations also can account for the moderate degree of price stickiness, measured by average price duration. Further, the dispersion in productivity across firms needed by the model to match the size and dispersion of price changes in the data does not seem overly large judging by the plant-productivity literature. In addition, the aggregate hazard functions generated by the benchmark models are rather flat, which is consistent with the data. This result occurs despite the fact that conditional on productivity, all hazards are upward sloping, a feature that appears to be consistent with micro hazard data from Japan and the United States. Dotsey and Wolman are able to trace out the way that aggregation works in their model, showing that flat hazards are a feature of the dispersion

and persistence of idiosyncratic productivity shocks. Additionally, their model is consistent with the autocorrelation of price changes and with the lack of correlation between the size of a price change and the time elapsed since the last price change.

Despite a relatively high degree of steady-state price flexibility in their model, two parameterizations produce significant nonneutrality of monetary disturbances (as stated above), but the magnitude and persistence of nonneutrality differ. These results reinforce the suggestion in Caballero and Engel (2007) that steady-state stickiness and the degree of nonneutrality are not as closely related in state-dependent models as they are in time-dependent models.

Finally, Dotsey and Wolman conclude that the steady-state sufficient statistic developed by Alvarez et al. (2016)—which determines the degree of nonneutrality in many models—does not predict the results produced by their model. Therefore, the extent of nonneutrality in heterogeneous firm models depends on a host of interconnected features for which a sufficient statistic remains to be found. Dotsey and Wolman’s model, however, can be solved by perturbation methods, holding out the hope that in the not-too-distant future, state-dependent pricing models can be estimated by Bayesian means, making them useful tools for analyzing the effects of monetary policy.

<https://doi.org/10.1111/iere.12420>

Rethinking Detroit

By [Raymond Owens III](#), [Esteban Rossi-Hansberg](#), and [Pierre-Daniel Sarte](#)

American Economic Journal: Economic Policy, May 2020, vol. 12, no. 2, pp. 285–305

<http://doi.org/10.1257/pol.20180651>

Editor’s Note: This article is substantially the same as the National Bureau of Economic Research working paper of the same title that was summarized in the *2017 Richmond Fed Research Digest*.

A Heterogeneous-Agent New-Monetarist Model with an Application to Unemployment

By [Guillaume Rocheteau](#), [Pierre-Olivier Weill](#), and [Tsz-Nga Wong](#)

Journal of Monetary Economics, forthcoming

<https://doi.org/10.1016/j.jmoneco.2019.11.006>

Editor’s Note: This article is substantially the same as the National Bureau of Economic Research working paper of the same title that was summarized in the *2019 Richmond Fed Research Digest*.

Diverging Trends in National and Local Concentration

By [Esteban Rossi-Hansberg](#), [Pierre-Daniel Sarte](#), and [Nicholas Trachter](#)

In the 2020 NBER Macroeconomics Annual, edited by Martin Eichenbaum and Erik Hurst. Chicago: University of Chicago Press.

<https://www.nber.org/chapters/c14475.pdf>

Editor’s Note: This article is substantially the same as the National Bureau of Economic Research working paper of the same title that was summarized in the *2019 Richmond Fed Research Digest*.

This issue of *Richmond Fed Research Digest* was compiled and edited by Karl Rhodes and John Mullin. It was published by the Research Department at the Federal Reserve Bank of Richmond and may be photocopied or reprinted in its entirety. Each individual summary also may be photocopied or reprinted in its entirety. Please credit the Federal Reserve Bank of Richmond and include the statement below.

Richmond Fed Research Digest summarizes articles and other works that reflect the views of their respective authors and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

