Coalition for Evidence-Based Policy

A NONPROFIT, NONPARTISAN ORGANIZATION

Demonstrating How Low-Cost Randomized Controlled Trials Can Drive Effective Social Spending:

Project Overview and Request for Proposals 2015

BACKGROUND AND PURPOSE

In response to the White House and Office of Management and Budget (OMB) call to action for evidence-based reforms across the federal government, the Coalition for Evidence-Based Policy has launched a competition to select and fund low-cost randomized controlled trials (RCTs) that seek to build actionable evidence about "what works" in U.S. social spending. This is designed as a high-visibility, three-year initiative, whose purpose is to demonstrate the feasibility and value of low-cost RCTs to a wide policy and philanthropic audience. This document contains the Request for Proposals (RFP) for the competition's second year (2015), in which we will select and fund 3-4 low-cost RCTs that meet the criteria for policy importance and other factors set out in the RFP.

The awardees in the competition's first year (2014) illustrate the type of studies we seek to fund:

- A large, multi-site RCT of Bottom Line, a program that provides one-on-one guidance to help low-income, first-generation students get into and graduate from college. This study is measuring college enrollment, persistence, and completion outcomes for a sample of nearly 1,400 students over a seven-year period, using administrative data from the National Student Clearinghouse. The total study cost is $159,000, of which we awarded $100,000.

- A large RCT of Durham Connects, a postnatal nurse home visiting program designed to improve child and mother health and well-being. The study is using hospital administrative records to measure program impacts on families' emergency department use and related healthcare costs through child age 24-months, for a sample of about 1,100 families in Durham County, North Carolina. The total study cost is $183,000, of which we awarded $96,000.

- A large, multi-site RCT of workplace health and safety inspections carried out by the federal Occupational Safety and Health Administration (OSHA). For a sample of about 29,000 business establishments eligible for a randomized inspection, the study is testing whether being randomly chosen for inspection affects establishments' subsequent injury rates and business outcomes (e.g., sales, business closures) over a multi-year period – all measured through administrative data from OSHA and other sources. The total study cost is $153,000, of which we awarded $96,000.

These and other examples were discussed at a July 2014 conference on low-cost RCTs, co-sponsored by the White House Office of Science and Technology Policy (OSTP) and our organization (see our overview, and White House OSTP summary). The conference is one of several recent Executive Branch efforts to advance low-cost, rigorous evaluations, described in the *2014 Economic Report of the President* (chapter 7) and the FY 2015 budget (chapter on performance and management, pp. 65-70). In an important related development, the National Institutes of Health recently launched an initiative to fund low-cost RCTs.

The Coalition is a nonprofit, nonpartisan organization that is unaffiliated with any social programs. This initiative is funded through philanthropic grants to the Coalition from the Laura and John Arnold Foundation, Annie E. Casey Foundation, and Overdeck Family Foundation.

This packet includes:

- **A brief concept paper on the initiative (three pages)** – *The Breakthrough: Low-cost RCTs are a recent innovation in policy research that can rapidly build the body of evidence about "what works" to address major social problems.*
- **The Request for Proposals, inviting grant applications for the second of three annual competitions (three pages).**

THE BREAKTHROUGH:
Low-cost RCTs are a recent innovation in policy research that can rapidly build the body of evidence about "what works" to address major social problems

I. Background: Well-conducted RCTs are regarded as the strongest method of evaluating the effectiveness of programs, practices, and treatments ("interventions"), per evidence standards articulated by the Institute of Education Sciences (IES) and National Science Foundation (NSF),[1] National Academy of Sciences,[2] Congressional Budget Office,[3] U.S. Preventive Services Task Force,[4] Food and Drug Administration,[5] and other respected scientific bodies.

Uniquely among study methods, random assignment of a sizable number of individuals[6] to either a treatment group (which receives a new intervention) or a control group (which receives services-as-usual) ensures, to a high degree of confidence, that there are no systematic differences between the two groups in either *observable* characteristics (e.g., income, ethnicity) or *unobservable* characteristics (e.g., motivation, psychological resilience, family support). Thus, any difference in outcomes between the two groups can be confidently attributed to the intervention and not to other factors. For this reason, recent IES and NSF research guidelines recommend that "generally and when feasible, [studies that measure program effectiveness] should use designs in which the treatment and comparison groups are randomly assigned."[1]

II. Breakthrough: Researchers have shown it is possible, in many instances, to conduct sizable RCTs at low cost, addressing a major obstacle to their widespread use, and building valuable evidence.

A. The low cost is achieved by –

1. **Embedding random assignment in initiatives that are being implemented anyway as part of usual program operations.** Government and foundations fund a vast array of strategies and approaches and, over time, new initiatives and reforms are often launched. Credible evaluations can be embedded in many of these efforts – for example, by (i) using a lottery process – i.e., random assignment – to determine who will be offered program services (since programs often do not have sufficient funds to serve everyone who is eligible); or (ii) randomly assigning some individuals to the program's usual approach (e.g., transitional jobs for ex-offenders) versus a revised model that is being piloted (e.g., transitional jobs plus drug treatment), to see if the new model produces better outcomes.

- and -

2. **Using administrative data that are collected already for other purposes to measure the key outcomes**, rather than engaging in original – and often costly – data collection (e.g., researcher-administered interviews, observations, or tests). In many jurisdictions, administrative data of reasonable quality are available to measure outcomes such as child maltreatment rates, employment and earnings, student test scores, criminal arrests, receipt of government assistance, and health care expenditures.

B. Such leveraging of ongoing efforts/resources enables many more RCTs to go forward, by reducing their cost as much as tenfold. Specifically, this approach reduces or eliminates what are typically the most costly and complex components of an RCT: collecting original outcome data from each sample member; delivering the intervention that is to be evaluated; and recruiting a sample of individuals or other units (such as schools) to participate in the study.

C. Low-cost RCTs thus offer a powerful new vehicle for evidence-building, and an important complement to traditional, more comprehensive RCTs as part of a larger research agenda. For example, low-cost RCTs can be a highly cost-effective tool for identifying interventions that show impacts and are therefore strong candidates for traditional RCTs. Traditional RCTs can then be used to generate valuable additional evidence about whether, under what conditions, and how to scale up the intervention so as to achieve optimal impact.[7]

III. Examples: The following are five sizable, well-conducted RCTs, in diverse program areas, that cost between $50,000 and $300,000 – a fraction of the usual multimillion-dollar cost of such studies. These studies all produced valid evidence of practical importance for policy decisions and, in some cases, identified program strategies that produce budget savings. (More details and citations for these studies are posted here.)

A. Child Welfare Example: Recovery Coaches for Substance-Abusing Parents

- **Overview of the study:** This Illinois program provided case management services to substance-abusing parents who had temporarily lost custody of their children to the state, aimed at engaging them in treatment. The program was evaluated in a well-conducted RCT with a sample of 60 child welfare agencies, working with 2,763 parents. The study found that, over a five-year period, the program produced a 14% increase in family reunification, a 15% increase in foster care cases being closed, and net savings to the state of $2,400 per parent.

- **Cost of measuring program impact: About $100,000.** The low cost was achieved by measuring study outcomes using state administrative data (e.g., data on foster care case closures).

B. K-12 Education Example: New York City Teacher Incentive Program

- **Overview of the study:** This program provided low-performing schools that increased student achievement and other key outcomes with an annual bonus, to be distributed to teachers. It was evaluated in a well-conducted RCT with a sample of 396 of the city's lowest-performing schools, conducted over 2008-2010. The study found that, over a three-year period, the program produced no effect on student achievement, attendance, graduation rates, behavior, or GPA. Based in part on these results, the city ended the program, freeing up resources for other efforts to improve student outcomes.

- **Cost of measuring program impact: About $50,000.** The low cost was achieved by measuring study outcomes using school district administrative data (e.g., state test scores).

C. Early Childhood Example: The Triple P (Positive Parenting Program) System

- **Overview of the study:** This program is a system of parenting interventions for families with children ages 0-8, which seeks to strengthen parenting skills and prevent child maltreatment. A well-conducted RCT evaluated the program as implemented county-wide in a sample of 18 South Carolina counties. The study found that the program reduced rates of child maltreatment, hospital visits for maltreatment injuries, and foster-care placements by 25-35%, two years after random assignment.

- **Cost of measuring program impact: $225,000-$300,000.** The low cost was achieved by measuring study outcomes using state administrative data (e.g., child maltreatment records).

D. Criminal Justice Example: Hawaii's Opportunity Probation with Enforcement (HOPE)

- **Overview of the study:** HOPE is a supervision program for drug-involved probationers that provides swift and certain sanctions for a probation violation. It was evaluated in a well-conducted RCT with a sample of 493 probationers, with follow-up one year after random assignment. The study found that the program reduced probationers' likelihood of re-arrest by 55%, and the number of days incarcerated by 48%, during the year after random assignment.

- **Cost of measuring program impact: About $150,000.** The low cost was achieved by measuring study outcomes using state administrative data (e.g., arrest and incarceration records).

E. Criminal Justice Example: Philadelphia Low-Intensity Community Supervision Experiment

- **Overview of the study:** This was a program of Low-Intensity Community Supervision for probationers or parolees at low risk of committing a serious crime (compared to the usual, more

intensive/costly supervision). The program's purpose was to reduce the cost of supervision to Philadelphia County without compromising public safety. The program was evaluated in a well-conducted RCT with a sample of 1,559 offenders, with follow-up one year after random assignment. The study found that the program caused no increase in crime compared to the usual, more-intensive supervision of such offenders, indicating that program is a viable way to reduce costs in the criminal justice system. Based on the findings, the county adopted this approach for all low-risk offenders.

- **Cost of measuring program impact: Less than $100,000.** The low cost was achieved by measuring study outcomes using county administrative data (e.g., arrest records).

IV. Why It Matters:

A. Progress in social policy, as in other fields, requires strategic trial and error – i.e., rigorously testing many promising interventions to identify the few that are effective. Well-conducted RCTs, by measuring interventions' true effect on objectively important outcomes such as college attendance, workforce earnings, teen pregnancy, and crime, are able to distinguish those that produce sizable effects from those that do not. Such studies have identified a few interventions that are truly effective (e.g., see Top Tier Evidence, Blueprints for Healthy Youth Development), but these are exceptions that have emerged from testing a much larger pool. Most, including those thought promising based on initial studies, are found to produce few or no effects – underscoring the need to test many. For example:

- **Education:** Of the 90 interventions evaluated in RCTs commissioned by the Institute of Education Sciences (IES) since 2002, approximately 90% were found to have weak or no positive effects.[8]
- **Employment/training:** In Department of Labor-commissioned RCTs that have reported results since 1992, about 75% of tested interventions were found to have found weak or no positive effects.[9]
- **Medicine:** Reviews have found that 50-80% of positive results in initial ("phase II") clinical studies are overturned in subsequent, more definitive RCTs ("phase III").[10]
- **Business:** Of 13,000 RCTs of new products/strategies conducted by Google and Microsoft, 80-90% have reportedly found no significant effects.[11]

B. The current pace of RCT testing is far too slow to build a meaningful number of proven interventions to address our major social problems. Of the vast diversity of ongoing and newly-initiated program activities in federal, state, and local social spending, only a small fraction are ever evaluated in a credible way to see if they work. The federal government, for example, evaluates only 1-2 dozen such efforts each year in RCTs that are usually specially-crafted projects, with research or evaluation funds often paying for delivery of the intervention, recruitment of a sample population, site visits, implementation research, and data collection through researcher-administered interviews, observations, or tests. The cost of such studies is typically several million dollars.

These studies produce important and comprehensive information, but – because of the cost and organizational effort – are far too few to build a sizable body of proven-effective interventions, especially since most find weak or no effects for the interventions being studied. For this reason, we believe such studies may be most valuable when focused on interventions backed by promising prior evidence that suggests impacts will found (e.g., findings from low-cost RCTs, as noted above).

C. Embedding low-cost RCTs in the myriad of ongoing social spending activities can dramatically accelerate the process, enabling hundreds of interventions to be tested each year, rather than a few. Often the key ingredient is creative thinking – i.e., figuring out how to embed a lottery or other randomization process into a particular activity, and measure key outcomes with an existing data source.

REQUEST FOR PROPOSALS:
A high-profile competition to select and fund low-cost RCTs designed to build policy-important evidence about "what works" in U.S. social spending

I. Overview:

A. This RFP invites grant applications for the second year of the competition, in which we will select and fund 3-4 low-cost RCTs – up to $100,000 each. The selected RCTs may fall within any area of domestic social policy; however, at least one will be in an area affecting children and families (consistent with the mission of the Annie E. Casey Foundation, as one of the initiative's funders). The Overdeck Family Foundation may provide funding for an additional low-cost RCT if there are sufficiently strong proposals addressing the Foundation's interest in early childhood, and/or K-12 education, programs.

There will be one additional competition, in 2016; we expect to fund a total of 9-10 low-cost RCTs over the three years of the competition (2014, 2015, 2016).

B. The Coalition will use an expert research panel to evaluate the proposals and select the awardees. We expect the panel to be similar in composition to the expert panel used in the Coalition's Top Tier Evidence initiative.

C. Per the high-visibility nature of this effort, we will invite awardees and finalists to participate in one or more meetings in Washington DC with senior policy officials, to discuss their studies and explore ways to advance wider use of low-cost RCTs in social spending.

II. Application Process and Selection Criteria:

A. The following table shows the requested application materials and timeline:

Stage of application process	Date
All prospective applicants are asked to submit a letter of interest (maximum three pages)	Deadline: February 13, 2015
Applicants will be notified whether they are invited to submit a full proposal (full proposals must be invited)	On or before March 20, 2015
Invited applicants submit a full proposal (maximum six pages)	Deadline: April 30, 2015
Applicants will be notified whether they have been selected for award	On or before May 31, 2015
Grants will be awarded	On or before June 30, 2015

B. Letters of interest and invited full proposals should address each of the selection criteria below, within three pages (for the letter) and six pages (for the proposal). Applicants may use their own format, with single or double spacing, and a font of 11 or larger. The page limit does not include attached letters or other documents specifically requested in this RFP. Please submit all items via email – to Kim Cassel (kcassel@coalition4evidence.org).

C. Selection Criteria – The review panel will consider the following factors in selecting awardees. For the letter of interest: While we ask applicants to address all four criteria, we do not expect applicants to have finalized all aspects of the study design and partnership agreements; therefore, reviewers will focus more on the other two criteria – "importance" and "experienced researcher" – in determining which applicants to invite to submit a full proposal. For the invited full proposal: Reviewers will consider whether all four criteria are satisfied.

- **IMPORTANCE: Whether the applicant is proposing to evaluate an intervention –**

  - **That is backed by highly-promising prior evidence, suggesting it could produce sizable impacts on outcomes of recognized policy importance** – such as educational achievement, workforce earnings, criminal arrests, hospitalizations, child maltreatment, and government spending. For example, we specifically encourage proposals seeking to replicate findings from prior rigorous evaluations that are exceptionally promising but not yet conclusive (e.g., due to only short-term follow-up, a single-site study design, and/or matched comparison groups but not randomization). As a threshold condition for "highly-promising" evidence, proposals should show that the intervention can be or (preferably) has been successfully delivered under real-world implementation conditions.

  - or -

  - **For which there is other compelling reason to evaluate its effectiveness – e.g., it is, or soon will be, widely implemented with significant taxpayer investment**, and its impact on its targeted outcomes is currently unknown.

- **EXPERIENCED RESEARCHER: Whether the applicant's team includes at least one researcher in a key role who has previously carried out a well-conducted RCT (even if not low cost)** – e.g., an RCT with low sample attrition, sufficient sample size, and valid outcome measures and statistical analyses.

- **STUDY DESIGN: Whether the applicant's proposed RCT design is –**

  - **Valid** – i.e., has a sufficiently large sample (as shown through a power analysis) and other elements needed to generate credible evidence about the intervention's impact on one or more targeted outcomes of high policy importance. Preferably, the design will measure such outcomes in both the near term and over a longer period, as appropriate for the type of intervention and study, to determine if the impacts are sustained. The review panel, in assessing an applicant's proposed design, will use the Top Tier Evidence initiative's RCT checklist as a reference.

  - and -

  - **Low cost** – for example, because it (a) will embed random assignment in an intervention that government or philanthropic organizations are already funding or planning to fund; and (b) will measure key outcomes using administrative data that are already collected for other purposes and are of reasonable quality.

- **PARTNERS: Whether the applicant's team includes all parties needed to do the low-cost RCT –** e.g., researcher(s), agency delivering the intervention, agency housing the administrative data. To verify the existence of such a partnership, the review panel will look for attached letters or other communication showing, for example, that (a) a social service agency that delivers the intervention has agreed to participate in the study, including random assignment; and (b) a data agency has agreed to provide the researcher(s) with access to the administrative data needed to measure study outcomes.

D. Other items to include in the letter of interest and invited full proposal:

1. **In addressing the STUDY DESIGN criterion, applicants should specify their primary outcome(s) of interest, how the outcomes will be measured and over what length of time, and what analyses are planned** (e.g., any subgroups to be examined, regression methods to be used).

2. **To address the EXPERIENCED RESEARCHER criterion, applicants should provide reports from 1-2 prior RCTs that the researcher has conducted.** Please submit the relevant study reports (no more than two in all) as attachments. Reviewers will rely primarily on these reports in assessing this selection criterion.

3. **Applicants should specify the amount of funding requested, up to $100,000**, and (for the full proposal only) attach a one-page project budget – with a 10% limit on indirect costs. If additional funding from other sources is needed to carry out the proposed study, we request that the applicant's budget also show the total study cost (including any in-kind contributions of researcher time or other inputs), and include an attached letter or other communication showing that the additional funding has been committed. In such cases, the total study cost – including the additional funding – should still meet the spirit of a "low-cost RCT."

4. **Applicants should briefly address how their study meets recognized ethical standards for research with human subjects.**

5. **In the full proposal, applicants should indicate whether we may share their proposal with others in the event it is a top candidate but cannot be funded by us** (due to limited resources).

III. What To Expect in the Grant Agreement: Awardees will be asked, as a condition of award, to –

- **Pre-register the study,** on the Open Science Framework (OSF) website, and upload a copy of the research and analysis plan in their proposal;

- **Provide us with brief quarterly or semi-annual updates on the study's progress, as well as concise reports on the impact findings at appropriate intervals** – reports that make it easy for readers to see the main results and gauge their credibility (e.g., by showing the similarity of the treatment and control groups in pre-program characteristics, the amount of sample attrition, and the statistical significance of the impact findings); and

- **Make their datasets and related materials (e.g., survey instruments, code used to clean and analyze datasets) publicly available on the OSF site, unless doing so would materially hinder study implementation or raise its cost.** Applicants will be asked to do this within one year of the last data collection, and only to the extent allowed under any confidentiality/privacy protections.

IV. Resources:

- **Please contact David Anderson, the Coalition's vice president, with any questions** (danderson@coalition4evidence.org, 202-239-1248).

- **We will host an optional webinar for prospective applicants in January**, providing an overview of the competition process, an opportunity for Q&A, and practical advice on developing a low-cost RCT from an awardee in last year's competition. A key goal of the webinar is to attract a diverse group of applicants to the competition, including researchers of color. We will send a webinar invitation to persons on our distribution list in early January, providing additional details including the date. Webinar details will also be posted on the competition's website.

- **On the competition's website, we are hosting a message board to facilitate partnerships among the various parties that need to come together to apply for an award** – e.g., (a) researcher, (b) social service agency, and (c) data agency, as described above. We anticipate that the message board may also be used to connect individuals who have promising ideas for low-cost RCTs that they do not wish to carry out themselves, with others who may wish to conduct the study. Individuals seeking others to partner with may submit a message for posting on the board to Kim Cassel (kcassel@coalition4evidence.org, 202-680-8210) that includes the following: (a) a short summary (no more than one page) of what you seek to contribute to an application, and the type of partner(s) you are looking for; and (b) your contact information, for potential partners to reach you.

References

[1] Institute of Education Sciences (of the U.S. Department of Education) and National Science Foundation, *Common Guidelines for Education Research and Development*, August 2013, linked here.

[2] National Research Council and Institute of Medicine, *Preventing Mental, Emotional, and Behavioral Disorders Among Young People: Progress and Possibilities*, Mary Ellen O'Connell, Thomas Boat, and Kenneth E. Warner, Editors (Washington DC: National Academies Press, 2009), recommendation 12-4, p. 371, linked here.

[3] *CBO's Use of Evidence in Analysis of Budget and Economic Policies*, Jeffrey R. Kling, Associate Director for Economic Analysis, November 3, 2011, page 31, linked here.

[4] U.S. Preventive Services Task Force, "Current Methods of the U.S. Preventive Services Task Force: A Review of the Process," *American Journal of Preventive Medicine*, vol. 20, no. 3 (supplement), April 2001, pp. 21-35.

[5] The Food and Drug Administration's standard for assessing the effectiveness of pharmaceutical drugs and medical devices, at 21 C.F.R. §314.126, linked here.

[6] In some RCTs, whole groups (such as schools or counties) – rather than individuals – are randomly assigned to treatment versus control conditions, but the same principle applies.

[7] Examples of additional evidence supplied by traditional RCTs include: (i) corroboration of the earlier impact findings in different samples and settings, thus building strong, replicated evidence of effectiveness; (ii) estimates of the intervention's effect on outcomes other than those measurable with administrative data; (iii) the subgroups and conditions in which the intervention is most effective; (iv) detailed information on the services received by intervention participants, and how they differ from any services received by the control group (so as to assess how much of a contrast in services is needed to generate a meaningful impact); (v) possible reasons why the intervention produced its effect; and (vi) how the intervention's benefits compare to its costs.

[8] Coalition for Evidence-Based Policy, *Randomized Controlled Trials Commissioned by the Institute of Education Sciences Since 2002: How Many Found Positive Versus Weak or No Effects*, July 2013, linked here.

[9] This is based on a count of results from the Department of Labor RCTs that have reported results since 1992, as identified through the Department's research database (link). We are preparing a short summary of these findings, to be released shortly.

[10] John P. A. Ioannidis, "Contradicted and Initially Stronger Effects in Highly Cited Clinical Research," *Journal of the American Medical Association*, vol. 294, no. 2, July 13, 2005, pp. 218-228. Mohammad I. Zia, Lillian L. Siu, Greg R. Pond, and Eric X. Chen, "Comparison of Outcomes of Phase II Studies and Subsequent Randomized Control Studies Using Identical Chemotherapeutic Regimens," *Journal of Clinical Oncology*, vol. 23, no. 28, October 1, 2005, pp. 6982-6991. John K. Chan et. al., "Analysis of Phase II Studies on Targeted Agents and Subsequent Phase III Trials: What Are the Predictors for Success," *Journal of Clinical Oncology*, vol. 26, no. 9, March 20, 2008. Michael L. Maitland, Christine Hudoba, Kelly L. Snider, and Mark J. Ratain, "Analysis of the Yield of Phase II Combination Therapy Trials in Medical Oncology," *Clinical Cancer Research*, vol. 16, no. 21, November 2010, pp. 5296-5302. Jens Minnerup, Heike Wersching, Matthias Schilling, and Wolf Rüdiger Schäbitz, "Analysis of early phase and subsequent phase III stroke studies of neuroprotectants: outcomes and predictors for success," *Experimental & Translational Stroke Medicine*, vol. 6, no. 2, 2014.

[11] Jim Manzi, *Uncontrolled: The Surprising Payoff of Trial-and-Error for Business, Politics, and Society*, Perseus Books Group, New York, 2012, pp. 128 and 142. Jim Manzi, *Science, Knowledge, and Freedom*, presentation at Harvard University's Program on Constitutional Government, December 2012, linked here.