Socio-Economic Weather Stations

John S. Ahlquist

December 19, 2018

Ahlquist is associate professor of political economy in the School of Global Policy and Strategy at UC San Diego. 2017-18 fellow at Stanford's Center for Advanced Studies in the Behavioral Sciences. www.johnahlquist.net. Email: jahlquist@ucsd.edu

What it means to "have a job" or "go to work" is changing. But the public policies designed to protect people from illness, injury, aging, and recession, are bound up in a mid-20th century understanding of employment and the household. Our administrative data and economic surveillance tools are not evolving. Consequently, efforts to better link administrative data across silos, while important, will be inadequate. We need a new data *platform* on which to build robust, granular, widely accessible, and secure applications enabling better policy, research, and business decision-making.

Using an analogy to weather station infrastructure, this memo outlines a flexible socio-economic data platform reflecting the work, households, and neighborhoods of today and tomorrow. Building this infrastructure will be a long-term, large scale project with multiple stakeholders. As a first step, we propose a multi-year "gap study" that systematically compares what we could learn from administrative data under optimal conditions against what we can observe using other tools. The study holds out the promise of learning more about how to use new tools at scale, including methods for eliciting collaborative engagement with data providers.

Introduction

There is now widespread acceptance—at least within academic and policy circles—of the need to better integrate our existing administrative data. Call this the "data linkage movement." But understanding our dynamic, technology-driven economy will require more than "just" using existing data better. I arrive at this conclusion from four premises.

> "Lesson 1. If a science-based agency...waits until it is close to becoming obsolete, it will require a complex and very expensive program to modernize."
>
> [National Research Council, 2012a, 75]

Premise 1: Commonly accepted, high-quality public data can enable better decision-making across the economy [PEW Charitable Trusts, 2018, OECD, 2004]

Premise 2: The structure of work, production, and employment is changing [Arrieta Ibarra et al., 2018, Weil, 2014, Ahlquist, forthcoming, McKinsey Global Institute, 2016].

Premise 3: Cultural and technological changes are eroding the usefulness of conventional surveys [National Research Council, 2013, Czajka and Beyler, 2016]

Premise 4: Existing public policy and our resulting data infrastructure assume a job contract, household, and neighborhood structure of the mid-20th Century [Weil, 2014, Harris and Kreuger, 2015].

Better linking of administrative data across government silos [Abowd et al., 2004, Card et al., 2010], while important, will not solve these problems. For example, LEHD[1] links unemployment earnings data with census data, allowing us to see whether (and where) those leaving (UI-covered) jobs in the last quarter found other (UI-covered) work. But only certain types of work contracts and employers appear in these data. Consequently we are unable measure with any certainty where workplace relationships are shifting, how quickly, and the consequences for human welfare [Abraham et al., 2017]. The extent to which more American families now fall through the cracks is hard to ascertain. Making and evaluating new policy is even more difficult. But to get useful information in the hands of policymakers, businesses, entrepreneurs, educators, and households we need a new public data *platform* on which to build robust, granular, widely accessible, and secure applications.

[1] Longitudinal Employer Household Dynamics [Abowd et al., 2004]

We require new tools and protocols from which we can build up from individuals and down from employers or clients, identifying the contracts where they meet over time. Figure 1 illustrates what I have in mind, where a person has variety of contractual arrangements with different employers doing different sets of tasks over time. This person is also offered a contract from employer3, which she rejects, and experiences a spell of unemployment. In such a world we need longitudinal data on people, employers, *and* contracts (jobs).

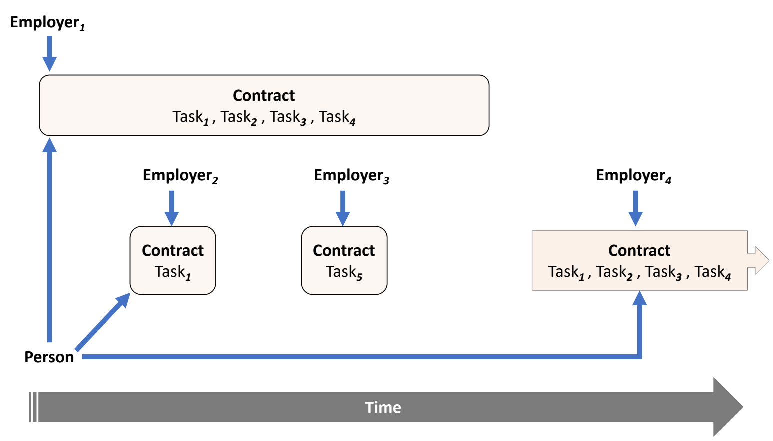

Figure 1: The new world of work: skills, tasks, and contracts over time

This may be controversial, so let me provide three brief examples of the types of changes we are currently unable to track consistently.

How big is the "gig economy"? This seemingly basic question has proved hard to answer convincingly, not least because our existing administrative and survey data are not designed to effectively measure serial and parallel "gig" work at scale and over time. As an illustration, Figure 2 displays a telling result from Abraham et al. [2017]. They do their best to measure the size of the gig economy using existing micro data found in Census (pink) and tax filing data from the IRS (black). Depending on the source used, the definition of "gig work" differs and the data tell different stories, both in levels and trends.

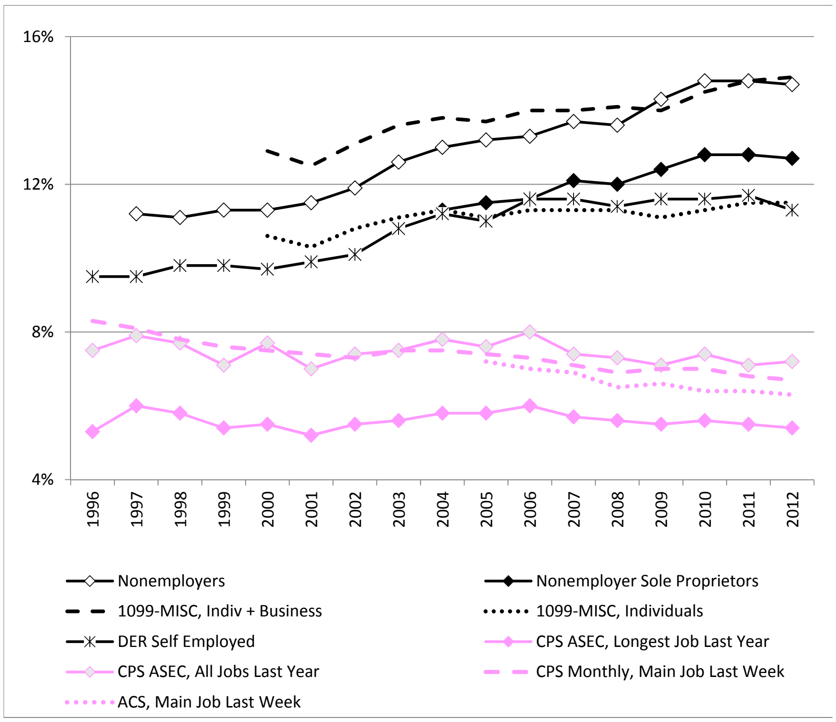

Figure 2: Measuring the "gig" economy with tax (black) and survey (pink) data sources. The vertical axis is self-employment rate. *Source:*Abraham et al. [2017].

Are workplaces "fissuring"? Weil [2014] argues numerous tasks that were formerly performed by direct employees of large firms have been converted into arms-length contracting arrangements, to the detriment of wages and worker health and safety. But Weil relies on case studies and journalistic accounts; we lack the data in the United States for documenting the "fissuring workplace." However, Goldschmidt and Schmieder [2017] use German administrative data to identify workers doing essentially the same sets of tasks at the same work sites over time but whose contractual status changes. They find that "domestically outsourced" German workers incur substantial wage and earnings penalties. Is it the same in the US?

WHAT IS A SKILL AND WHAT IS IT GOOD FOR? A thick market task-based contracting requires a skilled population capable of credibly signaling their capabilities. Where will worker training and expertise come from? Who will bear the risks of investing in skills that may become obsolete? Firms themselves face a "common pool" problem: they all benefit from a well-trained workforce but would prefer that someone else fund that education and take the risk of betting on a specific skill set. We have limited tools for measuring skills—especially those that are highly context-specific or lying dormant. For example, consider broadcasters on the Twitch real-time streaming platform to whom audiences donate large sums to watch them play video games online. Should we be measuring aptitude at a particular game as a skill? What happens when a broadcaster's game of choice loses popularity? Should unemployment schemes evolve to facilitate retraining? What are Twitch's incentives to train or manage broadcaster talent?[2]

[2] See Balter [2018] for an in-depth discussion of gaming-for-pay and the skills and risks involved.

WHAT COUNTS AS WORK MATTERS. We measure the size of firms based on the number of people working under certain contractual arrangements. Taxes, regulations, and the protections and benefits afforded workers depend on this number. Minimum wages are defined on an hourly basis. As a result, our tools are skewed to measuring only those workers—and the work they do—falling under particular contractual relationships. Not only do firms react strategically by adjusting how many worker they hire, under what contracts, and how they allocate work [Garicano et al., 2016]. It also means that we are partially blind to huge chunks of the work people are doing and the conditions under which it occurs. Firms and workers have an incomplete view of the opportunities available to them [Ahlquist, 2015].

EXISTING ANALOGIES DO RECOGNIZE THE EMERGENT FEATURES OF OUR SOCIETY AND ECONOMY. Some in the data linkage movement imagine data "mosaics" [Entwisle et al., 2017] or "tapestries" [na, 2017]. Neither of these analogies are compelling because, without a central designer who imposes her vision, mosaics and tapestries would be nothing more than a random scattering of tiles or weaving of threads. We are not seeking to assemble pieces to fit a pre-existing image. Rather we need a data infrastructure that produces a dynamic image of a large, complex, highly connected economy.

Another recent initiative is the NSF-funded "social observatories" [Social Observatories Coordinating Network] project. The social observatories initiative recognized a need for new tools, funded innovative projects, and produced an important series of essays on data infrastructure in the US.[3] But the observatories initiative was

[3] See the January 2018 and January 2017 issues of the *Annals of the American Academy of Political and Social Science*.

largely university-driven and research-focused, with no clear path to a broader and more accessible data infrastructure. More generally, the observatories analogy conjures images of expensive, bespoke instruments prone to obsolescence and only accessible to a few. It implies a distant, one-way, and purely academic relationship between "society" and the observers.

Weather stations and data platforms

We propose an analogy to weather stations as a better way to envision a new data infrastructure. Modern weather systems, illustrated in Figure 3, take data from a variety of sources—government-owned buoys, satellites, and observation stations as well as private airports, ships, aircraft, and local, citizen weather stations—and "ingests" them, based on a set of evolving models [National Research Council, 2012a,b]. The harmonized data feed into both government and private forecasting models and dissemination products. In other words, the weather station system represents an example of "government-as-platform" [O'Reilly, 2017], where other entities—government agencies, firms, researchers, enthusiasts—can both contribute to and build upon the system, providing broad dissemination in formats designed for specific use cases.

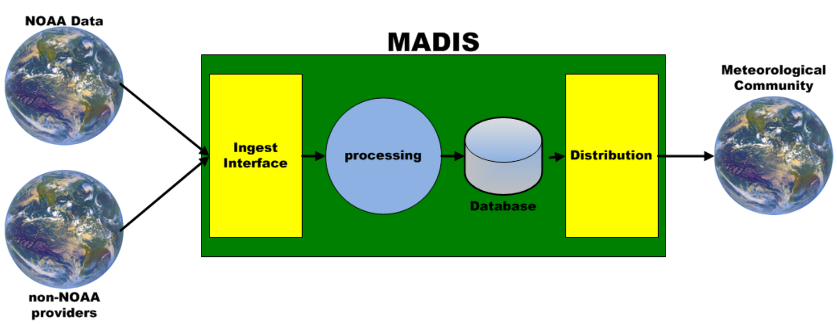

Figure 3: NOAA's Meteorological Assimilation Data Ingest System (MADIS). *Source:* NOAA.

I want to highlight the distributed, real-time private data infrastructure feeding in to the larger system. The Citizen Weather Observer Program (CWOP) and Automatic Packet Reporting System (APRS) allow individuals to collect and transmit local weather data into the larger system. In return the system provides contributors with specialized weather data and feedback on the quality of their measurements, reinforcing an enthusiast community. Figure 4 displays a snapshot of this system in action for the south San Francisco Bay region.

The system provides an interesting model for social data collection.[4] Imagine if individual workers, via a mobile app or other local

[4] CWOP is not critical for weather forecasting, in part because people and their structures are not located optimally with respect to weather forecasting needs (the problem changes when we think about social data). CWOP is important for validation.

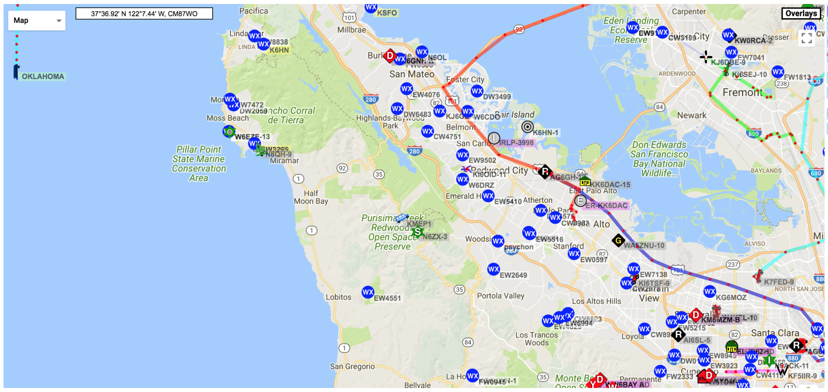

Figure 4: An example distributed collection of real-time weather data from APRS and CWOP. *Source:*https://aprs.fi/weather

interface, could choose to provide localized, intermittent data on their current location, activities, contractual status, and even the tools and teams they are working with. Imagine if this data could then be linked with both geo-referenced information and existing administrative data, such as the Quarterly Census of Employment and Wages (QCEW) as well a formal job or task characteristics as stated on LinkdIn, Indeed, or UpWork. Imagine if this data could be anonymized and processed such that other users could build it into research, policy-making, and even governmental program administration or business decisions. We would have something like a socio-economic version of the MADIS program.

THE WEATHER STATION ANALOGY OFFERS A SERIES OF DESIDERATA FOR A NEW SOCIO-ECONOMIC DATA PLATFORM:

1. High resolution in space and time
2. "Always on" data collection
3. Collaborative, relatively open platform
4. QA/data ingest with feedback to data providers
5. Link with existing administrative data
6. Incorporate drone/remote sensing capabilities
7. High-value, audience specific dissemination
8. Ongoing, evolving interplay between data sources, ingest, and analysis.

The last point is important. In the weather domain there appears to be consensus that the National Weather Service forecasts are inferior to European, British, and some commercial alternatives. This is largely the result of NWS relying on older models and processes for ingesting data from existing sources rather than lagging investment in more precise measurement tools [Behar, 2016, National Research Council, 2012b]. In the social data context, taking advantage of high-frequency but weakly structured data and non-random samples will require rapid, iterative advances in data ingest capabilities, not to mention data security and privacy (below).

Limitations

Like all analogies, the one to weather stations breaks down when pushed too far. Table 1 highlights several areas where this happens when applied to socio-economic data. The red text emphasizes two fundamental differences. First, unlike the weather, people usually *react*—perhaps strategically—to the process of measurement, especially when important decisions hinge on the data. Such tools may only attract participants who are highly motivated by extreme satisfaction or dissatisfaction with their workplaces. Firm might try and prevent workers or contractors from participating. Or the opposite may happen: some employers may want workers to participate so as to skew the data produced. Second, unlike the relatively stable set of 5-10 variables required for the weather, the attributes and behaviors we want to understand (e.g., tasks or skills) may evolve over time as new technologies, work arrangements, and products emerge. All the other distinctions ultimately flow from these fundamental differences.

The Strathern [1997, 308] paraphrase of Goodhardt's law: "When a measure becomes a target, it ceases to be a good measure."

Table 1: How the weather station analogy breaks down when applied to people.

	Weather	**People**
Variables & measurement	Stable, localized	**evolving, reactive**
Use cases	Well-defined	Broad, contingent
Representativeness, data ingest	Modeled	Can be modeled?
Forecasting	Hard	Wicked hard
Privacy & consent	What?	Critical
Data security	A problem?	Critical
Regional/Int'l Coordination	Strong, improving	Moderate

Identifying new categories from noisy, weakly structured data with an evolving feature set is a difficult problem. But machine learning tools, especially for text and image processing, are advancing rapidly. Managing human reactivity and protecting privacy and data security are much tougher.

DATA SECURITY AND THE TRUST OF DATA CONTRIBUTORS ARE CRUCIAL. Convincing both workers and organizations to share data on an ongoing basis will require three things: well-calibrated incentives for data providers, the ability to triangulate across multiple data sources, and a secure data environment that ensures participants' anonymity and control. Whether socio-economic weather stations succeed in becoming a new data platform will largely turn on our ability to solve these problems.

Table 2 outlines the needs and incentives from various stakeholders. Other actors include payment processors (ADP, Paychex, banks) and telecoms. They are important nodes in the web of economic transactions. Their participation would be transformative but creating the means and incentives for doing so is challenging.

Table 2: Costs and benefits

	Give	**Get**
Individuals	Data, time, trust	Feedback, monetary incentive, improved services
Employers	Access, data, trust	New/improved data, ability to shape tools, managed regulation, monetary incentives
Government	Time, staff resources, access funding	New/improved data, ability to shape tools
Researchers	Time, expertise, training, reputation/independence	Access, ability to shape tools, new/improved data
Civil society/funders	Funding, convening power incubation participants	Access, ability to shape tools, new/improved data

A proposed "gap study"

SOCIO-ECONOMIC WEATHER STATIONS ARE A LONG-TERM, INSTITUTION-BUILDING PROJECT requiring buy-in from policy-makers as well as active collaboration with government agencies, employers, and workers themselves. To be sustainable it will require both public support as well as demonstrable benefits to the private sector. We are not yet in a position to fully build out this infrastructure nor even to pitch a complete plan. Our motivating assumption is that we don't fully understand the nature of our ignorance.

WE PROPOSE A "GAP STUDY" AS A LOGICAL, FEASIBLE FIRST STEP. The core objective is to replicate, in miniature, the best we might achieve with existing administrative and survey tools and then compare it to the "ground truth," as measured using tools not commonly integrated with administrative data. Ideally we would track individuals, employers, and their connections over time (see Figure 1).

THE STUDY WILL BE LONGITUDINAL AND GEOGRAPHICALLY BOUNDED. Evolution over time is a key concern so we will need to follow participants over the course of at least one year. Given the heterogeneity across jurisdictions in administrative data quality and capacity, the study will likely be restricted to one state. California seems a natural candidate due to its size, diversity, innovation economy, and existing data capabilities.

THE STUDY HAS AT LEAST ONE ANCILLARY GOAL. We would like to study ways of improving data contributors' confidence, willingness to sustain participation, and willingness to provide high-quality data. To this end we propose including a set of experiments in which we randomly vary training protocols, the value and types of incentives offered to participants, and/or the ways in which participants can contol their data provision. The gap study also presents a "sandbox" in which to experiment with new data security technologies and tools for participants to control their data, including those based on blockchain.

WE PROPOSE FOCUSING ATTENTION ON WORK AND EMPLOYMENT, although there are numerous other areas where we might benefit from improved data infrastructure. We will evaluate measurement gaps around the concepts of skills, tasks/activities (both paid and unpaid), contracts/jobs, and occupation.[5] Concretely, we might focus on

[5] See Ahlquist [forthcoming] for definitions. See Autor [2013] for a summary of new theory in this domain.

- The activities engaged in for pay or with the expectation of earnings;
- The number of income-producing contractual relationships;
- The identity of the counter-party for each of these contracts;
- The activities performed under each of these contracts;

- The location of the activities, distance from home, and tools used;
- The ownership status of the tools used;
- The payment level and form, benefits, and degree of autonomy associated with each of these contracts;
- The consistency/predictability of earning opportunities associated with each contract;
- Subjective evaluation of contract stability/continuity and quality (e.g., whether it is meaningful, stimulating, or likely to improve);

The proposed study has four modules. The first module aims to replicate, in "miniature," what we would hope to achieve if we had all the currently available administrative data linked at the individual level. Modules 2-4 will use additional data sources—in-depth, qualitative interviews; smart phone-based micro-surveys; and access to job posting data—as indicators of "ground truth."

We anticipate that participant recruitment would start with a set of individual or household participants for two reasons. First, there are income-producing contracts and activities that would not be visible if we were to begin recruitment with firms or employers. Second, the tools for recruiting people into panels are well-established.

Module 1: the status quo

This module will have a household survey, administrative data, and, ideally, employer components. Some creativity will be needed to entice employer participation.

Module 1a: Standard household survey

The first module will involve a standard long-form survey that emulates an existing tool such as the American Community Survey, perhaps augmented with other questions or additional work-focused modules such as the Contingent Worker Supplement or the SIPP. This survey will be administered to all participants at the beginning and end of the study period.

Module 1b: Employers

Matched employer-worker data is critical. Inducing this matched data in the gap study will be difficult. Some possibilities for addressing this include:

- Asking individual participants to disclose their employers and then surveying those employers or enterprises directly;
- Surveying employers in the same region as individual participants;

- Starting with a diverse set of employers and then working to recruit their employees or contractors.

All have evident downsides.

MODULE 1C: LINKED ADMINISTRATIVE DATA

We will ask individual participants to give permission to link their administrative data, notably tax records for income and unemployment earnings as well as program participation, with the other data collected. We would like to link employers to relevant tax, regulatory, and economic census records. We will need to work with participants and the administrators of current data "silos" to make these non-trivial linkages in a secure fashion.

Module 2: In-depth interviews

A subset of participants (due to cost) will be asked to participate in extensive, in-depth personal interviews at regular intervals. Skillfully conducted interviews can result in important and surprising results not visible in surveys [Morduch and Schneider, 2017, Edin and Lein, 1997]; Abraham et al. [2017] show that limitations in the CPS interview prevent better measurement of "non-standard" work. Success of the interview component will be critical for triangulating the data from modules 1 and 3 as well as asking follow-up questions for better understanding what we see using other tools.

Module 3: EMA

EMA is the acronym for the (awkward) term of art "ecological momentary assessment." EMA consists of short micro surveys and other measurement typically administered using smartphones or simpler SMS [Salganik, 2018, Sugie, 2016]. A subset of respondents will be asked to participate in the EMA arm of the study.

EMA will allow us to ask participants short, contextualized questions as well as regular, slightly longer mini-surveys. EMA measurements can be staggered, randomly generated in time and space, and linked to phone metadata like location (when permission is given). Most importantly, we can ask participants to submit text, voice, image, or video data that can then be analyzed using machine learning tools. Using EMA in this context will require a purpose-built smartphone app as well as anonymization, security and encryption protocols. Participant incentives, question design and timing, and overlap with the qualitative interview component are all important areas for consideration.

Data control and trust experiment

EMA will allow us to investigate an area of broader interest: respondent control over the dimensions of participation. We will take advantage of the gap study to vary training and incentive protocols as well as how the respondents control to their data. As one example, we could randomly assign EMA participants to different "opt-in" conditions. In one arm the default participation level is to provide all the data we seek and they can then choose to opt out. In the other the situation is reversed: they actively opt-in to share particular data. We might also consider varying the training protocols as well the value and types of incentives offered to participants. Outcomes of interest would be participation and retention rate as well as data quality.

Module 4: Digital traces

The fourth module would incorporate a variety of digital traces, some individualized and some aggregated. At the aggregated level we can work with private sector digital media and data companies to link respondents to their local context. Housing turnover and price data from Zillow is one example. Geographically fine-grained (e.g., Census block group, stratified by gender) Google Correlate data is another. At the individual level, we would like access to participants' online job profiles, search, and employer contact record. In this way we can compare a worker's stated background and skills with those posted by an employer. Linking this data with surveys and administrative data is a type of "enriched asking" [Salganik, 2018].

What the gap study is not

The gap study is not a traditional research project in the sense that we are not beginning with a set of a priori hypotheses or scientific research questions that we seek to test. Rather the goal is systematic evaluation of existing data gathering practices when combined with newer tools. Project output will be take several forms, discussed below.

The project necessarily has an inductive component, given the semi-structured qualitative interviews and the (hopefully) surprising responses from EMA. We will need the flexibility to adapt and reorient both as the design evolves and during data collection itself.

Recruiting a representative sample and generalizing to a larger population are not the chief concerns. Rather, it is more important to secure participation of a wide variety of people facing distinct social and economic contexts. This way we are better able to map the social and economic terrain and discover where existing tools perform well and where new tools uncover things previously missed.

THE GAP STUDY IS *not* INTENDED TO BE A WEATHER STATION PROTOTYPE. Rather the, the gap study is meant to highlight what we can and cannot "see" with existing administrative and survey tools under good conditions. It may turn out that some of the triangulation tools used in the gap study prove useful and scalable. But it is neither assumed nor required that this is so.

THE GAP STUDY IS *not* INTENDED TO REPRESENT A REPLACEMENT FOR EXISTING DATA INFRASTRUCTURE. Quite the opposite. The gap study—and any future socio-economic data platforms—will work better the more we are able to integrate with long-running tools such as PSID and ACS and link to existing administrative data such as LEHD and tax data. Linking well-understood conventional data will be necessary to build functional data ingest and QA procedures for always-on data collection from disparate and non-randomly selected sources. Validation against well-understood existing data will be crucial for any "nowcasting" models we might hope to build.

Time line and Deliverables

Figure 5 outlines a proposed project time line. A conservative estimate for the gap study is three years, with the first year involving negotiation with public data providers, project design, recruitment, and testing. There is a year of data collection and a year of analysis. Notice the proposed overlap with the prototyping of a "weather station" system and the gap study.

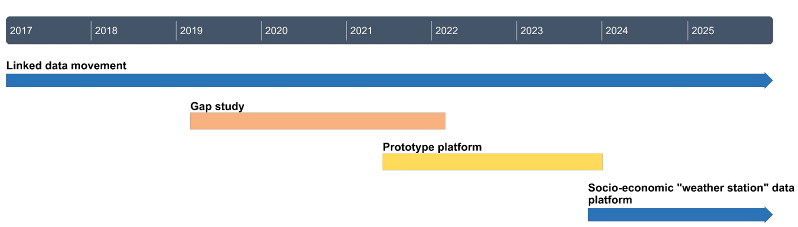

Figure 5: A potential time line for the weather stations project.

Project deliverables will take several forms, depending on the audience and objectives. Standard end-line formal research articles are most likely to emerge around the data trust and control experiments as well as the security and data management components. If we successfully architect a distributed and privacy-preserving system there may be software patents or open source protocols. Code for the EMA research app should also be mad available, either as open source code or a freely licensable tool.

Formal white papers integrating all data sources and reporting our gap findings for each of the identified concepts will emerge at end line. These are the key documents from the study and are aimed at government producers of data products as well as private sector and research community users of such.

Qualitative interviews are a key component of the study. Ethnographic analysis of interview data will be produced for research, program administration, and, perhaps, popular audiences. We might also develop a series of vignettes based on the interviews meant to illustrate a particular data gap that the study identifies. We also expect articles reflecting on the administrative and research challenges in integrating interviews, EMA, and other data sources.

During the course of project research reports on methodological topics are likely to emerge, especially around machine learning, data ingest, and modeling.

Finally, we envision a capstone report that synthesizes all the work product from the gap study, with a particular emphasis on surprises, lessons learned, innovations, and steps for the future.

Project structure

As the description of deliverables hinted, we imagine a modular structure for the project, with different but overlapping teams taking charge of different aspects. We expect that researchers in each of the teams would be able to develop work product independently, but that any dissemination prior to end-line would be subject to approval by the steering committee and the data security/privacy committee.

- Steering committee
- Team with expertise in administrative data, access and linkage
- Team with recruitment, panel management, and survey expertise
- Team with expertise in EMA and mobile data collection
- Team with expertise in in-depth qualitative interviews
- Team with expertise in data security, privacy, and blockchain.

Open questions for discussion

This memo has outlined the contours of a problem and sketched a solution, relying on an analogy to weather stations. There are several immediate areas for discussion:

- Are there any existing public/private data consortia not outlined here or in the appendix from which we could learn or collaborate?
- What are the actual security and privacy tools we might consider experimenting with here? Does this proposal present an interesting "sandbox" for data security researchers?
- Can we induce employers to participate?
- What about payment processors or telecoms?
- What are the legal or other hurdles that we might encounter? Are any of these deal breakers or things that require radical rethinking of the gap study?
- How do we reconcile the inherently inductive nature of gap study with expense and ambition?
- What is the best way to interface with government? Is California the best place to run this?

If we can solve the technical, political, and administrative challenges, the effort and resources put in to the gap study alone have the potential to not only tell us what we are currently missing but also revolutionize how we ask for and manage sensitive information. Building socio-economic weather stations will allow an unprecedented ability to make, evaluate and adapt granular investment and policy decisions.

References

J.M. Abowd, J. C. Haltiwanger, and J. I Lane. Integrated longitudinal employee-employer data for the united states. *American Economic Review*, 94(2), 2004.

Katherine G. Abraham, John C. Haltiwwanger, Kristen Sandusky, and James Spletzer. Measuring the gig economy: current knowledge and open issues. March 2017.

John S. Ahlquist. The future of work: Risk bearing and risk sharing, Sept. 3 2015. URL https://psmag.com/economics/the-future-of-work-risk-bearing-and-risk-sharing.

John S. Ahlquist. Research frontiers in the instutional analysis of work. In Claude Menard and Mary Shirley, editors, *The Research Agenda for New Institutional Economics*, chapter 22. Elsevier, forthcoming.

Imanol Arrieta Ibarra, Leonard Goff, Diego Jiménez Hernández, Jaron Lanier, and E. Glen Weyl. Should we treat data as labor? moving beyond "free". *American Economic Review Papers and Proceedings*, 108:38–42, 2018.

David H. Autor. The "task approach" to labor markets: an overview. *Journal of Labor Market Research*, 46(3):185–99, 2013.

Sam Balter. "I'm a professional gamer". Weird Work podcast, 2018. URL https://www.hubspot.com/weird-work-podcast.

Michael Behar. Why isn't the u.s. better at predicting extreme weather? *New York Times Magazine*, https://www.nytimes.com/2016/10/23/magazine/why-isnt-the-us-better-at-predicting-extreme-weather.html, 2016.

David Card, Raj Chetty, Martin Feldstein, and Emmanuel Saez. Expanding access to administrative data for research in the united states. Technical report, National Science Foundation, 2010.

Peter W. Cookson Jr., Kathryn Edin, and David Grusky. National poverty study. http://www.nationalpovertystudy.org.

John L. Czajka and Amy Beyler. Declining response rates in federal surveys: Trends and implications. Technical report, Mathematica Policy Research, 2016.

Kathryn Edin and Laura Lein. *Making Ends Meet*. Russell Sage Foundation, New York, 1997.

Barbara Entwisle, Sandra L. Hofferth, and Emilio F. Moran. Quilting a time-place mosaic: Concluding remarks. *Annals of the American Academy of Political and Social Sciences*, (669):190–198, 2017.

Luis Garicano, Claire Lelarge, and John Van Reenen. Firm size distortions and the productivity distribution: Evidence from france. *American Economic Review*, 106(11):3439–3479, 2016.

Deborah Goldschmidt and Johannes Schmieder. The rise of domestic outsourcing and the evolution of the german wage structure. *Quarterly Journal of Economics*, 132(3):1165–1217, 2017.

Seth D. Harris and Alan Kreuger. A proposal for modernizing labor laws for twenty-first-century work: The "independent worker". Technical report, Hamilton Project, 2015.

McKinsey Global Institute. Independent work: Choice, necessity, and the gig economy. Technical report, 2016.

Jonathan Morduch and Rachel Schneider. *The Financial Diaries*. Princeton University Press, Princeton, 2017.

Christina Pe na. From patchwork to tapestry: collaborating to maximize data utility. Technical report, Workforce Data Quality Campaign, 2017.

National Research Council. *The National Weather Service Modernization and Associated Restructuring: A Retrospective Assessment*. National Academies Press, Washington, D.C., 2012a.

National Research Council. *Weather Services for the Nation: Becoming Second to None*. The National Academies Press., Washington, D.C., 2012b.

National Research Council. *Nonresponse in Social Science Surveys: A Research Agenda.* National Academies Press, 2013.

OECD. Statistics, knowledge and policy: Key indicators to inform decision making. Technical report, 2004.

Tim O'Reilly. *WTF? What's the Future and Why It's Up to Us*. HarperCollins, New York, 2017.

PEW Charitable Trusts. How states use data to inform decisions, 2018.

Matthew J. Salganik. *Bit by Bit: Social research in the digital age*. Princeton University Press, Princeton, 2018.

Social Observatories Coordinating Network. URL `https://socialobservatories.org/`.

Marylin Strathern. 'improving ratings': audit in the british university system. *European review*, 5(3):305–321, 1997.

Naomi F. Sugie. Utilizing smartphones to study disadvantaged and hard-to-reach groups. *Sociological Methods and Research*, 2016.

David Weil. *The Fissured Workplace*. Harvard University Press, Cambridge, MA, 2014.

Appendix: Cognate studies and initiatives

- Longitudinal Employer-Household Dynamics (LEHD;[Abowd et al., 2004]). Job-based design frame. Based on state-level UI reports. starts in 2003. Quarterly Workforce Indicators. Links UI data, employer filings and supplementary census and surveys. Nothing about contracts, tasks, etc.

- Financial Diaries project [Morduch and Schneider, 2017]. Expensive, one-off, non-representative, bottom 1/2 of income distribution, only finances, nothing about work, tasks.
- National Poverty Study [Cookson Jr. et al.]. Poverty-focused; bottom of income distribution, not about work/occupations.
- Equality of Opportunity Project. Focused on "harnessing big data" for research and policy reports. Work, employment, and contracts not an explicit area of research.

Appendix: Agencies, Think tanks, and foundations in this space

Agencies

- Census (ADRN, CES)
- IRS, Social Security
- Department of Labor (CES, BED, QCEW)
- WIBs

Think tanks

- JPMorganChase Institute.
- McKinsey Global Institute
- RAND

Foundations

- Arnold
- Irvine
- Russell Sage
- Sloan
- Casey
- Upjohn
- Lumina
- US Chamber of Commerce