

August 31, 2014

NOTE: Neither IDinsight nor GiveWell has committed to conducting this project.

This note summarizes the current thinking regarding an initiative aimed at appropriately scaling up high-potential development programs. These thoughts are evolving, and questions and comments are highly welcome.

Executive Summary

Despite evidence from rigorous evaluations on the potential for specific anti-poverty programs to have cost-effective social impact, few such programs have been scaled-up, for a variety of reasons. This concept note proposes an initiative to address this challenge, and increase the number of high-potential development programs that are thoughtfully and appropriately scaled-up. The initiative would contain 5 key components: (1) identifying “high-potential” interventions¹; (2) identifying an appropriate context and implementer with the capacity for scale; (3) adapting the high-potential intervention to the new context and implementing partner; (4) rigorously evaluating the impact of the intervention in the new context; and (5) supporting scale-up if warranted based on the evidence generated in step (4). GiveWell has identified four possible “high-potential” interventions that could be part of such an initiative. They are described below.

This initiative seeks to fill an important gap in the global poverty space, although it is important to note that it also comes with significant risks. Identifying partners capable and willing to implement these interventions in appropriate contexts may be challenging. As such, this initiative will require willingness upfront to experiment, fail, learn and adapt. As such, a key component of this initiative will be ongoing and transparent reporting on the failures and successes of this approach, as well as flexibility to adapt the planned activities in the face of new information.

Background and Motivation

In recent years, rigorous impact evaluations in developing countries have drawn attention to several interventions with the potential for large-scale cost-effective social impact. In western Kenya, researchers found that school-based deworming reduced absenteeism by 25% for less than 50 cents per child per year ([Miguel and Kremer 2004](#)). In urban India, an iron fortification program for preschoolers was estimated to cause an additional 2.7 years of education per \$100 spent ([Bobonis et al 2006](#)). Studies such as these have inspired some organizations to scale promising interventions. Based on results from randomized evaluations, [Evidence Action](#) has distributed millions of deworming pills and point-of-source chlorine dispensers, and [TAMTAM](#) has delivered tens of thousands of free anti-malarial bed nets to pregnant women and children.

¹ “High-potential” intervention refers here to interventions that a rigorous impact evaluation has shown to be cost-effective and impactful in a particular context.

Yet the majority of high-potential interventions that have been shown by rigorous evaluations to be impactful in a particular context have not been scaled-up in those contexts. There are many possible reasons why action is often not taken based on the evidence generated, including a mismatch between evaluation priorities and timelines and those of implementing partners, a lack of capacity or interest among implementing partners to scale up based on rigorous evaluations, and a dearth of organizations to help implementing organizations operationalize study findings with on the ground support. Appendix 1 provides a stylized schematic of the connections required for evidence to be translated into at-scale action and potential weak links in that chain.

Furthermore, it is unclear whether the original results of a study would remain the same if interventions were conducted in a different place, with a different implementing organization, or at a different time. [Pritchett and Sandefur \(2013\)](#) show that context may matter more than statistical rigor for extrapolation of evaluation results: they find that non-experimental estimates from the right context are closer to the true effect of a program than experimental estimates from a different context. In Kenya, [Bold et al \(2013\)](#) find that a contract teacher program had large positive effects when implemented by NGOs, but no effect when the same program was implemented by the Ministry of Education.

This concept note proposes to address this challenge by identifying high-potential interventions, adapting and rigorously testing them in appropriate contexts with willing and capable implementation partners, and, if warranted by the evidence, supporting scale-up. The approach described below is intended to address concerns over the “external validity” of rigorous evaluations by testing whether a specific intervention implemented by a specific organization (with capacity and interest in scaling up based on evidence) is impactful in a specific geography and demography – and then supporting scale-up (if warranted by the evidence) of that intervention in that same geography with the same implementing organization. If the evidence does not warrant scale-up, reasons for lack of impact or cost-effectiveness will be assessed and disseminated to inform the global discourse on external validity and program prioritization.

IDinsight and GiveWell’s missions align well for this initiative. The proposed initiative will play an important role in IDinsight’s ongoing mission to “generate and use rigorous evidence to improve social impact” ([IDinsight website, 2014](#)). This initiative will also play an important role in GiveWell’s 2014 research priority “Experimental work to ‘seed’ potential additional top charities,” which includes identifying and funding high-potential interventions in partnership with Good Ventures ([GiveWell website, 2/26/14](#)).

Description of Initiative

This initiative aims to support the scale-up of high-potential interventions through the following steps:

1. *Identify an intervention with high potential for cost-effective impact and scale-up.* Conduct literature reviews, meta-analyses and cost-effectiveness analyses to identify interventions that (i) have rigorous evidence of cost-effective impact (perhaps limited to one or two studies) and (ii) have logistical and political viability to scale-up. Interventions will also be selected based on organizational constraints, such as IDinsight's experience and infrastructure in the program's geographic area. Four high-potential interventions, identified by GiveWell, are described in the next section.

2. *Identify an appropriate context and partner for the selected intervention.* Standardize a process of conducting needs assessments to identify a relevant context where the policy problem is endemic, and in which the theory of change of the original study is likely to hold in the new context. Identify a willing implementation partner who has the capacity to manage and scale the intervention, and meets GiveWell's requirement of being able to absorb funding of approximately \$2-5 million per year. If no organizations have sufficient capacity, then other options may include (i) piloting the intervention for a set period to build the capacity of the implementation partner, (ii) engaging multiple implementation partners with a third party, which would provide central management and oversight, or if necessary (iii) finding a new intervention. Given that finding a capable implementation partner may be challenging and infeasible in some contexts, IDinsight will plan to begin exploring alternative options 4-5 weeks into this phase in tandem with its search for a partner, if the prospects of finding an appropriate partner appear poor.
 - a. *Optional: identify two contexts and partners to simultaneously pilot and test the intervention.* To generate further learning from this initiative, it would be possible to test the intervention in two contexts simultaneously – one in the same geographic region as the initial impact study, and a second in another context where either (i) additional studies have found the intervention to work, or (ii) contextual determinants of program applicability appear to be met. While more expensive, this approach would offer the advantages of (i) increasing the chance of finding significant results in at least one context; (ii) allowing for the possibility of finding two contexts in which the intervention works, creating more room for funding than one alone; and (iii) generating valuable knowledge on the scope and limits of external validity, which may then inform subsequent iterations of this initiative and further work by GiveWell. This initiative is uniquely positioned to shed light on questions of external validity given that the intervention and evaluation design will be coordinated across contexts, and that the contexts may be purposefully selected in advance to maximize expected probability of success and gains to knowledge.²

² In addition to a separate impact evaluation, this option may require additional desk and field research to seek out new contexts in which an intervention may be effective.

3. *Adapt the intervention to the new context.* Assist the implementation partner in drafting guidelines that adapt the original intervention to the new context, being careful to retain the key components of the intervention that are necessary for impact. Conduct process evaluations to identify roadblocks in implementation, and use the results to refine the intervention. These process evaluations will be conducted on a small scale (not powered for statistical analysis), and may include multiple iterations.
4. *Rigorously evaluate the intervention in the new context.* Design an impact evaluation that tests the intervention in a representative subsample of the new context. Alternative design options, identified during process evaluations, may be tested at this time in order to determine the optimal program design in this context. The most statistically rigorous evaluation techniques will be applied, given cost, time and other logistical constraints. When feasible, a randomized evaluation design will be used.
5. *Scale-up or refine the intervention, depending on evaluation results.* Use the results of the impact evaluation to inform decisions for scale-up and/or refinement of the intervention. If the decision is made to scale-up, IDinsight would provide operational support to the partner organization, if necessary. This support may include longer-term assistance with developing a robust monitoring system to help maintain quality of implementation as the intervention is scaled. If the intervention fails to demonstrate cost-effective impact and if opportunities for refinement do not appear promising, assess lessons learned and disseminate findings to contribute to the broader discourse on program effectiveness.
6. *Repeat steps 1 through 5 with a new intervention.* Conditional on the experiences of all stakeholders, repeat the process with another high-potential intervention.

Transparency will play a key role throughout this initiative. Evaluation plans will be pre-registered, data will be published online, and all intermediate and final outputs will be made available for public scrutiny on GiveWell's website.

Possible Interventions

This initiative will identify interventions that (i) have rigorous evidence of cost-effective impact and (ii) have logistical and political viability to scale-up. Based on prior conversations, four interventions, described below, are recommended for consideration. IDinsight and GiveWell may consider other high-potential interventions than these four.

1. *Incentives for immunization*

- a. Policy issue: 2-3 million people die from vaccine-preventable diseases each year. 23 million infants are not reached by routine immunization services, a majority of whom live in India and Nigeria, despite the fact that vaccines are usually provided for free in these countries.

- b. Motivating study: Improving Immunization Rates through Regular Camps and Incentives in India ([J-PAL Overview](#), [J-PAL Policy Brief](#), [Academic publication](#))
- c. Intervention: Improving the supply of infrastructure for immunization camps, and simultaneously increasing demand through modest, non-cash incentives (the study separately measured the impact of improving supply alone and improving supply plus increasing demand).
- d. Evidence of impact: A small bag of lentils for each visit and a set of plates for the last visit increased immunization rates six-fold relative to the control group (39% versus 6% of children aged 1-3 fully immunized).
- e. Cost-effectiveness: Cost per fully immunized child were half as high in immunization camps with incentives than camps without incentives (\$28 versus \$56), since they were much busier. In a meta-analysis of 44 vaccine studies in low and middle-income countries, the cost per disability-adjusted life year averted was less than \$100 in a majority of studies ([Ozawa et al 2012](#))
- f. Potential issues with generalizability: This study has not been replicated in contexts outside of rural Rajasthan, India. The baseline rate of immunization was extremely low in this context: 6% in the control group, versus 80% worldwide. The cost-effectiveness of this intervention relies on finding a context with similarly low baseline immunization rates. The effect of incentivizing immunization without improving immunization infrastructure is not clear. Cultural factors may affect impact in other contexts, such as the appropriateness of food-based incentives and resistance to immunization.

2. Incentives for temporary labor migration

- a. Policy issue: Millions of households are forced into extreme poverty during the pre-harvest “hungry” seasons in rural Asia and Africa due to few employment opportunities and high food prices.
- b. Motivating study: Temporary Labor Migration as Mitigation: Strategies for Managing Seasonal Famine in Bangladesh ([J-PAL Overview](#), [Academic publication](#))
- c. Intervention: Providing cash incentives to enable household members to temporarily migrate to urban areas during the hungry seasons. The cash incentives (about \$11.50 per person) helped households overcome risk aversion and credit constraints associated with the migration.
- d. Evidence of impact: Over 40 percent of households in the treatment group migrated, compared to only 14 percent of households in the control group. Households with migrants increased consumption by 30% and caloric intake by 550 calories per person per day, and were 10 percentage points more likely to send a migrant again in the following season.
- e. Cost-effectiveness: The intervention induced consumption to increase at least \$40 per household in the first year, with additional long-term gains and welfare gains due to consumption-smoothing, easily exceeding the marginal cost of treatment (\$11.50).

- f. Potential issues with generalizability: This study has not been replicated in contexts outside of rural northwestern Bangladesh. The intervention is only appropriate in rural settings with a large number of households reliant on seasonal harvests, without mature credit, insurance or savings markets, and with relatively cheap access to nearby urban centers with robust employment opportunities (but sufficiently expensive that households cannot afford the cost of transportation out-of-pocket).

3. Eyeglasses and academic performance

- a. Policy issue: 10% of primary school children in developing countries have vision problems, which are preventing them from learning.
- b. Motivating study: The Impact of Eyeglasses on the Academic Performance of Primary School Students in China ([J-PAL Overview](#), [Academic publication](#))
- c. Intervention: Providing students entering grades 4 through 6 who have poor vision with appropriate lenses.
- d. Evidence of impact: Wearing glasses for one year increased average test scores by 0.15 to 0.22 standard deviations, the equivalent of one third to one half of an additional year of schooling.
- e. Cost-effectiveness: Eyeglasses cost about \$15. Since each additional year of schooling in rural China increases wages by at least 4.5% (Glewwe et al 2012, p. 28), this intervention has a present discounted value of about \$230 (assuming a discount rate of 5% and a linear Mincerian wage function).
- f. Potential issues with generalizability: This study has not been replicated in contexts outside of rural western China. The intervention is only appropriate in settings where vision problems are widespread and misunderstood, as in many parts of China where glasses are thought to cause vision to deteriorate more quickly, or where credit constraints or imperfect markets prevent families from purchasing glasses for children.

4. Commitment savings products

- a. Policy issue: Household savings are crucial for economic growth, yet many regions, like Sub-Saharan Africa, have had persistently low savings (less than 15% of gross income, [Loayza et al 2000](#)). One constraint to higher savings rates may be time-inconsistent preferences: individuals are more likely to be impatient for near-term tradeoffs than for long-term tradeoffs.
- b. Motivating study: Commitment Savings Products in the Philippines ([J-PAL Overview](#), [Academic publication 1](#), [Academic publication 2](#))
- c. Intervention: Providing bank clients with a commitment savings product, a “SEED” (Save, Earn, Enjoy Deposits) account, which prevents clients from accessing their savings until they reach a self-set goal.
- d. Evidence of impact: 28% of clients offered the account opened one. Average savings balances increased 82% more in the first year for clients assigned to the treatment group than for clients assigned to the control group.
- e. Cost-effectiveness: This intervention is potentially very cost-effective. The accounts are profitable for the bank to maintain, and the partnering

organization in this study has already launched the product in its other branches.

- f. Potential issues with generalizability: Many studies have documented time-inconsistent saving behavior in a variety of developing countries. However, this intervention is only effective for a very specific type of client: someone who has time-inconsistent preferences, but is sophisticated enough to recognize this, and disciplined enough to take action against her future self. This profile applied to less than 1/3 of bank customers in this context. Additionally, the intervention requires a setting where savings rates are low yet financial markets are mature and widespread.

IDinsight is operationally optimized to work in contexts where it has sufficient expertise and infrastructure (South Asia and East and Southern Africa). For these reasons, IDinsight is inclined toward options (1) and (2), which have been implemented and evaluated in South Asia, and therefore may require less time to adapt to a new context. If options (3), (4), or another option were selected, set-up costs and time may be greater.

Key Stakeholders and Roles

- **IDinsight**: IDinsight will identify a high-potential intervention, a relevant context and an implementation partner, with support from GiveWell. As part of this work, IDinsight will standardize a process for identifying contexts in which the findings of a given impact study may be replicable. IDinsight will manage all fieldwork and data analysis related to the process evaluations, impact evaluations and scale-up support. Contingent on funding, IDinsight will offer longer-term support and capacity building for the development of monitoring systems at scaled-up operations.
- **GiveWell**: GiveWell will help IDinsight identify a high-potential intervention. GiveWell will document all initiative activities on their website and, conditional on the results of the evaluation, add the selected intervention to its list of recommended charities. GiveWell will fund the initiative, including implementation costs during the evaluation (?).
- **Good Ventures**: Good Ventures may fund scale-up of interventions found to be impactful by the initiative.
- **Implementation partner**: The implementation partner will adapt the intervention to its context (in consultation with IDinsight and GiveWell), manage the intervention during process and impact evaluations, and, conditional on results, scale-up and/or refine the intervention.

Potential Timeline

Time period	Activities
Oct. 2014	<ul style="list-style-type: none"> • Conduct literature reviews, meta-analyses and cost-effectiveness analyses to identify high-potential interventions.
Nov. - Feb. 2014	<ul style="list-style-type: none"> • Identify an appropriate context and partner for the selected intervention.
Dec. 2014 – Feb. 2015	<ul style="list-style-type: none"> • Draft implementation guidelines with the implementing partner. Submit IRB.
Feb. – June 2015	<ul style="list-style-type: none"> • Pilot intervention on small scale, conduct process evaluations on the intervention, refine and iterate.
July. 2015 – April. 2016	<ul style="list-style-type: none"> • Conduct an impact evaluation on the intervention that also tests alternative design choices. Analyze evaluation results and make decision to scale.
April 2016 onwards	<ul style="list-style-type: none"> • Conditional on the results of the impact evaluation, scale the intervention and/or refine and conduct a new impact evaluation.

Appendix 1

Evidence-based policymaking: Evaluation to scale-up

