HOW AMERICANS WORK
USING IN-DEPTH INTERVIEWING TO RETHINK DATA INFRASTRUCTURE

CONFERENCE REPORT
JANUARY 17, 2020

UC San Diego

YANKELOVICH
CENTER FOR SOCIAL
SCIENCE RESEARCH

With the generous support of the Yankelovich Center, Prof. John Ahlquist of UCSD's School of Global Policy and Strategy convened a select group for a daylong exploratory workshop on January 17, 2020. The workshop was designed to provoke a conversation exploring how deep, qualitative interviewing and ethnography can be used to improve our public data infrastructure as we seek to understand the challenges Americans face in earning a stable living. The group included an interdisciplinary set of academics—communication scholars, economists, political scientists, and sociologists—alongside social entrepreneurs, government statisticians, and experienced field interviewers and research managers.

Premises and Goals

The inspiration for the workshop came from a variety of claims appearing in both the popular press and scholarly conversations about the rise of automation, the future of work and the nature of modern employment. Robust data capable of describing seemingly obvious trends —from the purported growth of the "gig economy" to the fissuring and restructuring of firms— is difficult to find and is particularly scarce in the United States, limiting the quality of existing policy debates. The workshop therefore began from the following premises:

1. Commonly accepted, high-quality public data enable better decision-making across the economy and undergird the usefulness of a variety of private sector data initiatives and business models;

2. The structure of work, production, and employment is changing;

3. Yet existing public policy and our resulting administrative data assume a job contract, and household from the mid-20th Century;

4. Cultural and technological changes are eroding the usefulness of conventional surveys;

5. If we are to understand how Americans work (and get paid), we will need to track the same households over time.

In short, our existing administrative and survey data are not good at telling us what we're missing. To chart a path forward, we need to compare what our existing data tell us with "ground truth," triangulated using multiple approaches, including repeated, in-depth interviews with the same people over time. Part of this will involve developing tools and methodologies for systematically and securely integrating qualitative, interview-based findings with other data sources.

There has been a recent flurry of excellent academic and journalistic reporting on issues of economic precariousness; perceptions of economic mobility; and changing skill requirements. This work has produced a trove of information and accumulated experience. The workshop took advantage of this situation by assembling some of the best people doing in-depth, interview-based data collection alongside social and data science specialists. We sought to address the following issues:

- What are the qualitative data telling us about how Americans work?
- What are some of the practical challenges in recruiting interview subjects, conducting interviews and connecting with other data sources, both administrative and survey-based?
- To what extent are people able to reliably report on the exact structure of their employment relationships? How much probing does this take?
- How do interviewers gain and maintain trust, especially as a team (as opposed to individuals)?
- What are the existing approaches to data security and privacy?
- What is the cost structure, workflow, and timeline for deep interview work? How does this vary across the population?
- Where do the major federal bureaucracies and data products fit in to all of this?

What we discussed

The conversation was organized into three main sessions. The first involved a detailed report from field interviewers and research managers from the American Voices Project (AVP)—an ambitious survey and interview project of 5,000 low- and middle-income American households currently in the field. The AVP team outlined their organizational structure and data gathering procedures in detail. The workshop heard about the challenges of recruiting and training field interviewers, the process for winning trust and securing interviews with randomly identified subjects, and the technological tools for running the whole enterprise. Of particular interest is the AVP's plan to allow the linking of deep interview and survey data with Census and other administrative data products. Negotiations for securing this linkage are in progress.

In the second session, the workshop turned its attention to digital platform-mediated labor. We were fortunate to hear from several sociologists who have conducted extensive interviews with workers about their lives in the era of platform- and application-mediated job contracting, especially with micro-tasks (mTurkers) and child and elder care (Care.com). Both types of jobs are far less concentrated in urban centers than ride hailing. These conversations highlighted the difficulty in finding and contacting workers based on their employment relationship with a specific firm or platform. Moreover, the prevalence of various scams on some high-profile gig-finding platforms emerged as an underappreciated challenge in securing interviews. The session also highlighted the multiplicity of gig types and gig-finding strategies that workers employ to make themselves "legible" to both the platforms and potential clients.

In the third session, the workshop heard from government statisticians working in the Bureau of Labor Statistics. In 2017 the BLS fielded an updated version of the Contingent Worker Supplement (CWS) to the Current Population Survey, designed to measure nontraditional and contingent contracting *in a worker's "main job."*[1] Ultimately, the 2017 CWS suffered from some weaknesses from which the BLS is attempting to learn. As a direct result, the Committee on National Statistics (CNStat) has convened a committee on measuring contingent work, with its

---

[1] The BLS defines contingent work as a job that a worker does not expect to last. https://www.bls.gov/cps/contingent-and-alternative-arrangements-faqs.htm

report expected in the spring of 2020. In the meantime, the BLS and Social Security Administration are expanding the Occupational Requirements Survey and plan to release the raw text of employers' responses for research purposes via secure US Census Research Data Centers.

The final session of the day focused on synthesizing the conversations and identifying key take-aways and directions for the future.

What we learned

Although workshop participants left with a wealth of detail, there are some conclusions worth highlighting here. First, there is broad consensus that interview-based work consistently reveals that job opportunities and incomes among low- and middle-income workers are quite volatile—even within the span of a few weeks. This volatility is hard to see in snapshot surveys or in data aggregated to the annual level. Workers are spending an increasing amount of (unpaid) time trying to manage this volatility, whether through searching for gigs or putting in the "cultural work" to make themselves legible and competitive in that new work-finding environment. This work could include taking attractive photos of oneself, reporting your location, or consistently managing a stream of text messages, emails, and reviews on a variety of online reputation management systems across platforms. Importantly, several major online platforms allow employers to review workers while preventing workers from reviewing employers. In some cases, *any* employer can rate a worker, even if they have not actually hired that worker.

Second, if there was any doubt, the household is the critical unit for understanding the choices people make about the type and diversity of work arrangements they entertain. Traditional surveys and administrative data make this harder to see.

Third, there are important differences between online, app-based work platforms. On-demand platforms (Uber, Lyft, Grubhub, etc.) typically process payments and take a cut of every transaction. Marketplaces, however (e.g., Care.com) rely on subscription or other fees for linking employers and workers, somewhat like a traditional temp agency. These different business

models have a variety of potential implications for the structure of work, the balance of power between workers, employers and platforms.

When it comes to effective interviewing there were several points of consistent agreement. First, there was broad consensus that starting an interview with an open-ended invitation for the respondent to tell her story is an excellent way to build rapport.

Second, experienced interviewers reported that middle- and professional-class respondents are generally more difficult to find and persuade to talk than working class and poorer individuals. This occurs for a variety of reasons. Some are practical: middle class Americans are more likely to live behind more layers of security and screening and higher income respondents tend to ascribe a higher value to their time. But there are also sociological barriers: middle-class respondents tend to be more skeptical of interviewers and concerned about scams. On the other hand, higher-income subjects are somewhat easier to track over time, as their addresses and other contact information tend to be relatively stable. But given the "hollowing out" of middle-income work and the variation in strategies for coping with these changes, finding effective ways to gather deep interview data from higher income participants will be crucial for developing improved data systems.

A third takeaway from our conversation involves the cost structure of deep interviews. We were unable to settle on an approximate dollar amount required for a standard interview, but the amount of work required to generate a single response is extraordinary, even before considering subject payments and the time of the actual interview. Constructing a random sample of respondents (i.e., a sample from which we might make credible inferences to a larger population) is particularly difficult and expensive compared to the convenience or snowball samples traditionally used in qualitative and ethnographic research. There is no purpose-designed technological platform ready for managing large-scale interview-based work, so interview teams usually cobbled together solutions from a variety of existing products built for other applications (e.g., political canvassing). Unsurprisingly, deep interviews and related techniques are not stable techniques, even if they will be critical for developing scalable processes for understanding the new world of work and employment.

What are the major outstanding issues?

As important as what we learned are the issues where we agreed we need more exploration. Here are major outstanding questions that various attendees at the workshop plan to pursue:

- What is the best way to ask survey respondents and interview subjects about their jobs and how they earn a living? How can we most efficiently elicit information on part-time, informal, and "gig" work?
- To what extent are workers themselves able to provide detailed responses to questions about job contracting and workplace "fissuring"? Does this vary with job or personal characteristics?
- How effectively can we use new ways to contact workers (e.g., Facebook groups) to construct linked worker-employer datasets?
- What are the best practices with respect to guarding respondent privacy while exploiting existing data?
- How much time does it take to generate a living? How many hours of unpaid work is needed to produce a paid hour?

Conclusion

Ultimately, we would like to supplement the traditional unemployment rate with new quantities: how many hours of unpaid work does a worker do for every dollar of income? Given a portfolio of jobs, can we describe how volatile a household's income is likely to be? We are only starting to get there, but the 2020 workshop illustrated the value and challenges of deep interviews. A series of new products, in particular the AVP and the Occupational Requirements Survey will provide a sandbox in which to test some new data science tools that may allow for the first step: richer data at scale. The Yankelovich Center is well-positioned to provide early support for social and data scientists to play in these new sandboxes. Similarly, the Center is well positioned to provide support as others around UCSD build the case for a Census Research Data Center in the San Diego region.

Further Reading:

Ahlquist, John S. 2019. "Socio-Economic Weather Stations." White paper, UC San Diego. http://johnahlquist.net/files/WeatherStationsMemo.pdf

American Enterprise Institute. 2019. "Can ethnographic research help shape public policies?" American Family Diaries. https://www.aei.org/american-family-diaries-3/

Irani, Lilly. 2015. "Difference and Dependence among Digital Workers: The Case of Amazon Mechanical Turk." South Atlantic Quarterly.114:1

Irani, Lilly and M. Six Silberman. 2013. "Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk." In CHI 2013

Ticona, Julia and Alexander Mateescu. 2018. "Trusted strangers: Carework platforms' cultural entrepreneurship in the on-demand economy." New Media & Society. 20(11):4384-4404

*****

American Voices Project: https://americanvoicesproject.org/

Canary: https://www.workwithcanary.com/

CNStat committee on Contingent Work and Alternate Work Arrangements: https://www8.nationalacademies.org/pa/projectview.aspx?key=51472