NBER WORKING PAPER SERIES

FROM MICRO TO MACRO DEVELOPMENT

Francisco J. Buera
Joseph P. Kaboski
Robert M. Townsend

Working Paper 28423
http://www.nber.org/papers/w28423

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2021

We thank David Lagakos, Jeremy Majerovitz, Jacob Moscona, and Steven Durlauf and the anonymous reviewers for helpful early feedback. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

From Micro to Macro Development
Francisco J. Buera, Joseph P. Kaboski, and Robert M. Townsend
NBER Working Paper No. 28423
February 2021
JEL No. O1,O11,O12,O2

ABSTRACT

Macroeconomic development remains an important policy goal because of its ability to lift entire populations out of poverty. In our review of the literature, we emphasize that the best way to achieve this objective is to embrace a synthesis of methods and ideas, with the science of experiments as a unifying feature. RCTs need representative data and structural modeling, and macro models need to be designed and disciplined to the realities and data of developing country economies. Macroeconomic models have key lessons for gathering and analyzing micro evidence and for moving to an evaluation of macro policy. Resource constraints, heterogeneity, general equilibrium effects, obstacles to trade, dynamics, and returns to scale can all play key roles. A synthesis for macro development is well under way.

Francisco J. Buera
Department of Economics
Washington University in St. Louis
One Brookings Drive
St. Louis, MO 63130
and NBER
fjbuera@wustl.edu

Robert M. Townsend
Department of Economics, E52-538
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
and NBER
rtownsen@mit.edu

Joseph P. Kaboski
Department of Economics
University of Notre Dame
3039 Nanovic Hall
Notre Dame, IN 46556
and NBER
jkaboski@nd.edu

1 Introduction

A chief reason why many remain poor in the world is simply the macroeconomy into which they were born. Thus, the macroeconomic question of what determines the aggregate level and distribution of income remains of utmost importance. The focus of this review is to examine how macroeconomic theory and micro empirical research can complement one another to increase knowledge and to improve macro development policy. By macro development policy, we mean simply any policy targeting economic development that is large enough to have important aggregate or economy-wide distributional and welfare implications. In principle, by this definition, macro development policies include important large-scale policies in a wide range of areas: taxation, transfers, and public spending; trade and industrial policy; roads and other physical infrastructure; financial products and access; education; and health.

Over the past 15 years, there has been a surge in empirical research based on randomized controlled trials (RCTs) and related work evaluating development policies at the small-scale level.[1] This literature has redirected research efforts toward micro-level program evaluations while often eschewing formal theory and macro-level policy modeling. Nevertheless, and perhaps paradoxically, evaluations of micro-level interventions are generally viewed as pilot studies with the policy goal of scaling successes to large numbers of people (Muralidharan and Niehaus, 2017; Banerjee et al., 2017b). Over the same period of time, a macro economic literature has made advances in building and solving models incorporating rich micro-structure, that is, with well-defined agent problems, with heterogeneity, and with contracting and market frictions. However this line of work has tended to rely on strong structural assumptions, e.g., assumptions on functional forms and distributions of unobservables, and on somewhat stylized calibration strategies, and thus economists often view it as disconnected from micro empirical research.

This article is about the interplay of economic theory and data as an experimental science and how economists can use it to formulate macro development policy. The implicit juxtaposition in the previous paragraph may suggest little overlap between micro development economics featuring RCTs and structural macro models. We emphasize the opposite: a much needed

[1] One measure of its impact is the award of the 2019 Nobel Prize in Economic Sciences to the leaders in the micro-experimental development field: Abhijit Banerjee, Esther Duflo and Michael Kremer.

synthesis is already underway, from bottom up to top down. RCT practitioners need and are using economic theory, and macro economists scrutinize micro underpinnings and are employing experimental evidence.

Almost all development policies studied in experiments have aggregate effects that go beyond their direct impacts on the individuals treated. To give just one example here, education policies affect the distribution of skills in the workforce which affect wages and incentives for investment and so on. Indeed, synthesis of micro with macro can be exemplified by efforts to scale up local experiments, that is, to assess what would be the impact if a policy were implemented at the national level.

We care about the welfare consequences of these policies, i.e., not just their treatment effects but how to interpret what these effects mean for human well-being. We therefore need to be able to combine the modern empirical toolkit for reduced-form estimation of treatment effects with tools that let us address both aggregate treatment and the distribution of welfare gains and losses. Modern macro provides just such a toolkit.

There is historical precedent for synthesizing structural macro and experimental methods, evidenced by the development of these methods. Modern macro was forged conceptually as an economic science within the framework of the Cowles Foundation: contributors were instrumental in merging advances in general equilibrium theory with those in econometrics, and they confronted identification issues in observational data. The development of the distinction between exogenous and endogenous variables, that is, externally versus internally-determined, variables and variation is an illustrative example. Moreover, efforts went beyond patterns and correlations to models, and then, as we have emphasized, to policy prescriptions. (Koopmans, 1947) described the business cycle empirics of Burns and Mitchell (1946) as "Measurement Without Theory". Koopmans emphasized that theories need to be micro-founded on economic behavior, especially if theories are to be used for policy, which he advocated as the key end goal. The debate was substantive and also impactful on institutional lines, with the Cowles Commission modifying their slogan from "Science is Measurement" to "Measurement and Theory" in 1952. These arguments were later echoed in Lucas (1976)'s critique of the policy reliability of large-scale macroeconomic models. Since that time, the macroeconomic methodological debate has overwhelmingly settled on the idea of micro-founded behavioral models.

Field experiments in micro economics also have a long, rich history. They date back at least to Fisher's work on crops, which were analyzed in the

1920s but based on experiments started in the 1840s (Fisher, 1921). Likewise there is early quasi-experimental work evaluating theory in agriculture of developing countries. For example, (Schultz, 1964) used the 1918-19 influenza epidemic as a natural experiment to test the doctrine of labor of zero value in agriculture.

The synthesis also aligns with an appreciation on both micro and macro sides of the lack of an ideal laboratory setting for testing theory and policy prescriptions in macroeconomics. For example, Angrist and Pischke (2010) were pleased to discover that Lucas (1988a) uses experimental language in trying to assess the impact of monetary policy; in lieu of experimenting on an entire economy, he imagined engineering a laboratory depression to test monetary theory by one day shocking the number of ride and game tickets granted per dollar in Kennywood amusement park. However, Lucas concedes that the informativeness of such an experiment would be nonetheless limited, "[w]e are not really interested in understanding and preventing depressions in hypothetical amusement parks. We are interested in our own, vastly more complicated society."

A bridge is needed to relate small-scale experimental work to the macroeconomy. We view the interplay between experimental empirics and macro theory as an important part of that bridge.[2]

To demonstrate, consider a thought experiment closer to the field of development than Kennywood Park: a laboratory of villages in Thailand with the introduction of microcredit as an experiment.[3] In our laboratory villages, most people earn their living through a combination of subsistence, day labor, and/or entrepreneurial activities. While they face a variety of risks to their income, productivity, wealth, and consumption needs, financial instruments like credit, savings, and insurance are limited. The most typical experiment might be to measure a treatment effect that randomly gives some villagers access to microcredit, while keeping others as controls. Some of treated villagers might borrow to invest, others to consume in various ways, others to

[2]Our call for a bridge between experimental micro work and structural macro analysis naturally relates to Heckman (2010)'s proposed third way to do policy analysis, combining structural and program evaluation approaches.

[3]Naturally this is an example dear to our hearts and own expertise, but we consider a wide variety of policies in the paper's body. Especially relevant here with its explicit general equilibrium foundations, is work on villages as entire economies, e.g., Townsend (1993, 1995), Udry (1995) and a large and growing literature which has followed. See recent theory based work on village network in Chandrasekhar et al. (2018).

re-lend to family or friends. Would this tell us what would happen if all villagers were given access? What if, instead, all *villages* were given access? What if persistent institutions were set up, with explicit subsidies funded through the national government budget? What might the long run and distributional impacts be on credit, income, savings, wages and prices, and how might these differ from alternative policies? These are the types of questions we are interested in answering, but clearly answering such questions will require more than empirical evidence, even though experiments can help greatly in identifying key mechanisms and parameters. It will also require models that are explicit about the basic decisions and frictions that villagers face, the relevant resource constraints and dimensions of heterogeneity, the level of integration of the various villages both amongst themselves and with Bangkok and the rest of the world, the dynamic considerations of various actors, and any potential sources of increasing or decreasing returns to scale.

Such a blend of models and data needs to appropriately balance parsimony and realism in order to be informative. Such work is a difficult process, requiring judicious decisions about theory and measurement. Although not definitive, this is the type of work that we feel will best inform macro development policy. We feature the details of various work in this essay. Several feature teams of researchers leveraging skills by integrating field and methods: theoretical, empirical, and computational.

From our review, we draw several lessons for researchers involving fundamental aspects of modern macro theory. *Aggregate resource constraints* on key inputs can make the aggregate impacts of scaled policies proportionately smaller than implied by RCT evidence. *Heterogeneity* means that the largest players have disproportionate impacts on aggregates, but these are often underrepresented in micro studies. Heterogeneity is also particularly important in economies with incomplete financial markets, prevalent market frictions, strongly nonlinear Engel curves, strongly gendered behavioral differences, and large variation in productive technologies, e.g., from traditional to modern. *General equilibrium* (GE) effects can redistribute, either reinforcing or counteracting partial equilibrium distributional impacts. Furthermore, spatial frictions and financial frictions, with heterogeneity in *obstacles to trade* can make the transmission of GE effects vary from across the populations and regions. Likewise, *capital accumulation* responds to prices. This can make the long run impacts substantially larger than from short run impacts, but there are key instances that go the other way. Finally, *economies of scale* can make large-scale policy proportionally more powerful than micro inter-

ventions, and can be a motivation for aggregate policy interventions. We stress that each lesson is a qualitative pattern with quantitative implications that researchers must discipline with models and data.

We proceed as follows. Section 2 underscores the needs for methodological integration and provides concrete examples of such integration. In Section 3, we address our principle lessons from macroeconomics for development policy at scale and discuss ways in which a micro-macro synthesis can better inform the key mechanisms.

2 The Micro-Macro Synthesis

We first discuss the relative strengths of micro experimental and macro structural work, underscoring the need for an integration of the two approaches. We then follow up with some concrete examples of such integration to inform macro development policy using various methodological approaches.

2.1 Micro: RCTs and Quasi-Experimental Techniques Need Theory

The obvious strength of experiments, whether natural or randomized, and instruments is that they can in principle produce credible exogeneity, ideally free from selection bias and other types of endogeneity. The controlled variation that is created is like a medical researcher injecting dye into a patient or an experimental physicist colliding particles, each to create new data in a researcher-controlled environment to shed new light on a phenomenon of interest. This explains the appeal of experiments for economics, by analogy.[4]

Valid natural experiments, as instruments, can produce an unbiased estimate of the local average treatment effect (LATE), the impact on those people induced into program participation (Imbens and Angrist, 1994). Similarly, micro randomized experiments can produce an internally valid statistical estimate of the average treatment effect (ATE) in the trial sample, since randomization is an instrument that induces participation across the full sample.

[4]There have been an active recent debate on the merits and limitations of RCTs. We refer the interest reader to recent reviews by Banerjee and Duflo (2009), Deaton (2009), Heckman and Urzua (2010), and Imbens (2010).

Treatment effects can be heterogeneous, however, as is often the case in economic applications. When the trial sample is not representative of the most relevant population, the ATE in an experimental setting could be different from the ATE for the population of interest. Moreover, in a heterogeneous population the ATE is only one of a host of other possible informative, relevant and distinct impact moments, e.g., the fraction of people who benefit or the median impact. These include the average effect of treatment on the treated (TT) or, important for predicting the impacts of an expansion, the average effect of treatment on the untreated (TUT). Heckman and Vytlacil (2005) introduce the unifying parameter (for a particular treatment) called the marginal treatment effect (MTE), the infinitesimal counterpart of the LATE, and show that other desired impact estimates are weighted averages of the MTE. Furthermore, when take up by the treatment group is limited, the focus is often on the average intent to treat effect (ITT) as distinct from treatment, actual voluntary participation in the policy or program.

The importance of heterogeneous impact effects is also a central motivation of the analysis of the conditions for an instrument to identify the LATE (Imbens and Angrist, 1994; Imbens, 2010). In the case of an instrument for treatment, the econometrician observes a variable that induces treatment for some individuals but does not directly affect the outcomes. In addition the instrument has to induce the treatment in a monotone way, i.e., it cannot simultaneously induce an otherwise treated person to abstain from the treatment. This assumption is needed precisely because of the heterogeneity, i.e., the possibility that treatment effects differ between those induced to treatment by the instrument and those induced to drop the treatment by the instrument.

To illustrate the importance of heterogeneous treatment effects, return to our example of an economy populated by individuals that differ in their entrepreneurial and worker abilities. Individual choices are potentially constrained by their financial resources, e.g., entrepreneurs' investment is constrained by their beginning-of-period wealth. Individuals choose their occupation and further, choose whether to pay a fixed cost to participate in a financial market, i.e., to save or borrow at a fixed interest rate. Imagine a policy-maker interested in measuring the benefits of fostering entrepreneurship or the effect of a policy that improves financial access. Townsend and Urzua (2009) use a version of this simple example as a laboratory to study the answers provided by alternative experimental and quasi-experimental methods.

Townsend and Urzua consider two experiments in the model that provide potential instruments to identify the LATE of entrepreneurship and the LATE of financial participation on income net of subsidies.[5] The first experiment consists of the random assignment of a lump-sum subsidy to individuals that become entrepreneurs. Importantly, it is assumed that the subsidy cannot be used to finance the capital of the business. Thus, the experiment provides an instrument that has a monotone effect on the probability that an individual becomes an entrepreneur, but it does not have a direct effect on the outcome of interest, income net of the subsidy. The estimate of the LATE of entrepreneurship on income in the model data is negative, reflecting the negative selection of entrepreneurs induced by the subsidy; they would have been wage earners, given their talent, if it were not for the subsidy. It is straightforward to calculate other treatment effects using the structure of the estimated model. Revealingly, in their application, the ATE of entrepreneurship is positive and the TT is twice as big.

In the second experiment, individuals in the model are randomly assigned a lower cost of accessing a financial market where they can save and borrow at a fixed interest rate. This intervention provides an estimate of the ITT of financial participation, as slightly less than 20% of the agents in the model choose to participate when offered a lower cost of participation. Equivalently, the intervention provides an instrument to estimate the LATE of financial participation. The estimated LATE of participation in the financial market is positive. In this case the ATE and TT of participation in the financial market are positive as well, although their relative magnitudes are substantially different. Interestingly, this second intervention does not provide a valid instrument to estimate the LATE of entrepreneurship, as it violates both the monotonicity and independence conditions. Monotonicity is violated because, while financial participation induces some productive individuals who are now able to take up loans to become entrepreneurs, it also induces some less talented entrepreneurs, who are now able to obtain a higher return on their assets, to close their unproductive businesses. Independence is violated because participation in the financial market affects the ability to finance the capital and, therefore, it has a direct effect on the income of entrepreneurs.

[5]In this application a pure experimental strategy can not be used to identify the ATE of entrepreneurship, as it is not realistic to randomly assign the occupational *choice* to an individual. Rather, an RCT can be designed to randomly assign monetary incentives that affect occupational choices. Thus, the experiment provides an instrument for entrepreneurship.

This simple example is intended to highlight the importance of prior theoretical knowledge in the design and interpretation of experimental and quasi-experimental empirical strategies. As Imbens (2010) points out, a priori theory is critical to determine the validity of instrumental variables. In the context of his argument, an atheoretical view of micro work is not valid.

2.2 Macro: Structural Analyses Need Measured Micro Foundations

Structural models utilize theory to link behavior with outcomes and welfare. The advantage of theory is that it provides a coherent framework through which to understand, interpret, and assess empirical results, incorporating them in an organized fashion. Macro theory in particular is our focus, as it is well suited for thinking about macro development policies, i.e., policies that can have sizable impacts on the aggregate economy, whether in levels or distribution.

The strength of macro models is helping researchers think clearly and quantitatively about the roles of aggregate resource constraints, general equilibrium effects, various potential obstacles to trade, forward-looking behavior and dynamics, scale economies and their implications for aggregates, distributions, and welfare, all through the lens of micro-founded behavior. Macroeconomists have much to add to the micro empirical literature in thinking through how a complex set of changes in an economy affect the economic outcomes of any given individual and the economy as a whole.[6]

There is nevertheless a need for strong microfoundations, appropriately tested and disciplined with data. Empirically, the work of Kydland and Prescott (1982) introduced calibration strategies that allowed the researcher to quantitatively discipline and evaluate models by comparing moments in simulated data to those in actual data. However, in the context of representative agent models, this analysis was done at the aggregate level. The twin criticism of this macro agenda was its harsh anti-econometric tone and the insufficiency of frictionless representative agent models to capture important economic realities, e.g., Summers (1986), Hansen and Heckman (1996).

The critiques are not new, but progress continues. Cambridge-Cambridge debates concerned whether an aggregate capital stock and aggregate (or representative) production function exist (e.g., Robinson, 1953, 1959; Solow,

[6]We elaborate on these claims in Section 3.

1957; Fisher, 1969). Aggregation research has been extended even quite recently (Baqaee and Farhi, 2019a). In the 1970s and 1980s, macroeconomics relied almost exclusively on representative agents, as noted, but seminal work by Bewley (1986), Aiyagari (1994), Hopenhayn (1992), Huggett (1993), and Krusell and Smith (1998) has allowed macroeconomists to model heterogeneous consumers and firms in an aggregate, general equilibrium economy. The point: these and related innovations allow macro models to consider richer sets of frictions and to increasingly embrace micro data to discipline them. Thus, while these critiques still apply to some work, virtually all these critiques have been incorporated to some extent in other contemporary macro work.

Closer to development, Banerjee and Duflo (2005) discussed the shortcoming of representative agent macro models, and the need to consider important empirical evidence on heterogeneity and frictions, including spatial frictions and financial frictions, even if accounting for aggregates were the focus. This critique proved influential in both the theory and empirical work on misallocation (e.g., Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009), one literature that is now salient in both development and in macro.

All models are abstractions, but when the assumptions built into the model veer too far from reality on important dimensions, the predictions of macroeconomic models and recommendations for policy can be off the mark. This is true both of the key assumptions but also the key parameter values, whether calibrated, estimated, or assigned. The full set of microfoundations and parameter values needs to be well justified, and this depends on the policy to be analyzed.

An illustrative example is recent work on agricultural misallocation. A series of macro development papers used detailed micro data and applied the framework of Restuccia and Rogerson (2008) and Hsieh and Klenow (2009) to agriculture, attributing large productivity losses to the misallocation across farmers of varying productivity. These exercises in Malawi (Santaeulalia-Llopis and Restuccia, 2017), Ethiopia (Chen et al., 2019), and China (Adamapoulos et al., 2017) attributed much of the productivity gap to losses from resource misallocation, especially misallocation connected with market frictions on land and other inputs. The work assumed a uniform, calibrated Cobb-Douglas production function in agriculture. Follow up work, however, by Gollin and Udry (2019) showed that similar levels of misallocation are observed in Tanzania and Uganda, but that this *apparent* misallocation is also observed across plots owned by the same farmer, who does not

face market frictions in allocating those internal resources.[7] Their analysis is also model-based, but more detailed, and it attributes much of this seeming misallocation to measurement error, heterogeneity in technologies, and ex post shocks to output.[8] Thus, a stark macro model with strong predictions yielded a faulty conclusion from the data because it didn't rest on the *appropriate* micro foundations.[9].

2.3 The middle ground between theory and data: how much structure is needed?

The more dynamics, frictions, heterogeneity and realism that macro models incorporate, however, the more difficult and challenging they are to solve and estimate. Basic intuitions and pen and paper solutions surrender to computational results, which can themselves be black box predictions. This is a valid criticism.[10]

There exists an important middle ground in terms of solving and stating models. Sometimes, depending on the policy and class of models under consideration, a given model does not have to be fully solved, nor does a particular model need to be chosen among a class. The field of macro trade, for example, has built explicit models but also relied on exact hat algebra results and sufficient statistics. Dekle et al. (2008) specify a trade model for the world economy in terms of changes from the current equilibrium, avoiding the need to assemble proxies for bilateral trade costs or inferring parameters of technology; Arkolakis et al. (2012) show that the welfare implication of a large class of trade model depend on only the trade share and the trade elasticity, very much in the tradition of public finance (see Chetty, 2009, for a review). Another example is aggregating firm-level distortions. Baqaee and Farhi (2019b) provide approximate formulas for the aggregate impacts

[7]Esfahani (2018) has related work on showing high factor ratio variation among same-owned plots in Ethiopia and Malawi.

[8]The macro development work on misallocation in manufacturing allows for heterogeneity in technology across industries, but measurement error also seems to play an important role in those findings, as well, see Bils et al. (2018).

[9]Related, a natural use of small experiments is to test and sometimes reject conventional theories incorporated in macro models. Behavioral economics has both old and new contributions on this front, including in development, e.g. Kahneman and Tversky (1979) and Kaur et al. (2015)

[10]Sometimes the predictions of models can be illuminated by comparative static simulations, robustness exercises, and the like.

of distortions on aggregate TFP and output in general equilibrium. The model allows for a fairly general input-output structure with a wide class of distortions. Related, Sraer and Thesmar (2018) propose a method to use experimental data to analyze the aggregate effects of policies. In particular, propose moments of the distribution of revenue-to-capital as sufficient statistics for aggregate outcomes.

Both papers have limitations of course. The frictions in Baqaee and Farhi (2019b) must be exogenous. Sraer and Thesmar (2018)'s frictions can be endogenous, but they must be homogeneous of degree one with respect to the wage and aggregate demand and firm-level technologies should be Cobb-Douglas. More importantly, they require experimental data about the *long run* effects of policies. Neither paper allows for firm entry.

An example spelling out an explicit non-competitive market structure is Joaquim et al. (2019), which develop an industrial organization model of imperfect competition in financial services with two components, a utility-profit contract frontier coming from mechanism design and allowing a variety of frictions versus the industrial organization part that delineates the number of bank branches and the extent of competition. It is shown that the contracting frontier can be identified with market share data, hence policies that impact competition do not need to identify the underlying frictions. They have not explored the issue of dynamics - both in contracting and competition of financial service providers, which may be relevant in particular applications

The overall lesson is that sufficient statistics can be utilized and useful in accounting for GE effects, but they are typically sufficient only for a well-defined class of models. Some policy welfare questions can be answered without much structure. The point also is to use theory to figure out when and how this can be accomplished.

2.4 From bottom up to top down, the literature is already integrated

We want to emphasize that an integration is already occurring, and so we offer some concrete representative examples. We organize these examples by their different methodological approaches, ordering the approaches from micro to macro but emphasizing the synthesis. For each approach, we offer a few salient examples to convey the flavor of the approach, rather than a

comprehensive review of the literature.

The first methodological approach involves integrating structural models with experimental evidence. Staying at the micro level, this structural synthesis nevertheless helps interpret and leverage experimental evidence. The studies of the negative income tax experiments to estimate the effect of taxes on the labor supply response of low-income workers in the US are early examples (see the review in Hausman, 1985, Section 7).

An early development policy example of the synthesis of structural and RCT work is Todd and Wolpin (2006), who use the randomized implementation of the Mexican program, *Progresa*, which offered conditional cash transfers and nutritional supplements to families who kept their children in school. Simulations of counterfactual policies using the estimated model establish that the program could have had similar effects at lower cost by eliminating subsidies. Attanasio et al. (2011) also examines *Progresa* with a structural model, but they use the experimental variation to estimate, rather than validate, their more general model.

Kaboski and Townsend (2011) evaluate a large microfinance intervention in Thailand, anticipated in the earlier discussion of village experiments. The village fund credit program placed a million Thai baht (roughly $25,000) in each village. They estimate their structural model of household investment and saving from pre-intervention data, and then use the Thai Million Baht village fund policy experiment to validate the model, nearly replicating the difference-in-differences estimates in Kaboski and Townsend (2012). Heterogeneity is key in distinguishing households, e.g., those on the margin of making an indivisible investment (for whom consumption actually drops); those who are credit-constrained hand to mouth (for whom consumption increases one-for-one), and those neither constrained nor investing (who increase consumption, no longer needing as much buffer stock saving with future credit availability now assured.)[11]

A second emerging methodological approach involves using structural macro models together with small-scale experimental and quasi-experimental evidence to evaluate the impacts of policies at an economy-wide scale. A broad range of different policies have been studied in this growing literature, at the heart of our envisioned synthesis.

[11] In the same spirit, Lagakos et al. (2018) evaluate transportation subsidies to seasonal migration for rural workers using experimental evidence in Bryan et al. (2014). RCT practitioners also bring structure into evaluations, typically focusing on estimation of preferences (for example, Kremer et al., 2011; Kreindler, 2018; Mahajan et al., 2019).

Buera et al. (2020) studies the macroeconomic impacts of universal access to micro-credit. The model is calibrated with macro and firm distributional non-experimental moments but is shown to match at the partial equilibrium level the micro evidence of impact from RCT and natural experimental variation, in India (Banerjee et al., 2015) and Thailand (Kaboski and Townsend, 2012), respectively. Buera et al. (2014) use a similar model and methods to examine the impact of cash grants to the poor. Both papers are clear that the partial equilibrium versus general equilibrium implications prices, wages, interest rates, TFP, saving and capital can be quite distinct, as will be the distribution of winners and losers.

These modeling efforts have a close parallel in Bergquist et al. (2019) who, as a team, are concerned with the impact of local experiments versus scaled-up agricultural policies. The do this within the context of general equilibrium model of farm production and, importantly, trade, featuring spatial trade frictions in the markets for agricultural outputs and inputs. The model is estimated with a multitude of micro data from Uganda. As in the first example, they find that the direction and magnitude of the GE effect is a non-linear and non-monotone function of the scale of the intervention.[12]

A third methodological approach is to empirically evaluate experiments that are actually introduced at scale, whether RCTs or quasi, natural experiments. This can provide valuable concrete evidence of impact that goes beyond what might happen through the lens of a model, to what actually happens, ideally using models to provide an interpretation.

For example, while Kaboski and Townsend (2012) analyzed the village fund credit program as an intervention at the village level, it was introduced in every village in the entire country. Ehrlich and Townsend (2019) establish that wages increase more for highly treated isolated villages, as modeled, due to production demand for inputs and limited labor inflow. Likewise, if the collection of nearby villages was highly treated, the wage in a given target village goes up, but this is due to higher out migration, as modeled.[13]

[12]Other examples include: Donovan (2020) who examines the impact of introducing farmers to a low-risk seed, validating the model with RCT findings on risk-reducing seeds from Emerick et al. (2016); Fujimoto et al. (2019) who examine the effects of free secondary schooling vouchers to talented, poor households matching aggregate moments and evidence from the secondary school scholarships experiment by Duflo et al. (2017); Greenwood et al. (2019) who examine the impact of circumcision on AIDS prevalence using experimental evidence and a general equilibrium search and matching model.

[13]Cai et al. (2016) is an example of a study of a large-scale village banking intervention

A well-known national anti-poverty program is India's National Rural Employment Generation Scheme (NREGS), which guaranteed 100 days of government funded employment every year to every adult willing to work. Muralidharan et al. (2017) evaluated in this context a electronic payment system introduced experimentally, to improve the program, at a scale of 19 million people. They find that 90% of the increase in income caused by the program comes from non-program earnings, driven by increases in both market wages and, somewhat surprising, private-sector employment. The further exploration of the positive effects on private-sector employment provide a good candidate for the use of an explicit theoretical framework to interpret experimental results.

A fourth approach is to utilize the natural experiments that are often gifts to researchers stemming directly, yet accidentally, from the details of macro policy, which is often implemented from the top down. These natural experiments can often allowing relatively transparent difference-in-differences analysis. This evidence is then used in structural modeling.

A prominent example is the study of the effects of transfers using tax rebates in the US with timing randomized, that is, based on the last few digits of social security numbers which are thus essentially random: Johnson et al. (2006) study the effect of income tax rebates of 2001, and Parker et al. (2013) study the effect of the stimulus payments of 2008. Kaplan and Violante (2014) use this evidence to motivate and calibrate a life cycle model with costly access to illiquid assets, yielding rich hand-to-mouth agents[14]

Other recent natural experiments in developing economies shed light on classical questions in monetary economics. Chodorow-Reich et al. (2019) recently analyze the 2016 Indian "demonetization", a dramatic initiative that made 86% of cash in circulation illegal tender overnight. Detailed data from the Reserve Bank of India allowed them to pin-point variation in the timing of the release of replacement notes. In one sense, this is a version of Lucas' imagined Kennywood experiment carried out in India in real time.[15]

---

in China. The migration outcomes are studied by (Cai, 2020).

[14]As in Sarto (2018), Wolf (2019) and many others, a stand must be taken on a class of models.

[15]Related, Alvarez and Argente (2019) estimate the private benefits for Uber riders of using alternative payment methods. They rely on a combination of large field experiments and quasi-natural experiments given by the ban of cash as a payment method in certain regions in Mexico.

3 Gains from Trade: Lessons for Modeling and Empirical Work

We now review some take away lessons for those interested in thinking about macro policy organized around five themes influencing modern macro theory: i) aggregate resource constraints; ii) heterogeneity; iii) general equilibrium effects; iv) obstacles to trade; v) dynamic optimization and capital accumulation; and vi) diseconomies and economies of scale. For each theme, we introduce a key lesson using a concrete example examining the scale up of a microintervention policy that has been evaluated using an RCT. We then extend the discussion with additional examples of macropolicy and additional opportunities for integrating microempirics with macrotheory.[16]

3.1 Aggregate Resource Constraints

A starting point for thinking about macroeconomic theory is the importance of resource constraints. Individuals and firms may be able to flexibly adjust scarce resources at the prevailing market prices, but in the aggregate these resources are typically constrained. Consequently, the aggregate consequences of a scaled policy can be much smaller than what is measured in a microintervention.

A concrete example of this is Buera et al. (2020)'s study of the aggregate effects of microcredit. In their model, a small-scale microcredit intervention leads credit to flow to some microentrepreneurs with lower marginal productivity. The impacts on those individuals with access include higher capital and higher income but those individuals are lower average productivity entrepreneurs. Without any external capital injection, however, aggregate capital is fixed in the short run. That is, capital is a scarce, constrained resource, and demand for microcredit capital must compete with other uses in the economy. As a consequence, the interest rate increases so that individual capital choices are consistent with the aggregate resource constraint. This

[16]The five themes in which we organize the lessons have somewhat artificial boundaries. For instance, the discussion of the (general equilibrium) effects of policies on prices will show up naturally when discussing the importance of resource constraints. Still, we choose to delay a careful discussion of general equilibrium effects until after highlighting the importance of heterogeneity, as this sets the stage to an analysis of how general equilibrium effects are central to understand the heterogeneous impacts of policies.

leads to much lower immediate impacts of microcredit on capital and income. Moreover, realized aggregate income comes from improved allocation of capital toward *higher* productivity entrepreneurs.

In the previous example, the important resource constraint is on capital, but aggregate constraints can be important for many different resources. When considering policy more broadly, a key question is which scarce resources are most important for success. In scaling micro interventions, the relevant scarce resources – e.g., physical capital, human capital, government capacity, firms and other organizational capital – may be quite specific to the intervention. Administrative capacity at the level of the organization or government may be an important constraint in scaling, for example.[17] One of the primary arguments for randomized controlled trials has been that, given scarcity in funding and administrative capacity, randomized assignment is a fair allocation rule. In some of the most famous RCTs (e.g., *Progresa*), randomized rollout was indeed justified as an equitable way to proceed given limited budget and administrative resources (Parker and Teruel, 2005).

Since scarce resources can be specific to the macro development policy in question, modeling relevant constraints requires case-by-case judgement based on local knowledge of the environment and program being scaled, economic theory, and prior empirical results as the following examples illustrate. In Donovan (2020), the key aggregation and scarcity in the analysis involve not only the labor, which can be assigned to produce agricultural and non-agricultural goods, but the output of the non agricultural sector, which can be used for final consumption or as an intermediate input into the production of agricultural goods. With a focus on risk aversion, scarcity of agricultural output plays a key role at the village level because of the incomplete markets and minimum consumption requirements. Bergquist et al. (2019) uses a similar commodity space, but they add land as a scarce factor for agricultural output. In the human capital model of Fujimoto et al. (2019), the direct costs of education need to be funded. The key dimensions of scarcity and aggregation involve the government budget constraint – where the real direct costs of education need to be financed by taxes – and the exogenous distribution of inherent ability. The labor market and governments are both national in their model, and that is the key level of aggregation.

[17]When the administrative capacity is the limiting factor, the equilibrium adjustment after scaling micro interventions may involve the rationing of these resources across competing programs instead of changes in market prices. Thus, aggregate resource constraints need not work through GE channels.

There are additional possibilities for integrating microempirics with macro models. Micro can learn from macro. For example, even micro-level empirics themselves can benefit from applying resource constraints. While budget constraints are rarely satisfied at the individual level (because of measurement error, for example), applying such constraints at a more aggregate level may be more reasonable. For example, Kaboski et al. (2019) analyze a cash grant experiment. They apply a budget constraint at the village level to estimations of impacts on income, consumption, savings, investment, lending, etc. to yield cross-equation restrictions that increase efficiency in estimation. They require that changes in investment, consumption, and savings equal changes in income and transfers (including the cash grant). The restriction cannot be rejected and yields more reasonable and easily interpretable estimates. A common practice in reduced form empirical work is to log outcome variables in order to reduce the influence of outliers. This practice can often increase significance in the estimation, but also leads to less obvious, nonlinear cross-equation restrictions. Coefficients don't "add up" in clear ways. Bonferroni type adjustments have become common in the micro-empirical literature when multiple outcomes are analyzed, but utilizing cross-equation constraints of aggregation have not been typically applied.

Macro models can also be better informed and disciplined by micro studies. Work that estimates the geographical level of impacts of programs (on individuals, local areas, and neighboring areas, for example) can give insights into the most appropriate level of aggregation. Similarly, take up rates from experiments can tell us which sub-populations of a distribution are more likely to be directly impacted by an intervention. That is, experimental results with extensive margin elasticities can be used as part of the specifications of macro models.

Microexperiments can also discipline key elasticities that determine the consequences of scarce resources. For example, Lagakos et al. (2018) use the estimated micro impacts of the migration subsidy in Bangladesh on migration and rural wages to estimate the labor demand elasticity in rural labor markets in the context of their Roy model framework.

3.2 Heterogeneity

Many micro experimental studies of policy interventions find that not everyone is impacted the same by treatment. Heterogeneity matters in macro models, both because of its distributional implications but also for aggregates

when decisions don't scale linearly, e.g., a rich person with double the wealth spends differently than the totals of two poor persons. Consequently, a key lesson for scaling interventions is the need to ask how a policy impacts the biggest players for aggregate effects, since they often have the largest weight in aggregates. Unfortunately, these players often don't show up in survey data. The largest firms are unlikely to show up in firm survey samples, but this is especially true when experimental work focuses on microenterprises.[18] The same issue holds for wealth and the impacts of policies on aggregate savings, or income and the impacts of tax policies on aggregate revenue.

An interesting example of this is the evaluation of village funds in Kaboski and Townsend (2011). One might think that considering large players would be less important in village economies when poverty is the main outcome of interest rather than national aggregates, but big investors matter in their study. Their structural model emphasizes the very heterogeneous responses to credit that different agents can have: some increasing consumption, while others decreasing consumption in order to increase investment. In their estimated model, there is wide variation in the the size of household's desired indivisible investment. Large investments have disproportionate weight on aggregate investments. In theory, there is always a chance that the availability of even a small amount of credit enables a much larger indivisible investment, and such cases, while rare, matter substantially for aggregate investment consequences of credit. Using simulations, the paper shows that small sample sizes are pooorly suited for capturing the impact on aggregate investment given the importance of large, infrequent indivisible investments.

In theory, heterogeneity is particularly important in economies with incomplete financial markets, prevalent market frictions, strongly nonlinear Engel curves, strongly gendered behavioral differences, large variation in productive technologies, e.g., from traditional to modern. Each of these are especially applicable in developing countries, and indeed there are important micro studies emphasizing dimensions of heterogeneity in these economies. These dimensions include wealth, productivity, ability, entrepreneurial background, education, and gender. Moreover, when we are especially concerned about inequality, those with less social standing in society, and the poorest of the poor dimensions like gender, education, and wealth can be especially

[18]The opposite problem occurs with some firm-level data from business registers in developing countries, which tend to report data for firms above a minimum size threshold (Bartelsman et al., 2004).

important from a distributional standpoint. Many microstudies find that such heterogeneity is critical.[19]

Macro models can now accommodate heterogeneity much more so than in the past, and good reduced-form empirical work (whether RCT or otherwise) can help direct and discipline heterogeneity in macro models. This opens up additional opportunities to integrate the two modes of research. We elaborate on three such opportunities.

First, RCTs can be used to help identify the important dimensions of heterogeneity that need to be modeled and to validate or reject particular model or distributional assumptions. For example, de Mel et al. (2008) evaluated randomized grants to non-employer entrepreneurs in urban areas of Sri Lanka. They found sizable impacts on investments and monthly profits with implied monthly return on the grants that substantially exceeded market interest rates. Yet the researchers emphasized the strong heterogeneity in returns to capital. those with disproportionately low levels of wealth and those with higher education or cognitive ability drove the overall impacts. In line with this evidence, wealth and talent are two important dimensions of heterogeneity that have dominated the theoretical and quantitative macro literature on financially-constrained entrepreneurs, which have been used to evaluate the impacts of the financial system more broadly.[20]

Second, combinations of experimental and observational data can be used to better discipline macro models. There are limitations associated with exclusive use of samples from micro RCTs, since evidence from RCTS often comes from non-representative populations. Consider the study by Muralidharan et al. (2017), which examines an experimental improvement in the payment system used for the India's NREGS employment guarantee program. On the one hand, the sample is impressively large and heterogeneous, involving 19 million people in India across 157 subdistricts. Yet, this sample may not be the most appropriate for disciplining a macro model for certain aggregate questions: by its very nature, the population is rural, so it may not be appropriate for thinking about urban labor supply. However,

[19] A comprehensive list would be impossible, but we note a couple examples. Banerjee et al. (2017a) show that microfinance impacts in India are concentrated among existing "gung-ho" entrepreneurs. Duflo (2003) finds pensions received by women in South Africa had a large impact on the anthropometric status of girls but little effect on that of boys.

[20] See for instance, Giné and Townsend (2004); Jeong and Townsend (2007); Townsend and Ueda (2006); Buera (2009); Banerjee and Moll (2010); Townsend and Ueda (2011); Buera et al. (2011); Midrigan and Xu (2014); Moll (2014).

non-representative data can be extrapolated to the relevant macro population using structural assumptions and more representative observational data. For example, to discipline labor supply from rural-urban migration, Lagakos et al. (2018) combine RCT data with representative household data in their evaluation of a scaled migration subsidy in Bangladesh. Parameters governing aggregate moments are targeted using the representative data, and yet the RCT data gives important evidence for disciplining the productivity parameters and disutility cost governing Roy-like selection into migration.

Third, macro models suggest the importance of attempts to sample more representatively. Samples of micro studies are often targeted toward specific populations (e.g., the poorest or least served, marginal populations in terms of take up), the populations that are easiest to survey (e.g., spatially clustered, those with lower opportunity cost of time, those available at the time of survey), or those that are easiest to incorporate into randomized evaluation methods (e.g., targeted expansion areas). Though targeting for poverty reduction makes sense as a laudable goal, this neglects the great force that macro growth and improved distribution can bring. For that we need data to estimate the underpinnings of macro development models. One way to balance these goals in collecting data is to combine surveying frequencies with representative national samples in order to obtain effective sampling weights that can be used to aggregate. That is, one can combine over-sampled populations of particular interest with partial samples from the full population.

3.3 General Equilibrium Effects

In a market economy general equilibrium (GE) effects are the price counterpart of aggregate constraints on quantities discussed in Section 3.1.[21] Nevertheless, we view them as conceptually distinct, especially when heterogeneity is important as discussed in Secton 3.2: GE effects on relative prices or wages can impact not only aggregates but the distribution of income. A key lesson here is that these GE effects can be important, either reinforcing or counteracting the partial equilibrium distributional impacts.[22]

[21] Acemoglu (2010) also emphasized the importance of GE and resource constraints as a reason that theory needs development.

[22] Many emerging empirical studies highlight the distributional consequences of GE effects. Cunha et al. (2018) show drops in prices of goods transferred in kind, especially in less developed villages. Similarly, Filmer et al. (2018) show that cash-transfers led to an increase in the prices of protein-rich foods, which affected the health outcomes of

An interesting example of this is the work of Donovan (2020), a macro study that examines the role that risk and the subsistence consumption requirements of poor farmers plays in the use of productivity-enhancing intermediates like high yield seeds and fertilizer. In the model, farmers with expected consumption near the subsistence requirement are reticent of the additional risk that intermediate expenditures entail. He uses the model to investigate the impact of the introduction of lower risk seed, which was evaluated in an RCT by Emerick et al. (2016). When done at scale, the adoption of this seed leads to higher output, and a GE decrease in the price of agricultural goods. This reduces the cost of subsistence in the event of crop failure, actually reinforcing the initial risk reduction from the seed, making the induced take up of intermediates and welfare gains larger in GE than in PE, especially for poor farmers.

In macro development models incorporating micro experiments, GE effects have been analyzed in many different settings and policy analyses. The experiment in Cai and Szeidl (2018) randomized loan takeup to small-medium enterprises (SMEs), both within (some firms receive the treatment) and across local markets (a higher fraction of firms receive the treatment in some markets) in China. The direct effect of credit access is to increase sales, profits, and factor payments, but the indirect effect is to reduce the sales, profits, and factor payments of one's competitors; a fair amount of the direct effect of treatment comes from a business stealing effect. They use a Melitz-type model to show how this works through prices. Hence, the GE effects in this case lead to smaller welfare impacts than the direct effects might indicate. In another recent paper, Fujimoto et al. (2019) examine the effects of free secondary schooling vouchers to talented, poor households in Ghana. They calibrate a model with human capital accumulation to match aggregate moments and evidence from the secondary school scholarships experiment by Duflo et al. (2017). They find important general equilibrium effects, which are key drivers of the fiscal and distributional effects of the program. The skill premium falls in consequence, and rich, skilled lose through lower high skill wages, while the poor, unskilled gain even though they do not attend school, as their productivity increases by working with more skilled, complementary workers. Examining the aggregate impact of circumcision on AIDS prevalence in Malawi, Greenwood et al. (2019) apply experimental evidence and a general equilibrium search and matching model, in which heterogeneous in-

non-beneficiary children.

dividuals selectively adopt different sexual practices while knowing inherent risk. They find that there are important GE changes in prices in markets for marriage, protected, and unprotected sex, although they have relatively small effects on behavior, since prices affect men and women's behavior in opposite directions.

Again, there are additional lessons and opportunities for integration. First, the experimental practitioner must always be concerned with GE effects; an ideal RCT requires an absence of any spillover of treatment onto non-treated groups, e.g., individuals, villages, and regions within a country. If treatment changes the prices that the nontreated face, that can impact the outcomes of the non-treated households thereby biasing the estimates.

Second, experiments themselves can intentionally generate exogenous price variation in either the prices that sellers face or the prices that buyers face. These can be used to illuminate the relevant elasticities of supply and demand that are key to macro models. Randomized price variation has been used inside some of the most influential randomized experiments, e.g., Fehr and Goette (2007), who estimate a labor supply elasticity by varying the wages workers received.

Third, because highly concentrated interventions can generate endogenous GE effects themselves, microevidence from these interventions can be used to discipline or validate the parameters governing the strength of GE forces in macro models. This can be done through estimation, calibration, or additional validation checks. For example, Akram et al. (2017) use variation in the intensity of their migration subsidy in Bangladeshi villages, i.e., the fraction of villagers receiving the subsidy, to show that the outflow of labor from local villages also increased wages. On the flip side, Breza and Kinnan (2018) use the impact of coordinated default on microfinance loans in Andhra Pradesh, India, the balance sheet impacts it had on microfinance banks, and the variation in pre-existing presence across other Indian provinces to show that a contraction in microcredit led to a drop in wages. Kaboski and Townsend (2012) similarly estimate positive impacts on wages using exogenous variation in treatment from the village fund. This evidence is used as an out-of-sample test on the reasonableness of the labor-entrepreneurship forces in Buera et al. (2020).

Finally, while localized GE mechanisms can be measured directly using cross-sectional or spatial variation, any impact that is truly at the aggregate level can only be quantified through the lens of a model (unless policies are randomized at the country level, for many countries!), and indeed models

are needed to interpret any local GE effects that happen through market segmentation. These analyses incorporate theory to varying degrees.

In some cases, assessing the welfare gains from price effects require only reduced form elasticities, while in others more explicit models and simulations are required. For example, the work of Arkolakis et al. (2012), shows that a reduced form elasticity is sufficient for calculating gains from a reduction in the price of imports in a wide class of tractable and relatively standard models. The analysis in Donaldson (2018) is a nice application of this in evaluating the benefits of infrastructure investment in India on the gains to agricultural trade. This follows the tradition of the public finance literature (Harberger, 1964; Chetty, 2009). For other questions, a full scale GE model with microfoundations may be necessary. Van Leemput (2016) looks at trade costs within India, but she does so within a two-sector model of agriculture and non-agriculture where agriculture is a necessity good as in Fieler (2011); the constant elasticity assumption is violated in the aggregate and moreover the model has heterogeneity. Van Leemput's welfare gains cannot, therefore, be reduced to a single summary statistic, and indeed they require quantitative simulation.

In sum, macro development policy needs to consider price effects because there is strong theoretical and empirical reasons to believe that price effects are important for aggregates, efficiency, distribution, and welfare. Moreover, incorporating and quantifying GE effects into analysis requires both models and data.

3.4 Obstacles to Trade

Spatial and other trade frictions are an important source of endogenous heterogeneity and can be intimately linked with GE mechanisms. The economic environments, available markets, and prevailing prices of those in remote areas of developing countries can be strikingly different from those in capital cities. We have emphasized the importance of careful modeling. A key lesson for extrapolating RCT evidence to scaled up policies for an entire nation is that the transmission of GE effects and the impacts of policy can vary considerably region-to-region and even person-to-person.

A concrete example of this is the paper evaluating the aggregate and distributional consequences of a scaled subsidy to agricultural intermediates by Bergquist et al. (2019). They develop a (static) trade model in which trade costs vary from household to household and map it to the Ugandan economy

using a clever combination of calibration and estimation.[23] Since markets are only partially integrated, wages vary from market to market and effective prices vary from household to household. They contrast a local saturation of the intermediates subsidy intervention with a nationwide saturation. One headline finding is that the intermediate subsidy causes larger GE effects on wages when nationwide than when just local, since labor is less elastic in aggregate, a point we have made in the sections on resource constraints and general equilibrium effects. More relevant to the point in this section, they find that some areas benefit considerably more at scale, while others benefit considerably less at scale depending on how integrated they are and the extent to which they specialize in particular crops. These patterns depend not only on factor endowments but the full set of trade costs, which they estimate.

Other examples incorporate simpler geographies to look at spatial frictions. Ehrlich and Townsend (2019) model not only trade but also migration flows using a geography of villages along a circle, while Lagakos et al. (2018) model migration using a simple two location urban-rural model. Naturally, there is a tradeoff between simplicity and the level of detail of predictions. Bergquist et al. (2019) can assess household level impacts, while Ehrlich and Townsend (2019) can assess village level impacts. Lagakos et al. (2018) can only show urban-rural distinctions, but this simplicity allows them to incorporate rich Roy-model heterogeneity and migration dynamics.

There are more opportunities for integration on these fronts, in part because of advances in modeling, improvements in data availability, and innovations in combining the two.

In the past, modeling such regional heterogeneity in ways that could be taken to data was not practical. However, in recent years, building on trade models, there have been important advances in not only proving the existence of spatial equilibria with mobility of goods and workers but also in developing methods to solve these models.[24] The models often use Frechet distribution assumptions on productivity to derive gravity-model patterns of trade, but they add in labor mobility. Researchers have used these models most directly to think about the spatial distributional effects of international

[23] Like Donovan (2020), they are concerned with the productivity benefits of agricultural intermediates and they too model nonhomothetic preferences, but they abstract from endogenous dynamics. The these two papers illustrate important trade-offs between richer model elements and tractability.

[24] Redding and Rossi-Hansberg (2017) provide an excellent review.

trade policy (e.g., Monte et al., 2018; Caliendo et al., 2018, 2019), but they have also considered the benefits of infrastructure investment (e.g., Allen and Arkolakis, 2014) and consequences of migration policy (e.g., Desmet et al., 2018b). All of these are important issues for developing countries.

Many of these models allow for a great deal of detail, but, similar to the sufficient statistics approach, do not require solving the full model in order to generate the predictions of interest. Instead, practitioners quantify the model using "exact hat algebra". The moniker was introduced by Dekle et al. (2008) to describe a development on the older "hat algebra" approximation that dates back to Jones (1965)'s contribution to general equilibrium theory. His "algebra" was the matrix algebra from the comparative statics exercise of log-linearizing a model and totally differentiating it. The resulting equations were functions of log changes in factor prices and endowments as well as elasticities and shares, which could be taken from either outside estimates (elasticities) or read from the data (shares). The log changes were eventually denoted by "hats" (or carats). "Exact" hat formulas occur when modeling choice allow such equations to no longer be approximations but exact, even for large changes. In this case, without solving the model or even specifying its deep parameters, counterfactual predictions can be calculated using the hat deviations from a benchmark together with knowledge of the relevant elasticities of substitution between goods and factors. Gravity models with trade/migration/capital flow costs typically require additional information on bilateral flows. In that sense, the elasticities, shares, and flows together constitute a sufficient statistics approach.

These models are especially appropriate for developing country analysis, since a robust and important feature of both the countries and their development policy is that regional heterogeneity matters. Development involves urban centers, peri-urban peripheries, and rural areas, and wide spatial variation. There are important questions involving both the international and within country flow of labor, goods, firms, and capital. Mapping models to developing countries requires nationally representative samples, however. Such samples exist for the largest countries (e.g., China, India, and Brazil) and are emerging for others, e.g., the World Bank's Enterprise Surveys for firms and Living Standard Measurement Surveys (LSMS) for households. The panel length of these data sets are limited but also typically not required for these spatial models, since they are static. However, one does need to augment these using knowledge of geographic location.

The methods of Bergquist et al. (2019) for Uganda are resourceful in lever-

aging limited data from disparate sources, and thus are particularly insightful.[25] They require key elasticities, shares, and trade costs. They combine multiple data sources. They use the detailed data of the LSMS samples to calibrate the relevant, region-specific, share parameters. They then combine these rich data with less detailed Ugandan census data and geocoded data on crop suitability from the Food and Agriculture Organization (FAO) in order to project shares from the representative LSMS sample to the full country. For key elasticities, they use RCT evidence from a nearby country to discipline the relevant elasticity of substitution governing demand and separate RCT evidence from another nearby country to discipline farmers' elasticity of substitution across traditional and modern (input-intensive) farming technologies. They then use global crop price variation as an instrument to estimate farmers' elasticity of substitution across crops. Identification of trade costs is likewise tricky since they lack the full set of household-household trade flows. They combine survey data on market-level flows and prices together with use household production and consumption decisions to infer farm gate prices, with price gaps across farms thereby identifying trade costs. Lacking data on distance at the household level, they project price gaps (trade costs) on observed demographic characteristics.

Another important example is the work of Townsend and coauthors, who have applied structural models of entrepreneurship decisions under financial frictions and mapped them to rich observational data from multiple datasets in Thailand. Financial frictions limit trade across time and states of the world and can also have a spatial dimension. For instance, the work of Paulson and Townsend (2004) and Paulson et al. (2006) shows regional heterogeneity in the underlying causes of financing frictions: moral hazard-driven frictions in the more developed Central region of Thailand and a limited commitment collateral constraint in the more rural Northeast in combination with other features. Ahlin and Townsend (2007) derive the signs of covariances among variables for a variety of models of limited liability lending, which allow them to rule out models of limited liability that are counter to these observations in the data, yielding a similar regional pattern. Finally,

[25] Other useful approaches applying regional variation and microdata to give potential insights for aggregate spatial models include: experimental methods, e.g., Egger et al. (2019); applying national income accounting methods to villages or regions, e.g. Paweenawat and Townsend (2012, 2019); using the spatial expansion of programs, e.g., Jack and Suri (2014), and using region-specific data to discipline region specific market imperfections, Moll et al. (2017).

Karaivanov and Townsend (2014) perform pairwise comparisons of the likelihoods of cross sectional and panel data on consumption and investment to distinguish between a variety of non-nested structural models. A rural-urban pattern emerges similar to that of Northeast vs. Central. This work leads to aggregate and regional policy analysis. For example, Moll et al. (2017) study migration and capital flows in a model with regional variation in financing environments. Consistent with earlier work, they impose that urban constraints be moral-hazard driven, while rural constraints derive from limited commitment. Mapping the data to unique regional flow of funds data, they show substantial rural-urban flows of labor and capital, and they show that policies to restrict these flows lead to substantial aggregate losses, but with heterogeneous impact across regions. This work is an example where model predictions together with observational data help us moves beyond reduced form modeling of financial frictions and distinguish between underlying mechanisms, and then showing their importance for policy and understanding the Thai growth.

The applications move beyond Thailand, however, as Dabla-Norris et al. (forthcoming) develop a general equilibrium framework featuring multiple realistic sources of financial frictions to study how different financial constraints interact in equilibrium. They show that financial inclusion policies should target the most binding constraint, and this can vary across countries in ways that may not be obvious, a priori. There are important tradeoffs between financial inclusion, GDP, and the distribution of income when conducting such policies. There are also inter-temporal tradeoffs and policy commitment issues, as some variables move in the short term in the opposite direction from their longer run changes.

3.5 Dynamic Optimization and Capital Accumulation

Dynamic considerations are indeed important, and dynamic models have been well analyzed in macroeconomics. At the most basic level, savings and investment decisions respond to changes in income and interest rates. Thus, dynamics are likely to be important any time GE effects are important, but dynamics and price effects can depend critically on the specifics of theory. A key lesson for scaling up the results of micro interventions is that long run impacts can differ substantially, in direction and magnitude, from short run impacts when laws of motion depend on prices.

Usually, long run elasticities are bigger than short-run elasticities, hence

the impacts on prices will be smaller in the long run but the impacts on quantities will be larger. A key example of this is the work of Fried and Lagakos (2020), who study the aggregate impacts of reliable electrity-generating infrastructure on development. Micro empirical work in developing countries focus on short-run impacts of power outages and find comparatively small impacts. In their model, for a given set of producers' capital and technology, they can replicate these short run impacts since power outages only lower productivity by leaving both existing output-producing capital and private generator capital idle. This lowers the return to capital and to the use of modern technologies, however, so investment falls both on the intensive margin (capital of existing establishments) and the extensive margins (new entry and the adoption of modern technologies among encumbents). In the long-run, both the capital stock and productivity of producers decline relative to a world without power outages and the aggregate impacts on output and productivity are many times larger.

In some cases, however, the long run impacts on quantities can be smaller than short run impacts. Recall the aggregate microfinance analysis of Buera et al. (2020) in a world with financial frictions in the form of collateral constraints discussed in Section 3.1. In that paper, microcredit increases interest rates both in the short and long run, but the capital stock still falls because the direct effect of the availability of microcredit is to reduce precautionary savings and to transfer income from individuals with high saving rates to those with low saving rates. Since savings has collateral value for acquiring capital, the capital stock also falls and this accumulates over time. Wage gains are experienced in the short run, but not the long run. Indeed, since it is the accumulation of savings that is so pivotal, the endogenous decline in capital can happen even when the country is modeled as a small open economy with a fixed interest rate, and especially so when microfinance capital is subsidized, where the wage can actually fall.

Other examples of the importance of dynamic considerations abound. In some cases, the laws of motion for dynamic models can lead to multiple dynamically stable steady states, where the less desirable steady states are referred to as poverty traps. The presence of such poverty traps – whether individual poverty traps, where these exist for the individuals' laws of motion, or aggregate poverty traps, where they exist in the aggregate dynamics – are of keen interest to development economists concerned with aggregate

and distributional outcomes.[26] An earlier theoretical heterogeneous agents literature in development emphasized the possibility for multiple equilibria and aggregate poverty traps (e.g., Banerjee and Newman, 1993; Galor and Zeira, 1993; Piketty, 1997), and the positive role of permanent redistribution policies (Aghion and Bolton, 1997). In these models, the wage or interest rate depend crucially on the distribution of income and/or wealth, and this feedback loop between distribution and aggregates is crucial. These models had policy prescriptions that opened up the idea that one-time redistribution might jolt the economy from the bad to the good equilibrium, improving aggregate surplus. However, the combination of high levels of heterogeneity, idiosyncratic shocks that cause churning in the distribution of entrepreneurial productivity, and forward-looking behavior can lead to a singular stationary equilibrium in related models of entrepreneurship and investment as in Buera et al. (2011), Buera and Shin (2013a), Buera and Shin (2013b), and Buera et al. (2014). While these models still yield poverty traps at the individual-level, one-time redistributions dissipate nearly completely within 10 years, and instead policies that continually redistribute are necessary to maximize aggregate output.[27]

Beyond these examples, there are again additional opportunities for integrating microempirics and macro models to better inform dynamic models and dynamic implications for policies more broadly.

Empirically, a number of long run experimental results are emerging in the RCT literature.[28] These studies are potentially highly informative for models. For example, focusing on the poverty trap dynamics, randomized recipients of one-time asset grants have been tracked by several studies. Research on cash grants in Uganda (Blattman et al., 2018) and Ethiopia (Blattman et al., 2019) shows no long term impacts on measures like earnings, consumption, employment, and health at nine years and five years, respectively, although the initial results were quite promising (Blattman et

[26]To be more precise, we refer to an individual poverty trap when there is a threshold wealth such that individuals that start with wealth below this threshold remain poor, while those that start with wealth above this threshold save to overcome poverty. An aggregate poverty traps arises when an economy is in a stationary with low per-capita income due to its initial distribution of wealth, but would transition to a stationary with higher per-capita income is a particular redistribution of wealth is implemented.

[27]Bowles et al. (2016) give a comprehensive treatment of poverty traps more broadly, including outside of development.

[28]Bouguen et al. (2019) combine a review of long term RCT evaluations together with many practical suggestions for addressing empirical issues, such as respondent tracking.

al., 2014; Blattman and Dercon, 2018, respectively). In contrast, two RCTs in South Asia find more sustained impacts from grants. de Mel et al. (2012) returned to the entrepreneurs in their de Mel et al. (2008) grant experiment in Sri Lanka and found sustained impacts on business survival and business profitability for male grantees. Similarly, Bandiera et al. (2013) studied ultra poor asset grants of livestock to poor women in Bangladesh. They find sustained poverty reduction and increased labor supply among the women after seven years. Our point is that not only is the evidence mixed, but that variations in underlying environments may lie behind the puzzles. Again, this evidence can help improve modeling, with assumptions appropriate to not only the environment but the policy considered. For instance, are the findings of the experiments consistent with the implications of quantitative models featuring financial frictions, and if not, what are the model elements or institutional details required to match this evidence?[29]

At the same time models might assist long run tracking of experiments in interpreting the data. Absent theory, the challenges of interpretation are formidable. Consider a "phase-in" RCT like *Progresa* described above, where one can always compare those who were treated early with those who were treated late. One cannot estimate the impact of long-term treatment, however. At best, one can estimate the long-term effects of difference in treatment timings or the long-term effects of a short difference in total amount of time the treatment was received. For example, if one were to observe a dissipation of treatment effects over time after phase in, it could reflect (i) the benefits of treatment falling over time, (ii) persistent benefits to treated communities, combined with catch up of later treated communities, or (iii) a weakening of the effective treatment through channels like migration, e.g., contamination of treatment and control areas/populations. The last possibility is particularly problematic; the assumption of no spillovers required for valid RCT comparison become increasingly dubious over long periods of time, and so the definition of treatment can become much more nuanced. An exacerbating source of spillovers is the presence of spatially-segmented GE effects, as described above. In principle, dynamic models based on sound theory could help both in the interpretation of the results of the few long

[29]Buera et al. (2014) is a crude first attempt at confronting the implications of quantitative macro development models featuring financial friction with the recent evidence from asset grants, and related historical evidence by Bleakley and Ferrie (2013). The aforementioned work by Kaplan and Violante (2014) is a more developed example showing how evidence from quasi-experiments can be used to identify important model elements.

term RCT studies, but also in the prediction of the long term consequences of policies at the macro level. One could model the interactions, dynamics, and spillovers directly, and simply use the differences in outcomes among treatment and control to identify key parameters. The long run empirics would then become additional identifying moments. The challenge here is to have dynamic extensions of models with realistic spatial frictions as in Desmet et al. (2018a).

Micro empiricists increasingly appreciate that aggregate shocks can be important, even for understanding and extrapolating RCT evidence, because they impact the interpretation of individual studies. For example, the results found in agricultural studies from a particular year may depend crucially on the weather experienced during that year. As a concrete example, Zhang and Carter (1997) argue that Lin (1992)'s estimates of the effect of agricultural reforms in post-Mao China were substantially overestimated because he failed to consider the effects of weather on agricultural output. Even if there is cross-sectional heterogeneity in weather patterns, this variation alone cannot be simply extrapolated to account for the aggregate shock, since the average or aggregate shock can affect the price of crops, labor, etc. in general equilibrium. Indeed, Rosenzweig and Udry (2016) make the point that aggregate shocks are important for determining the external validity of micro-level RCTs themselves, because of these aggregate channels. Naturally, this evidence calls for an integration of microevidence with quantitative macro models featuring aggregate shocks, a task that can be computationally challenging when considering environments with rich heterogeneity and various frictions.

While dynamic, general equilibrium models with heterogeneity, idiosyncratic shocks, and aggregate shocks can be difficult to compute and solve, new methods enable easier analysis. The models are difficult because investment responds to expectations of the future path of prices, but the state that determines prices at any point in time is a high dimensional object (the full distribution). Advances involve combinations of discretization and local linearizations. Krusell and Smith (1998) argued that most of the important aspects of the wealth distribution in their context could be captured with just a single moment, the mean of the wealth distribution. But more generally, the distribution of wealth matters. Recent work by Ahn et al. (2018) provides more efficient computational methods to solve models with heterogeneity and aggregate shocks that allow for larger numbers of individual state variables. The approach is to solve (an approximation to) the

global stationary equilibrium *without aggregate shocks* using efficient finite-difference methods, and then to add in the aggregate shocks using linear approximations. They provide a precoded toolbox to make these methods available for broader applications. Silva and Townsend (2019) propose a new form of solving for the transitional dynamics, by combining perturbation and finite difference methods to discretize the system of ODE/PDEs that does not assume aggregate or idiosyncratic shocks are small. These advances have in principle enabled macro models to more easily address the environments of developing countries with extensive heterogeneity across the macroeconomy, highly uncertain environments, incomplete markets, and important additional frictions. These richer models can be used, for example, to perform Monte Carlo analyses of estimates of the impacts of large scale policies in environments with large aggregate shocks or, more generally, to study the role of aggregate shocks in development (Aguiar and Gopinath, 2007).

Of course, assets and capital are not the only laws of motion that are of potential importance. We have already noted the distribution of entrepreneurial ability. In models with one-time set up costs for firms, the full distribution of firms can be an important state variable, while in the "ideas" models of Lucas (2009), Lucas and Moll (2014), Perla and Tonetti (2014), for example, the state space can be the distribution of productivity of all individuals in an economy. Human capital may be an important dynamic state, or locational distribution of workers or firms in the spatial models described above, including migration models. Culture itself can be a slow-moving state influenced by policy, for example sexual norms in the Bangladesh contraception intervention (Munshi and Myaux, 2006), and ignoring dynamic policy impacts on culture could lead to short-sighted policy.

In sum, quantitative models with rich state spaces provide a way of accounting for and understanding the multidimensional heterogeneity in treatment effects that RCTs often test for and find. Moreover, rich dynamics, anchored in theory, allow policy makers to account for changing dimensions of this heterogeneity, both the dynamics that are responses to the interventions, and those that are independent of the intervention and can therefore be taken as exogenous to the intervention but with endogenous dynamics nevertheless. Both long and short-run empirical work can be helpful in testing and disciplining theory.

3.6 Economies of Scale

Returns to scale are another important consideration for macro development policy and for extrapolating the results from micro experiments. The behavior of economies under aggregate decreasing returns to scale is very similar to the behavior under the presence of aggregate resource constraints (fixed factors are a key reason for decreasing returns), so we do not address this in more detail.[30] However, economies of scale can also exist and can open opportunities for development and have aggregate policy implications (e.g., coordination, Pigouvian subsidies). Focusing on scaling interventions, the key lesson is that impacts at scale can be larger because of economies of scale.

Here we turn to a concrete empirical example: mobile money, a banking and money transfer program. It has created a strong foothold in some countries, most famously M-PESA in Kenya, where it is almost universal, but it has far fewer users in other countries (Suri, 2017). Mobile money works through the mobile phone system and a network of existing retail establishments who act as agents. It is therefore much more useful at scale than in a small pilot and network effects are important. Indeed, Jack and Suri (2014) document the rapid rollout of M-PESA and the importance of the underlying retail network in adoption. There are at least two sources of increasing returns to scale at play. First, the more users, the more potential bilateral transfers, and that relationship increases with the square of the number of users. Second, the more users, the denser the network of agents, and the smaller the transaction costs of using the network.[31]

[30] Nevertheless, the channels through which smaller impacts under scale can arise are interesting and varied. For example, information can undermine scaled policy. For example, the experimental work of Pomeranz (2015) and Carrillo et al. (2015) who randomized warning letters of tax enforcement to firms that increased reporting, even though they were largely bluffs. If one were to implement a national policy based on this, it is likely that firms and tax lawyers would soon (perhaps even ahead of time) learn of the policy and learn to ignore the false letters. Othe surprising channels include bureaucratic and political opposition (Bold et al., 2018), counterproductive competition (Galle, 2019), and information.

[31] Batista and Vicente (2013) use an experiment in Mozambique to assess the impacts of mobile money, finding that it indeed increases a willingness to send remittances, but there is no assessment of the role of scale. Mobile money is interesting, not only because it is a scalable microintervention that can assist, for example, in the delivery of social service payments – the large-scale NREGS experiment in India by Muralidharan et al. (2017) is a nice example – but also because it has regulatory questions that relate to the more

Returning to theory, at the micro-level of firms and households, indivisibilities, learning by doing, or minimum efficient scales in production, consumption, and/or delivery may cause impacts to be larger when scaled up on the intensive margin. If there are regions of increasing returns to scale at this micro-level, one-time increases in scale can have persistent impacts. Beyond network effects, other potential theoretical mechanisms for external economies of scale can stem from critical mass, industrial concentration, technology spillovers, and external learning. On the firm side, many of these are the traditional agglomeration forces emphasized by Marshall (1890). Their external nature gives them particular policy relevance. Within the macro literature, external scale mechanisms played an important role in the endogenous growth theory of the early 1990s (Romer, 1990; Grossman and Helpman, 1991; Aghion and Howitt, 1992) in promoting policies of openness and subsidies to human capital and research and development expenditures.

The question of whether there are increasing returns to scale is therefore important in macro models, important for scaling policies, and important for guiding efficient growth policy. How can microstudies help inform us on this question?

Recent empirical work has used randomized short-term increases in demand to evaluate impacts. Carrillo et al. (2019) use the randomization of government procurement contracts in Ecuador to examine their impact on firms. They find essentially no evidence of increasing returns: firms expand by the amount of the government procurement contract and quickly contract after it is complete. In contrast, Atkin et al. (2017) randomize export orders to Egyptian carpet manufacturers and find evidence of improved technical efficiency over time, which they attribute to learning-by-doing mechanisms. Alfaro-Urena et al. (2019) use a quasi-experimental approach to look at the impact of multinational contracts on domestic suppliers in Costa Rica. They find firms grow 20 percent in response and TFP increases by 3-9%, depending on whether using a model-based measurement or not.

Another way to help uncover nonconvexities is to vary the endowment of treatment. Such is the approach in Balboni et al. (2018), who evaluate livestock grants in rural Bangladesh. The treated households in their experiment all receive the same sized grant of assets (livestock), but they vary in

traditional macroeconomic topics of money and banking. Electronic payments systems, for example, proved important in the 2016 demonetization policy in India (Chodorow-Reich et al., 2018).

their initial wealth. Focusing on asset dynamics, they show s-shaped returns consistent with poverty trap asset dynamics, and evidence of twin peaks in the asset distribution, consistent with multiple equilibria. Examining the same issue in Uganda, Kaboski et al. (2019) use a choice across different lotteries as an alternative way of combining intensive margin variation with randomization, and tying it to risk preference. When agents are in regions of non-convexities, they can be locally risk-loving. They offer a choice over cash lotteries to an underbanked population in Uganda and show that a substantial share of the population prefers a high-risk lottery even when it has a slightly lower expected payout. This study is unique in that it uses the macro theory to design the experiment, motivating the experiment by the local risk-loving that accompanies non-convexities. They attribute the non-convexity to the market for land, which has high returns.

The diffusion of technologies typically follow s-shapes, implying regions with increasing returns and regions with decreasing returns. The aforementioned recent macro literature that focuses on idea diffusion models thinks about the role of interaction in productivity growth (Lucas, 2009; Lucas and Moll, 2014; Perla and Tonetti, 2014), providing microfoundations for human capital externalities (Lucas, 1988b). Interactions between agents lead to the transfer of ideas and hence productivity increases for the less productive agents. The search externalities in these models allow for efficiency enhancing policy. These frameworks have been used to explicitly model at the micro level how macro policies like trade and openness to foreign direct investment impact the distribution of productivity across producers (Alvarez et al., 2013; Sampson, 2015; Buera and Oberfield, 2020; Perla et al., 2021), and how knowledge is diffused within and across firms through the reallocation of workers (Herkenhoff et al., 2018; Jarosch et al., 2019). In a recent paper combining macro models with experimental micro evidence, Brooks et al. (2018) use the evidence from an experiment combining novice entrepreneurs with more experienced mentors to quantitatively discipline the productivity diffusion process in Kenyan slums. The paper is quite innovative, although it leaves open the question of how to extrapolate narrow experimental evidence to the broader economy.

To summarize, economies of scale are important considerations for guiding macro development policy and interpreting the results of small-scale studies in light of macro policy. Nevertheless, both large scale and small-scale studies can help inform macro theory, especially when designed in conjunction with it.

4 Conclusions

We have discussed advances in empirical development and macroeconomics which we argue are best seen as complementary.

On the one hand, beginning at the micro level, grounded on new study designs, namely, natural and randomized experiments in conjunction with models, we are now able to have clearer snapshots and understanding of the realities in developing countries, and to identify causal effects of various micro policy interventions. There are obvious limits to a purely empirical strategy, which we discussed at length.

On the other hand, at the macro level, there have been parallel advances in the construction of quantitative macroeconomic models that incorporate rich heterogeneity across agents and space, and various contracting and trading frictions. We argue that the rich quantitative heterogeneous agents models are ideally suited to connect the dots of micro interventions and, therefore, evaluate (and devise) macro development policies. The recent wealth of micro evidence is central to better discipline the structure of the new generation of policy oriented macroeconomic models.

Although as we have shown that there is already much work integrating these approaches, this is not a review of a mature literature. Rather, we view the paper as equal parts review and call to arms for more work in an exciting and promising direction. The development problem is important but hard to crack, with no magic bullets. There are certainly no magic methodological approaches.

Successfully moving this agenda forward requires not only the integration of the seemingly methodological extremes, randomized experiments and macro modeling, but advances from other complementary areas in economics, from the collection and homogenization of representative country datasets to better understanding of the identification of structural models. The needs span across many fields in economics: field work, macro and growth modeling, econometrics, spatial methods, trade, industrial organization, labor, family economics, public economics, etc. The most important questions require people with diverse expertise and skills. We encourage macro or trade development economists to bring models to collaborate with micro empiricists who have collected RCT data to evaluate scaled policies, and we have cited many such examples. More ambitious, however, is for such collaborations to begin earlier in the process, with the macro question and conceptual model leading the design of experiments and data collection. The ground is

therefore fertile for both young researchers and for new collaborations.

References

Acemoglu, Daron, "Theory, General Equilibrium, and Political Economy in Development Economics," *The Journal of Economic Perspectives*, 2010, *24* (3), 17–32.

Adamapoulos, Tasso, Loren Brandt, Jessica Leight, and Diego Restuccia, "Misallocation, Selection and Productivity: A Quantitative Analysis with Panel Data from China," NBER Working Paper 23039, National Bureau of Economic Research January 2017.

Aghion, Philippe and Patrick Bolton, "A Theory of Trickle-Down Growth and Development," *Review of Economic Studies*, 1997, *64* (2), 151–172.

Aghion, Phillipe and Peter Howitt, "A Model of Growth Through Creative Destruction," *Econometrica*, March 1992, *60* (2), 323–351.

Aguiar, Mark and Gita Gopinath, "Emerging Market Business Cycles: The Cycle Is the Trend," *Journal of Political Economy*, 2007, *115*, 69–102.

Ahlin, Christian and Robert M. Townsend, "Using Repayment Data to Test Across Models of Joint Liability Lending," *Economic Journal*, 2007, *117* (517), 11–51.

Ahn, SeHyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry, and Christian Wolf, "When Inequality Matters for Macro and Macro Matters for Inequality," *NBER Macroeconomics Annual*, 2018, *32* (1), 1 – 75.

Aiyagari, S. Rao, "Uninsured Idiosyncratic Risk and Aggregate Saving," *The Quarterly Journal of Economics*, 1994, *109* (3), 659–684.

Akram, Agha Ali, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak, "Effects of Emigration on Rural Labor Markets," NBER Working Papers 23929, National Bureau of Economic Research, Inc October 2017.

Alfaro-Urena, Alfonso, Isabela Manelici, and Jose P. Vasquez, "The Effects of Joining Multinational Supply Chains: New Evidence from Firm-to-Firm Linkages," NBER Working Paper April 2019.

Allen, Treb and Costas Arkolakis, "Trade and the Topography of the Spatial Economy," *The Quarterly Journal of Economics*, 2014, *129* (3), 1085–1140.

Alvarez, Fernando E. and David Argente, "Consumer Surplus of AlternativePayment Methods: Paying Uber with Cash," BFI Working Papers 2019-96, Becker Friedman Institute September 2019.

_ , Francisco J. Buera, and Jr. Robert E. Lucas, "Idea Flows, Economic Growth, and Trade," NBER Working Papers 19667, National Bureau of Economic Research, Inc November 2013.

Angrist, Joshua D. and Jörn-Steffen Pischke, "The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics," *Journal of Economic Perspectives*, June 2010, *24* (2), 3–30.

Arkolakis, Costas, Arnaud Costinot, and Andres Rodriguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, February 2012, *102* (1), 94–130.

Atkin, David, Amit K. Khandelwal, and Adam Osman, "Exporting and Firm Performance: Evidence from a Randomized Experiment," *Quarterly Journal of Economics*, May 2017, *132* (2), 551–615.

Attanasio, Orazio P., Costas Meghir, and Ana Santiago, "Education Choices in Mexico: Using a Structural Model and a Randomized Experiment to Evaluate PROGRESA," *The Review of Economic Studies*, 08 2011, *79* (1), 37–66.

Balboni, Clare, Oriana Bandiera, Maitreesh Ghatak, and Anton Heil, "Why Do People Stay Poor?," 2018. Working Paper.

Bandiera, Oriana, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul, and Munshi Sulaiman, "Can Basic Entrepreneurship Transform the Economic Lives of the Poor?," IZA Discussion Papers 7386, Institute for the Study of Labor (IZA) May 2013.

Banerjee, Abhijit, Emily Breza, Esther Duflo, and Cynthia Kinnman, "Do Credit Constraints Limit Entrepreneurship? Heterogeneity

in the Returns to Microfinance," Working Papers 17-104, Global Poverty Research Lab 2017.

_ , **Esther Duflo, Rachel Glennerster, and Cynthia Kinnan**, "The Miracle of Microfinance? Evidence from a Randomized Evaluation," *American Economic Journal: Applied Economics*, 2015, *7* (1), 22–53.

_ , **Rukmini Banerji, James Berry, Esther Duflo, Harini Kannan, Shobhini Mukerji, Marc Shotland, and Michael Walton**, "From Proof of Concept to Scalable Policies: Challenges and Solutions, with an Application," *Journal of Economic Perspectives*, November 2017, *31* (4), 73–102.

Banerjee, Abhijit V. and Andrew F. Newman, "Occupational Choice and the Process of Development," *Journal of Political Economy*, 1993, *101* (2), 274–298.

_ **and Benjamin Moll**, "Why Does Misallocation Persist?," *American Economic Journal: Macroeconomics*, January 2010, *2* (1), 189–206.

_ **and Esther Duflo**, "Growth Theory through the Lens of Development Economics," in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 1A, Amsterdam: Elsevier, 2005, pp. 473–552.

_ **and** _ , "The Experimental Approach to Development Economics," *Annual Review of Economics*, April 2009, *1* (1), 151–178.

Baqaee, David Rezza and Emmanuel Farhi, "The Microeconomic foundations of Aggregate Production Functions," *Journal of the European Economic Association*, October 2019, *17* (5).

_ **and** _ , "Productivity and Misallocation in General Equilibrium*," *The Quarterly Journal of Economics*, 09 2019, *135* (1), 105–163.

Bartelsman, Eric, John Haltiwanger, and Stefano Scarpetta, "Microeconomic evidence of creative destruction in industrial and developing countries," Policy Research Working Paper Series 3464, The World Bank December 2004.

Batista, Catia and Pedro C. Vicente, "The Price of Political Opposition: Evidence from Venezuela's Maisanta," Working Paper Series, NOVAFRICA 2013.

Bergquist, Lauren, Benjamin Faber, Thibault Fally, Matthias Hoelzlein, Edward Miguel, and Andres Rodriguez-Clare, "Scaling Agricultural Policy Interventions: Theory and Evidence from Uganda," Unpublished manuscript, University of California at berkeley 2019.

Bewley, Truman, "Stationary Monetary Equilibrium with a Continuum of Independently Fluctuating Consumers," in W. Hildenbrand, A. Mas-Colell, and G. Debreu, eds., *Contributions to Mathematical Economics in Honor of Gérard Debreu*, North Holland, 1986.

Bils, Mark, Peter J. Klenow, and Cian Ruane, "Misallocation or Mismeasurement?," Unpublished manuscript, Stanford University 2018.

Blattman, Christopher and Stefan Dercon, "The Impacts of Industrial and Entrepreneurial Work on Income and Health: Experimental Evidence from Ethiopia," *American Economic Journal: Applied Economics*, July 2018, *10* (3), 1–38.

_ , Nathan Fiala, and Sebastian Martinez, "Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda," *The Quarterly Journal of Economics*, 2014, *129* (2), 697–752.

_ , _ , and _ , "The Long Term Impacts of Grants on Poverty: 9-year Evidence From Uganda's Youth Opportunities Program," Technical Report, National Bureau of Economic Research 2018.

_ , Stefan Dercon, and Simon Franklin, "Impacts of Industrial and Entrepreneurial Jobs on Youth: 5-year Experimental Evidence on Factory Job Offers and Cash Grants in Ethiopia," NBER Working Papers 25788, National Bureau of Economic Research, Inc April 2019.

Bleakley, Hoyt and Joseph P. Ferrie, "Up from Poverty? The 1832 Cherokee Land Lottery and the Long-run Distribution of Wealth," NBER Working Papers 19175, National Bureau of Economic Research, Inc June 2013.

Bold, Tessa, Mwangi Kimenyi, Germano Mwabu, Alice Ng'ang'a, and Justin Sandefur, "Experimental evidence on scaling up education reforms in Kenya," *Journal of Public Economics*, December 2018, *168*, 1–20.

Bouguen, Adrien, Yue Huang, Michael Kremer, and Edward Miguel, "Using Randomized Controlled Trials to Estimate Long-Run Impacts in Development Economics," *Annual Review of Economics*, 2019, *11* (1), 523–561.

Bowles, Samuel, Steven N. Durlauf, and Karla Hoff, *Poverty Traps.*, Vol. 8, Princeton: Princeton University Press, 2016.

Breza, Emily and Cynthia Kinnan, "Measuring the Equilibrium Impacts of Credit: Evidence from the Indian Microfinance Crisis," NBER Working Papers, National Bureau of Economic Research 2018.

Brooks, Wyatt, Kevin Donovan, and Terence R. Johnson, "Identification and Quantitative Implications of an Equilibrium Diffusion Model," Technical Report, University of Notre Dame 2018.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak, "Underinvestment in a Profitable Technology: The Case of Seasonal Migration in Bangladesh," *Econometrica*, 2014, *82* (5), 1671–1748.

Buera, Francisco J., "A Dynamic Model of Entrepreneurship with Borrowing Constraints: Theory and Evidence," *Annals of Finance*, 2009, *5*, 443–464.

_ **and Ezra Oberfield**, "The Global Diffusion of Ideas," *Econometrica*, 2020, *88* (1), 83–114.

_ **and Yongseok Shin**, "Financial Frictions and the Persistence of History: A Quantitative Exploration," *Journal of Political Economy*, 2013, *121* (2), 221 – 272.

_ **and** _ , "Financial Frictions and the Persistence of History: A Quantitative Exploration," *Journal of Political Economy*, 2013, *121* (2), 221–272.

_ **, Joseph P. Kaboski, and Yongseok Shin**, "Finance and Development: A Tale of Two Sectors," *American Economic Review*, August 2011, *101* (5), 1964–2002.

_ **,** _ **, and** _ , "Macro-Perspective on Asset Grants Programs: Occupational and Wealth Mobility," *American Economic Review*, May 2014, *104* (5), 159–64.

_ , _ , **and** _ , "The Macroeconomics of Microfinance," *The Review of Economic Studies*, 08 2020.

Burns, Arthur F. and Wesley C. Mitchell, *Measuring Business Cycles*, Cambridge, MA: National Bureau of Economic Research, 1946.

Cai, Jing and Adam Szeidl, "Interfirm Relationships and Business Performance," *The Quarterly Journal of Economics*, 2018, *133* (3), 1229–1282.

Cai, Shu, Albert Park, and Sangui Wang, "Microfinance Can Raise Incomes: Evidence from a Randomized Control Trial in China," Technical Report, HKUST 2016.

Caliendo, Lorenzo, Fernando Parro, Esteban Rossi-Hansberg, and Pierre-Daniel Sarte, "The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy," *Review of Economic Studies*, 2018, *85* (4), 2042–2096.

_ **, Maximiliano Dvorkin, and Fernando Parro**, "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 2019, *87* (3), 741–835.

Carrillo, Paul, Dave Donaldson, Dina Pomeranz, and Monica Singhal, "The Bigger the Better? Using Lotteries to Identify the Allocative Efficiency Effects of Firm Size," Technical Report 2019.

_ **, Dina Pomeranz, and Monica Singhal**, "Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcementa," *American Economic Journal: Applied Economics*, 2015, *9* (2), 144–164.

Chandrasekhar, Arun G., Cynthia Kinnan, and Horacio Larreguy, "Social Networks as Contract Enforcement: Evidence from a Lab Experiment in the Field," *American Economic Journal: Applied Economics*, October 2018, *10* (4), 43–78.

Chen, Chaoran, Diego Restuccia, and Raul Santaeulalia-Llopis, "The Effects of Land Markets on Resource Allocation and Agricultural Productivity," Working Papers tecipa-649, University of Toronto, Department of Economics September 2019.

Chetty, Raj, "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 2009, *1* (1), 451–488.

Chodorow-Reich, Gabriel, Gita Gopinath, Prachi Mishra, and Abhinav Narayanan, "Cash and the Economy: Evidence from India's Demonetization," NBER Working Papers 25370, National Bureau of Economic Research, Inc December 2018.

_ , _ , _ , and _ , "Cash and the Economy: Evidence from India's Demonetization*," *The Quarterly Journal of Economics*, 09 2019, *135* (1), 57–103.

Cunha, Jesse M, Giacomo De Giorgi, and Seema Jayachandran, "The Price Effects of Cash Versus In-Kind Transfers," *The Review of Economic Studies*, 2018.

Dabla-Norris, Era, Yan Ji, Robert M. Townsend, and D. Filiz Unsal, "Distinguishing Constraints on Financial Inclusion and Their Impact on GDP, TFP, and the Distribution of Income," *Journal of Monetary Economics*, forthcoming, *48* (2).

de Mel, Suresh, David McKenzie, and Christopher Woodruff, "Returns to Capital in Microenterprises: Evidence from a Field Experiment," *The Quarterly Journal of Economics*, 2008, *123* (4), 1329–1372.

_ , _ , and _ , "One-Time Transfers of Cash or Capital Have Long-Lasting Effects on Microenterprises in Sri Lanka," *Science*, 2012, *335* (6071), 962–966.

Deaton, Angus, "Instruments of development: Randomization in the tropics, and the search for the elusive keys to economic development," *Proceedings of the British Academy, 2008 Lectures*, 2009, *162*, 123–160.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum, "Global Rebalancing with Gravity: Measuring the Burden of Adjustment," *IMF Staff Papers*, July 2008, *55* (3), 511–540.

Desmet, Klaus, Dávid Krisztián Nagy, and Esteban Rossi-Hansberg, "The Geography of Development," *Journal of Political Economy*, 2018, *126* (3), 903–983.

_ , **Robert E Kopp, Scott A Kulp, Dávid Krisztián Nagy, Michael Oppenheimer, Esteban Rossi-Hansberg, and Benjamin H Strauss**, "Evaluating the Economic Cost of Coastal Flooding," Working Paper 24918, National Bureau of Economic Research August 2018.

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," *American Economic Review*, April 2018, *108* (4–5), 899–934.

Donovan, Kevin, "The Equilibrium Impact of Agricultural Risk on Intermediate Inputs and Aggregate Productivity," 12 2020.

Duflo, Esther, "Grandmothers and Granddaughters: Old-Age Pensions and Intrahousehold Allocation in South Africa," *The World Bank Economic Review*, 2003, *17* (1), 1–25.

_ , **Pascaline Dupas, and Michael Kremer**, "The Impact of Free Secondary Education: Experimental Evidence from Ghana," manuscript, MIT 2017.

Egger, Dennis, Johannes Haushofer, Edward Miguel, Paul Niehaus, and Michael W Walker, "General Equilibrium Effects of Cash Transfers: Experimental Evidence from Kenya," Working Paper 26600, National Bureau of Economic Research December 2019.

Ehrlich, Dan and Robert Townsend, "Spatial Spillovers and Labor Market Dynamics: Village Financial Interventions in Thailand," manuscript, MIT 2019.

Emerick, Kyle, Alain de Janvry, Elisabeth Sadoulet, and Manzoor H. Dar, "Technological Innovations, Downside Risk, and the Modernization of Agriculture," *American Economic Review*, June 2016, *106* (6), 1537–61.

Esfahani, Saraa, "Agricultural Misallocation: Mismeasurement, Misspecification, or Market Frictions?" - Manuscript sub- mitted for publication," Technical Report, University of Notre Dame, Department of Economics 2018.

Fehr, Ernst and Lorenz Goette, "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment," *American Economic Review*, March 2007, *97* (1), 298–317.

Fieler, Ana Cecília, "Nonhomotheticity and Bilateral Trade: Evidence and a Quantitative Explanation," *Econometrica*, July 2011, *79* (4), 1069–1101.

Filmer, Deon P., Jed Friedman, Eeshani Kandpal, and Junko Onishi, "Cash Transfers, Food Prices, and Nutrition Impacts on Nonbeneficiary Children," Policy Research Working Paper Series 8377, The World Bank March 2018.

Fisher, Franklin M., "The Existence of Aggregate Production Functions," *Econometrica*, 1969, *37* (4), 553–577.

Fisher, Ronald A., "Studies in crop variation. I. An examination of the yield of dressed grain from Broadbalk.," *Journal of Agricultural Science*, 1921, *11* (72), 107–135.

Fried, Stephie and David Lagakos, "Electricity and Firm Productivity: A General-Equilibrium Approach," Working Paper 27081, National Bureau of Economic Research May 2020.

Fujimoto, Junichi, David Lagakos, and Mitchell Vanvuren, "Aggregate and Distributional Eects of 'Free' Secondary Schooling in the Developing World," manuscript, University of California at San Diego 2019.

Galle, Simon, "Competition, Financial Constraints and Misallocation:Plant-Level Evidence from Indian Manufacturing," Technical Report, University of California, Berkeley 2019.

Galor, Oded and Joseph Zeira, "Income Distribution and Macroeconomics," *Review of Economic Studies*, 1993, *60* (1), 35–52.

Giné, Xavier and Robert M. Townsend, "Evaluation of Financial Liberalization: A General Equilibrium Model with Constrained Occupation Choice," *Journal of Development Economics*, 2004, *74* (2), 269–307.

Gollin, Douglas and Christopher Udry, "Heterogeneity, Measurement Error, and Misallocation: Evidence from African Agriculture," Working Paper 25440, National Bureau of Economic Research 2019.

Greenwood, Jeremy, Philipp Kircher, Cezar Santos, and Michèle Tertilt, "An Equilibrium Model of the African HIV/AIDS Epidemic," *Econometrica*, July 2019, *87* (4), 1081–1113.

Grossman, Gene M. and Elhanan Helpman, "Quality Ladders in the Theory of Growth," *Review of Economic Studies*, 1991, *58* (1), 43–61.

Hansen, Lars P. and James J. Heckman, "The Empirical Foundations of Calibration," *Journal of Economic Perspectives*, 1996, *10* (1), 87–104.

Harberger, Arnold C., "The Measurement of Waste," *The American Economic Review*, 1964, *54* (3), 58–76.

Hausman, Jerry A., "Taxes and labor supply," in A. J. Auerbach and M. Feldstein, eds., *Handbook of Public Economics*, Vol. 1, Elsevier, 1985.

Heckman, James and Sergio Urzua, "Comparing IV with structural models: What simple IV can and cannot identify," *Journal of Econometrics*, 2010, *156* (1), 27–37.

Heckman, James J., "Building Bridges between Structural and Program Evaluation Approaches to Evaluating Policy," *Journal of Economic Literature*, June 2010, *48* (2), 356–98.

_ **and Edward Vytlacil**, "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, May 2005, *73* (3), 669–738.

Herkenhoff, Kyle, Jeremy Lise, Guido Menzio, and Gordon M. Phillips, "Production and Learning in Teams," NBER Working Papers 25179, National Bureau of Economic Research, Inc October 2018.

Hopenhayn, Hugo A, "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, September 1992, *60* (5), 1127–1150.

Hsieh, Chang-Tai and Peter Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, November 2009, *124* (4), 1403–1448.

Huggett, Mark, "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies," *Journal of Economic Dynamics and Control*, 1993, *17* (5-6), 953–969.

Imbens, Guido W., "Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009)," *Journal of Economic Literature*, June 2010, *48* (2), 399–423.

_ **and Joshua D. Angrist**, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62* (2), 467–475.

Jack, William and Tavneet Suri, "Risk Sharing and Transactions Costs: Evidence from Kenya's Mobile Money Revolution," *American Economic Review*, January 2014, *104* (1), 183–223.

Jarosch, Gregor, Ezra Oberfield, and Esteban Rossi-Hansberg, "Learning from Coworkers," NBER Working Papers 25418, National Bureau of Economic Research, Inc January 2019.

Jeong, Hyeok and Robert M. Townsend, "Sources of TFP Growth: Occupational Choice and Financial Deepening," *Economic Theory*, 2007, *32* (1), 197–221.

Joaquim, Gustavo, Robert Townsend, and Victor Zhorin, "Optimal Contracting and Spatial Competition among Financial Service Providers," 2019. MIT.

Johnson, David S., Jonathan A. Parker, and Nicholas S. Souleles, "Household Expenditure and the Income Tax Rebates of 2001," *American Economic Review*, December 2006, *96* (5), 1589–1610.

Jones, Ronald W., "The Structure of Simple General Equilibrium Models," *Journal of Political Economy*, December 1965, *73* (6), 557–572.

Kaboski, Joseph P. and Robert M. Townsend, "A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative," *Econometrica*, 2011, *79* (5), 1357–1406.

_ **and** _ , "The Impact of Credit on Village Economies," *American Economic Journal: Applied Economics*, April 2012, *4* (2), 98–133.

_ **, Molly Lipscomb, Midrigan Midrigan, and Carolyn Pelnik**, "How Important are Indivisibilities for Savings and Development? Experimental Evidence from Uganda," manuscript, University of Notre Dame 2019.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–291.

Kaplan, Greg and Giovanni L. Violante, "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, July 2014, *82* (4), 1199–1239.

Karaivanov, Alexander and Robert M. Townsend, "Dynamic Financial Constraints: Distinguishing Mechanism Design From Exogenously Incomplete Regimes," *Econometrica*, 2014, *82* (3), 887–959.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan, "Self-Control at Work," *Journal of Political Economy*, 2015, *123* (6), 1227 – 1277.

Koopmans, Tjalling C., "Measurement Without Theory," *The Review of Economics and Statistics*, 1947, *29* (3), 161–172.

Kreindler, Gabriel E., "The welfare effect of road congestion pricing," Working Paper, International Growth Center 2018.

Kremer, Michael, Jessica Leino, Edward Miguel, and Alix Peterson Zwane, "Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions," *Quarterly Journal of Economics*, 2011, *126* (1), 145–205.

Krusell, Per and Anthony A. Smith Jr., "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 1998, *106* (5), 867–896.

Kydland, Finn E. and Edward C. Prescott, "Time to Build and Aggregate Fluctuations," *Econometrica*, 1982, *50* (6), 1345–1370.

Lagakos, David, Mushfiq Mobarak, and Michael Waugh, "The Welfare Effects of Encouraging Rural-Urban Migration," Working Papers 2018–002, Human Capital and Economic Opportunity Working Group January 2018.

Leemput, Eva Van, "A Passage to India: Quantifying Internal and External Barriers to Trade," 2016. International Finance Discussion Papers 1185.

Lin, Justin Yifu, "Rural Reforms and Agricultural Growth in China," *The American Economic Review*, 1992, *82* (1), 34–51.

Lucas, Robert and Benjamin Moll, "Knowledge Growth and the Allocation of Time," *Journal of Political Economy*, 2014, *122* (1), 1 – 51.

Lucas, Robert E., "What Economists Do," Unpublished Manuscript 1988.

_ , "Ideas and Growth," *Economica*, 2009, *76* (301), 1–19.

Lucas, Robert Jr, "Econometric policy evaluation: A critique," *Carnegie-Rochester Conference Series on Public Policy*, January 1976, *1* (1), 19–46.

Lucas, Robert Jr., "On the mechanics of economic development," *Journal of Monetary Economics*, July 1988, *22* (1), 3–42.

Mahajan, Aprajit, Christian Michel, and Alessandro Tarozzi, "Identification of Time-Inconsistent Models: The Case of Insecticide Treated Nets," Manuscript, University of California at Berkeley 2019.

Marshall, Alfred, *Principles of Economics*, London: MacMillan, 1890.

Midrigan, Virgiliu and Daniel Yi Xu, "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, February 2014, *104* (2), 422–58.

Moll, Benjamin, "Productivity Losses from Financial Frictions: Can Self-financing Undo Capital Misallocation?," *American Economic Review*, October 2014, *104* (10), 3186–3221.

_ **, Robert M. Townsend, and Victor Zhorin**, "Economic development, flow of funds, and the equilibrium interaction of financial frictions," *Proceedings of the National Academy of Sciences*, 2017, *114* (24), 6176–6184.

Monte, Ferdinando, Stephen J Redding, and Esteban Rossi-Hansberg, "Commuting, Migration and Local Employment Elasticities," 2018. American Economic Review, forthcoming.

Munshi, Kaivan and Jacques Myaux, "Social Norms and Fertility Transition," *Journal of Development Economics*, 2006, *80* (1), 1–38.

Muralidharan, Karthik and Paul Niehaus, "Experimentation at Scale," *The Journal of Economic Perspectives*, 2017, *31* (4), 103–124.

_ , _ , and Sandip Sukhtankar, "General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India," NBER Working Papers 23838, National Bureau of Economic Research, Inc September 2017.

Parker, Jonathan A., Nicholas S. Souleles, David S. Johnson, and Robert McClelland, "Consumer Spending and the Economic Stimulus Payments of 2008," *American Economic Review*, October 2013, *103* (6), 2530–53.

Parker, Susan W. and Graciela M. Teruel, "Randomization and Social Program Evaluation: The Case of Progresa," *The ANNALS of the American Academy of Political and Social Science*, 2005, *599* (1), 199–219.

Paulson, Anna L. and Robert M. Townsend, "Entrepreneurship and Financial Constraints in Thailand," *Journal of Corporate Finance*, 2004, *10* (2), 229–262.

_ , _ , and Alexander Karaivanov, "Distinguishing Limited Liability from Moral Hazard in a Model of Entrepreneurship," *Journal of Political Economy*, February 2006, *114* (1), 100–144.

Paweenawat, Archawa and Robert M. Townsend, "Village Economic Accounts: Real and Financial Intertwined," *American Economic Review*, May 2012, *102* (3), 441–46.

_ and _ , "The Impact of Regional Isolationism: Disentangling Real and Financial Factors," PIER Discussion Papers 109, Puey Ungphakorn Institute for Economic Research May 2019.

Perla, Jesse and Christopher Tonetti, "Equilibrium Imitation and Growth," *Journal of Political Economy*, 2014, *122* (1), 52–76.

_ , _ , and Michael E. Waugh, "Equilibrium Technology Diffusion, Trade, and Growth," *American Economic Review*, January 2021, *111* (1), 73–128.

Piketty, Thomas, "The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing," *Review of Economic Studies*, 1997, *64* (2), 173–189.

Pomeranz, Dina, "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax," *American Economic Review*, 2015, *105* (8), 2539–2569.

Redding, Stephen J. and Esteban Rossi-Hansberg, "Quantitative Spatial Economics," *Annual Review of Economics*, 2017, *9* (1), 21–58.

Restuccia, Diego and Richard Rogerson, "Policy Distortions and Aggregate Productivity with Heterogeneous Establishments," *Review of Economic Dynamics*, 2008, *11* (4), 707–720.

Robinson, Joan, "The Production Function and the Theory of Capital," *Review of Economic Studies*, 1953, *21* (2), 81–106.

_ , "Accumulation and the Production Function," *The Economic Journal*, 1959, *69* (275), 433–442.

Romer, Paul M., "Endogenous Technological Change," *Journal of Political Economy*, 1990, *98* (5), S71–S102.

Rosenzweig, Mark and Christopher Udry, "External Validity in a Stochastic World," NBER Working Papers 22449, National Bureau of Economic Research, Inc 2016.

Sampson, Thomas, " Dynamic Selection: An Idea Flows Theory of Entry, Trade, and Growth *," *The Quarterly Journal of Economics*, 09 2015, *131* (1), 315–380.

Santaeulalia-Llopis, Raul and Diego Restuccia, "Misallocation and Productivity," NBER Working Paper 23128, National Bureau of Economic Research February 2017.

Sarto, Andres, "Recovering Macro Elasticities from Regional Data," Unpublished manuscript, Princeton University 2018.

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

Silva, Dejanir and Robert Townsend, "Risk-taking over the Life Cycle: Aggregate and Distributive Implications of Entrepreneurial Risk," manuscript, University of Illinois at Urbana-Champaign 2019.

Solow, Robert M., "Technical Change and the Aggregate Production Function," *The Review of Economics and Statistics*, 1957, *39* (3), 312–320.

Sraer, David and David Thesmar, "A Sufficient Statistics Approach for Aggregating Firm-Level Experiments," NBER Working Papers 24208, National Bureau of Economic Research, Inc January 2018.

Summers, Lawrence H., "Some skeptical observations on real business cycle theory," *Quarterly Review*, 1986, *10* (Fall), 23–27.

Suri, Tavneet, "Mobile Money," *Annual Review of Economics*, 2017, *9*, 497–520.

Todd, Petra E. and Kenneth I. Wolpin, "Assessing the Impact of a School Subsidy Program in Mexico: Using a Social Experiment to Validate a Dynamic Behavioral Model of Child Schooling and Fertility," *American Economic Review*, December 2006, *96* (5), 1384–1417.

Townsend, Robert M., *The Medieval Village Economy: A Study of the Pareto Mapping in General Equilibrium Models.*, Princeton: Princeton University Press, 1993.

_ , "Financial Systems in Northern Thai Villages," *Quarterly Journal of Economics*, November 1995, *110* (4), 1011–1046.

_ **and Kenichi Ueda**, "Financial Deepening, Inequality, and Growth: A Model-Based Quantitative Evaluation," *Review of Economic Studies*, 2006, *73* (1), 251–293.

_ **and** _ , "Welfare Gains From Financial Liberalization," *International Economic Review*, 2011, *forthcoming*, p.

_ **and Sergio S. Urzua**, "Measuring the Impact of Financial Intermediation: Linking Contract Theory to Econometric Policy Evaluation," *Macroeconomic Dynamics*, 2009, *13* (S2), 268–316.

Udry, Christopher, "Risk and Saving in Northern Nigeria," *American Economic Review*, December 1995, *85* (5), 1287–1300.

Wolf, Christian, "The Missing Intercept: A Demand Equivalence Approach," Unpublished manuscript, Princeton University 2019.

Zhang, Bin and Colin A. Carter, "Reforms, the Weather, and Productivity Growth in China's Grain Sector," *American Journal of Agricultural Economics*, 1997, *79* (4), 1266–1277.