

How to avoid a 'winner's curse' for social programs

February 5 2024, by Peter Dizikes



Credit: Pixabay/CC0 Public Domain

Back in the 1980s, researchers tested a job-training program called JOBSTART in 13 U.S. cities. In 12 locations, the program had a minimal benefit. But in San Jose, California, results were good: After a

few years, workers earned about \$6,500 more annually than peers not participating in it.

So, in the 1990s, U.S. Department of Labor researchers implemented the program in another 12 cities. The results were not replicated, however. The initial San Jose numbers remained an outlier.

This scenario could be a consequence of something scholars call the "winner's curse." When programs or policies or ideas get tested, even in rigorous randomized experiments, things that function well one time may perform worse the next time out. (The term "winner's curse" also refers to high winning bids at an auction, a different, but related, matter.)

This winner's curse presents a problem for [public officials](#), private-sector firm leaders, and even scientists: In choosing something that has tested well, they may be buying into decline. What goes up will often come down.

"In cases where people have multiple options, they pick the one they think is best, often based on the results of a randomized trial," says MIT economist Isaiah Andrews. "What you will find is that if you try that program again, it will tend to be disappointing relative to the initial estimate that led people to pick it."

Andrews is co-author of a [study](#) that examines this phenomenon and provides new tools to study it, which could also help people avoid it.

The paper, "Inference on Winners," is published in the February issue of *The Quarterly Journal of Economics*. The authors are Andrews, a professor in the MIT Department of Economics and an expert in econometrics, the statistical methods of the field; Toru Kitagawa, a professor of economics at Brown University; and Adam McCloskey, an

associate professor of economics at the University of Colorado.

Distinguishing differences

The kind of winner's curse addressed in this study dates back a few decades as a social science concept, and also comes up in the natural sciences: As the scholars note in the paper, the winner's curse has been observed in [genome-wide association studies](#), which attempt to link genes to traits.

When seemingly notable findings fail to hold up, there may be varying reasons for it. Sometimes experiments or programs are not all run the same way when people attempt to replicate them. At other times, random variation by itself can create this kind of situation.

"Imagine a world where all these programs are exactly equally effective," Andrews says. "Well, by chance, one of them is going to look better, and you will tend to pick that one. What that means is you overestimated how effective it is, relative to the other options." Analyzing the data well can help distinguish whether the outlier result was due to true differences in effectiveness or to random fluctuation.

To distinguish between these two possibilities, Andrews, Kitagawa, and McCloskey have developed new methods for analyzing results. In particular, they have proposed new estimators—a means of projecting results—which are "median unbiased." That is, they are equally likely to over- and underestimate effectiveness, even in settings with a winner's curse.

The methods also produce confidence intervals that help quantify the uncertainty of these estimates. Additionally, the scholars propose "hybrid" inference approaches, which combine multiple methods of weighing [research data](#), and, as they show, often yield more precise

results than alternative methods.

With these new methods, Andrews, Kitagawa, and McCloskey establish firmer boundaries on the use of data from experiments—including confidence intervals, median unbiased estimates, and more. And to test their method's viability, the scholars applied it to multiple instances of social science research, beginning with the JOBSTART experiment.

Intriguingly, of the different ways experimental results can become outliers, the scholars found that the San Jose result from JOBSTART was probably not just the result of random chance. The results are sufficiently different that there may have been differences in the way the program was administered, or in its setting, compared to the other programs.

The Seattle test

To further test the hybrid inference method, Andrews, Kitagawa, and McCloskey then applied it to another research issue: Programs providing housing vouchers to help people move into neighborhoods where residents have greater economic mobility.

Nationwide economics studies have shown that some areas generate greater economic mobility than others, all things being equal. Spurred by these findings, other researchers collaborated with officials in King County, Washington, to develop a program to help voucher recipients move to higher-opportunity areas. However, predictions for the performance of such programs might be susceptible to a winner's curse, since the level of opportunity in each neighborhood is imperfectly estimated.

Andrews, Kitagawa, and McCloskey thus applied the hybrid inference method to a test of this neighborhood-level data, in 50 "commuting

zones" (essentially, metro areas) across the U.S. The hybrid method again helped them understand how certain the previous estimates were.

Simple estimates in this setting suggested that for children growing up in households at the 25th percentile of annual income in the U.S., housing relocation programs would create a 12.25 percentage-point gain in adult income. The hybrid inference method suggests there would instead be a 10.27 percentage-point gain—lower, but still a substantial impact.

Indeed, as the authors write in the paper, "even this smaller estimate is economically large," and "we conclude that targeting tracts based on estimated opportunity succeeds in selecting higher-opportunity tracts on average." At the same time, the scholars saw that their method does make a difference.

Overall, Andrews says, "the ways we measure uncertainty can actually become themselves unreliable." That problem is compounded, he notes, "when the data tells us very little, but we're wrongly overconfident and think the data is telling us a lot. ... Ideally you would like something that is both reliable and telling us as much as possible."

More information: Isaiah Andrews et al, Inference on Winners, *The Quarterly Journal of Economics* (2023). [DOI: 10.1093/qje/qjad043](https://doi.org/10.1093/qje/qjad043)

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

Citation: How to avoid a 'winner's curse' for social programs (2024, February 5) retrieved 27 April 2024 from <https://phys.org/news/2024-02-winner-curse-social.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.