

The Washington Post

[Back to previous page](#)

Stop guessing whether a bill will work. Instead, let's test it.

By [Dylan Matthews](#), Published: December 7

I'm a policy reporter. My job is to explain to my readers whether smaller class sizes help students learn, whether tax cuts boost economic growth and whether housing programs help families escape poverty.

In a perfect world, what I do would be a kind of science reporting. Just as my colleagues at the health desk often explain which medicines are effective and which are a bust, I'd ideally be able to describe what sociologists, economists and political scientists have discovered about which policies work.

With a few exceptions, however, I can't really do that — at least not with the precision my health colleagues often can. More important, neither can policymakers in Congress and in many regulatory agencies. The Food and Drug Administration has more information available in deciding whether to approve a treatment that a few thousand people will receive than Congress does in considering a bill that will affect every American.

Each year, hundreds of carefully controlled, double-blind studies are done to learn whether a given pill is better than a placebo or whether a new surgery leads to quicker recoveries. Many of these studies are funded by a single agency, the National Institutes of Health. By contrast, in a typical year, no such studies are conducted to evaluate social policy proposals.

That's not because such studies are impossible. In 1962, researchers in a small Michigan school district randomly selected 58 3- and 4-year-olds to enroll in a preschool program, then spent decades comparing them with a control group of 65 kids who didn't go to preschool. Those who enrolled learned more and made more money as adults. In 1976, the Chicago Housing Authority randomly placed public-housing residents in apartments either in the city or in the suburbs, and then tracked the two groups. Those given places in the suburbs did better on every metric from household income to their children's rates of college attendance.

Those studies had a big impact. The Chicago study, for example, is the main research cited in proposals to provide housing vouchers to poor families to break up pockets of "concentrated disadvantage."

But studies like those are very rare. They're expensive, which discourages universities and school districts (such as the one in Michigan conducting the preschool study) from doing them. And often, as in the Chicago case, they come about only because a court orders them.



The advertisement features a yellow background. On the left, two iPhone 5s are shown: one in black and one in silver. The text on the right reads: "Sprint" with the logo, "I am unlimited.", "Sprint is the only national carrier that offers Truly UnlimitedSM data for your iPhone 5.", and a "LEARN MORE" button. At the bottom right, it says "Restrictions apply." The Apple logo and "iPhone 5" are at the bottom left.

Because of this rarity, it's easy to pick nits. The preschool study involved only children with low IQs, critics noted. Maybe the results would have been different with children of average or above-average intelligence. The Chicago housing study was conducted during a time when crime was much worse than today. Maybe with safer inner cities, you wouldn't see similar gains from sending families to the suburbs.

Researchers have spent gallons of ink arguing over such caveats, trying to figure out what can and can't be inferred from the meager pool of good data with which they're forced to work. At no point does the straightforward solution present itself: run another study.

"Rigorous ways to evaluate whether programs are working exist," then-White House budget director Peter Orszag said in 2009. "But too often such evaluations don't happen. . . . This has to change."

As Orszag says, it's not that researchers don't know how to evaluate programs or that they don't want to. Indeed, researchers who focus on international development have been doing so in recent years, with very promising results. Economists such as MIT's Esther Duflo and Abhijit Banerjee and Yale's Dean Karlan, along with their research groups, the Jameel Poverty Action Lab (JPAL) and Innovations for Poverty Action (IPA), have run dozens of randomized experiments in developing countries to see which forms of aid work and which are worthless or counterproductive. The goal, as the JPAL puts it, is to "reduce poverty by ensuring that policy is based on scientific evidence, and research is translated into action."

With funding from individuals and foundations, but for the most part not governments, they have learned, for instance, that spreading information about the benefits of education keeps students in class; remedial tutoring doesn't. Giving away bed nets reduces malaria infections; charging even a small amount for them is much less effective.

The confidence with which development researchers can make these judgments is in stark contrast to the Talmudic reading of a handful of studies that characterizes debate about social policy in the United States. And that confidence means policymakers pay attention. Aid organizations have heaped praise on Duflo and company, with USAID chief Rajiv Shah declaring that "the whole movement that Esther and her colleagues at MIT and around the world have really spearheaded is so important in rethinking how we make aid work." The World Bank has teamed up with the JPAL to design better poverty-reduction programs.

This shouldn't be surprising. Information matters, especially when it comes from reliable, trusted sources such as Duflo's team. That's why lawmakers nervously await reports from the Government Accountability Office (GAO), the Congressional Research Service (CRS) and the Congressional Budget Office (CBO). They know that a bad CBO score can kill a bill and a good one can push it over the edge toward passage. A credible report suggesting that a policy would be effective, or would be cheaper than the alternatives, or wouldn't raise taxes on the middle class, is more persuasive than any floor speech.

But the GAO, the CRS and the CBO only do so much. Lawmakers have ways of knowing how much a bill costs (CBO), whether it's constitutional (CRS) and how it fares once enacted (GAO), but there is no agency that tests these proposals on a small scale ahead of time to see whether they will work. In other words, there's no JPAL or IPA for education, health care, housing or any number of other policy areas that Congress works on daily.

That's why policymakers have to turn to groups outside government to advise them on what the likely effects of legislation would be. The Center on Budget and Policy Priorities and its leader, Robert Greenstein, have become such respected sources on anti-poverty programs that when it came time to design the stimulus package, the Obama administration turned to Greenstein.

More nefariously, regulations of all kinds of products, from chemical additives to financial instruments, are

influenced strongly by industry lobbyists — not only because of their campaign donations but because they offer policy expertise that's hard to find elsewhere.

It doesn't have to be this way. Congress should establish a policy-evaluation office, modeled after the JPAL or IPA, to run randomized, controlled trials on social policies. The office should have broad authority to do test-runs of proposals of its choosing, operating under the same rules of informed consent used in medical studies.

Members of Congress should have the power to request studies, particularly when they bear on current debates and can be done relatively quickly. For example, the Obama administration is likely to push for immigration reform next year. In concert with the state involved, an evaluation office could randomly select one town and grant its illegal immigrants permanent residency, and randomly select another town and leave its undocumented residents in a legal gray zone. Within a few months or a year, researchers should be able to see whether bringing the immigrants out of the shadows hurts native-born workers' wages, reduces employer abuse or has any number of other consequences. That's a reasonable time for Congress to wait before adopting huge changes to the immigration code.

Other questions wouldn't be answerable that quickly. By definition, a study that seeks to find out whether access to preschool increases children's earnings as adults would take decades to complete. The office should engage in such projects — known as longitudinal studies — even without congressional prodding, providing updates along the way. The questions at stake are too important, and the amount of effort needed too great, to leave to the whims of Congress.

Ideally, the office would integrate itself into legislators' workflow. Rather than engaging in speculative debates about what a bill would do, members of Congress would request an evaluation upon introducing legislation, and within a few months or a year, they would have study results that would either seal support for a good idea or kill a bad one. With luck, only legislation that made it through the office's grinder would get to final passage. Being backed by experts at industry groups would no longer be enough.

Most people would never dream of taking a pill that lacks FDA approval — that hasn't made it through a randomized trial with a solid record. We shouldn't have to settle for less from federal policy. With a congressional policy-evaluation office, we wouldn't have to.

matthewsd@washpost.com

Dylan Matthews is a reporter for The Washington Post's [Wonkblog](#).

Read more from [Outlook](#), friend us on [Facebook](#), and follow us on [Twitter](#).

Sponsored Links

wheels when you want em.

Gas & insurance included. Join Zipcar now for \$50 free driving.
zipcar.com

LifeLock® Fraud Services

LifeLock® Credit Alerts For Fraud Protect Against Identity Fraud.
LifeLock.com

Track Your Investments

Organize your entire financial life with free tools from Personal Capital.
www.personalcapital.com

[Buy a link here](#)