Design of Impact Evaluation for Rural Roads Projects

A) Background

1. The Viet Nam Rural Transport I project (RTPI)

The Viet Nam Rural Transport Project I (RTPI) is a large scale rural roads rehabilitation project aiming to link communities up with markets and to reduce poverty. It was launched in 1997 for implementation in 18 poor provinces over 3 to 5 years, at a total cost of about $61 million. In each participating province, road links are identified for rehabilitation through least cost techniques. No ‘new’ roads are built although the prior existence of a road does not in any way imply that it was usable. In many cases bridges are missing and whole road sections are impassable by a motor vehicle much, if not all, of the year. A proposed road is eligible for the project subject to average investment costs being no more than $15,000 per km and the population served being at least 300 people per km. Bridges are also eligible for rehabilitation, based on the priority assigned to the road and construction costs being less than $50,000. In an effort to extend project benefits to low density, mountainous areas with concentrations of ethnic minority populations, twenty percent of each province’s rehabilitation funds can be set aside for roads not justified under the population and cost criteria. In practice, few roads have been chosen under these ‘social’ criteria.

2. Setting

Viet Nam has poor physical infrastructure and high levels of income poverty. Many argue that basic infrastructure investments will reduce poverty in Viet Nam. The country began its transition to a market economy in 1987 after a long period as a planned market economy. It has since been experiencing remarkable changes in all aspects of economic life. The last decade has seen considerable—though geographically unequal—growth. There have been rising opportunities and rising mobility. Labor and land markets are just now developing, as are the legal and judicial systems governing private property.

3. Evaluation objective

The aim is to evaluate the impacts of rural roads on both the mean and the distribution of living standards broadly defined. Specific questions of interest include:

- What are the impacts on living standards broadly defined to include health, education and security outcomes?
• What factors influence those outcomes?
• How much do benefits depend on other investments such as reflected in human capital endowments?
• In what ways do first-round impacts differ from longer term impacts?

A key issue for evaluation is to succeed in isolating impacts due to the road as opposed to the myriad other changes and shocks hitting the economy and simultaneously affecting living standards. This is the crux of the problem in all impact evaluations. But the magnitude of the difficulties can be expected to be larger in present day Viet Nam.

B) Thinking about the evaluation problem

1. What is different about roads?

It is useful to begin by reflecting on the nature of roads and what it implies for the evaluation problem.

• Roads are an intermediate good and their benefits are indirect and dependent on interactions with other investments, other social and physical infrastructure, geographical, community and household characteristics.

• Roads and road networks are widely believed to have economy-wide effects.

One must therefore control for the heterogeneity of factors that interact with roads to produce impacts. However, if there are indeed economy-wide effects from roads, then a lot of what one sees as potential controls for heterogeneity, or variables to match on, may well have themselves been determined by the road investment. Thus, traditional evaluation methods such as instrumental variables or propensity score matching ex-post, can not be relied upon since they may be contaminated by the effects of the project. Even when lots of data is available, much of what we see in the data is plausibly determined by the road. This leads to pessimism about conducting a worthwhile road impact evaluation without baseline data.

• Roads are not randomly placed, and it is highly likely that the factors that led to the road placement will also affect outcomes.

Roads may be targeted geographically to a community because it is poor, or on the contrary because the area is deemed to have economic potential or because it has a strong political constituency. Unless we can control for those reasons, impact measures will be biased. An endogeneity problem may also arise when looking at impacts at the individual level (e.g., the users), because there may be individual characteristics that are unobserved but geographically determined and correlated with the things that influence program placement.
The potential for randomization to evaluate roads and deal with these issues seems low, if for no other reason than that the full impacts of a road intervention may take a long time to play themselves out. A second phase of road rehabilitation might thus have to be delayed beyond what would ever be acceptable.

- Road benefits may be positive on average, but there may also be gainers and losers.

A specific concern is with the distributional impacts of rural road investments. It is key to understand who the losers are if one is to understand distributional impacts, and the heterogeneity of impacts at given levels of living. For example, if new roads lead to higher land values there may be a tendency towards land concentration and landlessness. Those with greater initial land, education, wealth or influence will be better able to take advantage of the changes. The distribution of current income and future income earning opportunities may widen. Here, one needs to make a distinction between short and longer term impacts. In the longer term even initial losers may win.

It is therefore also important to collect data that allows one to distinguish impacts across groups and to follow the experience of those groups long enough after the road is built so that the full effects can be understood.

In sum, we need a panel with a baseline, data on comparison non-project areas and sufficiently long project follow-up, appropriate controls for exogenous time-varying factors, and the ability to differentiate between welfare groups.

2. Learning from the ex-ante evaluation

Much can be learnt from the project itself. The ‘appraisal’ stage used to select the road links to be rehabilitated can be thought of as ‘ex-ante evaluation.’ Thinking about the evaluation from the project’s inception can help to understand biases in the ex-post evaluation and to define an appropriate counterfactual. The key questions that will help to better understand road placement and eventual impacts are:

- How were intervention areas picked? Why are certain provinces included, and within them certain districts and finally certain communes? Were poor areas targeted?

  There is little evidence of targeting to poor areas in RTPI. Choice of provinces is not well understood.

- How were road links chosen?

  Population density, the cost of the proposed investment and producer surplus.
C) Data and methods

The evaluation design was influenced by the above considerations, as well as those of costs and in-country survey capacity. Assistance from the General Statistical Office (GSO) was not an option. The study had originally hoped to piggyback on the second Viet Nam Living Standards Measurements Survey (VNLSS), contracting GSO to over-sample project areas in their normal implementation of the survey. The rest of the data would have provided comparison groups. For a number of practical reasons, including delay of the VNLSS by over a year to well after the project went into implementation, this option was ruled out. Without assistance from the GSO, the possibility of conducting a household level survey of adequate sample size and convincing coverage of household welfare measurement was deemed unrealizable. Outside GSO, there is little capacity and few institutions and no consultancy groups with the staff, equipment and competencies required to adequately carry out such a task. The risks and uncertainties associated with project timing also led to consideration of options under which the timing of the baseline data collection would be less critical—by relying in part on existing commune level data. Furthermore, data collection was designed to enable integration with a future VNLSS, and allow the possibility of drawing out implications for household level welfare indicators.

1. Survey of Rural Road Impacts in Viet Nam (SIRRV)

Road links all pass through communes, and a majority link up commune centers—where facilities and services are located—with the road network. Data is often, and more easily, collected at the commune level. For these reasons, the zone of a road’s influence was defined as the commune through which a road passes.

The "Survey of Impacts of Rural Roads in Viet Nam" (SIRRV) is a panel data set of pre-project baseline and post-project data for both project (“treatment”) and non-project (“comparison group”) areas that was especially created for the evaluation. A baseline was collected in the spring of 1997 in 200 randomly sampled communes in 6 participating provinces. A second round followed in the spring of 1999, a third in 2001, and ideally a fourth will follow two years hence.

The survey instrument consists of a detailed commune level survey, a household survey administered to 15 randomly sampled households in each sampled commune, a short district level survey and a project data base. To determine the household sample, a system of stratified sampling was used whereby 5 households were chosen from each of three lists, containing the poorest, middle and richest thirds of all households in the commune. The lists were based on a welfare ranking done by the commune authorities. This should assure a sample that is reasonably representative of each commune’s main socio-economic groups.

The commune questionnaire includes sections on general commune characteristics, infrastructure, employment, sources of livelihood, agriculture, land and other assets, living conditions, education, health care, development programs, community activities
and organizations, commune finance and prices. Households were asked about general characteristics, employment, assets and amenities, production and employment activities, participation in and access to education, health, markets, credit, community activities, social security and poverty programs, and transport. The main objective of the household survey is to capture information on household level access to various facilities and services and how this alters over time.

Heterogeneity in the condition of the roads pre-rehabilitation, as well as in their post-rehabilitation situation can be expected to matter to impacts. The project database (with information on the pre- and post-works road condition) can be used to take this into account.

2. Methods

The idea is to compare different methodological approaches. One approach is in the tradition of double differencing with matching methods. Propensity score matching (PSM) methods are used to select ideal comparison communes from among the one hundred sampled non-project communes. PSM basically tries to find a non-project commune that is similar in all observable characteristics to the project commune except that it does not have a project. This is the counterfactual. It does this by comparing the propensity scores – the probability of a commune getting a project – of project and non-project communes and picking the closest match for each commune.

The impact of the road infrastructure is then identified by the difference between outcomes in the project areas after the program and before it, minus the corresponding outcome difference in the matched comparison areas. This “matched double difference” estimate gives an unbiased estimate of project impacts in the presence of unobserved time invariant factors influencing both the selection of project areas and outcomes.

When turning to impacts at household level, only households in matched communes are compared. The VNLSS is used to estimate a consumption model on variables available in both data sets. The coefficients are then used to predict consumption for SIRRV households and rank them into national consumption expenditure groups. This allows a differentiation of impacts by welfare level.

A second approach uses a random effects model that is modified to deal with the potential endogeneity of project placement.

D) Implementation and Risks

1. Implementation

An independent consultant with an economics and research background in rural poverty and development was hired to be the in-country supervisor of the study. He and the Bank task manager worked closely to determine the evaluation design, the sampling and
choice of provinces and communes, devise the questionnaires, field test them, and plan implementation. He has hired and trained the team supervisors, organized all logistics, and closely supervised all data collection.

The total cost of three rounds of data collection has been $280,000.

2. Risks

There are many risks to conducting an evaluation of this kind. Some of the problems the study has run into include the following:

- One has little control over project timing. Far fewer projects had been completed by the time the second round was fielded than planned. This extends the number of rounds needed to have an adequate sample size.

- The Vietnamese government has a habit of periodically splitting administrative units such as provinces, districts and, most troublesome for the study, communes. This essentially means that the data on those communes is no longer representative and so the communes must be dropped from the study. This also affects the sample size.

- The weather can also present problems. The second round of data collection had to contend with very severe flooding in one province and a typhoon in another. This delayed data collection considerably in that year.

- Finally, it is not easy to maintain outside interest and enthusiasm for a long lasting study of this kind, where much work must be done before there is anything to show for it. Patience and considerable staying power is required.

References
