

2017

Erica Groshen Interview

Justin Carinci

W.E. Upjohn Institute, carinci@upjohn.org

Citation

Carinci, Justin. 2017. "Erica Groshen Interview." Kalamazoo, MI, W.E. Upjohn Institute for Employment Research.
<https://research.upjohn.org/externalpapers/76>

This title is brought to you by the Upjohn Institute. For more information, please contact repository@upjohn.org.



Do you see any role for private third-party data supplementing or supplanting official government labor statistics? If so, what it would be?

I absolutely do. I would actually use something in between supplementing and supplanting: I would say a symbiotic relationship is where we are heading. Administrative data, data that's created as a byproduct of automating various systems in our economy, is very useful, we learn a lot of things from it, but it doesn't have some of the features of official statistics. So in order to make the transition from this big data, this organic or administrative data to a statistic you can use, you almost always have to use official data, which is based on a representative sample that has known properties. So, the future I see is more and more use of these other private third party data as a way for people to get products very tailored for their own purposes that official statistics go only part of the way to address, so there will be an increasing need for official statistics at the same time you have this explosion of other sources as well.

The other way the statistical system benefits from having all this third-party data is there's a lot of innovation and exciting activity going on in that world and the statistical agencies can learn from. Things like autocoding, small domain estimation, highly computationally intensive work, are things the statistical agencies have been able to adapt in many cases for use within the system.

So private groups can bring innovation, but may lack rigor or consistency?

Private third party groups will never have the history that the statistical agencies have. They want to preserve their "secret sauce," so they're not going to be as transparent. They don't

have the requirement to document in the same ways the statistical agencies do. That stuff's expensive and they're just not going to do it.

But they can really augment what's being done in the statistical agencies and they can both use it for validation because one of the things you have to do with any statistic is try and figure out whether or not it makes sense in light of other things you think you know about what's going on. So, that validation work is very important as you get these other alternative measures, trying to bring them back and compare them is an important part of comforting yourself that you've found something—or discovering that you actually have not found something!

It is well documented that response rates to surveys have been dropping, especially for household surveys. What actions or considerations did BLS take to combat this decline? Some people have suggested greater use of administrative data, but what is your experience with the feasibility and cost-effectiveness of that approach?

I think we always have to remember, even when it's mandatory and especially when it's not, that participation in these surveys is a public service people perform for us and the statistical agencies are becoming more and more aware of the need to recognize and thank participants for participation. Of course, confidentially, because you have to preserve confidentiality. We have to remember that the survey response rates for federal surveys are higher than you find anywhere else, much higher than you find in the private sector. The pitch is always a three-legged pitch: it's really important, it's safe, and we make it as easy as possible. BLS has extensive regional network, has highly trained, ever more professionalized field staff that go out there and solicit participation. The key thing to maintaining it these days is lowering the burden and convincing people it's safe and counteracting any aspersions it's not worth it because the data's garbage or something. Currently, there are a lot of attacks on federal statistics as being useless.

So, the solution is to lower the burden and one of the things the BLS has been pioneering in is offering many response modes to lots of its surveys, particularly the employer surveys: so if you want to do it on paper, if you want to fax, if you want to call it in to a call center—offering many modes of response. Other things are to key off their own internal system as much as possible, so to work with intermediary companies such as payroll service providers and software providers.

Something BLS is working on now is how do you get it so you can say to a company: you want to participate and you use this payroll provider, “this is the button you push to produce the report that you then send us.” They don't have to do a bunch of additional work, they just have to elect to generate that report and transmit it to BLS.

There are also other methods using administrative data. Administrative data is a huge opportunity, but it's not a panacea. The quality may not be where it needs to be: if you're relying on fields that nobody cares about, chances are the information those fields is not very

good. The information in those fields may not be what the statistical agency really wants; it may be close but not exactly.

Then there's the problem of access. There are lots of administrative data the statistical agencies don't have access to legally or because of the turf consciousness. But it's a route and there are some great opportunities out there, particularly the unemployment insurance data, the wage records from those. There's sharing data across the statistical agencies, so right now census and BLS both have employer registers for the country and they're not the same.

Machine coding is a way of reducing burden on respondents as well. Instead of asking companies to give us your standard occupational codes, you can say to companies don't bother doing that, just give us your occupational title and BLS would use machine coding to convert that to SOC codes.

There are a lot of different methods. One other important thing agencies need to do is continually pay attention to their presence, make sure they communicate well about the importance and why the respondent can trust the agencies to do the right thing with the data, to protect it and to turn it into info that's really important to the respondents. To the extent you can use administrative data you should, but it will never answer all the questions you need. There is just some information you're not going to get if you don't ask people. You may be able to find out where people live and what they earn, but you won't find how they're using their time, you're not going to find out what their gender preference is, things like that, using administrative data.

There is growing attention being paid to the future of work, and concern that changing work relationships—for example, independent contractors, contract agency workers, gig workers, app-based workers, etc.—are evolving faster than BLS can develop the tools to measure. How has BLS considered collecting data to document these forms of work?

The main thing BLS has done is run a supplement to the current population survey, called the Survey of Contingent Work and Alternative Work Arrangements. BLS was doing this about every two years around 1995, then funding issues meant BLS has not been able to run it since 2005. It's been asking for appropriations to be able to do it, but not receiving them. But the Department of Labor stepped forward to fund it in 2017. It was fielded May 2017. The data will be out probably early 2018. The survey is pretty much the same as what was done in the past, so you'll be able to document the trends in it, which is really important. And there have been two questions added: was this contingent work or alternative work? Was it mediated through an electronic platform, with matching, or was the work done electronically, on a computer, so you didn't go anywhere, you just did it? (This is the Mechanical Turk kind of work and the first one is like Uber). Those questions are on it and that will be interesting. But that's a one-shot deal, and BLS is going to need to have the resources to follow it; that's a concern: that it's a one-shot deal.

There's one other case in which this has been documented to some extent. In the census of fatal occupational illnesses and injuries they added a field documenting whether the person who was killed was a contractor and from that we learned contractors have a much higher fatality rate than regular employees. What's interesting about that is that contractors probably don't have as much training and they probably have much lower tenure on any kind of work than a regular employee, so it does suggest they may well be doing more dangerous work but also their circumstances just inherently make things more dangerous because they're less likely to be trained.

The gaping hole is in understanding how employers see these alternative works and how they've changed their attitude toward how they obtain labor. It's very interesting to see it from the household side: what kind of work are people doing, what are the characteristics of the workers—but you also want to know which employers make the decision for what reasons and how the performance of firms changes when they adopt these strategies. And in order to do that, you need to survey employers about that. That's really nascent; it hasn't been done by BLS or very much by anyone else before. BLS is interested in doing that, and one step toward doing that is changing the Quarterly Census of Employment and Wages Annual Refiling Survey to enrich it to make it a random sample each quarter and then being able to add some questions about other things but also about these sorts of things, the employment practices of establishments. That, I think, will be very illuminating. BLS is making the changes in those directions but it will need additional funding to actually be able to go after this set of questions.

So the need increases and the funding goes away?

It's really unfortunate.

We have a puzzle with the current population survey. This has been going on in the same form for 50 years now, more or less the same questions and they're a set of questions that you would think would pick up a lot of this alternative work and contingent work. There's a question about do you work part time? Do you hold multiple jobs? There's a question about are you self employed, so you would think this set of questions would show if we're having this burgeoning of alternative and contingent work arrangements, you think that would see any one—or all three—of these growing and you don't.

So that's a puzzle that people have been looking at. Some of that work has been done right here at the Upjohn Institute by Susan Houseman, trying to get at, is there a lot of smoke but no fire? or are the respondents not understanding the questions in the way we think they should be, so they're not answering? Some of the preliminary evidence suggests that gig work itself may not be increasing but alternative work structures probably is increasing more, and it may not be so much through dual job holding but just from people being contracted out through companies. But there is some evidence that people aren't thinking about work for which they receive 1099 forms as work. They think of it as a way they got some more

money but it wasn't really a job. Cognitive research is needed to figure out how people are thinking about it so you can structure the questions properly to pick up.

Federal statistical agencies have experienced multiple budget freezes and cuts in recent years, even as the demand for better data has grown. If certain programs have to be cut, how are which ones to cut decided? Is there the possibility of cut programs being restarted in the future?

The basic thought process on eliminating a program at BLS is protection of the principal federal economic indicators. These are designated by the Office of Management and Budget as the main ways we track what's going on in the economy on the national basis. BLS does not cut those easily. It also protects programs that are written into law, so people have to use it for something.

Also state unemployment rates are, there are triggers written into extensions of unemployment insurance programs by the states by what's happening to the unemployment rate in the states, so if BLS stopped producing state unemployment rates ...

You have this set of programs that are either PFEIs or written into law, so the places you look to cut are none of those. Unfortunately, those tend to be programs on which there's a lot of research done: National Longitudinal Survey, the Job Openings and Labor Turnover Survey, the American Time Use Survey, and the Employer Benefits Survey are the four programs that are outside those categories.

For one thing, they're a very small proportion of what the BLS does. Well over 80 percent goes to the protected programs so there isn't that much left over. These other programs are programs where we develop a long-term picture of what's going on in the labor market, where we understand the big trends, and to lose them is pretty damaging.

That said, if BLS has to make cuts, it has proposed various times eliminating the National Longitudinal Survey, not because it's not hugely valuable, but because, of this set, it's a fairly large program, \$12-13 million whereas JOLTS is \$5 million, so you're going to get more bang for the buck cutting a large program than a small program. But defenders of the BLS have been quite vocal and have helped to protect it, which is great, because it is a really valuable program. But it does put BLS in a bind. So if you're going to look to cut the programs, the ones that are most vulnerable are those four, except that their supporters are quite energized to protect them because they know that they're close to the chopping block.

Then there's trying to shrink programs. There are two general ways to shrink a program. One is to cut the sample to make it smaller so then you lose detail and your standard errors rise. But that's one way. You often don't gain very much because the fixed costs of running the program are still there.

Another thing that's been done is to cut the periodicity, so National Longitudinal Survey used to be annual, now it's two years and there have been proposals to make it every three years, so that can save you some money. But both of these come at the cost of degrading the data. Sometimes programs have been combined when there's been an opportunity, but that's a long process. But has been done in the past.

What is very difficult to do is to end a program and then start it up again, because you lose that information in between, you lose the institutional knowledge about it and there's a lot of training that goes on with starting a program again. If you had a long gap you could save money in between. But if there's a short gap it costs money to run a program down appropriately so people can still get access to the data and then to get it started up again, you're hiring and retraining and all of that. So just mothballing is not easy.

You describe these programs as existing to provide better data to make better decisions. It would seem, then, that cutting them would result in increasingly worse decisions.

It's really tricky because statistics are—I know it's shocking to hear in a place like this that is founded on the opposite—but to most people, statistics are kind of boring. The elected officials, by and large, are not running on platforms saying “I'm going to make sure you have access to the statistics you need to make good decisions.” That's not a popular platform. What they're running on are, more, either shrinking government or government needs to train people or protect workers or provide good defense or good police force, whereas these kinds of public goods—this good government aspect of our political system—has not been the focus of the current generation of policymakers.

It's a classic public goods problem. Everybody uses it, it would be undersupplied by the private market, but few people are paying attention to ensuring that it continues to exist. Companies lobby all the time to ensure that they don't pay more taxes than they want to pay and they're not regulated in ways they think are inappropriate so they spend a lot of effort on that and very little effort on ensuring the data they use on a regular basis to make their decisions is as good as it should be. They just leave that to someone else they just assume that will happen. It's not front and center. It doesn't seem very burring to them and that's a mistake. I think that's a real problem. Statistics are relatively cheap, the whole budget of the BLS is \$600 million, it's less than \$2 per person in the economy