

The programs and their evaluations

1. Evaluations of large-scale programs rather than small-scale pilots
 - Colombia (Familias en Acción)
 - Argentina (Programa Jefes)
 - Mexico (PROGRESA-Oportunidades)
2. All three programs politically important and very visible
3. All three have had long-standing evaluation efforts, as opposed to a one-off evaluation

Findings from the evaluation

1. All three reported some positive impacts, and these have received much attention
 - PROGRESA, Familias en Acción: impacts on schooling, consumption poverty
 - Jefes: net income gains, equivalent to ~1/2 to 2/3 of value of cash transfer
2. Less clearly positive effects have received much less attention
 - Impacts of CCTs on child anthropometrics are mixed and even those for PROGRESA are fragile

Positive findings helped sustain and expand programs...

1. Argentina: “Evaluation provided credibility to the program and information about program implementation in a very unstable environment”
2. PROGRESA and Familias evaluations may have helped ensure the sustainability of programs across administrations
3. The PROGRESA evaluation also helped support the “CCT wave

...and led to much greater awareness of the importance of rigorous evaluations

1. The PROGRESA evaluation led to rigorous impact evaluations of CCTs in a number of countries
2. Systematic use of evaluations to draw policy conclusions
 - Upcoming Policy Research Report on CCTs
3. PROGRESA evaluation also provided part of the impetus for the creation of CONEVAL, the National Evaluation Council, with a mandate to rigorously evaluate other social sector interventions

What features of programs and evaluations made it more likely that evaluations would have policy impact?

1. External evaluation
 - Groundbreaking decision by PROGRESA to hire IFPRI—respected, and able to hire external academics of stature
 - Familias and IFS
 - Jefes and the World Bank Research Department
2. Data made available for verification
3. (and the fact that the evaluations found positive results, which was certainly helpful)

What features of programs and evaluations made it more likely that evaluations would have policy impact?

1. Thought experiment: What if the PROGRESA evaluation had found no positive impacts? Would it have been terminated?
 - Unlikely, at least during the Zedillo administration
 - And arguably what really changed President Fox's mind about PROGRESA was not so much the evaluation results, but the fact that he realized the *political* potential of the program

What features of programs and evaluations made it more likely that evaluations would have policy impact?

2. As it is, when evaluations find results that conflict with campaign promises or political imperatives, they are frequently brushed aside
 - Familias en Acción: rapid expansion into urban areas, despite the cautionary results: “the urban expansion, driven by political agenda, was done very quickly...”
 - BDH program in Ecuador:
 - Randomized evaluation showed no impacts among (relatively) better-off households
 - Widespread dissemination of results
 - Incoming Correa administration doubled the size of transfer to all households, including the better-off

What features of programs and evaluations made it more likely that evaluations would have policy impact?

3. Incentive problems are also important elsewhere
 - Donors and Bank TTL are less likely to enthusiastically welcome evaluation findings that show them that “their” projects have no impact
 - Publication bias: Difficulty publishing articles that show “zero effects”

So, what's to be done? (...other than to be modest and realistic)

1. Evaluate programs before they become nationwide
 - (i) Small-scale pilots before a program is begun
 - (ii) Adjustments to be overlaid on existing programs
 - One problem: External validity—are the results from these small-scale pilots, which often evaluate interventions in carefully-controlled “laboratory” settings, applicable to large-scale programs?
 - US literature: Head Start versus Perry preschool program
 - (iii) Evaluate programs and variants that are not yet nationwide, but are large enough to be realistic
 - Welfare reform in the United States

So, what's to be done? (...other than to be modest and realistic)

2. Make data available
3. Bring results from credible evaluations elsewhere to bear on policy dialogue
 - International public goods: role for multilateral agencies