FROM THE LAB

FROM THE LAB

Director's Message
Finn Kydland

This issue of *From the Lab* contains summaries of two conferences organized by LAEF. One was titled "OTC Markets and Securities," with program put together by Batchimeg Sambalaibat, Indiana University. The conference was designed to foster discussion and research on OTC financial markets. Its themes included the following topics: Why markets are organized and assets designed the way they are? Why certain assets trade in decentralized vs. centralized markets and in opaque vs. transparent markets? What do networks look like in the data, how do trading relationships form, and what is the optimal network structure in mitigating frictions and contagion? What are the macroeconomic and asset pricing implications of these institutional and microstructure frictions and features? We welcomed also research with the potential to be applied to the context of OTC market microstructure even if it does not, as is, pertain directly to market microstructure. Examples are new results in IO, contract and game theory, informational economics, and asset pricing.

In 1995, a volume titled "Frontiers of Business Cycle Research," edited by Tom Cooley, NYU, was published. In the decade leading up to the early 1990s, the field of aggregate economics (also called macroeconomics) had undergone a reorientation of methodology. Dominant until the 1980s was what one could call the system-of-equations approach. A model would include a consumption function, investment function, labor-demand and labor-supply equations, and so on. The belief was that, to make the models quantitative, statistical methods (econometrics) could reliably be used to determine the coefficients of the equations. An important problem with that approach transpired when, in 1973, Robert Lucas published a paper titled "Econometric Policy Evaluation: A Critique." Focusing on consumption and investment behavior as examples, he showed that the coefficients of such equations would not be invariant to policy change. In other words, that framework was inappropriate for policy analysis.

During the 1980s, an alternative framework had emerged, one in which a researcher's model was explicit about the intertemporal optimizing decision problems faced by the model inhabitants.

People and businesses were placed in dynamic, stochastic model environments as the questions being addressed called for, and through computational experiments one could observe how rational agents would behave in those environments. The models were calibrated to statistical moments (e.g., averages) that were known with high degree of accuracy in order to make the answers implied by the models as reliable as possible. The early applications focused on the role of the ebb and flow of technological change for the business cycle, hence the term real business cycle (RBC) analysis. The contrast between growth and business cycle theory became less pronounced. And as Tom Cooley and Gary Hansen, for example, showed, the same general approach would be used also for monetary analysis. Today, a more general term, Dynamic Stochastic General Equilibrium (DSGE) analysis, is often used.

Initially, only a limited number of researchers at a handful of universities fully embraced the new business cycle approach. So when it became clear that this would be the approach of the future, but it had not yet been incorporated in textbooks in any systematic way, Tom Cooley saw the need for a volume to flesh out details, mathematical and otherwise, involved in this modeling approach. He lined up 17 experts to write 12 chapters on different aspects of the approach.

Today, one hardly ever sees an aggregate business cycle model published that doesn't fall within this general framework. The difference is, modeling tools have developed, computers have become more and more powerful, and one can now incorporate all kinds of detail, including life-cycle behavior of generations of model inhabitants and, more generally, important sources of heterogeneity across model agents, features that would have been computationally infeasible before, say, the 1990s and 2000s. The idea of the LAEF conference titled "The New Frontier of Business Cycle Research," organized by Zach Bethune, Sophie Osotimehin, and Eric Young, all at University of Virginia, was basically to take stock of where we are today in terms of the subject areas covered by the original Cooley volume.

Over the Counter Markets and Securities 2nd Workshop

October 2017

Ana Babus – FRB of Chicago
Marco Di Maggio – Harvard Business School
Selman Erol – Carnegie Mellon University
Vyacheslav Fos – Boston College
William Fuchs – University of Texas, Austin
Benjamin Golub – Harvard University
Barney Hartman-Glaser – UCLA Anderson
Bernard Herskovic – UC Los Angeles
Burton Hollifield – Carnegie Mellon
Gregor Jarosch - Princeton
Finn Kydland – University of California, Santa Barbara
Ricardo Lagos – University of Pennsylvania
Richard Lowery – University of Texas, Austin
Semyon Malamud – EPFL
Peter Rupert - University of California, Santa Barbara
Batchimeg Sambalaibat– Indiana University
Semih Uslu – John Hopkins
Venky Venkateswaran – FRB Minneapolis
Colin Ward – University of Minnesota
Pierre-Olivier Weill – UC Los Angeles
Haoxiang Zhu – MIT Sloan

An Intermediation-Based Model of Exchange Rates

Semyon Malamud and Andreas Schrimpf

Research in international economics and finance has long sought to explain anomalies exhibited by international capital markets. Attempts towards explaining these puzzles often rely on building general equilibrium models which incorporate salient features of the market. In their paper, Malamud and Schrimpf contribute to this literature by creating a model of international financial markets with intermediation frictions. The authors seek to show that frictions created from the rent-seeking behavior of financial intermediaries can explain many of the abnormalities exhibited by these capital markets.

The basic model environment comprises a discrete-time international goods monetary economy with cash-in-advance constraints. There are multiple countries, each of which has integrated financial markets. It splits populations between two types of agents: customers and intermediaries. Customers can access domestic financial markets, but must use a price-setting intermediary to trade international financial instruments. Intermediaries face a complete,

dealer-to-dealer market through which they buy international financial assets. When households and intermediaries meet, the intermediaries extract rents in the form of markups based on customers' optimal demand for the international assets.

Under this model, specification the intermediaries play a key role in determining both international risk premia and exchange rates due to their role as the marginal investors in international financial markets. Their markups thus and influence pricing in the global market. It is through intermediaries and their markups that monetary and stabilization policies will affect state prices. Domestic monetary policy decisions result in customers adjusting their demand for insurance, leading intermediaries to shift their demand for foreign currencies and the change of state prices.

One of the most appealing features of Malamud and Schrimpf's model is the sheer number of financial puzzles it seems to explain using only the intermediation frictions. One such puzzle the model gives insight into is the reserve currency paradox and the safe haven properties of exchange rates. While frictionless monetary models predict currency depreciation for countries with large intermediation sectors during global downturns,

the authors can show that with intermediation frictions the so-called markup channel can overturn the risk sharing channel, leading to currency appreciation.

A safe haven currency may arise in the model as a response to noise in monetary and stabilization policies on exchange rates. Customers' expectations about future policy affect their demand for international financial instruments. Thus, customers in countries with noisy monetary policy will seek to insure against risk by using intermediaries, who will charge markups for this service, creating a situation in which these currencies will depreciate. Foreign currencies from less noisy countries will emerge as safe haven assets. These dynamics cause the US dollar to be an important priced risk factor in the pricing kernel and help explain deviations from covered interest rate parity.

Conference participants expressed a wish to see the empirical supporting evidence added as a supplement to the aforementioned model. In addition, one participant advocated for an increased focus on the interaction between monetary policy and the intermediary markups discussed within the paper.

Financial Industry Dynamics

Richard Lowery and Tim Landvoigt

It is well documented that firms in the financial sector display significant heterogeneity across a variety of dimensions. Much of this

diversity can be difficult to explain using existing financial models. For example, security issuance, and derivatives trading in OTC markets remains concentrated among a few firms, despite the large profit margins and lack of explicit barriers to entry. Landvoigt and Lowery, in their trying to explain much of the financial sector

heterogeneity, design a dynamic model of entry, exit and investment which incorporates different expertise.

Firms in this model undergo both security production and security trading. They may invest in one of two types of expertise: that which allows the firm to produce better securities, and that which allows them to improve

their terms of trade with competitors. Investing in either type of expertise imposes externalities on the market. If a firm invests in producing better securities than they create a positive externality if they sell the better security to another firm. In contrast, investments in the trade expertise impose negative externalities on one's competitors as improving the terms of trade for one party makes the other party worse off. Because trading expertise creates this negative externality on competitors, it may block entry for potential entrants.

To analyze the model, the authors do several numerical simulations using different parameter assumptions and industry configurations. The simulations show that this model can describe several stylized facts about the makeup of the financial sector. First, the authors find that simulations support the notion that sectors with informed incumbents, as represented by investments in trading expertise, have significant barriers to entry for entrants with little expertise. Also, it incentivizes incumbents to continue investing in trading expertise so as not to fall behind competitors. Thus, concentrated industries with high profit margins and no explicit barriers to the model can explain entry.

The authors' simulations suggest that industries with high security quality can be very fragile. Markets with many firms of low trade expertise can devolve into a state featuring few firms with high levels of trade expertise. As soon as firms invest in trade expertise, less informed competitors will drop out, leading to increased returns for further investment in trade expertise. In addition, the authors find that greater heterogeneity in expertise leads to greater price dispersion.

One conference participant pointed out that the model as presented does not allow for takeovers or mergers of firms. These actions may be important to include in the model as they play a role in determining industry size and concentration and occur. Participants questioned whether the trade expertise might create positive externalities through a price efficiency channel.

Network Hazards and Bailouts

Selman Erol
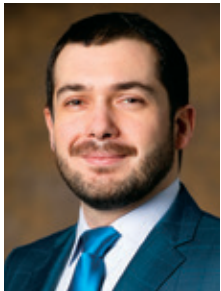
After the crisis, systemic risk became a prominent topic in economic research. Systemic risk can arise when the shocks to the smaller parts of the system can propagate through the interconnections and can damage the entire system. This paper addresses systemic risk in financial networks and asks whether anticipation of bailouts could create a moral hazard in the form of the financial networks which the author calls as the "network hazard".

Erol builds a model where large and small banks connect in a network. Each bank pair can make a joint investment called a link. They define the network as the set of banks and their joint investments. The model has three stages. In the first stage, they form the links between banks. These are the joint projects financed by borrowing from the households. There is no risk in terms of the return of the project, but a counter-party risk is that the counterparts may default since the project needs both counter-parties to put effort to mature. In the second stage, the banks get exogenous shocks, either good or bad. The bad shocks lead the operational costs to be too high such that the bank defaults on the project. After they realize the shocks, in the third stage, the banks decide which projects to continue and which projects to default depending on the counterpart's cost or decision to continue or not.

In a network setting, the banks may find it optimal to default if many of its counter-parties default even when they have good shocks. And first order counter-parties might default due to the cost placed by their own defaulting counter-parties which exposes the banks to the second order counter-party risk. This is a formal definition of network hazard and this creates contagion in the system. With absence of government intervention, the banks limit their exposure to second order counter-party risk by forming dense clusters.

The government intervention takes the form of transfers to the good banks which might default because of their bad counter-parties. With anticipation of bailouts, the banks will form an interconnected network as the second order counterparts risk is eliminated by the optimal intervention policy. This increases the potential contagion in the networks. When the banks are not ex-ante identical, a core-periphery structure emerges where high demanding banks becomes in the center of the network. This structure makes peripheral banks less resilient as the bad shocks to core banks may induce a contagion in the network. The opposite is also true, the core banks can serve a buffer against contagion. As a result, the terrible and the great outcomes become more likely which generates higher volatility in the system.

A participant asked whether the banks can communicate and adjust the payoffs in case of a bad shock. Erol

answered that the model does not incorporate renegotiation. However even if renegotiation is available, the insight about the network hazard will not change. The discussant, Babus, talked about the tradeoffs and implications of the model. With the help of simple examples, she showed that over-borrowing or under borrowing from households are substitutes for links in a network and banks can themselves stop contagion. In terms of the application of the model, syndicated loans can be thought in this context but banks don't actually default when the syndicate fails. If we were to think about venture capitals, then bailing out venture capitals is not something we observe. Erol answered that re-borrowing from households is an interesting extension to consider, but the goods are perishable in his model.

Illquid Spirals in Coupled Over-the-Counter Markets

Christopher Aymanns, Co-Pierre Georg and **Benjamin Golub**

Financial institutions often get funds by taking a loan using a security they own as collateral. Secured loans of this form have become a major source of funding liquidity for an array of financial institutions, including banks, money market funds, and security lenders. They are fragile in distinctive ways. A decline in funding liquidity could lead to a further decline in market liquidity. The basic principles of the price-mediated feedback are well-understood for centralized repo and collateral markets. Trade in both the repo and in the collateral market often occurs over-the-counter: each institution only trades with a subset of potential counter-parties fixed in the short run. For every intermediary whether the particular intermediaries it trades with be active is very important in this environment. This creates the potential for a different feedback loop, ceasing intermediation by the agents make others dependent on them for market access unwilling to lend, and this propagates through the network.

The authors investigate to what extent and how liquidity deteriorates in this way after the exogenous withdrawal of a subset of intermediaries. One of the main take-aways is that a financial framework in which they trade both repo and collateral over-the-counter is less resilient to such an exogenous shock than a system in which they trade collateral on an exchange. The authors development a model of intermediaries, banks for short, that interact in two over-the-counter markets. They study a coupled-network game in which banks be active or not, and their decisions are strategic complements. The authors show that solving the network game can be reduced to finding connected subnetworks comprising sets of banks which have at least one incoming directed link in both repo and collateral networks after the shock has disabled the banks. They investigate the effect of the exogenous shock and show an "illiquidity spiral" such that more banks withdraw from the collateral and repo markets after a first exogenous shock. Existence of these fragile connectors makes equilibrium liquidity fragile. In particular, the withdrawal of such a fragile connector can lead to sudden market freeze. In the absence of fragile connectors, the complementary nature of repo and collateral markets can amplify exogenous shocks and lead to illiquidity spirals. Replacing at least one OTC market by a centralized exchange reduces the extent of illiquidity spirals and improves the resilience of liquidity.

One discussant asked how the shock enters the model. The author replied that the shock is on one bank, and then it transmits through the network to the other banks. He mentioned that the reason for having two layers of networks in the model is to show that the links in different layers have a different nature. The author agreed with one discussant that if the market is complete in the model, there would be no externality.

Platform Trading with an OTC Market Fringe

Jerome Dugast, **Semih Uslu** and Pierre-Olivier Weill

Many financial instruments trade over-the-counter, that is, decentralized security markets with bilateral bargaining. A few examples are credit-default swaps, interest rate swaps, and asset-backed securities. Regulators have mandated centralization. For example, in 2009, G20 leaders agreed that standardized OTC derivative contracts should be traded on exchanges or electronic trading platforms. Since market participants showed a preference for OTC markets, a nature question is whether mandatory centralization improves social welfare.

The authors study the privately and socially optimal participation of investors in a centralized platform or in an over-the-counter (OTC) market. Investors incur costs trading in the platform, in the OTC market, or in both at the same time. Investors differ from each other in risk-sharing needs and OTC market trading capacities. One of the main results concerns the patterns of trade and participation. Banks with low risk-sharing need and high capacity go to OTC market only. Among them, low risk-sharing need and high capacity banks are intermediaries while high risk-sharing need and low capacity banks are customers. The banks with high risk-sharing need and low capacity choose Walrasian market only.

Whether mandatory centralization improves welfare depends on the banks' feature. The authors give two necessary conditions for customers' private incentives to be large relative to their social contribution. Mandating or subsidizing trade in a centralized venue can be welfare improving only if it satisfies these conditions. First, investors must differ in terms of OTC trading capacities. Second, participation costs must induce exclusive participation decisions. Based on the empirical trading patterns generated by closed-form examples of our model, they argue that the real-world OTC markets might satisfy the conditions under which mandating or subsidizing centralized trade is welfare improving.

A question was asked about defining capacity in the context. The author gave an example that capacity could be the collateral used in the OTC market. He added that capacity should be a fundamental characteristic of a bank. One discussant wondered what happens if a bank took part in both markets. The author noted the bank would choose a volume so that its exposure to risk are equal in both markets. He mentioned that only banks with high risk-sharing need and low capacity acted that way, and they become pure intermediary in the OTC market.

Mechanism Selection and Trade Formation on Swap Execution Facilities

Esen Onur, David Reifeen, Lynn Riggs, and **Haoxiang Zhu**

The Dodd-Frank Act includes several mandatory steps to make the OTC markets more transparent. The authors study swap execution facilities (SEFs) trading mechanisms and the associated behavior of market participants after Dodd-Frank. For this purpose, they analyze joint message-level data for index CDS traded on Bloomberg SEF and Tradeweb SEF in May 2016, on which customers have three options to execute their trades. They are all-to-all central limit order book (CLOB), 1-to-n request for quote (RFQ), and 1-to-1 request for streaming (RFS).

The authors see that both clients and dealers don't want to maximize competition on SEFs. The data show that customers expose their orders to few dealers, and trade with even fewer dealers. The data also show that dealers do not always respond to customers' inquiries, and they trade with a small fraction of customers on average. As a baseline model, the authors consider a winner's curse. In the first dealer-to-customer stage of their model, a customer selects multiple dealers in an RFQ, and dealers either respond with quotes or not. In the second interdealer stage, dealers trade in sequential double auctions. The winner's curse happens since each dealer in an RFQ is less likely to bid a high price for the customer's order to expect a lower interdealer price caused by higher total inventory. As an extension, the authors use the customer-dealer relationship. The model predicts positive relationships between a dealer's probability of responding to the RFQ and the number of active dealers and the gain from trade between the customer and dealers. It predicts negative relationships between the probability and the number of dealers in the RFQ, and the cost of holding inventory.

The authors find that a larger trade size makes the customers choose RFQ less than RFS, and, if the customer chooses RFQ, decreases the number of dealers queried in the RFQ. Besides, customers are more likely to send RFQs to their clearing dealers or to dealers with whom they have traded more in the past. The dealers' strategic responses to RFQs, the authors find that a dealer responds less to an RFQ as the number of dealers selected increases, notional quantity decreases, and the number of quotes available before the customer makes the order decreases. Having a clear relationship with the customer increases the dealer's response probability, but a higher trading volume with the customer in the past does not. With regard to dealers' pricing behavior, a higher notional quantity slightly increases dealers' quoted spreads. The clearing relationship reduces the quoted spread slightly only for investment grade contracts.

Several audience members asked about the properties of a transaction in SEFs such as observability, anonymity, and the standardized quantity of quotes. One concerned was whether every dealer has to trade in the interdealer stage. Another audience member was curious how quickly a dealer has to respond with a quote. Another audience member pointed out that some dealers systematically show patterns to respond less frequently than others. Finally, one audience member asked why it is better to have only five dealers regarding clients' gains from trade.

Sentiment, Liquidity and Asset Prices

Vladimir Asriyan, Brett S. Green and **William Fuchs**

Fuchs et al. develop a model where asymmetric information and agent sentiment result in frictions that allow for the existence of multiple equilibria.

The model comprises agents trading high value and low value assets based on the expected gains from asset productivity and the future resale value. The agents observe their own asset type but cannot determine the asset types of others. Each period, the assets each receive an idiosyncratic productivity shock and agents can buy or sell. Buyers decide their bid price based on the expected asset productivity and future resale value of the asset which both depend on the proportion of high value assets in the market.

If the proportion of high value assets is high, the buyers' willingness to pay lies above the reservation price of a shocked owner of a high value asset. Here, all shocked asset owners trade, asset allocation is efficient and asset values are high. If the proportion of high value assets is low, the buyer's willingness to pay will be below the reservation price of shocked owners of high value assets, and only low value assets trade in equilibrium. This allocation is inefficient and asset values remain low.

These thresholds are such that a range of proportions of high value assets that allow for multiple equilibria based on a buyers' beliefs about the quality of the assets available for purchase. Binary aggregate productivity shocks, or sunspots, with some persistence parameter shock the market, changing agents' sentiments. If there is perfect persistence of these sunspots, an efficient or inefficient constant price equilibrium results. If there is imperfect persistence, then multiple price equilibria can arise, provided the sunspot is persistent. Asset prices reflect fundamentals and no aggregate uncertainty, but asset prices, output, and welfare can fluctuate because of the sunspots.

During the presentation, a participant questioned whether agent heterogeneity influences the equilibrium outcome. For example, if the agents were better at selling, then this heterogeneity could affect the main predictions of the model. Fuchs replied that heterogeneity can be thought of as being incorporated into the idiosyncratic productivity shock, so agent heterogeneity should not change the results. Another participant suggested that contracting could overcome this problem of asymmetric information and agent sentiment. The presenter replied by saying that the model can be thought of as the residual after any contracting has been used to mitigate problems arising from imperfect information. A participant asked what the model predicted in the absence of asymmetric information. Fuchs answered that there would be a two-price equilibrium that would result in unshocked owners holding assets. A participant pointed out that leasing of assets might avoid problems arising from asymmetric information. The author confirmed this, but that leasing was rules out in their model.

Turnover Liquidity and the Transmission of Monetary Policy

Ricardo Lagos and Shengxing Zhang
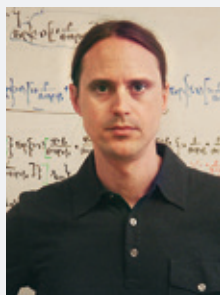
Monetary policy is used in most modern economies as an instrument to influence asset prices and the rest of the economy. Lagos and Zhang describe a novel mechanism of monetary policy that affects financial markets and equity markets in particular. They build a search-theoretic model of money and trade in over-the-counter markets to help explain the turnover liquidity channel of monetary policy. In their model, tight money used for payments, increases the opportunity cost of holding liquid assets. The implied scarcity of these instruments then reduces the turnover and re-salability of financial assets, which leads to a reduction in real asset prices.

Their first contribution is to construct a model of over-the-counter markets in which money is the medium of exchange between investors and dealers. Investors have stochastic and heterogeneous private marginal values for dividends. Investors take time to find a dealer. The existence of these frictions makes the shadow stock net present value higher than any investors' marginal net present value for holding an asset. Investors then use money to transact with dealers, which depreciates with inflation. Assets' prices are solved for using Nash bargaining. In equilibrium, asset prices and liquidity measures such as the spread, the trade volume, and the dealer supply immediacy are determined by the amount of money, market power of dealers and

the matching technology between investors and counter-parties. As a consequence of search and market frictions, asset prices exhibit a speculative premium or "bubble." An expected higher opportunity cost of holding money that reduces real balances and makes low valuation investors hold too many assets, and, therefore, can distort asset allocation diminishing stock's real returns.

The second contribution of this paper is documenting the response of real stock prices and turnover due to a tightening in monetary policy. Their sample includes between 1,300 and 1,800 stocks from the New York Stock Exchange in 2,014 trading days from 1994 to 2001. They use a heteroscedasticity-based approach similar to that of Rigobón and Sack (2004 to achieve identification of the effect of monetary policy on stock prices and turnover. A second approach is to use the change in the policy rate in a 30-minute window around the Federal Open Market Committee (FOMC) announcement as a high-frequency instrument. They use vector auto-regressions to estimate the persistence and dynamics between the nominal interest rate and stock prices. The authors conclude that a 25 basis points surprise increase in the policy rate reduces stock return between 1% and 2%, and it reduce the turnover rate between 17% and 30%. There is considerable heterogeneity in the response of the return and turnover rate across stocks with different liquidities. The return of the top 5% most liquid stocks falls two times more relative to median-liquidity stocks. They see a similar pattern for the turnover rate.

In a third contribution, the authors

calibrate the model to test how the theory can match the observed effects of monetary policy on asset prices in the short run. For instance, an unexpected increase in the policy rate of 1 basis point causes a decrease of between 4 and eight basis points in the stock market return on policy announcement days. They also associate unexpected increases in the opportunity cost of money with substantial and persistent declines in stock turnover and its effect on assets. A sudden tightening in the monetary policy generates a reduction in the return of a stock in the 95th percentile of turnover rates around twelve times larger than an asset in the 5th percentile on announcement days.

The discussant, Burton Hollifield, mentioned that this approach is a smart, new approximation of the role of money as a payment medium. He asserted that if money depreciates, then money crowds out the efficient allocation. Similarly, surprise inflation decreases stock return through lower turnover. He asked what would happen if they extend the analysis to other securities including corporate bonds, derivatives, currencies, etc. Is there a way to reconcile these results with a Harrison and Kreps mechanism? Is this model useful to understand the ex-ante premium and the volatility on meeting days? What if the investors are risk-averse? A member of the audience wondered why the equity premium is so high before monetary policy announcements. The speaker mentioned that is difficult to construct measures of beliefs for high-frequency variables.

Brokers and Order Flow Leakage: Evidence from Fire Sales

Andrea Barbon, Francesco Franzoni , Augustin Landier, and **Marco Di Maggio**

Barbon, Di Maggio, Franzoni, and Landier offer empirical evidence that massive institutional liquidations are predated upon by the best clients of aware brokers, which has a strong theoretical support in the literature. Predatory trading in sales happens when market participants anticipate other investor's orders and trade in the same direction. As a result, prices further decrease relative to the contra factual of no anticipation, which negatively affects the liquidating investors. Even though institutional investors split orders into smaller amounts over time to avoid market impact, such an event can take place when traders pass information to other traders. They might have incentives to do so to form a reputation for spreading the news. Alternatively, they might want to give liquidity to their long-term clients convincing other traders to be a counterpart of their sales.

The authors use trade data collected from 1999 to 2004 by Abel Noser Solutions, or formerly known as Ancerno Ltd -a firm that performs cost analyses for institutional investors. They argue that their sample represents volume and comparable to administrative data. The authors claim there is no self-reporting bias, survivorship and backfill biases. Earlier empirical research shows that this database includes stocks comparable to those in 13F mandatory filings with a volume traded in dollars of about 20% of that of 13F filings. Similarly, fund managers in the database are skilled in trading.

The researchers identify 400 episodes of massive institutional liquidations or fire sales as they call them. They define such transactions as five consecutive days of selling where the manager's Z-score falls below -0.25. They define stocks in fire sales as those where the volume ratio from the Center for Research in Security Prices (CRSP) is below a threshold. Sales must have a liquidity motive characterized by the managers performing such trades not buying back the same stocks, and by those stock prices not being affected permanently. In their sample, a distressed fund liquidates $380 million on average per fire sale. They take between 5 and 11 days to complete, and it involves 22 stocks on average. The volume of these funds is on average 15% of the CRSP volume.

The authors hypothesize and empirically test whether the best clients of aware brokers are more likely to engage in predatory behavior. They label aware brokers at the event, stock and day level and must have a significant fraction of the overall fire sale. If brokers have incentives to spread information, then the clients with the most robust relationships will capture the highest benefits. These relationships are measured by volume share, revenue share, and dependence of the manager on the broker. The authors regress a dummy of whether a broker's client trades in the same direction as the liquidating manager on whether the broker is aware. They find that their awareness measure is associated with an increase of 10% in the likelihood of predatory trading and positively associated with the volume of predatory shares. In a similar exercise, they find that the best-connected clients are 2%-3% more likely to sell their holdings of the fire-sale stock with the same broker during the first five days of the

liquidation period. They performed robustness checks for periods with no adverse events, bad news or negative momentum that might coincidentally increase the liquidation of stocks.

The authors also exploit a natural experiment that caused managers to sell their stocks. Specifically, in September 2003, the New York Attorney General announced the issuance of a complaint due to illegal practices by hedge and mutual funds managers. These made 25 fund families sell their holdings. The authors find that clients of the brokers who helped to liquidate those stocks were more likely to have taken a similar selling strategy.

Slava Fos, the talk's discussant, suggested including the interaction of broker and client fixed effects, the interaction of broker and manager fixed effects and the interaction of client and manager fixed effects in their baseline regressions. The discussant is skeptical of current claims of the paper including the amplification of price fluctuations mediated by brokers, and the importance for brokers to build a reputation and what mechanism to do so is more valuable for them. He pointed out that the paper relies on too many ad hoc assumptions which makes it less generalizable such as how to define a liquidation time window. Likewise, the sample is limited to brokers in mutual funds. He raised other questions such as what happens to other less connected clients and how the best-connected clients' strategies impact prices. Finally, he questioned the paper's contributions given a very similar paper in the literature by Marco Di Maggio and coauthors, entitled, "The Relevance of Broker Networks for Information Diffusion in the Stock Market."

The New Frontier in Business Cycle Research

March 2018

Zachary Bethune – University of Virginia
Vasco Carvalho – University of Cambridge
Tom Cooley – New York University
Joel David – University of Southern California
Emmanuel Farhi – Harvard University
Jesus Fernandez-Villaverde – University of Pennsylvania
Fatih Guvenen – University of Minnesota
Finn Kydland – University of California, Santa Barbara
Iourii Manovskii – University of Pennsylvania
Ellen McGrattan – University of Minnesota
Lee E. Ohanian – University of California, Los Angeles
Sophie Osotimehin – University of Virginia
Peter Rupert – University of California, Santa Barbara
Lukas Schmid – Duke University
Maxim Troshkin – University of Virginia
Guillaume Vandenbroucke – Federal Reserve Bank of St. Louis
Stephen Williamson – University of Western Ontario
Eric Young – University of Virginia

A Few Remarks

Tom Cooley and Lee Ohanian
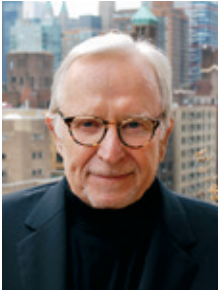
Modern business cycle theory has been among one of the most productive areas of economic research over the last 40 years. Originating from shocks that affect the economy, business cycle theory brings together the neoclassical growth framework and economic fluctuations associated with business cycles. The recent breakthroughs in dynamic economic theory and computational methods are covered in the book Frontiers of Business Cycle Research, edited by Thomas F. Cooley. During his talk, Dr. Cooley presented some of his latest joint work with Dr. Ohanian in macroeconomics and business cycle theory.

Although suggested by Lucas that in representative agent models welfare consequences of business cycle were likely to be small, Dr. Cooley argued that those issues should never be overlooked. A common first step in business cycles research reported in the book was to remove the trend, but this can be problematic for a variety of reasons. Not only did it leave behind important data, but it led to fruitless discussions about detrending, band-pass filters, co-integrations, unit roots, etc. More importantly, what was left out sometimes has significant welfare implications. So understanding how policies affect long-run aggregate performance has aroused great interest in studying economic transformation and stagnation.

To illustrate his perspective, Dr. Cooley discussed the post-World War II Western European economic expansion. In particular, why did growth in the EU and convergence with the US stall in the late 1970s? For instance, Italy was catching up to the US in GDP per capita until the mid-1970s. Since then, Italy has been experiencing negative TFP growth. They can even find similar catch-up patterns within Italy itself. Southern Italy's convergence in productivity to northern Italy stopped after the mid-1970s. The presenter made several speculations about the phenomenon. To begin with, a sharp growth in government spending on social welfare programs targeting the south created disincentives to work. The price level was much lower in the south, hence migration from the south to the north decreased after the 1970s. In addition, organized crime (much more prevalent in the south) imposed an alternative form of taxes. All the above hindered the catch-up process, despite the south continuing to catch up the north in labor productivity.

One audience member expressed concern about measurement issues with the unemployment data. The presenter acknowledged that unmeasured economic activities were larger in the south, but the source of distortions remained unclear. Another audience member wondered how Italy's example might resemble other European countries, such as Spain and France. Dr. Cooley compared tax collected as a fraction of the GDP among different nations to show how non-market allocations distort outcomes.

Intangible Capital and Measured Productivity Over the Business Cycle

Ellen McGrattan

Measured GDP does not account for all intangible investments, even though many firms invest in intangible assets at a similar rate to tangible assets, which changes measured GDP understate the real changes in GDP. The author maintains that understated changes in measured GDP leads to the observation of larger changes in hours worked and investment and little change in measured Total Factor Productivity (TFP). Using a dynamic multi-sector general equilibrium model and revised U.S. national accounts, which include expenditures on knowledge-related products as a part of investment, the author shows how shocks and industry correlation explain business cycle fluctuations better.

The author uses multiple sectors in the model to consider that there is a tremendous difference in intangible investment rates across industries. Firms in the model produce both tangible and intangible goods and services. The author allows industry correlation through purchases of both intermediate inputs and tangible or intangible investment goods. Business cycle fluctuations in the model are caused by shocks to industry and aggregate TFP. Since it is impossible to measure unobservable intangible capital stocks, the author estimates stochastic processes for the latent TFP using Maximum Likelihood Estimation (MLE) and Kalman filter.

The author finds that the model's prediction on fluctuations in total U.S. working hours and co-movements

of sectoral working hours surpasses the prediction from the one-sector model without intangible investments. The author's model recognizes three sizable booms over the 1985-2014 sample period and the predicted hours worked from the model is 65 percent of the actual data, the other predicts just 9 percent. After verifying the model's validity, the author derives MLE parameter estimates and time paths for latent TFP shocks and intangible investments by decomposition of the variances of U.S. data. The author finds that sector-specific shocks and industry correlation are important to account for business cycles. The author also finds that

intangible investments vary less over the business cycle than tangible investments and lag the cycle. As an extension, the author adds financial market disturbances in the same model. Between debt and equity, as a source of external finance, firms prefer debt to equity due to the tax benefit, but they face borrowing constraints subject to stochastic financing shocks. After constructing time series for ratios of tangible capital to output and debt to output, the author finds that unexplained labor wedges are smaller and less volatile than the wedge in Jermann and Quadrini's (2012).

Several audience members asked whether there is no separation

between durable goods and non-durable goods in the model and no depreciation issue. One audience member expressed concern whether the author considers a possibility that many R&Ds have internationally cooperated. Another audience member asked whether marginal productivity of labor is still different from the marginal rate of substitution when the author incorporates intangible capital in the model. Finally, the audience members were curious how to construct intangible goods and services shares and whether net debt includes global debt.

Financial Frictions and the Wealth Distribution

Jesus Fernandez-Villaverde

Before the Great Recession, economists have observed both an increase in credit and leverage and an increase in wealth inequality over decades. The authors analyze how much these two trends are related and which one affected more the subsequent slow output recovery. For these purposes, the authors use a continuous-time method which allows them to derive analytical equilibrium dynamics and to care about local derivatives only. Besides these benefits, the nonlinearity of the Perceived Low of Motion (PLM) followed by the continuous-time method makes the authors explore how aggregate shock affects the wealth distribution and how different wealth distributions change the way of aggregate shocks to affect the economy.

The authors use a continuous time stochastic general equilibrium model. There are three infinitely lived agents in their model: a representative expert, a representative firm, and heterogeneous households in their wealth and labor supplies. There are two types of assets: a risky asset and a risk-free one. A representative expert is an only agent to hold the risky assets. To finance risky assets, the expert issues bonds to households. The financial frictions in the model come from the fact that the expert has no source of insurance against the aggregate capital shock. It is assumed that the economy has a risky asset in positive net supply and a risk-free one in zero net supply.

Regarding the problem of an infinite dimension in households' income-wealth distribution, unlike Krusell and Smith (1998), the authors face the nonlinearities of the endogenous state variables. This leads the authors to use a neural network, one of the machine learning techniques, which can

estimate parameters in their model. The authors show the distribution of forecast errors from the neural network is not skewed, and it has a lower variance compared to the linear PLM.

The authors find that the Stochastic Steady State (SSS) has more debt, leverage, and wealth inequality than the deterministic steady state, since households save more in the world of uncertainty, and then, the economy has more wealthy households. From a Monte Carlo simulation, the authors find that the economy stays around the SSS, but it sometimes visits a high-leverage and a low-leverage region. The authors find that the recession is more persistent if the economy starts from the high-leverage region by computing the generalized impulse response functions and the distributional response function to a negative shock. This happens because wealthy households reduce consumption less in a high-leverage economy than in the SSS, which makes

the capital accumulation happen slowly, then produces a slow output recovery.

Several audience members were curious about the meaning of skewed capital around the SSS. One audience member asked whether there is an empirical analogy of the representative expert. Another audience member asked why the representative expert and households have different preferences. From the fact that the net wealth of the expert may be negative, there was an audience member to ask whether it is bounded below. Some audience members asked how much heterogeneity matters in the result of forecasting errors and whether the results in this paper could be done in discrete times.

Use it or Lose it: Efficiency Gains From Wealth Taxation

Daphne Chen, **Fatih Guvenen**, Gueorgui Kambourov, Burhan Kuruscu, and Sergio Ocampo-Diaz

The authors revisit the question of optimal capital taxation in an environment with two empirically motivated features. First, because wealth holdings are concentrated in the United States, a small fraction of the population pays most of the capital taxes. Second, it's plausible to conjecture that the mechanism by which it generates this concentration matters for the analysis of capital taxation. Despite the fast-growing literature on models with return heterogeneity, the implications of these models for capital taxation have not been quantitatively studied and thus are not well understood. To fill the gap, the authors analyze how taxing the income flow from capital differs from taxing the stock of wealth. One of the main conclusion of the analysis is that in the presence of return heterogeneity the two tax systems have a range of implications different from, and sometimes, the opposite of each other.

The authors study an overlapping generations economy inhabited by individuals who derive utility from consumption and leisure. A key ingredient of the model is a persistent heterogeneity in investment/ entrepreneurial skills, which, together with incomplete financial markets that prevents free flow of funds across agents, allows some individuals to earn higher returns on their wealth than others.

The model analysis produces three sets of results. First, the authors begin with a revenue-neutral tax reform that replaces the current US tax system of capital income taxation with a flat wealth tax, keeping taxes on labor and consumption unchanged. This reform raises average welfare significantly, and it disperses the welfare gains across the population.

Second, the authors move to an optimal tax analysis, in which a utilitarian government chooses linear taxes on labor income and on wealth to maximize the ex ante expected lifetime utility of a newborn. The authors repeat the same analysis, this time the government chose linear taxes on labor and capital income, and compare the implications of each optimal tax system to each other as well as to the current US benchmark. The main result from the first experiment is that a positive tax on the stock of wealth is optimal which reduces the more distorting labor income taxes. In the second experiment, the authors find that a negative tax on capital income is optimal, and the rate is high. This contrasts with the well-known results in similar life cycle models with incomplete markets where a large and positive tax rate on capital income was found to be optimal.

Third, the authors find that among the two optimal tax systems, the one with wealth taxes yields higher welfare than the one with capital income taxes. A decomposition analysis shows that the gains under wealth taxes come from both a rise in the level of consumption and a decline in the inequality of consumption and leisure. Thus, optimal wealth taxes yield both first- and second-order gains. Subsidies on capital, together with the small rise in consumption levels, leads to higher inequality yielding distributional losses, which offsets gains from levels, unlike under optimal wealth taxes.

One discussant asked the author how is the return on the same asset different. The author responded that the return in the paper is from a household perspective rather than focusing on a particular asset. Another discussant asked if the household holds just a single portfolio, would there be any efficiency loss? The author replied that there is still inefficiency in that situation. The author also noted that the misallocation implied by the model calibration is below the empirically estimated ones.

The Impact of Unemployment Benefit Extensions on Employment: the 2014 Employment Miracle?

Marcus Hagedorn, **Iourii Manovskii** and Kurt Mitman

The aim of the paper is to access the impacts of unemployment benefit extensions on the labor force and employment. Measuring the magnitude of these effects is important for understanding the economic consequences of this widely used policy instrument. Yet existing literature provides little information on the size or sign of these effects. Empirical micro literature has focused almost only on measuring the effects of benefit eligibility on the search effort of unemployed workers, which is too narrow to infer the total impact of benefit duration on employment. Estimates in the quantitative macro literature differ depending on the value of the parameters difficult to identity. The literature ignores the effect of policies on participation decisions of those out-of-the-labor force, which limits their ability to measure the total effect on employment.

The authors propose to sidestep these difficulties by measuring the employment and labor force impacts of a large nationwide cut in benefit duration in December 2013. The attractive feature of this quasi-natural experiment is that they can measure its effects using standard empirical techniques that do not require imposing assumptions of a particular labor market model on the data. The authors measure the impact of the December 2013 decision by Congress to end the Emergency Unemployment Compensation Act of 2008 which lowered benefit duration in all states to their regular duration of 26 weeks. The decision to eliminate benefit extensions was controversial, but the U.S. labor market performance in 2014 surprises many observers. Average employment growth was about 25% higher in 2014 than in the best of several preceding years.

To estimate the extent to which the elimination of unemployment benefit, extensions induced the aggregate labor market developments, the authors exploit the vast heterogeneity of the decline in benefit durations across states to identify the labor market implication of unemployment benefit extensions. They find that employment growth was much higher in 2014 in the border counties that experienced a larger decline in benefit durations relative to the adjacent counties. What makes this finding even more striking is that year after year before 2014 the relative employment was lower in the high benefit counties. The authors then use several empirical specifications to estimate the effect. All specifications imply that changes in unemployment benefit duration have a large and statistically significant effect on employment: a 1 percent drop in benefit duration increases employment by 0.0144 to 0.0233 log points across specifications. They find that more than half of the increase in employment attributed to the cut in benefits was due to an increase in the labor force. Analysis thus implies that not only did the unemployed not drop out of the labor force, but instead those previously not taking part in the labor market entered the labor force. The authors estimate that changes in unemployment benefits have a large and significant effect on job counts: a 1 percent drop in benefit duration increases the number of jobs by 0.010 to 0.0236 log points. Point estimates for the increase in the number of jobs is smaller than the estimate of the effect on employment, showing that a cut in benefits leads to an increase in the number of full-time jobs at the expense of part-time jobs.

One discussant asked whether benefit duration depends on the time of entry into unemployment. The author noted that it depends on the timing, but they guarantee 26 weeks of unemployment benefit. Another discussant asked whether many people moved to states with higher unemployment benefits. The author replied that they only observe a few such instances. One discussant asked whether we should focus the analysis on the minimum wage workers. The author noted that since unemployment benefit affects every worker in the economy via its effect on vacancy creation; they do not restrict attention to a subgroup of workers.

Fertility Shocks and Equilibrium Marriage-Rate Dynamics: Lessons from World War 1 in France

John Knowles and **Guillaume Vandenbroucke**

Theory suggests that marriage probabilities of single individuals depend crucially on the sex ratio among them. For instance, it expects a low male-to-female sex ratio to generate low marriage probabilities for single women. However, the authors document a somewhat contradictory pattern: marriage probabilities of single women (or female marriage hazard rates) were 50% higher in France after World War I (WWI 1914-1918) than before the war, even though the male-to-female ratio among singles was 33% lower. Marriage hazards remained fairly high until the mid-1920s. To understand this clear puzzle, a dynamic model of two-sided matching with stochastic aging is proposed. This competitive-search framework, à la Shimer (2005) is suitable for the analysis because it treats marriage matching as a dynamic process, and hence the computation of transition paths for marriage rates.

The equilibrium marriage model features infinitely lived men and women whose propensity to marry reflects optimal choices about when and whom to marry. It assumes that heterogeneity in the choices arises from traits that vary both within and across age-sex sub-populations of singles. Composition of any sub-population defined by observed variables, such as age, determines that group's marriage history. The key assumption is then heterogeneity in age because this dictates how individuals value single life. This means that in the model younger men avoid marriage while older are more willing to do so.

Given that a birth-rate decline that was both large and sudden, the model accounts for the potential dependence of marriage incentives on fertility intentions accompanied the war. It assumes that the benefits of marriage are the increasing numbers of children. At this point a participant expressed the concern of whether the sex ratio is important in the model, given the assumptions. The author mentioned that the model keeps many features of that using constant returns to scale, the sex ratio is important but will not be the only feature that affects marriage probabilities.

The authors calibrate the steady state of the model to pre-war French data (1900-1913) on marriage hazards by age and sex, and to the pre-war marital birth hazards by age. At this point participants discussed the appropriateness of using French data during and after World War I. First, the authors abstract from widows, divorce and remarriage because of the potential differences with singles. Second, most of the French men mobilized but did not fight in trenches, meaning they could get married during the war. Finally, another shock such as World War II is not a good candidate to study because economic conditions pre-war were not normal given the great depression in 1929. Many expected the war, and the bust in fertility during WWI changed the demographic composition of the 1930s.

They use the calibrated model to explore the dynamics of marriage hazards. It represents the war as a pair of temporary, unanticipated shocks of known duration: higher mortality for younger men, and weaker fertility incentives for married couples. Weaker fertility incentives mean that creation of babies become more costly exogenously and this reduces the value of married life. In response to these shocks, equilibrium hazard rates vary over time, driven by changes in both the current composition of the singles pool and in expectations about the future composition. They set the size of the wartime shocks so that the calibrated model matches the average wartime values for the aggregate birth rates and the mortality rate of men. Transition back to the steady state is then computed.

The model's transition path replicates to a large extent the post-war patterns of marriage hazards. A war-time marriage-rate decline that is roughly equal in size to that observed in the French data, a decline in the male-to-female ratio among singles that matches the empirical value as of 1920, and a large post-war boom in female marriage probabilities; about 63% of the observed peak for women aged 20-39. They then use this model to measure the relative importance of the two components of the war shock. The composition shock that results from the wartime marriage-rate decline implies an unusually high post-war abundance, relative to the stock of single women, of men with a relatively high propensity to marry. This effect turns out to be so strong that it outweighs the negative effects of the male mortality shock, via the decline in the male-to-female ratio, on the female marriage hazard.

Finally they presented additional counterfactual experiments. These allow the authors to conclude that the male-mortality shock alone would have yielded a 23% decline in women's marriage hazard after the war, as the standard model predicts. In the absence of the male mortality shock, the composition effect would have generated a 39% post-war increase in female marriage hazards, 33% higher than in the baseline experiment. The author closed the presentation invoking the need to understand the sex ratio and what causes a decline, but understanding its effect on the marriage market it may not be a sufficient statistic.

Low Real Interest Rates and the Zero Lower Bound

Stephen D. Williamson

How should monetary policy be conducted in the presence of the zero lower bound (ZLB), i.e., a liquidity trap? If the ZLB is optimal, is there any room for forward guidance to improve upon distortions? Earlier literature has analyzed the role of quantitative easing and large central banks balance sheets. This paper studies the consequences of further modeling the source of low nominal interest rates, and why the source of liquidity matters for monetary policy and forward guidance. This model does not consider issues related to negative nominal interest rates. Thus, the lower bound on the nominal interest rate is zero. In a world with liquidity premia on assets other than money, monetary policy might not achieve efficiency because of the presence of the ZLB.

The author proposes a standard New Keynesian model to discuss these issues for monetary policy. The model includes a Phillips curve tradeoff and sticky prices inefficiency. Zero inflation is optimal as long as the nominal interest rate policy is not constraint by the ZLB. The real interest rate is exogenous because of quasi-linear preferences in consumption and labor supply. Then the model shows Fisherian attributes; in particular, expected inflation is determined by current nominal interest rates. This is a feature shared among New Keynesian models.

The author develops a first basic model with sticky prices in a cashless economy à la Woodford with no role for assets. This model is like earlier literature where low real interest rates are connected to temporary reductions in the time preference. Then, in a second source of distortions, he assumes that they conduct transactions using secured credit. The only available collateral is government debt. If the real quantity of this collateral is small enough, the collateral constraint binds and the real interest rate is low. A liquidity trap problem arises as in the baseline model, but sticky prices and binding collateral constraints emerge as the two sources of inefficiency. Not only is there an inefficiency in the sticky-price goods market because of the sticky-price distortion, but inefficiency in the market for flexible-price goods generate the binding collateral constraint.

A third model nests the other two models and includes scarcity of cash. Under the presence of sticky prices, open market operations are not neutral, i.e., purchases of government bonds will lower the real interest rate permanently by tightening the collateral constraint. They do not maintain these results without sticky prices. In particular, when the collateral constraint does not bind, the nominal interest rate must be zero as in a Friedman rule. When the real interest rate is low and the collateral constraint binds, the nominal interest rate should be positive. The central bank faces a tradeoff between two sources of liquidity scarcity. An open market sale raises the nominal interest rate, which makes cash scarcer, but also reduces the safe collateral problem.

The model breaks down into multiple equilibria when the source of the low real interest rate is further modeled. A first equilibrium shows that a high inflation rate, a tighter collateral constraint, and a lower real interest rate are compatible. A second equilibrium includes a lower inflation rate, a less tight collateral constraint, and an interest rate that is not low. Not explicitly modeling the liquidity trap state thus result in different conclusions from those of previous literature that assumes a high discount factor. In other words, New Keynesian models do not work as advertised. Minor changes in seemingly innocuous assumptions in these models will change results dramatically.

The audience asked why the relative price of sticky-price to flexible-price goods in the model is so important for the central bank rather than inflation. The speaker explained that this variable captures the degree of price stickiness within the model, and that inflation is computed from this and other key endogenous variables. An audience member asked what would happen if central banks analyzed simulated data with these features. The speaker commented that the use of New Keynesian models for monetary policy might produce bad advice when they are taken into the data and used to analyze central bank policy. Particularly, movements in output, inflation, or interest rates could be under those models.

Bottom-up Markup Fluctuations

Ariel Burstein, **Vasco M. Carvalho**, Basile Grassi

The traditional view of imperfect competition in macroeconomics is modeled by monopolistic competition, the representative firm and sticky prices. This paper investigates if firm-specific productivity shocks drive aggregate markup fluctuations. In this framework, heterogeneous firms operate under Bertrand or Cournot oligopolistic competition with flexible prices. This paper constitutes a bottom-up framework that starts at the firm level and predicts markup fluctuations at the firm, sector and aggregate levels. The authors find that the largest firm is pro-cyclical regarding sector output, where the opposite is true for the smallest firm.

The model has three nested levels of aggregation; the firm, the sector and the economy levels. A double-nested CES demand system parameterizes preferences. There is a finite number of firms in each sector operating a CRS technology with heterogeneous productivity and with labor as the only input. Productivity follows a Markov chain over a discrete state space. The authors assume that firms consider the impact of their decisions on a sector's price and quantity but not on other sectors' or aggregate prices and quantities. This generates optimal pricing for the firm equal to a markup over its marginal cost. The markup is an increasing function of the firm's sales share. They pin the distribution of the sales shares in a sector down by the distribution of the productivity shock, which is the model's key state variable. The solution is not linear so the authors resort to a second-order approximation to the firm's pricing problem.

Then the authors build from the firm's decision to the sector and aggregate variables. For instance, they compute sector markup as sector price divided by sector marginal cost, while they calculate sector productivity as the sector price divided by sector wage. Specifically, the sector markup is a harmonic average of firms' markups weighted by their sales shares, while the sector productivity is the harmonic mean of firms' productivities weighted by their share in the sector output. The joint distribution of firms' productivities and sales shares then characterizes the sector markup and sector productivity. Sector markup determines the sector income distribution. The higher concentration in a given sector leads to a higher sector markup regardless of whether sector competition is Bertrand or Cournot. They compute aggregates given sectors' markup and productivities. They calculate wage, labor shares, markup, output, hours and productivity.

The authors use a second order approximation of the firms' pricing problem to study how idiosyncratic shocks to productivities affect firms, sectors, and aggregates. The mapping between productivities and sales shares plays an important role in such evaluation, but it has no closed-form solution. This approximation reduces the problem to track two sufficient statistics the sector average productivity and the sector productivity concentration. The assumption on the finite number of firms will affect this sector statistics. In the approximation, both the sector's markup and productivity are increasing functions of the concentration index. An increase in the productivity of a large firm causes a rise in the sector-level markup–the opposite is true for a small firm. Similarly, a positive shock on a firm's

productivity generates a lower sector-level price, which strengthens if the firm is small.

The relationship between markup and output is ambiguous at the sector and aggregate levels. Sector markups can go down while output can go up if the shock hits a follower–the opposite is valid if the leader or largest firm drives the shock. The aggregate-level markup cyclicality is ambiguous which depends on the market structure of sectors leading the aggregate movement and reallocation effects. They define the reallocation effect as the change in sector sales shares due to a change in relative prices between sectors. To sum up, signing markup cyclicality depends on the level of aggregation, the origin of the shock and the market structure within and across all industries.

The authors, in ongoing work, calibrate the model to the French manufacturing census data in 2007. They choose firm-level productivity processes to match Herfindahl indexes for all 5-digit manufacturing sectors and firm-level sales volatility. They estimate markups using a structural model of the firm's technology. A firm is defined as a company rather than an establishment, but could be described as the latter. They also match the size distribution of sectors and populate their calibrations with the same number of firms in each sector as they are in the data.

In a simulation exercise, the authors used the calibrated parameters and simulated the productivity path of each firm. Then they solve the firm problem at each period, which they do not base on an approximation but in a full nonlinear solution. They then aggregate these decisions to compute sector-level and aggregate-level variables. In ongoing work, the authors seek to match the simulated distribution of markups and sales

shares to real data. They also offer empirical evidence that firms markups are pro-cyclical regarding sector aggregates. In doing so, they use a sample of 7,824 publicly listed manufacturing firms in the US followed by Compustat, which are larger than the average firm in the population. They found that large firms' markups comove positively with sector output, while the average firm's markup comoves negatively.

The audience asked about the previous empirical evidence of the cyclicality of markups, which according to the speaker is difficult to sign because of the poor data quality. Likewise, the audience wondered how to identify the consumption elasticity of substitution between and within sectors. The speaker mentioned that it assumes them in the calibration; for instance, they choose such elasticity within a sector to be 5. The audience also asked about how to measure markup and aggregate profits, and the consequences of including the skewness of the productivity distribution. The speaker suggested that some statistics are endogenous to the model. It concerned the audience that the assumed Markov-type productivity process could not take large steps in a given period, even though it can generate a Gibrat's law in the firm-productivity distribution. Finally, the author mentioned that future work could include modeling intermediate inputs and capital accumulation.

Misallocation or Risk-adjusted Capital Allocation?

Joel David and **Lukas Schmid**

Recent studies have found that the marginal product of capital (MPK) varies across firms and industries. After accounting for the key factors such as adjustment costs and financial frictions, a significant part of the dispersion remains unclear. The authors develop a general framework, bridging misallocation across firms with systematic risks. Their theory suggests that risk exposures at the firm level help explain about 40% of total variations in MPK among US companies.

The authors investigate several important implications of the neoclassical investment theory. To begin with, exposure to common risk factors priced in asset markets is a key element in determining expected MPK. Firms with high MPK usually have high stock returns in expectation, thus higher MPK is prone to greater risk exposure. It relates changes in risk prices with cross-sectional variation in the predicted MPK. For example, credit spreads and the aggregate price/dividend ratio have successfully predicted fluctuations in MPK at the firm level. In addition, it links MPK dispersion to different market betas. Industries with high dispersion in factor exposures also have high MPK variation. Last but not least, changes in risk prices lead to fluctuations in MPK dispersion. Suggested by macroeconomic predictors, both MPK differences and stock returns comprise not only predictable but countercyclical factors. Using the US data on stock market and firm productions, the authors show their model is consistent with these empirical predictions.

The conference attendees had questions about identification issues. One participant had concerns about the robustness of the result when firms exit or merge. There was further discussion about the relationship between adjustment costs and expected excess returns.

To conclude, David, Schmid, and Zeke's paper sheds light on the sources of MPK dispersion in cross-sections. In allowing firms heterogeneous to risk exposures, the authors identify a substantial portion of MPK variations. Their results bring together aggregate shocks, risk premia, and resource allocations.

Productivity and Misallocation in General Equilibrium

David Rezza Baqaee and **Emmanuel Farhi**

This paper introduces the way to aggregate the microeconomic shock for economies with distortions. By exploiting the micro-envelope conditions, which says that all producers are cost minimizers, the authors introduce a new and interpretable decomposition of changes in aggregate TFP into "pure" changes in technology and changes in allocative efficiency. Their work provides a unified framework for analyzing the effects of distortions and misallocation equilibrium economies.

A "pure" technology effect, as the authors suggested, can be understood as, when given the initial distribution of resources, the changes in direct or indirect customers' output due to a producer's output level change. The effect of allocative efficiency comes into play when the shock shifts distribution of resources across producers. Resources reallocation may cause some producers' output increase while decrease for that of others and thus result in a change of the aggregate TFP. In efficient economies, changes in allocative efficiency are zero, while in inefficient economies, the effect is nontrivial.

Hulten (1978) shows that in efficient economies, the growth rate of GDP decomposes into the TFP growth in each sector and improvement in allocation of resources throughout the sectors. They treat Hulten's theorem as a benchmark result for the macroeconomic impact of microeconomic shocks in multi-sector models and models with production networks. The formula given by the authors coincides with the result of Hulten's for efficient economies where

the allocative inefficiency is assumed to be zero. In efficient economies, changes in allocative efficiency are zero to a first order, and so the overall effect characterized by Hulten (1978) boils down to the "pure" technology effect.

Besides theoretical contribution, the authors apply their framework by answering three different questions about the role of markups on aggregate productivity. Evidence shows that the average markups have increased over the past decades in the US. Focusing on markups as a source of distortions, the authors answer the questions: 1, "How changes in allocative efficiency contributed to measured TFP growth in the US over the past 20 years?" 2, "What are the gains from reducing markups in the US, and how have these gains changed over time?" 3, "How do markups affect the macroeconomic impact and diversification of microeconomic shocks?"

According to their result, 50% of the cumulated Solow residual is due to the improvement in allocative efficiency over the period 1997-2014. They find that in the US in 2014-2015, eliminating markups would raise aggregate TFP by about 20%. This result differs from the one given by Harberger (1954), 0.1%. The reasons for the difference are that the authors use firm-level data, where the dispersion of markups and elasticity of substitution are higher than the sectoral data used in Harberger (1954), and the Harberger's result focuses on manufacturing industry, while the authors for this paper can consider the general equilibrium input-output structure of the economy to aggregate the numbers in all industries. Further, the authors conclude that markups materially affect the impact of microeconomic productivity and markup shocks on output, both at the sector and at the firm level.

A conference participant questioned

the stability of their decomposition if considering factors like the fixed cost and free entry of the firms. The presenter answered that if assuming that firms have constant-returns-to-scale production functions, are ex-ante identical, markups are exogenous, and there is no return to product variety, then entry is socially wasteful, and their results will not change. The presenter also mentioned that they can build more extensions on their basic framework to handle more complications when the aforementioned assumptions do not hold, which will also be one of their focuses in a future study.

They present both ex-post and ex-ante results. Ex-post reduced-form results require no information about the microeconomic production functions besides input-output expenditure shares. The downside of these results is that they depend on observation of factor income shares before and after the shock. Ex-ante structural results require information about microeconomic elasticities of substitution. From the information associated with input-output expenditure shares, they deduce the implied changes in factor income shares. As a side benefit, their ex-ante results determine how factor income shares respond to shocks for a general neoclassical production structure, which is a question of independent interest in studies of inequality.

In inefficient economies, changes in allocative efficiency are non-zero. Solow (1957) defines for the first time the productivity change as the residual growth of real product not accounted for by the growth of real factor input. Hulten (1978) building on the work of the former, provides a rationale for using Domar aggregation to interpret the Solow residual as a measure of aggregate TFP. Hulten's

theorem has a non-parametric power, which comes from exploiting a macro-envelope condition resulting from the first welfare theorem. The authors, by instead exploiting micro-envelope conditions: namely that all producers are cost minimizers, they could integrate something. The paper's theoretical contribution is to characterize the macroeconomic impact of microeconomic shocks and their decomposition into "pure" technology effects and changes in allocative efficiency in inefficient economies.

Hulten (1978) shows that in efficient economies, the growth rate of GDP decomposes into the TFP growth in each sector and improvement of allocation of resources throughout

the sectors. The result in the paper generalizes Hulten's theorem beyond efficient economies, it provides an aggregation result for economies with arbitrary neoclassical production functions, input-output networks and distortion wedges. In efficient economies, changes in allocate efficiency are zero to a first order, and so the overall effect characterized by Hulten (1978) boils down to the "pure" technology effect. As in Hulten (1978), the authors use Domar aggregation to interpret the Solow residual as a measure of aggregate TFP.

The relevant elasticity of substitution is higher in our exercise than in Harberger's since it applies across firms within a sector rather than across

sectors. We consider the general equilibrium input-output structure of the economy to aggregate the numbers in all industries because Harberger focused on manufacturing and ignored input-output linkages.

Markups can amplify shocks and attenuate others. Unlike a competitive model, shocks to industries and firms have different effects on output, even controlling for the size. Firm-level shocks trigger larger reallocations of resources across producers than industry-level shocks (since firms are more substitutable). They find that output is more volatile than in a competitive model, regarding firm-level shocks.

On the Optimality of Periodically Reformed, Simple Fiscal Policies

Maxim Troshkin

The recent work in optimal taxation, known as dynamic Mirrleesian approach or New Dynamic Public Finance (NDPF) considers stochastic dynamic economies with information friction, such as hidden skills or productivity that agents know but the planner does not. Finding optimal policies in the setting are complex. One example of complexity is history dependence. The optimal policies that come out from the exercise depend on the stochastic process of the hidden state and are once and forever type of policies. If that stochastic process changes, one needs to recompute the problem again to find tax policies today.

The author explores alternatives to optimal policy design, different from the policies that are once and forever and assume perfect knowledge of the distribution of the stochastic process.

The work provides bases to answer the following questions: How to design an optimal policy in a world where pre-tax income distribution changes regularly and the change in the stochastic process is unknown from period to period. The framework can test the optimality of the policies we see are reformed from time to time.

The author studies optimal policy design in a dynamic Mirrleesian setting where information friction is modeled as 'Knightian' uncertainty. He finds that in a world where nobody is certain about the data generating process of the hidden state, optimal policies come out to be periodically reformed. The author shows that the policies will be weakly time consistent and efficient. The result holds even if the planner has access to the full commitment technology and even if on top of uncertainty one adds other information frictions.

To show, how the periodic renegotiation can happen in equilibrium the author develops a model where the data generating process for the stochastic skill process is unknowable.

Agent has pivotal statistical model about the future skill distribution but the agent distrusts, are uncertain about the model and considers other models around it. Aversion to uncertainty will drive the agent to rank allocations based on the worst-case outcome. They define the latter as an outside option, in a state where the agent is better off leaving the social insurance scheme. The author shows that in the framework for any pareto optimal complete contingent contract, there is a sequence of incomplete contracts that implement the complete allocation. Incomplete contract specifies allocation today and allocation in tomorrow's worst-case scenario. Once the economy moves forward and they have not realized the worst-case scenario the principle and the agent can improve upon the contract by renegotiating and writing new incomplete contract.

The author shows that this periodic renegotiation can be an outcome of an efficient allocation. Social planner who has access to the full commitment technology finds it optimal to reform

policies from period to period. The result is due to the 'Knightian' uncertainty. The social planner waits until she learns from the agents. Once the social planner moves forward, she re-learns from the agents' new beliefs and once the worst-case scenario has not been realized, the social planner updates the allocation and continues reforming repeatedly.

The author claims optimal policies that are reformed periodically leads to the simplified optimal policies, in the sense that the policies are redesigned once information arrives, without full backward induction, which is an important advantage compared to the dynamic Mirrleesian setting. However, periodic reforms do not necessarily result as simple policies as linear.

The audience asked questions about the importance of the information structure to produce equilibrium optimal polices that are periodically reformed. The author clarified that it generates the idea of the periodic reforms under various specification of the 'Knightian' uncertainty.

LAEF
Laboratory for Aggregate Economics and Finance

Finn Kydland
Director

Peter Rupert
Associate Director

Laurie Preston
Business Manager

Special thanks for their accurate and concise summaries of the presentations go to UCSB economic PhD students

Travis Cyronek

Yongwook Kim

Maria Kogelnik

Yizhou Liu

Miguel Mascarua (UVA)

Alec McQuilkin

E. Charlie Nusbaum

Jackson Pfeiffer (CMU)

Jaime Ramirez-Cuellar

Molly Schwarz

Ryan Sherrard

FROM THE LAB

Laboratory for Aggregate Economics and Finance
University of California, Santa Barbara
Santa Barbara, CA 93106-9215 U.S.A.

ADDRESS SERVICE REQUESTED

LAEF
Laboratory for Aggregate Economics and Finance

2112 North Hall
University of California,
Santa Barbara
Santa Barbara, CA 93106-9215
U.S.A.
Phone: (805) 893-2258
Fax: (805) 893-8830
Email: Laurie.Preston@ucsb.edu
www.laef.ucsb.edu