

Evaluation of World Bank Research

Francesco Caselli*

1. Investment Climate Assessment Program

This is by far the biggest of the research programs I have been asked to assess, and indeed it is an enormous undertaking involving a vast number of people, over many years. In practice, I understand that this project has absorbed a significant majority of the resources of the macroeconomics group over the last four or five years. The core of this program is constituted by a series of firm-level surveys carried out in more than 80 countries since 1999 and ongoing, so that the wealth of mounting information is truly massive. Firms in the survey provide a wide variety of data, including a wide range of inputs and outputs, performance, investment, product range, corporate governance, technology, as well as answers to a number of questions on the financial, infrastructural, institutional, or administrative constraints they face. The surveys, therefore, offer a detailed micro-economic picture of the characteristics of production units in developing countries, as well as the “investment climate” in which the firms operate, with a view to identifying reforms that are more urgent and more likely to lead to faster firm and productivity growth. Accordingly, as part of the output from this program there are several papers that analyze the data and provide policy advice.

I have no idea how much this cost, so in some sense I am unable to say whether engaging in this massive exercise has made sense or not. With that important caveat, I am a big fan of this initiative. It is entirely consistent and may even have contributed to a heightened interest in academic research on development on firm dynamics, as well as on firm heterogeneity. Macroeconomists interested in development increasingly recognize that it is difficult to think of aggregate investment, technological progress, occupation, and growth without understanding the details of firm-level decision-making, and particularly the constraints firms face. With its hugely influential Living Standards Measurement Surveys (LSMS) the World Bank has been a pioneer in generating internationally comparable surveys of consumers and workers in developing countries. The Investment Climate Surveys promise to be the firm-level equivalent, and as such they are warmly welcome. Incidentally, it is no surprise that both initiatives belong to the Bank. Only the World Bank has the resources, expertise, and infrastructure needed to carry out so many surveys in so many countries over a protracted period of time. Hence, these surveys show starkly that the World Bank has an irreplaceable role to play in research on development and development policy.

I think it would have been highly unrealistic to expect the people responsible for initiating this massive progress to get everything “exactly right” from the start, and accordingly there are some teething problems with the surveys. In particular, I learn from people who have looked at these and other data more closely that there are questions of representativeness for at least some of the country surveys. In some countries there seems

* I gratefully acknowledge the assistance of Zsofia Luca Barany and Ines Moreno de Barreda in preparing this report.

to have been no industrial census on which to base the sampling frame. Perhaps more worryingly, for those countries where censuses are available, several summary statistics (e.g. size distribution of firms) seem to differ substantially in the census data and in the IC survey. I think it is absolutely critical that future waves of the survey find ways of improving and validating the representativeness of the samples. In the meanwhile, I would strongly recommend that the public-use data sets made available to the public contain country-specific quality indicators reflecting the thoroughness of the underlying sampling framework (so that users can choose to work only with the countries featuring high-quality data), and — when available — sampling weights. Research using the data should also flag results obtained from samples whose representativeness is uncertain.¹

The survey instrument itself is, of course, somewhat tentative, but to the best of my judgement it appears to have been reasonably well thought through. For sure the information collected from respondents is truly, indeed astonishingly, comprehensive.

An impressive aspect of the project is the effort undertaken to make the data public. The authors have set up a web page where outside researchers can not only compute summary statistics from the data set, but indeed perform a variety of econometric exercises, such as linear regression, probit, and two-stage least squares. Of course if one wants to write an academic paper using these data, it is still likely that he or she will need to physically travel to DC (or get a World Bank co-author), but considering the obvious confidentiality issues involved with firm-level data, the web-page idea is in principle a very generous effort at diffusion. As far as I can tell, however, there is still a certain selectivity in the variables posted in the web page — I could not find some of the things I expected based on the questionnaire — so the generosity for now is more apparent than real.

As mentioned, the program extends beyond the data collection effort to a series of papers that analyze the data. I think it is important here to note that the papers I was given as samples from this effort reflect a very first pass at looking at the data. I think that subsequent researchers, with the benefit of the early experience, are figuring out more compelling ways of using the data.

Briefly, I see a tendency in the papers I have seen to (i) over-emphasize the answers to the self-assessment questions on the constraints faced by the firm, and (ii) rush into sticking such “investment climate” variables in the right-hand side of regressions for firm-level TFP or firm-level labor productivity. Many economists, including several of my fellow panel members, tend to dismiss questions that involve a large amount of subjective assessment and interpretation, and are therefore quite critical of the section of the questionnaire where respondents are asked, for example, to rank several problems by the severity of the constraints they impose on the firm’s growth prospects (e.g. corruption v. power failures). I am not as dismissive as my colleagues, and indeed I think these

¹ Given the existence of industrial censuses in many countries, there is of course a certain amount of duplication in the information collected in the IC surveys. But the information collected in the typical census is only a very small subset of the information collected in the IC program. And of course there are many countries without a census.

questions are interesting and potentially useful, as I will try to argue below. Still, the critics have a point, and it is therefore puzzling that authors who have had access to the data have focused nearly exclusively on these variables. There is so much more in the surveys than that!

Coming to the tendency to use the “investment-climate” variables on the right-hand side of productivity or growth regressions, as pointed out by Carlin et al. (2006), this is unlikely to generate particularly valuable policy advice. If the regression is run at the firm level, for example, one may get a “puzzling” positive coefficient on, say, corruption. Indeed such seemingly perverse results are pervasive. But this may just reflect the fact that more productive firms are more likely to chafe at the constraints imposed on them by corrupt officials (or by lack of infrastructure, high taxes, etc.). If instead the regression is run at the country (region, city) level, all the familiar endogeneity problems of cross-country empirical work raise their ugly head (as in: richer cities can fight corruption more effectively). Some of the papers try to obviate by using an intermediate strategy: firm-level left-hand-side variables, and regional or citywide right-hand-side ones. But the local averages on the right may (and my guess is that they will) still correlate with other local-level omitted variables.

This does not mean that the data are useless: far from it. Carlin et al. (2006) point to at least two ways in which the data can be made to be informative for the purposes of policy design. First and foremost, there is a tremendous amount of information in simple sample statistics. Firms are asked to rank the features of the investment climate in terms of the constraints they pose on their profitability. Hence, for each country we know if the average firm feels more constrained by, say, corruption, or by the lack of telecoms infrastructure. I think this simple information, coupled of course with an assessment on the relative costs of different reforms, is much more useful as an input into prioritizing a country’s reform strategy than the econometric evidence I alluded to before. Incidentally, this is an example where the journal-publication game reflects negatively on our profession: the authors may have thought that only papers with formal econometric work would be considered for publication. Unfortunately I am afraid they might be right about that.

The other potentially fruitful way of using the data illustrated by the Carlin et al. (2006) paper is to use the investment-climate indicators as left-hand-side variables in firm level regressions.² As they explain, this exercise tells us which type of firms are more likely to be most constrained by various systemic inefficiencies. They show, for example, that more efficient firms are more likely to feel constrained by customs regulations, while less efficient firms are more likely to complain about lack of access to external finance (which, as they point out, is exactly as should be). Knowing which reforms are most important for the most efficient firms may be another useful input in prioritizing reforms. Indeed, I think the two exercises performed by Carlin et al. could usefully be combined

² To be fair, one of the papers I read does indeed do some of this.

by simply constructing weighted summary statistics for the most important constraints, where the weights are proportional to indicators of firm efficiency.³

But, to reiterate, the biggest payoffs from the surveys may come from other sections of the questionnaire, where firms are asked to give detailed and objective answers on their operations and governance. As I mentioned, there is an increasing supply of models that build up to macro outcomes from a distribution of (typically heterogeneous) firms, whose behavior is micro-modeled. Firm level data sets from developing countries could help calibrate and test these models, and boost tremendously this line of research.

2. Program on Capital Subsidies and the Quality of Growth

This appears to be a conference book, as the three papers I have been given under this heading are titled Chapter 1, 4, and 5, respectively. I have a hard time evaluating this program as a whole because the chapters seem to be about very different topics, and I could not find an Introduction or explanatory note to gain a better picture of what the book is broadly intended to accomplish. Hence, I have to fall back on commenting on the merits of the three chapters in their own right.

Briefly, Chapter 1, “The Dynamics of Unbalanced and Sustainable Growth,” is a theoretical study of a multi-sector growth model, featuring a structural transformation out of agriculture. This is certainly an important topic, but the literature on it is enormous, and it is a pity the authors did not do more to relate their work to it and highlight what is new. I suspect one fairly new ingredient is that there are possible investments into replenishing “natural capital,” which may stand for policies aimed at restoring or protecting the environment. Nevertheless, I am frankly unable to say whether the theory develops any compelling original insight, because some of the assumptions are not clearly stated, others are quite dubious, the terminology is non-standard and confusing, and the derivations of the results are very hard to follow.

Chapter 4, “Why Governments Should Stop Non-Social Subsidies,” is also somewhat puzzling. This paper presents some cross-country estimates of the agricultural production function using a sample of Latin American countries. The focus is on publicly provided inputs, which are broken down into two broad categories: social subsidies and other government spending on the rural economy. Needless to say, the question of the relative productivity of different types of government spending is incredibly important. Unfortunately, the analysis here suffers from the standard identification problems associated with cross-country growth empirics. The authors do try to use instrumental variables, but the over-identification test rejects the validity of the instruments. Furthermore, interpretation of the results is made difficult by the fact that the authors

³ An important cautionary note is that the investment-climate survey, even when well used, cannot be the only source of information on institutional constraints faced by firms. In particular, almost by construction the ICS has virtually no information on barriers to entry or other costs to setting up a firm. Many economists believe that the limits to competition associated with barriers to entry are an important drag on growth and development.

enter total spending, and the share of non-social subsidies in total spending, in the regression. It is then not easy to map the coefficients into the elasticities of output to social subsidies and other spending, much less into marginal products. For these reasons, the conclusion that at the margin governments should cut non-social spending and increase social spending is unwarranted.

Chapter 5, “Distribution of Opportunities Key to Development,” appears to be intended as a discussion of existing evidence on cross-country patterns in the inequality of educational achievement. The central fact, according to the authors, is that the education Gini coefficient is strongly decreasing with the level of development. The authors assign great significance to this fact, but the paper itself suggests three reasons why one should be skeptical. First, a simple theoretical model shows that, if anything, the welfare-relevant measure of inequality of education is the variance, or standard deviation, not the Gini coefficient. Second, and related, the standard deviation behaves very differently from the Gini coefficient: it is inverted-U shaped in development instead of decreasing (the authors refer to this as a “myth,” but it is no more of a myth than the decreasing pattern of the Gini coefficient). Third, there are good reasons to believe that at least in part the observed patterns of the Gini coefficient may be driven by a mechanical association with the average level of education. Perhaps more importantly, the paper derives a series of policy implications from the evidence on the Gini coefficient that are entirely unwarranted.

3. Smaller Programs

3.1 New Dimensions in Public Expenditure Analysis

Under this heading I was given a paper titled “Evaluating Public Expenditure Management Systems: ...,” which has a very ambitious goal: to create an analytical framework that could serve as a common basis for comparisons between public expenditure reforms introduced in different countries. Establishing a more systematic way to analyze public-expenditure management systems is critical for policy making in developing and post-socialist countries, and hence a worthy goal for World Bank research.

The most useful part of the paper is the discussion identifying the key problems faced by the designers of public expenditure systems, highlighting the possible institutional arrangements that address these problems, and listing mechanisms that can enhance the accountability and the transparency of these arrangements. I learned quite a bit from this discussion, although in some cases some of the institutional mechanisms advocated in the paper are insufficiently motivated, and their downside or negative spillovers vis-à-vis other goals of the government are not fully recognized.

I was much less convinced by the part of the paper devoted to the construction of quantitative indexes aimed at evaluating expenditure-management systems in terms of their fitness to deliver on the goals set out in the first part. This exercise, together with the

application of the proposed index to the cases of Australia and New Zealand, made me even more skeptical of attempts to build time-series or cross-sectional summary measures of institutions than I already was. The authors are forced to make all sorts of hopelessly arbitrary assumptions, as well as impossible-to-defend choices in weighting and aggregating the data. Furthermore, they are also forced to resort to inevitably biased sources to fill in their questionnaires. I don't think this is helpful and I would not encourage the World Bank to generate a new summary "measure" of institutions that people can stick on the right-hand side of growth regressions.

3.2 China's Long-Run Fiscal Sustainability

Under this program heading I found a paper titled "Simulating the Long-term Effects of China's Expansionary Fiscal Policy in an Overlapping Generations Model." The main objective here was to quantify the growth and welfare effects of China's recently adopted expansionary fiscal policy. The most important question is how the chosen method of financing this policy affects the medium and long-run growth path of the economy, as well as welfare. The methodology is based on taking a leading academic framework (based on the work of Auerbach, Kotlikoff, and others) and applying it to the Chinese economy in order to simulate the effects of different budget balancing techniques. I think it makes a lot of sense for World Bank researchers to adapt leading academic frameworks to ask specific policy questions, as in this case, so I am very sympathetic to the spirit of the paper. Also, the topic and the method of the paper are of obvious relevance for the Bank, which needs to give policy advice to the Chinese fiscal authorities.

However, evaluating the success of the study is a difficult task. One problem is that the authors seem to have made insufficient efforts to adapt the original model, and its calibration, which was based on the US economy, to the special circumstances of China. Another problem is that the policy experiments examined with the help of the calibrated model are never explained with sufficient precision. Finally, I am under the strong impression that the ranking of policy options in terms of their welfare implications is built-in, and made rather trivial, by some arbitrary assumptions on the productivity of government spending and the relative degree to which different taxes are distortionary. This assessment is tentative because the paper is not clearly written.

3.3 Highlighted Research

Under this heading I was given to read a paper titled "Current Accounts in Debtor and Creditor Countries," which appeared in the *Quarterly Journal of Economics* in 2000. The paper presents a powerful critique of the standard approach to the current account in macroeconomic models. In standard models a country experiencing a positive, temporary income shock will save almost the entire amount of the windfall. If the productivity shock is short-lived, domestic investment will not respond very much, so that most of the additional saving will go abroad, and the current account surplus will increase.

This paper's basic observation is that, if domestic and foreign assets are imperfect substitutes (something not contemplated in the standard approach), such a response

implies a marked change in the risk and return structure of the domestic country's residents' portfolios. They argue, convincingly, that domestic residents will resist such a substantial change to their risk exposure. Pushing the argument further, the authors reach a startling conclusion. Under some assumptions, current account surpluses will only go up in response to a positive income shock in countries that already start out with positive net foreign assets, while they will deteriorate in countries that start out with a negative foreign asset position.

The intuition is that domestic residents will allocate their new savings so as to preserve their percentage asset allocation between domestic and foreign assets. This implies that if they were net creditors, they will be happy to continue to be creditors, which means that most of the extra savings are indeed invested abroad. But if domestic residents are net debtors to start with, they will not want their investments abroad to turn them into creditors, and will therefore invest most of the windfall domestically. Aside from a reasonably high degree of imperfect substitutability between domestic and foreign assets, the main additional requirement for this reasoning to go through is that the extra domestic investment is not associated with too rapid diminishing returns. An empirical section presents some results showing that the alternative framework proposed here works at least as well as the traditional one (admittedly, that is not a high standard to meet).

This is exciting work and the World Bank should obviously be proud of it. Of course this is highly abstract thinking, and we are far removed from direct policy applicability. Nevertheless, it is probably the case that — to effectively operate as a bridge between academia and the policy world — World Bank researchers, or at least some of them, should be able occasionally to produce papers of the highest academic quality. This paper certainly fits this bill.⁴

4. Flagship Publications

I was sent five book-length monographs under this heading. Two of them, *Globalization, Growth, and Poverty*, and *Assessing Aid*, are part of the Policy Research Report series; another one, *Inequality in Latin America*, is part of the Latin American and Caribbean Studies Series; and the last two, *Economic Growth in the 1990s* and *East Asia Decentralizes* seem to be stand-alone reports. While the books cover different topics, they share some common features. They tend to be written by teams of World Bank researchers or, when written by only a few authors, they still reflect the work and ideas of a much larger group. They present little new formal theoretical or empirical work. Rather, they combine overviews of existing work done inside and outside the Bank with informal case studies and examples drawn from the Bank's first-hand experiences in different countries. Finally, unlike the other projects reviewed here, they are written to be accessible to the policy community and the public at large. Hence, their purpose is partly to reflect and draw policy lessons from existing formal research and cumulated informal experience, and partly to educate the non-research community on the latest thinking in

⁴ My assignment also included a program on "Investment Patterns," but the paper I was given under this heading is the same as Chapter 1 under the "Capital Subsidies" program.

the area of development. In this sense I see them as very similar in spirit and purpose to the World Development Reports, which the Bank publishes yearly and I regularly read.

I like this format and I think, on the whole, the Bank does this kind of thing quite well. The books may not break much new ground in themselves, but being unconstrained by the strictures of formalism, they are able to confront academic research with direct experience on the ground, which is an indispensable step in improving both research and policy. The World Bank has a clear comparative advantage in this sense, as its staff straddles the academic and the policy world like very few other economists do. I come back to this point below.

4.1 Economic Growth in the 1990s

The most ambitious and wide ranging of the flagship publications I have seen is *Economic Growth in the 1990s*, the purpose of which is nothing less than to provide an overview of the lessons offered by the developing world's growth experience in the 1990s. Since in the late 1980s and throughout the 1990s many countries experimented with a variety of structural and institutional reforms, the goal here is to learn what works and what doesn't, and why. The book's style is informal, and based at least as much on case studies and anecdotal evidence as on background formal econometric work, which I find entirely appropriate for a study of this nature. Through its economists deployed in the field, the World Bank is a repository of hands-on direct policy experience and observation around the world. This is particularly true with the recent reform wave, since the World Bank was one of the key advisers on most such reforms. We can therefore not only learn about the reform outcomes, but also about the country context and the specific reasons that motivated the particular choice of reform path.

Although the reforms of the 1980s and 1990s were often sold as silver bullets for rapid economic growth, the results were very heterogeneous, and in many cases unexpected. In some countries the expected benefits largely failed to materialize, in others they exceeded all expectations. The book's main ambition is to try to shed light on the reasons for these different outcomes.

Somewhat inevitably, the main conclusion of the analysis is that there is no unique formula for policy making. The same reform could lead to different outcomes in different countries and different policies could achieve the same objective. Which reform to choose in each country depends on a number of factors, such as the quality of its institutions and its investment climate. The authors suggest that the best strategy is to discern at each moment what the binding constraints on growth are, and give priority to implementing policies that relax those constraints. One could argue that simple common sense would have come to the same conclusion, but the point here is that the knowledge of facts and country experiences afforded by books such as this one make the implementation of this principle much more feasible. The strength of the book is in the accumulation and divulging of experience, not in any particular formula.

The book is well organized. It first presents an overview of economic growth in the 1990s, giving a highly readable introduction to the main facts and controversies of this decade, their possible interpretations, and the main lessons learned according to the authors. The subsequent chapters tackle in detail the main institutional and policy objectives pursued by the reformers: Macroeconomic Stability, Trade Liberalization, Privatization and Deregulation, and Financial Liberalization. Each case is illustrated with a wealth of examples from disparate countries. More specific case studies are presented in the country notes.

4.2 Globalization, Growth, and Poverty

The purpose of this book is to examine the impact of worldwide economic integration on developing countries. The focus is on the effects of globalization on poverty, inequality, and cultural diversity. There is a strong emphasis on policy, as the book builds up to an agenda for actions that will enhance the potential benefits from globalization and mitigate the risks it generates. The book is based on previous studies of the World Bank related to these issues. The main goal is to collect the different results found in those studies and to compile the information in a book accessible to the broader public. Consequently, the book's style is informal, and only the results, and not the details, of the econometric studies the book is based on are presented.

The main thesis of the book is that global economic integration is a powerful force for poverty reduction. However, the authors argue that more could be done within and across countries to maximize the benefits of the global market for developing countries. As mentioned, the book closes with an agenda for actions that could make globalization work better: Promoting a "development round" of trade negotiation that focuses on market access; Improving investment climate in developing countries; Improving the delivery of education and health services; Providing social protection tailored for a more dynamic labor market in an open economy; Providing a greater volume of foreign aid and ensuring that it is better managed; Providing debt relief; and tackling greenhouse gases and global warming.

The book is very clearly written, and it organizes effectively a wealth of empirical results and case studies from different sources. Only the chapters "Globalization and Culture" and "Globalization and the Environment" are somewhat weak and superficially treated. The survey of empirical results from the literature is balanced and reflects well the consensus "point estimates" on most issues. But I am a bit concerned that the style of the presentation does not give the public an adequate sense of "standard errors," around its claims. This is particularly so given the peremptory tone of the policy proposals in the book's agenda for action. It does tend to suggest to the general public that we know all the answers. The danger of this is illustrated by my comments on the next flagship publication.

4.3 Assessing Aid

One big danger of this kind of exercise is that the Bank conveys to the public an inflated sense of how much we know about what works in development. This is most clearly illustrated by the report *Assessing Aid*. At the core of this very well written and highly informative book there is empirical evidence that foreign aid leads to faster growth in countries that score high on a number of institutional and policy dimensions. This is a claim that the international policy community has embraced with great enthusiasm, and which has arguably led to a renewed focus on aid as a poverty-fighting tool. No doubt the clarity and persuasiveness of books such as this one have contributed to these developments. The problem is that subsequent research by, among others, Easterly (2003) and Rajan and Subramanian (2005) has showed that the claim is not robust to using more recent and comprehensive data sets. It is not that the original research was not good: it was in fact very good. It is just that now we have better data.

So what do we do now? The whole aid industry juggernaut (including the debt-forgiveness part of it) is in full swing, especially in Europe, and major political figures have embraced aid among their flagship policy proposals. A reversal seems politically unfeasible, and it is fair to expect that aid will be at the center of development policy for several years to come. This seems unwarranted given what we know now, and I blame in part the excessive confidence of reports such as *Assessing Aid* for what may prove a costly detour on the road to a world without poverty.

4.4 Inequality in Latin America: Breaking with History?

The main objective of this report is to encourage Latin American countries to reduce inequality, and to give advice on the way this could be done. There is an in-depth and extensive analysis of the roots, the nature and the reproduction mechanisms of inequality in Latin America, as well as a fairly compelling overview of its negative effects.

The policy recommendations are organized into three groups according to the principal target they seek to reach: inequality in access to productive assets (democratizing education, changing the distribution of land, securing property rights, broadening the provision of infrastructure); inequalities arising from distortions in the operations of financial, product, and labor markets; and lack of true progressivity in the social protection system. The authors also emphasize that these policy reforms must be accompanied by political and social change (clientelistic and weak states, ethnic and racial divisions, attitude of the elites, etc.). The discussion is made in the context of a simple, general theoretical framework of the circular interaction between the distribution of assets and opportunities, outcomes, power and influence.

The authors have done a good job at creating a coherent theoretical framework within which to organize the vast existing empirical and historical literature. The arguments are lucidly and clearly presented, and the conclusions seem sound. The book is very well organized, one of the nicest features being that all the chapters have the same structure. This makes it easy to find specific topics of interest.

4.5 East Asia Decentralizes

This is clearly the most technical of the “flagship” publications and the narrowest in scope. Its main objective is to summarize the knowledge accumulated from the decentralization processes of East Asian governments, as well as to evaluate its success. A trend to devolve more power to local levels of government started in the last decade of the 20th century, and is continuing today. The assessment of ongoing or completed decentralization processes can therefore be highly useful for policy makers. Since existing theoretical work gives little guidance on the optimal path to decentralization, the information gleaned in the field by successful decentralizers can be invaluable.

This report is based on the case study of six East Asian countries that have embarked on the process of decentralization. The report is organized into 12 broadly connected chapters, of which the first two are more general, while the remaining ten delve deep into specific topics of decentralization, and their management in the East Asian countries that form the object of the study.

The first two chapters give a general overview of decentralization itself. The first one identifies the driving forces behind this phenomenon, describes the different approaches and trends, and finally emphasizes the main challenges a decentralizing government must face. The second chapter gives a country-by-country summary of many more specific topics (the origins of decentralization, the structure of the subnational governments, the implemented decentralization policies, etc.) with concise and informative tables.

The remaining ten chapters look at this process from different angles: fiscal sustainability, regional disparities, public expenditure management, own-source revenues, etc. Since these chapters were all written by different authors and look at different topics, it is perhaps not surprising that the style, the structure and the quality of the chapters show a high variance. Some of the chapters are very clearly organized, first presenting the theory and then analyzing the reality of specific countries. Other chapters are less systematic and harder to follow. It would have been a good idea to have a heavier editorial hand in the book, and perhaps to impose a common frame to all chapters.

As for all collections of case studies, the policy recommendations are mostly fairly vague and indistinguishable from what simple common sense would suggest, but the value of the exercise is to make practitioners much more aware of the myriad details that a sound decentralization policy needs to address. The report may also serve as a source of information on who are the “go to” people, inside and outside the Bank, on specific countries and policy issues covered in the book. I would expect anyone with an interest in decentralization, whether for research or policy reasons, to find this book useful.

5. General Thoughts on World Bank Research

As an organization producing and supporting research, there are (at least) three things the Bank does: data collection; formal theoretical and empirical work, which I will henceforth refer to as academic research; and informal policy discussions blending theory, empirics, case studies, and anecdotal knowledge collected “in the field,” which I will refer to as experience-based policy discussion. There were examples of all three in the material I read. It should also be clear that the Bank has an absolute advantage in the first and the third. It has an advantage in data collection because of its resources, expertise, local knowledge, access to government, and infrastructure around the world. It has an advantage in putting together academic research with informal practical knowledge because its economists are both academically trained and exposed to the day-to-day realities of development policy through the policy portion of their work.

The quality of the work in these two “absolute-advantage” areas is, to the best of my knowledge, very high. Indeed, in these areas I think the World Bank is the World leader. In terms of data collection, I can think of many economists who are entirely dependent on data provided by the World Bank for their research. World Bank data figure prominently in my own research. The IC survey discussed above is just the latest example of a collection exercise that will undoubtedly have a first-order impact on research and thinking. A recent new cross-country data set on the value of natural capital is based on superb work and it is already attracting quite a bit of attention. And since there is still a lot that we do not know, I am looking forward to further data collection initiatives by the Bank. I think the marginal dollar spent in this area at the Bank still has a very high return.

In the area of experience-based policy discussion the Bank is also a near-monopolist. The UN has a similar global reach as the Bank, but its staff does not have the same level of technical expertise. The IMF does not specialize in development, as the Bank does. Other multilateral institutions have at most a regional scope. Individual academic economists have at best detailed knowledge of a few countries and topics, and in any case their incentives are skewed against providing informal work accessible to the general public or the policy community. Hence, I think that World Bank’s publications, such as the World Development Report or other flagship publications such as the ones I have discussed above, play a crucial role in informing and educating both the academic and, more importantly, the policy community. As I indicated, I find the quality of the work and the analysis therein quite good in general.

I actually have some stunning evidence that this educational work is very effective. *The Financial Times Magazine* of 3/11/2006 features a story on the Bill and Melinda Gates foundation. According to the story, when asked about his motives for setting up the foundation, <<[Bill] Gates said his moment of truth came when he read the 1993 World Development Report, a dry 333-page document full of tables and statistics ...>>. According to the same story, the Gates foundation spends more money on health in the developing world than the World Health Organization. If both statements are even

partially true, this single fact is probably enough to justify the entire budget devoted by the Bank over many years to what I called experience-based policy discussion!

My only concern with this area of Bank research is that some of it sometimes conveys an inflated sense of the solidity of our knowledge in certain areas, and it consequently offers peremptory policy advice that in some cases turns out to have been unwarranted when better data or methods are applied to the same question. I have already discussed this point in section 4.3 so I will not belabor the point here.

The Bank's absolute advantage and excellent returns in the areas of data collection and experience-based policy discussion might suggest that it should focus its resources in these areas and leave the third area of present activity, academic research, as it were, to the academics. But I think this would be the wrong conclusion. Being in close touch with developments in academia is probably important to know which data to look for, and to discuss policy intelligently, albeit informally. And I think it is sensible to argue that one cannot be a fully astute consumer of academic research without also being a (competent) producer. After all this is why universities have the same people doing academic research and teaching. Therefore I think it is sensible that the Bank engages in all three types of research-related work I mentioned above.

Turning to the quality of the academic-style type of work I would say that the most striking feature of it, both in the sample of papers I was sent, and more generally in my experience with reading World Bank research, is its enormous variance. There are many papers written at the Bank (as well as in the sample sent to me) that are as good as the work produced in the top economics departments. There is much work that, without being stellar, is solid and informative. However, there also appears to be a pretty long and fat tail of work that is frankly of mediocre quality. The worst work is surprisingly bad: poorly written, poorly motivated, and poorly executed. Hence, I think one important question is whether the marginal dollar spent on academic-style work is yielding returns that are comparable to those in the other two areas. I suspect not.

Of course I am not advocating anything as coarse as simply slashing the budget for academic research, as this would not in itself necessarily lead to a change in the ratio of poor-quality work to good-quality work, and as I mentioned the good-quality work is really good. But there may be other ways to improve things. One area where I see potential is in the money spent on non-staff researchers. Some of the poorest papers I have read are actually not by World Bank staff. I think this particular part of the bottom tail should be relatively easy to cut off, perhaps by adopting some objective criteria for funding of external contributors (e.g. publication record) to ensure that their work tends to meet professional standards (provisions/exceptions would have to be made for young researchers).

Turning to work by staff members, the high variance of the papers may be a sign of managerial failure. I would have thought that a key managerial responsibility at the Bank would be to finely dose each staff member's assignments according to their comparative advantage. Those who are less suited for academic research should probably be steered

toward other, equally important, tasks. Perhaps once again some objective criteria, such as one's publication record, may be used more heavily for this purpose. Also, it may be worthwhile to review the internal incentive structure: it would for example be very problematic if writing papers of an academic nature was considered to be a key factor in promotion decisions. Finally, I wonder whether there may be issues with hiring. By way of providing a benchmark, it is fair to say that the worst LSE students produce work that is substantially better than the worst work produced at the Bank. Since our worst students cannot get jobs at the World Bank, I am left wondering whether some improvements in hiring policy may not help (an easy improvement is to automatically hire all LSE applicants :).

References

Carlin, W., M. Schaffer., and P. Seabright. 2006. "Where Are the Real Bottlenecks? A Lagrangian Approach to Identifying Constraints on Growth from Subjective Survey Data," working paper, UCL.

Easterly, W. 2003. "Can Foreign Aid Buy Growth?" *Journal of Economic Perspectives* 17(3), 23-48.

Rajan, R., and A. Subramanian. 2005. "Aid and Growth: What Does the Cross-Country Evidence Really Show?" IMF working paper.