World Bank Research Evaluation
Esther Duflo

Overall assessment

1. **In your area of expertise, has the Bank made a significant contribution?**

In the research that I was asked to evaluate, the most significant contribution, at least in terms of visibility, was made by the research on counting the poor in the world. While it remains controversial whether counting the poor is a necessary exercise, this research has had an important impact and is highly visible. The WDR 2004 on social service delivery, which was partly linked to the research I evaluated, also prompted a large amount of interesting research on various ways to improve social sector delivery, and contributed to shift the question from "how much money is needed" to "how should the money be delivered."

2. **In your area of expertise, has the Bank's research focused on the most important issues for developing countries? Do you feel that the Bank's researchers have appropriate incentives to allow identification of the most important issues?**

The research I was asked to evaluate is arguably all on important issues for developing countries (culture, social capital, poverty).

In practice, because some of these questions are difficult to address, the actual papers some times feel like they have only little relevance to how the Bank or the countries can address the problems of poverty. In particular I am not sure of the relevance of a significant part of the "culture" agenda for the World Bank's anti-poverty strategy. Some of the research has on the contrary focused on the most important issues (developing tools to measure poverty, and implementing them).

Generally, the Bank researcher's incentives may be a little too skewed toward (1) publishing in academic journals, and therefore producing the research that the researchers think will appeal to those journals –which leads them to focus on tools and models and not spend time on assembling and documenting useful data or (2) producing background material for the WDR.

The problem with (1) may be with the journals, more than with the Bank, but it sometimes leads to papers that are more often "stand alone," "cute papers." I should note that these problems may be stronger for the "second tier" academic journals than for the top field journals. All the papers I read that appeared in a top field journal had some interesting point to make. I think that World Bank researchers should be encouraged to do research that has more public good value (e.g. producing data sets that other people can use, or producing good quality, reliable impact evaluation even if it does not entail any methodological innovation) or that is of higher quality (e.g. can be published at least in the top field journals). Papers prepared in the second category (background for WDR) some times lead to great outcomes (like the data set on the 1 dollar a day poor created for the WDR 2000), but also sometimes leads to replace scholarship by models that can be explained simply (preferably in the form of flow charts), and by simple cross-country regressions that are more misleading than helpful.

3. In your area of expertise, has Bank research reflected awareness of substantive knowledge from other research available and in progress and sufficient knowledge of subject countries?

Some of the papers that I was asked to evaluate seem to be a little trapped in their own world. Cross-references to a very small number of authors in their own group are very frequent. Some of the more "theoretical" or "methodological papers" end up reinventing the wheel, sometimes introducing mistakes in the original arguments. When these papers are cited by their colleagues, the initial message sometimes gets really blurred. I don't think it necessarily reflects a lack of knowledge of the outside research, but more likely a tendency to be inward looking. But then, maybe this is a general bias and is not specific to the Bank. Many of the papers display good knowledge of the countries.

4. Has the Bank made appropriate selection of topics for data collection and surveys? Have data and surveys been well-maintained and disseminated?

Some of the data collection that has been undertaken for the studies I was asked to evaluate is incredibly useful, as I discuss in more detail below. In particular, in my view, the data assembled on the number of poor by countries is very useful (although it could be made easier to access by outsiders) and I was very impressed by the effort to introduce a high-quality mental health module in the LSMS. Improving and enriching the LSMS in a variety of areas should be a priority.

Some of the data have been collected for the purpose of evaluating particular programs, which is in principle also useful. Here, whether or not the data should have been collected is a function of whether they allowed for a good evaluation of the program or not. There are examples of data that should and should not have been collected in the research I reviewed.

Finally, some data were collected for the purpose of a particular research project, without being linked to the evaluation of a particular Bank program. There, I am worried that the data were just fitting the particular goal of this research project, and do not have much use beyond that project. In this case, there should be a procedure for the project to be competitively evaluated on the strength of the value of this particular piece of research, allowing open competition with outside researchers for this money. Otherwise precious public resources may be lost in pursuit of individual research agendas.

The dissemination procedures were not really described in the project, so I am not very sure how to answer this question. Some of the data do not appear to be public. At least they cannot be found very easily on the web. This should change. For example, World Bank funded data collection should come with the requirement to post a fully documented data set on ICPSR as well as on a World Bank website.

5. **In your area of expertise how could future Bank research better serve the developing country objectives?**

I think there should be five main areas:

1-Collection and diffusion of high-quality "general purpose" data on a variety of domains and areas (this could start by piloting new instruments, and then by implementing them widely). This implies continuous work on the LSMS (and fielding of more and more LSMS, if possible with larger samples). The work of Hammer, Rogers and others on absenteeism comes to mind, as does the work of Das on medical care and medical competencies. This could be replicated in many other countries.

2- Collection of data sets on the important changes in policies and institutions at the local (district, province) level in many countries, which could be linked with the LSMS. For example, someone could match the data on set-asides at the Gram Panchayat level in India to the LSMS for Bihar and UP at the village level. The World Bank would have a tremendous comparative advantage in doing so.

3-Helping the member countries strengthen their statistical apparatus to improve the quality of the data they collect, and requesting in exchange that the data be made available at some reasonable cost. Helping them to disseminate (the Australian Bilateral Aid agency has done that with the Bappenas in Indonesia; I think the World Bank played a role in the increased availability of the NSSO data in India).

4-Rigourous impact evaluation of a few of the World Bank supported programs.

5-Testing of the World Bank's most publicized claims (those that form the core of the WDR for examples) in randomized field experiments in several countries and settings.

6. **Particularly for ongoing projects you have evaluated, please comment on the proposal review, revision and selection process.**

The information for doing this was not provided.

7. **What is your overall assessment of Bank research?**

It is difficult to summarize a very diverse set of research projects. Generally, the World Bank's research is at its best when it does what no-one else has either the incentive or the means to do: assemble large quantities of data (pro-poor growth)—investing tremendous effort in developing a new data collection tool that will allow the collection of comparable data on new issues (cost of mental health—Das and Hammer); use fine details of a program to design credible identification for a project's impact (some of the poor area papers); use its leverage with the member countries to allow for randomizing either program placement or the details of the program rules (reaching the poor: Cambodia and fighting corruption in KDP).

It is at its worst when it follows some fad either in academia or from within the Bank and comes up with an isolated research project that has little scientific value and no external benefits; or when it collects expensive data sets to estimate program impacts even when it is really not possible to do so.

Comments on Individual Research Projects

Project title: Responding to the Value of the Poor

The only output in this research project is a training manual "which summarizes ways in which culture and poverty reduction activities harmonize and intersect." The objective of the training module is to train "you" (I think you is a World Bank official) to tailor poverty reduction activity to the group's values and aspirations. It is also meant to enable this person to be sensitive to power relationships in the group. It proposes a series of exercises that one can do in a village to understand a community's value (including "dimensions and matrices," "longitudinal exercises," "analysis of a common focal experience" and "interior reflection." The training manual uses as examples some of the Bank's failed and successful projects (where the failed projects did not incorporate culture and the successful projects did), and seems to propose exercises that have been suggested elsewhere in the literature.

I am not an expert in this area whatsoever. However, I am somewhat suspicious that (1) this qualifies as research (there are no data beyond anecdotes, no original theory, and the literature review is at a very basic level), (2) I would trust the assertions made in the manual, and (3) the training manual would be that useful if it were to be used to train anyone.

(1) May not be a problem to the extent this is an output of the research group that serves the rest of the Bank.
(2) Is a more serious issue. The training manual reads as if it just echoes an ideology. Beyond the anecdotes, there is no mention of any evaluation of the usefulness of any of this in getting projects to work. For example we are just told that "Manfred Max-Neef., a Chilean professor and activist, together with his associates, developed a matrix of human need" and that the matrix has been used in several countries.
(3) The description of the various methods is so abstract that I do not think anybody, even if they were convinced, could just start following instruction in the manual

> and get a PRA, or any other method, going. It is lacking the concrete details and instruction that one would need to really put these things to work.

On balance, I found this extremely disappointing. This sounds like it could have been written as a caricature of the "Voice of the Poor" approach. It is a collection of received wisdoms that do not come with much apparent groundings in any serious research. But again, I may have been missing something.

Project title: Linking Poverty and Culture in the Himalaya

The only out put of this research project is a 114-page long unpublished article that studies the extent to which Bhutan, Nepal and Ladakh have incorporated traditional architecture in modern architecture "to retain a sense of place." The main method of research was workshops where the officers discussed the topic, and case studies. Again, this is not my area of research at all, so I many not be the best placed person to evaluate this project. What seems fairly clear, however, is that the report has no bearing on the poverty reduction strategy of the World Bank. Despite the title, the report is mostly about architecture (although there are some boxes about NGOs that fight poverty in these countries here and there in the report). I do not understand why this project was funded by the World Bank.

Project Title: Culture and Development

This is another project on culture, but with this project, we enter the more comfortable zone (for me) of social science research based on data and hypothesis. The project displayed three articles and a book. The objective of the project is to produce scholarship on the relationship between culture and development policy. Two articles consider important (and, many would argue, inefficient) social phenomena (lavish expenditure on weddings by poor families and violence against women) and combine qualitative research, simple economic modeling, and econometric analysis to try and shed some light on those phenomena, both in the context of "signaling" games. (In the case of weddings, the hypothesis that is put forward is that the family is trying to signal to the village that the daughter is making a good marriage; in the case of violence, the hypothesis is that violence is used to force the bride's family to transfer more to the groom: violence increases when the dowry paid was lower–this is instrumented with Biju Rao's marriage squeeze instrument—and the bride's family is richer.) A third paper is a more descriptive exercise regressing measures of women's empowerment on cultural and economic variables, in a sample of villages from Karnataka and UP. I found the third paper to be a very interesting descriptive exercise.

I was a little less convinced by the first two. They are both on interesting questions and they put forward plausible and interesting hypotheses. They follow the accepted "rules of the game" in writing economic papers (some theory, some data, an instrument). In both cases, I was not sure the theory was adding much to the argument (the model is just a way of restating the main story and does not bring much new insight—in particular it does not really form the basis of the empirical specifications that are being tested; in some cases it ignores important pieces in the literature on the subject, i.e. Postelwaite's work on the same topic). The empirical work is not hugely concerned about the validity of the identification assumptions necessary for the instruments to work (although it is attentive about the main problem)—for example there is no discussion on whether the exogamic villages in Karnataka could have different norms about how much to spend on weddings as well. Overall they are perfectly acceptable pieces of research, but no breakthrough. The problem with the identification strategy not being perfect is that if the policy recommendations have to be more specific than "culture needs to be taken into account," it is important to be sure of the causal conclusions that are being drawn.

The most substantial piece (in terms of number of pages, anyway) of this research project is a book "Culture and Public Action," which brings together social scientists from various disciplines. The book is motivated in the introduction by how culture can help or hinder a Bank project, but most chapters are not really concerned about this issue at all. The contributors are a mix of very senior people (Amartya Sen, Mary Douglas) and quite unknown people, and the chapters also vary in quality. I was not very sure which niche the book was planning to occupy. It said that the objective was to be read by students and other scientists, but on balance I do not think it has the focus or standing that would lead to it being assigned in class. On the other hand it is not entirely clear that there are many lessons that can be useful to the Bank's managers.

The overall assessment of this research agenda is that it clearly makes great progress relative to the first two I reviewed, which were also on culture. We are now in the presence of serious scholarship. The overall topic is relevant for the World Bank's poverty reduction strategy. Still, this is a difficult subject, and I sometimes have the sense that the World Bank's economists are stepping outside their comfort zone when they discuss it. There is substantially better work done by economists on non-market forces (i.e. Kaivan Munshi's work on networks). One comparative advantage of World Bank's economists is that they have resources and contacts to collect data, and while this is exploited here (all the data used have apparently been collected for the authors, and with substantial inputs from them), this is a little buried under the effort to write models. It also appears that the data are a little too targeted to the authors' specific needs. That said, the data generated by this project look interesting. Are they available to other researchers? Where can the data be found?

Project title: Economic Cost of Mental Health in Post Conflict Bosnia and Herzegovina

The goal of this project is to collect and analyze data on mental health as part of an LSMS survey in Bosnia and Herzegovina. The project developed and piloted a mental health module to be added to the LSMS, and the research describes the process of doing that and validates the measures of depression that were generated by the survey. I am not a specialist in measuring mental health, and I cannot really assess how well this was done here, and whether the right choices were made, etc. But from what I can see, I find this research project tremendously useful. First, it generated new data that are presumably eventually going to be made available to others in the World Bank and outside of it.

These data are linked to the LSMS, another great contribution of the World Bank, and this creates an unusually rich data set on a central question. I am sure the data will be tremendously useful to many researchers and policy makers. Second, the papers provided describe in detail and very clearly how the data were collected, what problems were faced, what choices were made, and what are the success and limitations of the approach. It provides useful thoughts for how to do the same exercise in the future. It also turns out that the data proved largely useful, which is an added bonus. The analysis will be very interesting when it is performed. But even if the effort had failed, the returns to trying things like that and being forthcoming about its success are enormous.

Overall, this is great work, and I would love to see more effort in this direction: more experimenting with new modules to be collected in conjunction with the LSMS (on various topics), and more careful assessment of the lessons from these experiments. Collecting useful data that can be used by many to better understand poverty should be, in my view, one of the core missions of the World Bank's Research Department.

Project title: Social Capital

This project aims to explore the social dimensions of development. The research program is quite rich and diverse, and difficult to summarize in a few sentences. There are many publications listed, including a few highlighted as the "best output."

One theme that the team chose to highlight is the "participatory econometrics" idea (there is one "methodology paper" in the packet), and two applications (the papers under the projects "culture and development" are also applications). This is a methodological theme that is not necessarily linked to work on social capital.

The idea behind "participatory econometrics" is quite sensible: it argues that field visits and qualitative interviews can be used to generate useful ideas for how to ask questions, ideas of research questions, and ideas to generate instruments. All of this is clearly true, but I think this describes what all good data collection projects do, and I was a little confused about what is new here. A description of an actual process (like the one in the project on measuring mental health) would have been more useful than these general statements. The summary note for the research project claims that the theory and method papers in this research project have been widely cited, according to Google Scholar. I supposed this was meant to describe the theory piece here. In fact it is cited 16 times, including 13 times by members of the team (the most cited output in the highlighted project is cited 40 times). For comparison, there are other papers on mixed methods on the same page in Google Scholar that are cited more than 100 times. I think this reflects the fact that there is not a huge amount that is new in this piece.

Two applications are highlighted: an evaluation of the Social Fund program in Jamaica, and the paper "Governance in the Gullies." "Governance in the Gullies" is a largely descriptive paper on how the poor in Delhi slums access public officials. This is an interesting description, although the econometrics part suffers from not having a very clear hypothesis to test. It would have been interesting to insist more on the descriptive statistics. But the patterns in the data are interesting and new (to me anyway).

I have much stronger reservations about the paper that “evaluates” the Social Fund program in Jamaica. The impact of social fund programs is very difficult to evaluate, for various reasons: they are selectively placed in many different ways (villages need to be poor but they also need to be energetic enough to have a project), there is no clear outcome that they are trying to affect (except for poverty, but this is very long run). The paper address this problem by constructing five “matched” pairs with one village receiving JSDF (the program) and one program that did not receive it. The paper has some descriptions of the processes (which is fine as a process evaluation) in each of the five treated communities, and proceeds to run matching estimators on the data, which makes no sense, since there is no argument that the matched “controls” are comparable to the treated communities in terms of unobservable characteristics. The paper was used as one of the background papers for an OED paper that was very critical of the CDD projects. From what I can understand, this report was quite influential, but if it was based on similar studies in the other countries, this is actually quite misleading. The paper does acknowledge in the conclusion that “the results do not suggest either that JSDF is a failure or a success,” but in other parts it says that the project “did not succeed in what it intended to do.” Another core mission of part of the World Bank would be to set up the standard for what constitutes project evaluation within the Bank, and also to run some successful evaluations. This paper is quite counterproductive in this regard.

The other outputs of the research projects are a theoretical paper on the possible impacts of different forms of social capital on health, which I did not find illuminating, and an interesting paper that describes the failure of the “bureaucratic” solution to social service delivery, and provides a classification of the different “solutions that have been proposed in the literature.” This piece seems to have evolved in the conceptual framework of the WDR 2004 on social service delivery, which was very influential.

Overall, this research program seems to suffer from being “over-theoretical” and “over-reflexive.” The World Bank supports a series of projects that support, or leverage, social capital. These projects would be an opportunity to put some of these ideas to the test, either by rigorously evaluating these programs, or by varying the rules that govern the implementation of these programs (which can be done in a randomized framework). This was done by Ben Olken in the context of the KDP program in Indonesia, with the help of Scott Guggenheim and Lant Pritchett, from the World Bank. The project is not highlighted under the research program, perhaps because it ended up being funded entirely outside of it. I think these are exactly the type of studies that this group could and should undertake in larger number.

Project Title: International Development Goals

The objective of this research project was to provide analytical support to the World Bank’s work on the Millennium Development Goals. Two of the research projects provide cost estimates for how much money will be needed to do that, but in the process of doing that, the authors do realize that it is not only *how much money* is spent, but *how* it is spent that will determine whether or not the MDG are achieved. There is only one example of the importance of the organization of the service delivery in this packet ("Working for God"), but the note on the research project points out that subsequent major projects (such as the Bangladesh absence survey, and the Uganda health survey) were inspired by this work. Indeed, the same team was responsible for the WDR 2004 on social service delivery.

The two costing exercises rely heavily on cross-country growth regressions. One paper, "Goals for Development," uses cross-country growth regression of the impact of aid on growth on the one hand, and growth and poverty on the other hand to estimate how much aid money would be needed for reducing poverty by half. The second part of the paper adopts an "accounting" approach to estimate the cost of achieving the schooling and health MDG (i.e. take the number of children out of school and multiply it by the average cost of having a child in school. Another paper in the same packet points out that the accounting approach is misleading, as there is (in cross-country regressions) only a weak relationship between spending per capita and educational outcomes. It then produces estimates based on the elasticity measured in cross-country studies, which are of course much larger. The overall policy recommendation from these papers is not really that the World Bank should spend that much, but more that money is not sufficient. The WDR 2004 was an important sounding board for this message. While I buy the point, I find that making it on the basis of weak regressions, which can be (and have been) attacked by many people, is not a particularly reliable way to make progress.

The paper "Working for God" is completely different. It shows that faith based health providers in Uganda function better than public hospitals and that money given to them has a large impact on variables that can lead to better health outcomes. This is an interesting paper, although one must be careful about drawing policy implications (which the paper does not draw, to be fair). If one started giving more money to faith based health providers, the selection into who becomes a faith based health provider would very likely change rather dramatically.

Project title: Growth, Redistribution and Human Development

This project (a research proposal) starts exactly with the opposite stance as the previous project: one cannot use cross-country regressions to infer the impact of growth and inequality on human development indicators. Instead, it proposes a micro-decomposition approach to examine the determinants of aggregate health and other human development outcomes. This is a worthwhile research undertaking.

The plan is to regress human development outcomes on some external factors, and household income (instrumented; and where the estimation is non-parametric) and use this as a basis to calculate the impact of growth and inequality on these outcomes.

The application will use data from Vietnam and Morocco, although the primary aim of the project is methodological.

The exercise is interesting, but there is no discussion in the paper on whether the estimated non-parametric relationship between (instrumented) consumption and health or other HD outcomes will be indicative of a causal relationship. The instrumentation does not help here (it is mainly meant to deal with measurement issues). One could imagine that richer households have access to omitted infrastructure variables that will not improve with growth, or that there is a direct impact of one's rank on one's health (or desire to educate one's children), in which case a "tide that lifts all boats" will not have the impact on education that the cross-sectional relationship indicates. Finally there is the possibility of reverse causality.

It is surprising that this point is not really treated, given how much emphasis is put on other (relatively less important) technical issues. That said, the exercise will nevertheless very likely be instructive.

Project title: Reaching the Poor with Effective Health, Nutrition, and Population Services

The objective of this project is to present a series of case studies from various countries on the ability of health services to reach the poor.

Three of the papers are strictly focused on the question of the targeting of health and nutrition services: they use household surveys to calculate the share of households in different wealth quintiles that make use of various health and nutrition services, and show how this varies over time and across different types of programs. These are interesting descriptive papers.

Two papers are impact evaluations of two fascinating projects. In the first project, in Cambodia, the provision of health services was contracted out to selected NGOs in a small number randomly selected districts (8 out of 12 districts were initially randomly selected). In the second project, insecticide treated bed nets were distributed for free in Zambia and Ghana at the same time as a measles vaccination campaign took place. Ownership of the nets increased from 4% to nearly 100% in Ghana. While both of these programs are really interesting, and it is very refreshing to read papers that are mainly interested in the programs they look at, and the effects they might have, they both have serious problems.

The paper on Cambodia suffers from the fact that the randomization failed. Not all districts that were initially selected for the "contracting out" part were in fact contracted out (the others got no bidder). No data were initially collected from the "treated" districts where there was no bidder, and the analysis in this paper is therefore not valid (since it is based on a selected sample). This is not explained clearly in the report but subsequently, a wider group of authors (including Michael Kremer as well as the original authors of the study, with great help from Benjamin Levinsohn) collected a set of data covering the entire set of districts, and there is a new paper, which is now being circulated, that covers the entire set of districts.

The paper on Ghana and Zambia makes no attempt at a formal program evaluation (with a treatment and control set-up), although the set-up would have allowed it (it seems it would even have been possible to randomize the camps in which bed nets were distributed; but even without that, data on "control" districts should have been collected and analyzed). It would also have been interesting to look at the impact the program had on the take-up of the measles immunization itself. (Dupas (2005) shows very large impact on the take-up of pre-natal care services of free bed net delivery at pre-natal visits.) The findings here are really important, but an opportunity to do more seems to have been lost.

I did not fully understand what happened with these working paper series: most of the papers are apparently not written by World Bank researchers, and the level of technical expertise in this paper is weak, although the questions are fundamentally important. It seems like an error like the one in the Cambodia paper should not have been committed in the first place. Many in the World Bank complain that randomized evaluations are difficult to set up, that the governments are not keen, etc. But it seems that it should be the role of the research group to ensure that when an opportunity occurs, it is taken advantage of. The technical know-how is definitely present within the research group. It seems that there are many people who could have been consulted and would have been happy to help collect the data and work on the study. It would be nice if such opportunities were taken advantage of more routinely.

The research projects also provided a version of the findings for general distribution, which was very nicely produced. This is a very useful output of this project.

Project title: Social Exclusion and Poverty

The objective of this research project is to study how social exclusion (gender discrimination; discrimination on the basis of ethnic or caste group; exclusion based on income) and poverty interact and reinforce each other. Some of the inputs of the project were background work for the WDR 2000 on poverty, and the project seems to be related to the "voice of the poor" project. The two most influential outputs proposed by the projects are two chapters in the WDR 2000, one on "empowerment" and the other on "removing social barriers." Two articles are also proposed: one argues that son preference is largely cultural, and related to exogamy and patrilinearity more than to anything else. The other article studies the possibility between state-community synergies for community driven development.

In my view, the best article is the article on son preferences, which builds its argument based on detailed description of the institutions and the preferences for sons in China, India, and Korea. This is a very interesting, very well argued hypothesis.

The WDR chapters are disappointing. They argue most of their cases on the basis of cross-country relationships that do not make much sense (e.g. there is less bureaucratic delay in countries that have meritocratic recruitment procedures), or on the basis of anecdotes. They leave a little bit of a feeling of concluding that "good things are good" (i.e. social capital is good, but not when it excludes the poor). More description of the actual level of discrepancies between men and women (for example) in various countries, and the different policies that have been tried to remedy those would be more useful.

Project title: Pro-Poor Growth

The objective of this research project is to document where (and to some extent why) economic growth has led to sharper reduction in poverty, and what type of growth has led to a faster reduction in poverty. This is also the objective of an isolated paper, "Why has economic growth been more pro-poor in some Indian States than others," so this paragraph covers this article as well.

At the core of this research agenda are very important (both in terms of quantity and of consequences) data collection efforts. The "povcalnet" team has assembled a remarkable set of data on the number of poor people in many countries in the world, as measured in household surveys of comparable (and decent quality), and where the poor are measured in a uniform way (as those who live under a dollar a day, at PPP). This involved calculating consumption PPP for all countries, and using each national survey (including, but not restricted to, the LSMS) available to compute poverty rates. Much of this data are otherwise of restricted access, so this is a really valuable enterprise. The only caveats are that finding the data is really difficult (if you wanted to use PPP, the only way to get at it is through the povcalnet search engine, and someone knowledgeable will really need to show you the way), and I could not find a place where what was done was documented. Recent World Bank research also showed that some of the PPP may be widely off, but this is a much more general problem: calculating PPP is hard and no comprehensive exercise is likely to be 100% correct.

Likewise, the work on India is based on data sets that have been assembled from the NSS sample surveys (and have been distributed to many other researchers) and the work on China exploits data that are not available to outside researcher to construct poverty series for urban and rural areas (I am not aware if the aggregate time series could be diffused to outside researchers). This is unambiguously a very useful effort, for which Martin Ravallion really needs to be commended.

The papers in this research project run regressions that use the data to document how much inequality has hampered the poverty-reducing impact of growth (along the way, they make lucid (and welcome) arguments against misleading assertions in another World Bank research project–Dollar and Kray—which has had a quite devastating influence), using both the cross-country panel data set, a state-level panel data set in India, and a province level panel data set in China. The work is mostly descriptive, but these are very helpful descriptions. I think this is an excellent research project

Project title: Poor Areas or Poor People?

This project is a combination of papers on different topics. I should first mention that there are MANY papers in this project (and in the pro-poor growth project as well). Many of those are published in good field journals. Martin Ravallion, the head of this project, is clearly amazingly productive. Still, many papers make related points, and there may be a little too many papers here.

A few papers fit under the theme “poor areas or poor people.”

The general research question this project addresses is whether there are characteristics about areas that make people who live in them poor: in other words, can the differences in poverty from one region to the next be explained by the fact that people who live in the poor regions have characteristics that make them poor, or would people with exactly equivalent characteristics have different outcomes in some regions than in some others. One of the highlighted papers performs an Oaxaca decomposition of the source of poverty differentials across regions in Bangladesh, and finds that region dummies matter. It could be that many people with very good observed characteristics or with very high returns to one particular characteristic chose to locate in a particular region, so one should not over-interpret these results.

The second highlighted paper asks the same questions, but starts from a consumption growth model, where the growth in consumption from one period to the next is explained in part by fixed region-level provinces. They introduce individual fixed effects that may be correlated with the province fixed characteristics, but the model is still identified by letting the effect of these fixed effects on growth vary over time. I did not understand this paper very well. It seems somewhat related to Hausman-Taylor (which is not cited), but it is very hard to follow, and I did not fully understand that the identification assumptions made sense.

Some other papers fit under a general “program evaluation” header. They almost all use propensity score methods combined with DD.

The project highlighted a program evaluation of a development project in China. The paper is a DD with propensity score matching, and its originality is to show impacts both on savings and on poverty. The paper is perfectly reasonable, although given that the matching technique is well understood in general, and there is a large literature on it, it would have been useful to replace some of the detailed discussion of the technical aspects of the matching methods, rather by a discussion on whether the matching has really controlled for the unobserved variables, which is largely absent.

This last point is not a feature of this paper alone, but is present in most of the papers emanating from this group. This is a bit problematic, because of the influence that these particular papers have, on other researchers in particular: many researchers now seem to believe that matching corrects the omitted variables problems that OLS does not address.

This is obviously not true if the problem is that some important omitted variables are left out of the list. Matching just improves on the functional form. I think it would be very important for this group to be much more explicit and candid about this point, which they are clearly aware of, but others may not be.

Project title: What Can Ex-participants Reveal about a Program's Impact?

This paper was isolated, although it could have been combined with the previous research (which has several other papers on the same topic). The paper proposed to do matched DDD between leavers and stayers of a program (and then with one more level of difference between them and never-joiners of the same programs). I think this paper has some problems. It is very hard to read (in part because there are—I think—typos in the tables, so that what the table say is estimated is not what the text says—or maybe the text is unclear: in particular I could never understand whether the first DD in the tables is between stayers and never joiners, or whether it is between stayers and people who join in period 2). I also think that the point that leavers can be used to estimate program impact is generally not valid. First, people must leave the program endogenously, and in particular after having observed their gains in the program, so that the assumption that their leaving is independent of their growth trajectory is not tenable. Second, we have no idea about the dynamic effect of the programs. Note that contrary to what the text says, the identification assumption cannot be "tested," at best we can see whether there is some evidence that is consistent with it. In Argentina it looks like it would have been possible to construct an instrument for whether someone leaves or not by looking at whether their project closed down. It would be have been interesting to look at those people (although this does not address the second problem).