Impact Evaluation Design for MCC Connectivity Interventions in El Salvador

Elaborated by
Maximo Torero[1]

Submitted to the
Millennium Challenge Corporation
MCC-07-0124-CON-91

by Social Impact, Inc.
October 24, 2009

[1] Maximo Torero is Division Director of the Markets, Trade and Institutions Division at the International Food Policy Research Institute (IFPRI), email: m.torero@cgiar.org

Table of Contents

1. Background 1
2. The Connectivity Project 2
3. The impacts of the Connectivity Project 3
   3.1 What we know about the impacts of infrastructure 3
   3.2 What outcomes we expect as a result of the Connectivity Project 4
4. Impact Evaluation Approach 6
   4.1 General Approach to be applied in the Connectivity Project 6
   4.2 Specific Approach for the Connectivity Project: Northern Transnational Highway (NTH) 8
      4.2.1Regression discontinuity approach 9
      4.2.2Continuous treatment approach 10
   4.3 Calculating the cost of travel 12
5. The sample design 12
   5.1 Regression discontinuity approach 13
   5.2 Continuous treatment approach 13
6. Data collection 15
7. Timeline 16
8. References 17

Appendix 1 Sample Size and Power Calculation 20
Appendix 2 Calculation of the Number of Clusters per Condition for each of the Four Scenarios for the Variance 31
Appendix 3 Distribution Calculations 33
Appendix 4 Summary Tables for Other Outcome Variables: Number of Clusters per Condition 34
Appendix 5 Summary Tables for Other Outcome Variables: Number of Clusters per Condition 35
Appendix 6 Summary Tables for Other Outcome Variables: Number of Clusters per Condition 36
Appendix 7 Summary Tables for Other Outcome Variables: Total Sample Size 37
Appendix 8 Summary Tables for Other Outcome Variables: Total Sample Size 38
Appendix 9 Summary Tables for Other Outcome Variables: Total Sample Size 39

1. Background

In November of 2006, the Millennium Challenge Corporation (MCC) signed a five-year, $461 million Compact with the Government of El Salvador (GOES) to improve the lives of Salvadorans through strategic investments in education, public services, agricultural production, rural business development, and transportation infrastructure. The Government of El Salvador has set up a management unit called FOMLIENIO to implement the 5 year Compact from September 2007 to September 2012. Social Impact has been contracted by MCC to conduct an impact evaluation of the Connectivity Project of the Compact.

The Connectivity Project initially consisted of a network of connecting roads (NCR) and the Northern Transnational Highway (NTH) but because of the significant increase in construction costs and the existence of other interventions in connecting roads the project currently only focuses on the Northern Transnational Highway. Table 1 describes the evolution of the project components as well as the current investment. As a result of these changes, the current impact evaluation design focuses only on the evaluation of the Northern Transnational Highway.

Table 1. Project Components

Components	Expected number of beneficiaries	Initial Expected Effects	Initial Estimated Investment	Current Estimated Investment
Network of Connecting Roads (NCR)		Reduced transportation costs Time savings	$ 94 million (NCR)	Will not be build
	600,000 individuals (i.e. population 2km from either side of the road). This accounts for 70% of the total population of the Northern Zone	Increased prices for land along the roads		
Northern Transnational Highway (NTH)		Increased income by at least 5-6% for those households within 2 km of the NTH	$140 million	$233.6 million

The impact assessment will combine two parallel approaches. The first approach tries to take advantage of the sequence in which the different segments of the NTH is being constructed by combining a regressions discontinuity and pipeline design. The discontinuity is created by the boundaries of the different segments and the pipeline design by the different dates of construction. The second approach exploits the variation of intensity of treatment given the NTH will give over time different degrees of accessibility to the households along the NTH. As a result, a continuous treatment approach was implemented. In each of the comparisons between the control and treatment groups across both designs we will use propensity score matching (PSM), difference in differences (*as it allows controlling for the change in non-observable variables)*, and instrumental variables (*to control for the remaining potential sources of selection*

bias) to measure the change in household incomes within the "area of influence" defined as the area within 30 minutes of access to the NTH through existing means. Additional outcomes that will be evaluated include the reduction of transportation costs and transportation time, land values, access to public services and their impacts on health and education outcomes, changes in labor allocation between farm and nonfarm activities, and differentiated gender effects of road improvements.

This design report is divided in six sections in addition to this introduction. In section 2 a brief description of the Connectivity Project is presented, in section 3 the major outcomes to be evaluated are presented, in section 4 the impact evaluation approach is detailed, in section 5 the sample designs required to implement the proposed impact evaluation designs are explained, in section 6 the data collection is described, and finally in section 7 the timeline of the impact evaluation design is explained.

2. The Connectivity Project

The El Salvador Compact Program focuses on the Northern Zone of El Salvador, a region that includes one-half of El Salvador's poorest municipalities and that suffered most from the internal conflict during the 1980s. The Northern Zone has worse economic and social indicators than the national average. Fifty-three percent of the households (450,000 people) in the Northern Zone are poor and 28 percent of households (240,000 people) experience extreme poverty. Human capital development is also lower: the average level of schooling in El Salvador is 5.4 years, but the Northern Zone average is only 3.7 years. Additionally, the illiteracy rate of the Northern Zone is 1.6 times the national average (20.5 versus 12.9, respectively).

The goal of the Program is to reduce rural poverty through an increase in economic growth of the region. This goal will be achieved through a five-year program of strategic investments and technical assistance in various sectors. One of the major interventions is the construction of the Northern Transnational Highway. The NTH will serve as a transport artery within the Northern Zone and will potentially improve international connectivity with Honduras in the east and with Guatemala in the west. The activity will open, improve or rehabilitate 280.7 km of the NTH. The NTH will provide contiguous and reliable access to communities in the Northern Zone, as well as to main transport corridors, thereby enabling the Northern Zone to participate more fully in the national and regional economy.

The construction of the NTH had been split in the following sections detailed in Table 2.

Table 2
Timeline

Section	Scheduled Start Date	Scheduled End Date
T1	Will not be constructed	
T2	May 2009	October 2010
T3	December 2009	October 2011
T4	January 2010	January 2012
T5	October 2009	August 2011
T6	October 2009	August 2011
T7	Will not be constructed	
T8	Constructed	

3. The impacts of the Connectivity Project

3.1 What we know about the impacts of infrastructure

The importance of transportation infrastructure for market expansion and the division of labor was recognized as early as 1776 (Smith 1776). Similarly, early work on economic growth and development included physical infrastructure in the category of social overhead capital and highlighted the need for such capital as a basis for development (Hirschman 1959). Duly recognizing the role of infrastructure in economic development, Smith asserted that social overhead capital "is usually defined as those services without which primary, secondary and tertiary production activities cannot function" (Hirschman 1959).

Though the empirical literature on the impact of infrastructure only emerged in the mid-1980s, it has grown to be a large and diverse body of work. Perspectives range from the methodological, such as production function (Holtz-Eakin 1994; Garcia-Mila, McGuire, and Porter 1996) versus cost function (Nadiri and Mamuneas 1994; Morrison and Schwartz 1996); to levels of aggregation, such as firm-, household-, or village-level microstudies (Antle 1984; Ahmed and Hossain 1990; Dong 2000) versus state- (Aschauer 1989a) and country-level macrostudies (Aschauer 1989b); to sector specificities, such as manufacturing (Morrison and Schwartz 1996) versus agriculture (Binswanger et al. 1993; Paul et al. 2001); to regional studies, such as developed (Röller and Waverman 2001) versus developing countries (Fan, Hazell, and Thorat 1999). There are also empirical studies that examine the impact of a specific infrastructure, for example, roads (Jacoby 2000; Gibson and Rozelle 2003) and telephony (Norton 1992).

Antle's 1983 study pioneered the inclusion of infrastructure in explaining differences in agricultural productivity across countries. Using aggregate agricultural production data for 1965 from 47 developing and 19 developed countries, Antle (1983) utilized a Cobb–Douglas production function, including infrastructure as a production input. Infrastructure was defined as a gross domestic product of the country's transportation and communications industries, measured per square kilometer of land. Not surprisingly, infrastructure appeared to have strong positive impacts on agricultural productivity both in developed and in developing countries. In a similar study on Indian farmers, Antle (1984) found infrastructure capital to have a systematic effect on farm productivity.[2] A number of econometric problems are associated with this study, however. First, while it is tempting to infer a causal relationship from public capital to output, it is equally likely that the direction of causality goes from output to public capital. Second, common trends in infrastructure and output—that is, the estimated coefficient—may reflect a spurious correlation between output and public capital stock that is driven by a common time-trend and not by any underlying relationship between the two variables. Third, omitted variables could create a bias on the coefficient estimates.

The majority of studies recognize that road investment has a strong impact on rural incomes and especially on smallholders. However, this literature has not been completely successful in assessing the benefits and costs of alternative investments in roads or the causality of relations that generate higher rural incomes due to a better endowment of infrastructure services.

[2]The study included transportation costs proxied by the geometric mean distance to nearest bus, rail, and postal facilities (in kilometers) in a Cobb–Douglas production function.

The problem with the lack of causal relationship knowledge between the investment in infrastructure services and the increase of income generating opportunities and welfare benefits of rural populations is that the possibility of developing specific policy recommendations is very limited. This problem normally results in policy recommendations that are directed towards a general increase in public infrastructure investment but lacks opinions about appropriate intervention strategies for each specific context.

The Connectivity Project in El Salvador offers an opportunity to try to identify these causal relationships since it incorporates an impact evaluation strategy in its design.

3.2 What outcomes we expect as a result of the Connectivity Project

Road infrastructure, such as that being provided by the Connectivity Project, either aggregated as a single country-level stock or disaggregated into specific forms of village-level, has a significant positive impact on output. Though there are exceptions, most previous studies reported a positive impact on the outcome variables studied. Studies of the impact of road infrastructure on developing economies, however, are relatively few, partly because of lack of reliable data on infrastructure stocks and outcome variables.

This project expects that by reducing transportation costs and information asymmetry by providing accurate, timely and appropriate information as well as opening up opportunities to extend labor activities and diversify income sources through the access to electricity; improved access to transportation as a result of the Connectivity Project will increase the income of the participant households and will help increase their allocation of labor to more profitable non-farm activities. As a result, rural–urban trade should increase, markets become more integrated, and opportunities for rural income generation should develop. Also, as the NTH facilitates trade opportunities in the Northern Zone it is expected that households will have access to a greater selection of cheaper, higher quality goods.

Specifically, and to measure the effects on income, we know that income of a rural households can be expressed as the sum of the incomes that the household receives for the different activities they perform (farm and non-farm activities):

$$Y = \sum_{i=1}^{n} y_i$$

where Y represents the total income of the household and y_i represents the income in each farm and non-farm activities. Similarly, the income in each activity can be broken down into the total hours worked in the specific activity times the hourly wage in the respective activity (i).

$$Y = \sum_{i=1}^{n} l_i \left(\frac{y_i}{l_i} \right)$$

Finally the number of hours in activity *i* can be expressed as the product of the total hours worked (L) and the share of time allocated to activity *i*, therefore we can now express the total income as:

$$Y = L \sum_{i=1}^{n} Cl_i \left(\frac{y_i}{l_i} \right)$$

where L represents the total hours worked in a week by the household members, Cl_i is the distribution of their hours work between farm and non-farm activities and y_i/l_i is the average wage for each type of activity.

Our hypothesis is that given the access to the NTH and consequent reduction of transportation and transaction costs, this will result in a change of income for the rural households through an increase in the demand of rural products and through a change in prices for their farm and non-farm products. This change in income (ΔY), obtained by a household because of better access to transportation, can then be broken down in the following way:

$$\Delta Y = \left(\sum_i \Delta Cl_i \bullet \frac{y_i}{l_i} \right) \bullet L + \left(\sum_i Cl_i \bullet \frac{y_i}{l_i} \right) \bullet \Delta L + \left(\sum_i \Delta Cl_i \bullet \frac{y_i}{l_i} \right) \bullet \Delta L \qquad i = 1 \ldots n$$

where the first component shows the change in the proportion of time the household will assign between farm and non-farm activities given the changes in relative prices as a result of a better link to urban markets (composition effect), and the second component captures the effect of the income because of the increase in hours worked, for example because of access to electricity (employment effect) and the respective increase in demand for their products because of better linkages through better access to roads to the urban markets. Finally, the third component just captures the interaction effect between both components.

Based on this break down, the main outcomes of the reduction of transportation costs and transportation time we propose to evaluate are:

- Changes in agricultural and non agricultural income following the methodology previously outlined. That is, we will identify the price effect and the labor allocation effects between farm and non-farm activities.
- Change in land value
- Change in transaction costs as a result of better accessibility of traders and farmers to the market, for this we will follow the methodology developed by de Janvri, Alain, and Elisabeth Sadoulet (2001)
- Access to public services (healthcare, schools, etc) and their impacts on health and education outcomes
- Specialization in production
- Changes in consumption patterns (more varieties)
- Migration patterns and remittances
- Finally, we will also try to identify differentiated gender effects through a different set of variables as for example total changes in income, changes in hours work, changes in hours worked in non paid activities, changes in hours work in non-farm activities, hours spent in childcare and household chores, etc).

The expected effects from these outcomes are summarized in Table 3.

Table 3
Expected effects from the connectivity projects over major outcomes

Outcome indicator	Expected impact	Gender specific effects
Change in income and consumption patterns	Positive	Bigger positive effects on women
Number of hours worked	Positive	Bigger positive effects on women
Number of hours in leisure	Not clear	Less hours of leisure
Number of hours in non farm activities	Positive	Bigger positive effects on women
Number of hours taking care of kids	No effect	Potential reduction
Change in the value of land	Positive	No differentiate effect
Reduction in traveling time and access to education and health facilities	Better access to school and heatlh facilities	Bigger positive effects on women
Changes over transaction costs for producers	Significant reduction	No clear diffentiate effect
Specialization in production	Increase specialization	No clear diffentiate effect
Migration and remitances	Positive	No clear diffentiate effect

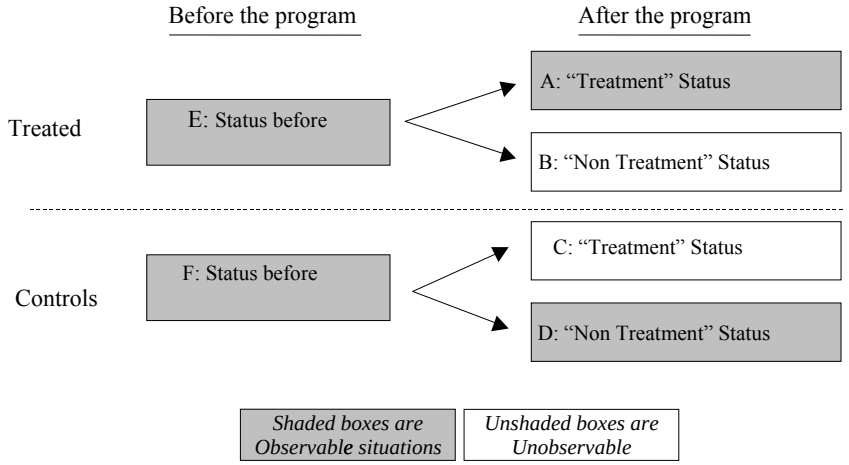
4. Impact Evaluation Approach

4.1 General Approach to be applied in the Connectivity Projects

By conducting an impact evaluation of a given program, we intend to quantitatively estimate the change in the situation of the population due to the execution of the program. Thus we can compare the population's situation given that the program has been executed relative to the population's situation if the program had not been implemented. In other words the basic principle that guides our approach is the comparison between situations "with" the program and "without" the program, also known as "treatment effect" (following Figure 1 what must be determined in an impact assessment is the difference between A and B). This is as opposed to merely comparing the "before" and "after" program implementation scenarios (i.e. assessing the change in the situation of the beneficiary between before and after (comparing A to E) or the difference between participants and non-participants (A to D)). Unfortunately, it is not possible to observe state B, what would have occurred if the participant did not participate. (In Figure 1, the un-shaded boxes represent unobservable situations.)

Figure 1

Possible situations for treated and control households



To address this problem, we need to identify a control group (D) that is as similar as possible to the treatment group, so that observations of D are a close approximation of B. The theoretically ideal approach to constructing a valid counterfactual is to use an experimental approach in which households are randomly assigned to treatment and control groups. Random assignment assures that the distributions of characteristics (both observed and unobserved characteristics) of the two groups are statistically indistinguishable. In our specific program, and generally in the provision of infrastructure services, random assignment of the non-treated is not feasible given that it could conflict with the deployment of the Northern Transnational Highway (NTH). Given this, we include in the design the use of regression discontinuities and continuous treatment in the NTH. Therefore, and given that a full randomization is not feasible, we will implement non-experimental assessment methods to solve this problem, although is important to mention that none of these methods offers a perfect solution (Ravallion 2005). As a result and to be able to control any possible selection variable in all of our designs we will include propensity score matching methods as detailed below.

Propensity score matching of households. We propose using a combination of methods to obtain valid inferences about the household level impacts of the Connectivity Project. These include selecting our analytical samples of recipients and non-recipients to be as similar as possible in terms of their observable characteristics prior to the program. The framework serving as a guideline for our empirical analysis is the Roy-Rubin model (Roy 1951; Rubin 1974, 1977, 1979; Rosenbaum & Rubin 1983). Once propensity scores (predicted probabilities) have been computed for each of the households in the first survey sample, these scores will be used to select pairs of treated and non-treated households that are as similar as possible in terms of their propensities to be in the program.

Information on characteristics of potential survey households before the program was implemented will be acquired from the baseline survey and this requires including as many questions as necessary to control for variables as much as possible. This information will be used to predict the propensity of being in a program using a binary probit model. In computing propensity scores, we also will use information from spatial (GIS) datasets for El Salvador and information from a community survey to take into account the agro-ecological conditions under which the households farm and their access to markets.

Although propensity score matching can ensure that treatment recipients are compared to non-recipients who are similar in terms of observable household characteristics, there still may be both observable and unobservable differences between recipients and non-recipients that may bias the results. For example, the quality and tenure of the plots used by recipients may differ from that of the non-recipients, and such information may not be available or usable for the propensity score matching and sample selection. This type of information will be added in the baseline surveys, and controlled for in the econometric analysis. Other explanatory variables collected in the survey will include household endowments of physical, human, natural, financial, and household access to markets, and services, and other factors influencing households' livelihood options and outcomes.

Even after controlling for such observable differences, there may be differences in unobservable characteristics of recipients and non-recipients that can also bias the results (called "selection on unobservables" in the literature (Heckman, et al. 1998)). Two methods will be used to address this problem. First, some of the relevant unobservables may be relatively similar across households within a village (e.g., level of trust or cultural variables). Inclusion of village-level

fixed effects in the econometric estimation can help to reduce any bias caused by such unobserved factors (Pitt and Khandker 1998).

Second, the double-difference (DD) estimator, which considers the difference between program participants and non-participants in changes in outcomes before and after the program ((A-E)-(D-F)) in Figure 1), will be used (this can be implemented in combination with an econometric approach controlling for differences between the participants and non-participants (Ravallion 2005)). Since we will have panel data collected, this method can be used based on differences in outcomes between the baseline and the follow up survey. This approach removes the effects of any unobserved fixed factors that differ between participants and non-participants, if those have a linear additive impact on outcomes (e.g., differences in abilities). However, these results may be confounded by other changes between the time of the baseline and the follow up survey besides the interventions program influencing changes in outcomes differentially between participants and non-participants (e.g., changes in access to other programs). Changes in such other factors will be accounted for to the extent possible. However, there still may be biases caused by changing unobserved factors that differ between program treated and non-treated households/individuals within villages (e.g., access to information). Instrumental variables (IV) estimation can be used to control for the remaining potential sources of selection bias.

4.2 Specific Approach for the Connectivity Project: Northern Transnational Highway (NTH)

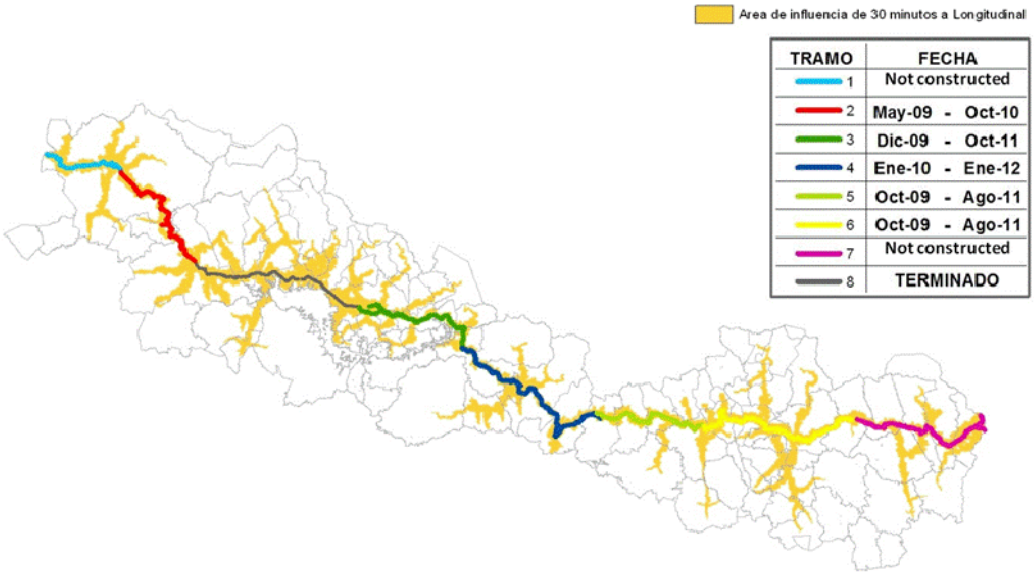
Since an experimental design is not feasible, a non-experimental design will be implemented as previously explained. "Roads are clearly not randomly placed, and it is highly likely that the factors that led to the road placement will also affect outcomes."[3] The proposed non-experimental approach first consists of restricting the area of influence of the impacts of the NTH. Defining an area of influence is of crucial importance given it is a way to identify households with similar characteristics in terms of their closeness to the main road (i.e. is equivalent to a one dimension first stage non-parametric matching). With this objective we implement an accessibility measure to identify the communities and households which are within 30 minutes from the NTH. This accessibility measure calculates the shortest time or distance from any household to a regional or local market using the distance traveled on different road surfaces combined with an impedance measure which reflects the traveling speeds on roads of different qualities and on the slope of the terrain through which the road passes. The resulting market access measure can be expressed as a weighted average of the distance traveled on each type of road, where the weights are proportional to the impedance factor. Figure 2 shows in yellow this area of influence.

Once the area of influence is identified the next step was to take advantage of the geographic location and the timeline of the implementation of the NTH to identify the potential control and treatment groups. We follow two approaches to evaluate the impact of the connectivity project. The first one is using the segments and the sequencing of the interventions to have potential neighboring control groups to the segments where the road is built first. This is similar to what is done under a regression discontinuity approach. The second approach is to implement a continuous treatment approach. In this later case, we don't have a control group but we

[3] World Bank, January, 2002. Van de Walle, Dominique and Dorothyjean Cratty: Impact Evaluation of a Rural Road Rehabilitation Project.

have‘treatment intensity.’ This is the change in time to get to the nearest market[4] caused by the improvement of the highway and the connecting road network.

Figure 2
Area of Influence targeted for the impact evaluation



4.2.1 Regression discontinuity approach

In this case, if a small variation in some explanatory variable (e.g. being part of one of the segments to be improved) produces a discontinuous change in eligibility for the program, given that each road segment will be improved at different times, then this discontinuity in differences in access will be used as an IV to identify the impacts, even if that variable is also a direct determinant of the responses and outcomes of interest (Ravallion 2005).

Therefore we will take advantage of the timeline of implementation of each of the sections of the division of the NTH and use the concept of regression discontinuity (Pitt and Khandker; 1998) where the “thresholds” will be defined by the dates of intervention so that control (segments implemented at a later date) and treatment groups (earlier implemented segments) can be defined. As a result, the following treatment-control groups are suggested and four different tests of the difference between these groups can be implemented:

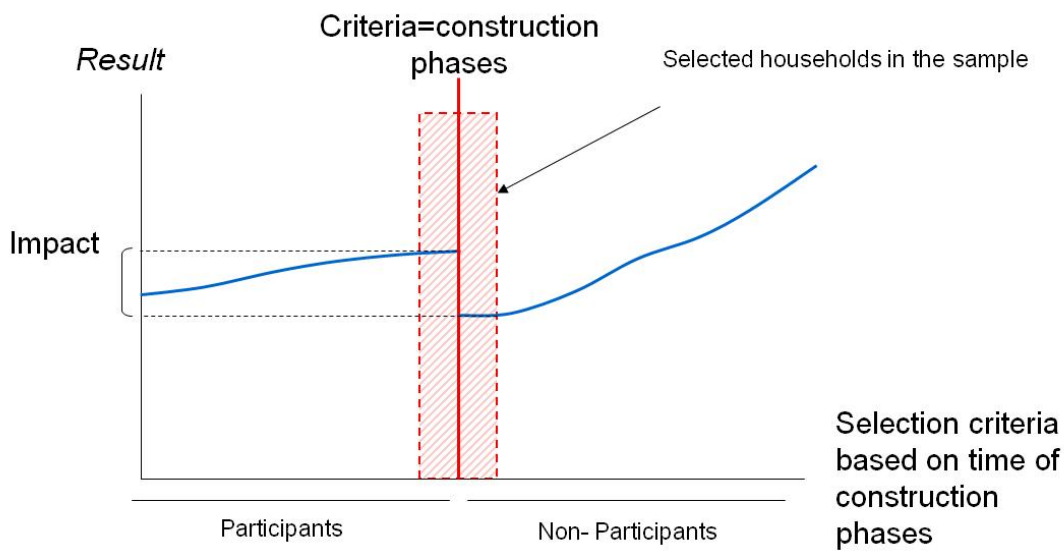
[4] Defined as any town/city with a population of 25,000 inhabitants or more

Table 4
Treatment and Control Groups

Test Number	Treatment Group	Control Group
1	T2	T1
2	T2	T3
3	T3	T5
4	T4	T5

For example, segment 2 (T2) will be implemented from May 2009 to October 2010 while segment 1 (T1) will not be implemented, making it an appropriate control for segment T2. Therefore we can argue that T1 could be the control group (non-participants as defined in Figure 3) of T2 (participants). The objective is to compare contiguous segments when the dates of intervention are sequential or don't overlap.

Figure 3
Regression discontinuity design



4.2.2 Continuous treatment approach

The impact of the highway can also be analyzed using tools of the recently developed literature on continuous treatment (Imbens 2000, Behrmann, Cheng and Todd 2004, Hirano and Imbens 2004, and Imai and van Dyk 2004). This section follows the exposition by Hirano and Imbens (2004). In this case, the treatment is the change in time to get to the nearest market[5] caused by the improvement of the highway.

[5] Defined as any town/city with a population of 25,000 inhabitants or more

Households from segment *j* will benefit from improvements in segment *i* (which may or may not be equal to *j*) if they use that part of the highway to send their products to the market. Therefore, the treatment can be interpreted as the change in time to the market.

Let T stand for the treatment (change in time to the nearest market), X for a set of covariates, and R for the Generalized Propensity Score. If $T_i \mid X_i \sim N(\beta_0 + \beta_1^{'} X_i, \sigma^2)$, we can define a Generalized Propensity Score (GPS) and estimate it via OLS or ML

$$\hat{R}_i = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\left(-\frac{1}{2\hat{\sigma}^2} \left(T_i - \hat{\beta}_0 - \hat{\beta}_1^{'} X_i \right)^2 \right) \ldots(7)$$

Next, we can estimate the conditional expectation of the outcome variable given the treatment and the GPS with the equation:

$$E\left[Y_i \mid T_i, R_i\right] = \alpha_0 + \alpha_1 \cdot T_i + \alpha_2 \cdot T_i^2 + \alpha_3 \cdot \hat{R}_i + \alpha_4 \hat{R}_i^2 \cdot + \alpha_5 \cdot T_i \cdot \hat{R}_i \ldots(8)$$

Finally,

$$\hat{E}\left[Y_i \mid T_i, R_i\right] = \frac{1}{N} \sum_{i=1}^{N} \left(\hat{\alpha}_0 + \hat{\alpha}_1 \cdot t + \hat{\alpha}_2 \cdot t^2 + \hat{\alpha}_3 \cdot \hat{r}_i(t, X_i) + \hat{\alpha}_4 \cdot \hat{r}(t, X_i)^2 + \hat{\alpha}_5 \cdot t \cdot \hat{r}(t, X_i) \right) \ldots(9)$$

The Average Dose-Response Function (ADRF) is obtained by estimating $E\left[Y_i \mid T_i, R_i\right]$ for every value of t (which implies re-estimating *r* in each stage). To test for the effect being zero, we need to test the following joint hypothesis

H_0: $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0$ against the alternative
H_A: ~ ($\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0$)

This allows calculating the joint significance of the coefficients. Under the alternative hypothesis (that at least one of the coefficients is not zero), the F statistic has a non-central F distribution. A non-central F distribution is the ratio of a non-central chi-squared and a (conventional) chi-squared random variable (for further detail, see appendix 3).

Three main conceptual differences from the regression discontinuity approach arise:

1. Now there is only one group, because all households are treated (there is no control group). What varies is the intensity of the treatment assigned to each household.
2. We are testing a joint hypothesis, therefore we need to use F statistics instead of t statistics
3. Since F distributions arise from the ratio of two chi-squared variables, their range is $[0, \infty]$, hence the tails are inherently one-tailed.

We propose using both approaches, i.e. the regression discontinuity and the continuous treatment approach. Both of these approaches complement each other and at the same time minimize the risks of the impact evaluation implementation in case there are important changes in the segments implementation which could damage the design of the regression discontinuity approach. At the same time, the regression discontinuity approach allow us to have local average treatment effects

which could capture important differences between the different sub-regions in the northern zone where the project will be implemented.

4.3 Calculating the cost of travel

The cost of travel will be the main treatment variable. The SI team proposes to geo-reference all households in the baseline survey using GIS data already collected in order to investigate the cost of travel in a more sophisticated way. With this data we can calculate the shortest time or distance from any village to a regional or local market and can express the resulting market access measure as a weighted average of the distance traveled on each type of road, where the weights are proportional to the impedance factor. [6]

However, there is one problem with this and other typical measures of market access based on time. They do not incorporate transportation costs which may vary differently with time than with distance and type of road surface. Where that is the case, the measure would be misleading: even though the market might be closer in terms of time, it may cost more to get there than it would to get to a different market. To address this, the SI team will use a measure that incorporates both time and costs by merging market distance data for each village with a matrix of transportation costs by truck (based on a costing model with inputs from transportation companies from El Salvador) on two different classes of roads (or by animal on trails where there are no roads). This resulting measure will give the SI team a better and more comprehensive way to evaluate the impacts of the intervention. The information needed for the transportation costs will mainly come from the baseline survey and information of fuel prices across the northern zone.

In summary, under the continuous treatment approach we will:

- Measure treatment as a continuous variable of change in cost to travel to fixed points of interest (this will vary by household and therefore we will call it δ_i) such as market places, health centers, schools, etc. The cost of travel will be calculated as previously explained. Initially, both the control and treatment households should have a similar distribution of δ_i.
- Distinguish the impact of "δ_i" (measured in dollars) on the impact indicators previously explained as income, health and other benefits, within the treatment households. This will allow us to measure the marginal gain within the treated communities comparing with the difference in "δ_i ". However, to capture the causal impact of the changes in "δ_i" we recommend measuring the differences between the treatment and control households using a difference in difference approach. This will allow controlling for any possible selection on observables, and also any possible selection on unobservable fixed over time.

5. The sample design

We assumed a clustered, quasi-experimental evaluation design with treatments administered at the cluster level and data collection before and after initiation of the treatments. With this design, impact estimates can be measured using the preferred approach of taking difference-in-differences or "double difference": the change in the outcome in the treatment group minus the

[6] Creating this measure involves using the distance traveled on different road surfaces combined with an impedance measure which reflects the speed one can travel on roads of different qualities and the slope of the terrain through which the road passes.

change in the outcome in the quasi-randomized control (or alternate treatment) group. The purpose of the sample size estimates is to determine the minimum impact, Δ, that can be detected for a given number of sampled clusters, g, and households per cluster, m, in each treatment for the evaluation sample.[7] If the impact of the treatment is at least as large as Δ, we will be able to detect it 80 percent of the time in a sample of total size *mg*. If the treatment impact is less than Δ, we are less likely to detect it, although it is still possible. Following the methodology outlined in Appendix 1 the sample size we propose for each of the two types of impact evaluation design are:

5.1 Regression discontinuity approach

Under the assumptions of Type I and II error rates of 5% and 20% respectively and a change in incomes of at least 20% based on estimations done with previous household surveys in El Salvador's Northern Zone . If the treatment is considered discrete, we recommend the following sample size:

Table 5
Recommended Sample Size for Regression Discontinuity Approach

Area	Intra-cluster Correlation	Observations in Treatment Group	Observations in Control Group	Total Observations in each area
Highway section				
T2 vs T1	0.040	233	233	466
T2 vs T3	0.117	414	414	828
T3 vs T5	0.144	432	432	864
T4 vs T5	0.103	674	674	1348
Total Sample				**3506**

5.2 Continuous treatment approach

Under similar assumption but for the continuous treatment approach we recommend the following sample size:

[7] In addition to *g* and *m*, the minimum detectable impact, Δ, is a function of the variance of the outcome variable, its intracluster correlation, and the area of influence of the highway being evaluated.

Table 6

Recommended Sample Size for Continuous Approach

	ICC	Sample size [1]	Implied t-value [2]
Highway sections			
1-2	0.00	647	3.67
3	0.01	432	2.82
4	0.03	674	2.89
5-7	0.13	2022	2.11
Highway as a whole	0.090	3775	3.87

[1] We input the sample size that arises from table 1 for each section except for sections 5-7, where we multiply times 3 the value of 674 obs.

[2] A t-value of 2.8 or greater guarantees power of at least 80%

[3] Assumptions:	ICC	0.12
	σ^2_T	213.84
	σ^2_Y	45,727.81
	$ß^2$	1.35
	ρ^2_{TY}	0.00018
Where	T: Treatment (reduction in time to market)	
	Y: outcome (household income)	
	ß: coefficient in the OLS regression of Y on T and a constant	

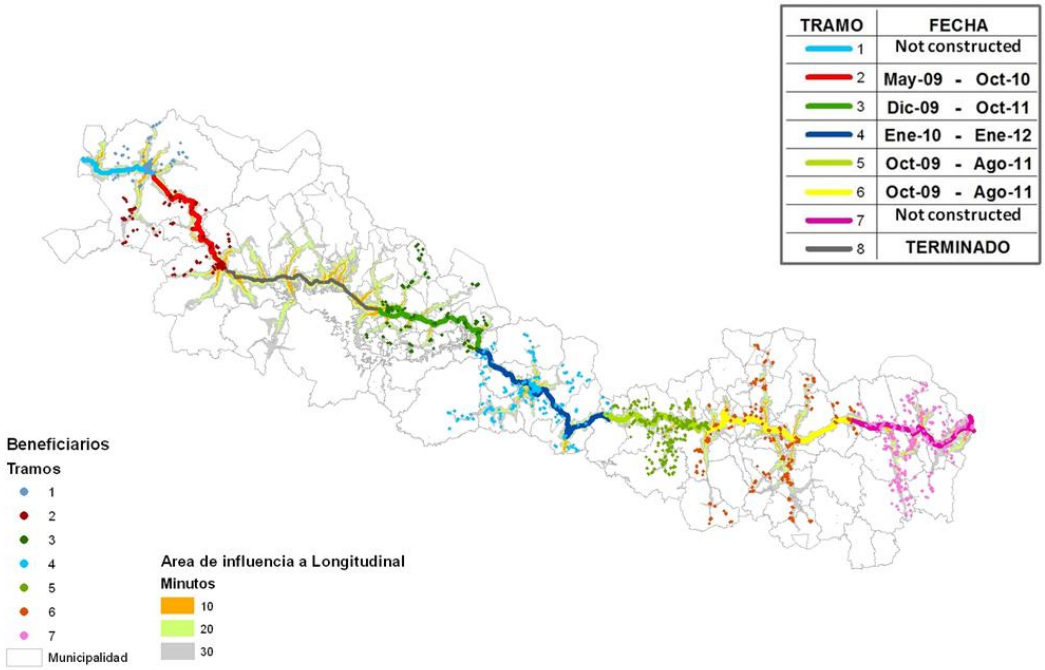
In summary, if we implement a survey of 3,775 households following the power calculations for the continuous approach, we would also comply with the sample requirements in the regression discontinuity case. This means that we do not need to add more assumptions about the variance of the outcome variable and its correlation through time and within individuals.

Table 7

Total Sample required for both approaches

Area	Required Observations
Highway section	
T1	233
T2	414
T3	432
T4	674
T5	674
T6	674
T7	674
Total Observations	3,775

Figure 3
Sample required for evaluating the connectivity project



6. Data collection

The evaluation will use household surveys and community surveys. To identify the households to be sampled a census of the selected areas was implemented and a sample framework was developed. From the sample framework a stratified random sample was selected[8].

The household survey will interview approximately 3,775 households as detailed in section 5. The questionnaire includes two sections – one that will be answered by the primary male household representative (including household income and agricultural productivity) and will be administered by a male enumerator, and the second which will be answered the primary female representative in the household (including household demographics, time allocation, and expenses) and administered by a female enumerator. The survey has detailed sections for each of the outcomes to be evaluated, both intermediate and final outcomes. In addition, and to be able to control for accessibility to markets, each of the survey households was geo-referenced. If both persons are not present at the time of the first visit, enumerators will attempt to make an appointment and return again to interview the appropriate person, provided that this return visit is possible

[8] For details on sample selection see sample selection final report.

within the time that the survey team will be in the area. When possible, a second adult can also be included in the interview process, particularly for the questions related to work and agricultural output. The survey is designed to take between 1 and 1 ½ hours for each questionnaire (i.e. male and female).

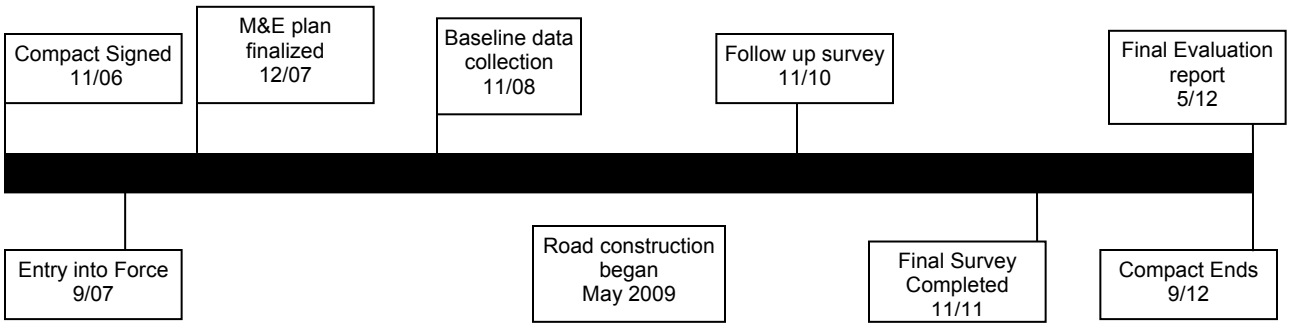
The community survey will be applied to key informants of communities where selected households live; each section of the survey was applied to the better informed informant. For example for issues related to the community infrastructure and prices to the mayor or community leader, for information on health of the community members to the health center director, and for issues related to education to the school director. This survey will gather information about the local economy; price levels for food, basic commodities, and water and sanitation related expenditures; community infrastructure and access key markets and social services. The goal of the surveys is to provide some context for the information gathered in the household surveys, to track community-level changes that may affect outcomes, and to reduce the required length of the household survey questionnaire.

The current data collection plan anticipates that each household will be surveyed three times, 1) baseline in November 2008, 2) follow up in November 2010 and 3) final in November 2011. This may change however if there are delays in the construction schedule.

7. Timeline

The timeline of the implementation of the design is described in Figure 5.

Figure 5
Timeline of Implementation



8. References

General

Ahmed, R., and M. Hossain. 1990. *Development impact of rural infrastructure.* IFPRI Research Report No. 83. Washington, D.C.: International Food Policy Research Institute.

Antle, J. M. 1983. Infrastructure and aggregate agricultural productivity: International evidence. *Economic Development and Cultural Change* 31 (3): 609–619.

_____. 1984. Human capital, infrastructure, and the productivity of Indian rice farmers. *Journal of Development Economics* 14 (1/2): 163–181.

Aschauer, D. A. 1989a. Is public expenditure productive? *Journal of Monetary Economics* 23 (2): 177–200.

_____. 1989b. Public Investment and productivity growth in the group of seven. *Economic Perspectives* 13 (5): 17–25.

Angrist, J., Imbens, G., & Rubin, D. (1996). Identification of causal effects using instrumental variables. Journal of the American Statistical Association , 91.

Angrist, J., Bettinger, E., Bloom, E., King, E., & Kremer, M. (2002). Vouchers for private schooling in Columbia: evidence from a randomized natural experiment. *American Economic Review* , 1535-1558.

Binswanger, H. P., S. R. Khandker, M. P. Rosenzweig. 1993. How infrastructure and financial institutions affect agricultural output and investment in India. *Journal of Development Economics* 41 (2): 337–366.

de Janvri, Alain, and Elisabeth Sadoulet (2001). ”Income Strategies Among Rural Households in Mexico: the Role of Off-farm Activities.” *World Development* 29: 467-480.

De Janvry, A., Sadoulet, E., & Villas Boas, S. (January 2006). *Short on shots: can self-restraint be effective in managing the scarcity of a vital good?* University of California at Berkeley.

Dong, X. 2000. Public investment, social services and productivity of Chinese household farms. *Journal of Development Studies* 36 (3): 100–122.

Duflo, E., & Saez, E. (2003). The role of information and social interactions in retirement plan decisions: evidence from a randomized experiment. *Quarterly Journal of Economics, 118* (3), 815-842.

Escobal, Javier, & Torero, Maximo. (2005). Measuring the impact of asset complementarities: the case of rural Peru. Cuadernos de Economia 42. Pp. 1-26.

Fan, S., Hazell, P., and Thorat, S. 1999. Linkages Between Government Spending, Growth, and Poverty in Rural India. Research Report 110, International Food Policy Research Institute, Washington, DC.

_____. 2000. Government spending, growth and poverty in rural India. *American Journal of Agricultural Economics* 82 (4): 1038–1051.

Garcia-Mila, T., T. J. McGuire, and R. H. Porter. 1996. The effect of public capital in state-level production functions reconsidered. *Review of Economics and Statistics* 78 (1): 177–180.

Gibson, J., and S. Rozelle. 2003. Poverty and access to roads in Papua New Guinea. *Economic Development and Cultural Change* 52 (1): 159–185.

Heckman, J, H Ichimura, J Smith, & P Todd. 1998. Characterizing selection bias using experimental data. *Econometrica*. 66: 1017-99.

Hirschmann, A. O. 1959. The strategy of economic development. 2nd Edition. New Haven, Conn., U.S.A.: Yale University Press.

Holtz-Eaakin, D. 1994. Public-sector capital and the productivity puzzle. *Review of Economics and Statistics*. Vol 75 (4):12–21.

Jacoby, H. 2000. Access to markets and the benefits of rural roads. *Economic Journal* 110 (465): 713–737.

Kling, J., Ludwig, J., & Katz, L. (2005). Neighborhood effects on crime for female and male youth: evidence from a randomized housing voucher experiment. *Quarterly Journal of Economics* , 87-130.

Levine, D., Hema, N., & Ramage, I. (July 2007). *Insuring health: testing for the effectiveness of micro-health insurance to promote economic wellbeing for the poor.* BASIS brief no 2007-05.

Morrison, C. J., and A. E. Schwartz, 1996. State infrastructure and productive performance. *American Economic Review* 86 (5): 1095–1111.

Nadiri, M. I., and T. P. Mamuneas. 1994. The effects of public infrastructure and R&D capital on the cost structure and performance of U.S. manufacturing industries. *Review of Economics and Statistics* 76 (1): 22–37.

Norton, S. W. 1992. Transaction costs, telecommunications and the microeconomics of macroeconomic growth. *Economic Development and Cultural Change* 41 (1): 175–196.

Paul, C. J. M., V. E. Ball, R. G. Felthoven, and R. Nehring. 2001. Public infrastructure impacts on U.S. agricultural production: A state-level panel analysis of costs and netput composition. *Public Finance and Management* 1 (2): 183–213.

Röller, L.-H., and L. Waverman. 2001. Telecommunications infrastructure and economic growth: A simultaneous approach. *American Economic Review* 91 (4): 909–923.

Pitt, M & SR Khandker. 1998. "The impact of group-based credit programs on poor households in Bangladesh: does the gender of participants matter?" *Journal of Political Economy.* 106(5): 958-996.

Ravallion, M. 2005. *Evaluating anti-poverty programs*. World Bank Policy Research Working Paper 3625. The World Bank, Washington, D.C.

Rawlings, L.B. and N.R. Schady. 2002. Impact Evaluation of Social Funds: An Introduction. *The World Bank Economic Review* 16 (2): 213-217.

Rosenbaum, PR & DB Rubin. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika.* 70(1): 41-55.

Roy, A. 1951. "Some Thoughts on the Distribution of Earnings". *Oxford Economic Papers.* 3: 135-145.

Rubin, D. 1974. "Estimating Causal Effects to Treatments in Randomized and Nonrandomized Studies". *Journal of Educational Psychology*. 66: 688-701.

Rubin, D. 1977. "Assignment to Treatment Group on the Basis of a Covariate". *Journal of Educational Studies*. 2:1-26.

Rubin, D. 1979. "Using Multivariate Matched Sampling and Regression Adjustment to Control Bias in Observational Studies". *Journal of the American Statistical Association.* 74: 318-328.

Roads

Bakht, Z. (2000). *Poverty Impact of Rural Roads and Markets Improvement & Maintenance Project of Bangladesh.* Paper presented at the World Bank South Asia Poverty Monitoring and Evaluation Workshop, June 8-10, 2000, India Habitat Centre, New Delhi.

Jacoby, H. C. (2000). Access to markets and the benefits of rural roads. *Economic Journal, 110*(465), 713 737.

Khan, Q. M. (1990). *Transaction costs approach for estimating development benefits of rural feeder roads.* Paper presented at the Transportation Research Record No. 1274. Transportation and Economic Development Conference, Williamsburg, Virginia; November 5-8, 1989.

Levy, H. (1996). *Morocco - Socioeconomic Influence of Rural Roads : Fourth Highway Project* (Impact Evaluation Report. Report No. 15808). Washington, D.C.: The World Bank. Operations Evaluation Department.

Windle, J., & Cramb, R. A. (1996). *Roads, remoteness and rural development: social impacts of rural roads in upland areas of Sarawak, Malaysia.* (Agricultural Economics Discussion Paper 3/96). Brisbane, Australia: Department of Agriculture. University of Queensland.

Appendix 1
Sample Size and Power Calculation

1. Introduction

The Longitudinal Highway runs from East to West in northern El Salvador, connecting the country's borders from Honduras to Guatemala. This highway is about to undergo a major program of road improvement, which can provide many valuable lessons. The power calculations lead to the conclusion that a survey must have a sample size of at least 3,775 observations to be able to detect the effect of the road improvement program on rural (total household) incomes, while also being representative of the highway's different sections. Due to the high intra-cluster correlations observed in variables as non-agricultural waged income or time allocated to non-agricultural non-wage labor, the power to detect differences in such variables will be lower, although it may still be possible to detect differences.

Section 2 deals with survey design issues like the intracluster correlation and assumptions in the variance calculation. The third section covers the main issues regarding power calculation when the treatment is discrete and continuous. Section 4 summarizes the findings and recommends specific sample sizes for each section of the highway.

2. Survey Design

We assumed a clustered, quasi-randomized evaluation design with treatments administered at the cluster level and data collection before and after initiation of the treatments. The purpose of the sample size estimates is to determine the minimum impact, Δ, that can be detected for a given number of sampled clusters, g, and households per cluster, m, in each treatment for the evaluation sample.[9] If the impact of the treatment is at least as large as Δ, we will be able to detect it 80 percent of the time in a sample of total size *mg*. If the treatment impact is less than Δ, we are less likely to detect it, although it is still possible.

2.1 Intra-cluster Correlation

The most controversial issue in sample design is the intra-cluster correlation, so we will proceed to make the calculation procedure explicit. DIGESTIC and MOP provided detailed GIS data on the road and the location of all the dwellings of northern El Salvador. In an ideal scenario, we would have the relevant socio-economic data from the census as well, but at the time of writing this was not available. The intra-cluster correlation of several variables (was calculated from the EPHM Survey 2007). Merging the survey and the GIS data, the cantons where the road goes through were identified and matched to the household survey data. The universe is constituted by the set of cantons that were identified in the census as having dwellings under different access time thresholds to the freeway to be developed (10, 20, and 30 minutes)[10]. The sub-set of cantons that were also included in the household survey constitutes "level 1".

---

[9] In addition to *g* and *m*, the minimum detectable impact, Δ, is a function of the variance of the outcome variable, its intracluster correlation, and the area of influence of the highway being evaluated.

[10] The accessibility measure is calculated using Geographic Information System (GIS). Digital information of roads, rivers and the slope of the terrain are used in order to build a measure of accessibility to the northern highway on El Salvador. The proper combination of this data allows the calculation of the least cost path surface (based on time) from any household to the nearest point of the northern highway.

For those cantons that were not included in the NHS survey, the municipality income and time allocation data was imputed. This group plus "level 1" constitutes "level 2". In turn, for those municipalities that were not included in the survey, the department data was imputed. These subset plus "level 2" conforms "level 3". For the 10 min threshold we recommend the use of level 1, and for the 20 and 30 minute threshold we recommend level 2 (canton plus municipality). Level 3 is not recommended because the department data may not be a good approximation to the households near the road.[11]

Several outcome variables were used in the analysis. We will summarize the results for overall household income, but the analysis also included weekly working hours, and both split by wage agriculture, non-wage agriculture, wage non-agriculture, and non-wage non-agriculture (see Appendices 5-9).

2.2 Scenarios for Variance Calculation

There are three important differences between the proposed sample for evaluation and the NHS sample, all of which are likely to affect the sample variance in the projected sample relative to that in the NHS sample: First, we estimate variance of the primary outcomes in the NHS using only one round of data collection, rather than two. The variance of the difference between the two measures depends upon the variance of each measure as well as the correlation over time between the two measures. We do not know this correlation; so we must make assumptions about it, which we vary below. Second, we stratified the sample for the collection of this data, in order to both balance the sample and reduce the sampling variance. The reduction in sampling variance will depend upon the variance between strata means; the larger the difference between the average outcomes across strata, the higher the variance reduction will be. Third, the NHS measures the variance of outcomes related to different levels of current access to roads; it is likely that the variance of baseline will be smaller given the assumptions of accessibility we are imposing.

Since the three differences between the proposed surveys and the NHS will certainly affect the variance of primary outcomes, we experiment with power calculations using several different variance estimates. First and most conservatively, we simply double the variance of the NHS outcomes; doing so assumes that the primary outcome will not be correlated across the two surveys, that each strata will have exactly the same mean outcome, and that the treatment will not affect the variance of the treatment. Second, we reduce the doubled variance by 10 percent, to simulate a significant decline in sample variance due to stratification. Third, we simply compute the power calculations using the NHS variance. Finally, we use the NHS variance less 10 percent, to account for gains from stratification, but also assume between-period correlation of 0.5 and a within-period correlation of -0.5. Since we also ignore the above assertion that the baseline variance in outcomes is likely to be smaller than the NHS variance, the fourth estimate is likely to be the most realistic and the one we propose to use.

3. Power calculations

3.1 Discrete Treatment

This model simulates the time it takes a person to reach the nearest point of the highway, assuming that people prefer to travel via highways, roads and trails, and also considering rivers or lakes as barriers if there is no bridge available. The model applies a cost-weighted distance algorithm is run in order to calculate the accumulated time required to travel from any location of the surface to the destination point. In this case, each household gets a weighted average of the time traveled to the highway.

[11] We would be able to include the cantons in "level 3" if the census socio-economic data were available.

The impact evaluation will be conducted with difference-in-difference estimators. This methodology requires repeat observations on members. Power calculations for this type of survey designs were based on Murray (1998, chapter 9). The main analysis is based on the following three equations (the equation number in Murray's book is in brackets):

$$g = \frac{2(1+(m-1)I\hat{C}C)(t_{\alpha/2}+t_{\beta})^2}{m\hat{\Delta}^2}\hat{\sigma}_y^2 \quad \ldots(1)\ [9.23]$$

$$\hat{\sigma}_y^2 = \frac{mg}{2(1+(m-1)ICC)}\hat{\sigma}_{\Delta}^2 \quad \ldots(2)\ [9.17]$$

$$\hat{\sigma}_{\Delta}^2 = 4\left[\frac{\hat{\sigma}_m^2(1-\hat{r}_{yy(m)})+m\hat{\sigma}_g^2(1-\hat{r}_{yy(g)})}{mg}\right] \ldots(3)\ [9.34]$$

Where:

g: number of clusters in each condition (treatment/control)
m: number of observations per cluster
ICC: Intracluster correlation
α: type I error rate
ß: type II error rate
$\hat{\sigma}_y^2$: estimated variance of the outcome variable
$\hat{\Delta}$: estimated change
$\hat{\sigma}_{\Delta}^2$: estimated variance of the change in the outcome variable
$r_{yy(g)}$: inter-period correlation
$r_{yy(m)}$: intra-period correlation

Replacing (2) in (1), we get:

$$1 = \frac{(t_{\alpha/2}+t_{\beta})^2}{\hat{\Delta}^2}\hat{\sigma}_{\Delta}^2 \ldots(4)$$

Inserting (3) in (4):

$$1 = \frac{(t_{\alpha/2}+t_{\beta})^2}{\hat{\Delta}^2}\times 4\left[\frac{\sigma_m^2(1-r_{yy(m)})+m\sigma_g^2(1-r_{yy(g)})}{mg}\right] \ldots (5)$$

Solving for g:

$$g = \frac{(t_{\alpha/2}+t_{\beta})^2}{\hat{\Delta}^2}\times 4\left[\frac{\sigma_m^2(1-r_{yy(m)})+m\sigma_g^2(1-r_{yy(g)})}{m}\right] \ldots(6)$$

Table A.3 presents the results for the different scenarios assumed in section 3.2.

Table A.3. Number of Clusters per Condition[1] and Total Sample Size[2] for Household Income[3] for each Scenario[4]

Test Number	Treatment Group	Control Group	Intracluster correlation[6]	Scenario 1 Clusters per condition	Scenario 1 Total sample size	Scenario 2 Clusters per condition	Scenario 2 Total sample size	Scenario 3 Clusters per condition	Scenario 3 Total sample size	Scenario 4 Clusters per condition	Scenario 4 Total sample size
m=25[5]											
1	T2	T1	0.00	20	1021	18	919	10	510	10	478
2	T2	T3	0.04	68	3405	61	3064	34	1702	17	851
3	T3	T5	0.06	88	4407	79	3966	44	2203	18	888
4	T4	T5	0.04	112	5586	101	5027	56	2793	28	1385
m=35											
1	T2	T1	0.00	14	1003	13	903	7	502	7	470
2	T2	T3	0.04	58	4026	52	3623	29	2013	12	836
3	T3	T5	0.06	77	5421	70	4879	39	2711	12	873
4	T4	T5	0.04	94	6612	85	5951	47	3306	19	1361
m=45											
1	T2	T1	0.00	11	994	10	894	6	497	5	466
2	T2	T3	0.04	52	4661	47	4195	26	2331	9	828
3	T3	T5	0.06	72	6450	65	5805	36	3225	10	865
4	T4	T5	0.04	85	7662	77	6896	43	3831	15	1348

[1] The conditions are "treatment" and "control". The number of clusters in each condition is equal
[2] Total sample size (treatment + control)
[3] The outcome variable is total monthly household income
[4] For the specification of each scenario see section 3.2, and for the formulae, see Appendix 2 and 3.
[5] Number of observations (households) per cluster
[6] Observed in the NHS with "level 2" households.
[7] α=0.05; ß=0.20; Δ =0.20

3.2 Continuous Treatment

The impact of the highway can also be analyzed using tools of the recently developed literature on continuous treatment (Imbens 2000, Behrmann, Cheng and Todd 2004, Hirano and Imbens 2004, and Imai and van Dyk 2004). This section follows the exposition by Hirano and Imbens (2004). In this case, the treatment is the change in time to get to the nearest market[12] caused by the improvement of the highway.

Households from segment *j* will benefit from improvements in segment *i* (which may or may not be equal to *j*) if they use that part of the highway to send their products to the market. Therefore, the treatment can be interpreted as the change in time to the market.

Let T stand for the treatment (change in time to the nearest market), X for a set of covariates, and R for the Generalized Propensity Score. If $T_i \mid X_i \sim N(\beta_0 + \beta_1^{'} X_i, \sigma^2)$, we can define a Generalized Propensity Score (GPS) and estimate it via OLS or ML

$$\hat{R}_i = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp\left(-\frac{1}{2\hat{\sigma}^2} \left(T_i - \hat{\beta}_0 - \hat{\beta}_1^{'} X_i \right)^2 \right) \ldots(7)$$

Next, we can estimate the conditional expectation of the outcome variable given the treatment and the GPS with the equation:

$$E\left[Y_i \mid T_i, R_i\right] = \alpha_0 + \alpha_1 \cdot T_i + \alpha_2 \cdot T_i^2 + \alpha_3 \cdot \hat{R}_i + \alpha_4 \hat{R}_i^2 \cdot + \alpha_5 \cdot T_i \cdot \hat{R}_i \ldots(8)$$

Finally,

$$\hat{E}\left[Y_i \mid T_i, R_i\right] = \frac{1}{N} \sum_{i=1}^{N} \left(\hat{\alpha}_0 + \hat{\alpha}_1 \cdot t + \hat{\alpha}_2 \cdot t^2 + \hat{\alpha}_3 \cdot \hat{r}_i(t, X_i) + \hat{\alpha}_4 \cdot \hat{r}(t, X_i)^2 + \hat{\alpha}_5 \cdot t \cdot \hat{r}(t, X_i) \right) \ldots(9)$$

The Average Dose-Response Function (ADRF) is obtained by estimating $E\left[Y_i \mid T_i, R_i\right]$ for every value of t (which implies re-estimating *r* in each stage). To test for the effect being zero, we need to test the following joint hypothesis

H_0: $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0$ against the alternative
H_A: ~ ($\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0$)

This allows to calculate the joint significance of the coefficients. Under the alternative hypothesis (that at least one of the coefficients is not zero), the F statistic has a non-central F distribution. A non-central F distribution is the ratio of a non-central chi-squared and a (conventional) chi-squared random variable (for further detail, see appendix 3).

Three main conceptual differences from section 1 arise:

4. Now there is only one group, because all households are treated (there is no control group). What varies is the intensity of the treatment assigned to each household.

[12] Defined as any town/city with a population of 25,000 inhabitants or more

5. We are testing a joint hypothesis, therefore we need to use F statistics instead of t statistics
6. Since F distributions arise from the ratio of two chi-squared variables, their range is $[0,\infty[$, hence the tails are inherently one-tailed.

Johnston and DiNardo (1997: 493-495) show that, under the null hypothesis:

- $(Rb-r)'[\sigma^2 R(X'X)^{-1}R']^{-1}(Rb-r) \sim \chi_k^2$
- $\dfrac{e'e}{\sigma^2} \sim \chi_{n-k-1}^2$
- and that these variables are statistically independent from each other.

Hence, under the null hypothesis:

$$\frac{(Rb-r)'[R(X'X)^{-1}R']^{-1}(Rb-r)/k}{e'e/(n-k-1)} \sim F\ (k, n-k-1) \ldots(10)$$

With 5 regressors, any sample size between 100 and 10000 observations, and 5% of type I error rate, the critical F value fluctuates between 2.2231 (for n=1000) and 2.2150 (for n=10,000). Hence, we will take $F_\alpha = 2.22$ as the critical value for the type I error rate.

Under the alternative hypothesis $E[Rb-r] = b_0 \neq 0$, hence:

$$\frac{(Rb-r)'[R(X'X)^{-1}R']^{-1}(Rb-r)/k}{e'e/(n-k-1)} \sim noncentralF(k, n-k-1, \lambda) \ldots(11)$$

One needs to assume a range for the non-centrality parameter $\lambda = \sum_{i=1}^{k}\left(\frac{\mu_i}{\sigma_i}\right)^2$, which can be proxied by the sum of squares of the t-statistics.

The following table shows the non-central F values for different assumptions about the non-centrality parameter. In first place, if only one coefficient is significant (and, in the extreme, all the others are zero) the non-centrality parameter will be 4. If there are only two significant coefficients the non-centrality parameter will be 8. If there are 3 or more, the non-centrality parameter will be higher than 10. These values are imputed as plausible values for the non-centrality parameter. Like in the case of the (central) F values, changing the sample size from 100 to 10,000 generates changes of under 0.001 in the non-central F values. Therefore, these values will be used for "any" sample size. The results are displayed below.

Table A4. Simulated non-central F-values under most plausible scenarios

If k=5 ; 100<n<10,000 ; power=80%	
non-centrality parameter λ	F_β
4	0.41
8	1.01
10	1.38

To ensure that the Type II error rate is ß, the probability that $F > F_\alpha$ must be $1-\beta$ for the expected value of the point estimate under the alternative hypothesis. This happens only when $F > F_\alpha + F_\beta$ or equivalently when $F > F_\alpha - F_{1-\beta}$.

$$Power = \Pr(\hat{F} \geq F_\alpha - F_{1-\beta})$$

Thus, in a simple randomized trial (i.e., with no clusters):

If $\lambda = 4$: $Power = \Pr(\hat{F} \geq 2.22 - 0.93) = \Pr(\hat{F} \geq 1.29) = 63.9\%$

If $\lambda = 8$: $Power = \Pr(\hat{F} \geq 2.22 - 1.48) = \Pr(\hat{F} \geq 0.74) = 96.3\%$

If $\lambda = 10$: $Power = \Pr(\hat{F} \geq 2.22 - 1.78) = \Pr(\hat{F} \geq 0.44) = 99.6\%$

This means that a non-centrality parameter of 4 would result in a low power. In this scenario, the probability of observing a difference and attributing it to chance is 36.1% (i.e., there is an important risk of not detecting the effect if there were one). However, this is a pessimistic scenario, where only one coefficient is barely significant at the 95% (t=2) and all the other coefficients are exactly zero. If they were higher than zero, then the non-centrality parameter would increase.

However, in a cluster-randomized trial it is necessary to include an estimate of the intra-cluster correlation. When one attempts to include the intra-cluster correlation in the same fashion as in section 1, the statistic is not an F distribution anymore because the denominator is not a chi-squared divided by its degrees of freedom[13].

$$\frac{(Rb-r)'[R(X'X)^{-1}R']^{-1}(Rb-r)/k}{[f(ICC)]e'e/(n-k-1)} \ldots(12)$$

To allow for the inclusion of the intra-cluster correlation we will test the coefficient of the treatment in a linear regression of income on the treatment and a constant only. This is only a preliminary approach, and a definite answer can only be given after careful study of the statistical properties of the distribution of (12).

Murray (1980) shows that with *many* degrees of freedom, $t = 2.80$ guarantees a power of 80%. Hence, we need to look for the conditions that would warrantee obtaining a t value of at least that magnitude.

$$\hat{t} = \frac{\hat{\beta}}{s.e.(\hat{\beta})} \geq 2.80 \ldots(13)$$

Hence, it is required that

$$\hat{\beta} \geq 2.80 s.e.(\hat{\beta}) \ldots(14)$$

[13] This statement can be easily proved by means of Slutsky's Theorem.

It can be shown that $\sigma_{\varepsilon}^2 = \sigma_{Y|T}^2$, and also that if Y|T~N, then $\sigma_{Y|T}^2 = \sigma_Y^2(1-\rho_{TY}^2)$. Hence, the variance of the OLS estimator of β is given by:

$$\text{var}(\hat{\beta}) = \frac{\sigma_{\varepsilon}^2}{\sum(T_i - \bar{T})^2} = \frac{\sigma_Y^2(1-\rho_{TY}^2)}{n\sigma_T^2} \quad \ldots(15)$$

In a similar fashion as in section 1, we start by asserting:

$$(t_{\alpha/2} + t_{\beta}) = \frac{\hat{\beta}}{\sqrt{\dfrac{\sigma_Y^2(1-\rho_{TY}^2)}{n\sigma_T^2}}} \quad \ldots(16)$$

Taking into account the intracluster correlation:

$$(t_{\alpha/2} + t_{\beta}) = \frac{\hat{\beta}}{\sqrt{\dfrac{\sigma_Y^2[1+(c-1)ICC](1-\rho_{TY}^2)}{n\sigma_T^2}}} \quad \ldots(17)$$

, where c is the number of clusters. The estimates for σ_T^2, σ_Y^2, ρ_{TY}^2, and *ICC* for each segment are obtained from the survey and are presented in the following table

Table A.5 Point Estimates of the Parameters in Equation 17[1]

Parameter	Point Estimate
σ_T^2	213.84
σ_Y^2	45,727.81
ρ_{TY}^2	0.00018
ICC	0.09
β^2	1.35

[1] The number of clusters is given by the results of Scenario 4 with 45 observations per cluster in Table A3.

Table A6. Sample Size Requirement for Continuous Treatment Approach under Scenario 3

	ICC	Sample size [1]	Implied t-value [2]
Highway sections			
1-2	0.00	647	3.67
3	0.01	432	2.82
4	0.03	674	2.89
5-7	0.13	2022	2.11
Highway as a whole	0.09	3775	3.87

[1] We input the sample size that arises from table 1 for each section except for sections 5-7, where we multiply times 3 the value of 675 obs.
[2] A t-value of 2.8 or greater guarantees power of at least 80%

[3] Assumptions:		
	ICC	0.09
	$ó^2_T$	213.84
	$ó^2_Y$	45,727.81
	$ß^2$	4
	$ρ^2_{TY}$	0.00018

Where
T: Treatment (reduction in time to market)
Y: outcome (household income)
ß: coefficient in the OLS regression of Y on T and a constant

The ICC is unusually high in sections 5 to 7 of the highway. This unusually high estimate can be due to sampling issues. To account for this possibility, we pool section 5, 6 and 7 together.

Under the assumptions of scenario 4, one would expect the sample sizes to be smaller. However, these sample sizes already are smaller than the requirement of scenario 4 for the sample size, so we can safely assume that a survey that complies with the sample requirements of scenario 4 in the discrete case would also comply with the sample requirements in the continuous case. This means that we do not need to add more assumptions about the variance of the outcome variable and its correlation through time and within individuals. One must also keep in mind that these tests focus on only one parameter, and if a more complex functional form were used, the chances of detecting an effect will be larger (the effect may not take a linear form).

4. Summary and Recommendations

The main assumptions are Type I and II error rates of 5% and 20% respectively and a change in incomes of at least 20%. If the treatment is considered discrete, we recommend scenario 4. These sample sizes are compatible with the requirements for tests with the treatment as a continuous variable.

Table A.8
Sample size requirements

Area	Required Observations
Highway section	
T1	233
T2	414
T3	432
T4	674
T5	674
T6	674
T7	674
Total Observations	3,776

These recommendations are valid if all the assumptions in the scenarios provided are met and if all the parameters obtained from NHS and the census are close to the parameters that will be obtained from the survey. Although the scenarios are plausible, they are not certain.

It is important to notice that the higher number of observations in segments T4 through T7 owes to the higher estimated intra-cluster correlation in the last 3 segments.

It must be kept in mind that any changes in the timing of the highway's sections (or in their length) may change the recommended sample sizes. In other words, if part of section T4 were passed to section T5, the recommended sample size will most likely be different. Also, changes in the start and end dates of the works in each section may impose further difficulties and may require a revision of the proposed sample sizes.

Appendix References

Behrman, J., Y. Chen and P. Todd (2004) "Evaluating Preschool Programs When Length of Exposure to the Program Varies: A Nonparametric Approach" *Review of Economics and Statistics* 86: 108-132.

Hirano, K., and G. Imbens (2004) "The Propensity Score with Continuous Treatments" in A. Gelman and X. Meng (eds.), *Applied Bayesian Modeling and Causal Inference from Incomlpete-Data Perspectives*, Wiley.

Imai, K, and D. van Dyk (2004) "Causal Inference with General Treatment Regimes: Generalizing the Propensity Score" *Journal of the American Statistical Association* 99: 854-866.

Imbens, G. (2000) "The Role of the Propensity Score in Estimating Dose-Response Functions", *Biometrika* 87: 706-710.

Murray, D. (1998) *Design and Analysis of Group-Randomized Trials.* New York and Oxford: Oxford University Press.

Appendix 2

Calculation of the Number of Clusters per Condition for each of the Four Scenarios for the Variance

Scenario 1:
Double the NHS variance[14]

$$g = \frac{2(1+(m-1)I\hat{C}C)(t_{\alpha/2}+t_{\beta})^2}{m\hat{\Delta}^2}(2\hat{\sigma}_y^2)$$

Scenario 2:
Reduce the doubled NHS variance by 10%

$$g = \frac{2(1+(m-1)I\hat{C}C)(t_{\alpha/2}+t_{\beta})^2}{m\hat{\Delta}^2}(0.9)(2\hat{\sigma}_y^2)$$

Scenario 3:
Use the NHS variance

$$g = \frac{2(1+(m-1)I\hat{C}C)(t_{\alpha/2}+t_{\beta})^2}{m\hat{\Delta}^2}(\hat{\sigma}_y^2)$$

Scenario 4:
Reduce the NHS variance by 10%, and assume between-period correlation of -0.5 and inter-period correlation of 0.5

$$g = \frac{2(1+(m-1)I\hat{C}C)(t_{\alpha/2}+t_{\beta})^2}{m\hat{\Delta}^2}(0.9)\hat{\sigma}_y^2$$

Using equations (2) and (3), we may express $\hat{\sigma}_y^2$ as

[14] The degrees of freedom for all the t statistics in every scenario are given by df = 2(g-1)

$$\hat{\sigma}_y^2 = \frac{mg}{2(1+(m-1)ICC)} 4 \left[\frac{\hat{\sigma}_m^2 (1-\hat{r}_{yy(m)}) + m\hat{\sigma}_g^2 (1-\hat{r}_{yy(g)})}{mg} \right]$$

With this,

$$g = \frac{2(1+(m-1)\hat{ICC})(t_{\alpha/2}+t_\beta)^2}{m\hat{\Delta}^2}(0.9)\frac{mg}{2(1+(m-1)ICC)} 4 \left[\frac{\hat{\sigma}_m^2 (1-\hat{r}_{yy(m)}) + m\hat{\sigma}_g^2 (1-\hat{r}_{yy(g)})}{mg} \right]$$

Cancelling out and solving for g we finally get

$$g = \frac{3.6(t_{\alpha/2}+t_\beta)^2}{m\hat{\Delta}^2}\left[\hat{\sigma}_m^2 (1-\hat{r}_{yy(m)}) + m\hat{\sigma}_g^2 (1-\hat{r}_{yy(g)})\right]$$

This way we can include the assumptions about the within- and between-period correlations (-0.5 and 0.5, respectively). With this, the final formula for the number of clusters per condition in scenario 4 is

$$g = \frac{3.6(t_{\alpha/2}+t_\beta)^2}{m\hat{\Delta}^2}\left[\hat{\sigma}_m^2 (1-(-0.5)) + m\hat{\sigma}_g^2 (1-0.5)\right]$$

Appendix 3

The χ_k^2 Distribution

If $\hat{\beta}_i \sim N\left(\beta_i, \sigma_i^2\right)\ \forall i = 1,\ldots,k$, then the random variable $\sum_{i=1}^{k}\left(\frac{\hat{\beta}_i - \beta}{\sigma_i}\right)^2 \sim \chi_k^2$.

The Non-central χ_k^2 Distribution

The non-central χ_k^2 distribution is a generalization of the χ_k^2 distribution. If $\hat{\beta}_i \sim N\left(\beta_i, \sigma_i^2\right)\ \forall i = 1,\ldots,k$, then the random variable $\sum_{i=1}^{k}\left(\frac{\hat{\beta}_i}{\sigma_i}\right)^2 \sim noncentral\chi_k^2(\lambda)$, where $noncentral\chi_k^2(\lambda)$ is the non-central chi-squared distribution with non-centrality parameter λ defined as $\lambda = \sum_{i=1}^{k}\left(\frac{\beta_i}{\sigma_i}\right)^2$.

The Non-central *F* Distribution

If $\sum_{i=1}^{k_1}\left(\frac{X_i}{\sigma_i^X}\right)^2 \sim noncentral\chi_{k_1}^2(\lambda)$, $\sum_{i=1}^{k_2}\left(\frac{Y_i - \mu_i^Y}{\sigma_i^Y}\right)^2 \sim \chi_{k_2}^2$, and X_i and Y_i are statistically independent, then

$$\frac{\sum_{i=1}^{k_1}\left(\frac{X_i}{\sigma_i^X}\right)^2 \Big/ k_1}{\sum_{i=1}^{k_2}\left(\frac{Y_i - \mu_i^Y}{\sigma_i^Y}\right)^2 \Big/ k_2} \sim noncentralF(k_1, k_2, \lambda),$$

where $noncentralF(k_1, k_2, \lambda)$ is the non-central F distribution with non-centrality parameter λ defined as $\lambda = \sum_{i=1}^{k}\left(\frac{\mu_i^X}{\sigma_i^X}\right)^2$.

Appendix 4
Summary Tables for Other Outcome Variables: Number of Clusters Per Condition (discrete treatment)[15]

	Scenario					Scenario			
m=25	1	2	3	4		1	2	3	4
T2-T1					T2-T3				
total income	20	18	10	10	total income	68	61	34	17
wage agr inc	21	19	11	4	wage agr inc	22	19	11	8
non wage ag inc	35	31	17	16	non wage ag inc	37	33	18	9
wage non ag inc	7	7	4	3	wage non ag inc	10	9	5	4
non wage non ag inc	11	10	5	5	non wage non ag inc	33	30	17	8
total hours hrs	21	19	10	10	total hours hrs	23	20	11	10
wage agr inc hrs	188	169	94	29	wage agr inc hrs	91	82	45	42
wage non ag hrs	814	733	407	108	wage non ag hrs	195	175	97	29
non wage ag hrs	57	51	28	21	non wage ag hrs	248	223	124	32
non wage non ag hrs	317	286	159	147	non wage non ag hrs	150	135	75	69
T3-T4					T3-T5				
total income	88	79	44	18	total income	112	101	56	28
wage agr inc	33	30	17	10	wage agr inc	64	57	32	15
non wage ag inc	66	59	33	11	non wage ag inc	76	68	38	16
wage non ag inc	12	11	6	6	wage non ag inc	38	34	19	7
non wage non ag inc	49	44	24	9	non wage non ag inc	44	40	22	7
total hours hrs	22	20	11	9	total hours hrs	49	44	24	11
wage agr inc hrs	81	73	40	38	wage agr inc hrs	219	197	110	41
wage non ag hrs	167	150	83	36	wage non ag hrs	341	307	171	62
non wage ag hrs	183	165	92	25	non wage ag hrs	116	105	58	20
non wage non ag hrs	164	147	82	74	non wage non ag hrs	180	162	90	63

[15] These tables make the same assumptions as Table A.3.

Appendix 5
Summary Tables for Other Outcome Variables: Number of Clusters Per Condition (discrete treatment)[16]

	Scenario					Scenario			
m=35	1	2	3	4		1	2	3	4
T2-T1					T2-T3				
total income	14	13	7	7	total income	58	52	29	12
wage agr inc	19	17	10	3	wage agr inc	16	14	8	6
non wage ag inc	24	22	12	11	non wage ag inc	32	28	16	6
wage non ag inc	5	5	3	2	wage non ag inc	7	6	3	3
non wage non ag inc	8	7	4	4	non wage non ag inc	28	26	14	6
total hours hrs	15	13	7	7	total hours hrs	16	15	8	7
wage agr inc hrs	173	156	86	20	wage agr inc hrs	64	57	32	29
wage non ag hrs	760	684	380	76	wage non ag hrs	178	161	89	21
non wage ag hrs	43	38	21	15	non wage ag hrs	231	208	115	22
non wage non ag hrs	223	200	111	103	non wage non ag hrs	105	95	53	49
T3-T4					T3-T5				
total income	77	70	39	12	total income	94	85	47	19
wage agr inc	27	24	14	7	wage agr inc	55	49	27	10
non wage ag inc	59	53	30	8	non wage ag inc	66	60	33	11
wage non ag inc	9	8	4	4	wage non ag inc	34	30	17	5
non wage non ag inc	43	39	22	6	non wage non ag inc	40	36	20	5
total hours hrs	17	15	8	6	total hours hrs	42	38	21	8
wage agr inc hrs	57	51	28	26	wage agr inc hrs	195	176	98	28
wage non ag hrs	145	130	72	25	wage non ag hrs	305	274	152	43
non wage ag hrs	169	152	85	18	non wage ag hrs	105	94	52	14
non wage non ag hrs	115	103	57	52	non wage non ag hrs	139	125	70	44

[16] These tables make the same assumptions as Table A.3.

Appendix 6
Summary Tables for Other Outcome Variables: Number of Clusters Per Condition (discrete treatment)[17]

	Scenario					Scenario			
m=45	1	2	3	4		1	2	3	4
T2-T1					T2-T3				
total income	11	10	6	5	total income	52	47	26	9
wage agr inc	18	16	9	2	wage agr inc	13	12	7	5
non wage ag inc	19	17	9	9	non wage ag inc	29	26	14	5
wage non ag inc	4	4	2	2	wage non ag inc	5	5	3	2
non wage non ag inc	6	5	3	3	non wage non ag inc	26	23	13	4
total hours hrs	11	10	6	5	total hours hrs	13	11	6	5
wage agr inc hrs	165	148	82	16	wage agr inc hrs	49	44	24	23
wage non ag hrs	731	658	366	59	wage non ag hrs	169	152	85	16
non wage ag hrs	35	32	18	12	non wage ag hrs	221	199	111	17
non wage non ag hrs	172	154	86	79	non wage non ag hrs	81	73	41	38
T3-T4					T3-T5				
total income	72	65	36	10	total income	85	77	43	15
wage agr inc	24	21	12	5	wage agr inc	50	45	25	8
non wage ag inc	56	50	28	6	non wage ag inc	61	55	31	9
wage non ag inc	7	6	3	3	wage non ag inc	31	28	16	4
non wage non ag inc	41	37	20	5	non wage non ag inc	38	34	19	4
total hours hrs	14	12	7	5	total hours hrs	39	35	19	6
wage agr inc hrs	44	39	22	20	wage agr inc hrs	182	164	91	22
wage non ag hrs	133	120	66	19	wage non ag hrs	285	256	142	34
non wage ag hrs	162	145	81	14	non wage ag hrs	99	89	49	11
non wage non ag hrs	89	80	44	40	non wage non ag hrs	117	105	59	34

[17] These tables make the same assumptions as Table A.3.

Appendix 7
Summary Tables for Other Outcome Variables: Total Sample Size (discrete treatment)[18]

	Scenario					Scenario			
m=25	1	2	3	4		1	2	3	4
T2-T1					T2-T3				
total income	1,021	919	510	478	total income	3,405	3,064	1,702	851
wage agr inc	1,074	967	537	190	wage agr inc	1,081	972	540	423
non wage ag inc	1,735	1,562	868	810	non wage ag inc	1,841	1,657	921	431
wage non ag inc	372	334	186	175	wage non ag inc	484	435	242	222
non wage non ag inc	541	487	271	251	non wage non ag inc	1,667	1,500	834	396
total hours hrs	1,034	931	517	479	total hours hrs	1,133	1,020	567	487
wage agr inc hrs	9,415	8,473	4,707	1,442	wage agr inc hrs	4,529	4,076	2,265	2,091
wage non ag hrs	40,708	36,637	20,354	5,416	wage non ag hrs	9,746	8,771	4,873	1,474
non wage ag hrs	2,826	2,543	1,413	1,069	non wage ag hrs	12,376	11,139	6,188	1,586
non wage non ag hrs	15,864	14,277	7,932	7,326	non wage non ag hrs	7,510	6,759	3,755	3,469
T3-T4					T3-T5				
total income	4,407	3,966	2,203	888	total income	5,586	5,027	2,793	1,385
wage agr inc	1,660	1,494	830	478	wage agr inc	3,183	2,865	1,591	737
non wage ag inc	3,280	2,952	1,640	547	non wage ag inc	3,790	3,411	1,895	788
wage non ag inc	608	547	304	277	wage non ag inc	1,898	1,708	949	353
non wage non ag inc	2,426	2,183	1,213	430	non wage non ag inc	2,214	1,992	1,107	374
total hours hrs	1,120	1,008	560	431	total hours hrs	2,449	2,204	1,225	535
wage agr inc hrs	4,034	3,631	2,017	1,887	wage agr inc hrs	10,965	9,869	5,483	2,027
wage non ag hrs	8,329	7,496	4,165	1,796	wage non ag hrs	17,059	15,353	8,530	3,098
non wage ag hrs	9,158	8,242	4,579	1,271	non wage ag hrs	5,814	5,233	2,907	979
non wage non ag hrs	8,186	7,368	4,093	3,700	non wage non ag hrs	9,000	8,100	4,500	3,144

[18] These tables make the same assumptions as Table A.3.

Appendix 8
Summary Tables for Other Outcome Variables: Total Sample Size (discrete treatment)[19]

	Scenario					Scenario			
m=35	1	2	3	4		1	2	3	4
T2-T1					T2-T3				
total income	1,003	903	502	470	total income	4,026	3,623	2,013	836
wage agr inc	1,355	1,219	677	187	wage agr inc	1,125	1,012	562	416
non wage ag inc	1,706	1,535	853	797	non wage ag inc	2,207	1,986	1,103	423
wage non ag inc	365	329	183	172	wage non ag inc	475	428	238	218
non wage non ag inc	532	479	266	246	non wage non ag inc	1,992	1,792	996	390
total hours hrs	1,017	915	508	470	total hours hrs	1,137	1,023	569	478
wage agr inc hrs	12,109	10,898	6,054	1,417	wage agr inc hrs	4,451	4,006	2,226	2,055
wage non ag hrs	53,228	47,905	26,614	5,323	wage non ag hrs	12,485	11,237	6,243	1,449
non wage ag hrs	2,983	2,685	1,492	1,050	non wage ag hrs	16,142	14,528	8,071	1,559
non wage non ag hrs	15,591	14,032	7,795	7,200	non wage non ag hrs	7,380	6,642	3,690	3,409
T3-T4					T3-T5				
total income	5,421	4,879	2,711	873	total income	6,612	5,951	3,306	1,361
wage agr inc	1,899	1,709	950	470	wage agr inc	3,820	3,438	1,910	724
non wage ag inc	4,149	3,734	2,075	537	non wage ag inc	4,638	4,174	2,319	774
wage non ag inc	598	538	299	272	wage non ag inc	2,363	2,127	1,182	347
non wage non ag inc	3,043	2,739	1,521	423	non wage non ag inc	2,794	2,515	1,397	368
total hours hrs	1,174	1,056	587	423	total hours hrs	2,971	2,674	1,486	526
wage agr inc hrs	3,965	3,568	1,982	1,854	wage agr inc hrs	13,667	12,301	6,834	1,992
wage non ag hrs	10,130	9,117	5,065	1,765	wage non ag hrs	21,318	19,186	10,659	3,045
non wage ag hrs	11,841	10,657	5,920	1,249	non wage ag hrs	7,342	6,608	3,671	962
non wage non ag hrs	8,046	7,241	4,023	3,636	non wage non ag hrs	9,745	8,771	4,873	3,090

[19] These tables make the same assumptions as Table A.3.

Appendix 9
Summary Tables for Other Outcome Variables: Total Sample Size (discrete treatment)[20]

	Scenario					Scenario			
m=45	1	2	3	4		1	2	3	4
T2-T1					T2-T3				
total income	994	894	497	466	total income	4,661	4,195	2,331	828
wage agr inc	1,639	1,475	819	185	wage agr inc	1,176	1,059	588	412
non wage ag inc	1,690	1,521	845	789	non wage ag inc	2,579	2,322	1,290	419
wage non ag inc	362	326	181	170	wage non ag inc	471	424	235	216
non wage non ag inc	527	474	263	244	non wage non ag inc	2,323	2,090	1,161	386
total hours hrs	1,007	906	504	466	total hours hrs	1,150	1,035	575	474
wage agr inc hrs	14,825	13,342	7,412	1,404	wage agr inc hrs	4,410	3,969	2,205	2,036
wage non ag hrs	65,826	59,243	32,913	5,273	wage non ag hrs	15,248	13,724	7,624	1,436
non wage ag hrs	3,160	2,844	1,580	1,041	non wage ag hrs	19,933	17,940	9,967	1,545
non wage non ag hrs	15,445	13,900	7,722	7,132	non wage non ag hrs	7,311	6,580	3,656	3,377
T3-T4					T3-T5				
total income	6,450	5,805	3,225	865	total income	7,662	6,896	3,831	1,348
wage agr inc	2,147	1,932	1,073	465	wage agr inc	4,469	4,022	2,234	717
non wage ag inc	5,028	4,525	2,514	532	non wage ag inc	5,498	4,949	2,749	767
wage non ag inc	592	533	296	270	wage non ag inc	2,834	2,551	1,417	344
non wage non ag inc	3,667	3,300	1,833	419	non wage non ag inc	3,381	3,042	1,690	364
total hours hrs	1,235	1,112	618	419	total hours hrs	3,502	3,152	1,751	521
wage agr inc hrs	3,928	3,535	1,964	1,837	wage agr inc hrs	16,403	14,763	8,202	1,973
wage non ag hrs	11,961	10,765	5,980	1,748	wage non ag hrs	25,627	23,065	12,814	3,016
non wage ag hrs	14,544	13,089	7,272	1,237	non wage ag hrs	8,886	7,998	4,443	953
non wage non ag hrs	7,970	7,173	3,985	3,602	non wage non ag hrs	10,546	9,492	5,273	3,061

[20] These tables make the same assumptions as Table A.3.