Evaluation of World Bank Research on Infrastructure
Michael Kremer

Individual Research Projects

Project: Infrastructure, Location & Development in Nepal (5 papers)

Total Funding: $ 323,200

Paper (1 of 5): The Spatial Division of Labour in Nepal (Marcel Fafchamps and Forhad Shilpi); JDS, 2003

A. Objectives

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

This interesting article examines how economic activity and market participation are distributed across space, with an application on Nepal. In particular, this paper examines how Nepalese households fit into the local economy as a function of their proximity to urban centers.

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

The authors argue that spatial effects are important for policy. Their results, for example, show that households located more than five hours away from towns and markets do not sell crops and do not consume fertilizer. Any effort to promote agricultural innovation would need to take these effects into account to be successful. It is clearly important to know whether and how nearby city size affects economic activities (e.g., whether people are employed in farm or nonfarm work) and whether and how nearby economic activities affect city size.

B. Design and Implementation

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

The authors use nonparametric methods on household data from the 1995/96 LSMS coupled with census data to investigate patterns of agricultural production, agricultural sales and purchases, and nonfarm work as a function of distance to the nearest urban center, measured by travel time. Dependent variables are the share of household labor going into nonfarm and farm activities, broken down by wage and self-employment in these sectors. Creating this extensive spatial data set is innovative and the use of nonparametric methods made detailed analysis possible – as their results show, the effect of distance is often highly non-linear.

The authors argue that the contribution is threefold: (1) by using several different measures of economic activity and market participation, they provide a detailed and comprehensive picture of the spatial division of labor in a poor country; (2) using nonparametric methods they obtain precise estimates of the distance at which various activities dominate; and (3), they show that spatial effects are large and that city size matters. For example, proximity to markets and towns fosters nonfarm activity, especially wage nonfarm employment.

On methodology, an important challenge is the correction for endogeneity of urban population and travel time. As the authors point out, it is conceivable that towns are larger whenever the surrounding countryside produces a food surplus. Observing that wards located closer to large towns sell more food would then be the result of reverse causation.

To address these concerns, they instrument city population and travel time using physical characteristics of the district in which it is located, such as size, arable land area, distance to nearest river, mean elevation, and whether it is mountainous. It seems, however, that these instruments would affect, for example, agricultural surplus and hence the share of household labor going into nonfarm and farm activities, net of urban population size, and as such would not meet the exclusion restriction.

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?

Yes, the literature review – both on the theory of agglomeration effects and relevant research on Nepal – is extensive.

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?

LSMS data were used which are detailed, extensive, and presumably reliable. Census information on town size was also used and travel time to nearby cities was calculated for all the homes in the data set.

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

The authors conclude that findings are largely consistent with the von Thunen hypothesis, that rural areas surrounding cities specialize in different agricultural products. This conclusion seems appropriate since this is what the data indicate. There was little discussion on the appropriateness of the instruments.

C. Accessibility

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

Yes.

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

There are few policy recommendations.

D. Results and Cost-effectiveness

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

The paper finds a strong spatial division of labor. Nonfarm employment is heavily concentrated in markets and in and around towns. Agricultural wage employment is concentrated in rural areas sufficiently close to cities that they can specialize in commercial crops but neither so close that nonfarm employment takes over, nor so far that they revert to self-subsistence. Crop choices also vary with distance to town and with their size. For example, vegetable production in both seasons is somewhat concentrated at the vicinity of markets and urban centers. Isolated households and villages essentially rely on self-subsistence.

It is clearly important for policy makers to know that there is a strong spatial division of labor, and have an interesting paper demonstrate this convincingly even if this would likely confirm the prevailing beliefs of most of the policy makers.

2. Were the results consistent with the costs? Did the project take adequate time to complete?

Yes, the results were consistent with the costs.

Project: Infrastructure, Location & Development in Nepal (5 papers)

Paper (2 of 5): Cities and Specialization: Evidence from South Asia (Marcel Fafchamps and Forhad Shilpi); Economic Journal, 2005

A. Objectives

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

This interesting paper examines the relationship between proximity to urban centers and the organization of labor using data from Nepal. It builds on a previous paper, which focused on exploring the relationship between proximity to urban centers and the sectoral division of labor (mainly agriculture vs. non-agriculture).

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

Yes, knowledge of the organization of labor has important industrial policy planning implications.

B. Design and Implementation

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

The methodology was appropriate and carefully executed, and the use of geographical data with detailed labor force survey data was innovative. The data used come from the Nepal Labour Force Survey of 1998/99, which covers 14,355 urban and rural households spread among 719 villages distributed over 73 of the 75 districts in the country. Respondents were asked how much time they worked over the seven days preceding the survey and how this was distributed over 16 activities that can be grouped into 3 categories: (1) market work; (2) subsistence related work largely for self-consumption; and (3) household chores. Respondents were also asked on the sector of employment. Information on firm size (number of employees) and employment type – managerial or administrative tasks – is used to identify the occupational structure of the labor force.

The authors use nonparametric methods to investigate the organization of labor as a function of distance to the nearest urban center, measured by travel time. In particular, they separate the investigation into specialization through the market and within firms. Dependent variables aimed at measuring specialization through the market are (1) the sectoral division of labor measured by whether an employee works in sector (i) agriculture, (ii) manufacturing and construction, (iii) private services, (iv) public services, or (v) trade, transport, restaurant; (2) individual and local specialization measured by a Simpson's index that takes on the value 1 if an individual or locality allocates all of his/its time on one activity, and using data on seven different production activities – wage work, nonfarm self-employment, agriculture, construction, food processing, handcrafts and other work; and (3) the division of labor and household chores measured by an individual's share on (i) shopping, (ii) cooking and cleaning, (iii) fetching water and firewood, (iv) caring for children and elderly, and (v) other chores.

Results on the sectoral division of labor are also reported in the authors' 2003 JDS article above. With regards to individual and local specialization, individuals residing close to large population centers have a higher specialization index, meaning that they are more focused on a few productive activities. Localities show the opposite pattern; they are more diversified in or near cities. The results on household chores are mixed. Dependent variables aimed at measuring specialization within firms are (1) whether a worker is salaried or self-employed (broken down by sector and for all sectors); and (2) firm size (broken down by sector and all sectors).

Results show (1) a strong relationship between proximity to cities and wage employment for all sectors combined, which is taken as preliminary evidence that town proximity is associated with larger firm size, but not within individual sectors except for agriculture; (2) firms are larger near cities – but only within the agricultural sector. In fact, the results indicate that cities have larger firms because they harbor more activity in nonagricultural sectors where firm size is larger. Within sectors, firm size is only weakly associated with distance to and size of urban markets.

The latter seems to be one of the most surprising things – when agriculture is left out, there is only a weak link between proximity to urban areas and self-employment and between proximity and firm size. As such, these findings fail to provide evidence of strong agglomeration effects near cities. Specialization and the division of labor are largely through the market, not within firms.

This raises the question – although answers are not much speculated on – of why firms aren't larger near cities. Human capital is higher near cities yet it does not give rise to more complex and bigger firms (at least not within sectors). Could moral hazard be an issue that larger firms run into but smaller firms can avoid?

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?

Yes, the literature review – both on the theory of agglomeration effects and relevant research on Nepal – is extensive.

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 were appropriate for the research question, and described in sufficient detail.

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

Yes, the conclusion is consistent with the research findings. I would have liked it if the authors had speculated more about the lack of evidence for agglomeration effects.

C. Accessibility

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

Yes.

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

Yes.

D. Results and Cost-effectiveness

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

See above.

2. Were the results consistent with the costs? Did the project take adequate time to complete?

Yes.

Project: Infrastructure, Location & Development in Nepal (5 papers)

Paper (3 of 5): Isolation and Subjective Welfare (Marcel Fafchamps and Forhad Shilpi); 2005

A. Objectives

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

This article investigates the relationship between isolation (measured by distance to nearest market, urban population within two hours, and population density) and subjective welfare (measured by answers to questions on subjective adequacy of food, clothing, housing, schooling, health care, and income).

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

Yes, this article suggests that measures to decrease isolation, for example by building roads, have welfare enhancing effects.

B. Design and Implementation

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

The methodology was appropriate and carefully executed, and the use of geographical data with detailed consumption survey data was innovative. The data used come from the Nepal 1995/96 Living Standards and Measurement Survey. The dependent variables were subjective measurements of welfare, proxies for utility. The topic of the paper, isolation & welfare, is novel, and the authors carefully addressed the theoretical approaches to interpret their results – demonstrating, for example, that their approach holds under

mobility or lack thereof, and that the use of subjective welfare measures such as consumption adequacy responses, although imperfect, has been shown elsewhere to correlate with real measures of well-being such as consumption and brain level activities.

The results obtained after controlling for consumption expenditures and prices (wages and rice) suggest that isolation has welfare costs over and above its effect on monetary consumption. Controlling for household mobility and various other controls leave the results unchanged.

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?

Yes, the literature review – both on the conceptual framework – individual location choice – and use of subjective measures of welfare are extensive and detailed. Knowledge of Nepal is extensive, as also demonstrated by previous papers using data from Nepal.

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 were appropriate for the research question, and described in sufficient detail. Data limitations were considered (see discussion above).

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

The conclusion is consistent with the research findings. Something perhaps surprising that was not so much emphasized by the authors is the finding that controlling for access to amenities such as travel time to nearest school does not substantively affect the distance coefficients. This is perhaps surprising given the model, which emphasizes that controlling for consumption, differences in subjective welfare should reflect differing amenities – with distance to nearby cities serving as a proxy. That access to basic amenities does little to affect the distance coefficient suggests that distance is capturing other important dimensions of isolation. Reducing isolation by building roads may therefore address the isolation aspect of welfare improvements more than building schools and health care facilities would.

An alternative interpretation of the results could be the following: Suppose people do not care about amenities. Instead they care about their income (consumption), which is a function of ability and of distance to nearest town, and they care about the income of others with the same ability elsewhere. Suppose also there is a high cost of moving. Then even conditioning on income, people living further away will be unhappy since they know their ability peers living nearer to cities earn more. Again, building roads would be welfare improving by reducing the cost of moving.

C. Accessibility

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

Yes.

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

Yes.

D. Results and Cost-effectiveness

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

The results obtained after controlling for consumption expenditures and prices (wages and rice) suggest that isolation has welfare costs over and above its effect on monetary consumption. Controlling for household mobility and various other controls leave the results unchanged. As the authors rightfully point out, these results imply that welfare assessments based on geographical poverty maps underestimate the subjective welfare cost of isolation since their focus is usually income and consumption. The authors also point out that in addition to the search for higher earnings and better employment, rural dwellers are likely also to migrate to reduce their isolation.

2. Were the results consistent with the costs? Did the project take adequate time to complete?

Yes.

Project: Infrastructure, Location & Development in Nepal (5 papers)

Paper (4 of 5): Subjective Welfare, Isolation, and Rivalry (Marcel Fafchamps and Forhad Shilpi); 2005

A. Objectives

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

This interesting paper uses LSMS data from Nepal coupled with distance data. Using answers to subjective consumption adequacy questions, the authors investigate whether people exhibit rival preferences, which are defined as preferences where someone's utility increases with own consumption but falls with the consumption of others. They also investigate the relationship between the presence of rival preferences and how

geographically isolated a person is, where the latter is defined as distance to the nearest market and the total urban population located within two hours of travel time from the locality.

They find evidence in favor of rival preferences, especially in geographically remote areas.

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

The paper has important implications for development policy. In particular, as the authors point out, recent years have witnessed a renewed interest in decentralization as a way of improving population participation in public affairs. This is especially true in developing countries where increased emphasis has been put on fiscal autonomy and local control over public expenditures (e.g., Besley, Pande, Rahman & Rao 2004; Besley & Burgess 2002). These efforts are often predicated on the assumption that social and political conflicts that plague decision making at the national or regional level are reduced at the local level because of greater social homogeneity; people in small communities have much in common, and therefore should be better able to organize the provision of public goods and the monitoring of government officials on issues that interest them directly. This is thought to be particularly true in a village context, where information circulates widely and repeated interaction should solve commitment problems.

The finding by Fafchamps and Shilpi that preferences are rival, particularly so in geographically isolated areas, is consistent with various examples in their review of the literature. As they point out, observers of decision making at the local level often bring back a less rosy picture from their fieldwork. In their book on the management of communal resources, Baland & Platteau (1995) provide numerous examples of small communities being unable to coordinate the provision of simple public goods. In their analysis, the authors repeatedly emphasize the deleterious effect of heterogeneity, a point that they revisit in subsequently published articles (e.g., Baland & Platteau 1998; Baland & Platteau 1997; Baland & Platteau 1999). Wade (1988) makes similar comments in his work on the management of irrigation systems in India.

These observations question how far decentralization can go as a way of improving the level and quality of participation in local government affairs.

B. Design and Implementation

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

The methodology was appropriate and carefully executed, and the use of geographical data with detailed consumption survey data to explore the relationship between rivalry and isolation is innovative. The data used come from the Nepal 1995/96 Living Standards and Measurement Survey.

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?

Yes, the literature review – both on the conceptual framework – rival preferences – and use of subjective measures of welfare is detailed. Knowledge of Nepal is extensive, as also demonstrated by previous papers using Nepal data.

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 were appropriate for the research question, and described in sufficient detail. Data limitations were considered in the discussion on using subjective questions.

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

The conclusion that Nepalese households evaluate the adequacy of their consumption by comparison with that of others like them is consistent with the research findings. The authors avoid making strong statements but rather cautiously couch their findings in terms of "we interpret this finding as evidence that…" or "..a finding consistent with the idea that.."

C. Accessibility

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

Yes.

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

Yes.

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 are that answers to consumption adequacy questions improve with own consumption but fall with the average or median consumption of others in the same ward. For food, clothing, and schooling, they cannot reject the hypothesis that surveyed households judge their consumption adequacy solely on the basis of relative expenditures – i.e. if an increase in household expenses is matched by an increase in ward average expenses, people do not respond being better off. For housing, a proportional increase

actually makes them worse off. The size of the 'rivalry' is surprising and has mixed implications for development policy. A dismal implication is that "lifting all boats" will not change responses to consumption adequacy questions. On the other hand, focusing on improving food, clothing, and schooling for the very poorest within communities can lead to considerable increases in consumption adequacy responses (among the poorest – it would decrease it for those who are better off and do not see their consumption increase relative to the community average) since it will increase individual consumption expenses more than the community average. Their result that rivalry is strongest in isolated communities suggests that the latter policy implication of targeting especially holds in those isolated communities. An alternative approach to increasing welfare is to reduce rivalry by reducing isolation through investments in roads, access to information, etc.

3. Were the results consistent with the costs? Did the project take adequate time to complete?

Yes.

Project: Infrastructure, Location & Development in Nepal (5 papers)

Paper (5 of 5): Gender, Generations, and Nonfarm Participation (M. Shahe Emran, Misuzu Otsuka, and Forhad Shilpi); 2003

A. Objectives

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

This paper uses the Nepal LSMS data 1995/96 to study intergenerational links in nonfarm participation with a focus on gender effects. In particular, how are the work choices – (i) participation in nonfarm vs. farm employment, and (ii) skilled vs. unskilled work – of sons and daughters correlated with those of their mothers and fathers? The authors find evidence that mothers' and daughters' choices are strongly correlated. Daughters whose mothers work in nonfarm employment are 200 percent more likely to be working in nonfarm employment. Daughters whose mothers work in skilled nonfarm employment are 1,200 percent more likely to be in skilled nonfarm employment. There is no significant effect on a son's participation in nonfarm employment when the father is in the nonfarm sector. A moderate positive correlation exists between fathers and sons for skilled jobs.

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

Yes, as the authors point out, poverty has a gender dimension with women among the poorest and chronically deprived segment of the population. Access to nonfarm employment can be a way out for poor and landless people in general, and women in

particular. Nonfarm employment by women may also increase their household bargaining power. Further, participation rates in the nonfarm sector are 45 percent for men and 22 percent women. It is important to understand how much of this gender bias is due to occupational immobility caused by gender-specific intergenerational linkages. If there are strong intergenerational linkages in women's nonfarm participation, then the long-run benefits from women's participation in the nonfarm sector are likely to be much higher due to the intergenerational multiplier effect. However, although it is an important research topic, it does not seem to directly inform policies on infrastructure.

B. Design and Implementation

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

The authors recognized the difficulty in identifying the effect of parent's occupation on children's occupational choices, and point out that (1) an intergenerational correlation may result from the fact that parents and children may face similar labor market opportunities (for example, they live in an area with better nonfarm opportunities), and (2) predictors of nonfarm employment such as education may be endogenous if they are simultaneously chosen.

The authors address the first concern by controlling for village fixed effects. To address the second, they employ a two-stage conditional maximum likelihood approach. They suspect two endogenous variables, education and assets, and use as instruments for education the village average distance to school, also interacted with ethnicity, and with inherited agricultural land; for assets they use inherited agricultural land and its squared term.

It seems, however, that inherited agricultural land would not be a valid instrument since it should be entered directly in the second stage regression predicting nonfarm employment. Further, the regressions include several potential endogenous variables, such as whether the household has a migrant and remittance amounts. This complicates interpretation of results.

Further, in light of the striking mother-daughter correlations, it would have been helpful to see if the intergenerational linkages also hold up in more narrow employment definitions, predicting the likelihood that the daughter/son follows the mother/father in choosing categories such as professional and technical worker, administrative and managerial worker, etc.

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?

Yes, the literature review and the conceptual framework are detailed.

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 were LSMS survey data, and thus reliable. Data limitations were considered in the conceptual analysis. Given the available data, these were appropriate.

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

The conclusion that the empirical results show strong intergenerational correlations, especially for daughters, in nonfarm participation, is consistent with the research findings. Unfortunately, answering the question most relevant to policy is difficult with these data: Do the intergenerational correlations persist if policies were designed that encouraged women whose mothers are currently in farm employment to choose nonfarm employment? In other words, would the granddaughter continue in nonfarm employment if the grandmother was in farm employment and the mother had been induced to choose nonfarm employment? Further, while nonfarm employment may currently be associated with higher earnings, it is not clear if policies that encouraged women to switch to nonfarm employment would be able to provide access to these higher paid jobs; it is possible that switchers would largely have access to low paid nonfarm employment only. Further, if large numbers of individuals switched, there would likely be general equilibrium effects reducing the benefits to nonfarm employment.

C. Accessibility

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

Yes.

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

Yes, with the above caveat.

D. Results and Cost-effectiveness

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

Please see discussion above.

2. Were the results consistent with the costs? Did the project take adequate time to complete?

Yes.

Project: Motorization and Urban Transport (5 papers)
Total Funding: $ 1,268,800

Paper (1 of 5): Traffic Fatalities and Economic Growth (Elizabeth Kopits and Maureen Cropper); Accident Analysis and Prevention, 2005

A. Objectives

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

Using panel data from 1963 to 1998 on the rate of motorization (vehicles/population or V/P) and on fatalities per vehicle (F/V), this paper examines the relationship between traffic fatality risk and per capita income and uses it to forecast traffic fatalities by geographic region through 2020.

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

From a public policy perspective it is clearly very important to have accurate predictions about traffic deaths. Global traffic deaths are estimated to rise from about 800,000 today to 1,200,000 in 2020.

B. Design and Implementation

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

While other studies have made projections of vehicle growth or the automobile fleet, these other studies relied on information for passenger cars and commercial vehicles only. For this article, the authors' motor vehicle counts include all buses and two-wheelers. Further, the data on traffic fatalities and vehicles used come primarily from the International Road Federation Yearbooks, which provide data that are more disaggregated for Sub-Saharan Africa than data used in other studies. For example, the Global Burden of Disease study by the World Health Organization (1996) relied for its 1990 estimates only on data from South Africa.

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?

Yes, the literature review is extensive.

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?

How reliable the data were is difficult to tell with such data. However, the authors mention to have cross-checked the data against various other sources.

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

The authors find that the per capita income at which traffic fatality risk (F/P) begins to decline is $8,600 (1985 international dollars) when separate time trends are used for each geographic region. This turning point is driven by the rate of decline in fatalities per vehicle (F/V) as income rises since vehicles per population never declines with economic growth. This is similar to findings from other studies.

Consistent with their findings, the authors conclude that, if developing countries follow historic trends, it will take many years for them to achieve the motor vehicle fatality risks of high-income countries. This is primarily due to the fact that developing countries are still many years away from reaching the $8,600 benchmark, if historic income trends continue.

It would have been useful to get a sense of the lower bound on traffic deaths / 100,000 people in developing countries for the period 2000-2020. For example, how would the interpretation of the findings change if developing countries would get access to the same technology and safety standards currently available in rich countries? And, which developing countries stand out as having low traffic fatalities per 100,000 and what can we learn from those countries in designing traffic safety policies in the future?

In fact, the policy value of this research could be greatly enhanced if the authors could identify specific policy instruments by linking these data with existing data sets on the regulatory framework governing traffic (e.g. speed limits, driving age, vehicle safety standards, etc.) and information on road infrastructure. If these data sets do not yet exist, the World Bank could leverage its comparative advantage and collect such cross-country data. Additionally, being able to attach prices to these intermediate variables would allow the authors to come up with cost estimates for reducing traffic deaths at various levels of economic development. In sum, although the current work is informative, the returns to going the extra step are large.

C. Accessibility

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

Yes.

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

The authors point out that the central question for policy in low-income countries is to identify the factors that underlie the decline in fatalities per vehicle and to implement

policies that are cost effective. Some suggestions are made, although the authors could have speculated more on these factors and appropriate policies.

D. Results and Cost-effectiveness

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

Please see discussion above.

2. Were the results consistent with the costs? Did the project take adequate time to complete?

Unknown – it is not clear what share of the $1.3m this particular research cost.

Project: Motorization and Urban Transport (5 papers)
Total Funding: $ 1,268,800

Paper (2 of 5): Urban Poverty and Transport: The Case of Mumbai (Judy Baker, Rakhi Basu, Maureen Cropper, Somik Lall, and Akie Takeuchi); Policy Research Working Paper, 2005

A. Objectives

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

This paper is part of a larger study. The goal of this study is to better understand the demand for transport services by the poor, the factors affecting this demand, and the inter-linkages between transport decisions and other vital decisions such as where to live and work.

The goal of this particular paper is to describe the residential and work locations and travel patterns of households in Mumbai using a survey of 5,000 randomly sampled households collected by the authors. The authors pose various questions to investigate:

• What is the spatial distribution of households by income (consumption) in the Greater Mumbai Region? How segregated are various income groups?
• How does the distance between residence and employment vary (a) by income group; (b) by location of residence?
• How do the number of trips made and modal split vary among income groups? What is the demand for trips other than trips to work? How do the poor utilize transport for daily errands such as shopping?
• How does access to basic social services such as education and health care vary across location and income? Is it the case that access is the major problem for the poor to use the service?

• Does the cost of public transit result in the lower mobility of the poor? This question is mentioned but actually intended for more thorough investigation in subsequent papers.
• Does low mobility imply that the poor are worse off than if they traveled more? This question is similarly mentioned but the authors suggest that evaluating the welfare effects of a reduction in the cost of transit is actually a subject for further research.

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

The authors rightfully point out that understanding the linkages mentioned above should ultimately help to design transport policies that will help the poor. The paper lacks, however, a conceptual framework that would help in interpreting the empirical analysis and would provide guidance on the evaluation of the welfare effects. Clearly, individuals benefit if there is a price reduction in transport, just as they benefit if prices of other goods are reduced. This, however, is not a sufficient rationale for policymakers to design policies, such as subsidies, that may lead to price reductions. Subsidies could certainly be justified for a number of reasons. The people in a country or area may find it socially desirable to reduce inequality by subsidizing transport frequently used by the poor. Or, if zoning regulations limit residential mobility in a way that industrial areas are too far (and hence too expensive) to be reached by the poor, subsidizing transportation may generate economies of scale or even learning externalities by linking employees and employers and allowing industrial development to occur. In other words, this question of how transportation prices and welfare are linked can be approached from various angles. Unless these angles are specified, interpretation of empirics is problematic.

B. Design and Implementation

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

Little is known about the transport behavior of the poor. Carrying out a survey of 5,000 households in Mumbai, India, focusing on this issue is certainly informative. The findings on the amount of income spent by the poor on transportation and the number of trips they make, as well as the information on distribution of commute distances, are interesting.

I would have thought that decisions about transport planning should be made primarily on efficiency grounds. I think that transport problems are issues for everyone in Mumbai, not just the poor. And while the poor no doubt suffer in particular from certain aspects of the situation, I think a more comprehensive look at this critical problem would have been appropriate.

Obviously it is possible that the poor tend to be neglected in deciding on social services, etc. and it is worth checking this. But I don't think it is reasonable to assume that there should be an attempt to equalize transport costs to work between rich and poor, for example. With current policies including rent control laws, difficulties with property rights, etc., it might be the case that Pareto improvements could come from a policy

where poor people can voluntarily sell rights to property near the center of town to wealthy people, and use the resources to purchase improved housing further away. There is some discussion of this, but something is lost in not having more discussions on those types of issues.

In fact, the paper could have been better had it been more structured, both conceptually and, related, in its presentation of results. It is full of tables (38 tables and 8 figures in a 47 page document detailing various dimensions of transportation behavior by the poor, such as commuting time by zonal area and income, the type of transportation used (also by various income categories, by distances, by zones), transportation used for non-work purposes, travel to school (by level, by private/public schools, by income, by age, by zone, by type of transport, walking times by these various categories), access to health services (by various types and breakdowns), distances to public transportation (by various detailed breakdowns), cost of railways and bus services by various distances, etc.

Given the volume of tables and figures, there is little room for discussion on how these tables fit into the larger picture. This makes it difficult to read and difficult to understand what lessons can be learned that can translate into policy actions.

The paper lacks a discussion on characteristics of self-employment vs. wage-employment and relation to transportation. Knowing more about employment patterns would have been helpful. For example, one of the main conclusions is that almost half of all people walk to work and more than 60 percent of the poor walk to work. How does this reflect employment patterns? Do the poor that walk work in different occupations than the poor that do not walk? Are poor men / women more likely to walk to 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?

I would have preferred more discussion of the extensive geography and economic development literature.

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 survey seems to have been carefully designed so as to be representative of the Greater Mumbai Region. The appendices to the paper provide a detailed account of the survey design. Information on refusal and/or error rates is lacking, however, making it difficult to judge the quality of the data.

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

The conclusion is as well organized as desirable. It highlights some of the main empirical observations, but does not attempt to synthesize the findings into a "conclusion."

C. Accessibility

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

The paper contains a great amount of information but due its poor presentation it is not as accessible as desirable.

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

The policy implications are deferred to future work.

D. Results and Cost-effectiveness

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

The main findings are that many people walk to work, poor people especially, and among those that do not walk, the train is a common means of transportation. It also finds that the poor in Zones 1-3 of Mumbai live closer to the non-poor than do poor households living in the suburbs.

Policy implications are not yet very clear. For example, since many poor people walk to work, perhaps one recommendation should be improvement of infrastructure for cycling. However, without knowing more about employment patterns, current cycling conditions, etc., it is not possible to make concrete policy recommendations.

The paper does discuss the issue of whether households would be better off commuting longer distances. Clearly they must be better off in some sense where they are by revealed preference. However, with different property rights arrangements they could be compensated for moving out.

2. Were the results consistent with the costs? Did the project take adequate time to complete?

The entire "Motorization and Urban Transport" project cost $1.3 million. If this were the cost of undertaking just the survey, this would amount to $260 per correspondent, which seems high for a survey in India. However, if other projects, senior research time, etc. are included, this could certainly be a reasonable amount for undertaking a survey of 5,000 respondents.

However, I am not sure that the result that came out of this paper was worth doing a study of 5,000 people, which presumably is fairly costly. I would have thought that the

spatial distribution of households by income or consumption could be obtained from other data sets but perhaps I am wrong. I also would have thought that it would be possible to find information on access to basic social services more cheaply from surveys of these facilities and from population surveys but maybe that's not the case.

I think my assessment of whether doing this report was useful depends largely on whether this information was asked for by the parties in Mumbai and needed by policy makers there. This study could have been very useful if there were lots of policy debate in Mumbai based on false premises about the access of the poor to various services and if this shed light on the actual facts. If that is the case, then it is quite possible that this survey paid for itself many times over. If not, given the costs, I am afraid that the return to undertaking this survey is quite low, also because I don't think it is that big a contribution to generalizable knowledge.

Project: Motorization and Urban Transport (5 papers)
Total Funding: $ 1,268,800

Paper (3 of 5): Welfare Effects of Slum Improvement Programs: The Case of Mumbai (Akie Takeuchi, Maureen Cropper, and Antonio Bento), 2005. Submitted to Review of Economics and Statistics

A. Objectives

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

The goal of this interesting paper is to evaluate the welfare effects of policies that would improve slum housing in situ – for example, by providing piped water and sewage connections – or that would relocate slum dwellers to better quality, low cost housing. For example, if finding a new job in a new location is difficult, in situ improvements may dominate relocation programs.

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

Yes, much of the developing world is urbanizing rapidly and slums are a common feature of urbanization. A better understanding of the cost and benefits of alternative approaches that can improve the living conditions of slum dwellers is critical. However, it is not clear to me how relevant the paper’s particular focus on evaluating in situ improvements against relocation programs is – while I am aware of in situ improvement programs of which the authors provide examples, I was under the impression that large-scale relocation programs are not commonly considered anymore. I would have thought that a policy of providing compensation or housing vouchers to people displaced by slum upgrading programs would clearly dominate one of relocating entire communities.

B. Design and Implementation

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

The authors argue that to evaluate the welfare effects of slum upgrading and resettlement programs, estimation of residential location choice (CV) models is required. In residential choice models, information is needed on the tradeoffs people are willing to make between commuting costs, housing costs and the attributes of the housing that they consume. It also depends on neighborhood composition such as ethnicity and caste (both in terms of preferences and economic value provided by social networks).

Using a data set of 5,000 Mumbai households, the authors estimate residential choice as a discrete choice problem in which each household's choice set consists of the chosen house plus a random subset of 99 houses that the household can "afford" (?) from the 5,000 houses in the sample. In one specification the employment location of the primary household earner is considered fixed (and distances are measured between the 99 house options and this location), and in an alternate specification, distance to the current workplace is replaced by an employment accessibility index.

The authors argue that they extend the existing literature on slum improvements in three ways. First, they introduce employment access as a factor influencing the choice of residential location. Second, they incorporate endogenous neighborhood amenities (the language and religion of one's neighbors). Third, they account for unobserved heterogeneity in housing and neighborhood characteristics.

There could be more detail, however, in the discussion on the (limitations of the) estimation strategy. For example, to what extent are unobserved household preferences for choice of residential location a problem for inference? Is the measure of employment access not endogenous? And there is little discussion on the limitations of the welfare analysis. Suppose, for example, that social networks are not only determined by ethnicity (which is controlled for) but importantly also the extent to which neighbors come from the same natal region. Relocation to a new area with the same ethnic composition can lead to significant loss if it means natal village networks are being disrupted. More detailed knowledge of such networks and of employment patterns more generally is necessary. Similarly, relocation may cause networks of friends to break down, possibly causing welfare losses due to the economic and social value such networks represent.

Further, one of the paper's main conclusions is that there is a welfare loss to individuals from moving into an ethnically/religiously heterogeneous neighborhood. However, I think more attention to the impact on others of this is needed. The welfare (and policy) implications, however, are complex. To take an example, one of the main ways in which the authors find their ethnicity and religion analysis mattering is when they find that moving a certain neighborhood would be welfare reducing because the Marathi speaking Hindu population who lives there would not want to live with Muslim and Hindi speaking neighbors. There is scope for debate about whether and how public policy

should take into account these preferences. It is not clear that levels of compensation paid to displaced people should take these preferences into account.

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 is some discussion on the literature evaluating slum improvements and types of slum improvements that have been undertaken. There could be more institutional detail, for example on the allocation of property rights and on the experiences of different types of slum improvements carried out in Mumbai and employment patterns among the poor in Mumbai. Information on the latter is important to evaluate which, if any, of the two measures of "distance to employment from house" accurately reflects the employment consequences of relocation programs. (The first measure of distance holds employment location fixed even if a household would relocate and the second measure assumes the distance will correspond to the average distance from the new location to the 100 nearest jobs in the worker's occupation, where the latter is based on five categories of work – unskilled, skilled, sales and clerical, small business, and manager/professionals.)

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?

There is little discussion on the survey itself, although the paper above is referenced as a source for more information. As mentioned in the discussion of the above paper, the survey design was carefully developed, although refusal rates are not known, making it difficult to assess the quality of the data.

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

The authors indicate that previous work has not sufficiently incorporated attributes that vary by households, such as distance to work or the percentage of neighbors similar to oneself. Endogeneity/identification issues aside, the empirical results indicate that these are indeed important characteristics. The paper also concludes that the model estimated can be of use in calculating the relative welfare gains from alternative slum improvement programs, and is also useful in predicting which households would be likely to participate in various programs, given the cost of participation. While this is true, providing more detail on the limitations and assumptions of the estimation strategy would facilitate interpretation of the findings and likely increase the willingness of policy makers to adopt this approach in policy design.

C. Accessibility

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

Yes, although more detail would aid, both if the intended audience is academic and if it consists of policy makers.

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

Please see (4) above.

D. Results and Cost-effectiveness

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

Please see (4) above.

2. Were the results consistent with the costs? Did the project take adequate time to complete?

Unknown – not clear what the costs were, nor if the findings are being used by Mumbai or other policy makers considering slum improvement programs.

Project: Motorization and Urban Transport (5 papers)
Total Funding: $ 1,268,800

Paper (4 of 5): The Effects of Urban Spatial Structure on Travel Demand in the United States (Antonio Bento, Maureen Cropper, Ahmed Mushfiq Mobarak, and Katja Vinha); The Review of Economics and Statistics, 2005

A. Objectives

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

The goal of this interesting paper is to measure how urban form – whether measured by the spatial distribution of population, the jobs-housing balance, or the supply of public transit – affects vehicle ownership and the number of miles driven by households in the United States.

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

Although the paper uses U.S. data, one might argue that lessons from the U.S. experience might be used in urban planning elsewhere. However, I believe that the World Bank should focus its efforts on developing country research, unless perhaps when data are not available or very difficult/costly to collect to answer an important generalizable question.

B. Design and Implementation

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

The use of data was innovative but limitations were not sufficiently discussed. The authors carefully carry out the analysis, paying close attention to appropriate measures of urban form. Before the presentation of results, the authors recognize that treating as exogenous to the individual worker measures such as road density and rail and bus supply is more difficult to justify than for measures of urban form such as city shape – although even the latter may be endogenous. However, this limitation is not revisited in the discussion of results. For example, as described below, the authors find that road density increases the likelihood that people drive to work. But without controlling for unobserved heterogeneity, it is not clear whether/how much of this effect is causal; people who enjoy driving will vote in politicians willing to spend more money on roads and less on public transport. And since variables such as road density are endogenous, it would have been helpful to see how the coefficients on arguably more exogenous city characteristics are affected when road density is excluded.

The authors construct various measures of urban form that are novel, arguing that several conventional measures violate the guiding principles used by the authors in deciding how to measure urban form: (1) the set of measures should capture different aspects of urban form (i.e. not be too highly correlated), and (2) it should be possible, conceptually, to vary one measure while holding the others constant.

For example, they construct a new measure of city sprawl in which the actual population at each distance from the CBD is weighted by distance – a person living on the city edge 10 miles from the CBD is weighted ten times as much as a person living 1 mile from the CBD. This may be justified by the fact that the person living 10 miles must travel 10 times as far to reach the city center. This so-called population centrality measure is only weakly correlated with, for example, average population density, a commonly used measure in this literature. Further, using zip code information, they also construct a Lorenz curve measure of job-housing balance (Massey and Doenton's Gini coefficient), with larger values of this measure implying a more uneven (less balanced) distribution of jobs versus housing.

Two sets of models are estimated. The first is a multinomial model of commute mode choice, in which four alternatives – driving, walking/bicycling, commuting by bus, and commuting by rail – are distinguished. This model is estimated using workers from the NPTS who live in one of the 26 cities in the U.S. that have some form of rail transit as well as data on measures of urban form. The indirect utility depends on income, cost per mile of driving, worker characteristics (age, race, education, gender, number of adults and children in household, household income), and measures of urban form and transit availability.

In examining the effects of urban form and transit supply, two results stand out. First, population centrality increases the chances that a worker walks to work. The same is true

for a more balanced job-housing balance. The second is that increasing the supply of rail (bus), increases their modal share, and increasing road density increases the likelihood a worker drives to work. Since these measures are likely endogenous, the causal effects are not clear.

The second is a logit model to explain whether or not a worker drives to work using data from 114 cities for which they have both urban sprawl and transit data. The second model explains the number of vehicles owned and miles driven per vehicle for households living in these same 114 urban areas.

Here, the main findings are that an increase in road density increases annual miles driven, as does a decrease in the jobs-housing balance. The more circular a city, the fewer the miles driven by one-vehicle households. An increase in rail route miles reduces annual miles driven. Similarly, since these measures are likely endogenous, the causal effects are not clear.

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?

Yes, the creation of new measures of urban shape is based on arguments that existing measures do not adequately capture important dimensions of urban shape.

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 authors combine the 1990 Nationwide Personal Transportation Survey (NPTS) data city-wide measures of sprawl and transit availability. These data are presumably reliable.

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

The authors conclude that even in a country like the U.S., which is heavily dependent on the automobile, urban form does affect travel demand. Although it would be difficult to argue that this conclusion does not hold (if in one city you have no option but to travel 10 miles to work while in another city only 5 miles, travel demand will look different), the analysis does not allow separation of the causal effects from other, indirect effects; e.g., people who enjoy driving cars are also people who like to live in sprawled out cities and elect politicians that will build the roads and create the zoning laws to achieve this.

C. Accessibility

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

Yes.

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

Please see (4) above.

D. Results and Cost-effectiveness

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

Please see discussion above.

2. Were the results consistent with the costs? Did the project take adequate time to complete?

It is not clear what share of the $1.3m this particular research cost. However, given its focus on the U.S., I do not think this is the best way for the Bank to allocate its scarce resources, if in fact this project was bank funded.

Project: Motorization and Urban Transport (5 papers)
Total Funding: $ 1,268,800

Paper (5 of 5): Does the value of a statistical life vary with age and health status? Evidence from the US and Canada (Anna Alberini, Maureen Cropper, Alan Krupnick, and Nathalie Simon); Journal of Environmental Economics and Management, 2004

A. Objectives

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

This interesting paper provides an empirical assessment of the effects of age and baseline health on willingness to pay (WTP) for mortality risk reductions by reporting the results of two CV surveys (Hamilton, Ontario, and national sample of U.S. residents). In particular, it estimates individuals' WTP for a reduction in their conditional probability of dying during the current period.

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

The theory and methodology on VSL can certainly be of importance for empirical applications using developing country data. This particular paper, however, attempts to provide an empirical assessment using newly collected U.S. and Canadian data, not necessarily providing a new model or empirical methodology. Hence, the objectives of this paper are not critical for policy in the developing or post-socialist countries.

In fact, I am not sure that this is really the type of research that the Bank should be funding if it was funded. First, I think there are large questions regarding contingent evaluation methods. Second, this research is done in the U.S. and Canada and while results may generalize, I think that in general the Bank should focus on research that is either in developing countries or of a cross-country nature.

There can be exceptions but I don't think this rises to that level. Research of this kind is a global public good, but not one of disproportionate value to the poor/developing world, and hence I don't think it should be a priority for the Bank. I think this research should be done. I am just not sure it's a priority for the Bank

B. Design and Implementation

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

Following a discussion about the respondent's (family) health history and exercises that acquaint the respondent with the concept of risk, survey respondents are being told the baseline risk of death over the next 10 years for someone of their race and gender, and are asked whether they would purchase a product (not covered by their health insurance) that would reduce this risk by either 1 in 1,000 or 5 in 1,000 at a stated price. The survey was designed keeping in mind the difficulties of getting realistic WTP for mortality reductions based on a non-existing product. The authors made sure that participants were being trained in probability assessment and collected information on respondents own risk assessment and whether they had doubts about the effectiveness of the product. Still, it is difficult to judge how realistic WTP estimates are when based on hypothetical scenarios.

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?

Yes, the paper demonstrates awareness of existing knowledge and literature and places the empirical application into a conceptual framework to aid its interpretation.

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?

There is considerable detail in the paper on the survey design, refusal rates, etc. This seems to have been a major data collection effort with considerable planning. Unfortunately, the Canadian data were plagued by high refusal rates; the response rate was only 26%, partly because respondents were asked to show up at a central facility and the CA$ 35 was not a great enough incentive. The authors could have mentioned more about the selectivity of their sample and how this may have affected their results.

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

The empirical results indicate (weakly) that older individuals and individuals with chronic diseases may be more willing to pay for a reduction in mortality risk. The authors evaluate this in light of the conceptual framework, which is based on the notion that individuals maximize their expected lifetime utility subject to the intertemporal budget constraint. According to this framework, older individuals would want to pay more for being able to give up the probability of dying in the current period, since this probability is higher for older people, but should be willing to pay less, since the expected value of their (shorter) remaining lifetime utility is lower. The authors conclude that the former must (weakly) dominate the latter effect. However, the empirical measure of risk reduction is not entirely comparable. Respondents were asked their WTP for absolute reductions in risk (1 in 1,000 or 5 in 1,000 annually). The utility value of this level reduction is indeed higher for younger individuals who have more years to live (reason 2). It is not entirely clear why older individuals would be willing to pay more for the same risk reduction received by young people for reason 1 outlined above.

Presumably one of the arguments against the idea that investing in reducing a particular cause of death for the elderly should be a priority is the competing risk argument, but it sounds to me like the way the questions were asked (focusing on an absolute reduction) abstracts away from that.

C. Accessibility

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

Yes.

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

The authors also conclude that the results support Health Canada's approach of using age-adjusted VSL estimates in its economic assessments, applying a lower VSL for people above 65 years. It is unclear how this follows since older people have a higher, if any, willingness to pay for given risk reductions.

D. Results and Cost-effectiveness

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

Please see above.

2. Were the results consistent with the costs? Did the project take adequate time to complete?

Unknown – it is not clear what share of the $1.3m this particular research cost. Further, it is not very clear how this particular empirical application of WTP relates to the overall "Motorization and Urban Transport" project.

Project: Infrastructure and Spatial Resource Allocation (1 paper)
Total Funding: $ 994,200

Paper (1 of 1): Agglomeration Economies and Productivity in Indian Industry (Somik Lall, Zmarak Shalizi, Uwe Deichmann); Journal of Development Economics, 2004.

A. Objectives

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

This interesting paper examines the following questions: (1) Does the magnitude and source of agglomeration economies vary between industrial sectors? (2) Does improved market accessibility enhance agglomeration economies and lead to increased plant level output? (3) Does a firm benefit from co-locating near other firms in the same industry? And, (4) does a firm benefit from being located in dense urban areas?

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

Yes, understanding the relationship between (spatial dimensions of) industrial composition and economic development helps inform industrial policies, urban planning, labor policies, etc.

B. Design and Implementation

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

As the authors point out, this is one of the first studies to use a combination of plant level and disaggregate physio-geographic data to examine the contribution of agglomeration economies to plant level productivity. Previous studies, according to the authors, have two limitations. First, the use of city and industry rather than plant level data introduces aggregation problems in the analysis, thereby biasing the returns to scale parameters upward. The second limitation is that agglomeration economies are often modeled as being Hick-neutral, a constraint that should be tested empirically and not imposed a priori. The latter is not imposed in this paper. Here the methodology is based on solving the firm's profit function.

The analysis is carried out by use of a combination of data sources (see below) for 11 industry sectors, grouping plants by their two-digit NIC codes. The authors carefully

develop various variables representing agglomeration economies. For example, a market accessibility index is created using GIS information on travel time and urban populations at different distances. Own industry employment in the district is used to measure localization economies. Urban population density, measured as the ratio of urban population to the urban area of the district, captures urbanization economies.

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?

Yes, paper places analysis in existing research and highlights where it departs.

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?

They combined existing data sources in a novel way. First, they use plant level data for 1994-1995 from the Annual Survey of Industries (ASI) conducted by the Central Statistical Office of the Government of India. The data quality has been examined by cross-referencing with standard growth accounting principles as well as by reviewing comments from other researchers who have used the data; the outcome of this exercise is not reported though.

These plant level data are supplemented by district and metropolitan area level demographic and amenities data from the 1991 Census of India and detailed information on the availability and quality of transport infrastructure linking urban areas.

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

The key findings of the study are that agglomeration effects vary considerably among sectors. The research finds that market access significantly increases plant level output in two sectors (machine tools, electronics and computer equipment), and reduces it for the beverage and tobacco industry. Distance to sea ports decreases output in three industries, (1) cotton textiles, (2) electronics and computer equipment, and (3) the beverage and tobacco industry, with the strongest effect for the latter.

Two sectors, (1) printing and publishing and (2) non-metallic mineral products, benefit significantly from localization economies (own industry employment in the district). The research finds only one significant effect (positive) of urban density (the ratio of urban population to the urban area of the district) for the cotton textile sector. These latter findings suggest that the benefits from locating in dense urban areas do not appear to offset the costs. This, according to the authors, is interesting as well as puzzling.

However, if firms can choose the initial location of their plant, one might expect that on the margin firms would be indifferent to locating in dense but higher cost urban areas than somewhere outside these areas.

Based on these findings, it concludes that a possible option for improving efficiency in industry location would be to improve the availability and quality of inter-regional transport infrastructure, linking smaller urban areas to the rest of the network.

Since isolating agglomeration and localization effects is very difficult in the absence of clear instruments, to corroborate the results it would have been important that the authors shed light on whether and why the agglomeration findings are consistent with their knowledge of the particular industries where significant effects are being found. For example, does it make sense that plant level output in the electronics and computer equipment industry would be increased as a result of increased market access but reduced by shorter distances to sea ports?

Without this type of information it is not possible to judge the extent to which endogenous placement of firms or agglomeration/localization effects are driving the results.

C. Accessibility

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

Yes.

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

On the whole, yes (building roads) – provided one is convinced that the causal effects are properly identified. It could have been helpful if the authors had taken one industry as an example and had indicated how their findings can be translated into policy recommendations for that industry. A policy brief based on this paper would open its findings up to a larger audience (perhaps such a brief already exists).

D. Results and Cost-effectiveness

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

Please see discussion above.

2. Were the results consistent with the costs? Did the project take adequate time to complete?

The project cost approximately US$ 1 million. Since the authors compile their GIS data from existing data sources, this seems high, but perhaps different project expenses were included too.

Project: Database on Infrastructure Privatization
Total Funding: $ 387,900

Paper (1 of 1): New Tools for Studying Network Industry Reforms in Developing Countries: The Telecommunications and Electricity Regulation Database (Scott Wallsten, George Clarke, Luke Haggarty, Rosario Kaneshiro, Roger Noll, Mary Shriley, and Lixin Colin Xu); Review of Network Economics, 2004

A. Objectives

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

The goal of this important research is to collect new data sets on telecommunications and electricity regulations. The database of telecommunications regulations includes 178 variables on regulatory governance and content in 45 countries. The electricity regulations database contains 374 variables in 20 countries. This paper motivates the data collection exercise, explains the survey design and provides a few examples of the types of information and results the data can yield.

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

Yes, as the authors point out, infrastructure industries underwent massive structural changes during the 1990s, including privatization, market liberalization, unbundling, and the introduction of new laws and regulations.

B. Design and Implementation

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

The authors point out that empirical research has focused far more heavily on privatization than it has on regulation. Yet, regulations, regulators, regulated industries, and politics interact in complicated ways that affect the development of the industry as well as the rest of the economy. One reason for the lack of research with a regulatory focus in developing countries is the lack of available data.

Collecting this new data set is innovative and leverages one comparative advantage that researchers at the World Bank have over researchers at universities. The World Bank has an enormous advantage in collecting comparable cross-country data given its presence across the world and its relations with client governments. Data sets such as those collected by the World Bank in collaboration with academics, for example on regulations

for business start-ups, have proven very valuable and collecting this kind of information is work that, in my opinion, the World Bank should do much more of.

The surveys were carefully designed to collect information on two important dimensions of the regulatory framework: (1) governance – rules that order the scope, configuration, and process of regulatory decision-making and the organization and procedures of regulatory agencies – and (2) content – the specific laws and policies toward pricing, competition, access, and investment that shape incentives facing firms and consumers.

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?

Yes, the authors put the survey design in the context of existing research on the topic and motivate the particular focus so as to fill a gap in the existing data available.

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 authors contacted 60 telecommunications regulatory agencies around the world and received responses from 45 countries (i.e. a refusal rate of 25%). They also contacted 46 electricity regulators and received responses from 20 (i.e. a refusal rate of 56%).

There is little discussion about difficulties encountered in the data collection process, other than stating the refusal rate. Why did some agencies refuse? Why did others not refuse? What is the average response rate on the questions among those who did not refuse?

It also lacks a discussion on the sample design. That is, what is the sampling universe? There is mention that every developing country that seemed to have a regulatory agency with an email address or fax number was part of the sampling universe. How representative is the sampling universe of the developing world more generally? And how representative is the list of countries that responded of this universe? What kind of selection effects may be at place?

This is critical information to judge the reliability and use of the data, but unfortunately is missing.

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

As mentioned above, the paper lacks detail on the problems encountered in the process of collecting the data. Although there is a little bit, the paper could also have had much more information on some of the basic findings – perhaps this is in a companion paper? This is unfortunate since such information could spur other researchers to use these data too.

C. Accessibility

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

Yes. The report is non-technical and contains links to the actual data set.

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

As mentioned above, the paper is sparse on findings.

D. Results and Cost-effectiveness

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

Not applicable – the paper lacks detail on findings.

2. Were the results consistent with the costs? Did the project take adequate time to complete?

The project funding was $387,900. Collecting cross-country data can be time consuming and expensive, but the returns can be large in spurring new research and generating new information that can lead to more effective development planning. In this case, 60 regulatory telecommunications offices and 46 electricity regulators were contacted by mail and phone for a self-administered survey. Based on the limited information on what may have been included in the cost of this data collection, this seems expensive though.

Project: MT-Biomass & Coal Utility Energy (2 papers)
Total Funding: $ 400,000

Paper (1 of 2): Romania Rehabilitation & Environmental Controls: Chubu Electric Power Inc. (Takashi Masaki), draft report 2005

This is a technical World Bank consultancy report – not a research paper.

A. Objectives

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

According to the authors, the purpose of this study is to review the planned environmental control projects associated with the Deva, Isalnita, Rovinari and Turceni power plants in Romania, and assess the appropriateness of the proposed scope, the reasonableness of estimated investments and the competitiveness of the plants after the rehabilitations/retrofits. This study is carried out in the context of a request from the

Government of Romania (GoR) to the World Bank to provide loans for some of the planned projects and is intended to assist the GoR and the World Bank to select the most appropriate projects to receive World Bank loans.

It is based on visits to the plants (April 3 to 14, 2005) to collect relevant information and meet with plant staff and other organizations. Then the study team (consisting of engineers from Chubu Electric Power Company of Japan and Stratos Tavoulareas, Energy Consultant) analyzed the alternatives and wrote the document. A list documenting the study team is missing.

B. Design and Implementation

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

Given the engineering nature of the document in an unfamiliar area, commenting on the execution of the work is difficult. It is not a research paper that leads to generalizable new knowledge.

2. Were the results consistent with the costs? Did the project take adequate time to complete?

The field visit took place in April, and the report was written in May, suggesting it took adequate time. On cost efficiency, $400,000 for these two reports (this one and the one below) seems high but without a better understanding of the cost structure and the composition of the study team, this is difficult to judge.

Paper (2 of 2): Turkey Assessment of Afsin-Elbistan Rehabilitation: Chubu Electric Power Inc. (Takahashi Masaki), report July 2004

This is a technical World Bank consultancy report – not a research paper.

A. Objectives

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

The Government of Turkey asked the World Bank to assist in planning the rehabilitation of Elbistan. Elbistan is one of the most (if not the most) important power plants in Turkey because of its size and utilization of local lignite, the largest energy resource of the country. However, Elbistan has experienced significant deterioration in its performance and reliability.

According to the author, the objective of this study is to carry out an assessment of the rehabilitation of Elbistan, which will include a preliminary definition of the scope of the project based on the urgent needs of the plant and the cost-effectiveness of rehabilitation options, and recommend additional assessments needed to plan the project.

The authors indicate that more detailed assessment (feasibility study) is required before the scope of the rehabilitation project is finalized. The recommendations of this study are designed to provide specific guidance to the consultant (an engineering company) regarding what should be included in its assessment/feasibility study.

The study team met with EUAS (?) in late February (2004) and visited the Elbistan power plant to collect information and meet with plant staff, and carried out a brief inspection (plant walk-through). A list documenting the study team is missing.

B. Design and Implementation

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

Given the engineering nature of the document in an unfamiliar area, commenting on the execution of the work is difficult. It is not a research paper that leads to generalizable new knowledge.

2. Were the results consistent with the costs? Did the project take adequate time to complete?

The field visit took place in February, and the final report was written in July, suggesting it likely took adequate time. Again, on cost efficiency, $400,000 for these two reports (this one and the one above) seems high but without a better understanding of the cost structure and the composition of the study team, this is difficult to judge.

Questionnaire on Strengths and Weaknesses of World Bank Research

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

Infrastructure is a critical issue in development and one that is very under-researched, both within the Bank and outside the Bank. The research I reviewed included a number of nice pieces, such as the Nepal works on specialization and the spatial division of labor, and on isolation, welfare, and rivalry. The collection of cross-country data on telecommunications and electricity regulations was promising. And although this was not part of the review papers, the recent work by Galiani, Gertler, and Schargrodsky on privatization of water infrastructure in Argentina is very good and important.[1] But, based on the research I was sent, I cannot say as a whole that the Bank has really made the contribution on infrastructure it should have. I sometimes had the sense that the people putting together the set of papers had to stretch to come up with enough papers to fit in the category. Thus, for example, two of the documents sent to me were not really research papers at all but rather documents related to proposed Bank projects, for example, about the rehabilitation work that would be necessary for a particular power plant to come into line with EU regulations. This may have been a very useful document but I don't think it constitutes research. (Maybe this reflects random selection from a list and things accidentally were put on a list that did not belong there.)

Another document dealt with whether it is appropriate to use procedures that assign lower weight to the statistical life of the sick and elderly in Canada and the United States. This is an interesting question of cost-benefit analysis and one that has applications to developing countries and also has applications to infrastructure. However, if I had to categorize this as most useful to a single field it would be health economics, not infrastructure.

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

Of the papers that were on research about infrastructure issues in developing countries, most were related to policy only indirectly. Thus, for example, the series of papers on Nepal, which I thought were very interesting, dealt with the impact of isolation on division of labor, subjective measures of well being, etc. That is good work, and while I think it most naturally fits in to an academic research program, I also think it is appropriate for the Bank to support some of this type of work as part of its portfolio. However, in reviewing the various projects, I have the strong feeling that the portfolio of research is seriously out of balance, and that policy is not getting enough attention and the potential comparative advantage of the Bank is seriously underplayed.

[1] Galiani, Sebastian, Paul Gertler, and Ernesto Schargrodsky (2005). "Water for Life: The Impact of the Privatization of Water Services on Child Mortality." *Journal of Political Economy* 113(1): 83-120.

I don't know enough about the incentive structure in the World Bank's Research Department to make strong recommendations regarding how incentives should be changed. However, I would like to make a few observations and discuss some of the relevant tradeoffs, as I see them. I think the Bank should strengthen the link between research, operations, and policy, and should focus more in those areas where it has a comparative advantage – in particular, carrying out impact evaluations of projects that have important policy implications and can bring new knowledge to the table, and working with governments to collect detailed comparable cross-country data using consistent definitions across countries.

One of the research projects reviewed constituted the collection of cross-country data on telecommunications and electricity regulations. I think this effort could have gone much further than it did, but I think the basic underlying idea is very good. With more support of country field offices, work of this type can collect detailed data from across the developing world on issues that are critical for policy. Provided a detailed and wide dissemination effort is in place, it can spur a great deal of research, act as a resource database for client governments, and be a base for more effective policies.

Similarly, rigorous and systemic evaluations also have the potential to leverage the impact of the Bank well beyond simply its ability to finance programs. Credible impact evaluations are international public goods: the benefits of knowing that a program does or does not work extend well beyond the organization or the country implementing the program. Programs that have been shown to be successful can be adapted for use in other countries and scaled up within countries, while unsuccessful programs can be abandoned. Through promoting, encouraging, and financing rigorous evaluations (such as credible randomized evaluations) of the programs it supports, as well as of programs supported by others, the Bank can provide guidance to governments, NGOs, and donors in the ongoing search for successful programs. Moreover, by credibly establishing which programs work, the Bank can address widespread skepticism about the effectiveness of aid spending and build long-term support for development. The Bank should be in a much better position to conduct careful impact evaluations than most researchers, because it is involved with various infrastructure projects and because of its ties with governments. It should be able to organize for data collection both before and after projects to allow for regression discontinuity approaches to estimating their impact. In many cases it may actually be able to arrange some level of randomization, for example randomization of the order in which programs are phased in to various geographic areas, so as to estimate program impact. Based on the material sent, it does not appear to have done that as much as it should. I would suggest that the Bank identify a set of questions of substantial policy importance; for these questions, it would be very useful to implement randomized evaluations in a number of countries.

Building links between operations and research in particular by conducting rigorous impact evaluations and collecting comparable cross-country data sets takes a lot of work on details. People may not have sufficient incentive to do them and instead undertake projects that can be done more easily. Moreover, the type of projects where the Bank has a comparative advantage require a lot of cooperation with operations and people in

operations may not all have incentives to put a lot of effort into this cooperation. Below are my suggestions for improving incentives.

One way to create these incentives would be to create dedicated trust funds for this purpose. For example, I think there should be a trust fund dedicated to randomized evaluations. Currently, a very small percentage of the Bank's projects are subject to randomized evaluations. These evaluations create a global public good that goes well beyond the particular project and the particular country, as the example of the PROGRESA evaluation makes clear. There is not sufficient incentive to conduct them for a task manager. Having a trust fund for which people could bid and which is limited to randomized evaluations would be appropriate and I believe spur a lot of interest from researchers to go out and forge links with operations to do this work. Another possibility would be to have a broader trust fund for impact evaluations of which a proportion would be reserved for randomized evaluations. Undertaking the detailed work necessary will require hiring people outside the Bank for data collection, and trust funds are useful for this.

Second, my understanding is that there is a perception of a significant gap between operations and research. Establishing a system in which the standard career path is for people to spend a short period in research and then move out to operations, perhaps to come back to research later, often in management, might help reduce this cultural gap and allow the Bank to capitalize on the comparative advantage discussed above.

There is an extraordinary interest among Ph.D. economics students in development economics, making it a buyers' market for new development Ph.D.s. Moreover, there is a lot of interest in policy issues among this group. Unfortunately, budget constraints have prevented the Research Department from hiring recently, or at least that is my understanding. One of the roles of a research department within the World Bank should be to assist in the recruiting of talented young people into the Bank so that they can contribute to the organization as a whole. I think that having a somewhat more dynamic system in which people moved from research into policy a bit more would help disseminate ideas from the Research Department and from academics into broader operations and vice versa and would also allow the Bank to bring new Ph.D.s into research even if budget constraints limit overall growth of the research group. I think the Bank might want to consider a set-up in which the norm would be that most people in the Research Department move into operations after 6 or 7 years, for example. Of course, certain senior people in the Research Department are contributing a fantastic amount to research and should certainly stay. This could be accomplished by having a certain number of senior research slots for which people compete on the basis of their publications in peer reviewed journals that were relevant to development. These people would have senior positions without being burdened with too much management/administration. My impression is that the Bank already does this to some extent. I just don't know how much.

I also think it would be helpful to bring visitors from the academic world to the Bank more often and perhaps also to create procedures to facilitate Bank staff visiting

academic institutions. I visited the Bank myself and my collaborations with Elizabeth King on Colombia and Benjamin Loevinsohn and Beth King on Cambodia came out of chance conversations while I was there. Other work where I have greatly benefited from collaboration with World Bank research and operations staff and from World Bank financing include randomized evaluations of Kenyan education and school health projects (e.g. deworming and HIV AIDS education), and cross-country work on teacher and health worker absenteeism. I might have gotten involved in these projects even if I had not visited, but I think visiting made a big difference.

The Bank could also consider having an annual conference not on a particular topic, but more open, with papers divided in some proportion with papers from the Bank and outside papers in which there would be a competitive selection of submissions both internally and externally, like the IMF research conference but without the special focus. The Annual Bank Conference on Development Economics is not playing this role, since there is not an open call for papers and since the presentations are not seminar style.

Another issue regarding the Research Department is the extent to which incentives should be tied to, for example, outside job offers. The clear disadvantage of doing this is that it creates incentives for people in the Research Department to focus on doing things that bring recognition externally but may or may not be useful to the Bank. On the other hand, the clear advantage of this is that, at least within the economics profession, there is a fair degree of consensus on quality of research output, so the information from outside offers is likely to be pretty good. My sense is that the Bank should probably use both external signals of academic research quality and information on performance relative to the Bank's research objectives to the extent these differ. I think people should be required to perform well on both metrics, just as academics are expected to perform well on both scholarship and teaching. I am not sure what the Bank is currently doing.

Based on the talent of people in the Research Department, I think the department should be generating more articles in general interest economics journals like the AER. The fact that these are rare suggests a problem with management and incentives in the bank.

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

Bank research does reflect awareness and sufficient knowledge of other research and the subject countries. This is an area where I have been impressed in my interactions with the Bank's staff.

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

In infrastructure, I think the Bank could collect more cross-country data sets and impact evaluations. In [5] below, I discuss some topics the Bank could examine.

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

The accompanying individual reviews give my opinions on the particular papers that I was asked to read, and the discussion above highlights recommendations to improve incentives. I thought it could also be useful to give examples of recent work that I consider very important in infrastructure. This is the type of work that I think the Bank could use as a model.

I would like to focus on two papers in particular: the paper by Paul Gertler, Sebastian Galiani, and Ernesto Schargrodsky on privatization of water in Argentina (2005), and Esther Duflo and Rohini Pande's work on dams in India (2005).[2]

The water privatization paper uses variation in ownership of water provision across time and space generated by the privatization process, and finds that child mortality fell 8 percent in the areas that privatized their water services and that the effect was largest (26 percent) in the poorest areas. To check the robustness of these estimates, they use information on cause-specific mortality; while privatization is associated with significant reductions in deaths from infectious and parasitic diseases, it is uncorrelated with deaths from causes unrelated to water conditions. I highlight this paper not only because it constitutes good careful research about a policy relevant issue in infrastructure in developing countries, but also because it provides strong evidence against a widely perceived view, in this case, that water privatization is bad for social outcomes.

The debate over dams is typically cast as one that pits economic development against the environment and against the rights of indigenous people displaced by dams. As with privatization, the debate is often in highly ideological terms. In "Dams," Esther Duflo and Rohini Pande develop a convincing empirical identification strategy to measure the impact of dams in India that takes account of the endogenous placement of dams. They find that districts located downstream from dams benefit from increased irrigation and hence increased agricultural production and reduced poverty. However, the districts where dams are built see no such increase in agricultural production but an increase in poverty. This makes logical sense, and the empirical work is very careful.

For me, the key finding of the paper is that the supposed economic benefits of dams are marginal at best. Estimated returns are very low. While the debate over dams is usually pitted as one of economics versus the environment and the rights of indigenous communities that will be displaced by dams, the paper suggests that the case for dams on purely economic grounds is not at all clear. The point estimate of the rate of return is 1 percent, even excluding the deadweight loss of taxation or the cost of increased labor usage on irrigated land.

The empirical identification relies on the fact that dams can be built where there are steep, but not too steep, river gradients. The regressions control for district fixed effects, state/year interactions, and the interaction of most district geography variables with

[2] Duflo, Esther, and Rohini Pande (2005). "Dams." Center Discussion Paper No 923, Yale University.

overall dam construction in the state. It is based on the interaction between the fraction of rivers in different gradient categories and overall dam construction in the state. Both the Gertler and Duflo and Pande papers directly evaluate particular policies of immediate interest to policymakers. Both have convincing identification strategies, are well executed, and find results that challenge the conventional wisdom. This makes me suspect that there are very high returns to this type of work. Bringing evidence to these issues could potentially lead to much better policy making.

Let me give some examples of the type of work that I would love to see. It would be great if, for example, when the Bank helps finance rural road construction, it could also finance baseline and follow-up surveys. Then it might be possible to estimate the impact of the rural road construction on incomes, business creation, health, migration, etc. One could look at the impact of fixing potholes on traffic accidents and on average vehicle speeds (maybe the road engineers already know all these things but I'm not confident of that). The Millenium Challenge Corporation is in the process of setting up a randomized evaluation of roads – so this type of approach is possible.

It would be good to see what predicts the amount of vehicle traffic roads will actually get. It would be good to have better information on the advantages and disadvantages of private toll road schemes. It would be great to have research on alternative approaches to reducing traffic congestion in big cities in developing countries. What is the impact of building ring roads or flyovers? Can road pricing be implemented effectively in developing countries? Are there ways in which bus service can be improved? What are the consequences of different systems of regulation of private bus and minivan service?

Controlling corruption in infrastructure programs is another important issue. The work by Olken (2005), done with World Bank support, can serve as a model here.[3]

Another area where more research could be done is in the electricity and power sector. What is the impact of extending electrical coverage to more areas on income, employment and education, health, etc.? This is something that could presumably be studied without too much problem using expansion of coverage over time. What are the advantages or disadvantages of various options for financing power generation? What's the optimal regulatory framework for electricity? How can losses due to unauthorized tapping into power systems be reduced? Should they be? What are international benchmarks for employment and revenues and pricing in the power sector?

Another area relates to telecommunications and information technology. What is the impact of different regulatory approaches to telecommunications on pricing and access? What is the impact of extending coverage with land lines or mobile phones? What are the benefits and costs of deregulating international telecommunications?

Shipping, rail, and airports are another area where I think more could be done. It would be useful to have research on the effects of rail and port tariffs and delays on exports and

[3] Olken, Benjamin (2005), "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." Harvard University and NBER.

imports. What are ways of bringing down air travel costs, which are tremendously high in many developing countries? Should countries be much more liberal about allowing landing rights unilaterally to other countries?

It should also be feasible to conduct randomized evaluations of many slum upgrading projects since such projects could be randomly phased into different communities over time. It would also be worthwhile looking at a number of approaches to implementation of these interventions. For example, the Bank could examine slum upgrading programs with different levels of community involvement or that put emphasis on different program features ranging from provision of sanitation or waste disposal to land titling programs. In Field and Kremer (2005),[4] we outline steps we think the Bank could take on evaluation of slum upgrading programs.

Finally, there is very little evidence from randomized trials on the impact of programs to improve water quality and quantity or sanitation, or the effectiveness of different approaches to maintaining water infrastructure. Randomized evaluations are similarly feasible since here too these can often be phased in over time. When this is not feasible, it should still be possible to use regression discontinuity design to evaluate them. In Kremer and Zwane (2006),[5] we review the existing research on the cost-effective prevention and treatment of diarrheal diseases, and identify research priorities in this area, including research on infrastructural investments, aimed at finding ways to reduce the diarrheal disease burden.

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

I don't have adequate information to comment on this.

7. What is your overall assessment of Bank research?

I have worked with people in the Bank's Research Department and have had very positive experiences. I am genuinely very impressed with the quality of researchers at the Bank, their knowledge of the realities on the ground in developing countries and of the academic literature and debate. But the quality and policy relevance of the work I have reviewed here does not meet the quality of the staff, which makes me wonder if Bank research could not be structured better.

While I have suggested that in general the Research Department could benefit from focusing more on policy, particularly the collection of cross-country data sets and impact evaluations, and that the link to policy could be improved, I do think that the Research Department – and in particular certain parts of it – has played a useful role.

---

[4] Field, Erica, and Michael Kremer (2005). "Impact Evaluation for Slum Upgrading Interventions." Harvard University, Mimeo.

[5] Kremer, Michael, and Alix Peterson Zwane (2006), "Cost-effective prevention of diarrheal diseases: A critical review." Report submitted to the Center for Global Development Global Health Policy Research Network

For example, the Bank deserves credit for moving earlier than most bilateral aid agencies and other multilateral organizations in supporting randomized evaluations. Nonetheless, I think the Bank could go much further in doing randomized evaluations, and that the Research Department could play a bigger role in this. As discussed in more detail in Duflo and Kremer (2005),[6] I think this is one of the most valuable things the Bank's Research Department can do. For a subset of policy relevant questions that have been identified as priorities by the Bank, undertaking randomized evaluations, which typically have results that are very transparent and are useful for policy makers, would be desirable. I think there should probably be some process to meet with people in operations to decide what questions are most important in each area that could potentially be subject to randomized evaluations and to find a way to begin gathering evidence on these questions. In some cases this will be a matter of testing the impact of a widely advocated popular approach where little hard evidence exists. In other cases it may be testing out various alternative implementation strategies.

[6] Duflo, Esther, and Michael Kremer (2005). "Use of Randomization in the Evaluation of Development Effectiveness," in George Pitman, Osvaldo Feinstein, and Gregory Ingram (editors), *Evaluating Development Effectiveness*, New Brunswick, NJ: Transaction Publishers.