Can Economics Become More Reflexive?

Exploring the Potential of Mixed-Methods

Vijayendra Rao

Abstract

This paper argues that Economics can learn from Cultural Anthropology and Qualitative Sociology by drawing on a judicious mix of qualitative and quantitative methods to become more "reflexive." It argues that reflexivity, which helps reduce the distance between researchers and the subjects of their research, has four key elements: *cognitive* *empathy,* the analysis of *narratives* (potentially enhanced by machine learning), understanding *process*, and *participation* (involving respondents in research). The paper provides an impressionistic and non-comprehensive review of mixed-methods relevant to development economics and discrimination to illustrate these points.

Can Economics Become More Reflexive?[1]

Exploring the Potential of Mixed-Methods

Vijayendra Rao

Introduction

There are artificial distinctions of method that limit us as social scientists. We economists think of ourselves as quantitative social scientists. Cultural anthropologists and some sociologists see themselves as ethnographers who rely on participant observation and qualitative methods. The reasons for this divide are a vestige of disciplinary history, how scholars are trained and assessed for tenure. By limiting ourselves to one set of methods, we limit the questions we can ask and constrain ourselves in the data available to answer them. The methods lead the questions, rather than the other way around.

It did not begin this way. Charles Booth, whose magnificent multi-volume *Life and Labour of the People in London* (1892-97) helped set the stage for all social science research, triangulated between data from household surveys, participant observation and open-ended interviews to create detailed "poverty maps" of London. These maps drew on data, both quantitative and qualitative, that Booth and his team collected from every one of London's four million residents to visualize where the poor and the rich lived. The maps drew a finely detailed picture of households and their living standards, the nature of their work, and many other characteristics that Booth grouped into three categories: Employment, "Circumstances" (illness, infirmity, family size, etc.) and "Questions of Habit" (what he called problems of "drunkenness and thriftlessness"). The task took seventeen years to complete (1886-1903), filled seventeen volumes and had a tremendous influence on economic and social policy, and on social science. (For more on Charles Booth's work, see Mary Morgan's (2019) excellent book.[2])

Over the course of the 20th century, most social scientists moved away from this grounded, multi-method process of enquiry. Economics, in its quest toward greater scientific precision, became almost entirely focused on the analysis of quantitative data. Other less homogenous disciplines like sociology had sub-disciplines that focused on either qualitative or quantitative methods, while anthropology split into (at least) two fields – cultural anthropology emphasizing participant observation and ethnography, and physical anthropology emphasizing the analysis of various types of quantitative data from archeological digs, genetic analysis, and experiments. Psychology moved from observational studies into an entirely experimental direction. These differences in method were a primary characteristic of how the social sciences distinguished themselves from one another.

However, even a cursory look at recent social science journals shows that things are changing. Disciplinary differences are increasingly getting bridged. Economics has taken a behavioral turn (Thaler, 2016), sociology and political science are getting increasingly interested in mixing qualitative and quantitative methods (Small 2011, Humphreys and Jacobs 2015), psychology is moving towards a mix of experimental, qualitative and survey methods (Paluck, 2010), and mixed methods are being actively used by development researchers and practitioners (e.g. Kanbur and Shaffer, 2007). We social scientists have expanded the questions we ask and seem to be moving towards a world where methods are subservient to questions. Consequently, we are increasingly mixing methods.

This paper provides a selective overview of mixing quantitative and qualitative methods – which is still a rarity in economics. I argue that qualitative methods offer a powerful set of tools that are of central concern to economists, and particularly for economists interested in questions of discrimination and development. They help reduce the distance between researchers and the subjects of their research and can help introduce an element of *reflexivity* into economics. The paper does not attempt a full-fledged review of mixed methods, whether in economics or in other disciplines, but rather tries to illustrate various ways in which a judicious combination of qualitative and quantitative methods and tools can result in a more empathetic and richer understanding of human behavior. The examples which I will employ are also impressionistic rather than comprehensive. I will hew closely to topics that are of interest

---

[2] The full set of maps and some of Booth's "notebooks" can be explored here: https://booth.lse.ac.uk/map.

to economists who study development (which is my primary focus of interest), and draw on work from across the world, including my own.

Economics, Ethnography and Reflexivity

At the risk of some oversimplification, we could argue that econometrics focuses on two broad issues. The first is causal inference, strengthened over the last three decades by the extraordinary advances made under the credibility revolution. The second, like most of applied statistics, is about understanding and overcoming the challenges of making robust inferences about larger populations from sample data.

But, despite the powerful tools that have emerged over the years under these two important umbrellas, something is left missing. The data employed by an econometrician is often collected by someone else. Even when the econometrician is responsible for collecting it, they usually delegate it to a survey firm focused on finetuning the questions, making sure the power calculations are done properly, and deploy research assistants to supervise and monitor the data collection and data entry process. All these steps are necessary to collect data at scale, but the very process of collecting information from many people results in a distance between the researcher and the subjects of their research. This accentuates the preexisting social and cultural distance that exists between the researcher and the researched, which is particularly acute for those who study discrimination or development where the researcher is almost always from a different social class, and usually from a different race, caste, and cultural background from those they research. The topics picked by a researcher, consequently, are more likely to reflect their lived reality – their reading of the literature, their notion of what kinds of topics will further their careers – rather than the perspectives of the people they are studying.

Ethnography is at the other of this spectrum. The focus is usually on one community, studied intensely by one person who actively participates in and observes the life of the community over an extended period. To illustrate this, let me begin with brief summaries of two ethnographies that employ participant observation – the primary method of cultural anthropologists, and of qualitative sociologists trained in the Chicago school of urban sociology.

The anthropologist Bhrigupathi Singh, in his wonderful ethnography, *Poverty and the Quest for Life* (2015) draws on one and a half years of fieldwork with the Sahariya, a community classified as a "primitive tribe" who live in a remote and extremely arid part of the Indian state of Rajasthan. The Sahariya are former bonded laborers who are considered much poorer than the Dalit ("Scheduled Caste") Chamar community with whom they sometimes share settlements. Singh delves deep into the spiritual lives of the Sahariya trying to understand and describe how they conceive of "well-being," and show how different it is from the view of philosophers and social scientists. His descriptions do not ignore the standard analytical categories of caste, class, power, the state, and structural inequality, but Singh attempts to draw the reader into questioning those categories as he tries to understand how the Sahariya see themselves. He questions the tropes describing them in previous work as bound by "superstition" and "subject to supernatural beliefs," which leads one government report to describe their lives as so desolate as to "make their living unthinkable." Singh's effort is to try to understand this better, to shed light on – as he says – "What makes life unthinkable?" (page 1). He shows us how it relates to how the Sahariya experience the government in everyday life, the prejudices that they encounter, and their response to this which draws on their rich spiritual life which is very different from dominant religious practices in the region. This leads them to be seen as different, and divergent from the development paradigm, which reinforces the discrimination they face both by society and the state.

Singh's description of his research process is worth quoting because it is a nice summary of the ethnographic method. (page 284-285) "Ethnographic method may be seen as a two-phase process….In

the first phase, one is a hunter-gatherer, pursuing targets, collecting impressions…The next stage of labor is of a more settled cultivator, as we move from impressions to expressions. You plow through what you have gathered, jostling with others who faced similar phenomena and arrived at different thoughts. When impressions are organized and attached to concepts, they turn into thoughts and expression. Concepts are seeds from which impressions grow into thoughts. Our concepts may also be limiting and force us to convey much less than what we saw and felt.” Thus, ethnography is a method that obsesses with process rather than outcomes, that seeks to describe rather than measure, and attempts to reduce the distance between researcher and researched via a process of deep empathetic understanding, rather than to maintain (what anthropologists see as) a fictive analytic distance.

This process of participating and observing has long been the fundamental tenet of ethnography, which I will next illustrate with an older and influential work by the sociologist Elijah Anderson. In his ethnography *Streetwise*, Anderson (1990) distills fourteen years of fieldwork[3] (from 1975-1989), to study interactions between whites and blacks in the Philadelphia neighborhood he lived in,[4] which he called the “Village,” and the neighboring, largely black, area that he called “Northton.” Drawing on Erving Goffman’s work on Stigma (Goffman, 1960) and strategic interaction (Goffman, 1969) he drew a picture of how the Reagan-era cuts to public spending, and the decline of the city’s industrial base, led to shifts in how the two neighborhoods interacted with one another.

The Village is racially mixed, mainly populated by middle-class whites with a few middle-class blacks – largely academics and researchers - and a few yuppies. Northton was almost entirely black consisting of “stable working-class black families, well-attended churches…and young people eager to ‘make something of themselves.’” (Page 238-239). But to people in the Village (both black and white), Northton had the reputation of being beset by “classic urban ills: drugs, crime, illiteracy, poverty…families on welfare…” Anderson says that while there was an outward culture of “civility” and tolerance in the Village, they were wary of “strangers with whom they must share the public space.” Whites were wary of blacks, and gave them “extra scrutiny,” and everyone saw the streets as a “jungle,” especially at night, when “all cats are grey” and everyone seems threatening. All of this challenged the Village’s vision of itself as a civil and tolerant community. Villagers had little “worthwhile knowledge” about black areas and tended to form stereotypes. Blacks in this “picture become the kinds of people they rather not have close to them,” and people whom they “try desperately to avoid.” This lack of familiarity with black culture resulted in “prescriptions and proscriptions” about public behavior that were structured to minimize conflict with blacks. And blacks, who happened to find themselves in the Village often tried to “prove themselves respectable to others they encounter.”

Anderson’s book provides a glimpse of a neighborhood just before the gentrifying processes of the 1990s began, which priced out black communities from their traditional neighborhoods, and shows how racial difference actually plays out among people who generally think of themselves as anti-racist and tolerant. Moreover, Anderson describes how poor black men must negotiate stereotypes as they navigate their way towards making a living in a world where jobs are scarce. It is entirely complementary to the large body of work on the economics of discrimination, both empirical and theoretical, and yet more alive, more grounded, more reflective of the lived reality of both blacks and whites. It describes how statistical discrimination might work but demonstrates that it has strategic and interactional implications which, at least at the time of the book’s writing, were not part of the economics literature. And it is a wrenching narrative of the disproportionate impact that Reagan-era cuts and industrial decline had on black communities socially, psychologically, and on their economic lives.

---

[3]Combining several hundred interviews with observations from the streets, bars, laundromats, brunches, and parties, among other places.

[4] I lived in the same neighborhood from 1985 to 1990, which is perhaps one reason why this work resonates so deeply with me.

It is impossible for me to do justice to these rich ethnographies in a couple of paragraphs, they are thick with narrative insight, description, ideas, and theory. What they have, which economics lacks, is a method that necessitates an intimate relationship with the field. Participant observation attempts to minimize the distance between the researcher and the subjects of their research, and to be acutely conscious of how the researchers' social class might position how they may see and interpret things, and to bring that *reflexivity* into their analysis. Central to this is the notion of what the sociologist Mario Small (2018) has called "cognitive empathy," "the ability to understand a person's predicament as they understand it."

What participant observation lacks, however, is representative data – it focusses on a case: a village, a neighborhood, a community, with which it deeply engages, but it does not have a robust way of demonstrating that the findings are representative of a larger population. For me, re-reading Anderson's book more than thirty years after I first read it, his findings are now part of received wisdom; they have been replicated many times and were a central theme even of iconic television shows like The Wire, which is a measure of the book's validity. But at the time when Anderson was writing the book, he was simply trying to do his best to describe what he observed and experienced in one community, at one period in its history, filtered through his sociological training and his reading of theory, rather than focus on making a generalizable statement about the urban United States.

There is, therefore, a trade-off between the finely filigreed detail of a good ethnography and the focus on a very small sample that it requires, and the strong emphasis in survey-oriented fields like empirical economics on making statistically generalizable statements by analyzing data from large representative samples. The sociologist Michael Burawoy (1998), in a brilliant paper on "The Extended Case-Study Method" argues that there are two models of social science, "positivist" and "reflexive". He says that in the positivist approach, within which economics would be placed, we limit (page 5) "our involvement in the world we study, insulating ourselves from our subjects, observing them from the outside, interrogating them through intermediaries. We keep our feet on the ground by adhering to a set of data collection procedures that assure our distance….We try to avoid affecting the situation we study, standardize the collection of data, bracket external conditions, make sure our sample is representative."

The reflexive approach characterized by ethnography and participant observation, on the other hand, is a "model...that embraces not detachment but engagement as the road to knowledge." Burawoy (page 16) calls dialogue "the unifying principle of reflexive science." It "starts out from a dialogue between us and them, between social scientists and the people we study." (page 7) and it (page 5) "embeds such dialogue within a second dialogue between local processes and extralocal forces." To Burawoy, "reflexive science" is premised on the idea that the interview is not a neutral entity, but an "intervention" where "by mutual reaction…we discover the properties of the social order." He argues that it requires the social scientist to understand the "process," context, and both the "narrative" and "nondiscursive" dimensions of interactions with the field of study and the respondents within it which may discovered "through participation, 'doing' things with, and to, those who are being studied." Understanding, and theory, thus emerge in a process of dialogue between participant and observer, who in turn engage in dialogue with the scientific community, which in turn re-enters the "wider world of participants."

Burawoy's purpose in this paper is to make the case for "reflexive science," arguing that it is at least as scientific as the positivist approach. He, however, sees these approaches as located in entirely separate worlds, so different that they could never meet.

My purpose, instead, is to see whether a marriage between reflexive and positivist science is possible, and whether a judicious integration of qualitative and quantitative approaches can make economics more reflexive.

What Can Economics Learn From Ethnography?

For over a century, the field of economic anthropology has employed ethnography and participant observation to study questions, such as those at the intersection between economic and social life, well before mainstream economics got around to them (e.g. Gregory and Altman 1989, Gudeman 2001). Economics is perhaps too embedded within the positivist paradigm to allow ethnography to become a primary tool. However, economics can come closer to ethnography and try to better incorporate at least four key inter-connected principles from it.

The first, to use Small's term, is *Cognitive Empathy*. To reduce the distance between the researcher and those they research, by trying hard to put the researcher in the shoes of the researched. In the positivist frame, a person who is being "researched" is usually called the "subject," a term which has implicit within it the vast power differential between scholar and subject. Reducing this power differential is part of the task of achieving greater cognitive empathy– though it may be impossible to put the researcher and researched on an entirely level playing field.

The second is the idea is that information does not come only in numbers, it can also come in words – via interviews, documents, text, images, video, and other *narrative* forms. The notion of "narrative economics" got an impetus in economics, particularly in macro and finance, after the publication of Robert Schiller's (2019) book with that title, but narrative data is still rarely analyzed in the field, particularly by micro-economists.

The third is the notion of *process.* What are the minute details, the slow day-to-day shifts, that make change happen – the close observation of which is so central to a good ethnography. Empirical economics, particularly since the credibility revolution, has been almost entirely focused on the measurement of outcomes. Understanding the mechanisms that lead to those outcomes has generally been relegated to theory. Structural economics, and some work in causal inference, does attempt to get at mechanisms, but this is usually done either via speculation – by relying on theory to create a story why a certain intervention may have resulted in a certain outcome, or in discrete pieces - by analyzing heterogeneity in impacts or by designing experiments with arms that can tease out some element of how different kinds of stimuli might result in different outcomes. What is not done often is to do the most obvious thing – to keep a close eye on the ground to observe how change happens by embedding oneself in a site over an extended period to create detailed observations of the process of change.[5]

The fourth is the principle of *participation*. Most economic research is primarily a process of extraction – we collect data from subjects, we analyze the data, write a paper and our respondents never hear from us again. The best ethnographies teach us that the simple act of treating the researched as a co-creator of research rather than a subject can sharply reduce the power differential with the researcher (note, however, that this is not true of all ethnographies which can be just as extractive as any other kind of research).

In the next sections of the paper, I will briefly sketch some of the work done under these broad categories to illustrate how economists might try to come closer to these principles in our work.

    A. Cognitive Empathy

---

[5] I remember being at an interdisciplinary seminar some years ago when an economist was presenting an econometric result and drawing on economic theory to understand how rational choice might explain it. A sociology grad student raised her hand and asked, "Why didn't you go back to your respondents and ask them?" The economist had no answer.

There is a long tradition in development economics of "village studies," where economists spend months, if not years, studying, engaging with, and collecting data from a small sample of rural villages. T. Scarlett Epstein (1962) spent several months in the late 1950s studying two villages in South India where she examined the links between the local economy and social structure, comparing a "dry" (arid), and a "wet" (irrigated) village. Her data included a mix of field observations, interviews, and survey data. Her approach to research integrated training in economics and anthropology; as a student of William Arthur Lewis she was deeply interested in patterns of employment, agrarian relations and migration, but as an anthropologist trained in the Manchester School she wanted to probe into the link between caste and social class and how these were affected by agrarian change. Her work on South India is therefore suffused by an interesting mix of qualitative and quantitative data and was widely influential.

Two economic theorists, Christopher Bliss and Nicholas Stern, inspired by Epstein, spent eight months between 1974-75 in the North Indian village of Palanpur to study a series of important questions: the nature of land tenure arrangements, labor and credit markers, risk and uncertainty. They were steeped in theory, but they let their conversations, experiences and observations inform the development of new theoretical models and condition the kinds of survey data they collected. Their book (Bliss and Stern, 1982) and the papers they wrote based on their first round of fieldwork, were followed by a series of repeated visits (primarily by Stern), their students and research collaborators that continues to this day (e.g: Himanshu, Lanjouw and Stern, 2018). Interestingly, while all the researchers involved in the Palanpur project are immersed in the conversations they had in the village, rarely are those conversations reported in their published research which is theoretical and econometric. Their work clearly displays a great deal of cognitive empathy, but it is almost entirely quantitative.

Several good economists cut their teeth on the Palanpur project. Among them is Jean Dreze, whom I think of as one of the most cognitively empathetic economists in the world. He has developed what he calls the "Research for Action" method (Dreze 2017) arguing that "Research can help with arguments and evidence that contribute to more effective action." To this end he has avoided any funding from the usual places (including my own institution – the World Bank) to avoid any semblance of bias. He lives for much of the time within the poor communities whose perspectives he is trying to represent. And, he has successfully advocated for policies that draw on this bottom-up perspective, most notably the Mahatma Gandhi National Rural Employment Guarantee Act which guarantees a minimum of a hundred days of employment for anyone in rural India. Despite all this I have rarely seen him explicitly draw on qualitative data, and yet every statement embodies the interests of poor and marginalized people in India.

In the late 1980s, to systematically explore what economists and anthropologists could learn from each other, Pranab Bardhan organized a pioneering conference that brought them into a "conversation" on measuring economic change in rural India and explore the "tension" between participant-observation and survey methods. Bardhan along with Ashok Rudra (Bardhan and Rudra, 1978) had himself previously engaged in field-based research that established the relational, socially embedded, interlinkages between land, labor and credit markers in rural India. Re-reading the edited volume that emerged from the conference (Bardhan, 1989), the anthropologist Arjun Appadurai's (1989) chapter which offers a critique of both anthropology and economics makes some points that resonate well today. He says that even if economists visit the field and engage in participant observation their work tends to be confined to outcomes rather than process, and they tend to collect data that is "distributional" rather than "relational" (in the sense of understanding the relationships between groups). Drawing on his own fieldwork Appadurai argues that forcing fuzzy concepts into survey-based measures can cause mismeasurement and that survey techniques need to be developed that can accommodate "fuzzy and approximate" quantitative responses. But, he also says that anthropologists tend to not care about bias – they tend to study more prosperous villages, focus on optimistic responses, and on respondents who are outliers rather than those

that are at the median. Interestingly, this is the only article in the entire volume that makes an explicit case for mixing qualitative and quantitative methods.

TN Srinivasan (1989) in his chapter makes an important point (which characterizes some of his own work and that of Bliss and Stern) that "there is nothing inherent in the survey method that precludes it from generating the same information as a village study based on participant observation." It should be noted that Srinivasan and Clive Bell had engaged in a long period of field work in rural India to study credit markets that resulted in seminal theoretical papers on the subject. Bell (2020) recently published an interesting monograph about their field work.

Christopher Udry, who was a student of TN Srinivasan, is another economist influenced by anthropology and ethnography who describes his method as "iterative field research" (Udry, 2003). He illustrates the method with an explanation of how he and Markus Goldstein conducted research in Ghana. To quote Udry "An initial hypothesis is refined and clarified through detailed observation, which informs the collection of appropriate data. As the economic environment is clarified during the course of fieldwork, the data collection procedure can be adjusted in response. Finally, the research proceeds to formal statistical analysis and, one hopes, to new hypotheses. This iterative process of moving between theoretical reasoning, informal observation and discussion, data collection and statistical analysis is the locus of creativity in this kind of field research and is its distinguishing feature." Note the striking similarity with the process of participant-observation described by Singh.

In the process of doing fieldwork in Ghana, Goldstein and Udry observed that there were distinct differences in the productivity of land across plots that were otherwise very similar. In particular, plots owned by women were less productive than those owned by men, and their respondents attributed this to the fact women were not able to "invest" in their land by keeping it fallow. They had to keep it constantly cultivated to not have it taken away from them. In the paper that followed (Goldstein and Udry, 2008) which was a seminal attempt to demonstrate that institutional factors affect economic efficiency, they demonstrate that this qualitative finding (which they never report in the paper) generalized to a result that less powerful people in the local political hierarchy have less secure tenure rights, and women – as among the least powerful in the community – are particularly hard-hit.

Thus, some economists, usually working in development, take inspiration from ethnography and spend a great deal of time in the field – developing insights, observing, conversing, revisiting theory, collecting several rounds of survey data. But they almost never conduct a systematic analysis of their open-ended interviews, group discussions, field notes, and other types of qualitative information. The research that results is almost always within the standard economics paradigm of theory and econometrics. In my view this results in an immense loss of valuable information, and the work of Bliss, Stern, Bardhan, Dreze and Udry, would be so much richer if qualitative data were integrated with their quantitative analysis.

My own initial experience with trying to do mixed methods work in economics may be instructive as an example of why economists do not publish qualitative analysis. I conducted six months of fieldwork in three villages in Karnataka state in South India between 1992-94 with the aim of understanding how socio-cultural and economic systems interact to affect living standards within families. The first round of fieldwork was conducted in 1992 with a small team of social workers. We conducted focus group discussions, engaged in several open-ended interviews, and participatory appraisal exercises, and administered a (relatively standard) structured quantitative questionnaire to every household in the three villages. In the process of one of our initial discussions with a group of women, one participant's husband dragged her by her hair out of the room where the discussion was being conducted shouting, "Why are you wasting your time with these people – lunch is not cooked yet!" Later we heard that she was severely beaten and faced such violence every day.

This led to the decision to make domestic violence (which was at the time not a topic studied much by economists) a focus of the analysis. In the process of trying to elicit responses on this sensitive issue, in the first week our respondents told us, basically, that life was tough but manageable. Their main problems were with the government – lack of good schools, lack of adequate drinking water, etc. After a week of staying in the village and continuing our interviews, one of the women finally opened up and said, "You have become our friends and we can't lie to you anymore. Let me tell you the truth. We feel like we are in jail. Our husbands beat us all the time and no one helps when that happens. They spend all the family money on alcohol, and when they come home – no one can help us." We then explored these issues further through in-depth interviews with men and women who outlined the cold rational calculation behind much of the violence.

Based on these discussions a few key questions on wife-beating were added to the quantitative survey instrument. The analysis of the qualitative and the quantitative data demonstrated the links of domestic violence to issues of control and power within the family, female sterilization (which led to an increased risk of being accused of infidelity), alcoholism, and dowry demands. Since the qualitative work demonstrated, among other things, that domestic violence was often a rational choice undertaken to forcibly extract money from the wife's family, it also led us to rethink how economists modeled both intra-household and transfers behavior. Without the fieldwork, we would have never thought of studying domestic violence. Being in the field and engaging in many months of participant observation allowed issues to be probed in the field the moment they were observed. This permitted 'surprises' to be easily incorporated into the data gathering process. These surprises were often everyday, even mundane, experiences in the lives of the rural poor and only surprising to relatively affluent outsiders. The malleability of "iterative field research" is the fact that reorienting the analysis based on observations in the field is a key element of the method, adds value to traditional econometric practice by discovering and locating important but understudied issues within the research discourse.

Initially my hope was to incorporate the qualitative work, the theory that emerged from it, and the econometric analysis into one paper. Where the qualitative would contribute to the standard positivist frame of generating insights, which would inform a theoretical model that would generate predictions that could be tested for their generalizability with econometric analysis (Rao, 1997b). However, my co-author Francis Bloch and I found it difficult to publish this version. Referees did not like the qualitative work, and neither did editors, and we were rejected by three journals. We finally decided to split it into two papers, a mixed-methods empirical paper (Rao, 1997a) published in a public health journal (which provides the full range of results) and an economics paper which introduced information asymmetry into a bargaining model, and explained why in that social context husbands might be able to treat their wives as hostages from whom they can extract higher post-marital dowries (Bloch and Rao, 2002).

The tendency in economics to treat qualitative work, from ethnographies, or open-ended interviews, or case-studies, as "anecdote" rather than "data" is, I think, deeply limiting. It constrains the use of information and limits the kinds of questions that can be asked. Nothing illustrates this more than Elinor Ostrom's Nobel Prize winning work (Ostrom, 1990) which is full of insights from careful case-studies and qualitative work that she conducted with her research collaborators all over the world. It is these case-studies that allowed her to gather the evidence to develop theories refuting the received wisdom that collective action was more likely to fail than succeed because individual interests would outweigh the gains from collective participation (e.g. Olson, 1965). Ostrom demonstrated in several insightful analyses of case-studies of communities, that norms for collective action evolve through repeated interaction, and this repeated interaction creates stable institutions for managing a wide range of common property resources including the management of water and forests. It is important to note that Ostrom was trained as a political scientist, taught in a political science department, and (to the best of my knowledge) did not publish her largely qualitative and experimental empirical work in an economics journal till after she won

9

the Nobel prize. Yet, her cognitively empathetic approach to understanding how communities managed common resources led to deep insights on human behavior that have had a transformative impact on the social sciences. I would venture to argue that she would never have been able to do this work had she had a job in an economics department.

Younger economists are increasingly spending time in the field to collect data and test interventions, and some display a great deal of cognitive empathy in how they approach their work. MR Sharan (2021) has written a marvelous account of ten years of travel in rural areas of the Indian state of Bihar, interacting closely with local activists and trying to understand the inner workings of village government. The book, which is meant for a general audience, is less an ethnography than a memoir but it brings into sharp focus the challenges faced by citizens, particularly those from disadvantaged groups in accessing government programs. This embedded, empathetic approach is apparent, albeit implicitly, in Sharan's research (Sharan and Kumar, 2021) where he and his co-author Chinmay Kumar analyze data from 100,000 village-level politicians in Bihar to examine the provision of public goods to lower castes. They find that leaders matter – when a lower-level lower caste representative reports to a high-caste representative, lower castes receive poorer access to public goods. However, a grievance redressal system instituted by the government is accessed more by lower-caste representatives to counter this discriminatory provision.

It is clear from this overview of field-oriented economics research, that Srinivasan was partially correct. Participant observation can be an extremely valuable tool for economists and can result in quantitative findings that are cognitively empathetic. It does not however, follow that this is true of all field-based research. Cognitive empathy requires a process of listening and learning from the people one is researching, and reflecting their interests, their insights, and their perspectives in the research through a process of iterative learning. Most work in field work in economics tends to be much more top-down, focused on testing an intervention or conducting an experiment which is not based on bottom-up learning. Similarly, not all qualitative work is empathetic – a great deal of it is unfortunately, perfunctory, and careless; employing delegated teams who summarize information from focus groups that replicate the prior biases of the researcher. In a lot of so-called mixed methods work, qualitative information of this kind primarily serves as means of providing color to "hard" quantitative analysis.

On the other hand, field-oriented economists (including me) seem to collect qualitative information that they do not report, or indeed fully analyze, because the profession does not treat it as data worth writing about. This is a mistake. Narrative insights from the deep, grounded interviews and field visits can greatly enrich and complement econometric work, help contextualize findings – they are not merely "anecdotes." I turn next to a brief overview of research that analyzes data from narratives - open-ended interviews, transcripts of group meetings, social media posts, and the analysis of newspapers, to demonstrate this point.

B. Narrative Data

People do not talk in numbers, they talk in words. A survey interview where the respondent is directed to take a vague notion and emerge with a precise numerical response, or pick one of a small set of multiple choices when none might exactly fit what they are thinking, is engaging in an unnatural interaction that can lead to a misinterpretation of what they are trying to communicate. Yet almost all the data analyzed by economists comes from interviews of this kind. If economics is to become more reflexive and learn from the people it is trying to study, it needs get more comfortable with analyzing information from narratives.

A nice example of the value of narrative analysis is the use of the technique of "financial diaries" by a team of authors that included an economist, an anthropologist, a development activist, and an expert in

finance, *Portfolios of the Poor* (Collins, Morduch, Rutherford and Ruthven 2009). Collins et al.'s goal was to understand "money management" of the poor; how people earning less than $2 a day in Bangladesh, South Africa and India managed their finances; juggling consumption, investments, credit payments, and emergencies. Instead of doing a conventional survey asking questions about credit, savings and assets, they conducted open-ended, narrative, interviews of 250 households in these countries. Each household was visited at least twice a month for a full year and the research team then reconstructed balance sheets and cash-flow statements from these interviews. This unusual technique revealed several patterns that were previously not understood with conventional survey methods, and they estimated that one-shot surveys were missing about half the financial activities of a household. They found that households had to cope with multiple, overlapping sources of uncertainty and risk, and that this required active financial management. A key to being able to cope with financial ups and downs was having the ability to be able to access relatively large sums of money to deal with emergencies and life events. However, respondents faced a "triple-whammy" – low incomes, irregular cash flow, and poor access to financial instruments that were unsuitable to deal with these conditions.

The key finding was that the "portfolios of poor households," unlike those of the rich, were "managed to ensure that money can be obtained in the desired amounts at the desired times." (page 61). The book argues this requires an entirely different set of financial instruments than those that are conventionally available in banks and micro-credit organizations. Jonathan Morduch, the economist in the Portfolios team, then teamed with a finance professional, Rachel Schneider, to use the same "financial diaries" method to examine the finances of poor Americans, also finding that they had precarious lives that were not captured in existing data (Morduch and Schneider, 2017). Both books draw on the diary technique to construct "hard facts" but couple this with qualitative analysis of the interview transcripts to provide powerful picture of how poor people in very different parts of the world cope with uncertainty and risk.

Sociologists routinely draw on narrative data from open-ended interviews to address important questions. An important example is Michele Lamont's book on *The Dignity of Working Men* (Lamont, 2000). One of her goals is to examine the "inner logic of racism" by understanding the "grammar of evaluation" used by employed, lower-middle class, white and black men in the United States, and white and North-African immigrants in France. Her data consisted of 150 two-hour open-ended interviews – 30 each with blacks in the US and North-Africans in France, and 45 with whites in both countries. She conducted all the interviews herself to avoid interviewer bias, and to reduce her distance from her respondents. Her key finding is that "workers judge members of the other group to be deficient in respect to the criteria they value the most." They do this "boundary work" by constructing "mental maps" of similarity and difference. Working class white American distinguish themselves from those from higher social classes by seeing them as less disciplined and lacking in "integrity and straightforwardness." And they see blacks as "lazy" and with "wrong values." African Americans, on the other hand, consider the upper-class as lacking the criteria they value the most - "the caring self." And they see whites as domineering and less compassionate. Lamont argues therefore that "moral criteria can generate strong intergroup boundaries," though, at the same time, it can help bridge differences with statements like "there are good and bad people in all races." In France, on the other hand, white workers, using the language of class solidarity, see the poor and blacks as "part of us," but say that North Africans lack civility and are "culturally incompatible with the French." Lamont thus argues that workers' definitions of "who is part of us" have little overlap with official, government categories. The construction of difference is thus "bounded by the differentially structured context in which people live."

Interviews are, of course, not the only form of narrative data. A project that I was involved with recorded, transcribed and analyzed 300 village meetings in South India (Sanyal and Rao, 2019). The context for this was the 73rd amendment to the Indian constitution that brought deliberative democracy to all 2 million Indian villages. Village meetings are thus supposed to be spaces where citizens could speak freely about public issues, come to agreement, and work with the village council to make decisions and

monitor village public goods and common property resources. The villages were sampled, following a natural experiment, to be matched by language and ecology across neighboring districts, which belonged to linguistically different states. Prior to 1956 these matched districts were part of the same state. Since the villages matched across the paired districts, spoke the same language, had the same ecological structure, the same caste structure, and had 300 years of shared history, they belonged to the same linguistic community. Everything else equal, they should have, therefore, had the same patterns of public discourse.

However, we found that deliberative meetings were very different in different states – which, because of the natural experiment, we could attribute to state government policy in implementing the 73rd amendment. Thus, we were able to show that deliberation – or what we call "oral democracy" to distinguish it from the kinds of deliberation theorized about in western countries – was possible in poor, highly unequal, contexts with low literacy, and that the quality of deliberation was less related to levels of inequality or literacy, but determined by state government policy. Moreover, our discourse analysis found that social distinctions were largely equalized within the context of these meetings because village elites were concerned about losing votes by suppressing voice. Thus, individuals from lower castes were just as likely to speak as those from upper castes, and topics raised closely matched the interests of the median voter in the village (Ban, Jha and Rao, 2012).

What characterizes these three research projects is their relatively small-N, which reflects the difficulty of analyzing narrative data. The transcripts of 300 village meetings in India, for instance, took me and my co-author ten years (with about two years of active work) to analyze; a process that required deep reading, careful coding, reflection, debate, analysis and writing. This trade-off is limiting, and in the next section I will discuss the advantages and disadvantages of machine learning tools, and whether they could help us analyze narrative data at scale.

C. Machine Learning and Natural Language Processing

The last decade has seen vast advances in using Machine Learning (ML) to analyze textual data, a field which is known as Natural Language Processing (NLP). ML and NLP pervade our everyday lives in ways that are seen and unseen – for instance they underlie how search engines work, target advertising from social media companies, filter spam emails, create virtual assistants to listen and respond to phone calls, and an infinity of other applications. The application of these methods by governments and firms have been shown to discriminate against minorities and women because they draw on training data sets that are largely populated by dominant groups and by men (Caliskan, 2021), which is not an indictment of the methods themselves but of how they have been used.

Researchers across a variety of fields, including Economics, have begun to use these tools to study a variety of questions (e.g. Gentzkow, Kelly and Taddy 2019). NLP vastly expands our ability make sense of high-dimensional textual data which opens whole new areas of research in the social sciences. In particular, it has the potential to revolutionize mixed-methods research by making very large amounts of narrative and textual data amenable to statistical analysis. The question is whether this can bring more reflexivity to economic analysis.

To answer this question, it might be helpful to contrast the "hand-made" analysis of village meetings discussed in the previous section, with a study that I was involved in (Parthasarathy et al, 2019) that used NLP methods to analyze 100 transcripts of village meetings recorded and transcribed in the Indian state of Tamil Nadu in 2014 (and, consequently, took six months rather than ten years to analyze). The NLP paper studies the degree of gender bias in the village meetings and whether having a woman village president can correct gender bias. The transcripts are analyzed using "topic-models," a method which

groups text into assigned numbers of "bags-of-words." This allowed us to examine which "topics" were more likely to be raised by women and which by men, and which by regular citizens and which by officials. We found that men tend to speak much more than women, but that citizens are more likely to speak than officials showing that gram sabhas were active spaces for democratic discourse. We also found that men dominated the discussion with the topics raised by them more likely to be taken up by the next speaker and responded to by officials. But having a (randomly assigned) woman president of the village corrected this gender discrimination.

Gentzkow et al (2019) also study transcripts of speeches, but in the US Congress, and develop a new estimator to measure the degree of partisanship in these speeches from 1873-2016. They find that partisanship sharply increased over the period, particularly after the Republican takeover of Congress in the 1990s and the election of New Gingrich as speaker. These patterns are very different from previous work on partisanship in Congress that used roll-call votes and did not find such a sharp change. Stephen-Davidowitz (2014) also studies US politics looking at trends in racial animus in Google searches from 2004-2007, and pinpoints the locations in the United States where this is more likely to be observed. He finds that areas that displayed racial animus during this period, were much less likely to vote for Barack Obama in the 2008 election relative to votes they cast for John Kerry in the 2004 presidential election. He estimates that racial bias might have cost Obama 4.2% of the popular vote.

Lieberman and Miller (2021) study 36,000 online newspaper articles, and 306,000 comments made on them, in Nigeria and South Africa – both multi-ethnic societies with a history of civil conflict. The question they are interested in is whether news media contribute to nation-building in multi-ethnic societies. Using word counts and topic-modeling, they find that when an ethnic group is explicitly named in the newspaper article, it generates more and negative references in readers' comments – referring either to the same ethnic group mentioned in the article or to another ethnic group. In other words, news media can negatively influence inclusive nationalist frames by triggering ethnicity specific reactions among their readers.

The question is if these NLP papers can be classified as "reflexive." They substantially expand the kinds of data that can be analyzed to study discrimination and polarization and contribute towards our understanding of bias in settings that have not been previously analyzed, but the data are analyzed in a manner similar to quantitative survey data without any explicit attempt at cognitive empathy, process understanding, or participation. Comparing my own work in Parthasarthy et al (2019) and Sanyal and Rao (2019) in analyzing the same topic – Indian village meetings, it is clear that the close reading that Paromita Sanyal and I did of the transcripts, followed by hand-coding, discussion and iterative analysis (i.e. traditional qualitative work) brought us into closer proximity with the people who participated in these village meetings, than when my co-authors and I let machines do the coding for us. This is, more generally, true of the work done using NLP to study social media. Matamorez-Fernandez and Farkas (2021) conduct a systematic review of 104 studies, across disciplines, that study hate speech, and racial and gender bias in social media. They find "a lack of geographical and platform diversity" and "an absence of researchers' reflexive dialogue with their object of study."

This is not by any means an indictment of machine learning, but simply emphasizes the point that a set of techniques by themselves cannot move us toward greater reflexivity. This is much more a function of the question being asked, the approach taken to the research, and the data being collected and analyzed. Take the wonderful recent paper on "Folklore" by Michalopolous and Xue (2021). They begin with an archive of thousands of motifs in folklore from 958 societies across the globe that were coded by the folklorist Yuri Berezkin. They classify the motifs into different "concepts" using a supervised NLP model using a dictionary created by the MIT Media Lab. They find that the concepts correlate with geographic attributes of the regions where they were created – for instance, earthquake prone regions are more likely to have motifs related to earthquakes. After validating the motifs, the examine the concepts to see if concepts that

emerge in these motifs are more likely to reflect contemporary realities. They find "a striking consistency between values derived from folklore and contemporary attitudes related to trust, risk-taking, and gender norms." Thus, they can relate a society's "ancestral" cultural heritage to its current norms and behaviors showing that narratives can demonstrate powerful path-dependency.

Another recent paper by Jayachandran, Biradavolu and Collins (2021) explicitly compares a reflexive, open-ended qualitative study on gender norms in the north Indian state of Haryana, to information collected using a structured questionnaire from the same communities. The main goal of the paper is to understand which questions on gender norms and women's agency, widely used in the literature, best capture the lived reality of women. For these authors, qualitative work is the "gold standard" for understanding women's agency so they begin with a careful qualitative study, where a team of researchers conduct open-ended interviews, and code these interviews to score 209 households across several domains by the degree to which women have independent agency. The authors treat this as a woman's "true agency." They then relate this score, using ML methods (Lasso and random forest) to find the best match between the qualitative score and outcomes measured in the structured questions to find an index of five measures of agency. This method builds on previous work by Blattman et al (2016) who also use open-ended interviews to build a validation method to understand the extent of measurement error in quantitative self-reported measures of sensitive questions about crime.[6]

There is a greater element of reflexivity in this work, and the Folklore paper, because of their attempts to use ML and NLP to bring reflexively collected data (qualitative interviews in the case of the gender paper, and a close anthropological reading of folklore narratives in the Folklore paper) into conversation with other kinds of more conventional quantitative data, and thereby enrich our understanding of important phenomena. The degree of reflexivity is even more acute in two projects that are underway that are conducting open-ended interviews at scale and then analyzing using them ML and NLP.

The first study, which I am involved with, piggybacks on an ongoing panel survey of Rohingya refugees and their Bangladeshi hosts in Cox's Bazaar in Bangladesh. We have been conducting 30-minute open-ended interviews with 2,000 respondents equally divided between hosts and refugees covering three domains – aspirations, well-being and "belongingness" to understand how the hosts perceive the refugees and vice versa. In the first paper out of this project we develop a method to extend a sub-sample of 400 human coded interviews to a much larger sample of documents using NLP (Ashwin et al 2022), thus attempt to circumvent the small sample bias associated with qualitative analysis. We apply the method to study parental aspirations for children, bringing in ideas from philosophy and anthropology to distinguish between "aspirations" (a process of reversing a "core value" that results in a "change in the self") and "ambition" ("a specic goal that can be fully grasped in advance of achieving it"), and to include the concept of the "capacity to aspire" which is a cultural and cognitive resource that allows people to navigate their way to a better future. We find that there are interesting and policy-relevant differences between general and educational ambitions, and secular and religious aspirations. We also find that while poorer and less educated parents (more prevalent among the Rohingya refugees) say that budget constraints limit their ability to help their children achieve their hopes and aspirations for them, refugees are more able to express clearly articulated navigational capacity than hosts which we think may due to the selection and learning effects associated with surviving violence and finding their way to the refugee camp.

The second study, "The National Poverty Study" (Alexander et al, 2017) is based in the United States and is conducting a "qualitative census of urban and rural poverty" by interviewing 5,000 poor and near-poor respondents, sampled to be representative of poor households in in urban and rural settings in the United

---

[6] See also this interesting effort to use video and audio images to visualize poverty in India: https://onehundredhomes.in/.

States. They are employing a semi-structured open-ended protocol that covers domains that include life-histories, social and familial relationships, and economic hardship. They plan to use a mix of traditional qualitative methods and NLP to analyze the large-N qualitative data.

D. Studying Process

Empirical economics, particularly after the credibility revolution, seems almost singularly focused on outcomes. Mechanisms are generally relegated to ex-post theory, rather than careful observation of the process by which change happens. Ethnography, participant-observation, open-ended qualitative interviews, are eminently suited to filling this gap, to understanding "how", as well "how much" (White 2008, Bamberger, Rao and Woolcock 2010). Yet, despite the increased of qualitative methods in randomized trials and experiments in adjacent disciplines such as psychology (Paluck 2010), they have not yet taken hold in economics. I will briefly review three studies that integrate serious ethnography with quantitative impact evaluations to demonstrate the added value of incorporating an element of reflexivity into causal inference and demonstrate that it can help both in improving scientific understanding and in informing policy-making.

Harkening back to Elijah Anderson's ethnography of Philadelphia, a team of public health researchers and ethnographers collaborated on a mixed-methods study to find ways of dealing with urban "blight" in poorer areas of the city (Branas et al 2018). An ethnography of a poor "micro-neighborhood" in the city found it "visibly impacted" by drug trafficking and gun violence. Overgrown vacant lots were centers of drug dealing and violent activity and were used to store drugs and guns. Virtually all the residents interviewed in this neighborhood, in comparison with a better-off neighborhood, were very supportive of an intervention to clean up and "green" vacant lots. The team designed an intervention which cleaned vacant lots, planted new grass and trees, and installed fencing. 541 vacant lots were randomly chosen from poor neighborhoods in the city, and then randomly assigned to treatment and control groups. After a 38-month period, they found that areas around the treated lots saw a statistically significant 36% drop in perceptions of crime, and a 76% increase in the use of these vacant lots for social activities. Overall crime dropped by 13% and gun violence by 29%. In this study, ethnography was used less to track the process of change, and more to understand community needs to design an effective intervention to address urban blight and associated gun violence and crime.

Blattman et al's (2021) recent paper also studies drug-ridden neighbourhoods in the city of Medellin in Colombia which has long had a reputation of being at the centre of Colombia's drug economy. The research team, which combined economists and ethnographers, spent four years conducting open-ended interviews with dozens of gang leaders, managers, and foot soldiers in thirty gangs asking questions about their 30 gangs on their organization, operations, and rule. They also interviewed several experts, police, prosecutors, and community leaders. Drawing on their field work they designed a quantitative survey on the services provided by drug gangs in neighborhoods, tax collection, extortion, and the extent to which these were perceived as legitimate by the community. They used a natural experiment where the city was reorganized into 16 wards called *communas* in 1987. Streets on either side of a *communa* border were very similar, but for the three decades since the reorganization were located at different distances from the police and administrative headquarters of the neighboring *communas*. Quantitative data were collected from 7,000 respondents and 223 low- and middle-income neighborhoods. The qualitative and quantitative data provide revealing insights into the workings of the drug economy. Gangs did not provide public goods, the government does, but gangs provide excludable goods such as security services, informal contract enforcement, and dispute resolution which complement the services that governments provide. Consistent with this, the quantitative data show that the services provided by drug gangs tend to be more frequently closer to *communa* headquarters. Blattman et al thus challenge the received wisdom that organized crime "fills a vacuum left by weak states" but argue instead "criminal governance can also be a strategic response to strong state presence."

Integrating an ethnography into the design of an impact evaluation can provide a robust way of examining processes and mechanisms. In 2007, I worked with colleagues to test an intervention to improve the quality of citizen engagement in Indian village governments (Rao, Ananthpur and Malik 2017). The intervention provided two weeks of training to village residents on participatory planning, village government regulations, budgets and administrative systems, which was followed by monthly follow-up visits by facilitators over a two-year period to mobilize citizens and translate their demands and needs into actions. We picked a hundred villages at random, and then randomly assigned 50 of them to the intervention. Quantitative data from surveys of villages, households and key informants were conducted prior to the intervention in 2007 and after the two-year period in 2009. A 10% sub-sample of the quantitative sample, five treatment and five control villages, were selected for the ethnography, which is a large sample for participant-observation. We developed an "embedded reporter" approach where five trained ethnographers were hired to live in these villages (each reporter was assigned to one treatment-control pair) to conduct participant-observations. They prepared monthly reports of everything they observed – including economic changes, social changes, and political events. These 240 monthly reports constituted our qualitative data. Early rounds of the qualitative data helped inform the development of the quantitative survey instruments.

 The quantitative data showed considerable improvement over time across a variety of indicators, but the difference in outcomes between treatment and control villages were not significantly different from each other. The ethnography allowed us to unpack the reasons why the intervention "failed" highlighting the role of variations in the quality of facilitation, lack of top-down support, and difficulties in confronting the stubborn challenge of persistent inequality. We found that while surveys have the advantage of being able to measure predictable outcomes with measurable effects, major events sometimes occur that are best investigated on the ground, during the moment they happen, by participant observers. Important shifts occur during points of conflict, or during periods of mobilization and coordination. These are not quotidian events and are very hard to predictably measure. Moreover, qualitative investigators can build strong relationships within communities that allow them to "see" differently and thus capture insights that a survey interviewer is unable to do.

These three studies, all of which integrate qualitative and quantitative methods into their design at the outset and not as an afterthought, clearly show how much these methods complement each other. Ethnographic work not only helps design better survey instruments, it can also help design better interventions and provide deep insights into the processes that underly the changes measured by quantitative data. This obvious point has been made several times, but ethnographic methods still have not become a routine part of the economist's tool kit, though recently a small number of impact evaluation papers have used open-ended interviews and focus groups to understand mechanisms (e.g., Muralidharan and Singh, 2020).

   E. Participation: Respondent as Analyst

If cognitive empathy is largely about reducing the social and political distance between the researcher and the researched, participation, or co-production, takes that a step further by actively involving communities in the research process. Robert Chambers and his colleagues at the Institute for Development Studies at Sussex have for several decades now advocated methods broadly labeled as "Participatory Rural Appraisal," that teach elementary graphic and visualization tools to communities so that can analyze and improve their own lives (Chambers, 1994) which have had a sizable impact in the practice of community-driven development.

Some scholars have made the case for co-produced research, where questions are designed with the input of the communities being studied, members of the community are involved in data collection, and research results are then fed back to the community, arguing that this improves the "rigor, relevance and reach of science" (Balazs and Morello-Frosch, 2013). The Centers for Disease Control released a "Tools and Techniques" note outlining a method for a more limited form of participation – where research findings are shared via a process of two-way dialogue with marginalized communities, arguing that this builds trust and improves the dissemination of public health advice (McDavitt et al, 2016). Development economists have made a similar case for presenting results from randomized control trials to communities, though some have cautioned that this could affect the integrity of the research process (McKenzie, 2011).

With all this, the fact remains that it is rarely done. Research in economics – even research that displays a great deal of cognitive empathy - remains largely extractive, and any benefits to the respondents who participate (other than the payments for participating in a survey) is usually indirect; via the long route of research outcomes affecting policy which in turn may have a possible impact on respondents. Recently several projects around the world have been involved in efforts to "democratize data," by involving people in designing, collecting and analyzing their own data tailored to answer their own needs (Ada Lovelace Institute, 2021). Yuen Yuen Ang (2020), building on Clifford Geertz description of ethnography as "thick description," has called this "thick data."

In 2014 some colleagues and I co-produced a method, that we called Participatory-Tracking, with tribal communities in South India to facilitate community-level decision making by democratizing the design, collection, and analysis of data (World Bank, 2021). Representatives of over 200 tribal villages engaged in several weeks of deliberations to think about what constituted the good life for them, turned those ideas into indicators measured with survey questions and then tested and adapted those questions in their villages so that they did not take more than 30 minutes to answer. The questionnaires were quite different from those generally used to track poverty with only a 17 percent overlap in questions with the Indian National Sample Survey. We incorporated the community designed questionnaire into a tablet-based electronic survey and used a system of video-based training to train community representatives to use the tablet to conduct surveys of their own villages. In our pilot, which covered one district, we were able to conduct a census of 32,000 households in the district in about six weeks. Once the survey was conducted the data were dispatched directly to a cloud server to prevent anyone from tampering with them. The same exercise was repeated the following year.

The goal was for villagers to use the data for two purposes, to track changes in the quality public services and in their living standards, and to use the data to make better decisions in village meetings (*gram sabhas*). This was challenging because levels of literacy were low (at about 65 percent), and numerical literacy was much lower. Thus, we could not use conventional data analytics and visualizations, and had to develop ways of showing the data that would be intelligible to someone who could not read or write. So, we co-produced data visualizations with the community, iterating a few times till we were confident that they were widely understood. These data were then presented at village planning meetings, including the village meetings described in earlier sections, and we found that they substantially helped improve the quality of discussion by allowing citizens and officials to focus on the issues rather than debate the facts. (More information is available here https://socialobservatory.worldbank.org/categories/democratizing-data.)

Efforts such as participatory tracking, which are still in their infancy, harness technology to empower communities to conduct their own research designed to study their own priorities. Professional researchers are largely on the sidelines serving mainly as advisors and facilitators. The rise in data literacy around the world, and technology to facilitate data collection, mapping, and visualization, have made such efforts a promising line of action. In effect, we have come full circle – and returned to Charles Booth and

his idea of using members of the community to collect data on themselves and their neighbors – with their full consent. And to use maps and visualizations to allow people to study to analyze the challenges that they face. The goal is to empower them to conduct research in a manner unmediated by professionals, and to act on the research findings to improve their lives.

Conclusion

Those of us who do research on topics such as development, poverty or discrimination study people who are almost always at a considerable economic, social and cultural distance from us. This raises the important question of "Whose Reality Counts?" (Chambers, 1997). Do we focus on trying to meet the current standards of rigor of our discipline, satisfying our peer researchers, and ensuring our objectivity by maintaining an arms-length distance from those whom we research? Or is it imperative upon us to minimize the gap by discovering questions that closely reflect the interests and perspectives of our research subjects via a process of dialogue, and use methods and modes of enquiry that as closely as possible reflect their lived reality?   Burawoy (1998) argues that research of the first kind is positivist, where objectivity and detachment is the organizing principle, and the second is reflexive where engagement, not detachment, is the primary organizing force. He further argues that "reflexive science" and "positive science" represent entirely different approaches to social science and live in separate spheres.

I disagree. I argue in this paper that there is a middle path between Burawoy's sharp divide between positive and reflexive science. While Economics is still a long way away from becoming a predominantly reflexive social science, there are four ways in which it can move towards greater reflexivity. All of them could be helped by a judicious mix of quantitative and qualitative methods: Greater *cognitive empathy*, incorporating qualitative and *narrative* information explicitly into our analytical toolkit, a focus on understanding *process* as well outcomes, and facilitating efforts towards greater *participation* of communities in our research. In this paper I have drawn on work in economics and the allied social sciences to provide examples of each modality and tried to show how mixing methods might add value.

However, there are a few caveats to note if we want to move further on the path towards greater reflexivity.

1) Cognitive empathy– as numerous economists have shown – can result in entirely quantitative analysis which emerges from a deep engagement with the field. Conversely, the analysis of qualitative, narrative data is not necessarily always reflexive. However, the explicit analysis of narratives and their incorporation into the empirical toolkit of economists can move us towards greater reflexivity by permitting the voices of respondents to be directly listened to and learned from, rather than being mediated via the artificial construct of a structured survey instrument. The reluctance amongst economists to treat open-ended narratives, unless done at scale, as anything other than colorful anecdata is a serious limitation.

2) Gathering good qualitative information, and analyzing it well, requires us to learn from the standards of rigor followed by the qualitative social sciences (See, for instance, Lareau 2021). Simply going to the field, conducting a few interviews and focus group discussions is not good enough. In particularly, there is a misconception among some economists that qualitative work is cheap, "quick and dirty," or easy. Good qualitative, and mixed methods, work can add quite a lot to the cost of a study.

3) The logic of sampling for qualitative data is not necessarily the same as for quantitative data. It depends on the nature and purpose of the study. For instance, when the purpose is to conduct case-studies without necessarily claiming broad representation, this has a different logic (Small 2009) than when

qualitative data is collected, for instance, to understand patterns of well-being across a large country which would follow the familiar logic of stratified probability sampling (Alexander, 2017).

4) Integrating qualitative and quantitative work can be done in different ways. Following classical inferential logic, small-N qualitative work can be conducted using the case study method to develop hypotheses, which can then be tested for their generalizability, possibly mediated by a theoretical model, with quantitative data collected from a representative sample of respondents (Rao, 1997b). The integrated collection and analysis of qualitative and quantitative information can also be analyzed using Bayesian inference (Humphries and Jacobs 2015).

5) Machine learning and Natural Language Processing are a double-edged sword. They offer tremendous advantages in moving us towards analyzing narrative data at scale. Yet, supervised methods that rely on biased training sets, such as sentiment dictionaries developed for western contexts applied to non-western linguistic cultures, can result in substantial bias. Furthermore, as Woolcock (2021) has argued, when machines are used for analyzing data, narratives have the danger of being analyzed out of context and without nuance, resulting in misinterpretation. In other words, relying on machines without sensitive human intervention has the danger of turning reflexively collected data into non-reflexive analysis.

6) Process, studied carefully with qualitative methods, can be extremely valuable to understand the mechanisms of change and thus complement time-variant quantitative studies, particularly impact evaluations. Understanding process matters not just for research but also for policy, where an exclusive emphasis on policies that can only be assessed using experiments or impact evaluations can sharply limit our capacity to imagine and create a better world (Rao, 2019).

7) The potential for the direct participation of "respondents," "beneficiaries," and "subjects" in research is vastly unexplored. If our purpose as researchers is to assist in the process by which people become better-off, then more can be done to design and share research with people, and even co-create tools so that they can conduct research on themselves without the mediation of experts. This has implications for many things, including the measurement of well-being and poverty.

References

Ada Lovelace Institute, 2021, *Participatory Data Stewardship: A Framework for Involving People in the Use of Data,* September, Creative Commons License

Alexander JT, R Andersen R, PW Cookson, K Edin, J. Fisher, D.B Grusky, M. Mattingly, C. Varner, 2017, A Qualitative Census of Rural and Urban Poverty. *The ANNALS of the American Academy of Political and Social Science*.;672(1):143-161.

Anderson, Elijah, 1990, *Streetwise: Race, Class and Change in an Urban Community*, University of Chicago Press, Chicago and London

Ang, Yuen Yuen, 2019, *Integrating Big Data and Thick Data to Transform Public Services Delivery*, IBM Center for the Business of Government, Washington DC

Appadurai, Arjun, 1989, "Small-Scale Techniques and Large-Scale Objectives," Chapter 11 in P. Bardhan (editor), *Conversations Between Economists and Anthropologists*, Oxford University Press, Oxford and Delhi

Ashwin, Julian, Vijayendra Rao, Monica Biradavolu, Arshia Haque, Afsana Khan, Nandini Krishnan and Peer Nagy, January 2022, "Qualitative Analysis at Scale: An Application to Aspirations in Cox's Bazaar, Bangladesh," mimeo, World Bank

Balazs, Carolina L and Rachel Morello-Frosch, 2013, "The Three Rs: How Community-Based Participatory Research Strengthens the Rigor, Relevance, and Reach of Science," *Environmental Science,* Volume 6, Number 1, Pp: 9-16

Bamberger, M., V. Rao, V. & M. Woolcock, (2010). Using mixed methods in monitoring and evaluation: experiences from international development. In A. Tashakkori and C. Teddlie (eds) *SAGE handbook of mixed methods in social & behavioral research* (pp. 613-642). SAGE Publications, Inc.,

Ban, Radu, Vijayendra Rao and Saumitra Jha, 2012, "Who has Voice in a Deliberative Democracy? Evidence from Transcripts of Village Parliaments in India," *Journal of Development Economics,* Vol. 99 #10, Pp: 428-438, November

Bardhan, Pranab (editor), 1989, *Conversations Between Economists and Anthropologists*, Oxford University Press, Oxford and Delhi

Bardhan, Pranab and Ashok Rudra, 1978, "Interlinkage of Land, Labour and Credit Relations: An Analysis of Village Survey Data in East India," *Economic and Political Weekly*, Vol 13, February

Bell, Clive, 2020, *Tales of Peasants, Traders, and Officials: Contracting in Rural Andhra Pradesh 1980-82*, International Development in Focus. Washington, DC: World Bank.

Blattman, Christoper, Gustavo Duncan, Benjamin Lessing, and Santiago Tobon, 2021, "Gang Rule: Understanding and Countering Criminal Governance," mimeo

Blattman, Christopher, Julian Jamison, Tricia Koroknay-Palicz, Katherine Rodrigues, Margaret Sheridan, 2016, "Measuring the measurement error: A method to qualitatively validate survey data," *Journal of Development Economics*, Volume 120, Pp 99-112,

Bliss, Christopher J and Nicholas H Stern, 1982, *Palanpur: The Economy of an Indian Village*, Clarendon Press, Oxford

Bloch, Francis, and Vijayendra Rao, 2002, "Terror as a Bargaining Instrument: A Case Study of Dowry Violence in Rural India," *American Economic Review*, 92 (4): 1029-1043.

Booth, Charles, 1892-97, *Life and Labour of the People In London,* MacMillan, London

Branas, Charles C, Eugenia South, Michelle C. Kondo, Bernadette C.Hohl, Philippe Bourgois, Douglas J. Wiebe, John M. MacDonald, 2018, "Citywide cluster randomized trial to restore blighted vacant land and its effects on violence, crime, and fear," *Proceedings of the National Academy of Sciences* March, 115 (12) Pp: 2946-2951

Burawoy, Michael, 1998, "The Extended Case-Study Method" *Sociological Theory*, Vol. 16:1, March

Caliskan, Aylin, 2021, "Detecting and Mitigating Bias in Natural Language Processing," *Artificial Intelligence and Emerging Technology (AIET) Initiative*, Brookings Institution, Washington DC, May

Chambers, Robert, 1994, "The origins and practice of participatory rural appraisal," *World Development*, Volume 22, Issue 7, Pp 953-969

Chambers, Robert, 1997, *Whose Reality Counts? Putting the First Last,* Intermediate Technology Publications, Great Britain

Collins, Daryl, Jonathan Morduch, Stuart Rutherford and Orlanda Ruthven, 2009, *Portfolios of the Poor: How the World's Poor Live on $2 a Day*, Princeton University Press, Princeton

Dreze, Jean, 2017, *Sense and Solidarity: Jholawala Economics for Everyone*, Permanent Black, New Delhi

Epstein, Trude Scarlett, 1962, *Economic Development and Social Change in South India*, Manchester University Press, Manchester

Gentzkow, Matthew, Jesse M. Shapiro, Matt Taddy, 2019, "Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech," *Econometrica* Vol. 87(4), Pp: 1307-1340

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy. 2019. "Text as Data." *Journal of Economic Literature*, 57 (3): 535-74.

Goffman, Erving, *Stigma: Notes on the Management of Spoiled Identity*, 1960, Prentice-Hall, New York

Goffman, Erving, *Strategic Interaction*, 1969, University of Pennsylvania Press, Philadelphia

Goldstein, Markus and Christopher Udry, 2008, "The Profits of Power: Land Rights and Agricultural Investment in Ghana," *Journal of Political Economy*, Vol. 116 (6).

Gregory, CA and JC Altman, 1989, *Observing the Economy*, Routledge, London

Gudeman, Stephen, 2001, *The Anthropology of Economy*, Blackwell Publishers, Malden and Oxford

Himamshu, Peter F Lanjouw and Nicholas Stern, 2018, *How Lives Change: Palanpur, India and Development Economics*, Oxford University Press, Oxford and New York

Humphreys, Macartan and Alan M. Jacobs, 2015, "Mixing Methods: A Bayesian Approach." *The American Political Science Review* 109, no. 4 : 653–73.

Jayachandran, Seema, Monica Biradavolu and Jan Cooper, 2021 "Using Machine Learning and Qualitative Interviews to Design a Five-Question Women's Agency Index," Northwestern Global Poverty Research Lab Working Paper no. 21-104, March

Kanbur, Ravi and Paul Shaffer, "Epistemology, Normative Theory and Poverty Analysis: Implications for Q-Squared in Practice," *World Development*, Volume 35, Issue 2, 2007, Pp: 183-196

Lareau, Annette, 2021, *Listening to People: A Practical Guide to Interviewing, Participant Observation, Data Analysis, and Writing It All Up*, University of Chicago Press, Chicago

Lamont, Michele, 2000, *The Dignity of Working Men: Morality and the Boundaries of Race, Class and Immigration*, Russel Sage Foundation and Harvard University Press, New York and Cambridge MA.

Lieberman, Evan and Andrew Miller, 2021, "Do Online Newspapers Promote or Undermine Nation-Building in Divided Societies: Evidence from Africa," *Nations and Nationalism* 27, Pp: 238-259

Matamoros-Fernández, A., & Farkas, J, (2021), "Racism, Hate Speech, and Social Media: A Systematic Review and Critique," *Television & New Media*, 22(2), Pp: 205–224.

McDavitt B, Bogart LM, Mutchler MG, Wagner GJ, Green HD Jr, Lawrence SJ, et al. 2016, "Dissemination as Dialogue: Building Trust and Sharing Research Findings Through Community Engagement," *Preventing Chronic Disease*, Center for Disease Control, Volume 13:150473.

McKenzie, David, 2011, "Should We Share Results with Respondents?" *Development Impact Blog,* October 3rd.

Morduch, Jonathan and Rachel Schneider, 2019, *The Financial Diaries: How American Families Cope in a World of Uncertainty,* Princeton University Press, Princeton

Morgan, Mary S, 2019, *Charles Booth's London Poverty Maps,* Thames and Hudson, London

Michalopoulos, Stelios, Melanie Meng Xue, 2021, "Folklore," *The Quarterly Journal of Economics*, Volume 136, Issue 4, Pp 1993–2046

Muralidharan, Karthik and Abhijeet Singh, "Improving Public Sector Management at Scale? Experimental Evidence on School Governance in India," NBER Working Paper 28129, Cambridge MA

Olson, Mancus, 1965, *The Logic of Collective Action*, Harvard University Press, Cambridge MA.

Ostrom, Elinor, 1990, *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge University Press, Cambridge and New York

Paluck, Elizabeth Levy, 2010, "The Promising Integration of Qualitative Methods and Field Experiments" *The ANNALS of the American Academy of Political and Social Science*. 628(1):59-71.

Parthasarathy, Ramya, Vijayendra Rao and Nethra Palaniswamy, 2019, "Deliberative Democracy in an Unequal World: A Text-As-Data Study of South India's Village Assemblies." *American Political Science Review* 113, no. 3: 623–40.

Rao, Vijayendra, Kripa Ananthpur and Kabir Malik, 2017, "The Anatomy of Failure: An Ethnography of a Randomized Trial to Deepen Democracy in Rural India," *World Development*, Volume 99, Pp:: 481-497

Rao, Vijayendra, 1997a, "Wife Beating in Rural South India: A Qualitative and Econometric Analysis," *Social Science and Medicine*, Vol. 44, # 8, Pp. 1169-1180, April

Rao, Vijayendra, 1997b, Can Economics Mediate the Relationship Between Anthropology and Demography, *Population and Development Review*, *23*(4), 833–838

Rao, Vijayendra, 2019, "Process-Policy & Outcome-Policy: Rethinking How to Address Poverty & Inequality," *Daedalus*; 148 (3): Pp 181–190.

Sanyal, Paromita and Vijayendra Rao, 2019 *Oral Democracy: Deliberation in Indian Village Assemblies*," Cambridge University Press, Cambridge UK

Schiller, Robert, 2019, *Narrative Economics: How Stories go Viral and Drive Major Economic Events*, Princeton University Press, Princeton.

Sharan, MR, 2021, *Last Among Equals: Power, Caste and Politics in Bihar's Villages*, Westland Press, Delhi

Sharan MR and Chinmay Kumar, 2021, "Something to Complain About: How Minority Representatives Overcome Ethnic Differences," mimeo, Department of Agricultural and Resource Economics, University of Maryland

Singh, Bhrigupathi, 2015, *Poverty and the Quest for Life: Spiritual and Material Striving in Rural India*, University of Chicago Press, Chicago and London

Small, Mario Luis. 2009, "`How Many Cases Do I Need?': On Science and the Logic of Case Selection in Field-Based Research." *Ethnography* 10, no. 1 (March): Pp: 5–38.

Small, Mario Luis, 2011, "How to Conduct a Mixed Methods Study: Recent Trends in a Rapidly Growing Literature," *Annual Review of Sociology*, 2011, Vol. 37, pp. 57-86

Small, Mario Luis, 2018, "Rhetoric and Evidence in a Polarized Society," Coming to Terms with a Polarized Society Lecture Series, Institute for Social and Economic Research and Policy, Columbia University, New York, March

Srinivasan, TN, "On Studying Socio-Economic Change in Rural India," Chapter 10 in P. Bardhan (editor) *Conversations Between Economists and Anthropologists*, Oxford University Press, Oxford and Delhi

Stephens-Davidowitz, Seth, 2014, "The cost of racial animus on a black candidate: Evidence using Google search data," *Journal of Public Economics,* Volume 118, Pp; 26-40

Thaler, Richard H. 2016. "Behavioral Economics: Past, Present, and Future." *American Economic Review*, 106 (7): 1577-1600.

Udry, Christopher, 2003, "Fieldwork, Economic Theory and Research on Institutions in Developing Countries," mimeo.

White, Howard, 2008, "Of Probits and Participation: The Use of Mixed Methods in Quantitative Impact Evaluation," *IDS Bulletin* 39:98-109

Woolcock, Michael, 2021, "On Folklore and 'Folklore' (QJE): Forging Dialogue, Reconciling Tradeoffs," *Development Impact Blog*, February 23rd

World Bank, 2021, "Helping Communities to Gain the Ability to Collect and Analyze Their Own Data," Spotlight 1.1. *World Development Report 2021, Data for Development.* The World Bank Group, Washington DC