HOW MANY PEOPLE PLAY THE LOTTERY?

The first thing to value about research is the reliability of the result, not its positive or negative implication.

STEPHEN WADE, RESEARCH AND DEVELOPMENT MANAGER, WASHINGTON'S LOTTERY AND PRINCIPAL, LOTTERY MANAGEMENT CONSULTING, LLC.

About half the population, according to several recent polls.[1,2,3] "Half the population plays with us" sounds very reassuring.

Early in my career, however, I sat down with a tough-minded lottery manager who let me understand that he put very little stock in these numbers coming out of survey research.

"Look" he said, "I have no trouble with the idea that about half the population will buy some sort of lottery ticket some time in a year. That's nice, politically. But what I really want to know is, "How many customers really are driving the business?" Some of these folks literally buy one ticket a year, and yet we know some play nearly every day. I would be much more interested to know something like "how many people buy a Lotto ticket at least once a week."

I had recently transitioned from my early career in the "hard" sciences where much of my daily effort went into rigorous effort to obtain reliable data. I had spent many hours operating delicate instruments to measure biochemicals, or following a precise study protocol to study animal learning. Here in the lottery world, obtaining data seemed trivially easy: just ask people what you want to know. I knew that I had inherited a tracking study survey that asked people precisely the sort of thing this manager wanted to know. Pretty soon, I was able to tell him, "Good news: twenty percent of the population plays Lotto every week!"

"Is that right," he said. "And how many people would that be, in this state?"

I did a quick calculation on the basis of the over-18 population (since that age criterion applied to both lottery play and participation in the survey), and came back with a figure of about one million.

"So if we have a million people who play every week, how come we never sell more than 200,000 tickets?"

"I didn't realize that," was all that I could say. "Let me get back to you on this."

Odd as it may seem, this is when lottery research became interesting for me. Developing methods for measuring things reliably had been a big part of my earlier career. The count of tickets in a drawing was a precisely known fact, even though it did not tell us anything about the players. The tracking study was meant to tell us something about the players, but if it produced an estimate of weekly players that was at least five times too high, it was clearly unreliable. This manager was justified in mistrusting the survey results. I took it as my task to understand where the tracking study was going wrong, and hopefully to fix it and restore trust in it. Either that, or to discredit it and abandon it. Preserving my personal credibility was my main concern.

Several years later, I am still using tracking studies to understand something about player behavior, and I hope I have preserved my personal credibility. I found some problems that I have tried to fix. The rest of this article deals with two important fixes: one in the survey, and one in its interpretation.

An important fix in the survey was in adding something that was missing. In jackpot games, like Lotto or Mega Millions or Powerball, we know that while some people play very consistently, others play sporadically, and still others in a way that depends on jackpots. My original survey had questions like "How often do you usually play Lotto?" It did not have questions like "At what jackpot level do you start to play Lotto?" You might expect that people who only play when the jackpot is very high would say, "I play less than once a month." But how about the person who only plays when the jackpot is very high, but then plays every draw until it is won? When we added a question like "at what jackpot do you start to play Lotto," we found that there were plenty of people who would say "I play more than once per week" and also say "I start playing when the jackpot is above $6 million." In Washington Lotto, this would lead to an inference that they play every week, but only in about one week out of five. The rest of

the time, the jackpot would be too low for their personal threshold.

Adding this question let me move closer to understanding how people really play. There was still a large gap, though, between any estimate based on the tracking study and the actual business result of ticket count. This led me to the second fix, in interpretation.

When we survey a few people, or a few hundred people, we seldom wish to find out only what those particular people think. Usually, we hope to find out something that can be applied to a bigger population. We make an assumption that the people we talk to somehow represent that bigger population. This is an important assumption, and one that is often dubious.

Consider how this assumption fell apart for telephone surveys in the early years of this century. At one time, Random Digit Dialing (RDD) was considered a good way to get a random sample of the whole U.S. population in contact with an interviewer. After all, practically every household had a telephone that was part of what we now call the "landline" network. Before the turn of the century, this was just called "the telephone network." Mostly, these phones just rang

when a call came in; they did not identify the caller. People tended to answer their phones. A skilled interviewer might be able to get something like a representative sample of people-with-phones to participate in a survey.

Many things have changed since then: landlines got abused by telephone marketers, people got caller ID and answered their calls selectively, "do not call" registries were established, mobile phones became widespread, and many households abandoned landlines altogether. People who answer landline telephone calls from unknown callers are now much less representative of the general US population.

Survey professionals, of course, have adapted to these largely technology-driven changes. Surveys are now often done with Web-based panels, or with a combination of landline and mobile phones, ensuring that the respondent initially contacted by the surveyor can be more representative of the target population. The surveyor always asks some questions up front to screen out respondents who are not part of the target population.

However, not everyone who is contacted and who passes the initial screen completes the survey. This is not a function of technology - it was true in the old RDD days and it is

Table 1

NAÏVE INTERPRETATION OF WASHINGTON FY15 TRACKING STUDY

Play Frequency	Survey Respondents	Each Respondent Represents	Per Capita Annual Spending	Group Annual Spending	Eligible Population %	Players
Weekly	329	1,561	$342.88	$176,291,245	9.1	514,150
Monthly	357	1,561	$61.82	$34,578,372	9.9	559,350
LT Monthly	299	1,561	$7.35	$3,445,670	8.3	468,950
Totals:				**$214,315,287**	**27**	**1,542,450**

true now. It is also not the fault of any particular survey provider. Rather, there is always some self-selection among the respondents as to who completes the survey. Skilled interviewers and thoughtfully-constructed incentives can partially mitigate this, but basically, people prefer surveys that are about something that interests them. The more demanding the survey, the greater the self-selection.

This is how I now understand the issue that manager identified for me: people who complete a lengthy survey around lottery play tend to be lottery players, and especially avid lottery players. Unless we somehow recognize this in our calculations, we may incorrectly assume that they are representative of the general population, and this may lead us to dramatically over-estimate their numbers. We may become victims of "survey participation bias."

After a few tries, I developed a way of dealing with participation bias in lottery tracking surveys that has served reasonably well over a period of 10 years or more. Basically, I use the self-reported survey data and assumptions about survey participation to estimate annual spending, and insist that this estimate must match what we actually sold. I find assumptions about survey participation that make this true.

To estimate spending, I rely on self-reported frequency of play (including jackpot sensitivity if applicable) and customary spending from each respondent, and I build up an estimate of annual spending across all the respondents on each game. Note that this person-level data comes exclusively from the tracking study.

I then extrapolate this spending from the respondents to represent spending across the whole population. If I use the naïve assumption that the respondents are a representative sample of the population, and my FY2015 survey of 3,618 respondents represents 5.65 million Washington adults, I would say: "Each survey respondent represents 1,561 potential players." The estimate of Lotto sales using the naïve assumption is high by about fivefold for Lotto: actual sales in FY15 were $44,462,100.

Now to apply the idea that the more avid the player, the more likely to complete the survey, I need to be able to recognize the more avid players. I choose to do this on the basis of play frequency, making groups of the players who play every week (weekly), less frequently than every week but most months (monthly),and less often than that (less-than-monthly, LT monthly). I can then account for self-reported spending from each of these groups independently.

The weekly players contribute the most spending by far – if I apply the naïve assumption that each respondent represents 1,561 similar people in the population, weekly players alone account for more than four times the amount of Lotto we actually sold in FY15. The monthly players, under the same assumption would account for nearly 80 percent of what we actually sold. The LT monthly players account for less than 10 percent of what we actually sold. **Table 1 summarizes this result.**

Clearly, in order to make the survey results correspond with reality, I could change my assumptions about how many similar people are represented by survey participants who played with weekly or monthly frequency. The LT monthly players do not account for much spending, and this is in line with the sense that they are not much engaged with the lottery and are not much more likely to complete the survey than non-players.

The weekly players are most engaged and are most likely to complete the survey – each of them represents some number of people, unknown but much smaller than 1,561. How much smaller? And what about the monthly players? Here is where I apply a little practical math to get useful estimates. I assume (somewhat arbitrarily) that the monthly players are just halfway between the extremes represented by the highly engaged weekly players,

Table 2

CALIBRATED INTERPRETATION OF WASHINGTON FY15 TRACKING STUDY

Play frequency	Survey Respondents	Each Respondent Represents	Per Capita Annual Spending	Group Annual Spending	Eligible Population %	Players
Weekly	329	194	$342.88	$21,850,205	1.1	63,798
Monthly	357	877	$61.82	$19,362,471	5.6	314,378
LT Monthly	299	1,561	$7.35	$3,429,424	8.3	468,950
Totals:				**$44,642,100**	**15**	**847,126**

and the LT monthly players. I express these assumptions mathematically and then solve for the number of people that each weekly player and each monthly player must represent, in order for their calculated spend to agree with what actually happened. The math is straightforward to do with in Excel. The result for Washington Lotto in FY15 is shown in Table 2.

This result is typical of what I have seen over ten years. Here are two practical implications:

Success of this game is narrowly based, with fewer than 100,000 players participating on a consistent weekly basis. This estimate agrees with the actual business result of ticket count: when jackpots are near the starting level, the number of tickets sold for the popular Saturday draw averages about 120,000. This average could result from all the consistent weekly players, and about 20 percent of the monthly players, buying a ticket for that draw. The calibrated tracking study estimates the size of the monthly player group, which should be a prime focus for marketing efforts to encourage more regular play. The tracking study may help us to understand the attitudes and behaviors of this group in order to better address them.

The overrepresentation of avid players will influence not only estimates of participation, but also accounts of perceptions and attitudes. Tracking Study results that are averaged over "all respondents" may be more positive than would be found in a survey that randomly samples the general population. Perception and attitude measures will be more meaningful if they are broken out by play frequency.

On the whole, I feel that I have increased rather than decreased the perceived value of survey research by exposing these issues. The first thing to value about research is the reliability of the result, not its positive or negative implication. ■

STEPHEN WADE
RESEARCH AND DEVELOPMENT MANAGER, WASHINGTON'S LOTTERY AND PRINCIPAL, LOTTERY MANAGEMENT CONSULTING, LLC

Notes:

1. About Half of Americans Play State Lotteries.
   Gallup Social Issues, June 22, 2016: http://www.gallup.com/poll/193874/half-americans-play-state-lotteries.aspx

2. Powerball Study by Leger available on NASPL Matrix (August 2016)

3. Kentucky TURF Analysis available on NASPL Matrix (August 2016)