

## **Evaluation of World Bank Research**

**Marcel Fafchamps**

### **General remarks**

Before embarking on the evaluation of individual projects, I would like to make a few remarks. First, we are asked to evaluate individual research projects, not individual researchers. This is different from the academic evaluation practice, which focuses on researchers. I confess I find this a bit difficult. In my case I was given four projects, two of which involve the same researcher. In evaluating the individual projects I find it difficult to abstract from what I know of the researchers themselves – including what is revealed about them by the four research projects that I am asked to review. It would worry me seriously, for instance, if the researchers involved in projects 116 and 117 did not have research output other than that reported in these two projects.

Second, although we are asked to assess whether the research projects are cost effective, we are given little information about funding. For instance, it is unclear whether the figures reported in the Total Funding column of the spreadsheet represent amounts budgeted or spent, whether the time spent by researchers is imputed in the total, whether dissemination expenditures (e.g., big conferences) are included, and whether data were collected that can be used by other research projects. Without that information it is difficult to assess the cost-effectiveness of the projects. A similar comment relates to the timeliness of the research.

Third, the research unit of the World Bank has a dual role: to produce good quality research and to support the operational work of the Bank. One could imagine a system in which each researcher devotes an equal amount of time and effort to both tasks. But it may be more cost-effective for the research unit to let particularly productive researchers specialize relatively more in research while somewhat less productive researchers specialize in operational support. As a result, some individuals will produce more published research than others. This may reflect negatively on those who have published little even though their input is essential to the work of the research unit as a team. In academia similar issues arise when faculty members specialize in teaching or administration.

Finally, we are asked to assess research output relative to its intended audience. Since, in selecting their best output, researchers have emphasized academic publications in top journals, their intended audience appears to be the research community at large. If their intended audience was policy makers, we would have been given policy briefs and other dissemination documents to look at. Having said this, I suppose that the research is also intended to assist policy makers. I keep this in mind in my evaluation, but it is important to recognize that it is extremely difficult to serve both audiences well. Some researchers may serve one audience better than the other.

## **Rural roads and welfare** – Dominique van de Walle

Note: Although Vinaya Swaroop is mentioned in the spreadsheet, this person appears to have not been involved in the project.

This is an impact evaluation study of a rural road rehabilitation project in Vietnam. Program placement is not random but targeted toward priority areas. Random samples of 100 project communes and 100 non-project communes were selected and surveyed in 1997 (baseline), 1999, 2001 and 2003. In each sampled commune a questionnaire was administered to 15 randomly sampled households. A project-level database for each of the project areas was also constructed. The ultimate objective of the study is to evaluate the impact of rural roads on living standards. The project has not been completed yet; the completion date is December 2007.

I examined two project outputs:

1. van de Walle, Dominique. 2002. Choosing Rural Road Investments to Help Reduce Poverty. *World Development* 30(4): 57589.
2. van de Walle, Dominique, and Dorothyjean Cratty. 2005. Do Donors Get What They Paid for? Micro Evidence on the Displacement of Local Resources, Policy Research Working Paper No. 3542, Development Research Group, World Bank, Washington, DC, March.

Output 1 is not a research paper but rather a practical discussion on how to select road investment sites in a world of imperfect information. The paper provides few links with the economic literature on social choice, public finance, and welfare economics. I did not find it particularly illuminating, but perhaps this is because I am not a policy maker having to decide where to locate roads.

Output 2 is a recent working paper testing whether World Bank financing of rural road rehabilitation targeted to specific geographic areas displaced local financing. This is accomplished by using propensity score matching to correct for non-random program placement. The paper finds that, without correction, project and non-project communes are not statistically different in terms of road rehabilitation. Once the authors correct for non-random placement, they find a statistically significant effect of the project on road rehabilitation.

I like the paper. It is clearly written and competently done. It is a pleasure to read and I learned a lot from it. Given that the ultimate objective of the study is to examine the impact of a World Bank project on living standards, the fact that project and non-project

communes are identical in terms of road rehabilitation must have come as a shock to the researchers involved in this research. Obviously, if the control group received the same treatment as the project communes, it is difficult to infer the effect of the treatment on living standards. This is why the researchers first have to ascertain why the two groups appear identical and whether, once we control for non-random project placement, the two groups are in fact different. Finding that they are must have come as a great relief to the researchers. But the difference is not large and the use of the parity bounds matching method implies that many communes have to be dropped from the living standard analysis, thereby seriously reducing the size of the useable sample – and the strength of the results.

This project perfectly illustrates the difficulty of impact assessment as a research tool: the researchers do not have complete control over the experiment. In this case, they could not prevent the Vietnamese government from diverting to non-project communes the road rehabilitation funds that had been saved by the project. It also appears that new roads were constructed in project communes, even though this was not supposed to take place. This makes it difficult to assess the impact of the project on the target population. Even so, I thoroughly enjoyed reading output 2, which I regard as a cautionary tale about impact evaluation but at the same time an upbeat story about how to make up for loss of control. The paper provides valuable lessons, especially in terms of how to prepare oneself against such eventualities. It also raises interesting questions about government compensatory behavior that apply to all projects and hence to all impact evaluation exercises – i.e., How can we be sure that the government does not compensate the control group in some other way, and is it feasible or even ethical to prevent such compensatory behavior?

Regarding cost-effectiveness, I am a little confused. The spreadsheet gives a total funding figure of 1058 (which I take to mean US\$1,058,000). But the project proposal attached to what I received has a price tag of US\$297,800 only. Perhaps there is more than one project under this heading. I guess one million dollars for a single paper is too much. But the project also collected data that are to be used for project evaluation, and the data analysis has not been completed. Timeliness is more of an issue: given that data collection was completed in 2003 and that impact evaluation is meant to assist project evaluation and subsequent project development, I would have expected faster completion.

Summary: Slightly less than average on research quality; much less than average on timeliness and cost-effectiveness.

<b>S</b>	Superior
<b>AA</b>	Above Average
<b>A</b>	Average
<b>BA</b>	Below Average
<b>U</b>	Unacceptable

	<b>S</b>	<b>AA</b>	<b>A</b>	<b>BA</b>	<b>U</b>
<i>Please rate the following aspects of the project outcomes (or in the case of ongoing projects, design and intermediate outputs)</i>					
<b>Topics</b>					
Importance of the issues addressed			<b>x</b>		
Clarity of the project focus and stated objectives			<b>x</b>		
<b>Analysis</b>					
Theoretical/conceptual framework				<b>x</b>	
Empirical application			<b>x</b>		
Statistical and econometric methods			<b>x</b>		
Use of existing knowledge and resources			<b>x</b>		
<b>Data : no information</b>					
Awareness of other data sources					
Data compilation, cleaning and cataloging					
Survey design and sampling					
<b>Output Quality</b>					
Writing quality (clarity, organization, etc.)			<b>x</b>		
Clarity of conclusions and recommendations				<b>x</b>	
Extent to which conclusions are based on analytic evidence		<b>x</b>			
Appropriateness of the recommendations					
Appropriateness of output form (working paper, book, database, etc.) for intended audiences			<b>x</b>		
Availability of translated outputs where appropriate					
<b>Extent to which research:</b>					
Increases knowledge and understanding of the issues				<b>x</b>	
Provides a sound basis for policy				<b>x</b>	
<b>Actual or likely impact of research on:</b>					
Government policy					<b>x</b>
Future analysis				<b>x</b>	
The development community in general				<b>x</b>	
<b>Overall Quality of Research: too little output</b>				<b>x</b>	

## Studies of Aggregate Agricultural Growth – Donald Larson

This ongoing project includes a cluster of research activities related to agricultural growth and productivity. The objective of the project is to identify key determinants of agricultural growth and productivity, with an emphasis on informing agricultural policy.

The first cluster of activities centers on the estimation of cross-country agricultural production functions. All activities in this cluster involve Yair Mundlak, emeritus faculty member from the Hebrew University in Jerusalem. Two outputs were submitted from this cluster of activities:

1. Yair Mundlak. 2001. "Explaining Economic Growth". *American Journal of Agricultural Economics*. 83(5): 1154-67.
2. Yair Mundlak, Donald Larson and Rita Butzer. 1999. "Rethinking Between and Within Regressions: The Case of Agricultural Production Functions." *Annales d'Economie et de Statistique*, 55-56: 475-501.

The second cluster of activities focuses on agricultural productivity within countries, with a specific emphasis on Ecuador and Nicaragua. Articles written in this cluster involve Cruz Aguayo, who I believe to be a graduate student in the School of Public Policy at the University of Maryland. One output was submitted from this cluster of activities (the name of Cruz Aguayo appears as author in the spreadsheet and on the summary statement provided on the CD but not on the paper itself; in contrast, the name of Mauricio Leon does not appear in the summary statement on the CD; this is confusing):

3. Donald Larson, Cruz Aguayo, and Mauricio Leon. 2005. How Endowments, Accumulation and Choice Determine the Geography of Agricultural Productivity in Ecuador. Working paper under journal review.

Other related research activities took place under the heading of this project, but do not form the core of the project. I comment only on the two clusters for which research output has been provided. More work is anticipated on cluster one and we are told that a revisit of the cross-country study is underway.

Output 1 is the text of a lecture delivered by Yair Mundlak to the AAEA annual meeting in August 2001. The lecture revisits a number of points, made by the author earlier in his career, on the estimation of production functions. The basic idea is that producers choose their production technique, a point that is, according to the author, often ignored in practice. The author argues that estimating Cobb-Douglas production functions is misspecified because it fails to recognize that factor elasticities vary with the technique of

production, which is endogenous. The author also discusses the issue of country fixed effects and emphasizes that within-country regressions yield different production function estimates from between regressions. These points are illustrated with a cross-country regression of agricultural production over time. Output 2 revisits this issue in more detail, but follows basically the same idea. We learn from the AJAE paper that these points were initially made by Mundlak in 1988 and 1993.

I confess I fail to see the value added from this research. The profession has long acknowledged the limitations of the Cobb-Douglas production function. More flexible functional forms have been introduced to remedy these limitations, a point that is ignored by the authors. A more intellectually satisfactory approach is to posit a general functional form that encompasses all the different technologies/techniques of production available at a point in time, and to estimate its parameters.

The literature on cross-country regressions has also long recognized that country fixed effects often capture most of the variation that the researcher is interested in, and has expressed much doubt regarding the usefulness of country fixed effects in this context. (See for instance Durlauf and Quah, *The new empirics of economic growth*, 1999, for a detailed discussion that pre-dates the research presented here.) The authors also ignore the discussion of factor-biased technological change, which attracted a lot of attention in the 1970s. (See for instance the work of Hayami and Ruttan on induced technological change in agriculture.)

On a more fundamental level, the estimation of a production function by definition requires that not all producers use the same technique of production. For instance, in a simple Cobb-Douglas framework with capital and labor, if all producers use the same capital-labor ratio, the labor and capital shares are not identified. Identification of a technique of production therefore requires that individual producers choose a different input mix. This in turn implies that they face different relative input prices. It is this realization that makes researchers usually reluctant to adopt a very flexible functional form because identification is problematic. This is particularly true when using cross-country aggregate data, which usually include only a small number of data points for each country. These concerns are simply not recognized by the authors in the two papers I was asked to examine. Moreover, although the authors make a big issue of the fact that the choice of technology is endogenous, they do not seem to realize that this by itself creates estimation problems.

Output 3 from the second cluster is different because it focuses on a single country. Using data from 128,000 farms in Ecuador, the authors estimate a flexible form efficient frontier production function containing 118 parameters. Inputs are regarded as exogenous in the estimation. The analysis focuses on input elasticities of production and on regional differences. The authors find support for their flexible functional form. They also find that large farms are more efficient. They discuss the implications of their research for policy and poverty alleviation.

While I find this output much more satisfactory, I am nevertheless left somewhat disappointed. The paper reads like a methodology in search of a topic. Estimating a complex frontier production function with over 100,000 observations is a massive effort. But the researchers seem to have little idea of what hypothesis it is they wish to test. Some of the same arguments found in outputs 1 and 2 are reproduced here to justify the use of a flexible functional form. This is fine. But since the reader is not told what hypothesis is being tested, it is very difficult to form an opinion about the paper: without knowing which variables are of interest, I cannot decide whether endogeneity is an issue, for instance. I also cannot suggest robustness checks to verify the validity of the test results.

To summarize, it is true that this research project produced published output. But I would argue that the published output is of questionable quality. The research is not very original, drawing on ideas and techniques developed a number of years ago. The authors do not appear to understand the need to clearly state what hypothesis they wish to test, how they intend to test it, and what robustness check they need to perform. Finally, the policy implications coming out of this research are entirely unclear. For these reasons I would argue that this project fails to deliver.

Summary: Below average on research quality

<b>S</b>	Superior
<b>AA</b>	Above Average
<b>A</b>	Average
<b>BA</b>	Below Average
<b>U</b>	Unacceptable

	<b>S</b>	<b>AA</b>	<b>A</b>	<b>BA</b>	<b>U</b>
<i>Please rate the following aspects of the project outcomes (or in the case of ongoing projects, design and intermediate outputs)</i>					
<b>Topics</b>					
Importance of the issues addressed				<b>x</b>	
Clarity of the project focus and stated objectives					<b>x</b>
<b>Analysis</b>					
Theoretical/conceptual framework				<b>x</b>	
Empirical application				<b>x</b>	
Statistical and econometric methods					<b>x</b>
Use of existing knowledge and resources					<b>x</b>
<b>Data</b>					
Awareness of other data sources					
Data compilation, cleaning and cataloging					
Survey design and sampling					
<b>Output Quality</b>					
Writing quality (clarity, organization, etc.)				<b>x</b>	
Clarity of conclusions and recommendations					<b>x</b>
Extent to which conclusions are based on analytic evidence				<b>x</b>	
Appropriateness of the recommendations				<b>x</b>	
Appropriateness of output form (working paper, book, database, etc.) for intended audiences			<b>x</b>		
Availability of translated outputs where appropriate					
<b>Extent to which research:</b>					
Increases knowledge and understanding of the issues					<b>x</b>
Provides a sound basis for policy					<b>x</b>
<b>Actual or likely impact of research on:</b>					
Government policy				<b>x</b>	
Future analysis					<b>x</b>
The development community in general				<b>x</b>	
<b>Overall Quality of Research</b>				<b>x</b>	

## Impact of Public Spending – Klaus Deininger

The project explores the determinants of effective public service delivery in Uganda. It uses the National Household Survey to assess the effectiveness of programs to establish universal primary education, to waive user fees for health services, and to provide agricultural extension services. As far as I can tell, this project did not collect data but uses household survey data collected by the Ugandan government.

Five papers were produced under this research project, three of which were selected for evaluation. They are:

1. Klaus Deininger and Paul Mpuga. 2005. “Does Greater Accountability Improve the Quality of Public Service Delivery? Evidence from Uganda.” *World Development*, 33(1): 171-91.
2. Klaus Deininger, Anja Crommelinck and Gloria Kempaka. 2005. “Impact of AIDS on Family Composition, Welfare, and Investment: Evidence from Uganda.” *Review of Development Economics*. 9(3): 303-24.
3. Klaus Deininger. 2003. “Causes and Consequences of Civil Strife: Micro-Level Evidence from Uganda.” *Oxford Economic Papers*. 55: 579-606.

Using data from a 2002 national integrity survey of 12,000 Ugandan households, Output 1 examines whether household knowledge about how to report inappropriate behavior by bureaucrats is associated with less corruption and better service quality. Variation in the knowledge on how to report inappropriate behavior is not randomly assigned. To correct for non-random assignment, the researchers instrument this variation using household characteristics. They then regress household-level incidence of corruption on instrumented knowledge of how to report bad behavior. They find predicted knowledge to be strongly significant in a number of public services. They then regress satisfaction with a whole range of public services on predicted knowledge and mean corruption in the village. They find predicted knowledge to be positive and significant in all regressions and mean corruption to be negative and significant in most. The authors also report descriptive statistics on the incidence of corruption in various public services as well as in the private sector.

Without random assignment of knowledge on how to report inappropriate behavior, it is difficult to fully eliminate endogeneity concerns: households that are more knowledgeable in general are probably smarter and may be better able to avoid corruption while obtaining better service. The authors do not have panel data, which would take care of such unobserved heterogeneity. Given data limitations, the authors do the best they can by instrumenting knowledge on how to report. A useful robustness check would have been to use the propensity score matching method. This would not have taken care of unobserved heterogeneity but it would have served as a check on the

validity of the instrumenting approach to address the selection issue. To summarize, this is not an earth shattering piece but it adds to knowledge on an important policy issue.

Output 2 uses a household survey panel with two years of data – 1992 and 2000 – to investigate the welfare effect of fostering a child. As a result of massive adult mortality due to AIDS, child fostering has become extremely frequent in Africa, as has been documented by a number of studies. Some 15% of the surveyed households fostered a child during the study period. Using household fixed effects to account for unobserved heterogeneity, the authors find that adding a foster child results in a significant reduction in per capita consumption, income, and household investment. This reduction is more pronounced for the poor. The authors also derive other results regarding foster children's access to education and health facilities.

Unlike Akresh, the authors do not have information on the household from which foster children originate. Consequently they cannot ascertain whether fostering represents an improvement or deterioration for the fostered child. The focus of the paper is rather on the fate of the receiving household, especially consumption, income and investment. The authors regard receiving a fostered child as an unanticipated increase in consumption needs – an additional mouth to feed – that reduces the household's capacity to invest. Fixed effects are used to control for unobserved heterogeneity. The authors show little concern for non-random placement of fostered children, as would be the case for instance if children (or their parents) could choose the fostering household, perhaps among a restricted set of relatives and friends. We do not know whether child fostering households differ in other respects from non-fostering households. For these reasons it is difficult to be certain that the observed effect should be regarded as resulting from an exogenous external shock. The authors also do not seek to check their results by testing whether, say, the birth of twins (a similar negative shock) or the early death of a child (a positive shock) has comparable effects on investment and consumption. To summarize, this is a valiant but only moderately convincing effort.

Output 3 is an ambitious attempt to identify determinants of civil strife and their effect on rural investment. In the 1990s Uganda suffered from guerilla warfare instigated by the Lord's Resistance Army (LRA), a vicious armed group long supported by Sudan and famous for its child abductions and forced enrollment. The guerilla has affected mostly the North of the country. Using Ugandan household data from 1992 and 1999, the authors start by constructing a community-level data set of civil strife and its potential determinants – e.g., distance from infrastructure, asset inequality, presence of cash crop, etc. They then regress civil strife on these potential determinants, using community-level data. The two surveys have a small panel component, which is used to investigate increases in civil strife over the 1990s. The authors used a reduced-form approach, basically using a range of observable community variables to proxy for factors thought to influence the supply and demand of civil strife. In the second part of the paper, the authors use household-level data to investigate the effect of civil strife on rural investment and non-agricultural enterprise startups. They find that civil strife has the anticipated negative effect.

This is one of very few papers on this important topic. It is obviously difficult to document the effect of civil strife given that the consequences often are disastrous. The special circumstances prevailing in Uganda in the 1990s make such investigation possible. Civil strife remained present throughout the decade but at a relatively low level. Although the LRA committed numerous acts of unspeakable barbarism (their leaders have now been indicted in front of the international penal court), the number of these incidents remained sufficiently small and the protection of the army was sufficiently strong to enable life to go on in the affected communities.

While these facts make identification of the effect of civil strife possible, they do not eliminate estimation difficulties – e.g., civil strife as a treatment is not randomly assigned, as the authors document in the first part of the paper, and some of the factors favoring investment (e.g., the presence of a cash crop) also seem to attract guerilla attacks. At the same time, other factors (e.g., distance from infrastructure) would probably discourage investment but are associated with more civil strife. The authors do not seek to address these difficulties but rather present the regression results as they come out of the data. I also find the paper insufficiently informative about the way the data are handled. For instance, it is unclear from Tables 4 and 5 what the sample population is – one set of regressions has 9,248 observations while the other only has 1,248, even though they are both supposed to be on the same household population. Without knowing more about the data, it is unclear whether more could have been done. But as it stands the paper reads a bit like a missed opportunity: much more could have been done to address endogeneity and selection issues.

Summary: Average on research quality.

<b>S</b>	Superior
<b>AA</b>	Above Average
<b>A</b>	Average
<b>BA</b>	Below Average
<b>U</b>	Unacceptable

	<b>S</b>	<b>AA</b>	<b>A</b>	<b>BA</b>	<b>U</b>
<i>Please rate the following aspects of the project outcomes (or in the case of ongoing projects, design and intermediate outputs)</i>					
<b>Topics</b>					
Importance of the issues addressed		<b>x</b>			
Clarity of the project focus and stated objectives			<b>x</b>		
<b>Analysis</b>					
Theoretical/conceptual framework			<b>x</b>		
Empirical application			<b>x</b>		
Statistical and econometric methods				<b>x</b>	
Use of existing knowledge and resources			<b>x</b>		
<b>Data</b>					
Awareness of other data sources					
Data compilation, cleaning and cataloging					
Survey design and sampling					
<b>Output Quality</b>					
Writing quality (clarity, organization, etc.)			<b>x</b>		
Clarity of conclusions and recommendations			<b>x</b>		
Extent to which conclusions are based on analytic evidence				<b>x</b>	
Appropriateness of the recommendations			<b>x</b>		
Appropriateness of output form (working paper, book, database, etc.) for intended audiences		<b>x</b>			
Availability of translated outputs where appropriate					
<b>Extent to which research:</b>					
Increases knowledge and understanding of the issues			<b>x</b>		
Provides a sound basis for policy			<b>x</b>		
<b>Actual or likely impact of research on:</b>					
Government policy			<b>x</b>		
Future analysis			<b>x</b>		
The development community in general				<b>x</b>	
<b>Overall Quality of Research</b>			<b>x</b>		

## Land Policy Studies – Klaus Deininger

This large research project spread out over a number of years examines a host of issues surrounding land policy in developing countries. The objective of the project is to update the received wisdom regarding the land market and land access. This project has produced an impressive number of published and unpublished articles (34) as well as a book.

In addition to the output geared toward a research audience, the project organizers point out that the project has also influenced the Bank's activities in a number of ways. They write: “The PRR is perceived by Bank staff and other stakeholders as a key statement of Bank policy in the land area and which has helped to provide inputs into a number of other documents, most recently the guidelines for acquisition of land in Bank projects. WBI has used the output from the activity to design, together with DEC, a 4-day training course on land policy and administration which was conducted 8 times to between 5-10 countries each by video-conference and which is now an element of WBI's regular course offerings. A number of papers targeting more specifically a policy making audience has been developed and the research has directly fed into three pieces of ESW and innovative operations in the regions, most recently a series of loans to improve land registration systems in African countries so as to create a more favorable environment for private sector development.”

Among the many published and unpublished outputs coming out of the research project (35 in total), five outputs were submitted for evaluation:

1. Klaus Deininger. 2003. *Land Policies for Growth and Poverty Reduction*, OUP/The World Bank. 250 p.
2. Klaus Deininger and Raffaella Castagnini. 2005. “Incidence and impact of land conflict in Uganda.” *Journal of Economic Behavior and Organization* (forthcoming).
3. Klaus Deininger and Songqing Jin. 2005. “Tenure security and land-related investment: Evidence from Ethiopia.” *European Economic Review* (forthcoming).
4. Klaus Deininger and Songqing Jin. 2005. “The potential of land rental markets in the process of economic development: Evidence from China.” *Journal of Development Economics*, 78(1): 241-70.
5. Hanan Jacoby, Guo Li and Scott Rozelle. 2002. “Hazards of Expropriation: Tenure Insecurity and Investment in Rural China.” *American Economic Review*, 92(5): 1420-47.

Output 1 is a single-authored book surveying the recent literature on land policy in developing countries. The book offers a detailed conceptual framework of the many efficiency and equity issues surrounding agricultural land. It presents an in-depth survey

of the empirical literature covering all regions of the world. The end of the book is devoted to a detailed discussion of policy issues. Much of the emphasis is placed on market-based solutions to allocation and equity issues.

It is well known that institutions affecting the management and distribution of agricultural land vary dramatically between different regions of the world. Prior to this book, much the literature was torn between regional viewpoints that were driven by the dominant issues in a specific part of the world. The main advantage of the book is that it covers all regions and is not overly influenced by one particular outlook. Land issues are presented in a way that can be equally understood by policy makers in Latin America, Africa, India or China. Furthermore, by drawing lessons from different regions, it is a valuable source of inspiration for all. The book helped redefine land policy for the Bank, bringing it back on the policy agenda after years of neglect. For this alone the book should be praised. Although the book does not make a theoretical or methodological contribution, the literature review that it represents can be used as a pedagogical tool in undergraduate and graduate development programs.

Outputs 2 to 4 are just a small selection of the many papers written, alone or with various co-authors, by the lead researcher in this project. Output 2 documents how a change in the Ugandan law failed to reduce the incidence of conflicts over land. It argues that female-headed households – divorced and widowed women – are the main victims of land conflicts as their right to land is not recognized in traditional law. Output 3 examines whether tenure security affects land investment in rural Ethiopia – primarily the planting of trees and the building of terraces. The authors find that more secure transfer rights are unambiguously investment-enhancing. They argue strongly in favor of the government providing more secure land tenure. Output 4 examines the allocation of agricultural land in China. It contrasts two methods: administrative reallocation and land rental markets. It concludes that both mechanisms reallocate land to those with lower endowments but rental markets are more effective in achieving an efficient allocation of land.

These papers have in common a strong belief in the efficiency enhancing virtues of market forces and therefore in the need for institutions securing property rights and reducing transactions costs in land transactions. They also expose some of the equity issues surrounding land, such as access to cultivable land by women and by landless farmers. As such they make a useful contribution to the literature.

Output 5 is a top journal publication on the effect of tenure insecurity on manure application in China. The authors show that higher expropriation risk reduces manure application, guaranteeing land tenure would yield only minimal efficiency gains. The paper is carefully done, relying on hazard analysis and fixed effects to identify the effect of tenure security on manure application. This is obviously a good paper, with a thorough treatment of the evidence and a balanced conclusion.

Summary: An excellent research project that has contributed to both knowledge and policy making in a major way.

<b>S</b>	Superior
<b>AA</b>	Above Average
<b>A</b>	Average
<b>BA</b>	Below Average
<b>U</b>	Unacceptable

	<b>S</b>	<b>AA</b>	<b>A</b>	<b>BA</b>	<b>U</b>
<i>Please rate the following aspects of the project outcomes (or in the case of ongoing projects, design and intermediate outputs)</i>					
<b>Topics</b>					
Importance of the issues addressed		<b>x</b>			
Clarity of the project focus and stated objectives		<b>x</b>			
<b>Analysis</b>					
Theoretical/conceptual framework			<b>x</b>		
Empirical application			<b>x</b>		
Statistical and econometric methods			<b>x</b>		
Use of existing knowledge and resources	<b>x</b>				
<b>Data</b>					
Awareness of other data sources		<b>x</b>			
Data compilation, cleaning and cataloging					
Survey design and sampling					
<b>Output Quality</b>					
Writing quality (clarity, organization, etc.)			<b>x</b>		
Clarity of conclusions and recommendations		<b>x</b>			
Extent to which conclusions are based on analytic evidence			<b>x</b>		
Appropriateness of the recommendations		<b>x</b>			
Appropriateness of output form (working paper, book, database, etc.) for intended audiences		<b>x</b>			
Availability of translated outputs where appropriate		<b>x</b>			
<b>Extent to which research:</b>					
Increases knowledge and understanding of the issues		<b>x</b>			
Provides a sound basis for policy		<b>x</b>			
<b>Actual or likely impact of research on:</b>					
Government policy		<b>x</b>			
Future analysis			<b>x</b>		
The development community in general			<b>x</b>		
<b>Overall Quality of Research</b>		<b>x</b>			

## Can Africa Claim the 21<sup>st</sup> Century?

This flagship publication is the outcome of a team exercise involving a large number of participants from within and outside the World Bank. The 250-page report identifies the challenges facing Sub-Saharan Africa at the beginning of the 21<sup>st</sup> century and proposes a multi-pronged approach to put Africa squarely on the path to growth and poverty alleviation.

There are many good things to be said about this report. First of all, it is the fruit of a collaborative effort between the World Bank and some of the best economic research institutions in Africa itself – namely AERC, UNECA, and ADB. The report is at the same time frank and upbeat, identifying the failures of the past and the numerous constraints of the present, but delineating a path that leads to an integration of the African economies in the growing international economy. The report clearly recognizes that poverty alleviation in Africa will not be accomplished without growth and that growth will not be accomplished without more emphasis on trade and foreign investment.

These are important achievements in their own right, as I can still remember the time not so long ago when Africa saw foreign investors as enemies and trade with the rest of the world as something to wither away from. Having the most important think tanks in Africa sign off on this agenda is essential for future policy implementation. The fact that this report could be written without all the recrimination that accompanied earlier efforts (e.g., the Berg report in the early 1980s and the sustainable development report in the early 1990s) is by itself a fantastic achievement. It signals the beginning of a new era. What we need now is for these recommendations to influence policy making, as they have begun to do in a number of countries.

Being essentially a document for public consumption aimed at donors and governments, it comes as no surprise that the report fails to satisfy the demanding standards of research output. While the report contains a wealth of useful information and tables, it remains a consensus piece and as such fails to have any of the sharp edges that researchers expect from exciting research. The list of policy recommendations, for instance, is as in previous similar reports a long laundry list of things to do, with little discussion of sequencing and little prioritization. It is doubtful that the poorest countries in Africa would be able to implement all the suggestions contained in the report. But the merit of the report, in my view, is to clearly state that poverty alleviation cannot be accomplished without fostering growth, and that this means investing in infrastructure, improving agricultural technology, and diversifying exports. This may be obvious to most economists, but it is refreshing to hear the Bank say this after a decade spent courting NGOs and investing heavily in welfare and micro-development programs.

While the authors of this report have made an outstanding effort to avoid the overtly neutral tone of earlier similar reports, it cannot entirely avoid paying homage to standard clichés. Undoubtedly this is the result of the process by which the report was written –

i.e., presentation in front of numerous audiences, each of which has a different constituency and opinion. Instead of exposing disagreement, the report adopts the standard solution that consists in mentioning all relevant opinions as valid. Sometimes the desire to satisfy everyone takes over.

Let me illustrate this by picking – a bit unfairly – on a single sentence. On page 2, paragraph 2, the authors state that “Women are one of Africa’s hidden growth reserves, providing most of the region’s labor.” I want to take issue with the statement. It is certainly true that African women – like women elsewhere in developing countries – work hard and make a fundamental contribution to the livelihood of their households. But to say that they provide most of the region’s labor is probably not factually correct. The claim has been around for a while but the reality is that there are enormous differences in the gender division of labor across African regions, so that this generalization is unwarranted. Moreover, it is implicitly based on treating household chores as work. While I would personally favor this approach, it nevertheless is in contrast with the way labor market participation is defined in developed economies – and hence could be misinterpreted. Finally I am not sure that African women are a hidden growth reserve: in most developed economies growth has initially coincided with a withdrawal of many women from the labor market so as to focus on the provision of home goods. Similar tendencies can be observed in South Asia, for instance. Personally I am strongly in favor of educating girls and boys equally, but it is not because I believe that all girls will productively use the skills they have learned – it is because I believe that equal education is essential for equity reasons. Letting the reader believe that Africa could grow by educating women is, at best, wishful thinking and, at worst, fallacious. I do not want to overlabor this point but it illustrates the difficulty of writing a consensus report that satisfies everyone. Having said this, I find that this particular report succeeds much better than other reports I have seen in avoiding as much as possible language of this kind.

Summary: An unusually good general consumption report that signals a new consensus on Africa – a real achievement.

## **Questionnaire for Flagship Projects**

### **A. Objectives**

1. Are the topics and objectives critical for policy in the developing or post-socialist country? *YES.*

### **B. Design and Implementation**

1. Does the project reflect awareness of existing knowledge from other research available at the time and does it adequately reflect a good understanding of the country(s) in question? *YES.*
2. Is the project design appropriately anchored in a sound and coherent analytical and empirical framework? *YES.*
3. Are the conclusions consistent with the research findings? *In general YES. Were problems/concerns noted and reviewed? See above.*

### **C. Accessibility**

1. Were project reports and publications presented in a manner appropriate for and accessible to the intended audience(s)? *YES.*
2. If applicable, are policy recommendations commensurate with findings? *Well, as always policy recommendations are speculative.*

### **D. Results and Cost-effectiveness**

1. What are the key findings of the study? *Given that the report is basically a literature review, there aren't new findings per se.*
2. How do they advance country policy in the field? *I have no idea since I have not been provided with an impact assessment.*

## Questionnaire for draft report on strengths and weaknesses of Bank research

1. In your area of expertise, has the Bank made a significant contribution? *Yes.*
2. In your area of expertise, has Bank research focused on the most important policy issues for developing countries? *Pretty much.* Do you feel Bank researchers have appropriate incentives to allow identification of the most important issues? *I think so.*
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? *Yes to both for a substantial proportion on Bank researchers, but obviously not for all of them.*
4. Has the Bank made appropriate selection of topics for data collection and surveys? *The Bank makes an essential contribution to data collection in developing countries. In my own work I have found it very receptive to new ideas and concerns. The main issue today is integration of random assignment and impact assessment in project evaluation. Have data and surveys been well-maintained and disseminated? Up to a point. Sometimes the Bank accepts countries' desire to keep the data inaccessible even when there is no reasonable ground for it.*
5. In your area of expertise, how could future Bank research better serve developing country objectives? *I am not sure what this question means because the objectives of developing countries often include things that the Bank – or the research community – is not keen to condone, such as military supremacy over neighbors, political control over the population, and preservation of the status quo. Much of the difficulties facing development organizations such as the World Bank are precisely that their objectives – e.g., to help the poor – are not aligned with those of the recipient government. Short of taking over the country or pushing for a change of government (as the donor community has sometimes done, for instance in Kenya, with doubtful results), I do not see what can be done about it. I have never seen conditionality as a solution because in the long term it weakens the recipient government and because donors are not accountable to the local population and thus cannot be held accountable if conditionality messes things up, as it regularly does.*
6. Particularly for ongoing projects that you have evaluated, please comment on the proposal review, revision and selection process. *These issues are addressed in the first part of my report.*
7. What is your overall assessment of Bank research? *The Bank's Research Department has a very difficult job to do: to produce research that is good enough to impress the academic community – and hence to have a durable effect on graduate teaching in top universities and consequently on long-term policy – but also to assist Bank operations in the short and medium term. It is constantly torn between these two objectives, with some staff members better at reaching out to the academic community and others better at reaching out within the Bank and to the donor community. My sense is that the Research Department is doing pretty much as good a job as could be expected, given the circumstances. Could it produce more top quality research? Probably, for instance by taking in PhD students on doctoral fellowships and involving them in data collection, as IFPRI does. It could also be more aggressive in hiring senior academic researchers into true research positions. It could also rotate less productive researchers faster toward Bank operations. The management of DEC will have a better sense than me of whether these options are feasible, the risk being that they undermine the legitimacy of the Research Department itself and hence its very existence. The one suggestion that seems to me both easy to implement and potentially useful is to create a doctoral fellowship*

*program. The other is to build more research into project evaluation itself, by emphasizing more random assignment and impact evaluation.*

wb18913

L:\DECRS\Evaluation of Research\Evaluator Reports\Evaluator Reports FINAL VERSION\Marcel Fafchamps.doc

11/21/2006 12:29:00 PM