

A Review of World Bank Research on the Financial Sector and Private Sector Development, 1998-2005

Jonathan Morduch

Financial Liberalization

Project 102

Main output: Book: Caprio, Honohan, Stiglitz, *Financial Liberalization: How Far? How Fast?* Cambridge University Press, 2001. (Esp. ch. 2, 6)

A. Objectives

1. What specific questions or hypotheses did the study seek to answer?

The volume investigates failed liberalizations (and one success: Uganda). The book questions whether the costs of laissez-faire approaches to financial markets (via the volatility induced) outweigh the gains in static efficiency. While liberalization is often necessary, the question is, with what speed and in what sequence?

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

The volume documents failed liberalizations in India, Indonesia, Korea, and Mexico, among other countries. The costs to their economies have at times been huge (the Russian failure was particularly devastating), both for businesses and poor households alike. This is a first-order problem for economic policymakers.

B. Design and Implementation

1. Was the methodology appropriate and well-executed? Was it innovative?

The studies employ varying methods, including country case studies and cross-country regressions. The case studies tend to be more useful, but the cross-country studies helpfully show increases in real interest rates that emerge broadly in the wake of failed liberalizations.

None of the studies really grapples with the political and institutional bases for failure. As such, they fail to explain the dynamics that drove the policy debacles – was it misguided thinking? Poor empirical predictions? Political dynamics? The way those questions are answered can help policymakers form priorities for avoiding future catastrophes.

2. 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?

The chapters implicitly (and sometimes explicitly) critique laissez-faire approaches that minimize the need for financial regulation. Given how seriously those laissez-faire ideas are taken (or were at the time), particularly among some at the IMF and other major financial institutions, it would have been helpful to have a more direct engagement with those ideas. The issues have been the source of considerable debate, and finding a way to present the opposing side in a serious and sympathetic way would have helped readers sort through the debates

3. How reliable were the data? If appropriate, were surveys properly designed and executed? Were data compiled properly from appropriate sources and aggregation? Were data limitations considered in the analysis?

The cross-country analyses are subject to all of the caveats attached to the sub-field; they lean on data from the IMF International Financial Statistics data, which is the obvious source. The work here is no better or worse than average in the field. The discussion and interpretation is particularly fine, though. The case studies use official statistics, and no sense is given of their reliability (I have no particular reason to question them, nor to trust them especially).

4. Are the conclusions consistent with the research findings? Were problems/concerns noted and reviewed?

To a large degree. The chapters point to the problems that emerge from rising real interest rates, but there will surely be debate (and has already been quite a lot) about what drove the increases. Cho, for example in chapter 6, places responsibility with short-term debt, but other observers of Korea point instead to the rise of commercial paper and unchecked corporate investment.¹

C. Accessibility

1. Were project reports and publications presented in a manner appropriate for and accessible to the intended audience(s)?

Yes, the papers are written for well-informed policymakers. The chapters are data-intensive, but not regression-intensive. They are concept-laden, but not overly formalized.

2. If applicable, are policy recommendations commensurate with findings?

For the most part. The policy recommendations are largely a plea for humility and a pitch for case-by-case analysis over simple universal approaches. The points are well argued.

The failure to step back and engage with the underlying sources of institutional and political failure, though, suggest that wise policies may not be embraced as a matter of

¹ E.g., Domenica Tropeano, "Review of Gerard Caprio, Patrick Honohan and Joseph E. Stiglitz (eds.), *Financial Liberalization. How Far, How Fast?*" *Financial History Review* 10, 2003.

course. In a single volume, taking on the economic debates is probably enough of a task, but it remains that the economic analysis is only one input into policy formulation. Much of the action rests with the dynamics of domestic businesses, political groups, and international actors—issues about which readers learn little here.

D. Results

1. What are the key findings of the study? How do they advance country policy in the field?

Some key arguments: Quick liberalization can get countries into trouble. Financial crises emerge through self-dealing, deliberate gambling, and over-optimism. Regulators have high challenges in limiting these behaviors. Liberalization has been driven by the imperatives of technological advances and the forces of globalization, as well as by ideological argument—and it won't be turned back, but it can be managed better. The volume is sympathetic to the cause of regulation, while mindful of possible side effects. This is a helpful antidote to those who argue for letting markets weed out problems on their own, and the case-study approach is enlightening.

The book empowers regulation-minded policymakers against their free-market colleagues. If the regulators have absorbed the subtleties of the book, all for the good. If they haven't, it's unclear whether it's best to throw one's lot in with the bureaucrats and politicians or to lean more heavily on the market.

Annex 3: Research Quality Indicator Form

S	Superior
AA	Above Average
A	Average
BA	Below Average
U	Unacceptable

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

103: Rural Finance Studies

5 must-read papers:

Pitt-Khandker, *JPE* 1998

Khandker-Samad-Khan,

Khandker-Faruqee

Jain-Mansuri, *JDE* 2003

Pitt-Khandker-Cartwright

A. Objectives

1. What specific questions or hypotheses did the study seek to answer?

Four of the five papers seek to measure the impact of financial access (microfinance specifically) on household-level and village-level impacts – income, employment, poverty, and “empowerment” of women. The applications are in Bangladesh (3 of 4) and Pakistan. The fifth paper turns to a specific aspect of microfinance contracts (regularly scheduled repayments) and seeks to explain its role.

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

There has been much debate about the impacts of microfinance and the mechanisms employed. Few studies employ convincing control groups, so debates persist. The questions asked are relevant and important, especially as interest around microfinance and financial access builds.

B. Design and Implementation

1. Was the methodology appropriate and well executed? Was it innovative?

The impact studies, for the most part, recognize and describe the basic identification problem essential to solve when assigning causality. It’s not an easy problem to solve in this case, and the authors probably feel more confident in their results than would critical readers.

There is often a tension with empirical projects: given an imperfect data set and an imperfect approach, do you forge ahead anyway—hoping to shed some light and doing your best to note the caveats along the way? Or do you wait for better data and methods in order to generate definitive, robust results? There can be a reasonable difference of opinion over the choice, and inevitably it becomes a question of *how* imperfect the inputs are. The impact studies here are mainly of the “forge ahead anyway” approach. As such, they are among the first to identify key questions and generate results.

With the luxury of hindsight, it's hard not to feel that it would have been better to wait for better data. Identification here rests too much on untested assumptions. In some cases, critical assumptions are so implausible as to render the papers uninteresting and perhaps misleading.

The paper on village-level impacts of microfinance in Bangladesh (Khandker, Samad, and Khan, *Journal of Development Studies* 1998), for example, attempts to estimate village-level impacts of microfinance using a cross-section of villages. This is an ambitious agenda in the best of circumstances. It seems close to impossible to get right with the data here. Because program rules bar households with over a half-acre of land from borrowing, there is a group in each village that can't borrow from the microlenders. (It turns out that there is a lot of leakage in the application of the half-acre rule, although this is not noted in any of the papers). Even if there was no leakage, this better-off group is essentially used as a control group here, although the microfinance programs determined their eligibility rules with the notion that villagers holding over half an acre are *structurally different* from poorer households. Identification of causal impacts rests on the use of village-level fixed effects, and that in turn hinges on the assumption of basic similarities across the landholding groups. A moment's reflection shows how tenuous this set-up is. The authors note that the methods are "less than ideal" (p. 99) but go on to argue that "Even if the method is not ideal it can shed some light on the basic question of whether there are net positive impacts on rural economy due to micro-credit programme interventions." Ultimately, it is not clear how much light is shed. The authors characterize and interpret their results as causal impacts but they are, in the end, only associations. The interpretations of the results are plausible, but ultimately they are only as useful as the underlying results—which are open to a host of methodological questions.

The paper by Pitt and Khandker (*Journal of Political Economy*, 1998) uses the same cross-sectional data set and is even more ambitious (it attempts to split out results by gender). It takes on valuable questions, but, not surprisingly, the reliance on fixed effects similarly undermines the usefulness of their results. [Pitt and I have had a few rounds on this, and his response is on the web; his response, in my reckoning, still leaves unaddressed the underlying problem with the fixed effects.]

The paper by Khandker and Faruquee on farm credit in Pakistan considers the impact of credit from a government bank. Because credit is in part allocated bureaucratically, the authors use as instruments (for the endogeneity of credit receipt) a series of averages of attributes at various levels (village, district, etc.). The paper describes the problem well, but it is impossible to feel comfortable with the exclusion restriction necessary to interpret the results meaningfully (i.e., one has to assume that average conditions in a region have no direct bearing on outcomes of interest—yet in a market system it is hard to imagine how this could be plausible). This method has been proposed previously and has never found wide acceptance. The paper's motivation and set-up are terrific, and it would be great to go back to the data with an alternative identification strategy--if one could be found.

Jain and Mansuri have taken up a question around credit contracts, and their contribution is mainly theoretical. They use some data that suggest the plausibility of their case, but

they are (appropriately) not pushed very hard. It would be interesting to see a stronger test were the right data available.

2. 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?

The impact studies are clearly written and quite wide-ranging. In many ways they show detailed knowledge of the countries. At the same time, though, the methodological assumptions made are so implausible that the country knowledge is countered by the simplifications inherent in the estimation. Ultimately, the estimation drives the papers.

The Pitt-Khandker paper, for example, uses the assumption that land markets do not exist in rural Bangladesh. Simply chatting with villagers would show how off base the assumption is. In subsequent work (not included in the materials here), Pitt shows the basic results are robust to a series of different assumptions about land holdings, and I have argued that in this particular case the assumption turns out to not pose a real problem in the set-up. But, that said, it is striking that the initial claim was ever put forward with seriousness and was never tested (despite ample data on land sales in the data set).

The Jain and Mansuri paper would have profited from more on-the-ground evidence. I don't find their story as plausible as other stories about frequent repayment schedules, and it would have been useful to be presented with corroborating evidence. The question they take up was highlighted in a paper I wrote with Beatriz Armendariz de Aghion (which they cite). We argue that regularly scheduled repayment patterns can be explained reasonably through savings constraints and the need to mesh the timing of cash flows with the timing of loan installments. Our argument comes from discussions with borrowers and lenders in Bangladesh and Indonesia and from a series of case studies of financial mechanisms globally. Jain and Mansuri provide an alternative explanation based on information signals—which is a contribution to the menu of possibilities. In the end, I find our evidence more convincing than theirs (I'm not unbiased, of course), and I would have hoped for some guidance in distinguishing between stories.

3. How reliable were the data? If appropriate, were surveys properly designed and executed? Were data compiled properly from appropriate sources and aggregation? Were data limitations considered in the analysis?

The Bangladesh data (to which, thanks to Shahidur Khandker, I have had access) were not collected particularly carefully. For example, in villages without microfinance programs, a sub-sample was collected from households holding less than an acre of land—these are households that in principle would have been eligible to borrow had Grameen or another microlender operated there. In the villages with microfinance access a similar subsample was collected, but about a quarter of these “eligible” households have over a half-acre. A proper pilot survey and discussion with program staff should have resolved issues around the leakage before the survey was put in place—given how critical these sub-samples and definitions were for the sampling strategy.

The data set is large and comprehensive and has the advantages and disadvantages of size. My experience is that the data were quite carefully compiled and documented,

though, and Khandker should be commended for making the data accessible to other researchers.

4. Are the conclusions consistent with the research findings? Were problems/concerns noted and reviewed?

As far as the impact studies go, I don't think that some deep problems were ever truly addressed. But that said, many problems were addressed carefully and responded to thoughtfully.

The findings by Pitt and Khandker are quite clear and nicely described in the paper. Khandker subsequently extrapolated from those findings (in his book, *Fighting Poverty with Microcredit* – which, surprisingly, is not included in the work under review here). There is a great deal to like about his book, but some of the extrapolations are based on implausible assumptions and yield implausible results--that have nonetheless been widely touted by advocates. Most notable is the claim that 5 percent of microfinance customers exit poverty each year thanks to microfinance, an extrapolation of a finding in the Pitt-Khandker paper. I don't believe that Khandker ever meant the result to be interpreted so strongly, but it was obviously too hard to resist citing it. Khandker's follow-up work with a later round of data showed how far off the earlier claim proved to be.

C. Accessibility

1. Were project reports and publications presented in a manner appropriate for and accessible to the intended audience(s)?

The Pitt-Khandker paper, the most celebrated of the group, is more forbidding than it needs to be. It is written for a top journal and is less forbidding than many papers in the journal (the *JPE*), but it is laden with unnecessary acronyms (WESML-FIML-FE) and lacks a clear intuition for where identification actually comes from. Given how much demand there is for these results among the practitioner community, an opportunity was lost in not making the results and method more transparent.

Overall, the papers are written clearly. Although the details will be lost on non-specialists, they'll be able to get the big picture.

2. If applicable, are policy recommendations commensurate with findings?

All seems consistent.

D. Results

1. What are the key findings of the study? How do they advance country policy in the field?

The findings of the impact studies show positive impacts of microfinance on household behaviors—and more for money lent to women than to men. There are some interesting findings on labor market effects – which it would be terrific to follow up on with better data and methods. Given the generally positive endorsements of microfinance already, the findings underscore the arguments made by advocates and reinforce policy directions that are already in place.

Annex 3: Research Quality Indicator Form

S	Superior
AA	Above Average
A	Average
BA	Below Average
U	Unacceptable

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

African Labor Markets

Project 105

Abigail Barr and Abena Oduro “Ethnic Fractionalization in an African Labour Market”
Journal of Development Economics 68, 2002.

A. Objectives

1. What specific questions or hypotheses did the study seek to answer?

The study explains earnings differentials between members of different ethnic groups and between employers’ relatives, unrelated co-ethnics, and other workers in the Ghanaian manufacturing sector. Specifically, the authors are interested in the role played by ethnic fractionalization in these earning disparities. In doing so, they use the firm-level results here to illuminate the macro-level results for Africa (the cross-country results highlight strong negative roles of ethnic fractionalization on economic performance; e.g., Easterly and Levine, “Africa’s Growth Tragedy,” *QJE* 1997).

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

Understanding labor market dynamics and discrimination is indeed critical. The topics are politically sensitive and economically meaningful, making them appropriate candidates for careful research.

B. Design and Implementation

1. Was the methodology appropriate and well executed? Was it innovative?

The method draws on earlier work by Altonji and Pierrot (*QJE* 2001) that focuses on statistical discrimination and learning by employers about worker quality. The framework is adapted in a straightforward way to address networking effects and to focus on the role of observability over time. It’s an appropriate tool and is well applied.

2. 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?

There’s an extensive literature review on discrimination that reaches back to the early papers by Becker and Arrow, but the review of issues related to labor market discrimination in developing countries (and Ghana in particular) is brief – and limited to a footnote on the second page (fn2, p. 356).

3. How reliable were the data? If appropriate, were surveys properly designed and executed? Were data compiled properly from appropriate sources and aggregation? Were data limitations considered in the analysis?

The data cover four cities, drawn from the 5th wave of the Ghanaian Manufacturing Enterprise Survey. There is no discussion of data quality or adjustments. On p. 364,

there is discussion that would suggest that the authors played a role in the data collection or carried out a complementary survey, but it is not made explicit and only mentioned there. As the authors note, defining ethnicity can be tricky, and the data here are self-reported.

4. Are the conclusions consistent with the research findings? Were problems/concerns noted and reviewed?

Yes, the conclusions appear consistent with the research findings. The Altonji-Pierrot test for statistical discrimination does not support the hypothesis that there is statistical discrimination. The main conclusion is that there is strong labor market segregation by ethnic groups and that this accounts for pay differentials. The authors acknowledge that the classification of employees into occupations was conducted through interviews with the workers and may in many cases be quite arbitrary, and the authors suggest that the results (at least for part of the analysis) should be treated with “considerable caution.” The caveats are not returned to later and do not affect the overall results.

C. Accessibility

1. Were project reports and publications presented in a manner appropriate for and accessible to the intended audience(s)?

While the description of the empirical model requires a certain level of sophistication with econometrics, the article presents background and conclusions in a manner appropriate for individuals without expertise in this area. When jargon is used, efforts are made to define technical terms before they are used. The paper in general is clearly written.

2. If applicable, are policy recommendations commensurate with findings?

There are no direct policy recommendations. The authors appear more interested in understanding the causal mechanisms involved in earning differentials in the Ghanaian labor market than in proposing methods to reduce them. One aspiration is to depoliticize issues of wage inequality by showing that patterns don't seem to be driven by discrimination (rather, they're driven by market segregation). Yet the kind of market segregation observed is surely a product of discrimination at a higher level – and, surprisingly, this receives no attention.

D. Results and Cost-effectiveness

1. What are the key findings of the study? How do they advance country policy in the field?

The article concludes that the Ghanaian labor market is indeed ethnically fractionalized and that is the cause of considerable earnings differentials between ethnic groups. There is little evidence to support that the differentials result from statistical discrimination by employers. Explaining the greatest portion of the earnings differentials is a strong tendency for workers to be employed by members of their own ethnic group and the fact that entrepreneurs from different ethnic groups run very different types of enterprises. The conclusions suggest that ethnic fractionalization in the African labor market is not necessarily the result of nefarious causes but may occur naturally and may not be easily changed by government policy. The question of what determines market segregation is left open.

Annex 3: Research Quality Indicator Form

S	Superior
AA	Above Average
A	Average
BA	Below Average
U	Unacceptable

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

Small Financial Systems Project 106

2 papers:

Biaggi Bossone, Patrick Honohan and Millard Long: "Policy for Small Financial Systems," Financial Sector Discussion Paper No. 6, February 2001

Patrick Honohan and Anqing Shi, "Deposit Dollarization & the Financial Sector in Emerging Economies," working paper, no date.

A. Objectives

1. What specific questions or hypotheses did the study seek to answer?

The two papers address issues specific to small financial systems. In Bossone, Honohan, and Long, the systems are defined to include the 60 countries like Sudan and Uganda with financial systems whose size (as measured by M2) is under \$1 billion; 30 are active Bank/IDA borrowers. About two-thirds of the 120 World Bank member states have M2 lower than \$10 billion, and 50 of them are active borrowers.

The Bossone, Honohan, and Long paper is synthetic and does not pursue particular hypotheses. Instead, they describe conditions, explore tensions, and advocate for solutions. It is a policy research paper, rather than an applied research paper.

The tension for a small financial system involves the inefficiency of small scale pitted against the risks inherent in fully integrating into regional economies. On the latter score, Honohan and Shi specifically address dollarization, among the most extreme forms of integration. Their paper addresses specific questions around dollarization. For example,

- a. Is faster exchange rate pass-through associated with increased dollarization?
- b. Does dollarization of deposits shrink credit availability? Does it increase interest spreads?
- c. Is dollarization associated with currency risk?

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

They're clearly vital. It's not obvious why these are issues should be taken up by the Bank rather than the Fund, but they are certainly important.

B. Design and Implementation

1. Was the methodology appropriate and well executed? Was it innovative?

My best response to the question is to cite footnote 23 of Honohan and Shi in its entirety (p. 13):

The regression approach we have adopted is designed to detect broad cross-country trends. The data is still not good enough to pretend to estimate structural models, and the relationships we report vary in their statistical robustness, as is

noted where appropriate. The regressions of annual relationships reported in this table represent a specification style that we have found useful through the remainder of the paper. Thus, we estimate a panel relationship with a single first-order autocorrelation correction to take account of detected serial correlation. The panel is usually between 7 and 11 years long, with many missing observations. Where the explanatory variable is country-specific, we rely on estimates made with a common intercept, as use of country fixed-effects in these very short duration panels tends to wash out any differential impact of the variable being examined. Where the explanatory variable is common across all countries, we may rely on estimates with country fixed-effects. In short, the data are far from perfect, and the authors do what they can with it. The methodology can only be pushed as far as the data.

2. 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?

The Bossone, Honohan, and Long study provides an impressively wide-ranging synthesis. The Honohan and Shi study is more focused but thoughtfully comprehensive.

3. How reliable were the data? If appropriate, were surveys properly designed and executed? Were data compiled properly from appropriate sources and aggregation? Were data limitations considered in the analysis?

The footnote cited above speaks in part to data reliability and availability. The general sense is that reliability remains a major concern.

4. Are the conclusions consistent with the research findings? Were problems/concerns noted and reviewed?

Problems were noted, and the conclusions are broadly consistent with the findings.

C. Accessibility

1. Were project reports and publications presented in a manner appropriate for and accessible to the intended audience(s)?

Neither appears to have been published. Perhaps they were eventually, or parts were, but as unpublished papers their accessibility is likely to be limited – unless they're posted and featured on the World Bank site, where they can get a fair amount of play in working paper form.

2. If applicable, are policy recommendations commensurate with findings?

Broadly speaking, yes.

Results

1. What are the key findings of the study? How do they advance country policy in the field?

The main contribution is to shine light on small financial systems. They reflect economies of importance to the Bank. And, from a methodological point of view, they contribute a fair number of data points in cross-country analyses. A study focused on small financial systems is welcomed.

The Bossone, Honohan, and Long synthesis achieves a good balance. It argues that World Bank financial policy is too un-nuanced – it fails to distinguish between small and large financial systems. The difference matters: opening to easy foreign entry can create grave problems in small financial systems. They are particularly vulnerable to contagion, for example, and often have difficulty attracting high-quality financial institutions. At the same time, the authors argue that, at present, most small financial systems are probably too closed, and they pay a high cost in inefficiency.

The studies thus provide a concrete, evidence-laden critique of Bank policy that is put forward in a way that seems constructive and likely to be heard internally.

Annex 3: Research Quality Indicator Form

S	Superior
AA	Above Average
A	Average
BA	Below Average
U	Unacceptable

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

Foreign Direct Investment (FDI), Financial Markets, and Growth

Project 107

Must-read documents:

Ross Levine and Sergio Schmukler, “Internationalization and Stock Market Liquidity,” working paper, December 2005

Clarke et al, “Bank Lending to Small Businesses in Latin America: Does Bank Origin Matter?” *Journal of Money, Credit, and Banking* 37, 2005

Rui Albuquerque, Normal Loayza, and Luis Servén, “World market integration through the lens of foreign direct investors,” *Journal of International Economics* 66, 2005.

A. Objectives

1. What specific questions or hypotheses did the study seek to answer?

The papers address how global financial integration is affecting domestic financial markets. Is it increasing FDI? Increasing the capital access of small businesses? Driving patterns in domestic stock markets?

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

Financial integration is a major item in conversations around macro policy, trade policy, and other system-wide features of economies. So, at one level, these issues are obviously of top importance. At another level, it's not clear that they matter for much of the Bank's sectoral work – nor is it entirely clear how (a large chunk of) the research applies to specific policy levers.

B. Design and Implementation

1. Was the methodology appropriate and well executed? Was it innovative?

The research leans heavily on cross-country analysis. New data are collected and indices created. Within the context of the cross-country literature, the work is well done. There's a good case to be made that alternative approaches (country-based studies in particular) would have provided useful context for the broad-brush approach favored here.

As elsewhere in the cross-country literature, there are difficult issues of endogeneity that the researchers cannot fully address. For example, the Clark et al paper on foreign banks in Latin America has an interesting finding that disaggregates the role of foreign banks with a small presence in a country versus those with a large presence. The reader, though, is left to wonder why a bank has a large presence—and whether the mechanism is correlated with outcomes of interest.

2. 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?

Much of the work relies on cross-country regressions. Being macro analysis, that's the natural level of investigation. But too much of the work begins and ends with cross-country analysis—and begins and ends with the paper's relevance relative to other papers written in the same vein, rather than to specific policy debates or country experiences. Other work done by DEC finance researchers more explicitly couches cross-country analysis within specific experiences of countries and draws links between historical analyses and the broader descriptions of world-wide patterns (i.e., cross-country work). That practice is, unfortunately, far less evident here—and those in the policy community lose out for it.

For example, for all of its merits, the Levine and Schmukler paper is motivated nearly entirely by the previous academic literature. The discussion centers on the results relative to findings by other researchers also working in the cross-country empirical context. While the authors may have a deep awareness of country experiences, virtually none of it is on display. The paper is not (apparently) motivated by a question that emerges from conversations within particular developing countries or within the community of policymakers; rather, the motivation arises (apparently) from conversations within the community of like-minded researchers. The research question is whether there are notable spillovers on the liquidity of domestic stock markets when some firms seek international capital (through trading equity, say). The point is not that the question is irrelevant. To the contrary, it's quite important to a wide range of actors. The main finding—that internationalization by some firms appears to harm others in the market—is a striking rebuttal to those who preach the merits of unmitigated openness. It is the relevance of the result that makes it especially notable that the paper remains so devoid of country-level context—and makes such little attempt to engage an audience beyond research economists.

The paper is destined for a top finance journal, and the mode of writing can be understood in that context. If publication in top journals is the only criterion for success, there is no incentive to do otherwise. Those who are concerned with actual policymaking, though, will need much more context and motivation. Other parts of the project could have presented the material accessibly (as has been done successfully in other DEC projects), but I don't see it here.

3. How reliable were the data? If appropriate, were surveys properly designed and executed? Were data compiled properly from appropriate sources and aggregation? Were data limitations considered in the analysis?

Levine and Schmukler put together a data set based on 1000 firms over 12 years and 31 countries (taken from DataStream). Those data are supplemented with information on the internationalization of equity and related metrics of internationalization (taken mainly from the Bank of New York and various stock exchanges). It is an impressive empirical effort, but no real sense of the representativeness of the data is given, or of the pros and cons of these sets.

Albuquerque, Loayza, and Serven put together a large cross-country data set that links globalization, capital market liberalization, and investment. It is impressive but, again, little sense is given of the underlying data and its limits.

4. Are the conclusions consistent with the research findings? Were problems/concerns noted and reviewed?

Like much of the cross-country literature, the conclusions are made boldly and applied universally. For example: “The paper finds that the internationalization of stock markets has a negative effect on the trading activity of domestic firms in emerging markets.” But the coefficient of interest is a weighted average across countries. There may be important exceptions, positive and negative, that are washed out of the average—and the exceptions can be as interesting as the central tendencies. No serious attempt is made to go beyond averages. So the conclusions are consistent in a broad sense, but there is little appreciation of the underlying heterogeneity.

It is perhaps unfair to isolate these particular papers for a practice that is widespread. But in a body of work that claims policy relevance, the leap of faith required in moving from results to conclusions should be on the table.

C. Accessibility

1. Were project reports and publications presented in a manner appropriate for and accessible to the intended audience(s)?

Yes, but, as above, the *intended* audience is a narrow one. It would seem important to consider ways to reach other relevant audiences.

2. If applicable, are policy recommendations commensurate with findings?

Yes and no. The main issues are described above in B4 and I won't repeat them.

D. Results

1. What are the key findings of the study? How do they advance country policy in the field?

The studies overall show strong (and growing) links between the globalization of the world economy and domestic financial outcomes. The findings are extremely useful in setting the context for debates on globalization, and I hope that they will get the attention they deserve.

Albuquerque, Loayza, and Serven set the scene by creating a globalization measure that can be related to economic variables in their cross-country/time-series data. Between 1985 and 1999, their globalization measure increased 10 times, and developing countries appear to be as exposed to the global economy as do developed economies. The global factors drive much of foreign domestic investment, and the authors take this as an argument for further (and urgent) capital market liberalization.

Clark and others show that foreign banks in four Latin American countries tend to lend less to small businesses relative to larger businesses (as a share of their lending portfolios). But, more optimistically perhaps, foreign-owned banks with a large domestic presence lend *more* to small businesses than do large domestically owned banks. While the results are nuanced, the bottom line is that there are advantages to encouraging foreign banks to take a large presence. An important (and un-noted) question concerns what domestic banks would make of that—i.e., a policy to encourage foreign banks would likely have deep political ramifications and would face considerable push back from domestic banks. So, in the end, it's not clear how policymakers should proceed once politics enters as a part of the equilibrium—but the paper is helpful in mapping out the economic side of the issues.

Levine and Schmukler consider what happens when some firms seek international capital. They document negative implications on other firms in the domestic market and a tendency for “migration” to international markets. Although the results are interesting, the paper, as written, does little to connect the results to country policy.

Annex 3: Research Quality Indicator Form

S	Superior
AA	Above Average
A	Average
BA	Below Average
U	Unacceptable

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

Taxation of Financial Intermediation

Project 108

The main output is a book edited by Patrick Honohan: *Taxation of Financial Intermediation: Theory and Practice for Emerging Economies*.

The highlighted chapters are a synthesis by Honohan, a discussion of inflation as a tax by Honohan, and a theoretical chapter by Broadway and Keen.

A. Objectives

1. What specific questions or hypotheses did the study seek to answer?

The book offers a comprehensive take on the pros and cons of taxing financial intermediation. Banks are easy (sometimes too easy) to tax, and a range of distortions often results. Given the importance of financial intermediation to growth and investment, Honohan and company step back and explore a wide range of issues pertaining to domestic banking. One particular set of questions surrounds VAT proposals that extend to financial intermediation.

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

They're not the grandest of policy issues, but tax policy is certainly relevant for policymakers.

B. Design and Implementation

1. Was the methodology appropriate and well executed? Was it innovative?

There is a range of studies. They draw on fairly well understood principles of public economics, and the main interest is in the application. They are thoughtful and relevant, but not particularly innovative.

2. 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?

The main chapters are not very rooted in country experience, but the book as a whole provides a deeper richness of context.

3. How reliable were the data? If appropriate, were surveys properly designed and executed? Were data compiled properly from appropriate sources and aggregation? Were data limitations considered in the analysis?

Not particularly relevant.

4. Are the conclusions consistent with the research findings? Were problems/concerns noted and reviewed?

Generally, yes. Honohan's chapters are particularly careful to note tensions.

C. Accessibility

1. Were project reports and publications presented in a manner appropriate for and accessible to the intended audience(s)?

Honohan's introductory chapter is characteristically lucid and balanced. The next chapter, by Broadway and Keen, lays out the theory well. I found it quite useful, but the algebra is thick and even many economists will find it a slog if they make it to the end. Policymakers won't know what to make of it. Still, it's useful to establish clear principles up front, and in this case the math helps.

Honohan's chapter on the inflation tax provides a clear statement of the issues and new analysis, easily accessible to economic researchers. The chapter is smart and is, unfortunately, unlikely to get the attention it deserves tucked away toward the end of this volume. It would have been more accessible if placed in a journal—or if a version of it was.

2. If applicable, are policy recommendations commensurate with findings?

Generally, yes.

D. Results

1. What are the key findings of the study? How do they advance country policy in the field?

The book investigates the possibilities for and difficulties with taxation of financial intermediation. It identifies three main criteria by which a tax system is judged acceptable: (1) minimizes distortion for a given amount of revenue collected; (2) corrects distortions due to asymmetric information; and (3) doesn't distort excessively relative to other sectors.

Beyond that, a few conclusions are highlighted: avoid big capital gains taxes, avoid big taxes on transactions, be careful when trying to correct market imperfections (damage may exceed harm), and watch out for arbitrage that shifts taxes. A particular argument, well made in chapter 13, is that inflation imposes a larger tax than direct taxes.

These are all sensible bits of advice from a public economics viewpoint; none is surprising.

The chapters recognize that taxing the financial sector may be one of the only options open to governments strapped for cash. If that's so, it's not clear that policymakers will be in a position to listen to the wisdom here. Further steps could be taken to profitably tackle second-best (or third-best, etc.) options directly.

Annex 3: Research Quality Indicator Form

S	Superior
AA	Above Average
A	Average
BA	Below Average
U	Unacceptable

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

Entry, Growth, and Entry of Firms

Project 110

This report focuses on two papers (there are no others in the sample):

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, Ashrei Shleifer (2003). “Courts,” *Quarterly Journal of Economics* 118 (2).

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes, Ashrei Shleifer (2001). “Regulation of Entry,” *Quarterly Journal of Economics* 117 (1).

A. Objectives

1. What specific questions or hypotheses did the study seek to answer?

Both studies seek to understand the structures and implications of legal systems; they both draw on new cross-country data sets, compiled by the authors, that provide evidence for a large number of countries. “Courts” takes up determinants of legal procedural formalism, and in turn the effect of formalism on court quality and efficiency (generally not great). “The Regulation of Entry” takes up the determinants and costs (mainly in terms of corruption) of regulation of entry.

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

These are enormous issues. Legal frameworks provide the parameters and incentives that govern investment, contract enforcement, and bribery, to name several issues of concern for well-functioning societies.

B. Design and Implementation

1. Was the methodology appropriate and well executed? Was it innovative?

The key innovation is data collection. It seems sensitive and well done, although one would have to be an expert in each country to be able to truly judge how coding choices were made.

2. 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?

The papers excellently incorporate competing theoretical viewpoints. They do not, though, lean heavily on prior empirical work, perhaps because it’s scant or inaccessible—or the work was simply not done.

3. How reliable were the data? If appropriate, were surveys properly designed and executed? Were data compiled properly from appropriate sources and aggregation? Were data limitations considered in the analysis?

The “Courts” article relies on data from attorneys who answered questionnaires on the legal contexts in 109 countries. Care was taken to have cross-checks within the firms completing the surveys, but there were no cross-checks via asking different firms to

provide information on the same countries. Readers are expected to take on face value that the questions were answered (and interpreted) in appropriate ways by the chosen attorneys. Readers are given no indication of who the attorneys are or how they came to be associated with the project.

“Regulation of Entry” also uses law firms to provide data (supplemented with official data) and to adjudicate differences.

The focus in “Regulation of Entry” is on a “standardized firm” that performs general industrial or commercial activities, operates in the largest city, does not participate in foreign trade, and has limited concerns with taxes.

“Regulation of Entry” notes an important caveat that likely spills over to the “Courts” essay: “Our data surely underestimate the cost and complexity of entry. Start-up procedures in the provinces are often slower than in the capital. Industry-specific requirements add procedures...”

This standardization is necessary and helpful, although it will necessarily lead readers to wonder what things look like in “non-standard” cases—which in fact are apt to be far more common than the “standardized” but artificial one taken as the benchmark.

4. Are the conclusions consistent with the research findings? Were problems/concerns noted and reviewed?

Both studies to varying degrees suffer from problems of endogeneity. For example, the authors argue in “The Regulation of Entry” that regulation creates corruption. This seems entirely plausible. But the reverse is also possible: to some degree regulation may be a response to corruption—an attempt to tamp it down. The “Courts” article recognizes the reverse causation explicitly and turns to instrumental variables strategies (instrumenting the endogeneity of the degree of procedural formalism with data on historical origins of the country’s legal system).

C. Accessibility

1. Were project reports and publications presented in a manner appropriate for and accessible to the intended audience(s)?

Overall, yes. Both articles build simple models that are empirically tested. Both studies rely heavily on exhaustive tables to present results (although simple summaries of the key results would have helped to streamline the presentation). The conclusions are neatly and clearly laid out.

2. If applicable, are policy recommendations commensurate with findings?

The results of both studies strongly support reducing regulation—i.e., less procedural formalism and fewer barriers to beginning a business. The findings are derived from cross-country analysis and do certainly support this conclusion (caveats noted above, though).

That said, without a deeper understanding of the situation in a given country (the history, politics, and economic context), simply applying the broad dicta may lead to wrong-headed and misguided policy.

D. Results and Cost-effectiveness

1. What are the key findings of the study? How do they advance country policy in the field?

“Courts” finds that procedural formalism is systematically greater in civil versus common law countries, and is associated with inefficiencies and corruption. “Regulation of Entry” determines that countries with heavier levels of regulation tend to be poorer and have higher levels of corruption and larger unofficial economies; nor do they enjoy a better quality of public or private goods. Both articles suggest that the costs of formal procedures in the legal system in the form of corruption and inefficiency may outweigh any benefits in the public interest.

These are strong conclusions and may or may not bear out in a particular context. The value of the article is in illuminating the potential importance of legal frameworks in shaping economic development. It has all the strengths and weaknesses of cross-country analysis and cannot substitute for well-executed country studies. Those country studies would begin by understanding the political actors and their stakes in reforming (or blocking reform of) the existing regulatory frameworks.

Annex 3: Research Quality Indicator Form

S	Superior
AA	Above Average
A	Average
BA	Below Average
U	Unacceptable

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

Ownership Changes and Corporate Governance in Chinese Firms

Project 112

This review focuses entirely on the paper by Colin Xu, Tian Zhu, and Yi-min Lin, *Politician Control, Agency Problems, and Ownership Reform: Evidence from China*, *Economics of Transition* 13(1), 1-24, 2005. [I found a related paper under this project, "Does Bribery Grease Corrupt Bureaucracies?" (Colin Xu, Wei Li and Ting Lu, mimeo, the World Bank), more interesting in many ways. It pits the hypothesis that bribes can enhance efficiency by acting as "grease money to speed up the wheels of commerce" versus the hypothesis that bribes are mainly harassment. Disappointingly (but not surprisingly), firm-level data are used to show that the harassment hypothesis wins out. But to keep things streamlined, I've focused on the paper that Colin Xu identified as being the "best" of the lot.]

A. Objectives

1. *What specific questions or hypotheses did the study seek to answer?*

This study aims to understand the impact of reforming Chinese state-owned enterprises. It's a tough empirical problem as reform is often endogenous – i.e., responsive to conditions of the moment and applied most assiduously to the most reformable of a group of candidate businesses. The questions here focus on reforms in authority allocation, ownership structure, and corporate governance mechanisms.

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

Yes, in general. But they are particularly critical in China, and that alone is enough to defend the worth of the investigation.

B. Design and Implementation

1. *Was the methodology appropriate and well executed? Was it innovative?*

The questions are already well plumbed. The advances here include access to a particularly large data set on firms, and a mostly convincing econometric strategy to address the endogeneity of reform.

2. *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?*

The paper is firmly located within a particular line of economic investigation with roots in corporate finance (Shleifer/Vishny, Aghion/Tirole). There's reference to other economic studies of Chinese reform, discussed at a general level. There is no reference to any particular firm or reform experience—no case study material, for example, not even a sentence or two, that would ground the study. There are scattered details of an aggregate nature scattered throughout the footnotes, but, while I believe that the researchers know the context intimately, the paper is written with little China-specific

context. In short, like most economics papers, the paper seems to be mainly written for other researchers.

3. How reliable were the data? If appropriate, were surveys properly designed and executed? Were data compiled properly from appropriate sources and aggregation? Were data limitations considered in the analysis?

Data on returns on assets are central in this study. They are notoriously difficult to collect for firms like this. There's little discussion of data reliability or a way to gauge potential biases. This is one of the least satisfactory elements of the study, and it contrasts with the data analysis, which is generally thoughtful.

4. *Are the conclusions consistent with the research findings? Were problems/concerns noted and reviewed?*

Yes and no. The conclusions are based on a model that takes into account the endogenous nature of ownership reform. This is an important and well-executed step. The authors acknowledge that the data used are cross-sectional in nature and hence there may be bias from unobserved heterogeneity of firms, but they do not spell out how to think about or sign the likely bias. The authors also note that the analysis is restricted to short-term effects, but again they give no sense of how (the more important) longer-term effects are likely to play out.

C. Accessibility

1. *Were project reports and publications presented in a manner appropriate for and accessible to the intended audience(s)?*

Yes, if the audience is mainly fellow researchers. The researchers note (in a separate note) that the working paper has been downloaded over 300 times from the SSRN online database, so it is certainly getting play with a broad audience. I expect that readership extends well beyond economists and those interested in corporate reform, but the paper is not written in a way that is particularly accessible for those audiences.

2. *If applicable, are policy recommendations commensurate with findings?*

There are no specific policy recommendations, but the paper is generally supportive of reform. The main conclusions tie to methodological concerns (the endogeneity problem, notably) rather than to substantive policy concerns.

D. Results and Cost-effectiveness

1. *What are the key findings of the study? How do they advance country policy in the field?*

The authors suggest that reformed firms perform better when managers have flexibility in labor deployment, when corporate governance mechanisms lead to a better alignment

between ownership and control, and when foreign ownership is higher. These results are robust across specifications and predicted by theory.

Additionally, and importantly, dispersed individual ownership and operating autonomy lead to worse performance, suggesting that agency problems are a serious concern. This is an important—perhaps devastating?—concern, and certainly one relevant for policy. Yet it is treated as an anomaly that simply falls out of the regression and is left hardly remarked on (despite a mention in the abstract). There would surely be much interest were the authors to tackle that piece independently.

Annex 3: Research Quality Indicator Form

S	Superior
AA	Above Average
A	Average
BA	Below Average
U	Unacceptable

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

Subnational Credit Markets

Project 137

Book: Freire Maria E, John Peterson, Marcela Huertas, Miguel Valadez (2004). *Subnational Credit Markets* -- published by Oxford and the Bank.

A. Objectives

1. What specific questions or hypotheses did the study seek to answer?

The book is pitched in the context of decentralization. It investigates how subnational governments can best use financial markets to manage cash flows, instill fiscal and financial discipline, and foster responsible credit access. It does not pursue a specific hypothesis but instead presents itself as a primer for practitioners and government officials.

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

Given how important decentralization has become, this is a natural and important set of questions.

B. Design and Implementation

1. Was the methodology appropriate and well executed? Was it innovative?

The methodology combines the framing of broad analytical issues (part 1) coupled with country case studies (part 2).

It is not particularly innovative. Is it appropriate and well executed? One of the first elements that struck me was that some chapters in part 1 made very little reference to country cases – few anecdotes or examples to draw out the analytics. The chapters worked better where analysis and evidence were mixed. Correspondingly, the case study chapters were not uniformly well conceptualized. The book is straightforward and clear, if a bit dry. It does not attempt to prompt any serious rethinking or point to any serious questions, but perhaps that's what target readers want.

The book is notable in *not* going over some of the basics of financial management—for example, swaps and other financial derivative instruments that are commonly used in the United States by subnational governments. The book is a public finance book, not a financial management book. That's a distinction that will matter greatly to some insiders. As such it has inherent limitations and may not provide the comprehensiveness that a manager needs—or, more importantly, may not point the manager to what she doesn't know but needs to.

2. 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?

Quite good on this score.

3. How reliable were the data? If appropriate, were surveys properly designed and executed? Were data compiled properly from appropriate sources and aggregation? Were data limitations considered in the analysis?

N/A

4. Are the conclusions consistent with the research findings? Were problems/concerns noted and reviewed?

The book is really a primer, rather than a document that draws conclusions from evidence. The book would have been more successful if key dictums had been defended more rigorously. But as an accessible guide, it meets the basic criteria for usefulness.

C. Accessibility

1. Were project reports and publications presented in a manner appropriate for and accessible to the intended audience(s)?

The writing is clear. The demarcation between the chapters in part 1 is relatively obvious even to the uninitiated. Overall, the book ranks highly on accessibility. The chapters are short and will not satisfy detail-oriented readers or those with substantial experience, but they will be enough, I'd guess, for most readers.

2. If applicable, are policy recommendations commensurate with findings?

The policy recommendations are very general. In that sense, they are commensurate. This is not a study that pushes boundaries; instead the book houses simple wisdom. No doubt experts will quibble over details, but the big picture seems levelheaded and unexceptional.

D. Results

1. What are the key findings of the study? How do they advance country policy in the field?

The main sentiment is that subnational governments need to be in the business of raising long-term capital. But it's a hard business, and many of the subnational governments are freshly minted. At this point, the main directives thus focus on transparency and prudence, rather than on anything more exotic. This is no doubt wise counsel.

The book is a primer. It will give readers a sense of deeper conversations, most of which go on in the context of particular country policies, but the case studies mainly skim on the surface, jumping from region to region with just enough detail for readers trying to acquaint themselves with the big issues.

Annex 3: Research Quality Indicator Form

S	Superior
AA	Above Average
A	Average
BA	Below Average
U	Unacceptable

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

Does Foreign Direct Investment Increase Productivity?

Project 179

Javorcik, Beata Smarzynska. 2004. "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages," *American Economic Review* 94(3).

A. Objectives

1. *What specific questions or hypotheses did the study seek to answer?*

This study is interested in finding positive externalities from foreign domestic investment on the productivity of domestic firms. In particular, it focuses on effects operating across industries, rather than within. The literature has focused in the main on within-industry spillovers, and this study suggests that the real action is from manufacturers to suppliers in different parts of the production chain.

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

Knowledge about production spillovers informs attitudes toward openness to trade and foreign direct investment, both critical policy questions.

B. Design and Implementation

1. Was the methodology appropriate and well executed? Was it innovative?

The paper is well done. It's straightforward, which in this case is an advantage. The main innovation comes in focusing squarely on vertical linkages.

2. 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?

The study focuses on Lithuania, and there's limited Lithuania-specific content, but the author demonstrates a broad and deep understanding of both the economics literature and the business literature. Anecdotal evidence is brought in to balance systematic empirical evidence in a way that both sharpens and rounds out the paper.

3. How reliable were the data? If appropriate, were surveys properly designed and executed? Were data compiled properly from appropriate sources and aggregation? Were data limitations considered in the analysis?

The data, an unbalanced panel from 1996-2000, were collected by the Lithuanian Statistical Office. The one important missing bit of information is firm identifiers that would allow identification of individual firms that supplied multinationals; proxies must be employed instead.

4. *Are the conclusions consistent with the research findings? Were problems/concerns noted and reviewed?*

Yes, the conclusions follow from significant and robust estimates of forward and backward linkages. The study also does an excellent job noting and reviewing concerns.

C. Accessibility

1. Were project reports and publications presented in a manner appropriate for and accessible to the intended audience(s)?

Overall, yes. The article does a good job of developing a model of knowledge spillovers and describes the strengths and weaknesses of the particular data set used. In addition, its conclusions are presented with little jargon, although the empirical section is jargon-laden. Still, the jargon, where used, is appropriate for the publication and the audience and simply reflects acquaintance with advanced econometric techniques.

2. If applicable, are policy recommendations commensurate with findings?

No specific policy recommendations are stated in the article. But the results showing productivity spillovers certainly support policies that open economies to foreign investment.

D. Results and Cost-effectiveness

1. What are the key findings of the study? How do they advance country policy in the field?

The key findings in this study are a strong link between foreign domestic investment and the productivity of domestic firms. Backward linkages (contacts between foreign affiliates and their domestic suppliers) are found to be the mechanism behind such effects. The findings of this study speak to a conundrum in this field: the inability to robustly identify knowledge spillovers from foreign investors. The author suggests such spillovers exist, but previous studies have been looking in the wrong place.

Annex 3: Research Quality Indicator Form

S	Superior
AA	Above Average
A	Average
BA	Below Average
U	Unacceptable

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

Saving in Developing Countries

Project 96

Five Papers from *Review of Economics and Statistics* 82 (2000).

Attanasio, Pucci, and Scorcu: Saving, Growth and Investment: A Macroeconomic Analysis Using a Panel of Countries

Bandiera, O., G. Caprio, P. Honohan, and F. Schiantarelli: Does Financial Reform Raise of Reduce Saving?

Deaton and Paxson: Growth and Saving Among Individuals and Households

Loayza, N., K. Schmidt-Hebbel, and L. Servén: What drives Private Saving Across the World?

Lopez, N., K. Schmidt-Hebbel, and L. Servén: How effective is Fiscal Policy in Raising National Saving?

A. Objectives

1. *What specific questions or hypotheses did the study seek to answer?*

Attanasio, Pucci, and Scorcu: This study aims to establish the “stylized facts” that link saving, growth, and investment. In addition, it provides a review of the estimation techniques and methodological issues involved with analyzing macroeconomic relationships in a large panel of countries.

Bandiera, Caprio, Honohan, and Schiantarelli: This study focuses on the effect of financial liberalization on national saving.

Deaton and Paxson: The authors attempt to solve a conundrum: the lack of empirical support for the lifecycle theory’s prediction that changes in an economy’s rate of growth will affect its aggregate saving rate by changing the lifetime resources of younger people relative to older people. The authors suggest that it is the traditional method of estimating age-saving profiles using cohorts that is introducing error in the model.

Loayza, Schmidt-Hebbel and Servén: Using a new data set with panel information not previously as complete or extensive as this one, the authors seek to establish the stylized facts that determine private saving rates across the world.

Lopez, Schmidt-Hebbel and Servén: This article examines the role different fiscal policies play in raising national, and particularly, private saving.

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

Saving clearly matters, and governments are often particularly interested in raising national savings rates as a way to increase investment.

B. Design and Implementation

1. Was the methodology appropriate and well executed? Was it innovative?

Methodological issues are highlighted in all five papers and each represents an advance relative to previous work.

2. 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?

The Deaton-Paxson paper is the only one that focuses on micro data. The others are based on cross-country evidence and have relatively little to say about particular country experiences.

3. How reliable were the data? If appropriate, were surveys properly designed and executed? Were data compiled properly from appropriate sources and aggregation? Were data limitations considered in the analysis?

Data issues are noted elsewhere in this report. Savings data (particularly household and private saving) are particularly subject to measurement error.

4. Are the conclusions consistent with the research findings? Were problems/concerns noted and reviewed?

Attanasio, Pucci, and Scorcu: The conclusions do appear consistent with the research findings. Across multiple specifications, the authors find that lagged savings rates are positively related to investment rates, that investment rates Granger-cause growth rates with a negative sign, and growth rates Granger-cause investment rates with a positive sign. In addition, the authors review several causes of bias that they believe are not important enough to invalidate the results.

Bandiera, Caprio, Honohan, and Schiantarelli: The conclusions are broadly consistent with the findings. The authors do not address the strengths and weaknesses of using constructed indices as key independent variables of interest.

Deaton and Paxson: The conclusions are strongly in line with the research findings. The authors find that the relationship between growth and saving differs across the two countries in the study, and hence conclude that the positive relationship between growth and saving seen in cross-country studies is far from obviously consistent with the micro evidence. The authors thoroughly note problems and concerns, particularly problems of self-selection and economies of scale in household saving.

Loayza, Schmidt-Hebbel and Serven: The authors acknowledge that all of the variables may be fraught with measurement error. The authors estimate multiple specifications, and the conclusions they make appear to be robust across specifications.

Lopez, Schmidt-Hebbel and Serven: The key concern with macroeconomic empirical research of this sort rests with endogeneity. The authors address the problem by using

instrumental variables procedures based on estimates of the autoregressive parameters. However, unlike some of the other articles in this project, the authors call their estimates “quite satisfactory,” and do not identify any uncorrected problems.

C. Accessibility

1. Were project reports and publications presented in a manner appropriate for and accessible to the intended audience(s)?

Attanasio, Pucci, and Scorcu: This study’s accessibility depends on the group defined as the intended audience. It is highly unlikely that policymakers without a significant background in macroeconomics and econometrics will be able to understand this article. The article had the tendency to use terminology like “Granger causation,” without first defining the term or the concept behind it. However, trained economists should find the presentation informative, and it’s certainly appropriate for *Review of Economics and Statistics* readers.

Bandiera, Caprio, Honohan, and Schiantarelli: This study makes a decent attempt at describing the relationships between financial liberalization and saving. On pages 244-245, it describes clearly the theories as to the effect of the main components of liberalization on saving. However, much of the article is spent defending the empirical methodology and results, and requires a reasonably sophisticated knowledge of statistics to assess.

Deaton and Paxson: The article should be accessible to both statistical experts and policymakers without a statistical background. For econometricians, there is a mathematical derivation of the techniques used in this study. For statistical laymen, the language is clear and with little jargon used and multiple illustrative figures.

Loayza, Schmidt-Hebbel and Serven: This paper should be accessible to econometricians and academics and policymakers alike. For the mathematically savvy, the methodology is derived and explained in detail. The choice of each variable is explained without jargon, and the conclusions are summarized in everyday English.

Lopez, Schmidt-Hebbel and Serven: This article is clearly intended for an audience interested in advanced quantitative methods, as 80% of the article involves a description of the empirical methodology used in this study. As a result, it is unlikely that anyone without a statistical background will understand most of the article. However, the conclusion neatly summarizes the findings in an easy to understand manner.

2. If applicable, are policy recommendations commensurate with findings?

Overall, yes. Although, it’s unclear to what degree a particular country would be wise to put its faith in evidence drawn exclusively from cross-country data. The results here are helpful by shining light on methodological issues and showing general tendencies. Readers are given little guidance about how to extrapolate to any specific place or time.

Attanasio, Pucci, and Scorcu: The recommendations in this article are not policy recommendations; instead the focus is on empirical recommendations for estimation of large macroeconomic panel data sets.

Bandiera, Caprio, Honohan, and Schiantarelli: Yes. The authors argue that nations should not rely on an increase in saving as the channel through which financial liberalization can be expected to increase growth. The argument is in line with their result that financial liberalization may have a negative effect on saving.

Deaton and Paxson: This study does not attempt to make explicit policy recommendations, but aims to put forth an empirical model of savings that it believes is an improvement on preceding models. Resuscitating the life-cycle model, though, has important implications.

Loayza, Schmidt-Hebbel and Serven: There are few policy recommendations in this paper. But one key recommendation, encouraging private saving through economic development, appears to be commensurate with findings as growth is found to have a substantial and positive effect on private saving. The findings are derived from a cross-section of countries and reflect average tendencies across countries. The findings are useful for a particular country if there's some basis for believing that the country will follow the norms.

Lopez, Schmidt-Hebbel and Serven: One of the key policy insights that the article offers is that temporary expenditure cuts are more effective than temporary tax hikes in raising national saving. This recommendation is associated with results that indicate temporary expenditures do not alter saving behavior for constrained or unconstrained individuals. As above, the findings come from a cross-country data set and rest on all of the caveats entailed.

D. Results and Cost-effectiveness

1. What are the key findings of the study? How do they advance country policy in the field?

Attanasio, Pucci, and Scorcu: The key findings in this study are a solid link between lagged saving rates and investments rates, and support for a conclusion that investment rates Granger-cause negative growth rates, and growth rates Granger-cause positive investment rates. In terms of the empirical review aspect of the paper, the authors suggest researchers think about the number of time periods instead of the number of observations. The authors do not make any policy recommendations in relation to their findings.

Bandiera, Caprio, Honohan, and Schiantarelli: The key finding of this study is that financial liberalization is more likely to reduce private saving than to increase it. Hence, its effectiveness as an economic development policy is limited.

Deaton and Paxson: One of the key findings of this study is that results predicted by the lifecycle model can be achieved with the use of an empirical model that estimates

individual-level saving using household data. When applied to Taiwan and Thailand, the model finds vastly different relationships between growth and saving.

Loayza, Schmidt-Hebbel and Serven: There are several key results from this study. First, the authors conclude that growth has a significant, substantial, and positive effect on private saving. Therefore, instead of focusing policies on encouraging private savings, it may be more effective to focus policies on economic development. In addition, the authors conclude that the determinants of private saving are very slow to have an effect, although increasing public saving does have moderate positive effects on saving. The results indicate that deficits can be harmful to private saving.

Lopez, Schmidt-Hebbel and Serven: There are four key findings in this study:

The response of national saving to fiscal policy changes differs vis-à-vis developed and developing countries.

The effect of fiscal policy also differs based on whether it is permanent or temporary.

Temporary expenditure cuts are more effective in raising national savings than temporary tax hikes.

The impact of fiscal disturbances on national saving also depends on the factors that govern the time path of private consumption

Annex 3: Research Quality Indicator Form

S	Superior
AA	Above Average
A	Average
BA	Below Average
U	Unacceptable

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

Bank Supervision and Regulation

Project 97

Main outputs:

- Database on Bank Regulation and Supervision (posted on website).
- “The regulation and supervision of banks around the world - a new database” by James R. Barth, Gerard Caprio, Jr. and Ross Levine. May 2001.
- “Bank regulation and supervision: what works best?” by James R. Barth, Gerard Caprio, Jr., and Ross Levine, Updated December 2001. Published in the *Journal of Financial Intermediation*, 2003.
- Book: *Rethinking Bank Regulation: Till Angels Govern* by James R. Barth, Gerard Caprio Jr. and Ross Levine (Cambridge University Press, 2006). (*Note: Not available for this review.*)

A. Objectives

1. What specific questions or hypotheses did the study seek to answer?

Ultimately, the project seeks to identify the modes of bank supervision and regulation that lead to efficiency and robustness in a cross-section of countries.

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

Many developing countries are moving to models of prudential regulation of banking that follow the examples of developed countries. There is a broad range of advice offered about how to proceed, most of it based on expert judgment. The project seeks to provide evidence-based advice—or at least to enable it by collecting better data. This is a helpful step in the regulatory process. The project is motivated by the desire to replace existing (and sometimes contradictory) “Best Practices” that are based on anecdotes or an advisor’s experiences in their home country. The paper by Barth, Caprio and Levine (2003) lays out the issues well and motivates the project in terms of debates in the wake of Basel II.

B. Design and Implementation

1. Was the methodology appropriate and well executed? Was it innovative?

The heart of the project is data collection. The Bank is in a good position to take on this sort of project, as it involves surveying regulators worldwide. The method is not particularly innovative—but its straightforwardness is in many ways its virtue. The aim is simply to find out what’s going on by surveying regulators.

In relating regulation to bank outcomes, there is an obvious simultaneity problem – which is noted and addressed with a GMM instrumental variables method. The instruments for regulatory variables are an eclectic lot:

- Religious composition (following an argument that Catholic and Muslim religions “tend to generate hierarchical bonds of authority that shape the structure of government institutions;” B-C-L, p. 37)

- Latitudinal distance from the equator (because “countries in poor climates—tropical climates—tend to produce exploitative political regimes that gear governmental institutions toward protecting a small elite.”)
- Legal origin variables (“civil law and socialist law countries will tend to support stronger governments relative to private property to a greater degree than common law countries”).

The variables have explanatory power in the first stage and pass the over-identifying test, but it is a weak test (as the authors acknowledge). In addition explanatory variables are introduced sequentially in these tests, rather than being entered together (this was done for lack of more than 3 plausible instruments). In the conclusion, the authors note (appropriately) that “causality issues remain.”

The data are a cross-section of country-level variables, and there are also concerns about omitted variables, although these get scant attention.

All in all, the evidence is tentative, and is treated as such.

2. 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?

The papers build on the database. They are not a particularly rich source of country-specific data. The database, though, is freely accessible on the web and offers responses to about 175 questions for 107 countries.

3. How reliable were the data? If appropriate, were surveys properly designed and executed? Were data compiled properly from appropriate sources and aggregation? Were data limitations considered in the analysis?

It’s hard to say. The data were collected from experts, and the experts did not always agree in their assessments of conditions in a given country. That’s to be expected in this kind of exercise, but readers are not told who exactly completed the surveys. I assume that the main respondents were from the staff of regulatory agencies—and the sources of differences are unclear.

A second concern involves data completeness. Not all questions were answered by each respondent, and it’s unclear how this affects the analysis by skewing the composition of the data in the analyses.

One nice step is that part of the survey was reconciled with data from a survey conducted by the U.S. Office of the Comptroller of the Currency.

4. Are the conclusions consistent with the research findings? Were problems/concerns noted and reviewed?

Generally, but as noted above, there are (un-noted) concerns with potential biases from missing data and a set of (un-noted) caveats that are inherent when drawing conclusions from cross-country regressions when underlying behaviors are heterogeneous.

C. Accessibility

1, Were project reports and publications presented in a manner appropriate for and accessible to the intended audience(s)?

The project's commitment to accessibility should be applauded. The reports are in books and academic journals at which a regulator might not look, but the working papers should be widely available. The Cambridge University Press book should be accessible too. Most important, the base data are freely available on the web, and that should increase the flow of subsequent analyses on these themes.

2. If applicable, are policy recommendations commensurate with findings?

Broad policy recommendations are drawn from the cross-country analyses. It is unclear, though, whether results that hold on average across countries will hold for any given country. This is a generic issue and not specific to this project.

D. Results

1. What are the key findings of the study? How do they advance country policy in the field?

The studies find that the private monitoring of banks is associated with improved bank stability and financial performance. Generous deposit insurance schemes are associated with worse bank performance and fragility. (This might be moral hazard – but I'm left wondering if there's also simultaneity bias.) Government ownership is a bad sign too (again, can simultaneity bias explain this in part—governments are more likely to step in where the hurdles faced by the private sector are high?)

The study also finds that easing restrictions on bank activities enhances bank performance and stability by diversifying income streams. It makes sense—but again there is a question about simultaneity—are banks more likely to be allowed to diversify in places where everything else is going quite smoothly?

Caveats aside, the results systematize knowledge and show difficulties with the existing set of Best Practices—even if it's not entirely certain what a better set should include.

Annex 3: Research Quality Indicator Form

S	Superior
AA	Above Average
A	Average
BA	Below Average
U	Unacceptable

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

Growth and Development in Africa Project 99

5 projects, including a book:

- David Dollar , Shantayanan Devarajan , Torgny Holmgren, *Aid and Reform in Africa*
- George R. Clarke and Claude Menard, “A Transitory Regime: Water Supply in Conakry, Guinea.” Book Chapters from *Thirsting for Efficiency*.
- Alan Winters and Zhen Kun Wang, “Africa’s Role in Multilateral Trade Negotiations: Past and Future,” *Journal of African Economies*, vol. 7 Supplement 1, June 1998, pp. 1-33
- Shantayanan Devarajan, William Easterly, and Howard Pack, “Is investment in Africa too high or too low? Macro- and micro-evidence,” *Journal of African Economies* 10, Supplement 2:81-108, September 2001.
- Shantayanan Devarajan, William Easterly, and Howard Pack, “Low Investment Is Not the Constraint on African Development,” *Economic Development and Cultural Change* 51(3), April 2003.

A. Objectives

1. What specific questions or hypotheses did the study seek to answer?

The pieces of the project are united in seeking to understand the relationship between aid and growth in Africa. They are interested in the political and economic dynamics that can make aid be a successful part of reform—or that can contribute to aid delaying reform or compounding crises.

The Dollar, Devarajan, and Holmgren book, *Aid and Reform in Africa*, which is the major output of the project, asks (p. 4; this is verbatim):

- a. Are there common characteristics of successful and failed reformers that enable us to understand better the political economy of reform?
- b. Do donors tailor their assistance to different types of countries and to different phases in the reform process, or just provide undifferentiated assistance?
- c. Have large amounts of financial aid to countries with poor policy sustained the bad policies?
- d. Has aid played a useful role in the successful reformers? If so, how?

Devarajan, Easterly and Pack ask whether low investment is the root cause of Africa’s growth failures—and argue that it is not.

Clark and Menard investigate a water project in Guinea to assess the role of the private sector in a context of weak institutions. They seek to know whether the private sector can provide solutions when governments are dysfunctional—the answer is “yes, but...”

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

They’re clearly important for Africa--at the heart of debates on development policy.

B. Design and Implementation

1. Was the methodology appropriate and well executed? Was it innovative?

The method for *Aid and Reform in Africa* is country case studies, and the authors are an impressive lot. The case studies provide richness and historical context that is helpful. The syntheses, on the other hand, rely on various cross-country regressions, which are generally less helpful.

The case studies can be dry, and the overview (chapter 1) plays an important role. In many ways I wish there were a larger, richer overview (and that it was more analytical).

The two Devarajan, Easterly, and Pack papers are very close to each other and are essentially the same output. The paper (either one, really) relies on two sources of information: cross-country analysis and a study of the Tanzanian manufacturing sector. It is a clear, well-written paper, but comes to quite different (and more pessimistic) conclusions from the book. Is it because of the different methodology? Readers can't say since the methods (cross-country analysis vs. case studies) are so different. At a broad level, this is a methodological tension for the entire project.

Devarajan, Easterly, and Pack recognize the endogeneity of private investment in growth regressions. They instrument using the level of private growth at the start of the period. It's a common method, but statistically minded readers will ask whether instrumenting in this way solves the problem given the high (presumed) autocorrelation in private investment.

The Clark and Menard paper on water in Guinea is nicely done – a good narrative structure and well-told story.

2. 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?

The case studies are dense and rich. The Devarajan, Easterly, and Pack paper tends to paint with a broad brush, and, except for the exploration of manufacturing in Tanzania, offers relatively little country-specific detail. Attention is instead on broad regression results.

3. How reliable were the data? If appropriate, were surveys properly designed and executed? Were data compiled properly from appropriate sources and aggregation? Were data limitations considered in the analysis?

The cross-country parts of the analysis rely on data that is known to be weak, but it's the best available.

4. Are the conclusions consistent with the research findings? Were problems/concerns noted and reviewed?

The temptation to make strong, universal-sounding declarations based on evidence taken from cross-country regressions is not resisted here. The regressions help in providing a broad picture and describing tendencies in the data. But, at the same time, the underlying heterogeneity of country experiences and constraints ought to be borne in mind. That does not seem to happen here (or elsewhere) as much as it should.

In the Clark and Menard case study from Guinea, the authors are refreshingly open in discussing both successes and failures of the private sector's involvement in water provision. Yet their conclusions are surprisingly booster-ish for the role of the private sector. Given that prices rose far higher than expected and that problems are ongoing, the positive tone in the introduction and conclusion is jarring given the unraveling of the underlying story sandwiched in between. The introduction and conclusion read as if the message was spun to accentuate the positives. Perhaps it was not deliberate, but the moments of advocacy are unnecessary: the positives come through on their own and take particular force in the context of a decidedly difficult experience.

C. Accessibility

1. Were project reports and publications presented in a manner appropriate for and accessible to the intended audience(s)?

The book is clearly written and well organized, as are the other pieces.

2. If applicable, are policy recommendations commensurate with findings?

My assessment here follows from my comments in B4 above re the conclusions. In general, policy recommendations are commensurate with the papers' conclusions from the analysis. I do, though, have questions about how generalize-able the basic conclusions are.

D. Results

1. What are the key findings of the study? How do they advance country policy in the field?

One of the positive aspects of this project is that the papers are not in full agreement. At the same time, it would have been helpful to see some recognition of the disagreements and attempts at reconciliation.

The Devarajan, Easterly, and Pack paper argues that investment has very low returns in Africa, and, given that, investment levels are surprisingly high. The finding runs counter to a thread in the cross-country growth literature that locates investment rates as a strong correlate of growth.

It also runs counter to some of the positive sentiments in *Aid and Reform in Africa*, where a case is made for productive investment in countries that have successfully reformed.

The main finding in *Aid and Reform in Africa* is that domestic political economy matters more than international actions. Crisis tends to lead to either radical reform (Ghana, Uganda) or radical decline (Nigeria, Democratic Republic of the Congo).

Giving aid hurt countries with weak policies, delaying reform—while it helped to support countries committed to reform. Technical assistance matters early, and conditionality can be effective – but much less so in later stages of reform. (It should be noted that Easterly has argued against these kinds of conclusions in subsequent work, and the debate is unresolved—but the results here do seem plausible.)

Another important finding of the book is that success does not seem to hinge so much on formal democratic institutions and good economic policy as much as a political process that yields consensus.

The Winters and Wang paper on multilateral trade negotiations fits in least well with the other papers here. It uses a bit of data, but is mainly a smart synthesis and analysis of Africa's stake in future trade talks. The data are used mainly to round out the arguments, rather than being the central object of investigation. The most important conclusion is that countries in Sub-Saharan Africa are due some consideration and are in a position to argue for concessions—a cause that economists can aid by helping to develop the analytical case.

Annex 3: Research Quality Indicator Form

S	Superior
AA	Above Average
A	Average
BA	Below Average
U	Unacceptable

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

Overall Assessment

Overview

This note reflects on 13 World Bank research projects completed between 1998 and 2005. They each focus on financial and private sector development and were randomly selected from the larger portfolio of Bank research. Some are large, comprising books and related papers, while others involve just a single, focused study. The topics are central to the Bank's mission and of keen interest to many member countries. Projects range from an edited volume on aid and reform in Africa, a new database on bank supervision and regulation, and a series of studies of microfinance impacts.

In the main, the research is thoughtful, imaginative, and shows impressive initiative. Many of the working papers are destined for quality peer-reviewed academic journals or are already published there. Many studies rise to a high level of technical sophistication and empirical firepower. Few seem overtly driven by ideology. It should be said that nearly all stick close to mainstream economic thinking as taught in leading US and British universities, but that's overall a strength. A substantial minority question existing or past World Bank policy.

At the same time, taken as a group, the projects hold too little immediate relevance for policymakers. Many projects remain at too high a level of aggregation to speak to country-specific debates. Even the country-specific work is nearly entirely divorced from the political concerns of implementers. As such, conclusions are apt to be only modestly useful—and sometimes misleading—for those who most need the results.

There is a viable model in which World Bank research focuses on high-level broad brush analyses while others, located in developing countries perhaps, adapt the knowledge to specific country circumstances. My (untested) premise is that this final step does not occur often enough—and thus the tenor of this note.

The broad pattern of lost opportunities suggests that the failure to generate the most relevant analyses does not rest with individual researchers. The pattern suggests that outcomes result from the incentives and culture created by Bank management.

Introduction

A friend is a financial reporter who occasionally reviews off-Broadway plays for the *Financial Times*. He complains about the *New York Times* drama critics; they're too dismissive, he says. They criticize plays for not rising to some notion of what might have been--if only the *Times* critic had been given the chance to re-write, re-conceptualize, or re-cast the play. The sometime-*FT* critic cries foul, instead judging the plays on their own terms and by their own ambitions, and he often finds more to like. The initial choice of vantage is the largest determinant of whether he picks or pans.

I find myself in a similar position. Judged on their own terms and by their own ambitions, there is a great deal to like here. The projects are, for the most part, clearly conceived, conscientiously executed, and well written. Some are staggeringly good from an academic perspective; even the weaker ones have redeeming attributes. The World Bank has drawn on an impressive group of scholars, and the results move academic debates forward and, in some cases, generate new data and define new avenues for inquiry.

The picture shifts sharply, though, when the research is judged by what it is not—or by what it might have been. My comments fall into 7 bins:

1. Too much exploitation of comparative advantage: too exclusive reliance on cross-country analysis

Researchers at the Bank are well positioned to complete projects involving complicated data collection and compilation. The LSMS household surveys are the clearest example, and the research reviewed here shows other types of examples. Specifically, the research teams have compiled large cross-country or institution-level data sets of various kinds. For example:

- Bank supervision and regulation (107 countries, 175 questions, 1999)
- Internationalization and globalization of financial sectors
- Legal procedural formalism in adjudicating business disputes (109 countries)
- Costs (mainly in terms of corruption) of regulation of entry by firms.

The data provide new and comparable evidence on a range of important issues. They allow the description of broad patterns across countries, which can be very helpful in framing policy debates.

The analyses, though, are too often pushed harder than they should be. The country-level data reflect considerable heterogeneity of history, geography, constraints, and opportunities. The coefficients in the resulting cross-country regressions are (complicated) weighted averages of the behavioral responses of the hundred or so countries in the data sets. They capture general tendencies, and there are apt to be important exceptions—or, as importantly, the suspicion of exceptions. The problem is that we learn little about the exceptions. For country-specific policy, though, the ability to resolve debates hinges on determining how exceptional (or unexceptional) your own country is apt to be.

The issue of heterogeneous responses is a universal one in empirical economics. But in household-level analyses the policy implications of heterogeneity tend to be far less severe. If you run a regression on data from Tanzania, for example, the regression coefficients will in general reflect an average of behavioral relationships in Tanzania. For a policymaker in Tanzania, the broad tendency in the country is most relevant. There is less concern that Mr. X in village Y has behaved somewhat differently.

The issue is different when it comes to policy implications drawn from cross-country data. The data points are countries, and a policymaker in country X cares indeed that behavior is different at home than in some broad mass of countries for which a researcher could find data. The exceptions matter.

A nice example of the importance of looking carefully at exceptions (drawn from DECRG, but not from the articles I reviewed) is Martin Ravallion's 11/01 *World Development* article, "Growth, Inequality and Poverty: Looking Beyond Averages" (which plays off another DECRG work by Dollar and Kraay: "Growth is good for the poor"). Ravallion shows that focusing on average responses of poverty to broad-based

growth obscures contradictory and quantitatively important scenarios that suggest a wider range of policy options.

A second concern rests with the coarseness of policy variables. Most are at such an abstract level (an index of financial liberalization, for example), that mapping into specific (and imperfect) policy levers is not straightforward.

A third concern is methodological and well-rehearsed by others: too few solid ways to address policy endogeneity (although some clever ideas are implemented in the papers I reviewed); too great a concern with omitted variables; too few relevant degrees of freedom; issues around heterogeneity (as noted above); and a fear that results are sensitive to the inclusion or exclusion of particular control variables.

All that said, when addressing major macro issues, cross-country analyses are often the only sensible strategy. The papers here show the power of cross-country analyses and they illustrate striking patterns. In the end, it's a question of balance: cross-country analyses have a place, but they cannot be the last word and they cannot substitute for careful studies of individual countries or regions.

2. Too little exploitation of comparative advantage: too little leverage of Bank operations

One of the natural questions that arises in reading the articles is: Why was this paper written by someone sitting at the World Bank rather than by one of my colleagues at NYU or any other quality academic site?

One answer lies in the type of work: as noted above, many of the papers take advantage of a comparative advantage in data collection.

But I would have thought (had I not known better) that the papers would also draw heavily on the researchers' other comparative advantage: being housed in the same building as thousands of Bank operational staff. As far as I can tell, the work here could have been written equally well if the Bank researchers were instead cocooned off-site.

Again, Ravallion's work on poverty and safety nets serves as a counter-example—many of his papers speak directly to Bank projects and draw on their experiences. Indeed, much of the work on poverty and safety nets by others (e.g., Harold Alderman's work on education and nutrition) similarly draws on Bank operations. Perhaps the apparent lack of synergy is specific to the projects reviewed here, but I would be surprised that this is so.

While there may have been much informal discussion with Bank operations staff, little of it is on display in the final products. This seems like a puzzling lost opportunity.

3. Too little applied theory?

The next comment somewhat contradicts the two above. All of the papers and studies draw on economic theory in one way or another. But there's surprisingly little straight economic theory here (i.e., where theory is the main contribution)—apart from a paper on microfinance contracts by Sanjay Jain and Ghazala Mansuri and a chapter on tax theory in a volume on taxing financial intermediation.

The Research Department should be driven by empirical work. On the other hand, the empirical work is interpreted through the lens of off-the-shelf theorizing that has been mainly developed for Western economies. Innovative theory can clarify concepts and address constraints particular to the contexts of developing economies. I'd opt for more theorizing—clarifying political equilibria around failures of policy reform, for example—but empirical work should still dominate.

4. Peer-review is an asset, but it rewards taking motivation from the academic literature

Much of the work here aims for publication in academic journals. At one level, publishing in peer-reviewed journals is the right answer to a complicated series of incentive problems. It lets hard-thinking external peers place the stamp of rigor on the work, mitigates the role of personality in promotion, and provides a natural outlet for polished papers. The quality and quantity of peer-reviewed output from DECRG is impressive. Relying on external review may be the best of an imperfect set of options.

The imperfections are nonetheless evident. Peer reviewers don't necessarily share the objectives of research managers or Bank leadership, and that creates its own set of incentive problems.

One of the most striking outcomes is that papers are more likely to be motivated exclusively (or near-exclusively) from the perspective of the prior academic literature—rather than from pressing, real-world problems.

Many papers that are published in top journals take that strategy (“This paper extends the work of Blah, 1999, etc.”). But there is a good argument that researchers at DECRG should aspire to different starting points. I say this because, by the end of several (quite good) papers here, the academic literature has also become the end point as well. This occurs despite the possibility of re-shaping the papers to speak more closely to worldly concerns.

I'm not sure how much to emphasize this point since readers can always do their own mapping of results to questions of interest—the issue is mainly one of exposition rather than statistical method. But I fear that by not insisting on more grounded motivation, papers risk slipping further toward practical irrelevance. I also suspect that the academic journals would become more lively places if there were fewer “extensions of Blah” and more “this paper addresses a puzzle in Latin America...”

5. A rush to judgment

There are many big issues about which we know too little (the impact of banking access on employment, to take a concrete example). Sometimes we have data that allow us to shine a bit of light, however imperfectly (in this case, identifying causal links is tricky). So when data or methods are imperfect, we take care to underscore the necessary caveats when drawing conclusions (the authors did so carefully in this case).

In the end, though, a host of imperfect results do not aggregate up to one pristine result (in which causality is truly nailed). With imperfect results, it's seldom possible to gauge accurately whether they're apt to be just a bit off or if they're likely to be way off.

Results may get published in peer-reviewed journals, owing to the importance of the topic and the effort expended to squeeze out answers, but it's still not clear what we have learned in the end.

Ultimately, my sense is that it's generally better to be less ambitious or to drop the research project entirely when we know from the start that methods or data are fundamentally flawed. It's not an easy mandate for a place like the World Bank, especially since decision-makers often can't wait for better answers. Sometimes knowing something can be better than knowing very little. But my experience is that we often learn much less through rough-and-ready or nearly-good-enough approaches than we imagined when starting the projects.

6. Gaps

The biggest gap has to do with politics, and I discuss it in the next section. Here, instead I focus on microfinance.

For the past few years, my work has focused on microfinance. During most of this period, researchers at DECRG have been nearly completely absent from microfinance conversations, despite the rising attention to microfinance globally. Shahidur Khandker and his colleagues undertook early studies (with important contributions also from economists working on rural finance, although not microfinance *per se*), but it is only in the past 2 years (with the UN Year of Microcredit 2005 and Stanley Fischer's suasion) that work on microfinance has picked up. I'm happy to see it—a DEC perspective is much needed. Khandker's book, *Fighting Poverty with Microcredit*, asks all the right questions--even if the answers are still up for grabs.

Without interventions by researchers at a DECRG level, microfinance conversations have been dominated by advocates and practitioners of various stripes. They have put a great deal on the table, but logical leaps and poorly done empirical studies have gone unchecked. More important, interesting (and policy-relevant) questions have gone unasked.

The failure of economists at DECRG to focus squarely on issues of financial access for the poor is a critical lost opportunity. Microfinance is being addressed now within DECRG, but there was a signal absence in 1998-2005, a period in which focused microfinance research could have made a deep impact on policy directions.

I don't know all of the reasons for the absence of attention, but I suspect that the reasons would generalize to other areas.

I suspect that one reason for ignoring microfinance for so long was skepticism within the Bank about the viability of the movement and distrust of claims made by early advocates. Another is discomfort, in the 1990s at least, at engaging in policy conversations dominated by NGOs rather than academics and government officials. Neither seems like a particularly good reason for staying out.

I suspect that the attention gap was also a product of the ambivalence of microfinance policymakers to actively engage with Bank counterparts. The microfinance movement was bent on re-imagining policy-making by putting a strong emphasis on the private

sector, and I suspect there was ambivalence about engaging Bank staff who tend to think more comfortably in terms of traditional public finance frameworks. There were some contacts, but not strong bonds. I suspect that there was also reluctance on the part of microfinance policymakers to encourage evaluations of microfinance.

I should add that some important empirical work on financial access has been undertaken, but not (I believe) by researchers at DECRG. The studies of financial access by Anjali Kumar, Priya Basu, and others have helped fill knowledge gaps—with more work to come.

Within DECRG, recently Patrick Honohan has written a smart synthesis on the sector, and Xavier Gine, Thorsten Beck, Maria Soledad Martinez Peria, Asli Demirgüç-Kunt, and Robert Cull, among others, have made important contributions. In operations, Marilou Uy has emerged as an articulate speaker on financial access and a respected Bank representative at microfinance conferences.

7. Political analyses

The work here is dominated by economists, with the exception of the volume on aid and reform in Africa, which has a strong politics thread. That book shows the value of having more political scientists at the table.

One of the big gaps in the research here surrounds implementation (or the lack of it). Much of the research argues for specific policy interventions, based on correlations of observed policy positions and subsequent economic outcomes. But little attention is paid to policy endogeneity. I should correct that: the problem of policy endogeneity is noted clearly in motivating instrumental variables methods to correct for resulting biases.

What's most missing is the flip side: rather than just searching for exogenous variation (important as it is for identification), it would be helpful to also understand the nature of endogeneity itself: Why do some governments manage to implement reforms? Why do some policies emerge in one place but not in others?

Taking up these kinds of questions more frequently would help in the mapping of policy discourse to public action. It would sharpen analysis and help focus on feasible solutions that stand some chance of meaningful implementation.

Concluding thoughts

It has, for the most part, been a pleasure to engage with the projects here. The World Bank has managed to hire an extremely talented group of researchers, and they have been productive and creative. They navigate constraints that other researchers don't have to, and meanwhile keep a focus on big questions of wide interest. World Bank research, taken as a whole, has had powerful impacts on policy around the world. At the same time, there have been lost opportunities and omissions. Some are small, but some are fundamental. The last issue above, ignoring politics, is the most troubling gap.