Rose, Nancy L.

Article

Industrial organization

NBER Reporter

Provided in Cooperation with:
National Bureau of Economic Research (NBER), Cambridge, Mass.

Suggested Citation: Rose, Nancy L. (2012) : Industrial organization, NBER Reporter, National Bureau of Economic Research (NBER), Cambridge, MA, Iss. 4, pp. 1-6

This Version is available at:
https://hdl.handle.net/10419/103232

NBER Reporter

NATIONAL BUREAU OF ECONOMIC RESEARCH

Reporter OnLine at: www.nber.org/reporter

2012 Number 4

IN THIS ISSUE

Program Report

Industrial Organization

Nancy L. Rose*

The NBER's Program on Industrial Organization (IO) begins its third decade with a core of 60 program members, including 15 whose primary affiliations are in another NBER Program. The Program's meetings attract submissions from a large and diverse set of researchers, and are lively sessions with 75 to 90 scholars typically in attendance. The IO Program produces important applied research on a broad range of industries and topics, increasingly at the intersection with such other NBER Programs as Environmental and Energy Economics, Productivity, and Health Care. That commonality is recognized with frequent joint program meetings and contiguous Summer Institute sessions with other NBER groups. In 2012, IO Program members Aviv Nevo and Ariel Pakes delivered the annual Summer Institute Methods Lectures, focusing on the econometrics of demand estimation and related methodologies.

This report describes work in just three of the Program's areas: modeling consumer choice; the industrial organization of the digital economy; and lessons for designing government auctions. Readers interested in exploring the broader range of NBER work in IO are encouraged to visit https://www.nber.org/papersbyprog/IO.html

Consumer Choice

Empirical economists in the field of IO have devoted substantial attention to modeling the determinants of demand across a variety of settings. For some time, NBER researchers have been active in the design, innovation, and evaluation of methods to estimate demand based on neoclassical theories of consumer utility maximization. Nevo and Pakes discussed this in their 2012 Methods Lectures[1] and dozens of NBER Working Papers have been published in this area.[2] In recent years, empir-

**Rose directs the NBER's Program on Industrial Organization and is a Professor of Economics at MIT.* *The numbers in parentheses throughout this report refer to NBER Working Papers.*

NBER Reporter

ical researchers increasingly have turned their attention to analyzing the underpinnings of individual choice, for example characterizing the implications of deviations from standard neoclassical models of optimization behavior and the role of information in markets.

Consumer Behavior

The detailed microdata that are the mainstay of much empirical IO research have proved useful for identifying departures from conventional models of consumer utility maximization. A body of work in this area has looked at automobile purchases, one of the most significant consumer purchase decisions for most households. Meghan Busse, Florian Zettlemeyer, and co-author Duncan Simester (13140) document consumer responses to "price cues" in the context of a Big Three automaker "Employee Discount Pricing" promotion in the summer of 2005. They find that consumers responded to this promotion with unprecedented increases in new car purchases, even though prices during the promotion were not substantially lower than immediately prior to it. Indeed, sales increased even for some models with higher prices during the promotion. While the researchers point out that this behavior can be consistent with rational reliance on (noisy) price signals, their results are cautionary for those who would model consumers as responding primarily to observed prices. In another paper on auto purchases, Nicola Lacetera, Devin Pope, and Justin Sydnor (17030) look at heuristic information processing in used car purchases. They find that sale prices drop discontinuously at exactly 10,000 mile odometer readings, consistent with customers focusing on the leftmost digit of the odometer reading rather than incorporating the full odometer reading into their valuation. They estimate $2.4 billion of mispricing as a result. Busse and Pope and their co-authors (18212) use a sample of 40 million vehicle purchases and 4 million house purchases to explore the role of projection bias — the tendency to over-predict the degree to which one's future tastes will resemble one's current tastes — in purchasing behavior. They find that weather at the time of purchase overly influences purchase decisions for these major durables. They meticulously explore alternative explanations for this finding, and their results

rule out explanations grounded in neoclassical utility maximization. For example, spring or fall days that are unusually warm and sunny induce additional convertible sales, which are not merely time-shifted. Moreover, the convertibles purchased on such days are more likely to be traded in quickly, consistent with misestimating future tastes.

Justine Hastings and Jesse Shapiro (18248) analyze "mental accounting" in household purchases of gasoline. Their results consistently reject the null hypothesis that households treat spending on gasoline as fungible with other income. Instead, when gasoline prices rise, consumers disproportionately substitute to (less expensive) lower octane gasoline, far more than the substitution that occurs for similar income effects from non-gasoline price sources; the converse is true when gas prices rise. This complements work that Hastings has done with other collaborators (13614) on how households adjust grocery purchases when gasoline prices change. Andrei Shleifer and his collaborators (17947) develop a model of context-dependent consumer choice focused on "salient attributes" that is consistent with this mental accounting behavior, and use their model to study discounts in a variety of settings.

Better understanding of consumer choice is an important input to modeling firm decisions. Julio Rotemberg (13754) models the implications for firms and policymakers of consumers who do not make effective use of price information, and then suffer ex post regret or anger as a consequence. Hastings and Shapiro (18248) argue that supplier response to the consumer octane adjustment behavior they find in gasoline markets may contribute to an observed inverse relationship between gasoline prices and retailer markups over time.

The Role of Imperfect Information

The rich theoretical literature on markets with imperfect information recently has been married to an increasing body of empirical work exploring how imperfect information affects markets. In one paper, Ginger Jin and David Dranove (15644) review the theoretical and empirical research on product quality disclosure and certification. Jin and her collaborators (14252) also explore how information about the properties of a new pharmaceutical is diffused to doctors and patients, and they consider a range of information sources including academic articles, advertising, media reports, FDA updates, and individual patient experiences. How information is presented, in addition to its content, can have substantial impacts on consumer responses: Hastings and various collaborators have shown this in the context of workers' choice among pension investment options in Mexico's privatized social security system (14538) and in parents' decisions on public school choice programs (12995). Similarly, Phillip Leslie and Alan Sorensen's work with Bryan Bollinger (15648) on posting calorie counts in restaurant chains demonstrates the importance of how information is presented to consumers.

In many markets, search does not appear to be effective in matching consumers to the lowest-price or highest match quality product. Glenn Ellison's research with Alexander Wolitzky (15237) argues that this may in part reflect actions taken by firms to impede effective search. In their model of "obfuscation," firms selling homogeneous goods find it individually rational to invest in actions that make it more difficult for consumers to learn about their product and full product price, because this reduces competition across firms and sustains higher mark-ups. Bruce Carlin and coauthor Florian Ederer (17895) model oligopolists' product proliferation responses to the possibility of consumer search fatigue, the notion that search is not only costly but also tiring, potentially leading consumers to break from searching in some periods. Robert Hall and Susan Woodward (16007) argue that mortgage broker decisions by borrowers suggest substantial deviation from optimal search behavior, and are indicative of buyer confusion, not only on how to assess complex menus of broker charges but even about the potential benefits of search among brokers. They conclude that current disclosure policies have done little to mitigate that confusion.

Industrial Organization of the Digital Economy

The digital economy has exploded in the two decades since the IO Program's January 2001 conference on e-commerce,[3] along with economic research on its characteristics and the implications for firm strategies and traditional retail markets. Jonathan Levin (16852) examines the literature in this area and describes the economic implications of key features of the digital economy: an unusual combination of substantial economies of scale with customer personalization; the ability to collect large volumes of detailed data about customers, their behavior, and preferences; and the rapid pace of innovation facilitated by seller experimentation. He notes the critical role of economic theory in the design and analysis of these markets, and the platform that these markets offer for empirical research on the digital marketplace and as a setting in which to test models of imperfect competition. Below are results from just three strands of IO research that explore these and other themes in online markets: the design of online strategies by firms; mechanisms to address asymmetric information about online seller quality; and the implications of digital distribution for producers and "bricks and mortar" retailers.[4]

Designing Online Strategies

As Levin notes, digital marketplaces offer new challenges and new possibilities for firms. David Reiley and his collaborators analyze online auctions in their chapter for the *Handbook of Economics and Information Systems* (12785), focusing on the theory, experimental research, and empirical analysis of online retail auctions such as eBay.[5]

This work describes the responsiveness of bidder strategies to seller strategies, and its implication for optimal design of online auction markets. It also addresses endogenous innovation in markets such as eBay, highlighting the importance of considering dynamic implications of auction design for the viability of platforms.

One of the most active online markets involves "position auctions" which are conducted by search engines such as Google, Bing, or Yahoo to allocate to advertisers the "sponsored link" positions on a search response page. Susan Athey and Ellison (15253) emphasize the two-sided market aspect of these auctions. Bidders (advertisers) care about how consumers respond to advertising links, and those responses in turn are affected by the mechanism that sellers (search engines) use to allocate advertisers to positions. Enriching the analysis to include consumer search behavior yields a number of insights not present in conventional auction models, such as the benefits of high reserve prices to exclude lower match quality ads and the informational inefficiencies that can be induced by weighting bids by customer click-through rates.

In online markets, experimentation is facilitated and rewarded. Search engine firms rely on substantial experimentation, in addition to economy theory, to enhance profits through the design of their auctions. Liran Einav, Levin, and their collaborators (17385) document the activity of eBay sellers to improve their strategies through both active and passive experimentation. The ease of experimentation online is a boon to researchers as well. Reiley et al. (12785) and Levin (16852) describe a number of academic studies that have taken advantage of online platforms to construct field experiments to investigate consumer behavior, pricing strategies, advertising effectiveness, and the implications of auction design, some of which are described below.

Asymmetric Information on the Internet

The growth in online markets has elevated interest in the effect of asymmetric information on seller quality, and has provided new tools for its empirical investigation. While online markets may reduce search costs and offer greater apparent pricing transparency, their heightened anonymity of exchange exacerbates the problem of asymmetric information between buyers and sellers, particularly with respect to seller quality or trustworthiness. Seller reputation can mitigate asymmetric information, and often is established online through buyer feedback mechanisms, such as eBay's well-studied feedback system. Third-party certification provides an alternative to feedback or reputation mechanisms. Jin and her co-authors (17955) study the effectiveness of certification authorities for online pharmacies, used by many consumers to reduce drug acquisition costs. For four of the five popular brand-name drugs they ordered from online pharmacies, drugs labeled as branded were authentic versions for all delivering pharmacies, whether certified or not, but prices at certified U.S. pharmacies were roughly 50 percent higher than were prices at non-U.S. certified pharmacies. This suggests considerable cost to consumers from complying with FDA warnings to avoid all foreign websites, perhaps without concomitant consumer benefit. For the fifth drug—Viagra—certified pharmacy prices and quality were roughly identical regardless of country, while uncertified pharmacies offered both a lower price and a lower probability of receiving an authentic branded product. This suggests a potential advantage to buying only from certified websites; the authors' online survey of over 2500 consumers suggests that is what more than 40 percent of consumers who purchase drugs online do. Jin and collaborators study the role of price signals and regulation across international pharmaceutical markets (16854 and 18073).

Erzo Luttmer, Asim Khwaja, Rajkamal Iyer, and Kelly Shue (15242), as well as Jin and co-author Seth Freedman (16855), use data from the online peer-to-peer lending platform Prosper.com to investigate the ability of lenders to screen the creditworthiness of prospective borrowers. Iyer et al. find that lenders respond to coarse information in the Prosper.com profiles in order to infer much of the information that would have been accessible from (unreported) individual-level credit scores. Freedman and Jin report that lenders on Prosper.com generally underestimate the credit risk of borrowers, but learn significantly from their own experiences on the site. Newer cohorts of lenders underestimate less, suggesting some diffusion of learning across cohorts. Of particular interest is the convergence the researchers note between online and more traditional offline sources of credit: as lower quality subprime borrowers have been increasingly excluded from funding on Prosper.com, the site has competed more directly with conventional lending institutions such as banks. Competition between online and offline outlets is also the subject of other work by Freedman and Jin on peer-to-peer lending, and is the topic of a broad research agenda by other NBER researchers, to which I turn next.

Interactions with Offline Markets

The rapid growth of the internet-based economy over the past 15 years has dramatic implications for both producers and "bricks and mortar" retailers. Early research in this area focused on pricing impacts of online search and e-commerce. More recent research has highlighted the impact of the internet on the allocation of sales across retailers, and the entry/exit decisions of firms, and product choice decisions by producers.

Einav, Levin, and their co-authors (18018) explore the impact of sales taxes on consumers' choices of online retailers, which is of considerable policy interest. Their analysis of eBay customer responses to sales taxes suggests considerable sensitivity: a single percentage point increase in a customer's home state sales tax implies an increase of nearly 2 percent in online purchasing from other states, and a decline of roughly 3 to 4 percent in online purchasing from home state sellers. The authors also note

increased density of sellers on the low-tax side of state borders.

Ali Hortaçsu and Chad Syverson and their collaborators (14166) [6] examine impacts on physical retailers for three of the sectors they expect to be most affected by the internet: auto dealers, bookstores, and travel agencies. They find that online shopping has shifted the distribution of revenues among physical bookstore and auto dealers from small retailers to larger retailers, and that smaller retailers disproportionately have exited as the fraction of consumers using online shopping increased. Travel agencies experienced the same reallocation away from small outlets, but for that industry the trend appeared to be national, a function of changes in airlines' distribution systems and not dependent on local consumer online shopping patterns.

Igal Hendel, Nevo, and co-author Francois Ortalo-Magne (13360) compare the impact on home sellers of using conventional versus online sales outlets in a study of the 2004 housing market in Madison, Wisconsin. They find no sales price difference across houses sold through traditional realtors using the Multiple Listing Service (MLS) and those sold by owners using the online FSBO.com website. However, houses on the MLS are both more likely to sell and are quicker to sell, conditional on a transaction, which is consistent with improved matching on the still-larger MLS network.

The internet facilitates search not only on price but also on provider and product quality. And, online review systems allow consumers to register feedback on physical outlets. Jin and her collaborators (18567) analyze restaurant ratings on Yelp.com, focusing on the optimal way for a review site such as Yelp to construct aggregate ratings from individual feedback. The usefulness of user feedback depends on its credibility, though, and assessing credibility can be difficult. Judith Chevalier, Dina Mayzlin, and Yaniv Dover (18340) investigate the incidence of review manipulation by comparing the distribution of hotel reviews on Tripadvisor.com, which allows anyone to post a review, to those on Expedia.com, which restrict reviews to consumers who have made a booking at that hotel through Expedia. They find that on TripAdvisor.com, small independent hotels have more five-star reviews, and their neighboring hotels have more one- and two-star reviews, consistent with their predictions of ex ante incentives for review manipulation. While these results suggest that review manipulation may be economically significant, the authors note that the overall level of manipulation is relatively low, ensuring that the platform still communicates useful content.

The music and video industries have been among those argued to be most affected by the internet, in large part because of the producers' greater difficulty in enforcing intellectual property protection of their content online. In principle, unauthorized distribution of online content may have both demand contraction effects (by substituting for purchased content) and demand expansion effects (by increasing potential consumer awareness of the product, effectively advertising). Joel Waldfogel (13497) explores these twin effects on television viewing in a study of unauthorized (primarily YouTube) and authorized (primarily network) web distribution of television shows. Using a survey of university students, he finds that internet access induces a modest substitution away from traditional television which is more than offset by a strong demand expansion effect: overall time on network-controlled sites (television and network websites) increases by 1.5 hours/week. Julie Mortimer, Sorensen, and co-author Chris Nosko (16507) find that musical artists have reacted to the decline in album sales that is associated with unauthorized file-sharing by increasing their live performances. Less well-known or popular musicians among the more than 1800 artists they study experience significant increases in concert revenue in the post-Napster era, in part offsetting the lost album revenues, and perhaps reflecting greater awareness of their music by potential fans. Waldfogel (16882) assembles a novel dataset to explore whether reduced album revenues have led to reductions in the production of new albums, and he concludes that there is no discernible decline in quantity or quality post-Napster. That is consistent with lower costs of bringing new works to market and growth of independent labels. Finally, Leslie and Sorensen find that the expansion of ticket resale markets for major rock concerts, facilitated by online resale sites, improves the allocation of tickets to high valuation buyers. However, half of the gains are dissipated through higher transactions costs, so resale buyers end up with little of the potential surplus (15476).

Designing Government Auctions

Governments and quasi-public agencies use auctions in a wide variety of settings, including: competitive procurement; sales or leases of publicly-owned assets, such as mineral and timber rights on public land and spectrum allocation; wholesale electricity purchases and sales; and the allocation of pollution permits under some cap and trade programs. Economic theory has made fundamental contributions to the design of many of these auctions, and empirical research has contributed to evaluation of their operation and guiding improvements in their execution. NBER research has played a role in both fronts, and the NBER Working Group on Market Design, led by Athey and Parag Pathak, focuses on these and related issues.

Since 1994, Federal Communication Commission spectrum auctions have been used to allocate billions of dollars in spectrum rights. Patrick Bajari and his co-author Jungwon Yeo (14441) describe how FCC auction design has evolved over time to mitigate concerns about tacit collusion by bidders. The researchers examine patterns in the bidding data from four large spectrum auctions and conclude that later auctions do, in fact, exhibit fewer examples of strategies most likely associated with potential collusion. Their analysis gives a flavor of the considerable complexity that is involved in bidding in spec-

trum auctions. Jeremy Bulow, Levin, and co-author Paul Milgrom (14765) describe the potential for economic and game-theoretic modeling to help bidders devise successful strategies in the face of that complexity. They illustrate such potential by demonstrating how a new entrant used it in the 2006 90MHz auction, which contributed to the firm's success in purchasing nationwide spectrum coverage at one-third the price paid by incumbents, thus saving more than a billion dollars.

Much of the empirical work in government auctions done by NBER researchers has focused on U.S. Forest Service timber auctions, which can generate more than one billion dollars annually. James Roberts and Andrew Sweeting (17624) focus on when sellers should use auctions, comparing the expected relative performance of a simultaneous bid auction to a setting in which sellers invite buyers to make offers sequentially. Athey and Levin work with various collaborators to analyze the design and operation of timber auctions. In one paper (14590) they compare performance under two different auction formats: sealed bid auctions, which attract more small bidders, and open outcry auctions. Their calibrated model suggests that sealed bid auctions generate greater expected revenue for the Forest Service, and it focuses attention on bidder competitiveness as a critical choice in auction format. In a more recent paper (16851) they turn to the set-asides and subsidies that the government frequently uses for preferenced bidders, most often small or minority-owned businesses, in procurement or natural resource auctions. Their analysis shows that restricting entry to small businesses is associated with significant revenue and efficiency costs; replacing the restriction with a bidder subsidy would increase revenue, efficiency, and the profit of small bidders, with minimal impact on large firm profitability.

Highway construction procurement contracts are a significant state level activity, imposing substantial direct costs to finance road construction and repair, and substantial indirect costs on drivers who are subject to delays and longer commutes while construction projects are underway. Bajari and Greg Lewis have developed a research agenda that investigates how to design procurement contracts to more effectively align the incentives of contractors with those of the highway department and drivers. In one paper (14855), they evaluate scoring auctions used by the California Department of Transportation to provide explicit time-to-completion incentives in contract awards. They estimate substantial welfare gains from the incentive contracts, although direct outlays by the Department of Transportation also increase through their effect on the winning bid. Their model suggests even larger potential gains from an optimally designed policy. In a more recent paper (17647), they develop a model of contractor adaptation to productivity shocks, incorporating time incentives in an optimal contract design. They combine this with day-level information on work plans, progress, and delays for Minnesota highway projects to explore empirically the role of adaptation and delay, and illustrate the impact of alternative incentive structures on outcomes.

---

[1] *Video and slides can be viewed at www.nber.org/econometrics_minicourse_2012/*

[2] *See for example, recent contributions by program members Patrick Bajari, Steven Berry, Jean-Pierre Dube, Jeremy Fox, Philip Haile, Christopher Knittel, Julie Mortimer, Aviv Nevo, and Stephen Ryan, among others.*

[3] *Published as a special issue of the* Journal of Industrial Economics, *Vol. 49, no. 4, December 2001.*

[4] *Many IO program members are active in the Productivity Program's new Economics of Digitization and Copyright Initiative, which brings together researchers from a diverse set of fields to study this sector of the economy. This report focuses on work in this area by members of the IO program; additional working papers can be found on the NBER's website.*

[5] *In* Handbook of Economics and Information Systems*, Terrence Hendershott, ed., Elsevier Science, 2006.*

[6] *Published as Maris Goldmanis, Ali Hortaçsu, Chad Syverson, and Önsel Emre, "E-Commerce and the Market Structure of Retail Industries," NBER Working Paper No. 14166, July 2008, and* Economic Journal, Royal Economic Society, *vol. 120(545), (2010), pp. 651–82.*

Cross-Border Capital Flows, Fluctuations and Growth

Sebnem Kalemli-Ozcan*

What is the extent of international financial integration, and how does such integration affect economic fluctuations and growth? Does the effect differ during tranquil times versus times of financial crisis? Does financial integration transmit shocks across the globe and lead to contagion? In recent research, together with my co-authors, I search for answers to these and other related questions, using both macro-level country data and micro-level firm data.

Capital Flows: Where and Why?

One common definition of international financial integration is the amount of cross-border capital flows. These flows can take the form of foreign direct, portfolio equity, and debt investment, constituting the financial account — the mirror image of current account in the balance-of-payments statistics. Figure 1 plots the average current account balance with reverse sign as a measure of total net capital flows from more than 150 countries, together with different types of flows.[1]

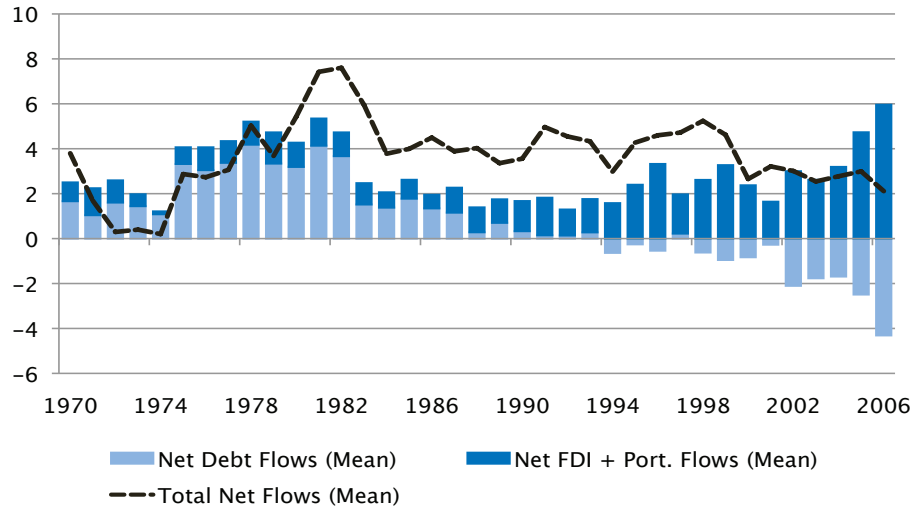
The black dashed line shows that the world is running a current account deficit, roughly around 4 percent of GDP, implying positive net capital flows on average since the 1970s.[2] Since the 1990s, however, countries seem to be net borrowers in FDI and equity investment and net lenders in debt instruments.[3] This simple plot hints that current account may not be informative in terms of testing the predictions of certain classes of models for the amount and direction of capital flows and their implications for economic fluctuations and growth. The appropriate definition (FDI versus debt, public versus private, or net versus gross flows) must be used depending on the question asked.

For example, the neoclassical model predicts a large amount of capital flows based on return differentials from capital-abundant rich countries to capital-scarce poor ones. The lack of such flows in the data is known as the Lucas paradox. The recent period of global imbalances has seen a related paradox, where capital flows in the reverse direction (when measured from current account), from "still poor but growing fast" countries such as China to "rich but not growing" countries such as the United States. Laura Alfaro, Vadym Volosovych, and I have investigated the reasons for both of these phenomena.[4]

Our results show that in a sample of developed and developing countries, the positive correlation between capital flows and GDP per capita (that is, the Lucas paradox) during 1970–2000 goes away once we account for the effect of institutional quality: rich countries receive more foreign investment because they have better institutions. Exogenous variation in institutional quality, measured by the historical determinants of institutions, is the most important determinant of capital flows, causally explaining the Lucas Paradox.[5]

If capital is flowing to productive places in the long run, where long-run productivity is proxied by institutional quality, then why do we worry about capital flows from China to the United States, where the latter clearly has higher quality institutions? We worry

Figure 1: Net Capital Flows (% of GDP) by Type of Flow, 1970–2006. All countries



*Kalemli-Ozcan is an NBER Research Associate and a Professor of Economics at the University of Maryland. Her Profile appears later in this issue.

because the standard models imply that China must have received more capital flows than, say Zimbabwe, in a sample of developing countries, given China's faster catch-up productivity growth to the United States. This does not seem to be the case. The fast growing countries accumulate a large amount of reserves and export capital to slow growing countries, causing global imbalances.

In our 2011 paper, we decompose international capital flows into public and private components (private debt, FDI, equity flows versus public flows). We focus on a sample of developing countries since the 1980s and measure the amount of private and public capital flows these countries have received in relation to their catch-up growth. It turns out that upstream flows and global imbalances are manifestations of the same underlying phenomenon: sovereign-to-sovereign flows in the form of government debt from official lenders, reserve accumulation, and aid. International capital flows net of government debt and/or aid are positively correlated with growth. Government debt flows are negatively correlated with growth only if government debt is financed by another sovereign and not by private lenders, where government debt from private lenders also flows in the right direction.

As we document in these works, there is much more nuance to the direction of capital flows than is commonly appreciated. Standard model's predictions are best tested by using private capital flows, because government is outside this model. In relation to figure 1, China had a current account surplus and was a net lender overall. But this is because China is a net lender in what we call "sovereign-to-sovereign" flows. During the same period, China was a net borrower in terms of private flows, as it should be according to the neoclassical model.

Do Capital Flows Transmit Shocks? Tranquil Times versus Crisis Times

Private capital flows go in the right direction to productive countries but still might bring instability, especially short-term debt flows intermediated via banks. In fact, since the 2008 crisis, global banks have been seen as the "bad guys" who spread the crisis from the United States to other financial systems.

Does financial integration transmit shocks? Academic research has not been helpful in answering this question, because the theory suggests that financial integration may lead to a higher level of business cycle synchronization but can also cause a "decoupling" of business cycles between inter-connected economies. The key issue seems to be the nature of shocks.[6]

Take for example the case of two integrated economies, where one is hit by a negative shock. If the shock hits the banking sector, then problems in one country will likely spread to the other, as banks operating in both countries pull funds from the non-affected country, making the business cycles of the interlinked countries more synchronized. If, however, the negative shock hits the productivity of firms in a country while banks stay healthy, then return to capital falls and banks pull funds out of the affected country, amplifying the initial shock; this in turn makes the business cycles of financially interconnected economies diverge.

In recent work, Elias Papaioannou, Jose-Luis Peydro, Fabrizio Perri, and I examine the role of banking integration on business cycle synchronization since the 1970s for the OECD countries, explicitly allowing for the possibility that its effect may differ in tranquil versus financial crisis times.[7] Our results provide first-time evidence for the above theory.

We document that across country-pairs, there is a significant positive correlation between cross-border banking exposures and output synchronization. This should come as no surprise. The U.S. business cycles are both more synchronized and more financially linked with Canada than with France. There can be many reasons for such a result, including socioeconomic ties and less distance between certain pairs than others.

In sharp contrast to the positive cross-sectional correlation, in examining the *within* country-pair response of output synchronization to banking integration before the 2008 crisis we find a significantly negative association. This implies that in tranquil times, increases in banking integration within country-pairs over time are associated with more divergent output cycles. The negative association between bilateral financial linkages and business cycle co-movement is in line with the standard models summarized above, where in the absence of financial shocks, banking integration causes divergence in output cycles. Yet, this negative correlation between financial integration and output synchronization turned positive during 2008–10. This result is again in line with the above models showing that during crisis, financial integration facilitates co-movement via contagion.

These findings bridge two bodies of research in international macroeconomics and finance on the implications of financial integration: one looks at its effect on international business cycles and another focuses on financial contagion. The results imply that conduct of monetary policy becomes significantly harder within financially integrated currency areas. Financial integration magnifies output fluctuations across countries, thus making it difficult to conduct monetary policy in all regions. This problem is clearly illustrated nowadays in the euro area.[8] The high degree of integration has amplified country-specific shocks, leading to divergence in economic activity between countries of the south and the core; and, as global banks pull capital out of the periphery, the low policy rate of the ECB is not channeled to the south.

Do Capital Flows Bring Growth and Welfare?

The textbook case for supporting international financial integration is well known. In spite of possible contagion during major crisis, integration ultimately will bring growth and welfare. The standard model implies that when capital flows from low return to high return countries, the cost of borrowing will go down, boosting investment and growth. FDI will bring better technology and

know-how, together with financial stability, because FDI tends to be long-term, enhancing growth and welfare.

Given the large increase in FDI and portfolio equity flows (Figure 1) during the last decade, it is important to study the effect of such flows on both growth and volatility during tranquil times, for the developed and emerging countries. Using country-level data at the macro level is not appropriate for such a study, because many policy changes occur simultaneously with financial integration, growth, and volatility. That makes the identification of the individual effects harder.

In joint work, Bent Sorensen, Volosovych, and I use a novel dataset of firm-level balance sheets and foreign investment from 25 European countries for the period 1996–2006 (AMADEUS) to study the effect of foreign direct investment and portfolio equity investment on output volatility.[9] Our dataset differs starkly from Compustat or Worldscope-type data, because in our data 99 percent of the firms are privately held. We document a positive, highly robust, relationship between firm-level foreign investment and output (value added) volatility in cross-section and over time, both for emerging and developed Europe.

One theory that can explain this stylized fact is as follows: foreign investment brings technology and know-how to the target firm and improves diversification of ownership of capital. After receiving investments from multinationals, target firms might choose to invest in risky but high return projects that jointly increase their productivity and output volatility.[10] In joint work, Chris Fons-Rosen, Sorensen, Volosovych, Carolina Villegas-Sanchez, and I ask whether firms that receive foreign investment are becoming more productive and whether they create spillovers for domestic firms.[11] Notice that spillover effects from foreign to domestic firms are essential to realizing any aggregate effects of foreign investment, and they are the main reason behind the big policy push for FDI over last two decades. We extend our firm-level data to 60 countries worldwide (ORBIS), where we have more than 30 million shareholder/subsidiary links with sector and nationality of the investor.

Selection is a problem. Multinationals are likely to buy local firms with high future growth potential. Foreign firms may select themselves into high productivity sectors and drive weak domestic firms out of business. In such a case, domestic firms in the foreign activity sector might become more productive on average, but not because any single firm has become productive. Since our data encompasses many countries and sectors, we can control selection through the use of firm and sector-year effects. By exploiting the difference between financial and industry investment as exogenous variation, we control dynamic selection at the firm level.

Our results show that the positive correlation between foreign investment and productivity growth in developed countries is driven by selection. Multinationals target more productive companies but do not contribute to further increases. In emerging markets, there is evidence of productivity enhancing effects of multinationals on targets but the effects are small. For domestic firms, in both developed and emerging countries there are negative spillover effects from direct competitors. Only domestic firms with high initial productivity and suppliers of foreign owned firms benefit from knowledge spillovers. The effects are too small to contribute to aggregate productivity.

Summary

Overall, my research shows that the neoclassical model is alive and well in terms of predicting where and why capital should flow if we measure capital flows as private flows (FDI, portfolio equity and private debt). When it comes to the effects of capital flows in terms of further increases in productivity and knowledge spillovers, these are harder to find. Most of the positive correlations between foreign investment and growth at the aggregate level initially are explained by the fact that foreign capital is attracted to high productivity firms, sectors, and countries. Firm heterogeneity in terms of foreign investment and initial productivity are important for the realization of positive effects of FDI on productivity at the disaggregated level.

---

[1] *Notes: The data is from IMF, IFS. Net capital flows represent average net flows of FDI and portfolio equity investment, and debt (portfolio debt investment and other investment) divided by nominal GDP in current dollars, based on WB and IMF data and corresponding to the sum of the flows of assets (outflows) and liabilities (inflows), because assets have a minus sign as BOP convention. Total capital flows are represented by the negative of total current account flows. The data for current account is available for 186 countries; the data for FDI and portfolio flows and debt flows is available for 179 and 178 countries respectively, varying across years. FDI assets and liabilities correspond respectively to Direct Investment Abroad (line 78bdd) and Direct Investment in Reporting Economy (line 78bed). They include equity capital, reinvested earnings, other capital, and financial derivatives associated with various intercompany transactions between affiliated enterprises. Portfolio Equity Investment assets and liabilities correspond to Equity Securities Assets (line 78bkd) and Equity Securities Liabilities (line 78bmd). They include shares, stock participations, and similar documents that usually denote ownership of equity. Debt assets and liabilities include Debt Security Assets (line 78bld) and Debt Security Liabilities (line 78bnd), which include bonds and money market or negotiable debt instruments; Other Investment Assets (line 78bhd); and Other Investment Liabilities (line 78bid), which include all financial transactions not covered by direct investment, portfolio investment, financial derivatives, or other assets. The current account total corresponds to the Current Account excluding Exceptional Financing (line 78ald).*

[2] *This pattern of "borrowing from space" is related to the fact that countries' liabilities are better measured relative to their assets.*

[3] *If we divide the sample between rich and poor countries — where "rich" is defined as GDP per capita higher than 15,000 USD*

in 2000 dollars on average throughout the period — we see that poor countries receive more capital flows than rich, with a 5 percent current account deficit relative to a zero balance for the rich, on average. But this is not true when we look at FDI and equity investment, where poor countries receive less. Gross flows (sum of assets and liabilities) amount to 150 percent of GDP for rich countries and 25 percent of GDP for poor countries, on average.

[4] *L. Alfaro, S. Kalemli-Ozcan, and V. Volosovych, "Why does not Capital Flow from Rich to Poor Countries? An Empirical Investigation," NBER Working Paper No. 11901, December 2005, published in* Review of Economics and Statistics, *90, 2008, pp. 347–68, and "Sovereigns, Upstream Capital Flows, and Global Imbalances," NBER Working Paper No. 17396, September 2011.*

[5] *We used settler mortality rates to account for endogenous institutions, based on work by D. Acemoglu, S. Johnson, and J. Robinson in "The Colonial Origins of Comparative Development: An Empirical Investigation,"* American Economic Review, *91, (2001), pp. 1369–1401.*

[6] *Most theoretical works focus on one type of shock, that is, either tranquil times or financial crises. For theoretical models where both mechanisms are in place, see B. Holmsrom and J. Tirole, "Financial Intermediation, Loanable Funds, and the Real Sector,"* Quarterly Journal of Economics, *112, 3, (1997), pp. 660–91; F. Perri and V. Quadrini, "International Recessions," NBER Working Paper No. 17201, July 2011, and S. Kalemli-Ozcan, E. Papaioannou, and F. Perri , "Global Banks and Crisis Transmission," NBER Working Paper No. 18209, July 2012, forthcoming in* Journal of International Economics.

[7] *S. Kalemli-Ozcan, E. Papaioannou, and J. Peydro , "Financial Globalization, Financial Regulation, and the Synchronization of Economic Activity," NBER Working Paper No. 14887, April 2009, forthcoming in* Journal of Finance, *2012; S. Kalemli-Ozcan, E. Papaioannou, and F. Perri, "Global Banks and Crisis Transmission," NBER Working Paper No. 18209, July 2012, and forthcoming in* Journal of International Economics.

[8] *See R. Mundell, "A Theory of Optimum Currency Areas,"* American Economic Review, *51, 4, (1961), pp.657–65, for the classical treatment of the issue.*

[9] *S. Kalemli-Ozcan, B. Sorensen, and V. Volosovych, "Deep Financial Integration and Volatility," NBER Working Paper No. 15900, April 2010.*

[10] *This argument is detailed in M. Obstfeld, "Risk-Taking, Global Diversification and Growth,"* American Economic Review, *84, 5, (1994), pp. 1310–29, and D. Acemoglu and F. Zilibotti, "Was Prometheus Unbound by Chance? Risk, Diversification, and Growth,"* Journal of Political Economy, *105, 4, (1997), pp. 709–51, who show how diversified capital ownership allows firms to choose riskier projects leading to higher growth and volatility.*

[11] *C. Fons-Rosen, S. Kalemli-Ozcan, B. Sorensen, V. Volosovych, and C. Villegas-Sanchez, "Quantifying Productivity Gains from Foreign Investment," forthcoming as an NBER Working Paper.*

Kidney Exchange

Alvin E. Roth*

More than 90,000 patients are on the U.S. waiting list for a kidney transplant from a deceased donor, and only 11,000 or so such transplants are accomplished each year. So, the waiting is long and costly, sometime fatally so. But healthy people have two kidneys and can remain healthy with only one, which also makes it possible to receive a kidney from a living donor — around 6,000 such transplants were accomplished in 2011. Nevertheless, someone who is healthy enough to donate a kidney may be unable to donate to his or her intended recipient because of various types of donor-recipient incompatibility. This is the origin of kidney exchange. In the simplest case, two incompatible patient-donor pairs exchange kidneys, with each patient receiving a compatible kidney from the other's donor. The first kidney exchange in the United States was performed at the Rhode Island Hospital in 2000, when doctors there noticed two incompatible patient-donor pairs who could benefit from exchange. Shortly after that, Tayfun Sonmez, Utku Unver, and I proposed a way to organize a multi-hospital kidney exchange clearinghouse[1], and began discussions with Dr. Frank Delmonico of Harvard Medical School, that soon led to the founding of the New England Program for Kidney Exchange.[2] Together with Itai Ashlagi, we have since

**Roth is a Research Associate in the NBER's Program on Labor Studies and a visiting professor of economics at Stanford University. His Profile appears later in this issue.*

assisted in the formation and operation of other kidney exchange networks operating around the country.

In the United States and most of the world it is illegal to buy or sell organs for transplant.[3] As Jevons (1876)[4] noted, one obstacle to two-way barter exchange is the need to find a counterparty who has what you want and also wants what you have. One way to reduce the difficulty of finding these double coincidences is to assemble a large database of interested patient-donor pairs. Another is to consider a larger variety of exchanges than those between just two pairs: for example, a cycle of exchange among three pairs, or a chain that begins with a donation by a non-directed donor (such as a deceased donor, or an altruistic living donor) to the patient in an incompatible patient donor pair, whose donor "passes it forward" to another such pair or ends the chain with a donation to someone on the waiting list for a deceased donor (that is, the chain ends when a donation is made to a patient who does not have a willing but incompatible live donor).

Our 2003 paper proposed kidney exchange that integrated cyclic exchanges of all sizes and chains beginning with a non-directed donor and ending with a donation to someone without a living donor. We focused on two kinds of incentive issues that seemed likely to be important in a mature system of kidney exchange, both concerned with aligning incentives so as to make it safe and simple to participate. First, we showed how exchanges could be arranged so that they would be in the core of the game, which means that no coalition of patient-donor pairs could go off on their own, or to a competing exchange, and do better than to accept the proposed exchanges. Second, we showed how this could be accomplished in a way that made it a dominant strategy for patients (and their surgeons) to reveal the medical information that determined the desirability of each potential transplant. It is worth noting that the tools we used built on theory that was initially proposed in a very abstract setting: Shapley and Scarf (1974) studied a "top trading cycle" algorithm for trading indivisible goods without money and showed that it produced an allocation in the core[5], and Roth (1982)[6] showed that the top trading cycle algorithm made it a dominant strategy for traders to reveal their true preferences. Abdulkadiroglu and Sonmez (1999)[7] extended this model to deal with assignment of dormitory rooms when some students already had rooms, some did not, and some rooms might be vacant, so that assignment would involve chains as well as cycles.

We observed that the efficient chains and cycles in kidney exchange mostly would be short but occasionally would be long, which presented a logistical problem, since, for incentive reasons, all surgeries in a given exchange would be performed simultaneously (because contracts can't be written on kidneys). This means that even an exchange between two pairs requires four operating rooms and surgical teams, for the two nephrectomies (kidney removal from the donor) and two transplants. A three-way exchange would require six. When we presented this initial proposal to our surgical colleagues, led by Frank Delmonico, they felt it was a critical problem—the prospect of four simultaneous surgeries was daunting enough. They asked us to present a proposal with the more modest aim of organizing exchanges involving only two-way exchanges.

Our new, more limited proposal[8]and the accompanying software formed the basis for organizing the New England Program for Kidney Exchange,[9] and was widely shared and explained and soon adapted for use elsewhere. Almost simultaneously, we began exploring with our surgical colleagues the possibilities of including larger exchanges and chains.[10,11,12] (It speaks volumes about the relative publishing speed of Economics and Medicine to note that the follow-up paper which reported in the *American Journal of Transplantation* how longer exchanges actually had been carried out was published a year later than the publication of the original 2005 NBER Working Paper analyzing such exchanges.)

Although the three-way chain reported in that AJT paper was performed simultaneously (and hence involved six operating rooms and surgical teams), the paper also proposed that chains that begin with a non-directed donor might not need to be performed simultaneously. The argument was a simple cost-benefit analysis. The reason that cyclic exchanges are performed simultaneously is that if they were not, some patient-donor pair would have to give a kidney before getting one, and if the cycle were to be broken subsequently, that pair would suffer a grievous loss. The donor in the pair would have undergone a nephrectomy that yielded no benefit to the recipient in the pair, and there would no longer be a kidney with which to participate in a future exchange.

Now consider a chain that begins with a non-directed donor, who donates to some incompatible patient-donor pair under the understanding that they will subsequently donate to another, and so on. Every pair in this chain will receive a kidney before they donate one. If the chain is broken, then the pair that was scheduled but fails to receive a kidney will be disappointed, but not grievously harmed. They are not worse off than they were before the non-directed donor came forward, and, in particular, they still have a kidney with which to participate in some future exchange. Hence the cost of a broken link in a chain initiated by a non-directed donor is much less than that of a broken link in an exchange among a cycle of patient-donor pairs.

In 2007, Mike Rees, a pioneer of kidney exchange and the founder of the Alliance for Paired Donation, which is one of the most active networks, began the first such non-simultaneous chain. It was reported on in Rees et al. (2009), at which point it had accomplished ten transplants (and 20 surgeries), many more than could have been done simultaneously.[13] Since then, non-simultaneous non-directed donor chains have become the fastest growing part of kidney exchange, even though the number of non-directed donors is small. In some cases a non-directed donor has initiated a chain of more than 30 transplants.

Ashlagi and I have worked to under-

stand why long chains are so useful, and how to structure them. As kidney exchange has grown and become a standard tool of transplantation, hospitals are more able to do some exchanges among their own patients. This means the players in the kidney exchange game have changed: where it used to be enough to think of the incentives of patients and donors and their surgeons, now the directors of transplant centers are players, and they see many patient-donor pairs. Their strategy sets now include which pairs to show to a centralized exchange. The present organization of kidney exchanges gives them some incentives to withhold their easy-to-match pairs. This could be fixed by taking account of which hospitals enrolled easy-to-match pairs and using this information (in a sort of "frequent flier program") to give some increased probability of matching to patients at those hospitals.[14] But this faces important political obstacles and has so far not been adopted. Partly as a result of the withholding of easy-to-match pairs, the percentage of patients enrolled in kidney exchange networks that are hard to match, even to a blood-type compatible donor, has skyrocketed.

We can organize patient and donor data in a compatibility graph, in which each node represents a patient and her incompatible donor(s), and an edge goes from one node to another whenever the donor in the first node is compatible with the patient in the second node. As patients have become harder to match, the compatibility graphs have become sparser, that is, they contain fewer edges. When we look at the data of the kidney exchange networks with which we work, there is a densely connected sub-graph of the relatively few fairly easy-to-match pairs, and a sparse sub-graph of many hard-to-match pairs (this is joint work with David Gamarnik and Mike Rees). Within the easy-to-match sub-graph, many patients could be transplanted with the aid of two-way or three-way exchanges, but within the sub-graph of hard-to-match pairs, only long chains offer the chance of transplanting many patients.[15] Non-directed donors have a chance of starting those long chains, and the presence of easy-to-match pairs allows more hard-to-match pairs to be included.

Despite the growing success that kidney exchange has had in facilitating transplants from living donors, the list of people waiting for kidney transplants from deceased donors continues to grow. Deceased donor organs are a scarce resource of an unusual kind, because their supply depends on decisions to donate made by potential donors (while still living) and their next of kin (immediately afterwards). Consequently there are market design issues associated with how donations are solicited, and how organs are allocated, both of which may influence the donation decision and hence the supply. Judd Kessler and I have begun to investigate this:[16] we begin with an experimental investigation motivated by a priority allocation scheme just put into place in Israel, in which people who have registered as donors will be given some priority in case they need to receive an organ for transplant, and so will members of their immediate family.

While it is natural that economists should investigate institutions that facilitate exchange, many people (including some economists) find it surprising that economists should be helping to design the institutions of kidney exchange. This is a natural outgrowth, however, of two strands in modern economics: market design in general[17], and the study of matching markets. Matching markets are those in which price does not do all the work of determining who gets what, and they include some of the important passages in our lives, from school choice and college admissions to marriage and labor markets. In none of these can you simply choose what you want — you also have to be chosen. In some of these, economists have begun to help design the matching institutions.

Economists should welcome opportunities to learn how to be engineers.[18]

---

[1] *A.E. Roth, T. Sonmez, and M. Utku Unver, "Kidney Exchange," NBER Working Paper No.10002, September 2003, and* Quarterly Journal of Economics, *119 (2) (May 2004), pp. 457–88.*

[2] *A.E. Roth, T. Sonmez, and M. Utku Unver, "A Kidney Exchange Clearinghouse in New England,"* American Economic Review, Papers and Proceedings, *95 (2) (May 2005), pp. 376–80.*

[3] *A.E. Roth, "Repugnance as a Constraint on Markets," NBER Working Paper No. 12702, November 2006, and* Journal of Economic Perspectives, *21 (3) (Summer 2007), pp. 37–58.*

[4] *W.S.Jevons,* Money and the Mechanism of Exchange, *New York: D. Appleton and Company, 1876.*

[5] *L. Shapley and H. Scarf, "On Cores and Indivisibility,"* Journal of Mathematical Economics, *1(1) (1974) pp. 23–37.*

[6] *A.E. Roth, "Incentive Compatibility in a Market with Indivisible Goods,"* Economics Letters, *Vol. 9 (1982) pp.127–32.*

[7] *A. Abdulkadiroglu and T. Sonmez, "House Allocation with Existing Tenants,"* Journal of Economic Theory, *LXXXVIII (1999), pp.233–60.*

[8] *A.E. Roth, T. Sonmez, and M. U. Unver, "Pairwise Kidney Exchange," NBER Working Paper No. 10698, August 2004, and* Journal of Economic Theory, *125 (2) (December 2005) pp.151–88.*

[9] *A.E. Roth, T. Sonmez, and M. U. Unver, "A Kidney Exchange Clearinghouse in New England".*

[10] *S. L. Saidman, A. E. Roth, T. Sönmez, M. U. Ünver, and F. L. Delmonico, "Increasing the Opportunity of Live Kidney Donation By Matching for Two and Three Way Exchanges,"* Transplantation,*81(5) (March 15, 2006) pp. 773–82.*

[11] *A. E. Roth, T. Sönmez, M. U. Ünver, F. L. Delmonico, and S. L. Saidman, "Utilizing List Exchange and Undirected Good Samaritan Donation through 'Chain' Paired Kidney Donations,"* American Journal of Transplantation, *6, (11) (November 2006) pp. 2694–705.*

[12] *A. E. Roth, T. Sonme, and M. U. Unver, "Efficient Kidney Exchange: Coincidence of Wants in Markets with Compatibility-Based Preferences," NBER Working Paper No. 11402, June 2005, and* American Economic Review, *97 (3) (June 2007) pp. 828–51.*

[13] *M. A. Rees, J. E. Kopke, R. P. Pelletier, D. L. Segev, M. E. Rutter, A. J. Fabrega, J. Rogers, O. G. Pankewycz, J. Hiller, A. E. Roth, T. Sandholm, M.U. Ünver, and R. A. Montgomery, "A Non-Simultaneous Extended Altruistic Donor Chain,"* New England Journal of Medicine, *360(11) (March 12, 2009) pp. 1096–1101.*

[14] *I. Ashlagi and A.E. Roth, "Individual Rationality and Participation in Large Scale, Multi-Hospital Kidney Exchange," NBER Working Paper No. 16720, January 2011.*

[15] *I. Ashlagi, D. Gamarnik, M. A. Rees, and A. E. Roth, "The Need for (long) Chains in Kidney Exchange," NBER Working Paper No.18202, July 2012.*

[16] *J. B. Kessler and A. E. Roth, "Organ Allocation Policy and the Decision to Donate," NBER Working Paper No. 17324, August 2011, and* American Economic Review, *102 (August 2012). pp. 2018–47.*

[17] *A.E. Roth, "What Have We Learned from Market Design?" NBER Working Paper No. 13530, October 2007, and* Economic Journal, *118 (March 2008) pp.285–310.*

[18] *A.E. Roth, "The Economist as Engineer: Game Theory, Experimentation, and Computation as Tools for Design Economics,"* Econometrica, *70 (4) (July 2002) pp. 1341–78.*

The Sources and Consequences of Productivity Differences

Chad Syverson*

Economists have consistently found both large and persistent differences in measured productivity across producers, even within narrowly defined industries. The size of these differences is striking: for instance, within U.S. 4-digit SIC manufacturing industries (such as saw blade manufacturing), the plant at the 90th percentile of the industry's productivity distribution typically obtains almost *twice* as much output with the *same measured inputs* as the plant at the 10th percentile of productivity. (These figures, and all those described below, use total factor productivity measures. They reflect the amount of output that a producer obtains from a given combination of labor, capital, and intermediate inputs.) And U.S. manufacturing is not exceptional in this regard; in fact, researchers have documented even larger dispersion in other sectors and countries.

The observed persistence of producers' productivity levels indicates that industries typically contain both firms that appear to have figured out their business and those that are woefully lacking in such knowledge. Far more than bragging rights are at stake, because higher productivity producers are more likely to survive than their less efficient industry competitors.

The discovery of these ubiquitous, large, and persistent productivity differences has shaped research agendas in a number of fields, including (but not limited to) macroeconomics, corporate finance, industrial organization, labor, and trade. I have studied various aspects of the sources and consequences of productivity dispersion as a part of my research agenda; this essay summarizes that work.

Two Sources of Productivity Differences

In a recent survey article, I review the research over the past decade that has sought to explain the sources of observed productivity differences.[1] I split the explanations into two categories. One includes factors that operate within the plant or firm and which directly affect productivity at the producer level. These are the "levers" that management or others potentially can use to influence productivity. The second category includes forces that are external to the firm: elements of the industry or market environment that can induce productivity changes or support productivity dispersion. I have researched factors in both categories.

Levers that Influence Productivity

On the "lever" side of the ledger, Steven Levitt, John List, and I look at the mechanisms that underlie learning by doing — productivity gains achieved through the very act of producing.[2] Using extremely detailed data from an assembly plant of a major auto producer, we find that productivity gains from learning arrive quickly and in force. Defects per vehicle fall by more than 80 percent in the first eight weeks of production. Interestingly, when the plant's second shift comes on line at this point, the learning process does not begin again. Instead, the second shift actually comes on line at defect rates lower than the first shift's contemporaneous rates, despite the first shift's two month head start in production. And, while worker absenteeism statistically affects defect rates, its impact is economically small. Furthermore, the hundreds of assembly processes on the line have

**Syverson is an NBER Research Associate and a Professor of Economics at the University of Chicago's Booth School of Business. His profile appears later in this issue.*

highly correlated defect rates across shifts, even though the workers completing these tasks are different. Taken together, these patterns illustrate one of our main findings about the learning mechanisms at the plant: rather than remaining with workers, much of what is learned very quickly becomes embodied in the plant's physical or organizational capital. This finding is consistent with the institutional processes that plant management puts in place to encourage knowledge dissemination.

In a series of papers, Enghin Atalay, Ali Hortaçsu, and I examine the connections between firms' vertical structures and their plants' productivity levels.[3] We find that vertically integrated plants have higher productivity levels than their non-integrated industry cohorts. However, the evidence suggests that little of this difference is related to the firms' vertical structures per se, but rather to other factors correlated with integration status. In fact these productivity differences — and the firm's decisions about whether to have a vertical structure in the first place — are not usually related to the movement of goods along the production chain. Using detailed shipment-level data on the flow of goods throughout the economy, we find that vertically integrated firms' upstream plants ship a surprisingly small amount to downstream plants in their firm (that is, small relative to both the firms' total upstream production and their downstream needs). Almost half of upstream plants report no shipments to downstream units inside their firm. About 90 percent of upstream plants ship less than a third of their output internally. These patterns suggest that vertical ownership is not usually about moderating goods transfers along production chains. We propose and find suggestive evidence that the primary purpose of integration instead is to facilitate within-firm transfers of intangible inputs (for example, managerial oversight or intellectual capital).

External Factors that Influence Productivity

My research on the external factors shaping productivity has looked at the roles of both competition and regulations in influencing producer productivity levels. Most models of competition among heterogeneous-productivity producers share a prediction that a greater ability or willingness of consumers to substitute across producers either will induce low productivity suppliers to improve their efficiency or will force them to exit. Either effect truncates the market's equilibrium productivity distribution from below, thereby raising average productivity and reducing productivity dispersion.

I test this prediction in studies looking both across industries and across markets within an industry. The across-industry analysis uses producer-level data from 443 U.S. manufacturing industries and finds that industries with more substitutable output — measured in several ways, including aspects of spatial, physical, and brand-driven differentiation — have less productivity dispersion and higher median productivity levels.[4] The within-industry investigation focuses on the ready-mixed concrete industry.[5] The industry's homogeneous product and very high transport costs make the density of concrete producers in a market a primary determinant of the intensity of competition (that is, substitutability). There too, the predicted truncation effect of substitutability is observed in the data. Markets with denser construction activity (an exogenous shifter of concrete producer density) have higher lower-bound productivity levels, higher average productivity, and less productivity dispersion. In follow-up work, I demonstrate that these patterns of competition-driven selection on costs also are reflected in ready-mixed prices.[6]

My recent work with Michael Greenstone and John List considers regulation's effect on plants' productivity levels.[7] We use detailed production data from nearly 1.2 million plant observations from the 1972–93 Annual Survey of Manufactures to measure the economic costs of the Clean Air Act Amendments. We track productivity growth at plants from heavily polluting industries that are located in counties declared by the EPA to be in nonattainment with the Act's pollution limits, a determination that subjects those plants to command-and-control-style abatement mandates. We compare productivity growth at these plants to their industry cohorts located in counties that are in attainment with the Act's provisions, and to plants in non-polluting industries that are free from regulation in all counties. We find that for surviving plants in heavily polluting industries, a nonattainment designation and its associated abatement mandates result in an average 4.8 percent decline in the plants' total factor productivity. In plain language, this means the amount of output that the plants are able to produce from a given amount of inputs (that is, labor, capital, and materials) is 4.8 percent lower than before the abatement mandates. This output loss corresponds to an annual economic cost from the regulation of manufacturing plants of roughly $21 billion in 2010 dollars, about 8.8 percent of average annual manufacturing sector profits over the sample period.

Productivity vs. Demand

While productivity is typically thought of as a feature of the production technology, as actually measured in producer micro-data it generally reflects more than just supply-side forces. Much of the work I've just described, and most of the broader literature investigating productivity differences among businesses, uses revenue to measure output because business-level price indexes are rarely available. This means that within-industry price differences are embodied in output and productivity measures. If prices reflect in part idiosyncratic demand shifts or market power variation across producers — a distinct likelihood in many industries — then high "productivity" businesses may not be especially technologically efficient.

A new strand of research has begun to extend the productivity literature to also explicitly account for such idiosyncratic demand effects. Lucia Foster, John Haltiwanger, and I have been active in this area. We take advantage of the avail-

ability of physical output data for a select set of "commodity-like" product industries (for example, cardboard boxes, white pan bread, and sugar). This lets us measure not just the standard revenue-based productivity metric, but also its two components: physical-quantity-based productivity (number of units of output per unit input, reflecting more closely the pure supply-side concept of productivity) and average unit price. We show that there are important differences between revenue and physical productivity.

In one paper, we consider the separate roles that supply- and demand-side fundamentals play in driving selection and survival in heterogeneous-producer industries.[8] We show that physical productivity is inversely correlated with price while revenue productivity is positively correlated with price. This means that previous work linking (revenue-based) productivity to survival has confounded the separate and opposing effects of technical efficiency and demand on survival, understating the true impacts of both. Perhaps most strikingly, we find that even in these near-commodity industries, a producer's demand is particularly important for its survival prospects. A given-sized shift in a producer's demand level has four times the effect on its likelihood of surviving as does the same-sized shift in its physical productivity.

A second paper looks at the role of demand in explaining the well documented fact that new businesses on average are much smaller than their established industry competitors, and that this size gap closes slowly.[9] We show that these patterns are not a result of physical productivity gaps, but instead reflect differences in demand. Even though new producers are technically more efficient, they sell only a fraction of the output of their more established competitors. Estimating a dynamic model of plant expansion in the presence of a demand accumulation process (for example, building a customer base), we find that this accumulation results mostly through businesses' active investments in building demand, rather than through passive processes tied simply to the passage of time. We also show that within-firm demand spillovers, like those conferred by established firms on their new plants, affect plants' initial demand levels but not their growth.

---

1 *C. Syverson, "What Determines Productivity?," NBER Working Paper No. 15712, January 2010, and* Journal of Economic Literature, *49(2) (2011), pp. 326–65.*

2 *S. D. Levitt, J. A. List, and C. Syverson, "Toward an Understanding of Learning by Doing: Evidence from an Automobile Assembly Plant," NBER Working Paper No. 18017, April 2012.*

3 *A. Hortaçsu and C. Syverson, "Cementing Relationships: Vertical Integration, Foreclosure, Productivity, and Prices," NBER Working Paper No. 12894, February 2007, and* Journal of Political Economy, *115(2) (2007), pp. 250–301, and E. Atalay A. Hortaçsu, and C. Syverson, "Why Do Firms Own Production Chains?" NBER Working Paper No. 18020, April 2012.*

4 *C. Syverson, "Product Substitutability and Productivity Dispersion," NBER Working Paper No. 10049, October 2003, and* Review of Economics and Statistics, *86(2), May 2004, pp. 534–50.*

5 *C. Syverson, "Market Structure and Productivity: A Concrete Example," NBER Working Paper No. 10501, May 2004, and* Journal of Political Economy, *112(6) (2004), pp. 1181–1222.*

6 *C. Syverson, "Prices, Spatial Competition, and Heterogeneous Producers: An Empirical Test," NBER Working Paper No. 12231, May 2006, and* Journal of Industrial Economics, *55(2) (2007), pp.197–222.*

7 *M. Greenstone, J. A. List, and C. Syverson, "The Effects of Environmental Regulation on the Competitiveness of U.S. Manufacturing," NBER Working Paper No. 18392, September 2012.*

8 *L. Foster, J. C. Haltiwanger, and C. Syverson, "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" NBER Working Paper No. 11555, August 2005, and* American Economic Review, *98(1) (2008), pp. 394–425.*

9 *L. Foster, J. C. Haltiwanger, and C. Syverson, "The Slow Growth of New Plants: Learning about Demand?" NBER Working Paper No. 17853, February 2012.*