

USING FIELD EXPERIMENTS TO UNDERSTAND THE IMPACT OF INSTITUTIONS ON ECONOMIC GROWTH

May 2024

Omar Al-Ubaydli, Faith Fatchen, and John A. List¹

ABSTRACT

Field experiments are a useful empirical tool that can be deployed in any sub-discipline – including institutional economics – to enhance the sub-discipline’s empirical insights. However, we here argue that there exist fundamental barriers to the use of field experiments in understanding the impact of institutions on economic growth. Despite these obstacles, we present some significant scholarly contributions that merit exposition, while also proposing some future methods for using field experiments within institutional economics. While field experiments may be limited in answering questions in institutional economics with macroeconomic outcomes, there is great potential in employing field experiments to answer micro founded questions.

Keywords: institutional economics; economic growth; field experiments.

JEL: B1, B5, B52, C93, D2

INTRODUCTION

In terms of the questions it seeks to answer, economics as a discipline has developed a high degree of intellectual diversity, spanning important queries such as: what is the optimal exchange rate regime for a country to pursue (Mundell, 1961)? How much redistributive activity should governments undertake (Piketty, 2014)? And what is the nature and extent of discrimination in markets (List, 2004a)? However, an area of constant scrutiny and enlightenment since the late 18th century concerns the determinants of the wealth of nations: what factors allow an economy to realize a persistently high rate of growth (Barro, 2003)? Moreover, this inquiry is decidedly motivated by practical considerations that extend well beyond an intellectual desire to know how the world works. In other words, scholars seek knowledge of the underpinnings of economic growth because they wish to use that information to help economies grow and become more prosperous.

¹ Affiliations: Al-Ubaydli: Bahrain Center for Strategic, International and Energy Studies (Derasat), and George Mason University; Fatchen: University of Chicago; List: University of Chicago and NBER. We wish to thank... for helpful comments.

Institutional economics is a notable link in a long chain of methodological and topical evolutions within the broader economics discipline. While “institutions” is a somewhat amorphous concept in economics, new institutional economics (NIE), is structured around the notion that institutions, humanly devised rules, and norms – both formal and informal – are critical to understanding economies (Menard and Shirley, 2022). NIE encompasses a diverse set of ideas, reflected by its most famous contributors include – among others – Douglass North, Oliver Williamson, and Elinor Ostrom.

Built upon the basis that institutions are fundamental, NIE approaches a diverse set of questions, diverging from old institutional economics in both the set of question and methods leveraged by scholars. Unlike old institutional economics, NIE embraces some of the assumptions of neoclassical economics – subject to augmentation of behavioral assumptions and transaction costs. NIE additionally leverages the tools of econometrics to causally study questions of economic organization and development without ignoring institutional contexts (Menard and Shirley, 2022). New institutionalists approach questions at many different levels – micro and macro – to understand the role of institutions. It is intellectually adjacent to – and intersects with – public choice economics, which analyzes many similar factors but with a less ambitious selection of outcome variables. Thus, for example, where a public choice scholar may study the impact of voting rights on proximate tax policy, an NIE scholar wants to know what the voting rights imply for long-run economic growth. In seeking to answer this question, they may wish to understand what are the formal or informal norms in place that lead to well-functioning voting system.

To some degree, the genesis of institutional economics reflects the feeling among its practitioners that classical and neoclassical growth theory have failed to yield satisfactory answers to the core questions that Adam Smith and other founders of modern economics posed. In particular, it has brought to the attention of mainstream scholars the importance of a series of previously overlooked factors, such as legal systems, social norms, power structures, unique historical contexts, and so on.

Despite the rich and ultimately constructive exchanges between institutional economics and other sub-disciplines within economics, a separate literature has advanced contemporaneously and virtually independently: the use of experimental methods in economics. For those unfamiliar with experimental economics, the method can be succinctly summarized as the estimation of causal effects of groups of explanatory variables on dependent variables through the investigating scholar’s direct manipulation of the explanatory variables, most frequently (though not definitively) via the use of randomized control. In short, experimental economists control the assignment mechanism, allowing them to answer heretofore quite complex questions using causal inference. This latter feature brings economics in line with the epistemological standards of the natural sciences by helping scholars to overcome the problems of endogeneity caused by relying on observational data (Al-Ubaydli and List, 2013).

Thus, for example, a non-experimental study of the causal effect of the interest rate on the demand for corporate credit would involve gathering data on realized interest rates and demand for corporate loans and inferring the statistical relationship between the two, aided by gathering data on control variables such as the business confidence and profit margins. In contrast, a study controlling the assignment mechanism (an experimental study) would witness the researcher directly and randomly manipulating the interest rate on corporate loans and observing the impact on demand, safe in the knowledge that any extraneous variation in factors that affect the demand for corporate loans is orthogonal to the interest rate variation and therefore does not bias the estimates of the causal effect. In the special case of natural field experiments, the environment in which the experiment takes place is the real, natural setting, and the participants are unaware of the fact that an experiment is under way (Harrison and List, 2004). This enhances the generalizability of the findings through a series of channels, most notably by eliminating the impact of Hawthorne effects, more formally referred to as experimenter demand effects (Levitt and List, 2007) as well as by controlling selection (Al-Ubaydli and List, 2013).

Having offered simple definitions of institutional economics and field experiments, it is clear that there are barriers to combining the two. The causal effects that institutional economics is interested in estimating involve explanatory variables that are generally very difficult for a researcher to manipulate, such as legal systems and social norms. Moreover, the outcome variables imply causal chains that may take years or decades to materialize, such as the impact of property rights on economic growth. However, it is possible to use field experiments to understand elements of the large puzzle that links institutions to economic growth. Yet, in doing so we caution the reader and note that we must understand the complete set of mechanisms before generalizing the insights from a part of the puzzle to the whole. Indeed, without a keen sense of how the whole puzzle fits together, including mediation and moderation, one should take great care in generalizing and scaling ideas (List, 2024). This is especially important considering the many number of causes of variation in the outcomes that accompanies many questions within the institutional economics corpus. The overarching message is that greater trust should be given to a policy that is well understood theoretically and one that has a simple mediation path. In addition, measuring key surrogates – proxies for long run outcomes, and establishing relevant surrogacy conditions, can open an entirely new area of institutional economics research that leverages field experiments.

With those goals in mind, in this chapter, we survey a selection of studies, while noting that these ultimately represent watered down versions of institutional economics. We focus on three main areas: legal systems, political systems, and social systems, where the latter can be thought of as informal social norms that also include religion (since there are only a handful of theocracies in the world at present). Considering the contextual sensitivity of the findings in these studies, and the breadth of inquiry, the literature does not yield a sufficient density of results relating to a certain question that allow for generalizable conclusions. Instead, our survey works more as an exposition of creative ways in which field experiments can be used under the umbrella of growth-centered institutional economics.

Given that our background is in field experiments, we conclude the chapter by presenting a series of ideas of future field experiments that could help scholars estimate the impact of institutions on economic growth. In some cases, our suggestions are based on adapting a previously-executed laboratory experiment to a field setting, while others involve totally novel designs. We hope that this class of open brainstorming can provide our fellow practitioners with inspiration for their own future research.

As a final remark, this chapter is written for people who have a limited background in at least one of institutional economics and field experiments, and possibly both. Given our general argument about the distinct nature of these two sub-disciplines, we believe that a large proportion of our colleagues satisfy this criterion.

1. PROLOGUE: RUNNING FIELD EXPERIMENTS IN INSTITUTIONAL ECONOMICS

1.1. DEFINITIONS

Normally, an economics paper would not require a lengthy section on semantics. In the case of this chapter, it is likely that readers lack a background in at least one of the two concepts of field experiments and institutional economics, and a full understanding of the two is required for our main argument. Readers familiar with the definitions can skip to later sections.

1.1.1. Field Experiments

Field experiments are an epistemological method that spans the social sciences, while institutional economics is a combination of (1) a field of research that interests economists and (2) an associated set of epistemological techniques. *Prima facie*, field experiments can be used in institutional economics in the same manner that cross-country regressions can be used to study monetary policy, or surveys can be used to study how much people value protecting the environment. As we demonstrate below, upon careful consideration of the defining characteristics of institutional economics, it becomes evident that field experiments can rarely be applied to the study of institutions in the case of macro outcomes such as economic growth, but are of use to new institutionalists seeking to understanding the micro-foundations of institution centered questions.

We begin by offering a clear definition of field experiments, based on Harrison and List (2004). In the social sciences, an empirical causal investigation is a process of gathering data to estimate the effect of varying a group of explanatory variables on the value of a dependent variable, where at least some of the explanatory and dependent variables are a form of human behavior. The effect being estimated is known as the casual effect of those explanatory variables on the dependent variable (Heckman, 2005).

In general, there are two types of data: observational and experimental (Al-Ubaydli and List, 2013). In the case of observational data, the data are generated organically and independently of the investigating researcher's efforts. In the case of experimental data, the variation in the explanatory variables is induced by the investigating researcher, who has control over the assignment mechanism; that is, who is placed in treatment and control. In this manner, an experiment usually involves the researcher inducing random variation in the explanatory variables (known as a randomized control trial), but this is not necessarily a defining characteristic; the key is the researcher's control of the explanatory variables, and their exercising that control as part of an effort to acquire knowledge. For example, in many of Vernon Smith's seminal laboratory experiments he controlled the environment but did not use randomization to identify a treatment effect. In this case, he used theory to generate the relevant counterfactual rather than data.

The simplest setting that social scientists can use for running experiments is a conventional laboratory, where the human behavior being studied is that of college students, the decision-making environment is abstract, and the rules governing their behavior are imposed. Conventional laboratory experiments transition into artefactual field experiments when the participating students are replaced by a non-standard subject pool, such as a representative sample of the general population. This then mutates into a framed field experiment when the decision-making is given real context. Critically, in all three of these classes, the participants – students or otherwise – consciously enroll in the study and are aware of the unnaturally high levels of scrutiny of their behavior by the overseeing research team.

The key distinction to be drawn is between these types of experiment and natural field experiments, commonly shortened to “field experiments”. In a natural field experiment, the setting is the natural environment where participants occupy natural roles within the setting, and – most importantly – the participants are unaware of their participation in an experiment and therefore there is no selection into or out of the experiment. While each class of experiment possesses a set of advantages and disadvantages relating to factors such as feasibility, ethicality, and epistemic value, a consensus is emerging that natural field experiments yield data that is more likely to provide generalizable insights. That means conclusions that apply to the real, high-stakes contexts that social scientists are most interested in studying, especially when their goal is to acquire knowledge that can be leveraged to manipulate the environment to the benefit of humans (Al-Ubaydli and List, 2013).

While a full exposition of the reasons for this is beyond the scope of this chapter, it can be concisely stated as follows: enrolling people to take decisions in abstract or unnatural settings coupled with unnaturally high levels of scrutiny can lead to distorted behavior that yields limited insights about behavior in natural settings over the entire population of interest with natural (and therefore low) levels of scrutiny (Levitt and List, 2007). A simple illustration is running an experiment to estimate the impact of an unexpected positive wealth shock on immediate charitable contributions. In a conventional laboratory experiment, people will donate an average of around 40% of their

fortuitously acquired wealth, which is a large overstatement of donations in an isomorphic real setting (List, 2007).

The ability of experiments to overcome the problem of endogeneity has made them an increasingly popular tool in social sciences. In the case of using field experiments within institutional economics, we aim to delineate where precisely we see potential applications.

1.1.2. New Institutional Economics

The precursor to defining new institutional economics is defining institutions. The most famous definition is provided by Douglass North: “Institutions are the rules of the game in a society or, more formally, are the humanly devised constraints that shape human interaction” (North, 1990). Acemoglu (2008) expands by noting that this definition of institutions implies two things: they are humanly designed (as opposed to geographical phenomena, for example); and the constraints they impose on human behavior operate through incentives. Some of the core institutions that economists study include property rights, which fall under the broader umbrella of legal systems. Other formal institutions include markets, auction houses, online brokerages, and other systems for exchanging goods and services. One can add to these informal institutions, such as social norms.

This brings us to the definition of NIE. As alluded to in the introduction, institutional economists are united on the fundamental statement that institutions are critical to understand economies and societies (Menard and Shirley, 2022). However, the approach and methods differ in important ways from what is referred to as old institutional economics. Both centered around the importance of institutions, but NIE leverages the theoretical and analytical tools of neoclassical economics. With such tools at their disposal. NIE scholars position themselves for empirical testing.

To complete the definition of new institutional economics, it is important to expand upon the methods within the field. New institutionalists, unlike old, leverage formal modeling, rejecting some assumptions of neoclassical economics by introducing three tenets: transaction costs, bounded rationality, and opportunism (Menard and Shirley, 2022). Notably, the field of institutional economics exhibits a commitment to narrative arguments like those seen in non-quantitative approaches to anthropology, history, political science, and sociology, featuring extensive expositions of instructive anecdotes, as well as formal case studies. One of the leading names associated with this approach is Elinor Ostrom, whose ethnography lies far away from the core methods found in Economics 101. However, as a whole, NIE embraces more of the core methods found in other circles such as formal modeling and modern econometric techniques.

1.2. BARRIERS TO RUNNING FIELD EXPERIMENTS

The protracted nature of these definitions is necessary for one of the main arguments that this chapter makes, which is that field experiments are of limited use in understanding the macro-level

outcomes often associated with institutional economics. The primary cause of this purported incompatibility with the “bigger” questions is that if one looks at an institutional economics paper dealing with such outcomes, and one identifies the key explanatory and dependent variables, a field experiment is usually infeasible, as controlling the assignment mechanism is difficult. In this regard, it is similar to many questions within mainstream macroeconomics.

Running field experiments in institutional economics is difficult for a several reasons. First, the nature of the explanatory and dependent variables in institutional economics research typically makes a field experiment infeasible. In the case of the explanatory variables, many are too complex to be manipulated by a researcher, or require an unrealistic volume of financial resources. For example, scholars cannot – in a field setting – engineer wholesale changes in property rights (Davidson et al., 2018), judicial independence (Voigt et al., 2015), constitutional checks and balances (Fatás and Mihov, 2013), and so on. This is especially true because the sorts of institutions being studied are the ones that have an impact on some aggregate measure of societal welfare, and so represent pillars of society and the economy.

In the case of the dependent variables, the arguments inherent in an institutional economics paper usually imply outcomes that take years, decades, or even centuries to materialize, which would call upon the use of surrogates. As List (2025) discusses, the feasibility of a field experiment in such longitudinal settings revolves around generating treatments that affect surrogates and are the sole mediators of the treatment effect. In cases with high causal density, this is a difficult task because, as explained above, institutional economics is concerned with macro-societal outcomes such as the level of economic growth over a long period of time, the level of per capita income, the level of inequality, and so on. Thus, even if the requisite control over the explanatory variables were to be secured, the practical constraints of generating useful surrogates is a difficult chore.

Furthermore, when the outcomes of interest are aggregate measures of societal welfare creates problems of statistical power that are often insurmountable. Researchers cannot hope to have an institutional intervention where the unit of observation is a city or country, and to have control over the explanatory variables in enough cities or countries to generate satisfactory levels of statistical power. In fact, in many conventional institutional economics studies, the sample size is exactly one, as in the case study of Haiti in the period 1796-1820 (Ferrero, 2020). Moreover, when operating with units of observation as large as cities or countries, the problem of extraneous noise is amplified, making the need to secure a large sample more acute at a time when the feasibility is low. For example, culture and beliefs are critical mediating variables for the impact of institutions on economic growth, yet they can rarely be controlled for in a study because they are intangible and difficult to observe.

The challenge of sample size is made even more difficult by the importance of historical context to the arguments that some scholars make in institutional economics. A unique event in a certain location in history is often argued to be the driving force behind a pivotal institutional development, with the concomitant impact of economic growth, as in the case of England’s Glorious Revolution

of 1688 (North and Weingast, 1989). Reproducing these unique historical circumstances in a contemporary field experiment is impossible.

Finally, if one were to overcome the aforementioned feasibility constraints, there remain considerable ethical ones because when answering these macro questions the unit of analysis is generally an entire society. The process of randomization will mean that unlucky societies – be they cities or even entire countries – are potentially on the path of relatively low aggregate welfare outcomes for a significant period of time, which is morally unacceptable without research equipoise. For example, in the case of a study of judicial independence (Voigt et al., 2015), are we to condemn certain cities or countries to a system of kangaroo courts for the purposes of demonstrating how useful independent courts are? Informed consent is ruled out by natural field experiments, but even if informed consent were somehow compatible with a natural field experiment, it is unclear how one would secure it for an entire society.

Combining these factors highlights that the researcher interested in conducting field experiments in the area of institutional economics faces deep challenges, much as in many areas of macroeconomics (yet, for good examples of lab experiments in the area of macroeconomics please see the work of John Duffy and Charles Noussair, and coauthors). Phenomena such as inflation, unemployment, exchange rate regime, fiscal balances, and so on are ill-suited for field experiments for the same reasons as in institutional economics: they are not feasible, one cannot collect data, and they are ethically unacceptable. Lest we be accused of hypocrisy, two of us wrote a review paper on field experiments in unemployment (Al-Ubaydli and List, 2019), yet if one examines the papers cited within the review, all of them are field experiments looking at a microeconomic aspect of unemployment. None have a dependent variable of aggregate unemployment, meaning that they are not macroeconomic papers. This foreshadows our next section, which argues that field experiments are of use in understanding such micro-foundations in institutional economics.

We close this section with a small anecdote that offers some validation for our claim. The experimental economics community (Economic Science Association) has an email listhost that members use to pose questions, advertise conferences, and so on. Members regularly send emails wherein they request references in a certain topic because they are authoring a survey paper and want to ensure maximum coverage. Fortunately, members are very cooperative, sharing the papers that they are familiar with. (The social norm is for the source of the original query to collate the suggested papers and then share a final list with the listhost.) In our case, we sent an email indicating our writing this chapter and asking for references. We received one reference only.

1.3. HOW FIELD EXPERIMENTS CAN CONTRIBUTE TO INSTITUTIONAL ECONOMICS

Despite the barriers that we argue that the scientist must overcome to conduct field experiments in institutional economics, we believe that there several approaches in which field experiments can contribute to our understanding of institutional economics. We focus on two here.

The first is that – as intimated above – there exists a small but not insignificant minority of research questions in institutional economics that are potentially suitable for field experiments. In particular, the two scholars who shared a Nobel prize for institutional economics, Elinor Ostrom and Oliver Williamson, have conducted significant research where the dependent variable is a microeconomic phenomenon, and where the explanatory variables can be feasibly manipulated by an investigating researcher.

For example, Williamson (1967) analyzes optimal firm size as a function of various parameters within the organization, such as the firm’s hierarchical structure and levels of compliance. Williamson (1988) examines how the characteristics of assets determine the optimality of debt versus equity finance for projects. Schlager and Ostrom (1992) studies how the specific ownership structure of common-property resources affects usage incentives and outcome sustainability. Dietz et al. (2003) proposes resolving non-local problems through a variety of strategies including dialogue between stakeholders and allowing for experimental solutions.

Naturally, one could make the argument that knowing how to structure a business well or how to manage a common-property resource effectively incrementally adds to our knowledge of how to make the economy grow. This is exactly what institutionalist studying micro foundations do. This branch of NIE is critical to the fusion of NIE and field experiments.

The leads us to the second way in which field experiments can contribute to institutional economics is by using them to understand microeconomic elements of the main economic growth-related institutions that scholars study in institutional economics, either in spirit or from a technical perspective. As mentioned above, this is the approach that we used in Al-Ubaydli and List (2019), and it will constitute the bulk of the survey that we will conduct in the next section. In this regard, Figure 1 and Figure 2 provide a good overview of what some of those key growth-related institutions are, including property rights, special interest groups, the judiciary, state capacity, and so on.

At this point, it is worth clarifying what our approach has in common with the recently published (at the time of writing) paper by Callen et al. (2023), and how our approach differs. Callen et al. (2023) is a survey of field experiments *on* institutions, which reveals an immediate difference: we are surveying field experiments in institutional economics. Their survey of the non-experimental literature leads them to the conclusion that certain institutional changes have been credibly demonstrated to have had persistent effects on the path of long-run development. Their focus then turns to using field experiments to study what determines institutional change, with a special emphasis on the value of measuring beliefs. We will cite many of the papers that Callen et al. (2023) cover in their highly valuable survey.

2. A SURVEY OF SELECTED FIELD EXPERIMENTS IN INSTITUTIONS-MEDIATED ECONOMIC GROWTH

At the outset of this survey, we drew readers' attention to the fact that we plan to cover selected papers based on our subjective view on what will be most beneficial to readers. This survey is not a systematic or comprehensive literature review.

2.1. LEGAL SYSTEMS

When analyzing the impact of legal institutions on economic growth, the most frequently cited element is property rights. As we argued above, randomizing property rights appears to be infeasible and morally questionable in certain cases. Nevertheless, Fabri (2021) documents how this actually happened in rural Benin: the government formulated land tenure reforms that consisted of:

“... formalizing customary tenure over land characterized by collective property and informal possession. This formalization was achieved by recording the set of rights each individual holds over land parcels, creating precise boundaries demarcation, and generating official land maps stored in a public repository. With the reform, the right-holders acquired formal, legally recognized use rights that can be traded, used as collateral, and defended in court against contenders. Thus, the reform produced a substantial shift toward a system of well-defined and individually assigned land rights.”

Note that the unit of randomization was a rural village. As expected, almost a decade after the experiment took place, the study was unable to link the institutional reform to a central macroeconomic variable such as economic growth. However, Fabri (2021) ran a series of artefactual and framed field experiments and found that people benefiting from the reforms exhibited more cooperative and trusting behavior within those experiments. These findings are consistent with the broad-based positive effect of property rights that scholars in institutional economics often argue.

Other studies are unable to randomize property rights, but they do randomize awareness and enforcement of property rights. Subramanian and Kumar (2024) randomly nudge Indian farmers into acquiring formal land titles, and they find that it increases agricultural investments, thereby raising various agricultural performance measures. The benefiting households also experience a net increase in per capita consumption. Aberra and Chemin (2021) randomize access to the free services of a lawyer for two years in rural Kenya. The result was more secure property rights, thereby inducing greater investment and access to credit. Similar results were found in the context of a Western country (the US) by Greiner and Pattanayak (2011), who randomized access to legal assistance for people facing eviction. Blair et al. (2019) is a field experiment in Liberia where teams of superior (better trained and equipped) policy officers had recurring patrols randomized

by rural village, resulting in more secure property rights and less crime, though without any significant impact on trust in the police, courts, or government.

In the adjacent domain of common property resources, Cavalcanti et al. (2010) conducts a field experiment in six fishing communities in Brazil to investigate the impact of having a participatory management approach on people's willingness to individually contribute to solving the problems. One of the structural challenges they face is a small number of units of randomization: they randomly allocate three of the six participating villages into a participatory approach to managing the common resources, and three into the control. They find the participatory approach – via a variety of mechanisms including beliefs about the willingness of others to cooperate – results in higher levels of cooperation.

Exploring beyond property rights to broader studies of the legal system, field experiments have been deployed quite widely (Green and Thorley, 2014), though primarily as a means of assessing the impact of various factors on the performance of the legal system, without necessarily linking the performance of the legal system to aggregate societal welfare. For example, many field experimental studies demonstrate the important result that increasing perceived law enforcement for certain crimes (in this case via visually salient enforcement) leads to a decrease in the incidence of those crimes (Dur and Volland, 2019). Peyton et al. (2019) use a field experiment with a large urban police department and find that the alternative to hard enforcement – community-oriented policing – is effective in improving police legitimacy and residents' willingness to cooperate with police. Findley et al. (2015) use a field experiment to vary the properties of a request for anonymous incorporation, and they find that offers to “pay a premium” (illicit payments) reduce compliance, while alerting companies to international law has no effect. Kassin et al. (2019) find that recording crime suspects during interrogations had no effect on the important outcome variables such as the propensity to confess or to waive Miranda, demonstrating the importance of features in which a formal system is operating in.

The courtroom also attracts the attention of researchers using field experiments. What must surely be one of the earliest legal field experiments was conducted by Heuer and Penrod (1989), which found that instructing jurors prior to the evidence part of the trial and providing them with a written copy of the judge's instructions has no discernible impact on a variety of outcome measures, such as ability to recall the judge's instructions, trial length, or juror satisfaction. The effect of juror note-taking was investigated in a field experiment by Heuer and Penrod (1994) and they also find no effect. Thompson et al. (2022) randomly vary the content of legal Wikipedia articles to show that judges apparently use Wikipedia to inform their judgements. Ho and Ouellette (2020) show that the expansion of scientific peer review in patent determinations slightly improved search efforts and citations to the non-patent scientific literature, albeit at a considerable resource cost.

One of the oft-proposed methods for improving the functioning of the legal system – namely decreasing corruption – is greater transparency. Peisakhin and Pinto (2010) induce random variation in access to ration cards to demonstrate that India's freedom of information law is almost

as effective as bribery in helping the poor secure access to basic public services. Worthy et al. (2017) compare freedom of information requests in English parish councils to simple asks and find that the former are significantly more effective. Karim (2020) induces transparency via household visits by police in Liberia, and she uses a field experiment to demonstrate a positive effect on trust in the legal system.

One huge sub-literature that intersects with field experiments in legal systems is field experiments on discrimination. Bertrand and Duflo (2017) provide a comprehensive survey, including several law-related findings such as: people illegally discriminate, sometimes consciously, other times not; the factors that contribute to higher levels of illegal discrimination; and so on. Another comparable large sub-literature examines tax compliance, but we defer an exposition to the section below that examines state capacity.

There are many other legal field experiments one could expand upon, but the pattern is clear. First, field experiments about legal systems are extremely diverse, and there is a high degree of context dependence. Accordingly, it is difficult to derive *general* conclusions beyond the ones that theory would point to *ex ante*, such as higher enforcement works, and transparency helps. Field experiments have helped rigorously demonstrate the sensitivity of the behavior of judicial actors (judges, jurors, lawyers, etc.) to seemingly innocuous factors, though running these kinds of experiments involves traversing an ethical minefield given the stakes associated with erroneous judicial decisions. In the context of property rights, the limited evidence does support the Northian argument that property rights are conducive of economic growth.

Though these papers provide little to no direct evidence of the effect of institutions on macroeconomic variables such as economic growth, they do provide micro-founded insights such as how specific institutions work. These examples are exemplary as to where we see field experiments contributing to institutional economics.

2.2. POLITICAL SYSTEMS

The attributes of the political system have been a very popular area for field experimentation and the literature is several orders of magnitude larger than that in the legal arena, possibly because of the greater ability to experimentally manipulate key explanatory variables. However, just as in the case of legal systems, the studies – by our criteria – drift from the core of institutional economics because they are unable to link the interventions to aggregate societal welfare measures via surrogates of economic growth.

Following Fukuyama (2011) and Callen et al. (2023), when considering aggregate welfare, we divide the political system’s characteristics into two primary spheres: those that relate to the state’s capacity, and those that relate to accountability within the state. Starting with the former, field experiments have studied two main determinants of state capacity.

The first is a functional one, which is the government's ability to collect tax revenues, which it can then allocate to the various responsibilities that the state assumes. The reviews by both Green and Thorley (2014) and Callen et al. (2023) provide excellent thorough surveys, which we selectively explore here. The simplest interventions include sending letters that increase the salience of audits (Kleven et al., 2011), or that remind people of their responsibility to pay taxes (Hallsworth et al., 2017). This latter example and others, such as Torgler (2013), build on the larger social sciences literature whereby moral suasion in the absence of material incentives is used to induce prosocial behavior (Ferraro and Price, 2013). In general, these informational interventions find statistically significant but economically modest positive impacts, though given that the cost of the intervention is usually trivial, it typically passes a cost-benefit test.

Notice that feasibility and potential ethical concerns limit scholars' ability to directly investigate the most obvious way of increasing compliance, i.e., increasing enforcement. Instead, as described above, they typically use workarounds, such as making existing enforcement levels more cognitively salient. A notable exception is Andrade et al. (2013), which explores various ways of motivating informal businesses in Brazil to formally register (a stepping-stone toward paying taxes). They find that the informational treatments have no effect, while increasing inspections does substantially increase formal registration of businesses.

The second vein of state capacity-related field experimental research concerns increasing citizens' trust in government, as trust is widely regarded to be a great economic lubricant (Fukuyama, 2011). When people trust the government, a smaller volume of costly coercive effort is required to induce cooperation with the state, and to thereby effect the policies that the state prefers. Callen et al. (2023) report the findings of several field experiments where researchers worked with governments to increase citizens' trust in their governments. Haim et al. (2021) estimated the impact of an outreach program that helped Philippine citizens positively update their beliefs about the competence of their government, resulting in them providing the state with more timely information about the Covid-19 pandemic.

In another interesting piece of research in the second vein, Blattman et al. (2023) used a field experiment to demonstrate – in villages with certain attributes – improvements in government legitimacy and decreases in emergency calls in response to the state deploying full-time liaisons to assist with intra-community conflict resolution and to foster communication with the government. Likewise, Blair et al. (2022) study a separate intervention in Colombia designed to “... replace rebel governance by harnessing complementarities between state and communal authorities and by improving security and justice provision in areas once dominated by FARC, the country's largest rebel group”. They also reported improved local conflict resolution and increased citizens' trust in some state institutions.

We next consider the other sphere of political systems, accountability, which has been a particularly vibrant area of research for scholars using field experimental methods. The two relevant review papers we came across, Green and Thorley (2014) and Callen et al. (2023),

together cite well over 20 studies conducted during the last 15 years examining a wide variety of aspects of accountability. Interestingly, the designs are fundamentally homogeneous: the scholars conjecture that a certain factor makes people more willing and able to hold their principals accountable, or makes the principals more willing and able to be held accountable by the people they serve. This factor can be individual level or group level, and it is randomly varied, with the outcome variables being either accountability activity (voting in elections, effort to communicate with senior officials, etc.) or some measure of satisfaction with the accountability process (trust in principals, quality of elected officials, etc.).

For example, Casey et al. (2021) intervene in elections in Sierra Leone by randomly changing the degree to which candidates are directly nominated by voters (at-large nominations) versus nominated by party officials. The latter, indirect route leads to candidates who are less likely to satisfy voters' preferences. These results echo the earlier findings of Beath et al. (2016), which randomly varied the electoral institutions of 250 Afghan villages to compare at-large elections with district-bound ones, finding that the at-large system produced superior outcomes in terms of politician quality (as measured by having completed high school). Olken (2010) drew comparable conclusions based on a field experiment in Indonesia.

Novel technological interventions are a popular choice for field experiments because when one starts with a baseline where the new technology is absent, and the technology is likely to be helpful, it is ethically palatable to randomly introduce that new technology due to research equipoise. Grossman et al. (2014) do this by making it cheaper for Ugandan citizens to send text messages to their representatives, and they find a high level of take-up that allowed marginalized populations to gain better access to politicians than would have been possible via traditional channels. Marx et al. (2021) also experiment with text messages, but as a voter mobilization tool in Kenya, and they find increased turnout but decreased trust. This latter finding is important in demonstrating the context-sensitivity of interventions in environments as complex and riddled with path-dependencies as national elections.

One recent area of research is the effect of partisan media on the quality of political institutions. Conroy-Krutz and Moehler (2015) compare home partisan to cross-cutting (opposing partisan) radio broadcasts in Ghana, finding that the latter were persuasive and worked to moderate voter views. Gerber et al. (2010) induce partisanship in the US by reminding voters that they need to be registered with a party to vote in that party's primaries. The result is a strengthening of partisan identity and more partisan voting behavior. Porter and Rogowski (2018) demonstrate that partisan email writers during local elections were more likely to elicit a response from local election clerks. Both of these latter partisanship studies further show how complex political systems are in practice, as it is unclear whether the behavioral shift induced by partisanship is something that improves or diminishes the performance of political institutions.

There are also many papers that examine the effect of civic education on the political system, again due to the relative absence of feasibility and ethical barriers to such experiments. Some occur in

developing economies, such as Finkel and Lim (2021) and Finkel et al. (2023); while others occur in advanced economies, such as Claassen and Monson (2015) and Green et al. (2011). Brock-Petroshius (2023) is an interesting addition to this latter group that experimentally explores the impact of deep canvassing and finds that is effective in changing attitudes. In general, studies of the effect of civic education on the performance of the political system tend to find positive effects on self-reported measures of civic-mindedness, and on forms of political engagement.

The sub-topics we have covered here are a mere selection of what can be found in a diverse literature. We encourage those interested to explore the additional topics in Callen et al. (2023), such as political inclusiveness, political accountability interventions relating to politicians, the role of social media, participation on protests, and so on.

2.3. INFORMAL SOCIAL SYSTEMS INCLUDING RELIGION

Today, in early 2024, theocracies play a very limited role in the global political system, yet religion remains important to daily life, thereby transforming it into a quasi-informal social system. Institutional economics pays great attention to informal social systems, with one of the most famous examples being the Weberian argument regarding the persistent effect of the Protestant work ethic on productivity in the 20th century, well after the formal separation of church and state.

Before exploring the empirical evidence, it is worth briefly recapping the theory that guides how we should expect religion to affect aggregate welfare. In primitive societies, the belief in supernatural deities that reward righteousness and punish anti-social behavior can help overcome social dilemmas such as the free-rider problems, as adherents will fear divine retribution lest they behave in an anti-social manner (Fukuyama, 2011). Religion is also typically the first source of rule of law, where rule of law is defined as senior officials (including the head of state and head of government) being subject to the same restrictions as ordinary people; as these senior officials who have dominion over the masses must abide by religious restrictions. At various points in history, religious institutions have amassed political power independently of the state, helping them to act as a check on the state abusing its power. However, the key weakness that religion suffers from as an institutional source of superior performance is that it creates an in-group bias and exclusionary tendencies toward out-group members that may ultimately result in a net negative impact on aggregate societal welfare (Norenzayan and Shariff, 2008). This ambiguity – which puts religion in stark contrast to many other important institutions studied by institutional economists – enhances the value of empirical research in understanding its impact, and makes field experiments a potentially invaluable tool.

The main problem field experimental researchers face, however, is that adherence to a religion cannot be readily manipulated, nor can other important explanatory variables such as faith, edicts, acts of worship, and so on (Wright, 2018). Researchers are restricted to interventions such as priming religion in an individual and seeing how it affects their behavior, or manipulating the

observability of an individual's religion and seeing how it affects how others behave toward them. This latter approach has generated many studies of religious-based discrimination, such as Wright et al. (2013) and Valfort (2020); see Bertrand and Duflo (2017) for a more comprehensive list. These types of studies make salient the aforementioned in-group out-group problems, as religion ends up working as a vehicle that creates grounds for socially destructive animus in a wide variety of settings including labor markets, housing, socializing, charitable contributions, and so on. It is fair to say that religious discrimination studies account for a very large proportion of religious field experiments, primarily due to the ease of executing the basic audit design to which they mostly adhere.

The priming design is often used to demonstrate the potentially positive effect of religion on prosocial behavior. For example, Duhaime (2015) finds that the Muslim call for prayer induces a transient increase in prosocial behavior in Moroccan markets, while Al-Ubaydli et al. (2023) find that religious text messages induce higher levels of prosocial conservation in electricity consumption among residents of Qatar. Feldhaus et al. (2022) demonstrate that announcing the Catholic church's support for climate protection efforts induces higher charitable donations in the German Catholic Convention.

A unique study that likely has limited generalizability but that is highly instructive within the original context is the field experiment conducted by Clingingsmith et al. (2009), which involved randomizing participation in the greater pilgrimage in Islam, the *hajj*. The design leverages the fact that most Muslims only perform it once (if at all) during their lifetimes, and that it constitutes an extremely significant component of their cumulative religious activity. The authors find that it has a lasting, positive effect on tolerance, thereby inducing prosocial behavior which captures the importance of taking part in the institutional proceeding.

When it comes to informal social systems, institutional economics looks well beyond religion, too, especially as the world becomes more secular. There are far too many social norms to enumerate, though some of the more salient ones (Elster 1989) include reciprocity (rewarding those who treat you well and punishing those who treat you badly), kin selection (treating your close kin better than non-kin members), in-group favoritism (behaving more favorably toward fellow members of a group you identify with, such as an ethnic or confessional group), and inequality aversion (being averse to an unequal distribution of resources within your group). As with the specific case of religious-based social norms, social norms in general are difficult for investigating researchers to manipulate because they are formed over decades (and sometimes even millennia), forcing researchers to use priming and salience. Yet, social norms acting as moderators of effects represent a key area of scholarly research (Levitt and List, 2007).

Additionally, it is worth affirming the potential value of experimental empirical work in studying social norms, as there exists theoretical uncertainty over the growth implications of several key social norms. Perhaps the most famous is the perpetual debate about whether collectivist or

individualist cultures yield superior economic performance (Kyriacou, 2016). Some of the key social norms that have been studied using field experiments are as follows.

Bergquist et al. (2020) is one of many field experiments that leverages social norm-based moral suasion to induce prosocial behavior. It follows the typical design of providing descriptive and injunctive social normative information, in this case relating to the importance of protecting the environment. The result was higher pro-environmental donations. Ferraro and Price (2013) provide information on people's water consumption compared to that of their neighbors and find that this creates an incentive for people to conserve water. Senderey et al. (2022) leverage social norms of responsibility toward others to induce higher vaccination rates in a field experiment. These examples are merely illustrative, as there exist numerous papers with analogous designs, many of which fall under the banner of "nudge" economics (Thaler and Sustein, 2009).

Yoeli et al. (2013) takes an indirect approach to studying the social norm of indirect reciprocity, i.e., the propensity to help people who help others, and punish those who mistreat others. A key moderating variable that determines the effectiveness of indirect reciprocity as a social norm is the observability of people's actions. The authors use a field experiment to manipulate observability in a utility company's blackout-reduction program, and they find increasing observability has a strong positive effect on people's prosocial behavior. Mujcic and Leibbrandt (2018) also use a field experiment to study indirect reciprocity, and show that it is effective in inducing prosocial behavior in the context of drivers in traffic permitting others to switch lanes.

Bandiera et al. (2020) use a field experiment to estimate the interaction between the degree of individualism-collectivism and the effectiveness of piece-rate incentives offered to workers. They find that piece rates are more effective for people who come from societies with an individualistic social norm, but it should be noted that individualism (or any prime thereof) is not randomized in this field experiment; instead, it is the strength of the incentives, and so this is technically not a randomized control field experiment in the context of estimating the impact of a social norm.

Scholars have also used informational interventions in field experiments to help correct misconceptions that might perpetuate out of date social norms. Bursztyn et al. (2020) do this in the case of the social norm of women staying out of the labor force in Saudi Arabia in accordance with their husbands' wishes. The authors show that when young married men learn that they have systematically underestimated acceptance of women working among their peer groups, they themselves become significantly more supportive of their own wives seeking employment. In the long-run, this can yield significant aggregate economic benefits as the labor market potential of women is unleashed.

The previous section on political systems also features several interventions designed to improve social norms related to conflict management in populations that may suffer from weak conflict management norms and/or weak state enforcement. For example, Blattman et al. (2014) showed how community education helped to establish better dispute resolution norms, though in such

contexts, it is difficult to distinguish between a social norm (people have an improved mutual understanding about how to resolve conflicts) and human capital (people have acquired hard conflict resolution skills).

This selection of studies is – as in the case of political systems – far from exhaustive due to the wide range of social norms that have been scientifically investigated using field experiments. Though the range of social norms that could plausibly cause higher economic growth is considerably narrower, institutional economists have allocated a smaller volume of their research efforts to studying such norms, as compared to formal legal and political institutions. This means that there are many opportunities to deepen our understanding of the effect of social norms, a point that we return to below in the epilogue. From the field experiments that we have surveyed, the breadth of inquiry and the importance of context together work to limit the ability to draw general conclusions, especially when combined with the structural barriers to running field experiments that we covered in previous sections. There is suggestive field experimental evidence that religion can be a force for the good, but the research designs preclude the general equilibrium analysis required to explore the major weakness that religions suffer from as a system for organizing behavior: the bias against out-group members. Continuing to explore these issues as the tools for running field experiments improve holds much promise.

3. EPILOGUE: WHERE DO FIELD EXPERIMENTS HOLD MOST PROMISE FOR INSTITUTIONAL ECONOMICS AND ECONOMIC GROWTH?

As field experimentalists, despite arguing that “true institutional economics” is difficult to address with field experiments, we remain hopeful that some of the barriers that we described can be overcome with a combination of technological advancement and lateral thinking. In this section, we first consider ideas grounded in institutional economics laboratory experiments, and then proceed to consider some more novel designs. We conclude the section by discussing the issue of mixed methods.

3.1. IDEAS GROUNDED IN LABORATORY EXPERIMENTS

Within institutional economics there is an evergreen debate concerning individualism versus collectivism, decentralization versus central control, markets versus hierarchies, socialism versus capitalism, and so on. Narrative, ethnographic, and empirical studies based on observational data have failed to resolve the debate, while field experiments have (to the best of our knowledge) failed to make a contribution. However, many relevant laboratory experiments have been run, including Al-Ubaydli et al. (2013) and Choi and Storr (2022), which supraliminally prime market experience among experimental participants and then observe the effect on trust. With a sufficient budget,

these studies could be translated into field experiments. While the impact on truly aggregate welfare outcomes could not be observed and surrogates might be difficult to deliver, the design could be applied to small groups that have to live in comparative isolation for extended periods of time, such as military squads during training, generating informative data. More generally, as seen in the religion-related field experiments, any laboratory experiment that relies on priming or salience can typically be translated into a field experiment in a straightforward manner.

A deep behavioral economics literature that has developed concerning decentralization versus central control revolves around whether preferences are defined over consumption levels or changes in consumption (Kahneman and Tversky, 1979). In a positive sense, the value disparity has considerable relevance. For example, such preferences call into question commonly held interpretations of indifference curves, make cost-benefit analysis illegitimate, change the procedure necessary to resolve damage disputes, and cause a reconsideration of several of our other deep intuitions, such as the invariance result of Coase which relies on Hicksian equivalent surplus and Hicksian compensating surplus to be roughly equivalent (List, 2004b). While the theoretical paper of Kahneman and Tversky (1979) helped to set up the field of behavioral economics, Kahneman et al.'s (1990) laboratory experiment struck a deep empirical chord. The work has spawned several laboratory experiments that can be translated into a field experiment in a straightforward manner, and indeed that process has already begun (List, 2004b).

As an example, scholars took the baton from the important laboratory work in the 1980s and 1990s and used field experiments to show that the endowment effect is not a fundamental characteristic of preferences that is immutable by market experience (see, e.g., List 2003). Lab and field studies replicated List's experience results (e.g., Engelmann and Hollard 2010; Sutter 2021) and new insightful theories were introduced (e.g., Kőszegi and Rabin 2006).

Another leading issue in institutional economics is the impact of property rights on long-run growth. Kimbrough et al. (2008) explore this issue in a laboratory experiment featuring a somewhat abstract trading game, where inducing a history of secure property rights encourages more impersonal trade in the long run and hence higher levels of aggregate societal welfare. Here, the nature of the intervention is a protracted historical experience, and so it is much more difficult than priming experiments to translate into a field experiment. One option is massively multiplayer online role playing games (MMORPGs), which can act as a bridge between the realism of the real world and the pliability of virtual worlds. Experiments relating to trade have been run in such environments (De Sousa and Munro, 2012), and while they are unquestionably less generalizable than field experiments in non-virtual settings, these are the sort of epistemic tradeoffs that inevitably arise when tackling difficult questions. More generally, virtual worlds are a likely underleveraged source of scientific knowledge in economics, and their potential value will probably increase as virtual reality technologies continue to improve.

3.2. GREENFIELD FIELD EXPERIMENTS

The difficulties described above in conducting field experiments relating to informal social systems, especially religion, are not insurmountable, and Wright (2018) provides some sage strategies for overcoming the feasibility and ethical challenges relating to the manipulation of social norms. The most promising is randomizing access to oversubscribed programs relating to the social norm, as Clingingsmith et al. (2009) did in the case of the pilgrimage that Muslims perform typically once during their lives. Ideally, as in the case of Islam's greater pilgrimage, the program would represent a highly significant component of the formation of that social norm. For many social norms, including religious ones, no intervention approaches the significance of the greater pilgrimage in Islam, but there are some unexplored (to the best of our knowledge) options.

One example is the programs designed to socialize immigrants so that they adopt the social norms of the country they are emigrating to. For example, Norway has a cultural orientation program that helps foster the integration of resettled refugees in local communities. This multi-month program is ripe for a field experiment, especially if one is to collect data on dependent variables of import. These would include individual level outcomes for the immigrants, and society level ones for the settlements that receive many immigrants. Norway likely does not have enough settlements for a high-powered study but were a similar program to be introduced to the US, then over time, enough data at the aggregate level could potentially be collected that ultimately adds to our knowledge about the impact of informal institutions on economic growth.

If one affords us with a slightly expanded definition of institutional economics, then for our second greenfield example we would advocate going down the human capital route. This is because human capital is important for long run growth, and schools are institutions that provide human capital. Field experiments have been used to explore the education production function and early skill formation, which are both linked to long run growth of an economy.

Consider Schweinhart et al.'s (1993) analysis of the Perry Preschool Project, a preschool program launched in the 1960s. While the Perry Preschool Project began as early childhood framed field experiment to understand whether high-quality education benefits at-risk children, it soon became a longitudinal study documenting the lasting impact of the intervention on adult outcomes. Researchers, including James Heckman and colleagues, have reported that the project produced dramatic benefits for academic achievement, graduation rates, and earnings. It also lowered rates of crime, delinquency, teenage pregnancy, and welfare dependency when compared to the control group. We view future work using field experiments to provide such insights as key contributors to a more broadly defined institutional economics.

One feature that this example raises that has been mentioned in various places of this chapter is how to provide scientific insights into long-run problems. As aforementioned, economic experiments typically focus on estimating short run substitution effects. Yet, in many cases understanding the long-term effects of an intervention is critical from a policy and scientific

perspective. This is especially true within institutional economics, where key questions revolve around effects measured in the long run, such as economic growth. In such cases, it is fundamental to understand causal pathways.

To leverage causal pathways in the long-run, a first key consideration is an outcome's causal density. The causal density of an outcome variable refers to how well the determinants of the outcome are understood. An outcome with low causal density is one where the factors determining the outcome are well understood. For example, farmers have a near-perfect understanding of the determinants of crop yields, yielding a low causal density. In contrast, the determinants of individual labor market outcomes or economic growth are much less well understood and thus have high causal density. For example, future employment and wage outcomes are a function of the numerous and interconnected inputs by multiple people and institutions.

When exploring outcome variables with low or high causal density, a key consideration is whether a feasibly observed proxy variable, called a statistical surrogate, can be measured in the short-run that estimates long-run outcomes. While the use of surrogates within economics is a developing area of research, we view its development as a key unlock that can usher in a new wave of field experiments in institutional economics. This is because development of attractive surrogates has the potential of moving the research agenda from one of short run measurements to a viable approach to examine macro outcomes of interest in institutional outcomes, such as long term economic growth, or some alternative measure of aggregate welfare, such as per capita income, societal inequality, or political stability. As soon as relevant surrogates are developed, we envision field experiments as an attractive tool that will experience substantial growth amongst institutional economists.

3.3. MIXED METHODS

While we are optimistic about the progress that can be made in using field experiments based on introspection about new designs and by leveraging new technologies, we remain realistic that the fundamental difficulties discussed throughout this chapter are unlikely to be overcome in the near future. Accordingly, we wish to extoll studies that use mixed methods to try to minimize the weaknesses associated with each type of data.

Orthodox economists have – in our opinion – been needlessly hostile to narrative and ethnographic arguments made within the profession and by other social scientists. While we are sympathetic to the view that a rigorous randomized control field experiment adds more to our knowledge of institutional economics than a purely narrative argument, we firmly subscribe to the view that a combination is superior. This is especially important because – unlike the physical phenomena that natural scientists study in the laboratory – the human phenomena at the heart of institutional economics exhibit very high degrees of context dependence. Accordingly, those looking to advance their knowledge would benefit considerably from understanding the historical context.

Ethnographic work in tandem with field experiments, especially in-depth interviews, can be vital for generating hypotheses to test with field experiments, and for solving empirical puzzles.

In this regard, we are particularly supportive of taking the ethnographic and narrative work of scholars such as Elinor Ostrom and combining it with the sort of field experiments we surveyed in this chapter, i.e., the ones that do not present complete answers to institutional economics questions, but that contribute part of the answer based on a watering down of the original research question. Surveys are also growing in popularity in economics generally and can play an important role as a complement to imperfect field experiments in institutional economics. Blattman et al. (2021) use a variety of empirical methods – quantitative and qualitative; interviews, surveys, observational data and quasi-experiments – all conducted in the field to study institutions in Medellín, Colombia; and they combine this with theoretical modeling. We regard this approach to be a snapshot of the future of empirical institutional economics.

CONCLUSION

Institutional economics is field of economics that gets to the heart of what the discipline is ultimately about: helping societies to improve their living standards. As with many foundational research questions in economics, there are no easy answers, and scholars are forced to combine a selection of imperfect epistemic tools to advance our understanding of the phenomena that underpin economic growth. The importance of historical context means that for the most part, the literature has relied on narrative and ethnographic methods, combined with the econometric analysis of observational data. While we collectively know a lot more today than a century ago about how institutions such as property rights, rule of law, and cooperative norms impact economic growth, there remain huge gaps in our understanding. The problems are made more challenging by the fact that the world is constantly evolving, and the determinants of economic prosperity are always mutating.

For the most part, as stand-alone contributions, field experiments face difficulties in institutional economics at the macro level: they are frequently infeasible, potentially unethical, and impossible to statistically power and find useful surrogates. Nevertheless, field experiments are a promising tool for understanding small pieces of a bigger puzzle such as economic growth with the caveat that the researcher must ensure an understanding of how the puzzle fits back together (List 2024), and as already demonstrated by Williamson and Ostrom for understanding explanatory variables and outcomes that are feasibly manipulated.

The last decade in particular has seen an explosion of field experiments in the study of relevant factors such as property rights, state capacity, political accountability, and so on. Despite this progress, the literature remains nascent, and the studies remain context dependent, limiting the reach where we can make general conclusions or offer a list of “regular” empirical findings. There

also remains a significant gap to bridge regarding the ultimate impact of field experimental interventions on aggregate welfare metrics, most notably economic growth.

For this reason, we urge scholars to overcome any excessive affinity they might have with certain methodological approaches, and to embrace mixed methods that include field experiments and complementary approaches as the ideal path forward. For example, insisting on randomized field experiments to study the critical question of how property rights affect development yields limited work, and we discuss one study conducted in Benin. Rather than putting our proverbial intellectual eggs all in the experimental basket, when it comes to institutional economics, we argue that it makes sense to place weight on the conclusions drawn from narrative, ethnographic, and observational studies as well.

REFERENCES

- Aberra, A. and Chemin, M., 2021. Does legal representation increase investment? Evidence from a field experiment in Kenya. *Journal of Development Economics*, 150, p.102612.
- Acemoglu, D., 2008. Growth and institutions, in Durlauf, S. and Blume, L. (eds) *The New Palgrave Dictionary of Economics*, 2nd edition.
- Al-Ubaydli, O., Cassidy, A.W., Chatterjee, A., Khalifa, A. and Price, M.K., 2023. *The power to conserve: a field experiment on electricity use in Qatar* (No. w31931). National Bureau of Economic Research.
- Al-Ubaydli, O., Houser, D., Nye, J., Paganelli, M.P. and Pan, X.S., 2013. The causal effect of market priming on trust: An experimental investigation using randomized control. *PloS one*, 8(3), p.e55968.
- Al-Ubaydli, O. and List, J.A., 2013. *On the generalizability of experimental results in economics: With a response to Camerer* (No. w19666). National Bureau of Economic Research.
2019. How natural field experiments have enhanced our understanding of unemployment. *Nature human behaviour*, 3(1), pp.33-39.
- Bandiera, O., Dahlstrand-Rudin, A. and Fischer, G., 2020. Incentives and Culture: Evidence from a Multi-Country Field Experiment. *Manuscript in preparation*, pp.917-962.
- Barro, R.J., 2003. Determinants of economic growth in a panel of countries. *Annals of economics and finance*, 4, pp.231-274.
- Beath, A., Christia, F., Egorov, G. and Enikolopov, R., 2016. Electoral rules and political selection: Theory and evidence from a field experiment in Afghanistan. *The Review of Economic Studies*, 83(3), pp.932-968.
- Bergquist, M., Nyström, L. and Nilsson, A., 2020. Feeling or following? A field-experiment comparing social norms-based and emotions-based motives encouraging pro-environmental donations. *Journal of Consumer Behaviour*, 19(4), pp.351-358.
- Bertrand, M. and Duflo, E., 2017. Field experiments on discrimination. *Handbook of economic field experiments*, 1, pp.309-393.
- Blair, R.A., Karim, S.M. and Morse, B.S., 2019. Establishing the rule of law in weak and war-torn states: Evidence from a field experiment with the Liberian National Police. *American Political Science Review*, 113(3), pp.641-657.

- Blair, R.A., Moscoso-Rojas, M., Castillo, A.V. and Weintraub, M., 2022. Preventing rebel resurgence after civil war: A field experiment in security and justice provision in rural colombia. *American Political Science Review*, 116(4), pp.1258-1277.
- Blattman, C., Duncan, G., Lessing, B. and Tobón, S., 2021. *Gang rule: Understanding and countering criminal governance* (No. w28458). National Bureau of Economic Research.
- Blattman, C., Duncan, G., Lessing, B. and Tobón, S., 2023. Civilian alternatives to policing: Evidence from Medellín's community problem-solving intervention Operación Convivencia.
- Blattman, C., Hartman, A.C. and Blair, R.A., 2014. How to promote order and property rights under weak rule of law? An experiment in changing dispute resolution behavior through community education. *American Political Science Review*, 108(1), pp.100-120.
- Brock-Petrossius, K., 2023. *Changing Dominant Carceral Attitudes: A Community Organizing Field Experiment* (Doctoral dissertation, UCLA).
- Bursztyn, L., González, A.L. and Yanagizawa-Drott, D., 2020. Misperceived social norms: Women working outside the home in Saudi Arabia. *American economic review*, 110(10), pp.2997-3029.
- Callen, M., Weigel, J.L. and Yuchtman, N., 2023. *Experiments about Institutions* (No. w31964). National Bureau of Economic Research.
- Casey, K., Kamara, A.B. and Meriggi, N.F., 2021. An experiment in candidate selection. *American Economic Review*, 111(5), pp.1575-1612.
- Cavalcanti, C., Schläpfer, F. and Schmid, B., 2010. Public participation and willingness to cooperate in common-pool resource management: A field experiment with fishing communities in Brazil. *Ecological Economics*, 69(3), pp.613-622.
- Choi, G.S. and Storr, V.H., 2022. The market as a process for the discovery of whom not to trust. *Journal of Institutional Economics*, 18(3), pp.467-482.
- Claassen, R.L. and Monson, J.Q., 2015. Does civic education matter?: The power of long-term observation and the experimental method. *Journal of Political Science Education*, 11(4), pp.404-421.
- Clingingsmith, D., Khwaja, A.I. and Kremer, M., 2009. Estimating the impact of the Hajj: religion and tolerance in Islam's global gathering. *The Quarterly Journal of Economics*, 124(3), pp.1133-1170.
- Coase, R. H., 1937. The nature of the firm. *Economica*, 4(4), 386-405. <https://doi.org/10.1111/j.1468-0335.1937.tb00002.x>

- Conroy-Krutz, J. and Moehler, D.C., 2015. Moderation from bias: A field experiment on partisan media in a new democracy. *The Journal of Politics*, 77(2), pp.575-587.
- Davidson, S., De Filippi, P. and Potts, J., 2018. Blockchains and the economic institutions of capitalism. *Journal of Institutional Economics*, 14(4), pp.639-658.
- De Andrade, G.H., Bruhn, M. and McKenzie, D.J., 2013. A helping hand or the long arm of the law? Experimental evidence on what governments can do to formalize firms. *Experimental Evidence on What Governments Can Do to Formalize Firms (May 1, 2013). World Bank Policy Research Working Paper*, (6435).
- De Sousa, Y.F. and Munro, A., 2012. Truck, barter and exchange versus the endowment effect: Virtual field experiments in an online game environment. *Journal of Economic Psychology*, 33(3), pp.482-493.
- Dietz, T., Ostrom, E. and Stern, P.C., 2003. The struggle to govern the commons. *science*, 302(5652), pp.1907-1912.
- Duhaime, E.P., 2015. Is the call to prayer a call to cooperate? A field experiment on the impact of religious salience on prosocial behavior. *Judgment and Decision making*, 10(6), pp.593-596.
- Dur, R. and Vollaard, B., 2019. Salience of law enforcement: A field experiment. *Journal of Environmental Economics and Management*, 93, pp.208-220.
- Elster, J., 1989. Social norms and economic theory. *Journal of economic perspectives*, 3(4), pp.99-117.
- Engelmann, Dirk, and Guillaume Hollard. 2010. "Reconsidering the Effect of Market Experience on the 'Endowment Effect.'" *Econometrica* 78 (6): 2005–19. <https://doi.org/10.3982/ECTA8424>.
- Fabbri, M., 2021. Property rights and prosocial behavior: Evidence from a land tenure reform implemented as randomized control-trial. *Journal of Economic Behavior & Organization*, 188, pp.552-566.
- Fatás, A. and Mihov, I., 2013. Policy volatility, institutions, and economic growth. *Review of Economics and Statistics*, 95(2), pp.362-376.
- Feldhaus, C., Gleue, M. and Löschel, A., 2022. Can a Catholic institution promote sustainable behavior? Field experimental evidence on donations for climate protection. *Journal of Behavioral and Experimental Economics*, 98, p.101855.
- Ferraro, P.J. and Price, M.K., 2013. Using nonpecuniary strategies to influence behavior: Evidence from a large-scale field experiment. *Review of Economics and Statistics*, 95(1), pp.64-73.

- Ferrero, M., 2020. Accidental socialism: a natural experiment in Haiti 1796-1820. *Journal of Institutional Economics*, 17(3), pp.393-409.
- Findley, M.G., Nielson, D.L. and Sharman, J.C., 2015. Causes of noncompliance with international law: A field experiment on anonymous incorporation. *American Journal of Political Science*, 59(1), pp.146-161.
- Finkel, S.E. and Lim, J., 2021. The supply and demand model of civic education: evidence from a field experiment in the democratic Republic of Congo. *Democratization*, 28(5), pp.970-991.
- Finkel, S.E., Neundorff, A. and Rascon Ramirez, E., 2023. Can online civic education induce democratic citizenship? Experimental evidence from a new democracy. *American Journal of Political Science*.
- Fukuyama, F., 2011. *The origins of political order: From prehuman times to the French Revolution*. Farrar, Straus and Giroux.
- Gerber, A.S., Huber, G.A. and Washington, E., 2010. Party affiliation, partisanship, and political beliefs: A field experiment. *American Political Science Review*, 104(4), pp.720-744.
- Grabowski, R., 2008. Modes of long-run development: Latin America and East Asia. *Journal of Institutional Economics*, 4(1), pp.25-50.
- Green, D.P., Aronow, P.M., Bergan, D.E., Greene, P., Paris, C. and Weinberger, B.I., 2011. Does knowledge of constitutional principles increase support for civil liberties? Results from a randomized field experiment. *The Journal of Politics*, 73(2), pp.463-476.
- Green, D.P. and Thorley, D.R., 2014. Field experimentation and the study of law and policy. *Annual Review of Law and Social Science*, 10, pp.53-72.
- Greiner, D.J. and Pattanayak, C.W., 2011. Randomized evaluation in legal assistance: What difference does representation (offer and actual use) make?. *Yale Law Journal*, 121.
- Grossman, G., Humphreys, M. and Sacramone-Lutz, G., 2014. "I wld like u WMP to extend electricity 2 our village": On Information Technology and Interest Articulation. *American Political Science Review*, 108(3), pp.688-705.
- Haim, D., Ravanilla, N. and Sexton, R., 2021. Sustained government engagement improves subsequent pandemic risk reporting in conflict zones. *American Political Science Review*, 115(2), pp.717-724.
- Hallsworth, M., List, J.A., Metcalfe, R.D. and Vlaev, I., 2017. The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of public economics*, 148, pp.14-31.

- Harrison, G.W. and List, J.A., 2004. Field experiments. *Journal of Economic literature*, 42(4), pp.1009-1055.
- Heckman, J.J., 2005. The scientific model of causality. *Sociological methodology*, 35(1), pp.1-97.
- Heuer, L. and Penrod, S.D., 1989. Instructing jurors: A field experiment with written and preliminary instructions. *Law and Human Behavior*, 13(4), pp.409-430.
- Heuer, L. and Penrod, S., 1994. Juror notetaking and question asking during trials: A national field experiment. *Law and Human Behavior*, 18(2), pp.121-150.
- Ho, D.E. and Ouellette, L.L., 2020. Improving Scientific Judgments in Law and Government: A Field Experiment of Patent Peer Review. *Journal of Empirical Legal Studies*, 17(2), pp.190-223.
- Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2): 263–91. <https://doi.org/10.2307/1914185>. 1990. Prospect Theory: An Analysis of Decision under Risk. *Rationality in Action: Contemporary Approaches*. New York, NY, US: Cambridge University Press.
- Kahneman, D., Knetsch, J.L. and Thaler, R.H., 1990. Experimental tests of the endowment effect and the Coase theorem. *Journal of political Economy*, 98(6), pp.1325-1348.
- Karim, S., 2020. Relational state building in areas of limited statehood: Experimental evidence on the attitudes of the police. *American Political Science Review*, 114(2), pp.536-551.
- Kassin, S.M., Russano, M.B., Amrom, A.D., Hellgren, J., Kukucka, J. and Lawson, V.Z., 2019. Does video recording inhibit crime suspects? Evidence from a fully randomized field experiment. *Law and human behavior*, 43(1), p.45.
- Kimbrough, E.O., Smith, V.L. and Wilson, B.J., 2008. Historical property rights, sociality, and the emergence of impersonal exchange in long-distance trade. *American Economic Review*, 98(3), pp.1009-1039.
- Kleven, H.J., Knudsen, M.B., Kreiner, C.T., Pedersen, S. and Saez, E., 2011. Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, 79(3), pp.651-692.
- Kőszegi, Botond, and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences*." *The Quarterly Journal of Economics* 121 (4): 1133–65. <https://doi.org/10.1093/qje/121.4.1133>.
- Kyriacou, A.P., 2016. Individualism–collectivism, governance and economic development. *European Journal of Political Economy*, 42, pp.91-104.
- Levitt, S.D. and List, J.A., 2007. What do laboratory experiments measuring social preferences reveal about the real world?. *Journal of Economic perspectives*, 21(2), pp.153-174.

List, J. A. 2003. “Does Market Experience Eliminate Market Anomalies?” *The Quarterly Journal of Economics* 118 (1): 41–71. <https://doi.org/10.1162/00335530360535144>.

2004a. The nature and extent of discrimination in the marketplace: Evidence from the field. *The Quarterly Journal of Economics*, 119(1), pp.49-89.

2004b. Testing Neoclassical Competitive Theory in Multilateral Decentralized Markets. *Journal of Political Economy* 112 (5): 1131–56. <https://doi.org/10.1086/422564>.

2007. On the interpretation of giving in dictator games. *Journal of Political economy*, 115(3), pp.482-493.

2024. Optimally generate policy-based evidence before scaling. *Nature*, 626(7999), pp. 491-99.

2025. Experimental Economics: Theory and Practice. University of Chicago Press, forthcoming.

Marx, B., Pons, V. and Suri, T., 2021. Voter mobilisation and trust in electoral institutions: Evidence from Kenya. *The Economic Journal*, 131(638), pp.2585-2612.

Ménard, C. and Shirley, M.M., 2022. *Advanced introduction to new institutional economics*. Edward Elgar Publishing.

Mujcic, R. and Leibbrandt, A., 2018. Indirect reciprocity and prosocial behaviour: evidence from a natural field experiment. *The Economic Journal*, 128(611), pp.1683-1699.

Mundell, R.A., 1961. A theory of optimum currency areas. *The American Economic Review*, 51(4), pp.657-665.

Norenzayan, A. and Shariff, A.F., 2008. The origin and evolution of religious prosociality. *science*, 322(5898), pp.58-62.

North, D. C., 1990. *Institutions, institutional change and economic performance*. Cambridge University Press.

North, D.C. and Weingast, B.R., 1989. Constitutions and commitment: the evolution of institutions governing public choice in seventeenth-century England. *The journal of economic history*, 49(4), pp.803-832.

Olken, B.A., 2010. Direct democracy and local public goods: Evidence from a field experiment in Indonesia. *American political science review*, 104(2), pp.243-267.

Peisakhin, L. and Pinto, P., 2010. Is transparency an effective anti-corruption strategy? Evidence from a field experiment in India. *Regulation & Governance*, 4(3), pp.261-280.

- Peyton, K., Sierra-Arévalo, M. and Rand, D.G., 2019. A field experiment on community policing and police legitimacy. *Proceedings of the national Academy of sciences*, 116(40), pp.19894-19898.
- Piketty, T., 2014. *Capital in the twenty-first century*. Harvard University Press.
- Porter, E. and Rogowski, J.C., 2018. Partisanship, bureaucratic responsiveness, and election administration: Evidence from a field experiment. *Journal of Public Administration Research and Theory*, 28(4), pp.602-617.
- Schlager, E. and Ostrom, E., 1992. Property-rights regimes and natural resources: a conceptual analysis. *Land economics*, pp.249-262.
- Schweinhart, Lawrence J. 1993. “What the High/Scope Perry Preschool Study Reveals about Developmental Transitions and Contextual Challenges of Ethnic Males.” <https://eric.ed.gov/?id=ED371257>.
- Senderey, A.B., Ohana, R., Perchik, S., Erev, I. and Balicer, R.D., 2022. Encouraging COVID-19 vaccination through behaviorally informed reminders: Results from a national randomized field experiment in Israel. *Behavioral Science & Policy*, 8(1), pp.25-32.
- Subramanian, A. and Kumar, P., 2024. Property rights, factor allocation and household welfare: Experimental Evidence from a Land Titling Program in India. *Journal of Development Economics*, 167, p.103238.
- Sutter, Matthias. 2021. “Does Market Experience Eliminate Market Anomalies? (By John A. List).” In *The Art of Experimental Economics*, 176–81. Routledge.
- Thaler, R.H. and Sunstein, C.R., 2009. *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.
- Thompson, N., Flanagan, B., Richardson, E., McKenzie, B. and Luo, X., 2022. Trial by Internet: A Randomized Field Experiment on Wikipedia’s Influence on Judges’ Legal Reasoning.
- Torgler, B., 2013. A field experiment in moral suasion and tax compliance focusing on underdeclaration and overdeduction. *FinanzArchiv/Public Finance Analysis*, pp.393-411.
- Tucker, C.E., 2014. Social networks, personalized advertising, and privacy controls. *Journal of marketing research*, 51(5), pp.546-562.
- Valfort, M.A., 2020. Anti-Muslim discrimination in France: Evidence from a field experiment. *World Development*, 135, p.105022.
- Voigt, S., Gutmann, J. and Feld, L.P., 2015. Economic growth and judicial independence, a dozen years on: Cross-country evidence using an updated set of indicators. *European Journal of Political Economy*, 38, pp.197-211.

Williamson, O.E., 1967. Hierarchical control and optimum firm size. *Journal of political economy*, 75(2), pp.123-138.

Williamson, O.E., 1988. Corporate finance and corporate governance. *The journal of finance*, 43(3), pp.567-591.

Worthy, B., John, P. and Vannoni, M., 2017. Transparency at the parish pump: A field experiment to measure the effectiveness of freedom of information requests in England. *Journal of Public Administration Research and Theory*, 27(3), pp.485-500.

Wright, B.R., 2018. Field experiments in religion: A dream whose time has come. *Journal for the Scientific Study of Religion*, 57(2), pp.193-205.

Wright, B.R., Wallace, M., Bailey, J. and Hyde, A., 2013. Religious affiliation and hiring discrimination in New England: A field experiment. *Research in Social Stratification and Mobility*, 34, pp.111-126.

Yoeli, E., Hoffman, M., Rand, D.G. and Nowak, M.A., 2013. Powering up with indirect reciprocity in a large-scale field experiment. *Proceedings of the National Academy of Sciences*, 110(supplement_2), pp.10424-10429.