

Policy-Based Evidence Making

Oren Cass

“EVIDENCE-BASED POLICYMAKING” is the latest trend in expert government. The appeal is obvious: Who, after all, could be against evidence?

Most EBP initiatives seem eminently sensible, testing a plausible policy under conditions that should provide meaningful information about its effectiveness. So it is not surprising to see bipartisan support for the general idea. Speaker of the House Paul Ryan and Senator Patty Murray even collaborated on the creation of an Evidence-Based Policymaking Commission that has won praise from both the Urban Institute and the Heritage Foundation.

But the perils of such an approach to lawmaking become clear in practice. Consider, for instance, the “universal basic income” campaign. Faced with the challenge of demonstrating that society will improve if government guarantees to every citizen a livable monthly stipend, basic-income proponents suggest an experiment: Give a group of people free money, give another group no money, and see what happens. Such experiments are underway from the Bay Area to Finland to Kenya to India.

No doubt many well-credentialed social scientists will be doing complex regression analysis for years, but in this case we can safely skip to the last page: People like free money better than no free money. Unfortunately, this inevitable result says next to nothing about whether the basic income is a good public policy.

The flaws most starkly apparent in the basic-income context pervade EBP generally, and its signature method of “controlled” experiments in particular. The standard critique of overreliance on pilot programs, which are difficult to replicate or scale, is relevant but only scratches the surface. Conceptually, the EBP approach typically compares an expensive

OREN CASS is a senior fellow at the Manhattan Institute.

new program to nothing, instead of to alternative uses of resources—in effect assuming that new resources are costless. It emphasizes immediate effects on program participants as the only relevant outcome, ignoring systemic and cultural effects as well as unintended consequences of government interventions. It places a premium on centralization at the expense of individual choice or local problem-solving.

Politics compounds the methodological shortcomings, imposing a peculiar asymmetry in which positive findings are lauded as an endorsement of government intervention while negative findings are dismissed as irrelevant—or as a basis for more aggressive intervention. Policies that *reduce* government, when considered at all, receive condemnation if they are anything other than totally painless. Throughout, the presence of evidence itself becomes an argument for empowering bureaucrats, as if the primary explanation for prior government failure was a lack of good information.

The common thread in these shortcomings is an implicit endorsement of the progressive view of the federal government as preferred problem-solver and a disregard for the entire range of concerns that prevent conservatives from sharing that view. Like Charlie Brown with his football, conservatives repeatedly lunge with enthusiasm at the idea that evidence will hold government accountable for results, only to be disappointed. Lauded as a tool of technocratic excellence, EBP more often offers a recipe for creeping statism.

Even in this era of populist uprisings and collapsing faith in institutions, confidence has never been higher that the challenges of government can be “solved” with the right studies. This confidence is misplaced; to the contrary, there is no evidence that our investments in social science and EBP have improved our capacity to govern wisely and effectively at all. For evidence to play a valuable role in the policymaking process, social scientists designing their research and policymakers using the results need a better understanding of what evidence provides—and what it cannot.

EVIDENCE PRODUCTION

At the zenith of recent public-policy research sits the Oregon Health Insurance Experiment, a randomized controlled trial that sought to establish the benefits of Medicaid coverage.

In 2008, the state of Oregon found itself with the resources to expand its Medicaid program to reach some, but not all, of the low-income

adults without children in the state who had not previously been eligible for coverage. The state decided to fill the available slots by lottery, randomly selecting 40% of the 75,000 eligible adults for Medicaid enrollment. This allowed researchers to conduct a randomized controlled trial, the gold standard for quantifying a policy's effects, in which the health outcomes of otherwise similarly situated recipients and non-recipients could be compared.

Medicaid costs the United States more than \$500 billion annually, accounting for the majority of all safety-net spending nationwide. Its rapid historical growth has placed enormous pressure on state budgets, and its rapid forecasted growth is one of the major drivers of the nation's long-term fiscal imbalance. The debate over how best to ensure that low-income Americans have access to health care in the most cost-effective way possible is one of the most controversial and complex policy quandaries in our politics. Yet the researchers providing the evidence on which to base policy were investigating whether the value of Medicaid is larger than *zero*.

Proponents of Medicaid expansion understandably delighted in this framing, which established a bar of "not worthless" for the program. When initial results showed that recipients felt better about their health, used more health-care services, and experienced less financial stress, the *New York Times* gloated: "[C]onservative critics of Medicaid — and of health care reform's requirement to expand it — have made the outlandish claim that it provides such poor care that enrollees would be better off having no coverage. . . . Those claims have now been refuted."

But no one should be surprised to learn that when Oregon offered to pay for the health care of some residents — a benefit worth more than \$5,000 per person in that state, among the nation's highest per-capita levels — those residents did in fact consume more health care, keep more of their own money, and feel better about their health than did people who received nothing. If that were the basis on which to embrace public policies, the government would quickly find itself giving piles of money to every resident for every imaginable purpose.

When framed this way, the study's guiding question seems absurd. But this experimental design is not an outlier. It is very nearly the archetype. In his own discussion of EBP, Harvard Kennedy School professor Ricardo Hausmann uses a parallel, smaller-scale hypothetical: testing the value of tablet computers in classrooms by randomly assigning half

of a sample of 300 schools to receive them, and then comparing their test-score progress with the half that didn't. Whether free resources offer any immediate value to direct recipients has somehow become a useful litmus test for increased government spending.

Tellingly, while Hausmann is interested in describing the “problem with evidence-based policies,” holding new programs to an adequate standard seems far from his mind. To the contrary, his concern is with moving faster toward higher spending because it is obvious to him that all classrooms should have tablets. “[T]he real question we wanted to answer,” in his view, “was how tablets should be used to maximize learning,” but traditional experiments “do not permit testing of more than two or three designs at a time.” Thus, he suggests giving all classrooms tablets to “free all teachers to experiment with different software, different strategies, and different ways of using the new tool. . . . Over time, we will observe some teachers who have stumbled onto highly effective strategies.” Certainly, this will expand school technology budgets; whether it cost-effectively improves learning is apparently beside the point.

Concluding with a vision of “creating organizations that learn how to learn” and an analogy to “lean manufacturing,” Hausmann accidentally highlights his own confusion: He links to a source defining “lean manufacturing” as “creating more value for customers with fewer resources,” by “creat[ing] processes that need less human effort, less space, less capital, and less time to make products and services at far less costs.” Indeed, few private-sector organizations could survive by distributing ever-greater resources in hopes that they will be usefully deployed. The lean mindset seems useful, but also like the opposite of Hausmann's and of EBP generally.

Another way of understanding the problem with EBP is to disaggregate two policy components that EBP invariably treats as a unit: the redistribution of resources, and the allocation of those resources to a particular use. Consider an alternative experimental design in which participants themselves are required to pay for their program participation. “Do families benefit when forced to purchase tablets for their children to use at school?” is a useful policy question; likewise, “do low-income households benefit when forced to allocate \$5,000 per person toward Medicaid coverage?” If the result were positive, it would seem to offer strong evidence for such paternalism.

Of course, low-income households in particular often cannot afford such expenditures, even if they would be worthwhile. Nor are most EBP practitioners interested in dictating how people spend their own money; they are interested in launching new programs. But just as in the case of individual households, government resources are not infinite. If pilot programs and experiments merely assume the availability of resources, a condition that would not hold for wide-scale implementation, then the resulting evidence is inapplicable. This evidence may also attribute to the program benefits that are merely the result of the redistribution itself: Are Medicaid recipients benefiting from the health care they receive, or from the redistribution of resources — which Medicaid may in fact dilute?

Unfortunately, EBP is inherently ill-equipped to evaluate the wisdom of redistribution because such policy decisions do not depend upon their immediate effects on direct recipients. The effects of redistribution ripple outward through society and across generations. Determining when the state should take property from one group and give it to another is a philosophical and cultural question as much as an economic one, so it is impossible to judge scientifically. Predicting and evaluating the long-term response by those in the receiving group is likewise beyond the power of any experiment. If the issue were primarily economic, one might well anticipate that resources have higher marginal value to poorer households and thus expect redistribution to continue improving society until all were made equal. Surely, evidence showing that a small group of recipients benefits over the experimental period from such transfers would not offer meaningful support for full-fledged socialism.

A ridiculous example, perhaps, yet how far is it from what the basic-income researchers are attempting? A basic income is, by definition, the redistributive step disaggregated from the allocative step; money is redistributed, but government does not choose how to spend it. No matter how exciting and positive the result of giving randomly selected individuals free money for a limited time, the experiments in no way simulate the changes to society that might ensue if the core obligation to provide for oneself were shifted to the government. In turn, the behaviors of the studied recipients offer little information about how recipients of a fully implemented basic income might live.

Beyond redistribution and allocation, broad implementation of policies can have a range of cultural and systemic consequences that

override the initial benefits that an experiment might identify. Evidence that moving low-income households into higher-income communities can promote upward mobility, for instance, says little about the widespread implications for either the departed or receiving communities if such movement were made a centerpiece of anti-poverty policy. College-tuition subsidies might be highly beneficial for a limited set of recipients, whereas if they became widely available then tuition might quickly rise to absorb them. Such concerns are not hypothetical: When Denmark began offering generous maternity leave, so many nurses made use of it that mortality rates in nursing homes skyrocketed.

In essence, EBP free-rides on a broader society not yet subject to the policy. Pilot programs have their cake and eat it too, claiming the narrow benefits of government intervention without triggering the broader costs that so often frustrate a policy's aims. Policy advocates will not only ignore these challenges but actively discourage their consideration and undermine anyone who raises them. Then, when the evidence-based policy fails to deliver, it becomes merely yet another cautionary tale of "unintended consequences."

SELECTIVE READING

Whatever residual benefits of EBP survive its methodological flaws face quick erasure upon contact with the political process. Not only are EBP studies built to produce positive results, but those negative outcomes that slip through are explained away and rejected. The most obvious illustration of this point is, once again, the Oregon Medicaid study. The positive, self-reported benefits identified in its first year were met with great interest and enthusiasm. But attitudes shifted sharply in year two, when a broader range of outcomes was revealed.

As the researchers reported in the *New England Journal of Medicine*, "This randomized, controlled study showed that Medicaid coverage generated no significant improvements in measured physical health outcomes in the first 2 years." This was an extraordinarily poor result that stunned most experts (a reaction that underscores the lopsided nature of the experiment and its seeming emphasis on validating rather than scrutinizing the program).

But it is not in the nature of EBP to accept such findings. Hausmann acknowledges as much in his tablet example:

Suppose there are no significant differences, as has been the case with four [randomized controlled trials] that found that distributing books also had no effect. It would be wrong to assume that you learned that tablets (or books) do not improve learning. What you have shown is that that particular tablet, with that particular software, used in that particular pedagogical strategy, and teaching those particular concepts did not make a difference.

No doubt the reliability of the test scores used to measure progress would come under such fire as well.

The Medicaid study's results were condemned on similar grounds: The sample size was too small, the time period too short, the indicators not the right ones. What did not happen was any reconsideration of the value of Medicaid. The *New York Times* editorial board chose not to revisit the question, but the *Times's Upshot* health-care correspondent, Margot Sanger-Katz, offered this Orwellian summary: "There were some important questions the study never answered: It didn't find any evidence that patients' physical health improved over the two years." Apparently a question can be deemed "answered" only if the answer is the right one.

This evidence ratchet, in which findings can promote but not undermine a policy, is common. National Public Radio recently reported on a pilot program testing the effectiveness of police body cameras in Washington, D.C. The results are not yet in, but, as the story concluded, "[N]o matter what this study ultimately shows about the cameras, one thing seems already clear, [the police officer] says: 'They're here to stay.'"

The poor results of the Oregon study were equally unlikely to change the conversation. The value of Medicaid has become especially salient, as efforts to repeal the Affordable Care Act intensified in the spring of 2017. The ACA achieved its insurance-coverage increases largely through Medicaid expansions. If that expansion was unlikely to improve health outcomes, the harm of repeal would be low. But the Oregon "evidence" has faded quickly to the background.

Instead, ACA proponents have warned that its repeal would kill countless Americans annually: 36,000, according to Senator Bernie Sanders; 44,000, according to CUNY professors David Himmelstein and Steffie Woolhandler (who, interestingly, cited "scientific consensus" on the issue). Back at the *New York Times*, Paul Krugman asserted without

reference that “there is no real question that Obamacare is saving tens of thousands of lives every year.”

A genuine interest in evidence should also have led enterprising policymakers and journalists to investigate how the ACA was in fact performing. By 2017, multiple years of nationwide data were available. But the story those data told repeated the Oregon findings: Since Medicaid expansions began in 2014, the U.S. mortality rate had halted its decades-long decline and managed to increase.

Further, because each state had the choice whether to expand its Medicaid population, ACA implementation provided a useful natural experiment. In fact, mortality rose *faster* in the states that pursued Medicaid expansion. Was two years too soon to expect results? Critics had leveled that charge at the Oregon study, but this was exactly the time period over which the expansion of private insurance coverage in Massachusetts had produced substantial mortality gains in 2008-09. *That* result was widely cited as proof that expanding insurance coverage works.

In any case, time doesn't gain undesirable evidence any more traction. Consider the welfare reform enacted at the federal level in 1996, among the most dramatic policy changes of recent decades and one of the few that *cut* a major federal program. If the EBP mindset had any prospect of meaningfully guiding major policy decisions, the reform's 20th anniversary in 2016 should have been an opportunity to dispassionately assess the results. On the other hand, if policymakers cannot use evidence to reach agreement 20 years *after* the fact about whether a policy was advisable, it will never play a credible role at the outset of debate.

Predictably, the anniversary was marked not by consensus sociological assessment but rather by a contentious re-litigation of the reform along precisely the same ideological divide that had separated those who favored or opposed reform in the first place. The sides largely agreed on the *data*, but not at all about what it was evidence of. The right pointed to increases in workforce participation and decreases in government dependence as signs of success. The left emphasized the limited magnitude of these changes and lamented the dire straits of those no longer receiving cash assistance.

This last concern is particularly strange and casts new light on the issue, examined above, of what metric EBP sets for success. A new program faces the low bar of achieving more than nothing. For the elimination of an existing program, demanding that nothing change is

an outrageously high standard. *Of course* those who relied on welfare's cash benefits saw a decline in cash benefits when welfare was cut. That is little more than a restatement of the policy. If a cut can show "evidence" of success only if all former beneficiaries are made whole, evidence will be in short supply.

Surveying 13 different studies on the effects of welfare reform, the Center on Budget and Policy Priorities found that work requirements increased employment and reduced poverty. The effects were not enormous and only some were statistically significant. But whereas one test found a statistically significant drop in poverty, 11 found smaller declines, and one found no change, *none* found that poverty increased. This would seem a major victory for work requirements. Certainly, if one reverses the default and reads the studies as a test of *no* work requirements, the result would be striking: loosening welfare not only fails to reduce poverty but seems in many cases to increase it. But that is not how CBPP saw things; their headline read, "Work Requirements Don't Cut Poverty, Evidence Shows."

One lesson from this experience could be the importance of setting clear goals and metrics in advance. Had welfare-reform supporters established in 1996 that their goals were particular changes in workforce participation and a higher ratio of earned income to transfer payments, perhaps they would be on firmer ground claiming success today. But even that level of rigor is unlikely to escape the black hole of partisan politics. The argument merely shifts to a fight over baselines and outside forces: It was the late-'90s economic boom, not welfare reform, which produced positive changes. Or so the argument goes.

The same game is played when the left evaluates its own policies, though in that case, the outside forces take the blame instead of the credit. Early in 2009, promoting its stimulus package, President-elect Obama's team released a chart showing the forecasted trajectory of unemployment in subsequent years with or without the proposal. Krugman applauded the analysis and declared, "[I]t will be a joy to argue policy with an administration that provides comprehensible, honest reports, not case studies in how to lie with statistics." Ultimately, unemployment *with* the stimulus climbed far higher than Obama's economists said it would reach *without* the stimulus. Some might consider this evidence of policy failure.

Not Krugman, who reported that "it is clear to all fair-minded economists that the stimulus did work." Jared Bernstein, co-author of the

initial report, explained, “At the time, we simply didn’t have good information. . . . So, the statistics were way off at the time. . . . I mean, it wasn’t just that we were wrong. Everyone was wrong.”

Krugman and Bernstein’s claims may in fact be correct—certainly, the effect of the American Recovery and Reinvestment Act is beyond the scope of the current discussion. But “evidence” seems unlikely to provide a satisfying conclusion.

The right is by no means immune to evidence’s seductive charm. In his “A Better Way” anti-poverty agenda in 2016, House speaker Paul Ryan jettisoned the more aggressive, substantive reforms that he had championed as chair of the House Budget Committee, replacing them with tweaks to existing safety-net programs and then a wide-ranging endorsement of EBP. Bullet number eight of his 41-point agenda captured the essence: “Pay More for the Good Stuff, Less for Everything Else.” At least four proposals sought to reduce “duplication” and “redundancy” while increasing “coordination,” “transparency,” “flexibility,” and “accountability.” Nine others emphasized “evidence,” “metrics,” “evaluat[ion],” “data,” “technology,” “research,” and “standards.”

But is the problem with the federal government that it has never thought of these things, or that it cannot implement them effectively, interpret them intelligently, and then adjust accordingly? If legislators will not act on the evidence and bureaucrats will not adhere to it, then centering a policy agenda on its creation is a counterproductive dodge, not a responsible approach to reform. The result is likely to be an expansion and empowerment of government, but EBP cannot be trusted to ensure that the expansion will deliver positive results.

EVIDENCE AND DATA

While both the production and consumption of policy evidence will always warrant a healthy skepticism, certainly it would be desirable for policymakers to have access to useful evidence alongside an understanding of the limits of its usefulness. Some changes to how we approach EBP could help. Most important, studies seeking to validate particular policy interventions should acknowledge resource constraints, which can be done in several ways.

To start, in designing experiments, new interventions can be substituted for existing programs instead of being layered on top of them; such a substitution-based model of control would better reflect real-life

constraints. For example, instead of merely giving tablets to a set of classrooms, give them tablets but cancel art class. Perhaps this sounds cruel, but it reflects the budgetary tradeoffs that school districts will face if they choose to buy tablets. Or, viewed from another perspective, many districts have already been forced to cut enrichment programs; if new funding were to become available for tablets, should the superintendent campaign to reinstate art instead?

This kind of thinking need not be limited to experimental design. Policymakers can also seek this type of evidence from natural experiments provided by the broad range of resource allocations pursued across states and localities. A recent study by the Yale School of Public Health provides a powerful illustration: Rather than giving a group of people new health-care benefits, researchers compared health outcomes across states that allocated relatively more or less of their social-services spending toward health-care programs. They found that “states with a higher ratio of social to health spending (calculated as the sum of social service spending and public health spending divided by the sum of Medicare spending and Medicaid spending) had significantly better subsequent health outcomes for the following seven measures: adult obesity; asthma; mentally unhealthy days; days with activity limitations; and mortality rates for lung cancer, acute myocardial infarction, and type 2 diabetes.”

A critic might object that this study lacks the “gold-standard” design of randomization and control. Perhaps states spend less *because* their populations are healthier, or perhaps states with better hiking trails have both healthier populations *and* a preference for other forms of social spending. Reasonable concerns all; certainly, this study does not prove conclusively that states should shift significant funding away from health care to improve health. But it offers a place to start, and each objection points toward additional analyses that could be done from historical data and new experiments that might be constructed. At a minimum, it offers policymakers at least as much guidance as the Oregon experiment (whose positive outcome would have been unremarkable and whose negative outcome was dismissed) possibly could.

Another refinement of EBP would create more useful control groups. Rather than taking something away from program participants, a control group could be given something of potentially comparable value. For instance, in Oregon, researchers might have given subsets of the control group housing vouchers, a wage subsidy, or a lump-sum cash

payment. To justify expansion, Medicaid should show that it performs *better* than these other approaches.

Notably, the Oregon researchers attempted after the fact to impute the monetary value that beneficiaries received from enrollment, concluding that each dollar spent had delivered 20 to 40 cents of value. The magnitude of value destruction is remarkable, even for government. Across a broader range of programs, such effects—and opportunities to improve welfare while reducing cost by reallocating resources—will present themselves only if researchers seek them out.

A related form of empirical research would investigate at what income levels households allocate resources toward what purposes, or in which programs households enroll if forced to choose. Even simple surveys of household priorities, targeted at the prospective recipients of support, would be valuable. Not only would these types of data inform policymakers about what recipients want, but, perhaps more important, the heterogeneity of preferences would remind them that no amount of evidence can find a single program that is “best” for everyone. And divergence between reported or expressed preferences and experimental evidence would force them to grapple with the limits of their knowledge: Who are they going to believe, regression coefficients or their lying constituents?

Lastly, beyond better testing of new programs, EBP researchers should give equal focus to cuts. The substitution-based model of control, described above, is a start. For every program deemed more effective, one is deemed less effective as well. But a greater emphasis should go toward validating smaller government directly. Arbitrarily cutting services might be unethical, but governments in general and individual programs face countless situations in which they have no choice but to make cuts. These situations offer opportunities to develop evidence by assigning the cuts to a random subset, just as Oregon’s circumstances invited its random Medicaid expansion. Assigning such cuts randomly, instead of spreading them evenly, is no less fair than Oregon’s decision to focus all of its new Medicaid resources on a chosen few.

Better evidence will help only to the extent that policymakers—and politicians in particular—are willing to use it responsibly. In some respects this is impossible, and politicization is an exogenous reality that inherently limits what evidence can contribute. But a common framework and vocabulary would at least help policymakers stop talking past one another.

The most important distinction to be drawn is between strategy and tactics. Strategic questions about government's role in society and its approach to solving public problems cannot be resolved—or even meaningfully informed—by the experimental evidence that the social sciences can produce because that evidence does not speak to the social consequences of government intervention. If computer models could accurately simulate entire human societies and test various approaches to government across generations, the results would be powerful. But that is not possible. Whether small groups enrolled in pilots display certain behaviors provides little meaningful guidance, and reliance on such data as representative is more likely to distort than enhance decision-making.

Conversely, where policymakers have already decided that government should act and on the form that action should take, good evidence can inform tactical choices about program design. If all indigent defendants will receive free representation, should the state provide this via a public defender, a court-appointed attorney, or a voucher that allows the accused to shop around? Where broader cultural and economic effects are held constant, EBP's tools do help to identify the more effective programs. But even here one must be careful: The results might show which model costs least, produces highest defendant satisfaction, or yields the lowest sentences. They will not identify the “best” or “most just” approach.

Thus, policymakers would also do well to distinguish between “data,” which *informs*, and “evidence,” which *proves*. Better data helps to define problems and formulate solutions by establishing a shared and accurate view of past experience and the status quo. Policymakers need to know how many people are unemployed, how that number is changing, and how it compares to historical levels. They need to know who lacks health insurance and why, where and when atmospheric concentrations of particulate matter are highest, and what curriculum each school district uses. But they should not expect these facts to answer the question of which policies to pursue.

Most of the evidence promoted by EBP would be better understood as data. Perhaps when a limited number of subjects receive basic-income payments for several years, their average working hours increase. That would be an interesting addition to the policy fact base. But it is not evidence that a basic income would increase work hours. Perhaps states

that have pursued recent Medicaid expansions have seen mortality rates rise most quickly. That does not prove that the ACA is killing people.

A STRANGE TIME FOR CONFIDENCE

Policymakers' faith in evidence is ascending just as shortcomings are becoming apparent in the evidence-based model of medicine on which EBP is based, and as skepticism surges about the unreplicable findings of the social science that EBP most resembles.

Embracing new drugs and procedures on the basis of statistical significance turns out to be an often costly and inefficient approach to health care. Data alone cannot dictate the ideal mammogram frequency, for instance. There are tradeoffs. Approving or rejecting a drug on the basis of its aggregate effect on a large, heterogeneous population may not lead to the best treatment regimes. A drug failing to show statistically significant results on that basis may in fact be highly effective for identifiable subpopulations, while a seemingly successful one may be wasteful or counterproductive for others. Flexibility, creativity, and judgment — to which federal control has proven an active obstacle — are critical.

Social scientists, meanwhile, are discovering that they cannot replicate many of their own findings. The incentives of the profession weigh heavily toward designing experiments that will produce a positive result, analyzing and reanalyzing data until one appears, and then interpreting it as broadly as possible. That kind of research produces headlines, but not knowledge — a fact that social scientists are beginning to acknowledge. With their enthusiasm for EBP, policymakers are running headlong into a collapsing structure, pushing past inhabitants fleeing in the other direction.

If EBP is unreliable, and especially if its failures point consistently and uncritically toward more government, then embracing it represents an abdication of responsibility and signals a loss of will and courage to defend ideas. Deferring to evidence becomes a crutch, transforming public statesmen into the satirical corporate lackeys who declare at every meeting, "I'm for what works."

In *The Closing of the American Mind*, Allan Bloom observed:

When the liberal, or what came to be called the utilitarian, teaching became dominant, as is the case with most victorious causes, good arguments became less necessary; and the original good

arguments, which were difficult, were replaced by plausible simplifications—or by nothing. The history of liberal thought since Locke and Smith has been one of almost unbroken decline in philosophic substance.

EBP represents a logical step in this progression.

In the case of a universal basic income, assessment should begin from a philosophical inquiry into the proper role of the state and its relationship to the development of healthy families and communities. It should incorporate economic analysis of labor-market incentives and, yes, data from any relevant studies. It should draw on historical examples or, lacking those, on analogs in private behavior: Do most parents promise their own children guaranteed incomes for life? Our culture's view of "trust-fund kids" would suggest that generations of accumulated wisdom may take a dim view of providing an automatic and unearned income for all.

Grappling with such questions is frustrating and unsettling, as the policymaking process should be. It encourages humility and demands that the case for government action clear a high bar. EBP does just the opposite, giving policymakers a false sense of confidence—like a stiff drink that at once fortifies the courage and weakens the judgment. One should avoid making serious decisions in that state.