

# Comparative Cost-Effectiveness to Inform Policy in Developing Countries

---

Iqbal Dhaliwal  
Director of Policy, J-PAL  
Department of Economics, MIT

*Exec-Ed, Cambridge, May 26, 2011*

# Cost-effectiveness Analysis (CEA) and Comparative CEA

---

- Summarizes complex program as simple ratio of costs to effects
  - Measure the cost for a given level of effectiveness: e.g. cost to increase school attendance by 1 year
  - Or, measure the level of effectiveness for a given cost: years of additional attendance induced by spending \$100
  - Must accurately reflect the program based on costs and impacts actually observed
  - Sensitivity analyses provide additional information
- **Comparative** cost-effectiveness then compares this cost-effectiveness ratio for multiple programs
  - Must compute costs and benefits using similar methodology for all programs
  - If done correctly, it helps policymakers choose between multiple programs addressing the same problem – e.g. what are the most cost-effective ways to increase school attendance?

# Outline

---

## 1. Motivation

- Bridge the gap between research and policymaking in developing countries
- Propose a standardized approach to conducting CEA
- Provide calculations on the web to encourage policymakers to use research
- Create worksheets for gathering costs as part of future research

## 2. Quantifying Impacts

## 3. Quantifying Costs

## 4. Common Units across Countries and Time

## 5. General Issues in CEA

# Cost-Effectiveness (CEA) vs. Cost-Benefit Analysis (CBA)

---

- CBA compares monetary value of benefits against costs:
  - As a result CBA can show ratio of all benefits to costs of a program
  - CBA can also deliver absolute judgment on whether program is worth the investment, CEA cannot
- But this means that CBA requires a number of assumptions about the monetary value of all the different benefits
  - This requires value judgments (cost of life, disability, lower crime among school kids)
- Advantage of CEA is its simplicity:
  - Allows user to choose an objective outcome measure (e.g. cost to induce an additional day of schooling) – no need for making judgments on monetary value of that schooling
  - Easier for policymakers to compare programs when they are primarily concerned about one outcome of interest (e.g. increasing school attendance, not child health)

# CEA is Rarely Done for Developing Countries

---

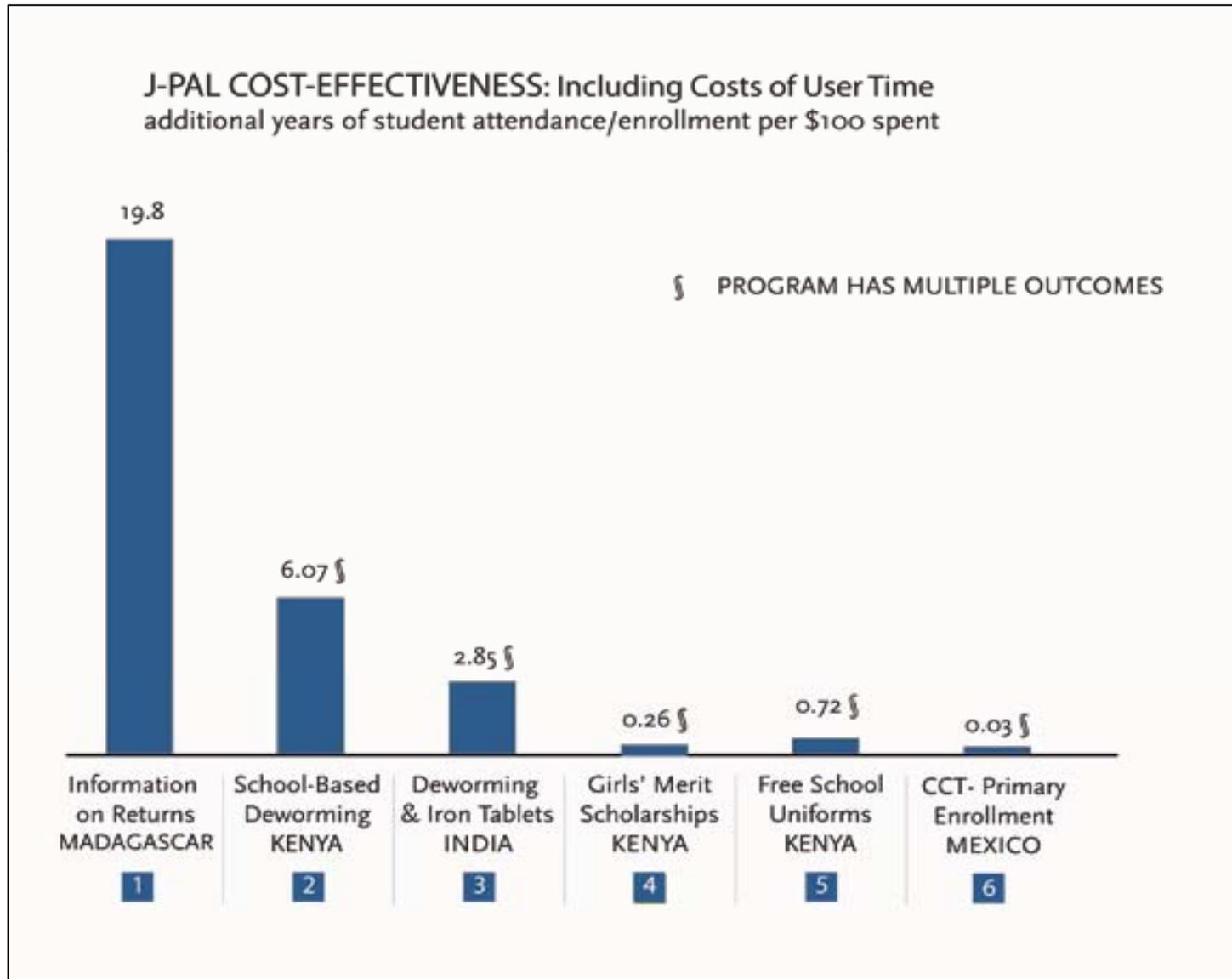
- Cost-effectiveness analysis is very widespread in health studies, but much less so in education
  - Henry Levin and Patrick McEwan (2001) provide an excellent overview in their book
- Very few **comparative** cost-effectiveness of different programs
  - Kremer, Miguel and Thornton (2005)
  - Banerjee, Cole, Duflo, Linden (2007)
  - But all were done for the same country / organization
- Even rarer to see comparative cost-effectiveness studies **across** developing countries because:
  - Difficult to get data on costs, especially from other authors
  - Difficult to compare studies - exchange rates, inflation, discount rates, differing factor endowments, efficiency of implementers
  - Absence of incentives to do CEA - academic and publication (what happens when the program you evaluated is effective but not really cost-effective, no editorial demand to show CEA)

# But CEA can be a Vital Tool for Policymakers

---

- Contrast the lack of incentives for researchers to do CEA to the needs of policymakers
  - Policymakers are either funders or implementers in developing country governments and NGOs, foundations, and international development agencies
- Limited time and resources to analyze hundreds of research studies in academic journals (and growing rapidly)
  - As a result research results are often completely ignored
  - Or used selectively to support “instincts”
  - And policymakers get into a continuous “state of trial and error” – but easier to roll out new programs than to scale them back later
  - Egg inspectors! – more frequent than you would think
- CEA makes research more accessible to policymakers and allows them to make more informed choices

# But CEA can be a Vital Tool for Policymakers



# Outline

---

1. Motivation

**2. Quantifying Impacts**

- Spillover Effects
- Programs Achieving Multiple Impacts
- Imprecision in the Estimation of Impacts
- Aggregating Impacts

3. Quantifying Costs

4. Common Units Across Countries and Time

5. General Issues in CEA

# Sources of Impact Estimates

---

- Randomized evaluations
  - Done by J-PAL's 50+ affiliates (more than 250 evaluations in 42 countries)
  - Any other rigorous RCTs that report cost data or where authors are willing to share cost data not published in the paper
- Non-RCTs not included in our analysis, but can be easily added to our worksheets provided on the web

# Spillover Effects

---

- Effects which accrue to individuals not directly treated by program
  - e.g. Deworming schoolchildren reduces disease transmission, decreases likelihood of infection among non-treated children
- Two kinds of spillover effects possible in randomized evaluation
  - Effects on untreated individuals in ITT group - e.g. Children who refused to take deworming pills
  - Effects on individuals in control group - e.g. Children in schools randomly assigned to receive treatment next year
- Which benefits would be present in scaled version of program?
  - Some children will remain untreated at larger scale, so include
  - No control group when program implemented on larger scale, therefore exclude

# Multiple Impacts

---

- Many anti-poverty programs have multiple impacts on lives of poor, cost-effectiveness can only compare costs to effect on one outcome
  - e.g. Giving schoolchildren free meals incentivizes attendance, but also improves nutritional status
- No easy resolution within CEA:
  - Separate out costs which relate to that one impact (e.g. cost of health subsidy component in CCTs)
  - But overhead and administrative costs are much harder to separate
  - *Alternative*: Allocate costs equally between major multiple outcomes – if the cost/effect ratio is comparable to best alternatives to reach those outcomes then it is cost-effective (e.g. deworming students costs \$12.17 per DALY saved and \$8.23 per additional year of schooling when costs split evenly)
  - Make clear in the analysis that program has multiple outcomes
- But CEA very useful as often policy is made with one objective in mind – e.g. education department or foundation wanting to maximize school attendance, irrespective of health benefits

# Imprecision in Estimating Impacts

---

- Some impact results not significant- may represent two things:
  - Point estimate is precisely measured and close to zero – want to show this because absences of impact is very useful information too
  - Point estimate is moderate or large, but imprecisely measured
- Even if we establish a significance “cutoff”, problems remain:
  - Two CE estimates may appear different, but not differ significantly when variance around point estimates is taken into account
  - What do we do with “insignificant” results?
- We take middle ground:
  - Show bars for only point estimates that are significant at 10% level. Place other programs on the graph, but without bar for point estimate and with note
  - Provide table with information on confidence intervals around impact estimates
  - Too much information can obscure main message for non-research audience

# Imprecision in Estimating Impacts

---

- **One idea:** Rank programs based on their cost-effectiveness using point estimate of impact, and then for each program...
  - If point estimate of impact puts CE above 75<sup>th</sup> percentile of various programs being compared, but the lower bound puts it below the 25<sup>th</sup> percentile of CE, it is too imprecisely measured to be included
  - Tradeoff between providing all possible information and keeping the comparison simple enough to be useful

# Aggregating Impacts

---

$$\text{Total Impact} = \text{Impact (per unit)} \times \text{Sample Size} \times \text{Program Duration}$$

- **Sample Size:** If impact is multiplied by correct sample, ITT and ToT estimates should give same total impact
- **Time:** Make sure impact is from the same period of exposure to program as costs are:
  - E.g. care needed to ensure not using costs over one year vs. impacts based on several years of exposure to the program
- May need to translate proximal impacts (e.g. increase in chlorination of water) to final impacts (e.g. reduction in diarrhea incidents) using estimates from meta-analyses

# Outline

---

1. Motivation
2. Quantifying Impacts
- 3. Quantifying Costs**
  - Gathering Cost Data
  - Modeling Costs at the Margin
  - Free Goods and Services
  - Costs of User Time
  - Ingredients with Overlapping Uses
  - Transfers
  - High-Level Management Overhead
  - Experimental vs. Scalable Costs
4. Common Units Across Countries and Time
5. General Issues in CEA

# Gathering Cost Data

---

- Ingredients method (Levin and McEwan 2001)
  - **Academic Papers:** Good listing of ingredients, but rarely have cost figures
  - **Budgets:** Have accurate costs but need to link to program (e.g. acquisition of assets vs. depreciation, aggregate costs of inputs but many used for other purposes too)
- Gather cost data from multiple sources:
  1. Academic paper for description of program structure, ingredients
  2. Talk to researchers for additional ingredients, their costs and budgets
  3. Program staff and field research staff for unit cost data
- Much easier when program costs tracked with CEA in mind
  - General worksheet for tracking costs being developed, seeking feedback

# Modeling Program Costs at the Margin

---

- What is meant by “the program” (and therefore the margin)?
  - Assume program will be added to existing government or NGO infrastructure
  - When multiple variants examined in same evaluation, be specific about which variant is being analyzed as this affects calculation of costs/impacts
  - Example remedial teaching vs. remedial teaching with new pedagogy
- Be specific about the margin onto which the program is being added (e.g. Contract teachers in Kenya)
  - Program gave schools funds to hire a contract teacher to teach an *additional* first grade class → impact = contract teacher + smaller class size
  - Evaluation allows isolation of impact of contract teacher vs. civil service teacher since both teach different classes
  - Is “the program” the replacement of civil service teachers with contract teachers? Or the addition of contract teachers to create smaller classes?
  - The cost of the first (replacement) will be lesser than that of the second

# Another Example (Marginal Program Costs)

---

- Computer-assisted learning

- If computers already present but under-utilized, wouldn't have to pay for them
- If computers not present, or used at capacity, would have to pay for them

$$\text{Benefit-Cost Ratio} = \frac{[\text{Test Scores with Program}] - [\text{Test Scores Without Program}]}{[\text{Costs With Program}] - [\text{Costs Without Program}]}$$

- Specify “**Comparator Case**”, i.e. the margin on to which the new Program is being added

- This defines the basic level of costs (teacher salaries, classroom supplies) and achievements (student attendance, test scores) that exist even in absence of program
- Comparator case must be well-specified, consistent across all programs in an analysis

# Free Goods and Services

---

- Some goods and services provided for free in evaluations
  - e.g. Deworming drugs, donated community labor
- Costs of goods and services provided for free in evaluation should be included in estimation of costs
  - Many ingredients not free when program replicated elsewhere, at larger scale
  - If assessing cost-effectiveness to society, costs of free ingredients should be included
  - Program funders or implementers may choose to take this out in their calculations
- Apply ingredients method using market costs of the ingredients
  - Program staff often do not keep a tally of this therefore important to emphasize ahead of time
  - For labor costs estimate what it would cost to get same work done in the market

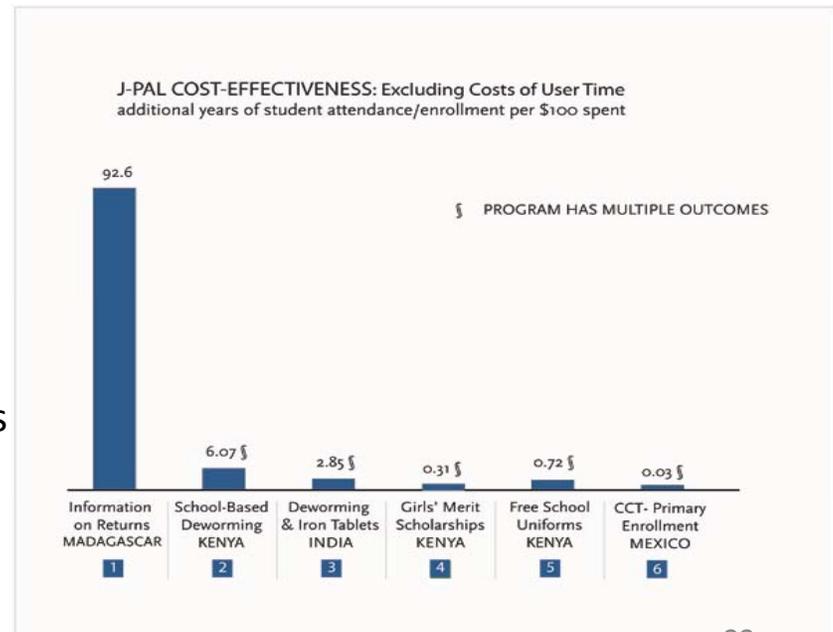
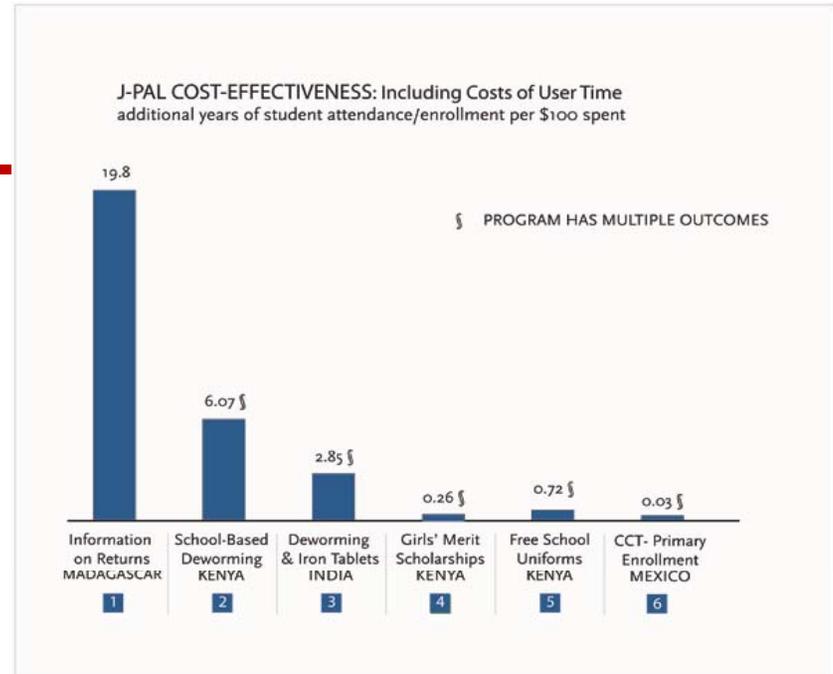
# Costs of User Time

- Users must spend time participating in program

- E.g. parental consent meetings
- We care about costs to beneficiaries, since programs designed to help them
- Costs should be included if user time is not voluntary but required
- Many funders may choose to exclude these

- Costs of user time

- Most studies report average household wages for treatment and control groups
- Local wages usually small relative to NGO wages or material costs
- Including/excluding user costs usually doesn't change relative ranking of programs



# Ingredients with Overlapping Uses

---

- New programs may use ingredients whose uses overlap with other existing programs, or with future programs
  - Example is targeting surveys to identify beneficiaries for PROGRESA or for a Public Distribution Systems for subsidized food grains (PDS)
  - If the ingredient was already paid for before the program started, then do not include in the cost (similar to the Marginal Program Cost analysis)
- But if the ingredient was first acquired as part of the program, then include it in cost even if it may be used in other future programs as:
  - It is hard to model in the uncertainty of shared future use,
  - Subsequent programs will not include it in their CEA,
  - In replications in similar situations this cost is likely to be incurred, and
  - If it is essential for the success of the program then the funder has to budget for it
  - Providing underlying calculation worksheets gives the policymaker an opportunity to see the impact of the ingredient's cost and exclude it if appropriate for context

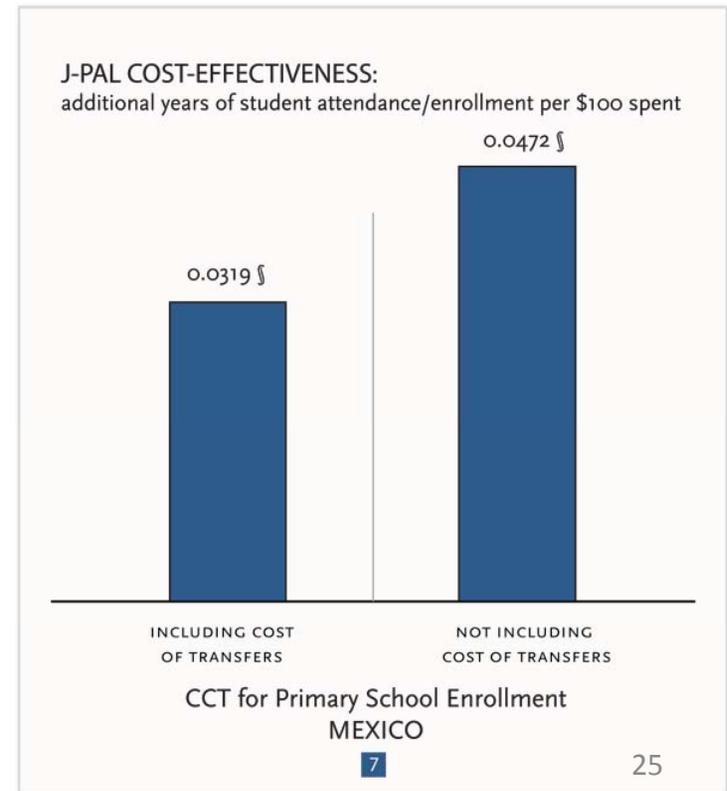
# Transfers

---

- Two types of costs associated with Transfer Programs:
  - Administrative costs of getting transfers to users are real resource cost and should be included
  - Actual cash transfer represents an accounting cost to the organization funding the program, but not a cost to society. Should we count cost of transfers themselves as cost of program?
  - Would not be a problem in CBA – shows up as both a cost and a benefit (ignore deadweight loss of increased taxes to fund transfers)
- We do not net out costs of non-cash transfers (e.g. free school meals)
  - To net out costs would have to assume a “value” for the benefit
  - Implementer may have different valuation of good than recipient, after all transfer is given to deliberately skew consumption

# Cash Transfers may be viewed as Multiple Outcomes

- Should we net out costs of cash transfers (e.g. PROGRESA, in student attendance analysis) ?
  - Value of transfer is not “used up” in incentivizing attendance, merely redistributed
  - But cash transfer is a separate outcome from student attendance, similar to school meals, but just easier to estimate monetary value of
  - Cash transfers are special case of multi-outcome program in which value of outcome is easier to measure
  - Moreover for budget constrained organizations, this is a relevant cost to consider
  - So we include costs of transfers as program costs
  - When costs are disaggregated, relatively easy to show cost-effectiveness with and without transfer costs



# High-Level Management Overhead

---

- Additional time, effort, and cost of high-level administration very hard to observe
  - e.g. a program requiring protracted renegotiation of union contracts
  - Organizations rarely track or quantify such overhead, especially at pilot or evaluation stage when they represent small amounts within large organizational budgets
- In many cases, costs of high-level overhead likely to be similar across programs
  - Inclusion/exclusion should not affect relative cost-effectiveness, especially when comparing programs at pilot or evaluation stage
  - If one program thought to have significantly higher costs of overhead, could choose some “overhead rate”, say 15% to adjust costs based on consultation with implementer
  - For scale-ups, organizations should make this adjustment in CEA worksheets

# Experimental vs. Scalable Costs

---

- Programs intended for evaluation often purchase goods and services at retail prices, administer program using smaller infrastructure
  - Scale costs could be lower, if goods and services are purchased in bulk and there are economies of scale to administration
  - Scale costs could be higher, if it is much harder to administer and supervise a program to a larger number of dispersed beneficiaries
  - For this reason, stick to costs reported from research under most cases
- Exception is when scale economies are very obvious and guaranteed to be realized, these may be used to estimate costs of program
  - Example is purchase of school inputs like textbooks or flipcharts
  - Must be judicious in applying scale costs - concern that while costs might go down with scale, so would quality and impact of program.
  - When changes are made to program design in scaled version, best to conduct pilot evaluation to test that costs and benefits are comparable in new environment

# Outline

---

1. Motivation
2. Quantifying Impacts
3. Quantifying Costs
4. **Common Units Across Countries and Time**
  - Time Discounting
  - Inflation
  - Currency and Exchange Rates
  - Order of Operations
5. General Issues in CEA

**Base Year:** *First year that program is implemented, year in which present value is taken.*

**Year of Analysis:** *Year in which cost-effectiveness analysis is undertaken, or choice between various programs is made.*

# Time Discounting to Base Year

---

- Needed when costs and impacts distributed over different years
- Discounting **benefits** is representative of how users would trade off between use of services this year vs. next year (time preference)
  - Discount rate calculated as Social Rate of Time Preference (SRTP), rate at which users would trade off one unit of consumption today vs. tomorrow
  - Little empirical data on time preference in the developing world, not a practical option for choosing a discount rate
- Discounting **costs** is representative of opportunity cost of funder of spending money today vs. investing for positive rate of return
  - Discount rate calculated as Social Opportunity Cost of Capital (SOC), foregone rate of return on capital markets
  - Less variance than SRTP and better documented for developing countries
  - Developing countries use 8-15% compared to 3-7% in developed countries
  - We use 10% which is around the median rate (Asian Development Bank)
  - Rate choice only affects long-run programs with different benefit-cost streams

# Inflation

---

- Depending on how costs reported, need up to two adjustments for inflation:
  - If a stream of costs is reported in nominal prices, must deflate back to real prices in terms of program's base year
  - If base year of program differs from year of analysis, must inflate forward to year of analysis prices
- Should be no difference between converting to USD then inflating or vice versa
  - But distorted exchange rates may not always capture inflation adequately, so must be consistent in order of operations (inflation then exchange or vice-versa)

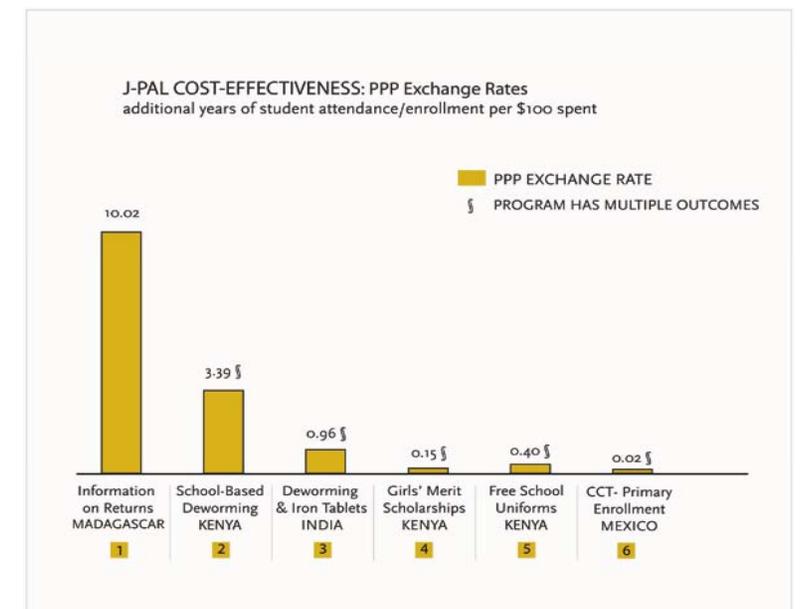
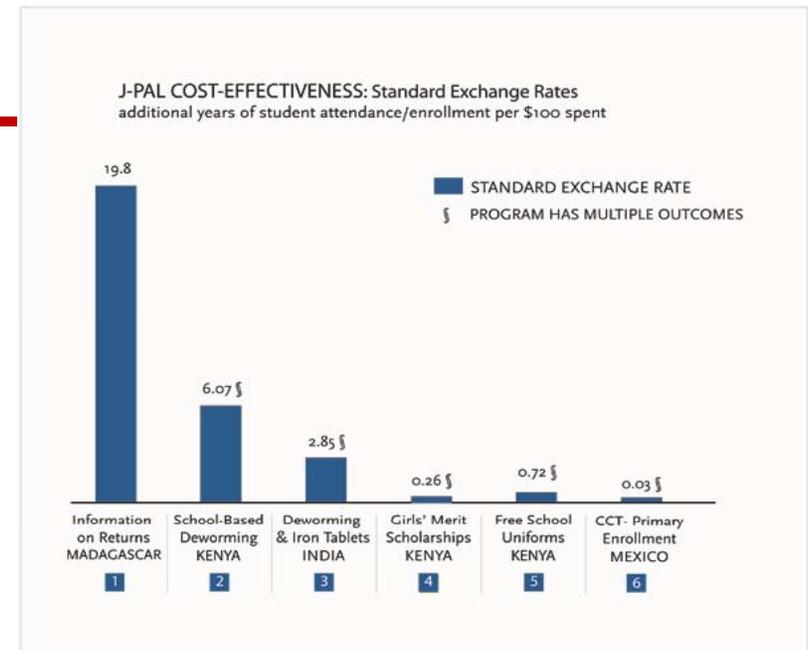
# Currency and Exchange Rates

---

- Costs reported in variety of currencies which must be exchanged into common unit - USD
  - Not a problem when comparing two programs run in the same country
  - But can be a problem when comparing across different developing countries
- When standard exchange rates are used, CEA figures do not reflect real purchasing power differences across countries that are driven mostly by higher prices of non-tradables in richer countries
- But even PPP exchange rates only somewhat adjust for different price levels across countries based on a standard basket of goods and services:
  - But don't completely adjust for different relative prices of goods and services used in a particular program that may be different due to that country's factor endowments
  - e.g. Skilled labor is much cheaper in India than in Mexico, so a highly labor-intensive program piloted in Mexico will appear less cost-effective even when adjusted by PPP

# Exchanging Currency

- Another problem with PPP rates is that readers often misinterpret them as standard rates:
  - Since PPP rates for many developing country currencies are higher than for standard rates, it makes all programs in an analysis look less cost-effective
  - Possibility of discouraging potential donors or implementers
  - We will present both PPP and Standard Exchange rate version (with the latter as default)
  - Example suggests that in most cases PPP vs. standard rates doesn't change relative cost-effectiveness of programs



# Standard Order of Operations across all CEA

---

Step	Operation	Unit of Currency
1.	Gather cost data	2004, 2005, 2006 MXN
2.	Exchange into USD using year-specific exchange rates	2004, 2005, 2006 USD
3.	Deflate nominal costs back to real value in base year (2004) prices using avg. annual US inflation over time between base year and incursion of costs	2004 USD (incurred in 2004, 2005, 2006)
4.	Take present value of cost stream	PV of cost stream in 2004, in 2004 USD
5.	Inflate forward to year of analysis (2009) using avg. annual US inflation over time between base year and year of analysis	PV of cost stream in 2009, in 2009 USD

# Outline

---

1. Motivation
2. Quantifying Impacts
3. Quantifying Costs
4. Common Units Across Countries and Time
- 5. General Issues in CEA**
  - Partial vs. General Equilibrium
  - Initial Levels of Underlying Problems
  - Generalizability of Costs

# Impact Size in Partial vs. General Equilibrium

---

- Even when evaluations designed to pick up long-run effects, impacts may be slightly different at general equilibrium
  - e.g. Graduates of a vocational education program may get higher returns to education when program is small and their knowledge is scarce
  - Or graduates could get higher returns when program is scaled up, due to complementarities among skilled workers
- Not practical to try to estimate general equilibrium effects in a cost-effectiveness model, must focus on what has been rigorously measured
  - When presenting results, acknowledge that effects of certain outcomes (e.g. school attendance or test scores) may vary when program more widely adopted

# Initial Levels of Underlying Problems

---

- Baseline rate of problem will vary by context, but impact may also vary due to declining marginal benefit from intervention
  - e.g. Deworming more effective in increasing attendance where worm load is high
  - e.g. Raising attendance from 85% to 90% may be tougher than from 50% to 55% (last mile problem)
  
- But we do not try to adjust programs to assume some average baseline rate of problem:
  - Whether there is a last mile problem and to what extent is an empirical question that is hard to resolve a priori
  - A program is likely to be implemented in an another area where the same problem it addresses is similarly salient
  - While not perfect, providing underlying CEA calculations will give an opportunity to conduct sensitivity analysis along some of these parameters like baseline prevalence, population density and other demographics, and impact (using lower and upper bounds)

# Generalizability of Costs

---

- Costs can also vary across different contexts within even the same country due to demographic factors:
  - e.g. Per unit costs of certain interventions may be lower if population density is higher
  - e.g. The number of houses a marketer can visit in a day may vary depending on household size, literacy, population density, transportation infrastructure etc.
  - e.g. Availability of volunteer tutors may be lower within a country in less literate states compared to more educated states, driving up their wages in the former
- When cost-effectiveness particularly sensitive to unit costs, it can be useful to do sensitivity analysis varying this figure

# Conclusion

---

- CEA has long been known to be an effective tool to compare different programs and can help bridge the divide in developing countries between a large and growing body of rigorous research and policy
- However CEAs have failed to become widespread due to:
  - The absence of reliable cost data,
  - The time and effort required to conduct a thorough CEA because of the large number of issues around impact, cost and making the units comparable, and
  - The absence of academic or publication incentives to undertake CEA
- These barriers can be overcome by:
  - Agreeing upon a standard methodology for conducting CEA by field,
  - Integrating collection of data on costs and impacts into the research projects,
  - Including cost-effectiveness analysis as part of the published paper, and
  - Making available the underlying assumptions and worksheets so policymakers can perform sensitivity analysis

MIT OpenCourseWare  
<http://ocw.mit.edu>

Resource: Abdul Latif Jameel Poverty Action Lab Executive Training: Evaluating Social Programs  
Dr. Rachel Glennerster, Prof. Abhijit Banerjee, Prof. Esther Duflo

The following may not correspond to a particular course on MIT OpenCourseWare, but has been provided by the author as an individual learning resource.

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.