

Federal Reserve Forum on Consumer Research & Testing: Tools for Evidence-based
Policymaking in Financial Services, November 9, 2010
Panel Two: Exploring research methodologies for consumer testing and studying consumer
behavior

Joe Garrett:

Good morning. I'm going to talk a little bit about some advantages of -- maybe let me adjust this here. I promise I'll put it back down for everyone else -- some advantages of online probability-based surveys. And, you know, it's not that there's other modes such as mail, phone, mail out, mail back, those are pretty well known, but I wanted to just share with you some advantages as I see this so that you can put this in your toolbox of techniques. I am not an economist. I'm not a financial expert. I have many financial documents in my house that can attest to that last statement. I am a statistician. But I want to talk first about some caveats before I get to advantages.

When you talk about online surveys, there's basically two approaches to inference. There's what I will call a design-based approach, where you attempt to design a probability-based internet panel. You use sample design theory and methods for not only the selection of the panel, but for also the estimation and the variance estimation procedures. And you try to cover the target population as best you can, and/or a representative of that approach. But there is another approach, and I'll refer to that as a model-based approach. And there's many firms that use this approach. You have a volunteer or an opt-in panel, usually of internet users and then, when you get to the inference stage, what you try to do is you correct for any representational bias because you didn't have a formal sample design infrastructure using propensity scoring or other weighting methods. And, as I said, there's many online panels that are that way. My comments today pertain only to the former, not the latter design-based panels. So, you know, for these remarks, I think the latter may have some inferential disadvantages that may outweigh some of the advantages of online, but here I'm assuming that there is a probability-based online panel. You've resolved most of the statistical coverage problems. And what that means is you're doing random recruitment from a frame that essentially covers the US population. When we first built our panel, we used random digit dialing. We have since moved away from that for continual maintenance and sample replenishment for some reasons that I'll just mention quickly, here. But, you also have to have some provision if you're doing online surveys in people's households for what you do for what I will refer to as the non-internet households. We resolved that by giving people a laptop and paying the internet service provider. Jeanne had mentioned earlier about, you know, many people do telephone surveys. I've done many telephone surveys myself. One of the problems with RDD telephone surveys, increasingly, is how do you cover the cell phone only households. I don't know. It's been interesting, speakers talking about their kids. My kids, I don't think, will ever be land line households. When they went to college, they were cell phone only. They're now out of college and have a job. But they still have cell phones. You know, they do

not have land lines. And that's about 25% of US households. Non-internet households is about 30% of US households. So you've got to make some provision. Now, what we've accounted for the cell phone only households is we've switched to a frame using the US Postal Service's delivery sequence file, which is essentially how you get your mail. It's a pretty complete list of all households in the US. And then in this country, increasingly you should be recruiting at least in two languages, in English and Spanish. I'll give a reference here, um, there is a good paper that was done by John Krosnic [assumed spelling] and Chang from Stanford University, where they compared what I have referred to as a design-based to the model-based. And they looked at RDD, which is probability-based, they looked at knowledge networks, which is probability-based and then he compared that to five or six or seven, what I'd call model-based approaches. I won't summarize their paper other than to say that design-based out-performs the others, and sometimes markedly so when you look at benchmark measurements.

Now, the second page of caveats is smaller. I just want to mention. I'm going to talk about some advantages just so that you may have an idea in your own research where this might be useful, where it might be used. But these advantages sometimes apply only to certain modes. It may apply relative to mail or it may apply relative to phone. It is not a blanket statement that it applies to all modes. And some advantages could be experienced in other modes, but having an online mechanism may maximize that benefit. This is not the best taxonomy. It's a useful taxonomy. I've put the advantages in these five buckets. I'm going to talk about survey instrument complexity, data quality, cost effectiveness, schedule efficiency and respondent experience.

Some speakers previously have talked about disclosure materials and other text type. One advantage of an online survey, certainly relative to phone, but you can do this in mail, is you can embed pictures or you can embed print materials in the survey instrument itself. You can also, if you have -- if you're trying to see if there are certain explanations of complicated features that may resonate better with respondents, you could also do audio or video clips. You can do mouse-overs. This is somewhat more difficult in a phone survey world where you have complex definitions going on, and you know, in the finance world, there are a lot of terms that are not commonly understood. One of the advantages in an online world is you can have the mouse where you can program it so the definition comes up. It's available for people. If you have a series of discrete choice experiments, which may be useful, where you're randomizing the treatment, and the trade off options are somewhat complicated and may be complicated to do by phone, programming in an online mode is relatively easy to accomplish that for things like discrete choice.

These last three things I'll cover very quickly, but they're very similar to phone surveys, or at least computer-assisted telephone interviewing surveys. You can build in automatic edits, you know, if numbers are supposed to sum up to 100 percent, you can make sure that's true. You can keep people from proceeding. You can have logic checks. And you can also, obviously, build in skip patterns, like you would in CATI, but somewhat more complicated in print.

One of the things I think is useful for surveys in the finance area is this notion of social desirability bias and different people in different fields may refer to this as a different term. What I'm referring to is if there is kind of a socially desirable answer, and you have an interviewer present, do you tend to bias the result? If you ask someone: Well, do you feel comfortable that you will have enough money to retire in a certain year? That may not be the best example, but you may get bias in that answer when you have an interviewer present as opposed to something like mail or in an online mode. I'm giving one reference.

Now, I don't know whether -- I know we signed speaker release forms, so my comments will be available. Jeanne, will the slides also be available at some point? Okay. Here's a reference that you can get when those are available if you wanted to look at social desirability bias when they look at computer-assisted telephone interviewing versus interactive voice response and web surveys.

Another advantage, I think, is you know, if you're doing surveys where you want to people to consult records, financial records. They may need to go look at 401K statements. This would allow time within some reasonable period of time. When we do our surveys, you know, we do not require them to respond right away like you may do in a phone survey even though you can have call backs in the phone surveys, but this allows people time to consult records if needed. What I'm calling message testing, but it could be more broadly defined of any time where you have a narrative text that you want people to react to. One of the things we can now do is, we're getting ready to do a survey now not on financial areas, but for Yale University on global climate change messaging where you can highlight, in a certain color, messages that resonate positively with you or you cognitively understand those well. And you can also highlight in a different way those that aren't resonating. When you have a panel, you can do longitudinal surveys. In our case, it's probably useful to have these longitudinal surveys where the waves are not 10 years apart. There is attrition in a panel. But if you're doing a longitudinal survey and you're going to do a baseline and then maybe if there's an intervention, and then you're going to come back and do a survey in six months or nine months or 12 months, it's a very useful. We can keep track of who's completed the survey, et cetera. And then, I had mentioned this earlier under survey instrument complexity, but it also enters in under data quality.

Now, in cost effectiveness, phone surveys and even mail surveys are getting more costly to do, and I was talking with someone just right before I came up here who is doing surveys of the unbanked and it would seem to me that that would be a difficult population to get by phone where you may have to go through screening to identify those people who are unbanked and then to get the response that enables them to respond to the survey. Most online survey providers have what, at least we refer to as profile data and you can target the cases you want to look at. One of the requirements for joining in our panel, for example, is that you have to give these data right here. I don't know whether you can see this pointer, but we call this our core demographic and socioeconomic data. This is not a complete list, but it's kind of the major one. So, if you wanted to do a survey in financial issues where you wanted to look at, say, low education, or if

you wanted to look at low income, we collect these data annually. You wouldn't need to go through screening procedures to identify them to a phone survey. We collect household income. This is some of the data we collect annually in our financial services profile. In fact, the person I mentioned I was talking to right before I came up here to speak, we have just added a question for her on the unbanked, using, I think, the FDIC wording of that question so we can capture that information. But we have this information also that would allow you to get to the panelists. And again, it's cost effective in the sense that you don't need to screen for this. Online, you know, compared to mail, not compared to phone, but compared to mail, can be cost effective in the sense that, you know, there are no paper costs.

Mailing cost, some of the recent mail surveys that I've looked at, I'm surprised just how expensive mail surveys have become in terms of sending out documents, sending out reminder postcards, sending maybe even out the instrument again. So there is some cost savings relative to that. This is probably more true to in person surveys. I know not too many people do in person surveys. When I started out in the survey research business back in the mid-70s, there were a lot of in person surveys. There are still major federal surveys that are in person, but with an online mode, geography costs go away and geography costs can be a big issue with in person surveys.

Survey follow-up costs. I think this is an advantage relative to both telephone and mail. It's much cheaper to send an email than it is to call someone or to send them a postcard reminder. This advantage is probably relative to mail. I've had experiences in mail surveys where we've done an instrument, you find out at some point after you've mailed out the instrument you need to make a change in the instrument. That's somewhat complicated on many levels, and it requires re-sending the instrument. In an online environment, and probably in a phone environment as well you can do this, but you can stop people from accessing the survey. You can make the instrument change and then you can open up access to the survey and they can proceed on.

We are starting to get, and I think other online firms are starting to get, increasing requests to do not just quantitative surveys, but also qualitative surveys. Doing focus groups online, you know, you miss the in-room dynamic of the conversation. You can, though, benefit by not doing Denver, Pittsburgh, San Francisco and Dallas all in one week as you're travelling on focus groups. It can be cheaper. You can record, you can do instant polling. In an online environment, you can also do larger than say the typical size focus group or nine or ten or 12 people. We've done them as large as 100 or 200. If you do it in that realm, it helps to have more structure to do that, where you're polling people through a structured instrument and then polling for the results.

I have to apologize for this next set of bullets. This is what happens when you have your graphics department review your slides. They saw all that -- this is supposed to be a race where everyone's running real quickly and then they wanted to get across this idea that this is speed. So, as a statistician, I can guarantee you, I'm not jazzy enough to come up with something like

that, so, I left it in. But unlike phone where you ramp up with a cadre of interviewers and then work through the respondent list, if you wish, this survey can start instantly. Email invitations can go out to thousands of people in the survey, or you can do a more replicated approach. Maybe you want to see what the first 100 responses look like. So you release an initial sample, you look at this, you see if you're going to make any modifications. But, I think online, one of the advantages is that data collection can be shortened. Data cleaning is faster because you have these -- somewhat like CATI -- where you have, you know, these online checks and skip patterns, data processing is faster. And this is the last one, I promise. These all means you can get data delivery faster.

A typical schedule for us is about four weeks from the time we receive a survey instrument to the time we give you a data file. We've done surveys as fast as one hour. We did all the polling for CBS after the presidential debates. When the debate ended, we did a survey. They had an analysis program that ran after that. We had to get them the results within the hour. I don't necessarily recommend doing surveys in an hour. You can do this with planning, but you also -- you lose a lot of sleep over those things.

Just respondent experience, I don't know how you are, but whenever I get a telephone survey, you know, being in this business, I try to respond to these surveys. I mean, but I can tell you, it's not always the most convenient time I'm contacted. And I realize they ask: Well is there a more convenient time? And usually the answer is: Probably not, we might as well do it now. And so, but I do think there is this decreased social receptiveness to having this, what I will call an interference at home, being called on the phone, getting something in the mail. So, you can kind of avoid this interference. One of the things about randomly recruited panels that are design-based or what I referred to as design-based, they have an expectation of being surveyed. This isn't a cold call. So, you know, and we don't give people unlimited time, our times that we allow respondents to complete a survey is based upon the client request. But we would typically give people 10 to 21 days. That's a typical field period. We get a lot of respondents doing things over the weekend. And, you know, so I do think when you can choose when you want to respond to the survey, there is this kind of convenience issue that I think is useful for a respondent's um -- I think you get an overall better survey experience. I think when you start thinking of the capabilities of the web in terms of -- I've only mentioned font sizes and colors here -- but one of our speakers will talk about how you can have some software that actually shows you the person now, and what you might look like when you get older and how that may influence some of your decisions. And I think this is another slide I've got to apologize for. They slipped in here again. But it is, for those of us who care about kind of going green, it is a greener way of doing surveys, certainly than mail.

But, in closing, let me also mention some disadvantages. I think non-scientific online surveys proliferate. In fact, sometimes people may think of an online survey as being non-scientific, and that goes back to, then, my earlier caveat. I think this statistical inference problem is a big issue. I think you do need to be design-based. I think you, if you want to do statistical significance

testing or if you want to develop confidence intervals, you know, those things just don't apply to non-randomly selected samples. Another problem we have with a panel, our panel has 50,000 people. You know, it's a lot of people, but obviously, it's still finite. If you do very low incidence surveys, let's just take 2%. 2% of 50,000 is, say, a thousand, you get below 2%, it starts cutting into your sample size. It may not be enough for it to satisfy, you know, the research objectives. Small areas in geographies are another thing. If you wanted to do a survey, say, in Fairfax County or Montgomery County near here or in Washington, DC, a nationally representative panel has the number of cases it has in these that's going to be roughly proportional to population, but we probably aren't going to be able to take a 2,000-person survey in Fairfax County. Sometimes you need an interviewer and for those surveys, online would be at a disadvantage. If you need an interviewer present, you know, if it's somewhat cognitive-based or you're looking at user type things.

So, I just wanted to close with some of the disadvantages, give you some idea of what I think some of the advantages might be and kind of ask you as you think about your own research needs and also here at the Federal Reserve, some of these things may be useful for some of the times when you wish to collect data at the consumer.