Abstract: Empirical Evidence and the Case for Foreign Aid

This paper addresses an important methodological question for a recent debate in global justice: What types of data are necessary for settling normative debates about foreign aid? Recently, several philosophers have considered the case for foreign aid and have concluded that foreign aid is either ineffective or counter-productive. This paper considers what kinds of evidence those doing applied philosophy must use to support different claims about aid's efficacy. Then, using some of the best available data, this paper makes a strong case for at least some foreign aid. This paper's methodological lessons, however, are quite general. It considers what types of data are necessary for establishing the kinds of empirical claims often relied upon in the political philosophy and public policy literature.

Empirical Evidence and the Case for Foreign Aid[i]

1. Introduction

In his groundbreaking article, "Famine, Affluence, and Morality," Peter Singer gave the following argument:

1) Suffering and death from lack of food and shelter and medicine are bad.
2) If we can do something to help prevent suffering and death from lack of food and shelter and medicine without sacrificing anything morally significant (or of morally comparable worth) then we should.
3) So we should help prevent this suffering and death by giving foreign aid.[ii]

Recently, several philosophers have questioned Singer's implicit assumption that we can do something to help prevent suffering and death from lack of food and shelter and medicine by giving foreign aid.[iii] These critics argue that foreign aid is either ineffective or counter-productive.[iv] This paper is not concerned to address the details of particular critics' arguments (the critics appeal to a mixture of fundamental principles and empirical considerations).[v] Rather, this paper examines the kind of empirical justification necessary to arrive at sound conclusions regarding aid's efficacy.[vi] Using the best available data, the paper then makes a strong case for at least some foreign aid programs. That aid can do some good may not be very surprising. Those who deny this claim must show that aid *generally* does more harm than good.[vii] Those who endorse it need only show that aid *can* do some good. The heavy burden is on the negative.[viii] Since, however, critics have recently argued that Singer is wrong in claiming that aid *can* do some good, it is worth defending Singer on this point.[ix]

This paper's methodological lessons are, moreover, quite general. It considers what types of data are necessary for establishing the kinds of empirical claims often relied upon in the political philosophy and public policy literature. This is important as such empirical claims underlie debates about everything from free trade, to immigration policy, and taxation to population control and many of the same problems that beset critics of aid also undermine arguments in these other areas.[x]

Before making the case for some foreign aid programs, the next section will consider the different kinds of -- micro- and macro- level -- empirical data available.[xi] Macro-level data tells us how aid affects (or types of aid affect) all countries (or at least large regions). Micro-level data tells us how particular aid projects or small subsets of projects are doing. So the best data to use if one wants to conclude that aid is generally good or bad is macro-level data. Micro-data is better if one wants to know whether a particular aid program is good or bad.

2. The Macro-Level Data

There have been roughly three waves of macro-level work on foreign aid's efficacy.[xii] Studies in all three look, primarily at aid's impact on growth.[xiii] Over time these studies have become more sophisticated testing more complicated models of how different kinds of aid impact growth.[xiv] Some surveys of the literature, such as that of C. Michalopoulos and V. Sukhatme and H. White, suggest that the evidence is ambiguous.[xv] One of the latest and most comprehensive reviews by H. Hansen and Finn Tarp suggests, however, that 'aid works, even in countries hampered by an unfavorable policy environment.'[xvi] They question the robustness and appropriateness of models underlying many of the studies which deny this conclusion.

Unfortunately, the evidence about aid's impact on growth cannot tell us what impact aid has on the global poor.[xvii] Even if aid increases growth, it may not reduce poverty. Growth can increase even if the poor become poorer as long as the rich or middle class gain more than the poor lose. Furthermore, different causes of growth affect the poor differently.[xviii] The growth impact of aid alone does not tell anything about how aid impacts poverty.[xix]

Fortunately, there are some relevant studies (though different ways of measuring aid's impact on poverty yield different results).[xx] Peter Boone, for instance, shows that non-military aid did not reduce infant mortality, increase life expectancy, or increase primary schooling in the 1970's and 1980's.[xxi] More recently, however, Paul Mosley, John Hudson, and Arjan Verschoor find that aid positively impacts pro-poor spending in low income countries and usually reduces income poverty and infant mortality.[xxii] They find that aid has an especially large positive impact on poverty when there is not too much inequality or corruption. Boone's study does not consider aid's impact on infant mortality *in low income countries* in the 1970's and 1980's. So, even if Boone is right, Mosley et. al.'s study gives us reason to think aid generally reduced infant mortality in low income countries in the 70's and 80's. When inequality and corruption are low, there is also reason to think aid has a positive impact on income poverty in these countries.[xxiii] Others replicate some of these results finding even greater impacts of aid from European nongovernmental organizations (NGOs) on infant mortality than Official Bilateral Aid which affects public sector spending.[xxiv] A recent study by Karuna Gomanee and Oliver Morrissey looks at aid's impact on public expenditure, infant mortality, and the Human Development Index (HDI). Looking at a panel of countries from 1980-1998 Gomanee

and Morrissey find that aid increases public expenditures and decreases mortality rates (increasing the HDI).[xxv] Though there may not be enough macro-level data to conclude that aid generally reduces poverty, there is evidence in favor of some kinds of aid.[xxvi]

Nevertheless, trying to make a general case for aid may be like trying to make a general case for investment.[xxvii] If Singer is right, what we really need to know is what investments to make and what aid to give.[xxviii] So the next section will turn to the micro-level data to get more information.[xxix]

3. The Micro-Level Data

Micro-level data can be *experimental*, *quasi-experimental*, or *non-experimental* (so can macro-level data, though most of it is quasi- or non-experimental). Most micro-level data is non-experimental. Non-experimental evidence can include, for instance, historical records, observational studies, and anecdotal evidence. Both quasi-experimental and experimental evaluations help test the causal efficacy of aid programs; they help insure a study's *internal validity*. A study has high *internal validity* when it captures the causal relationships between the particular program, policy, or process being evaluated and the particular outcome observed (no matter how unique the circumstances of the study). In *experimental* studies people (or other units of analysis) are assigned randomly to treatment groups (e.g. those receiving aid) and comparison groups (e.g. those not receiving aid).[xxx] This helps isolate a program's impact on participants. Q*uasi-experimental* studies do not use random assignment to insure internal validity. Instead, researchers try to minimize *selection bias* (the bias that results from differences between the treatment and the comparison groups) in other ways.

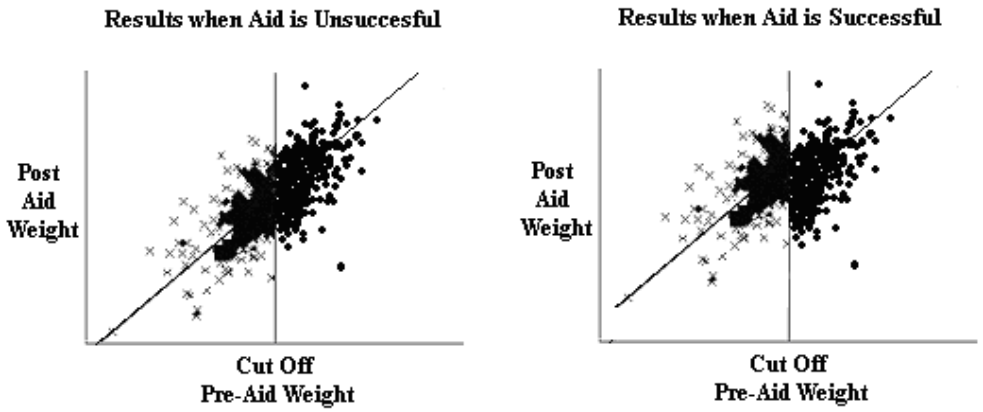
Experimental data is often best for insuring internal validity.[xxxi] Random assignment to treatment and comparison groups ensures that, on average, there will be no relevant differences between participants and non-participants. With a quasi-experimental design participants are more likely to differ from those in the comparison group in important ways.

An example will help illustrate the advantages and limitations of experimental and quasi-experimental studies. Suppose we want to evaluate an aid program to reduce malnutrition in a particular village. If villagers or researchers decide who gets to participate, there may be a selection bias. Many participants, for example, may be well nourished. Even if participants do better than non-participants researchers will not know if the program was successful.

Using a quasi-experimental method may help. Consider, for example, one such method -- *regression discontinuity design*. With regression discontinuity design, researchers use explicit selection criteria to select participants to receive aid. Aid might, for example, only be offered to people who are underweight for their age. The comparison group would then be made up of those who are just over the weight limit. Next, researchers see if there is a discontinuity in how people fare just above and just below the cutoff point for inclusion in the program.

Consider a graphical illustration of the results. The x's indicate those who start out underweight and so receive aid; the o's indicate those who start out over-weight. In the first graph aid has no effect. In the second aid has a good effect. On average, those receiving aid do better than those not receiving aid at the cut-off point (notice the

discontinuity).



Unfortunately, there can be differences between those just above and below the cutoff point and this can cause problems for this kind of quasi-experimental evaluation. Those just above the cutoff point might, for example, be able to participate in another aid program. If so, creating a comparison group made up of people right above the cut off point will introduce selection bias. Contrasting the treatment with the comparison group will not tell us whether the aid program is successful; we will not know how the malnourished would have fared without aid.

True experiments better prevent selection bias; randomization gives us reason to think the treatment group is relevantly similar to the control group. It will not matter if those above the weight limit can participate in another aid program. With proper randomization, the comparison group will be made up of people who, like those receiving aid, are below the limit. So, it should be easier to conclude that a perceived effect is due to aid.[xxxii]

Experimental evaluations can help us determine the efficacy not only of particular programs but of different aid processes and policies. Many evaluations of microfinance test processes like particular loan or savings products, for instance.[xxxiii] Other evaluations the efficacy of different policies regarding, for example, the distribution of things like bed nets.[xxxiv] And it is possible to use experimental micro-level studies to predict responses to macro-level policies (though larger samples are necessary to do this sort of research). For instance, researchers have recently evaluated how young women in Kenya respond to an AIDS awareness campaign that explains how older partners are more likely to have acquired HIV. They predict that the program will reduce HIV transmission rates amongst teenage girls in Kenya as it encourages young women to sleep with those in their age cohort.[xxxv]

It is important, however, to be careful not to generalize beyond what the data warrants. A good program can fail because the staff, beneficiaries, culture, or environment changes. Test projects can be more exciting and, so, more successful than normal programs just because they are test projects; this is called the Hawthorne effect. Fortunately, there are ways to deal with such problems. Increasing the time span and scale of the evaluation reduces the Hawthorne effect, for example. Testing programs in multiple locations at a large scale strengthens the case for replication or scaling up.

Often, however, experiments are hard to perform. Participants can drop out of treatment and those in the control group can sometime gain access to treatment illicitly. If, for instance, the program to combat malnutrition was offered at randomly selected schools, parents might move their children to schools offering the program (or away from such schools). This may bias results. Similarly, researchers may select a biased sample.

If, for instance, researchers randomly select people to participate in a study from a non-representative subgroup of a population (say, students participating in an after-school program), they may inadvertently end up with results that will not generalize well to the other segments of the population.

There are ways of dealing with some of these problems. Double-blinding is a traditional solution to minimizing selection problems. If neither researchers nor participants know who is receiving treatment that will prevent much intentional manipulation. But it may be hard to find an appropriate 'placebo' in the context of an aid program (parents and students usually know what kinds of programs are being carried out in local schools). Worse, inadvertent selection biases can be just as problematic as intentional selection effects. People may move out of a study area for reasons other than their desire to remain in or leave the program. Furthermore, the above example illustrates how researchers may select biased samples precisely because they do not recognize potential problems.

Yet more problematic is the fact that, because researchers are selecting only a small number of participants (and doing so from a small population), we can expect that even truly randomly selected groups will exhibit some potentially biasing regularities. One group may, for instance, include more boys than the other group. Theoretically, randomization guarantees that there are no differences between the treatment and control groups. But, this is only guaranteed if there is an infinite sample from which to choose. Obviously, in the real world, samples are not infinite. So, simply due to chance, the two groups may differ in significant ways.

To deal with potentially problematic differences between treatment and control groups due to chance, participants may be matched along potentially distorting dimensions (e.g. their sex) before being randomized. One person from each stratified matched pair would receive treatment. The other would become part of the control group. Alternately, there are ways of taking into account potentially relevant factors econometrically. In effect, however, using these methods or stratified matching, makes the resulting evidence quasi-experimental rather than fully experimental. (One must assume that there are no other factors that explain observed differences that have not been taken into account).[xxxvi] Nevertheless, randomization is an important tool (amongst others) that helps us take into consideration potential differences between these groups. So, this paper will consider experimental as well as quasi-experimental micro-level evidence regarding aid's efficacy below. This evidence will let us conclude that there are lots of good aid programs that can probably be successfully replicated and scaled up.

4. Making the Case for Some Foreign Aid

Both advocates and critics of aid are concerned about how aid impacts the poor. Few explicitly consider how we should measure aid's impact on the poor. So before arguing that there are lots of good aid programs, this paper will consider what a *good* program looks like.

There are many ways to measure poverty. We might, for instance, use an assortment of indicators such as education and caloric intake. Alternately, we might use a unitary measure of poverty. Unitary measures either specify a single formula for combining many disparate indicators of poverty or specify a single indicator (like income or Gross Domestic Product (GDP)).

In discussing the macro-level data, this paper suggested that, if we care about how aid impacts the poor, growth in aggregate GDP is not a good (unitary) indicator of aid's success. It cited studies that used a mixture of indicators. Better unitary measures are the Human Development Index (HDI) and the World Bank's poverty lines. The HDI combines (the logarithm of) GDP per capita at purchasing power parity (PPP), literacy and primary, secondary and tertiary enrollment rates, and life expectancy at birth into a single indicator.[xxxvii] The World Bank uses (PPP) income-based measures of poverty.[xxxviii]

The HDI includes more than income. This is a mark in its favor. Unfortunately, it has some problems. Poverty may be correlated with GDP per capita at PPP, literacy and primary, secondary and tertiary enrollment rates, and life expectancy at birth. But, it is not clear that the HDI is a better measure of poverty than the alternatives (education, income, and life expectancy may each be correlated with poverty in different ways and one of these may, alone, provide a better proxy for poverty than the combination).[xxxix] We need a philosophical account of poverty to decide which way of measuring poverty is best. But neither the United Nations Development Program nor Amaryta Sen who helped develop the measure has specified what set of basic functionings people need to be able to avoid poverty.[xl] The point is not just, as Sen recognized, that the HDI does not provide a complete measure and conceals important dimensions of poverty.[xli] The point is that some philosophical justification is necessary to establish that the HDI is even a reasonably good indicator.

Worse, the HDI is relies on average income levels, so it cannot tell us how many poor people there are within a country.[xlii] A country where half of the population is well off, and half very poorly off, may have the same HDI as a country where everyone is

doing equally, and moderately, well.[xliii] Germany, for instance, has the 14th most equal income distribution, while Hong Kong is 84th. [xliv] But Germany's HDI of .930 very close to Hong Kong's .916. [xlv] This last problem is so severe, it basically renders the HDI useless for our purposes.[xlvi]

Because the World Bank's poverty lines avoid this problem, we might use one of these indicators instead. But, the Bank's poverty lines share another problem with the HDI that stems from the PPP exchange rates they rely upon.

PPP exchange rates make incomes comparable between countries and individuals in different countries. Consider a simple example of how to calculate such an exchange rate between two countries for a single good. Suppose it costs one dollar to get a bag of corn in the US and two pesos to do so in Mexico. The PPP exchange rate for corn would be 2:1 (.5 pesos to a dollar or 2 dollars to a peso). PPP exchange rates are calculated for many countries to find relative prices for many of the goods and services that make up GDP. They usually express the result in terms of US$ equivalents.[xlvii]

There are several problems with relying on PPP exchange rates in measuring poverty. First, the data these measures rely upon is questionable. PPP measures are based on the Penn World Tables (PWT) and the International Comparison Project (ICP) surveys. These surveys are of variable quality and often measure different things (e.g. income vs. consumption) they must be adjusted significantly to make them consistent.[xlviii] Furthermore, these surveys do not have adequate coverage.[xlix] In 2005 China was included for the first time and India for the first time since 1985.[l] So the surveys may not provide good estimates of poor countries' incomes or the incomes of the poor within those countries.

Furthermore, the most common methods of comparing purchasing power make it seem like the poor are doing better than they are. The Geary-Khamis method, for instance, basically averages international price differentials for all commodities weighting 'each commodity in proportion to its share in international consumption expenditure.'[li] Essentially it considers how much it costs in each country to purchase the average 'basket' of goods consumed in all countries. The problem is that this 'basket' includes many services, for instance, that the poor do not buy but that are relatively cheap in developing countries.[lii] This makes it seem like the poor in these countries are doing better than they are. For, while food is relatively cheap in developing countries, it is not as cheap as PPP estimates suggest.[liii] Services have also come to make up a lager proportion of the international consumption 'basket' over time. This makes it seem that the poor are doing better simply because of a change in the rich's consumption patterns.

Fortunately, there may be some ways of ameliorating the problems with PPP measures. For calculating the World Bank's poverty lines, it might help to compare purchasing power over a representative basket of goods the poor consume, for instance.[liv]

Even if such changes were made, however, it is not clear that the World Bank's poverty lines would accurately capture the amount of poverty in a situation. For, more than money matters to the poor and different people need different amounts of money to avoid poverty. Subsistence farmers whose states provide health care and education, for instance, may not need much money. So if aid harms the poor financially but brings them greater benefits of a different type, it may decrease poverty. Alternately, aid may bring small financial benefits to the poor but harm the poor greatly in other ways and so

increase poverty. Income is just one indicator of poverty. It may not be the best indicator. Having more than one indicator may be best.

So, this paper will consider what micro-level studies tell us about how aid impacts individuals' access to things like education, health care, food, water, and shelter. It will assume that if aid helps people get these things, it reduces poverty.[lv] Of course, different indicators of poverty (like health and education levels) can exhibit opposite trajectories. Without a unitary measure, we may not be able to tell whether an aid program is making things better or worse overall. It may also be impossible to tell how much things are getting better or worse if, for instance, we find that an aid program helps some people get education but makes it less likely that others will get sufficient food. Sometimes, more than one thing may also be necessary to decrease poverty. Some medicines, for instance, may not work without food. Still, if a program helps people get basic water, food, shelter, health care, or education without making it harder for people to access these things that is evidence that the program reduces poverty. After all, helping people secure some of these things may help them secure others. If, for instance, children are well-nourished, they may learn more in school. At least I hope the advocates and critics of aid can agree on this much. On this assumption, this paper will look at programs funded by many sources governmental and non-governmental.

About ¼ of the world's population is infected with worms like schistosomiasis and hookworm.[lvi] These infections can cause protein deficiency, anemia, and other kinds of malnutrition, which can interfere with children's schooling. In parts of Kenya, more than 90% of school children are infected with worms.[lvii] Internationaal Chistelijk Steunfonds Africa, a Dutch non-governmental organization and the District Ministry of

Health started de-worming children in Kenyan schools. Researchers wanted to carry out an experimental evaluation of the program in Kenya's Busia school district. Unfortunately, administrators would not allow random numbers to determine which schools would get the medication. So the researchers used a heuristic for randomness based on the alphabet.[lviii] They found that the program decreased absenteeism by 25 percent or more. Because the program lowered illness transmission rates, absenteeism even decreased in nearby schools.[lix] The program was replicated in India. It was extended to provide iron supplements since lots of the Indian children were anemic. One year later, 'researchers found a nearly 50% reduction in moderate to severe anemia, large weight gains, and a 7% reduction in absenteeism among 4-6 year olds.'[lx]

Most of the world's poor live in rural areas and farming is essential to their survival. So, aid is often directed at increasing agricultural yields and, thus, food supplies. In the rural Busia District of Western Kenya a Dutch NGO International Child Support decided to start a seed and fertilizer program. Parents of students in the local school were randomly selected to participate in the program for six years. The farmers were given Calcium Ammonium Nitrate fertilizer to apply to a randomly selected plot of land and hybrid maize to plant along with Di-Ammonimum Phosphate fertilizer for another randomly selected plot. A comparison plot was randomly selected for traditional seed and fertilizer. The NGO also provided help in the first applications, tracked farmers' progress, and assisted with harvesting and weighing the crops. The quantities of fertilizer provided varied between growing seasons. Researchers found that the right amount of fertilizer and seeds could increase yields by over 90%.[lxi] Further, the evidence suggested that well-

timed price reductions induce some farmers to purchase and apply fertilizers themselves.[lxii]

Globally only five out of six boys and four out of five girls are in elementary school.[lxiii] The Inter-American Development Bank along with the Mexican government, created a conditional cash transfer project called Progresa/Oportunidades to help the poorest educate and get medical care for their children.[lxiv] Mothers of participating children had to attend nutrition and health programs (e.g. prenatal care, nutrition monitoring, well-baby care, immunization, supplementation, and preventive care programs).[lxv] Their children received scholarships to go to school -- with larger scholarships going to girls than boys. When the government wanted to expand the program half of the 506 eligible communities were randomly selected to participate. Comparing educational and health outcomes in these communities to those in the comparison group, outside experts at the University of California Berkeley showed that children who stayed in the program for two years were about 40% less likely to fall ill than children who did not participate.[lxvi] They were also about 25% less likely to be anemic, grew quicker, and returned to school more frequently after emergencies.[lxvii] On average, the percentage of children enrolled in first through eighth grades increased 3.4% in participating communities.[lxviii] The percentage of girls who completed sixth grade rose 14.8%.[lxix] Even adults benefited. Parents took 19% fewer sick days, on average.[lxx] Progresa/Oportunidades was so successful that it was extended to urban schools throughout Mexico. Similar evaluations show that conditional cash transfer programs are successful around the world.[lxxi]

In India only 35% of females and 61% of Males were literate in 1990.[lxxii] In 1994, Pratham, an Indian non-governmental organization supported by several aid agencies (including OXFAM, NOVIB, and AIF), began a remedial education program. Pratham hired local women to tutor children who were doing poorly in elementary school. An experimental evaluation of the program randomly assigned tutors to half of the 98 eligible schools Vandorara city (a similar experiment was performed in the L-ward of the Mumbai schools). The evaluation showed that the program 'increased student test scores by .39 standard deviations.'[lxxiii] Moreover, the improvements were 'largest for children at the bottom of the distribution.'[lxxiv] The program resulted in gains, per dollar, ten times greater than hiring new teachers and its returns improved over time.[lxxv] Pratham now runs similar programs in twenty cities for 161,000 children.[lxxvi]

There are many other successful programs that help people -- from microfinance to school voucher programs.[lxxvii] Many of these programs have been successfully replicated and scaled up.[lxxviii] We can do *some* good with foreign aid.

5. Conclusion

In arguing that we are obligated to aid the poor, Peter Singer assumes that we can use foreign aid to successfully ameliorate poverty.[lxxix] Recently, several philosophers have questioned this assumption. Many of aid's critics fail to provide the requisite kind of data to make their case.[lxxx] So, this paper has canvassed the strengths and weaknesses of different kinds of empirical evidence for establishing different conclusions about aid's efficacy. In doing so, it addressed the more general methodological question: What types of data are necessary for establishing normative conclusions in debates in political philosophy and public policy? It defended one answer to this question. It suggested that

macro-level empirical data is necessary to establish general claims about the efficacy of different policies, while micro-level experimental data is generally best for establishing claims about the efficacy of particular programs. It, thus, provides some important guidance for resolving debates about everything from free trade and immigration to taxation and population control policy.[lxxxi]

With regard to the debate about foreign aid, in particular, the paper argued that macro-level data is necessary to make a strong case regarding aid impact on the poor in general. Unfortunately, most macro-level studies of foreign aid do not address the question of whether aid generally reduces poverty. Nevertheless, this paper canvassed some relevant macro-level studies suggesting that certain kinds of aid reduce poverty. Furthermore, it argued that there are many good micro-level studies suggesting that at least some foreign aid programs are successful and can be replicated and scaled up.

There are many questions for further research. It would be great to know, for instance, what factors contribute to successful aid programs, policies, and processes and how aid is distorted differently by different factors.[lxxxii] The practical importance of that knowledge may, however, be limited. For, this paper has done enough to establish Singer's implicit empirical premise that aid can ameliorate poverty. So, if the rest of his argument is right, we can reach the important practical conclusion that we are obligated to provide aid that *does* work. Even if aid is generally a bad idea, we should not neglect the good we can do for some, even if we cannot completely ameliorate poverty. We need not only see the forest for the trees, we must not neglect the water for the seas.

i The author would like to thank Thomas Christiano, Gillian Brock, Michael Gill, Jerry Gaus, Allen Buchanan, Mathias Risse, Leif Wenar, Avram Hiller, Dale Jamieson, Monica Hlavac, Robert Keohane, Michael Woolcock, Clark Glymour, Alex London, Jay Kadane, Rick Furtak, Ben Fraser, Adam Cureton, Bill Oberdick, Chris Brown, Jason Matteson, Nathan Lubchenco, S. Submaranian as well as audiences at Duke University, the University of Washington, the University of Arizona, and Carnegie Mellon University for helpful comments, suggestions, and encouragement. She apologizes to anyone she has so carelessly forgotten to mention.

ii Peter Singer, (1972) 'Famine Affluence and Morality', *Philosophy and Public Affairs* 1.

iii There are, of course, other philosophers who advocate aid, but this paper is primarily concerned only with the debate regarding the empirical premise in Singer's argument. See, however: Thomas Pogge, (2002) *World Poverty and Human Rights: Cosmopolitan Responsibilities and Reforms.* Cambridge, Polity Press. William Aiken and Hugh LaFollette (eds.). 1977. *World Hunger and Moral Obligation*. New Jersey, Prentice Hall. Gillian Brock and Darrel Moellendorf, (2005) *Current Debates in Global Justice. Studies in Global Justice,* Vol. 2 (Norwell, MA: Springer). Deen K. Chatterjee (ed.), (2004) *The Ethics of Assistance: Morality and the Distant Needy,* (Cambridge: Cambridge University Press). Also see: Nicole Hassoun, (2008a) 'World Poverty and Individual Freedom*,' American Philosophical Quarterly*, 45, 2: 191-198.

iv David Schmidtz, for instance, says aid in Africa is often counter-productive and implies that we would do better just to try not to harm people. Dale Jamieson suggests making it easier for poor people to meet their needs by reducing trade barriers against developing countries instead of giving aid. Andrew Kuper argues that it may be better to purchase goods the poor make and vacation in poor countries. See: David Schmidtz, (2000) 'Islands in a Sea of Obligations: Limits on the Duty to Rescue', Working Paper 18 (Oakland: The Independent Institute). This is the open access version of the published article: David Schmidtz (2000) 'Islands in a Sea of Obligation: Limits of the Duty to Rescue.' *Law and Philosophy,* 19, No 6: 2. Dale Jamieson (2005) 'Duties to the Distant: Aid, Assistance, and Intervention in the Developing World', *The Journal of Ethics* 9. Andrew Kuper (2002) 'More Than Charity: Cosmopolitan Alternatives to the 'Singer Solution'', *Ethics & International Affairs* 16.

v If this paper is correct, however, it should be clear to those familiar with the critics' arguments, that many of them rely in part on anecdotes and/or statistical evidence that does not support the requisite causal inferences. Consider, for instance, one of the anecdotes Schmidtz provides in questioning the empirical premise of Singer's argument. Schmidtz says 'When I recently crossed the border from Zimbabwe into Zambia, a large sign warned that bringing second-hand clothing into Zimbabwe from Zambia is prohibited. Puzzled, I wondered whether second-hand clothing might carry some disease. When I passed through the town of Livingstone, just north of the border, I asked what the sign was about, and I was told by three different sources (two white men, one black woman, all local residents) that Livingstone had until recently been the hub of Zambia's textile industry. Cotton was grown, processed, and woven into cloth there. However, a few years ago, in the wake of a severe and highly publicized drought, international relief agencies decided that what Zambia needed was planeloads of second-hand clothing. Livingstone manufacturers could not compete with free clothing, though. Today, the unemployment rate in Livingstone is ninety percent.' Schmidtz, 'Islands in a Sea of Obligation: Limits of the Duty to Rescue.' p. 2. The story Schmidtz relates is probably true, in fact this kind of story has probably been repeated in many places around the world. But, this paper will argue that even this fact would do little to support undercut Singer's argument as this kind of evidence cannot establish the conclusion that aid generally does more harm than good. Similarly, Kuper defends his claim, that we should embrace a structural approach to dealing with poverty instead of giving foreign aid, with anecdotes (implying that aid is generally harmful or at least that we cannot know that aid will do any good). He says: 'Consider, most starkly, the perpetuation and intensification of the Rwandan conflict and the human misery aggravated by aid agencies that sustained refugee camps. In spite of the camps becoming bases for militiamen and incubators for cholera, the prospect of international NGO aid encouraged people not to return to their homes even when it was safer to do so, thus intensifying and prolonging the conflict. Consider also the 'food relief' of the 1970s that so damaged the situation of developing world farmers and their dependents.' Andrew Kuper (2002) 'More Than Charity: Cosmopolitan Alternatives to the "Singer Solution."' *Ethics & International Affairs,* 16: 113-114. Kuper suggests that the ways our actions impact the poor are complex and multifaceted and that the best way to relieve poverty is not obvious. He's probably right about that, but this paper will argue that the anecdotal evidence Kuper provides in questioning Singer's empirical premise cannot establish this. It will argue that statistical data is necessary to make a general case for or against aid. That is not to say that all

statistical evidence is equally good. For the paper's argument should also make it clear that sometimes Singer's critics also cite inappropriate statistical data for making their points. In support of the claim that development assistance is *usually* ineffective, for instance, Jamieson says '[T]here is little empirical evidence that [development assistance]… has substantially improved the welfare of the poor. A recent report from the Commonwealth Secretariat claims that although more than $1.2 trillion was spent on official development assistance between 1950 and 2000, the gap between the incomes of people in developed and developing countries has widened.' Jamieson, 'Duties to the Distant: Aid, Assistance, and Intervention in the Developing World.' p. 160. And he cites similar data from the UN. But inequality may have widened for many reasons and inequality can get worse even as poverty declines if the rich gain more than the poor gain. This paper will argue that the kind of statistical evidence Jamieson provides (even in conjunction with the examples he gives) provides little support the conclusion that any kind of aid generally does more harm than good.

vi Gillian Brock and Darrel Moellendorf, (2005) *Current Debates in Global Justice. Studies in Global Justice,* Vol. 2 (Norwell, MA: Springer).

vii Even if they succeed in doing this much, however, they must say more to show that we should not still implement those aid programs that do more harm than good.

viii Garrett Cullity (2004) *The Moral Demands of Affluence*, (Oxford, Clarendon Press). Though Jamieson points out that Singer also makes general claims about aid's efficacy he says 'that providing development assistance is… "usually the better long-term investment."' Jamieson, 'Duties to the Distant: Aid, Assistance, and Intervention in the Developing World.' p. 158. This claim likewise requires more evidence. From: Peter Singer, ''Famine Affluence and Morality,'' in W. Aiken and H. La Follette (eds.) World Hunger and Moral Obligation, First Edition (Englewood Cliffs: Prentice-Hall, 1977), p. 35. Leif Wenar (2007a) 'The Basic Structure as Object: Institutions and Humanitarian Concern', *Canadian Journal of Philosophy* sup. 31.

ix Leif Wenar, 'The Basic Structure as Object: Institutions and Humanitarian Concern'.

x Gillian Brock (2009) *Global Justice: A Cosmopolitan Account*. (Oxford: Oxford University Press). Thomas Pogge (2002) *World Poverty and Human Rights* (Cambridge: Polity Press). Nicole Hassoun, (Forthcoming a) 'Free Trade, Poverty, and Inequality.' *The Journal of Moral Philosophy*. Mathias Risse (2008) 'On the Morality of Immigration,' *Ethics and International Affairs* 22, 1: 25-33.

xi This paper does not tell us what to do in the absence of good evidence. But, even if one could show that foreign aid does not work, we might have other obligations to the poor. Nicole Hassoun, (2008a) 'World Poverty and Individual Freedom*,' American Philosophical Quarterly*, 45, 2: 191-198. Nicole Hassoun, (2008b) 'Free Trade, Poverty, and the Environment,' *Public Affairs Quarterly,* 22, 4: 353-380. Nicole Hassoun, (2009a) 'Freedom and Autonomy,' Carnegie Mellon University Working Paper. Available at: http://www.hss.cmu.edu/philosophy/hassoun/papers.php. For discussion of how we might go about fulfilling these obligations in theory see: Nicole Hassoun, (2009b) 'Meeting Need,' (2009c), *Utilitas*, 21, 3.

xii Leif Wenar is one of the few philosophers who has looked at the empirical data on the impact of foreign aid. Wenar concludes that we do not have the requisite data to come to firm conclusions about these matters. While I agree that the macro evidence is certainly not conclusive, we can learn a lot by considering it.

xiii Clemens et. al. argue that aid can increase growth. Clemens, M., S. Radelet, et al. (2004). 'Counting Chickens when they Hatch: The Short Term Effect of Aid on Growth.' Working Paper Number 44 (Washington D.C, Center for Global Development). Mosely et. al. failed to find a positive relationship between aid and growth. Paul Mosley, John Hudson, and Sara Horrell (1987) 'Aid, the Public Sector and the Market in Less Developed Countries', *The Economic Journal* 97. Collier and Dollar argue that aid increases growth in 'good' policy environments (e.g. when countries have a fiscal surplus, low inflation, and are open to trade). Burnside and Dollar 'Aid, Policies and Growth'. Paul Collier and David Dollar (2004) 'Development Effectiveness: What Have We Learnt?' *The Economic Journal* 114. Hansen et. al. counter that, if aid has diminishing returns, these results may not hold. Henrik Hansen and Finn Harp (2001) 'Aid and Growth Regressions', *Journal of Development Economics* 64. Easterly worries that the definitions of 'aid,' 'good policy,' and 'growth' have not been sufficiently well justified. William Easterly (2003) 'Can Foreign Aid Buy Growth?' *Journal of Economic Perspectives* 17. Dalgard et. al. argue that climatic conditions actually explain the relationship between aid and growth better than policy environment. Carl-Johan Dalgaard, Henrik Hansen, and Finn Tarp (2004) 'On the Empirics of Foreign Aid and Growth', *The Economic Journal* 114.

xiv Problems with defining foreign aid infect the macro-level data. Many researchers use an overly expansive definition of 'foreign aid' that captures almost everything besides normal business transactions and loans. See, for instance: Clemens, M., S. Radelet, et al. (2004). 'Counting Chickens when they Hatch: The Short Term Effect of Aid on Growth.' Working Paper Number 44 (Washington D.C, Center for Global Development). Even Official Development Assistance includes things many probably do not think of as aid like concessionary loans. Different studies also look at aid's impact on different populations etc. Some studies look at 'program' vs. 'project' aid. George Mavrotas (2003) 'Assessing Aid Effectiveness in Uganda: An Aid-disaggregation Approach,' processed (Oxford: Oxford Policy Management). Tito Cordella & Giovanni Dell-Ariccia (2003) 'Budget Support versus Project Aid,' International Monetary Fund Working Paper 03/88 (Washington D.C.: International Monetary Fund). Others consider grants and official development assistance excluding things like technical cooperation and food aid. Karuna Gomanee, Sourafel Girma, and Oliver Morrissey (2002) 'Aid and Growth in Sub-Saharan Africa: Accounting for Transmission Mechanisms' CREDIT Research Paper No. 02/05, Center for Research in Economic Development and International Trade, University of Nottingham. Some look at development vs. humanitarian aid. Trudy Owens & John Hoddinott (1999) 'Investing in Development or Investing in Relief: Quantifying the Poverty Tradeoffs Using Zimbabwe Household Panel Data,' Working Paper 99-4, Centre for the Study of African Economies, Department fo Economics, Oxford University. Others create their own categorization e.g. 'aid that could plausibly stimulate growth in four years.' Clemens, M., S. Radelet, et al. (2004). 'Counting Chickens when they Hatch: The Short Term Effect of Aid on Growth.' Working Paper Number 44 (Washington D.C, Center for Global Development). It is not clear that one can conclude that foreign aid increases or decreases growth based on a survey of studies looking at all of these different things. Furthermore, it is important to pay attention to the specific data sources each study uses as even those who say they are talking about the same kind of aid (e.g. program aid) may mean different things by these terms.

xv C. Michalopoulos & V. Sukhatme (1989) 'The Impact of Development Assistance: A Review of the Quantitative Evidence' in A. O. Krueger ed. *Aid and Development*, Chapter 7: 111-124. (Baltimore: John Hopkins University Press). H. White (1992). 'The Macroeconomic Impact of Development Aid: A Critical Survey.' *Journal of Development Studies* 28, 2: 163-240.

xvi H. Hansen & F. Tarp (1999) 'The Effectiveness of Foreign Aid.' Development Economics Research Group Mimeo, University of Copenhagen: 1.

xvii Different authors give different accounts of when aid is morally permissible or required. For views that contrast with Singer's see, for instance: David Reiff, (2002) *A Bed for the Night: Humanitarianism in Crisis.* (New York: Simon & Schuster). Jamieson 'Duties to the Distant: Aid, Assistance, and Intervention in the Developing World'.

xviii Arne Bigsten and Jorgen Levin, (2004) 'Growth, Income Distribution, and Poverty: A Review', in *Growth, Inequality, and Poverty,* ed. Anthony Shorrocks and Rolph van der Hoeven (Oxford: Oxford University Press), Raghuram G. Rajan and Arvind Subramanian, (2005) 'Aid and Growth: What Does the Cross-Country Evidence Really Show?' International Monetary Fund Working Paper 127 (Washington, D.C., International Monetary Fund), Rasmus Heltberg, (2004) 'The Growth Elasticity of Poverty', in *Growth, Inequality, and Poverty,* ed. Anthony Shorrocks and Rolph van der Hoeven (Oxford: Oxford University Press). Henrik Hansen and Finn Tarp, (1999) 'Aid Effectiveness Disputed' CREDIT Research Papers. University of Nottingham: Nottingham, UK, World Bank, (2002) *World Development Indicators 2002*, (Washington D.C., World Bank).

xix Furthermore, even the available macro-economic studies of aid's impact on the poor have significant problems. Some of these problems stem from the fact that the most common measures of poverty are inadequate. Sanjay Reddy and Thomas Pogge, (2006) 'How Not to Count the Poor', in *Measuring Global Poverty*, ed. Sudhir Anand and Joseph Stiglitz (Oxford: Oxford University Press). The World Bank's method of measuring poverty changed in the 1990's and different poverty lines are incommensurable. The survey data underlying the purchasing power parity (PPP) exchange rates that are used in calculating poverty internationally is not representative. And, the PPP exchange rates may over-estimate the purchasing power of the poor. Nicole Hassoun, (Forthcoming a) "Free Trade, Poverty, and Inequality*," The Journal of Moral Philosophy*. Nicole Hassoun, (Forthcoming b) "Global Poverty" *The Encyclopedia of Global Justice*, Deen Chaterjee ed. Oxford University Press, Oxford.

xx This should become obvious below, but this paper will say more about different measures of poverty in subsequent sections.

xxi See: Peter Boone (1996) 'Politics and the Effectiveness of Foreign Aid', *European Economic Review* 40. Statistics on life expectancy and infant mortality rates, like GDP, are aggregate statistics.

xxii See: Paul Mosley, John Hudson, and Arjan Verschoor (2004) 'Aid, Poverty Reduction and the 'New Conditionality", *The Economic Journal* 114. Mosely et. al. construct a PPE index for how spending impacts the poor.

xxiii Because both Boone and Mosely et. al.'s analyses are complicated this is all a bit imprecise. Boone, for instance, excludes Israel and those countries with high aid to GNP ratios from his analysis because he is interested in testing the impact of political structure on aid's distribution. Readers interested in the full story should refer to the relevant articles. There are also some worries about interaction variables that might apply to Boone's analysis (though I know of no papers analyzing Boone's paper's use of such variables). See: Catherine Pattillo, Jacques Polak, and Joydeep Roy, (2007) 'Measuring the Effect of Foreign Aid on Growth and Poverty Reduction or The Pitfalls of Interaction Variables.' IMF Working Paper. International Monetary Fund: Washington D.C.

xxiv Nadia Masud and Boriana Yontcheva, (2005) 'Does Foreign Aid Reduce Poverty? Empirical Evidence from Nongovernmental and Bilateral Aid,' International Monetary Fund Working Paper (Washington D.C.: International Monetary Fund).

xxv Karuna Gomanee and Oliver Morrissey, (2002) 'Evaluating Aid Effectiveness Against a Poverty Reduction Criterion,' DESG Conference, Nottingham, April 2002.

xxvi Leif Wenar, 'The Basic Structure as Object: Institutions and Humanitarian Concern'.

xxvii I would like to thank Clark Glymour for the analogy.

xxviii This is the point made in: Jeffrey Sachs. 2005. *The End of Poverty: How We Can Make it Happen in Our Lifetime*. London, Penguin Press. For discussion that seems to miss this point see: William Easterly. 2006. *The White Man's Burden: Why the West's Efforts to Aid the Rest Have Done so Much Ill and so Little Good.* New York, Penguin Press. For equally famous but not particularly sophisticated work by social scientists on aid see: Peter Bauer. 1972. *Dissent on Development,* (Cambridge, Harvard University Press). Peter Bauer. 1975. "N. H. Stern on Substance and Method in Development Economics," *Journal of Development Economics*, 2: 387-405. Easterly, William. 2003. "Can Foreign Aid Buy Growth?" *Journal of Economic Perspectives*, 17: 23-48. Teresa Hayter. 1971. *Aid as Imperialism.* (London, Pelican Books). Theresa Hayter and Catherine Watson. 1985. *Aid: Rhetoric or Reality?* (London, Pluto Press). For more recent work, see: Richard Black and Howard White (eds.). 2004. *Critical Perspectives on the Millennium Development Goals*, (London: Routledge).

xxix For an interesting simulation showing that, theoretically, redistribution can reduce poverty see: Hulya Dagdeviren, Rolph van der Hoeven, and John Weeks, (2004) 'Redistribution Does Matter: Growth and Redistribution for Poverty Reduction', in *Growth, Inequality, and Poverty*, ed. Anthony Shorrocks and Rolph van der Hoeven (Oxford: Oxford University Press).

xxx I know of no analogue to a placebo in tests of foreign aid, though it might be possible to develop an analogue in some cases.

xxxi Sometimes randomization is not the best way to achieve independence between the criteria for allocation to treatment and control groups and the hypothesized effect of aid. This paper cannot consider the conditions under which an alternative is better here, however.

xxxii Researchers have developed ways of addressing this possibility.

xxxiii Dean Karlan and Nathanel Goldberg, (2007) 'Impact Evaluation for Microfinance: Review of Methodological Issues,' Innovations for Poverty Action Working Paper. Yale University: New Haven.

xxxiv Jessica Cohen and Pascaline Dupas, (2007), 'Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment.' Brooking Global Economy and Development Working Paper 11. For criticism that suggests the need for replication in multiple locations see: Roderick, 'The New Development Economics: We Shall Experiment, But How Shall We Learn?'

xxxv Pascaline Dupas, (2006) 'Relative Risks and the Market for Sex: Teenagers, Sugar Daddies and HIV in Kenya,' Dartmouth College Working Paper. Dartmouth College: Hanover, New Hampshire.

xxxvi A final point worth mentioning is that to generalize to a larger population (to show that experimental results are significant), researchers must make some assumptions about the way the treatment effects the group that are not experimentally verifiable. And, while similar assumptions are necessary for non- and quasi-experimental designs, which requires the most faith depends on the particular design. For more on this point and the importance of theory to development economics see: Angus Deaton, (2009) 'Instruments for Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development,' The Keynes Lecture, British Academy, October 9th, 2008. For a thoughtful consideration of criticism of experimental methodology see:

Abhijit Banerjee and Ester Duflo, (2009) 'The Experimental Approach to Development Economics,' Abdul Laff Jameel Poverty Action Lab Working Paper. Massachusetts Institute of Technology: Cambridge.

xxxvii This section draws on Nicole Hassoun, (Forthcoming a) "Free Trade, Poverty, and Inequality*," The Journal of Moral Philosophy*.

xxxviii There is also a Human Poverty Index (HPI) but I know of no empirical work measuring the general impact of aid on the HPI perhaps because it is a relatively new index. Nicole Hassoun, (Forthcoming a) "Free Trade, Poverty, and Inequality*," The Journal of Moral Philosophy*.

xxxix Income is discounted at all levels but at an increasing rate. Rich countries appear less developed than they would if this scaling was not done. As the average income level rises, poverty appears to be less and less affected by increases in average income. But, because the HDI does not take into account distribution within countries, it is not clear that the HDI can tell us whether or not this is really the case in any given country. Nicole Hassoun, (Forthcoming a) "Free Trade, Poverty, and Inequality*," The Journal of Moral Philosophy*.

xl Although Sen has written a lot about capabilities and uses examples throughout his work, he refuses to provide a comprehensive list of basic capabilities.

xli One might also worry about whether the weight given to each of the HDI's components is appropriate (implicit weights are given to components when they are normalized though each is supposed to account for 1/3 of the indicator). See: Sen. 2006. 'Human Development Index', *in The Elgar Companion to Development Studies*, ed. D.A. Clark, (Edward Elgar: Cheltenham).

xlii Although, as Sen points out, there is some attempt to account for distribution in giving income declining marginal utility, the examples below indicate that the weighting is inadequate. See: Ibid.

xliii Due to changes in the methodology, HDI figures cannot be compared between years – though this is also the case for the World Bank's poverty lines.

xliv Ibid

xlv United Nations Development Program, (2005) *Human Development Report,* (New York: Oxford University Press).

xlvi For discussion of other problems with the Human Development Index see Kate Raworth and David Stewart, (2003) 'Critiques of the Human Development Index', in *Readings in Human Development: Concepts, Measures and Policies for a Development Paradigm*, ed. Sakiko Fukuda Parr and A.K. Shiva Kumar (New Delhi: Oxford University Press).

xlvii International Comparison Program, (2008) "Part I: Purchasing Power Parities and 2005 ICP Results." World Bank: Washington D.C. Available at: < http://siteresources.worldbank.org/ICPINT/Resources/icp-final.pdf>.

xlviii Angus Deaton, (2000) "Counting The World's Poor: Problems And Possible Solutions," Research Program in Development Studies Working Paper. Princeton University: Princeton.

xlix Robert Wade, (2004) 'Is Globalization Reducing Poverty and Inequality?" *World Development* 32: 572.

l International Comparison Program, (2008) "Global Purchasing Power Parities and Real Expenditures: 2005 International Comparison Program."World Bank: Washington D.C. Available at: <http://siteresources.worldbank.org/ICPINT/Resources/icp-final.pdf>. Urban regions were over-represented in both countries' surveys.

li Sanjay Reddy and Thomas Pogge, (2006) "How Not to Count the Poor," in S. Anand and J. Stiglitz eds. *Measuring Global Poverty* Oxford University Press: Oxford: 1. Available at: <www.socialanalysis.org>.

lii International Comparison Program, (2008) "Part I: Purchasing Power Parities and 2005 ICP Results." World Bank: Washington D.C. Available at: < http://siteresources.worldbank.org/ICPINT/Resources/icp-final.pdf>.

liii Looking at some of the poorest countries included in the 1985 ICP survey, for instance, the prices for basic food stuffs "Breads and Cereals" averaged 111 percent higher than consumer prices generally. Ibid.

liv Ibid.

lv There are many definitions of poverty and it clearly not the case that every program that helps people attain these things will ameliorate poverty (on any reasonable definition). However, this supposition will do for the present purpose.

lvi Michael Kremer, (2005) 'The Role of Randomized Evaluations in Making Progress Towards Universal Basic and Secondary Education', American Academy of Arts and Sciences Working Paper (Cambridge: American Academy of Arts and Sciences).

lvii Ibid

lviii Administrators might have used the same heuristic to determine which schools received other sorts of benefits biasing the results. Still, the quasi-experimental design provided some evidence in favor of the program. Deaton, 'Instruments for Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development.'

lix All programs I discuss were funded at least in part by either private or public foreign aid. I do not distinguish private vs. public funded programs here although one might fruitfully compare the efficacy of programs distinguished in this way. (Miguel & Kremer 2003) cited in: Esther Duflo and Michael Kremer, (2003) 'Use of Randomization in the Evaluation of Development Effectiveness', in World Bank Operations Evaluation Department Conference on Evaluation and Development Effectiveness (Washington D.C., World Bank).

lx (Bobonis, Miguel, & Sharma 2002) cited in: Ibid.

lxi Ester Duflo, Michael Kremer, and Jonathan Robinson, (2009), 'How High are Rates of Return to Ferlizer? Evence from Field Experiments in Kenya,' MIT Working Paper.

lxii Ester Duflo, Michael Kremer, and Jonathan Robinson, (2009), 'Nudging Farmers to Use Fertilizer: Evidence from Kenya,' MIT Working Paper.

lxiii United Nations, (2006) 'The Millennium Development Goals Report', (New York: United Nations).

lxiv Emmanuel Skoufias et al., (2001) 'Is Progresa Working? Summary of the Results of an Evaluation by IFPRI,' FCND Discussion Paper No. 118 (Washington D.C., The International Food Policy Research Institute).

lxv Ibid

lxvi This program was funded in part by the International American Development Bank. Similar programs include Argentina's Plan Familias, Brazil's Bolsa Família, Chile's Chile Solidario, Colombia's Familias en Acción, Costa Rica's Superémonos, the Dominican Republic's Solidaridad, Ecuador's Bono de Desarrollo Humano, El Salvador's Red Solidaria, Honduras's RAF, Jamaica's PATH, Nicaragua's Red de Protección Social and Peru's Juntos. Inter-American Development Bank Press Release, (2006) 'Programs of Conditional Cash Transfers Seen as Effective in Reducing Poverty and Exclusion'. (Washington D.C., Inter-American Development Bank).

lxvii Ibid

lxviii (Shultz 2001) cited in: Duflo and Kremer, 'Use of Randomization in the Evaluation of Development Effectiveness'.

lxix Ibid

lxx (Gertler et. al. 2005) cited in: Duflo and Kremer, 'Use of Randomization in the Evaluation of Development Effectiveness'.

lxxi World Bank (2004a) *Assessing Aid: What Works, What Doesn't and Why*, (Oxford, Oxford University Press), World Bank, (2004b) 'The Impact of Conditional Cash Transfer Programs: A Review of Evaluation Results', Second International Workshop on Conditional Cash Transfer. Panel 3. Workshop Proceedings (São Paulo, World Bank), Laura Rawlings and Gloria Rubio (2005) 'Evaluating the Impact of Conditional Cash Transfer Programs', *The World Bank Observer* 20.

lxxii United Nations Educational Scientific and Cultural Organization, (2006) Education in India, (New York: UNESCO Institute for Statistics).

lxxiii The authors of the Pratham study probably generalize beyond what their data justifies; they suggest that the Pratham program can be replicated widely in India. (Banerjee, Cole, Duflo & Linden 2003) cited in: Duflo and Kremer, 'Use of Randomization in the Evaluation of Development Effectiveness'. We can conclude, however, that the program can be replicated or scaled up within the city of Vandorara and the L-ward of the Mumbai schools.

lxxiv "Children in the bottom third gained 0.6 standard deviations after two years." Ibid

lxxv Ibid

lxxvi (Banerjee, Cole, Duflo & Linden 2003) cited in: Duflo and Kremer, 'Use of Randomization in the Evaluation of Development Effectiveness'.

lxxvii Nava Ashraf, Dean Karlan, and Wesley Yin, (2006) 'Female Empowerment: Impact of a Commitment Savings Product in the Philippines,' Center for Global Development Working Paper No. 106, Robert Cassen, (1986) *Does Aid Work?* (Oxford, Clarendon Press), Jonathan Isbam, Deepa Narayan, and Lant Pritchett, (1994) 'Does Participation Improve Project Performance: Project

Performance in the Rural Water Supply Projects Establishing Causality', Policy Research Working Paper 1357 (Washington D.C., World Bank), World Bank, World Development Indicators 2002, Al Kehler, (2004) 'Humanitarian Exchange', Humanitarian Policy Group No. 27 (London: Overseas Development Institute), Oxfam, (2005a) 'Food Aid or Hidden Dumping? Separating Wheat from Chaff', Oxfam Briefing Paper 71 (Oxford: Oxfam GB), Esther Duflo, Rachel Glennerster, and Michael Kremer, (2007) 'Using Randomization in Development Economics Research: A Toolkit', Center for Economic Policy Research Working Paper 6059 (Washington D.C., Center for Economic Policy Research).

lxxviii Duflo and Kremer, 'Use of Randomization in the Evaluation of Development Effectiveness', Hassan Zaman, (1997) 'Poverty and BRAC's Microcredit Programme: Exploring Some Linkages', BRAC Working Paper Number 18 (Bangladesh: BRAC), Jonathan Morduch, (1998) 'Does Microfinance Really Help the Poor? New Evidence from Flagship Programs in Bangladesh', Harvard University Department of Economics and HIID Working Paper (Boston, MA: Harvard University), Mark Pitt, (1999) 'Reply to Jonathan Morduch's 'Does Microfinance Really Help the Poor? New Evidence from Flagship Programs in Bangladesh'', Brown University Department of Economics Working Paper (Providence, RI: Brown University), BRAC, (2005) 'Brac at a Glance: Alleviation of Poverty and Empowerment of the Poor' (Bangladesh, BRAC), Center for Global Development, (2005) 'Making It Pay to Stay in School', Center for Global Development Notes (Washington D.C., Center for Global Development).

lxxix Singer, 'Famine, Affluence, and Morality'.

lxxx Kuper 'More Than Charity: Cosmopolitan Alternatives to the 'Singer Solution'', Schmidtz, 'Islands in a Sea of Obligations: Limits on the Duty to Rescue', Jamieson 'Duties to the Distant: Aid, Assistance, and Intervention in the Developing World'.

lxxxi Gillian Brock. 2009. *Cosmopolitan Justice and Patriotic Concern.* Oxford: Oxford University Press. Thomas Pogge. 2002. *World Poverty and Human Rights: Cosmopolitan Responsibilities and Reforms*. Polity Press: Cambridge. Nicole Hassoun, (Forthcoming a). "Free Trade, Poverty, and Inequality." *Journal of Moral Philosophy*. Mathias Risse and Michael Blake. 2009. "Immigration and Original Ownership of the Earth" (co-authored with Michael Blake), *Notre Dame Journal of Law, Ethics, and Public Policy.* Vol. 23, No, 1 (special issue on immigration): pp 133-167.

lxxxii This paper gives us reason to be skeptical of many other papers' claims to have isolated the factors underlying good programs. Recall that the majority of macro-level evidence looks at aid's impact on growth. See ft. nt. xv.