NORC
A national organization for research
at the University of Chicago

Design Report

Impact Evaluation Design and Implementation Services – Benin

Contract No. MCC-05-0195-CFO
Task Order No. 01

Presented to:
Millennium Challenge Corporation
Attention: Alexandra Schlegel
875 15th Street, NW
Washington, DC 20005

Presented by:
National Opinion Research Center (NORC)
at the University of Chicago
1155 East 60th Street
Chicago, IL 60637
(773) 256-6000
(773) 753-7886 (Fax)

NORC Project No. 6375 **July 31, 2006**

TABLE OF CONTENTS

MCC IMPACT EVALUATION DESIGN AND IMPLEMENTATION SERVICES – BENIN

DESIGN REPORT

A. INTRODUCTION

NORC is pleased to be working with MCC in designing and helping to implement the many evaluations that are now proposed for Benin. The evaluation design is intended to address the key questions posed by MCC about the impact of three of the activities funded under the Benin Compact: Access to Land (ATL), Access to Financial Services (AFS), and Access to Justice (ATJ). The central question concerning these activities is "What is the impact on incomes, investment employment, and attitudes about justice of these activities?"

The Task Order suggested several methodologies for evaluating the impact of each component. The evaluation team considered these methodologies in light of our understanding of the assignment, including the Benin context, MCA Benin's implementation plans for the activities, data availability, stakeholders' interests and their capacity for contributing to a rigorous impact evaluation. Much of the information contained in this report comes from an intensive two-week field visit in August 2006 by the evaluation team as well as follow-up phone calls and communication with MCA Benin and stakeholders in the following months. We summarize our conclusions for each activity below as an introduction to the chapters that follow for each activity.

Access to Land. The ATL project provides expanded formal land rights through the creation of a *Plan Foncier Rurale* (Rural Land Plan). The impact evaluation seeks to measure the ATL project activities' contribution to changing household income in the ATL project areas and the total value of additional investment in targeted rural and urban land parcels. The Task Order suggested an experimental approach to the impact evaluation, using random assignment of the 600 target villages into treatment and control groups of 300 each. We also sought to take advantage of MCC's support of the EMICoV national household survey by INSAE (*Institut National de la Statistique et de l'Analyse Economique*) that will collect data on many of the factors of concern to the impact evaluation.

Given that the EMICoV represents a powerful source of data in which MCC has invested heavily, the proposed design seeks to reconcile an experimental approach with the national random sample of households embodied in the EMICoV survey. Thus, rather than randomly assigning villages selected to participate in the ATL project, we have proposed using the villages embodied in EMICoV sample as the basis for creating treatment and control groups. The ATL implementing agency, ProCGN, will use villages in the EMICo sample as the target villages for the project during the period 2007-2009 (the time it will take ProCGN to process the eligible villages in the EMICoV sample through the activity); those villages qualifying for the project will then be assigned to either treatment or control groups so that half the villages (representing half of the eligible EMICoV sample) will be randomly assigned to each group. The EMICoV survey and its subsequent follow-up rounds, supplemented with a small survey carried out when each village is assessed and assigned to treatment/control to assess conditions at entry into the project, will then be used to evaluate the impact of the project based on differences between initial and final indicators between the treatment and control groups.[1]

Access to Financial Services. The AFS project is directed at poverty reduction through the creation of jobs and increased incomes as a result of expanded activity by micro-, small-, and medium-sized enterprises (MSMEs). The activity to be evaluated is the Challenge Facility, which will make grants to microfinance institutions (MFIs), business development service (BDS) and other institutional strengthening providers, and other MSME organizations (such as producer associations or cooperatives) for activities that address AFS goals. The impact evaluation will seek to measure these results through the level of profits, wages, and number of jobs created in MSMEs and changes in the access to and level of credit extended to MSMEs. The Task Order notes that the nature of these grant-funded activities could be wide-ranging and that the number of institutions receiving these grants could be relatively small, making a randomization approach that yields statistically significant results potentially infeasible.

Because the Challenge Facility will not solicit its first grant applications until later in 2007, our suggested design approach outlines alternatives that can be used depending on whether or not the

[1] A similar approach is planned for ATL activities in urban areas, but the feasibility of using randomization will depend on being able to link the EMICoV sample and urban *quartiers* targeted for the activity. This condition can only be verified once the implementation of the urban ATL has been developed further. If a randomization approach is not feasible, matching of treatment and comparison groups (using propensity scores) can be used to evaluate activity impacts.

circumstances of the grant programs to be evaluated allow a randomized approach. Where the number of participating MSMEs is large enough and the activity being funded by the grant can be randomly assigned to participating MSMEs—for example, training offered to an existing client base where trainees can be randomly selected—a randomized approach can be used and outcomes comparing treatment and control groups utilized.[2] If randomization is not possible—for example, where improved services will benefit all of a grantee's clients—then a comparison group will be necessary in order to measure the impact of the grantee activity on its clients.

Where randomization is not possible, it will be necessary to construct a comparison group using propensity score matching (PSM). To do so requires a source of data with sufficient observations and data elements that will allow efficient matching with the treatment group, from which data can be obtained as a condition of their participating in the program. Potential sources of data for matching and outcomes of the comparison group include both the EMICoV survey (for informal sector enterprises) and a survey of business enterprises to be conducted by the *Direction du Développement Industriel* (Directorate of Industrial Development, or DDI) at the Ministry of Industry. The DDI survey is planned to be conducted in the second quarter of 2007, so coordination with DDI will be necessary in early 2007 to ensure compatibility of the survey with evaluation purposes.

Access to Justice. The ATJ project [SUMMARY NEEDED]

Data Sources. A rigorous impact evaluation of the MCA Benin projects as described above requires more than simply collecting data pre- and post-treatment. Rather, it will be highly preferable to collect time-series data because the effects of the treatments are likely to change over time. For example, in the case of the ATL project, the impact of the treatment on households (who are primarily engaged in agriculture) is likely to be affected by changes in the agricultural production context (rainfall, weather, crop prices) so that comparing only two points in time (e.g., baseline and post-treatment) may dramatically over- or underestimate results. In addition, it is expected that some

[2] We have generally assumed that because of the small number of potential grantees, the evaluation unit of analysis will be the MSME clients of the grantee. However, it is possible to extend the proposed evaluation design to grantee organizations if the appropriate circumstances for randomization obtain—i.e., if sufficient qualifying grant applications are received (from sufficiently similar organizations or from a large enough network of organizations) from which treatment and control groups can be constructed by random assignment that will allow statistical inferences of sufficient power to be drawn.

effects will change over the course of the treatment; for example, as other MSMEs see the results of projects funded by the Challenge Facility or as the courts and ADR system begin to function more effectively.

We plan to rely primarily on existing data collection mechanisms (with some enhancements for the purpose of the evaluation) by INSAE, DDI, and other Beninese government agencies. The main data sources identified are shown in the box at right.

We also propose to implement the use of Geographic Information Systems (GIS) as an important tool to support the evaluation of all three MCA Benin projects.[3] The GIS will be used specific purposes beyond simply as a spatial display tool of compact elements. These purposes are outlined in the box below. A more detailed technical discussion of the use of GIS in impact evaluation and availability of GIS data for Benin is provided in Annex A of the report.[4]

Available Data Sources

- ***EMICoV Household Survey.*** The EMICoV survey is a national household sample survey that is being conducted for the first time in 2006. The survey collects comprehensive data on household income, expenditure, employment, economic activity, and investment. The EMICoV sample covers approximately 10% of the ZDs (zones dénombrement, or census tracts) in the country.
- ***DDI Survey of Business Establishments.*** This survey was last carried out in 2000, but has been postponed since then for budgetary reasons; it is now scheduled to be conducted in 2007. The DDI survey in the past has used as its sampling frame a census of businesses also carried out by DDI. In 2007, DDI has decided to combine the survey with its regular survey on the economic climate and at the same time update its full census of enterprises in all the major cities and towns of Benin.
- ***Ministry of Justice Satisfaction Survey.*** This survey was last carried out in 2005 and will continue to be administered by the Ministry of Justice every two years. Currently, a quality control test is being conducted to verify the reliability of data that was collected in 2005. Additional data on the survey (sample design, survey questionnaire content, and scheduling of future rounds of data collection) needs to be obtained.
- ***Court Case Records.*** Court case files are currently maintained by the courts in archives and in the offices of court clerks. The case records contain data needed to assess changes in the time required to settle cases, the types of cases, and data on plaintiffs and defendants.

[3] This GIS will also coordinate closely with and support the MCC Benin land cadastre program, including the integration and sharing of digital spatial land data.

[4] Annex A to the Access to Land design subreport provides additional detail on the use of GIS, an approach we see of general value to MCC in its evaluation work. Please note that we propose that the GIS database will be turned over to MCA-Benin and/or to Benin government agencies as a form of capacity building at the termination of the evaluation. Also, beginning immediately and during the evaluation process, we propose that as a form of capacity building that NORC collaborate with MCA-Benin or Benin agencies during the GIS database construction process. The on-going process of building the Benin GIS database (which will contain geo-locations of all Benin political administrative entities (villages, ZDs, communes, quartiers, etc.), road and infrastructure networks, and environmental data such as soil quality, rainfall variation, topography and land cover) can be shared as it is constructed with MCA Benin or Benin government ad this could help to support integration of GIS technology in the Benin government as well as improve the power and quality of the evaluation.

Use of GIS in MCC Benin Impact Evaluation

- ***To Qualify Observations.*** We will use the GIS to qualify Benin observations (villages, ZDs, households, etc.) with physiographic and spatial access variables. This will increase the power of our sampling in certain cases, through delineation of control and treatment groups, and by bringing into our statistical analyses important variables that would otherwise not be considered.
- ***To Control for the Effects of Other Projects.*** The GIS can be used to calculate measures of accessibility to the location of other projects (for example other on-going infrastructure improvements). This in turn provides a method for statistically controlling for the affects of these projects, to disentangle their effects from MCC program effects. It is difficult to do this rigorously without the use of GIS for certain types of interventions.
- ***To Predict Future Hypothetical Impacts and Outcomes.*** In conjunction with previously measured outcomes and statistical impact models in Benin, the GIS can then be used as a prediction tool to predict the likely influence or outcomes of future projects, or future investment, including spatial display of those outcomes. Using spatial or physiographic variables, it can predict the likely impact of a program in a specific geographic area as a function of the geographic characteristics of that area.
- ***In Conjunction with Statistical Analysis.*** GIS can be used in conjunction with numerous statistical techniques, such as PSM, etc. Variables calculated by the GIS can be exported directly to statistical processing software.
- ***Coordination with the MCC Benin Cadastre Program.*** The GIS will coordinate and support the on-going MCC Benin land cadastre project, and vice-versa. The Benin cadastre will be generating digital spatial GIS files in any case, and they can be imported directly into the evaluation GIS and vice-versa.
- ***As a Spatial Display Tool.*** The GIS can be used in powerful ways for MCC Benin programs, specifically to provide spatial display of MCC programs, or the location and extent of the influence of these programs. Access to all spatial displays and data (including viewing spatial data) can be setup to be done remotely through a web page or internet connection. Such displays could be considered when designing future MCC programs.
- ***As a Central Database.*** The GIS is a database, and as such it can store and integrate both spatial and non-spatial data, serve as a central repository of this data, provide querying across both spatial and non-spatial data types, link tabular data to geo-locations, and it can be accessible remotely through the web or through a internet connection.

In the remainder of the report we proceed as follows: After this introduction (Section A), we begin in Section B with some context setting about the nature of the Benin geography, its political and topological regions, roads, population centers, and villages. Next, in Sections C-E, we examine each of the MCA Benin projects in turn.

For each project, we set out our understanding of MCC's evaluation goals and the challenges that need to be addressed to complete a rigorous evaluation. How we address these challenges come next, with our approach and methodology being selected with particular attention to available data sources and coordination with stakeholders and project implementers. We also highlight how GIS can be used to support and enhance the impact evaluation for each project. Each project section closes with discussion of remaining issues for which additional information is required and/or decisions will need to be made by MCC and MCA Benin. Finally, each project section concludes with a summary of our recommendations for next steps, including a timeline for 2007 and beyond.

A preliminary budget estimate for the proposed impact evaluation is provided in Annex B of the report.

B. BENIN BACKGROUND AND GEOGRAPHIC OVERVIEW

Geography. Geographically, Benin ranges from the Bight of Benin in the south to the Niger River in the north. Most of the population lives in the southern coastal plains, where Benin's largest cities—Porto Novo (the political capital) and Cotonou (the commercial capital)—are located. The north of the country consists mostly of treed grasslands and semi-arid highlands, with the exception of the northwest of the country, where a range of mountains runs along the northwest border and into neighbouring Togo.

Figure B.1 — Benin Map & Satellite Image

Economic Overview. A significant share of the Benin economy is in subsistence agriculture, with maize (corn), beans, rice, peanuts, cashews, pineapples, cassava, yams, and other various tubers--all are grown as subsistence crops. Cotton is the major export crop, making up 40% of GDP and more than three-quarters of exports. The private commercial and agricultural sectors are the principal contributors to economic growth.

A significant share of the Benin economy is in subsistence agriculture, with maize (corn), beans, rice, peanuts, cashews, pineapples, cassava, yams, and other various tubers grown as subsistence crops. Cotton is the major export crop, making up 40% of GDP and more than three-quarters of exports, but its position has been declining over the last two to three years. The manufacturing sector is composed predominantly of light industry, focused mainly on the processing of primary products and the production of consumer goods. The service sector has grown recently, stimulated by

economic liberalization and fiscal reform. Services (especially telecommunications) and the agricultural sector are currently the principal contributors to economic growth.

Few Beninese have access to justice because of archaic laws, a lack of trained judges and courthouses, deficient administrative capacity and poor access to legal information. Conducting legal affairs is expensive and time-consuming. Further, the risk of corruption means that the justice system produces highly uncertain results. In this climate, private sector growth is hampered; and MSMEs engage mostly with known customers/suppliers, limit transactions to safe payment modes, and avoid investment, business expansion and business partnerships.[5] Moreover, the high incidence of land disputes – which in turn dampen economic growth - increases the need for more timely and less costly access to dispute resolution, such as Alternative Dispute Resolution as well as improvements in the traditional justice system.

MCA Benin Program Overview. The Benin MCA Program comprises four projects: (1) Access to Land, (2) Access to Financial Services, (3) Access to Justice, and (4) Access to Markets. The Task Order covers design and implementation of impact evaluations for the first three projects; the Access to Markets project, where no obvious candidates for legitimate comparators exist, is not included in the scope of the Task Order.

The Access to Land (ATL) project aims to support policy and legal reform for more sustainable and efficient land registration services. Currently, the title registration system is expensive, slow and complex; only 1% of households hold formal title to their land and a majority of the rural population relies on oral customary land rights. In urban areas, individuals occupy land under weak administrative permit while enterprises occupy state land by concession. Land disputes are widespread and are estimated to comprise more than 70% of all civil court cases in Benin.

Benin has a shallow financial sector that provides limited services to MSMEs, particularly those that are involved in the production of goods for Benin. The high cost or unavailability of credit and other financial services, including savings, limit the capacity of small businesses in Benin to expand production and employment, to respond to business opportunities and to manage risk. The MCC Access to Financial Services (AFS) project aims to alleviate these constraints and contribute to

[5] Description from Task Order Statement of Work.

creating a broader and deeper financial sector in order to improve the sustainability of MSMEs and, in turn, increase incomes of the poor that own, are employed by, or do business with MSMEs.

The Access to Justice (ATJ) project will support expansion of the Center of Arbitration, Mediation and Conciliation at the Chamber of Commerce and improve the network of Business Registration Centers (Guichet Unique). The project will also work to improve services of courts including capacity building and training for judges, court personnel, legal aid, and Inspection General service; creation of a legal information center; development of a public awareness campaign; and the construction of new courthouses. The ATJ project aims to improve the ability of the justice system to enforce contracts and reconcile claims through increased efficiency and improved services of courts and arbitration center, increased access to the court system, and an improved enterprise registration system.

C. ACCESS TO LAND PROJECT

C.1 Evaluation Goals

The ATL project provides expanded formal land rights through the creation of a *Plan Foncier Rurale* (Rural Land Plan). The impact evaluation seeks to measure the ATL project activities' contribution to changing household income in the ATL project areas and the total value of additional investment in targeted rural and urban land parcels. Improvements in tenure security are expected to increase investment incentives, lower transaction costs, and improve access to credit. Specifically, it is hypothesized that:

- Households will invest in making their property more productive (without fear of not being able to recoup that investment because of losing access to the land, as might be the case without secure tenure).
- Productivity should also rise to the extent that enhanced tenure security facilitates land transactions from less efficient producers to more efficient producers.[6]
- Where the owner does not currently have the capital to make the investment, the owner can pledge the land as collateral to gain access to credit to finance the investment (which also relates to the other components of the Compact dealing with finance and enforceable contracts).[7]

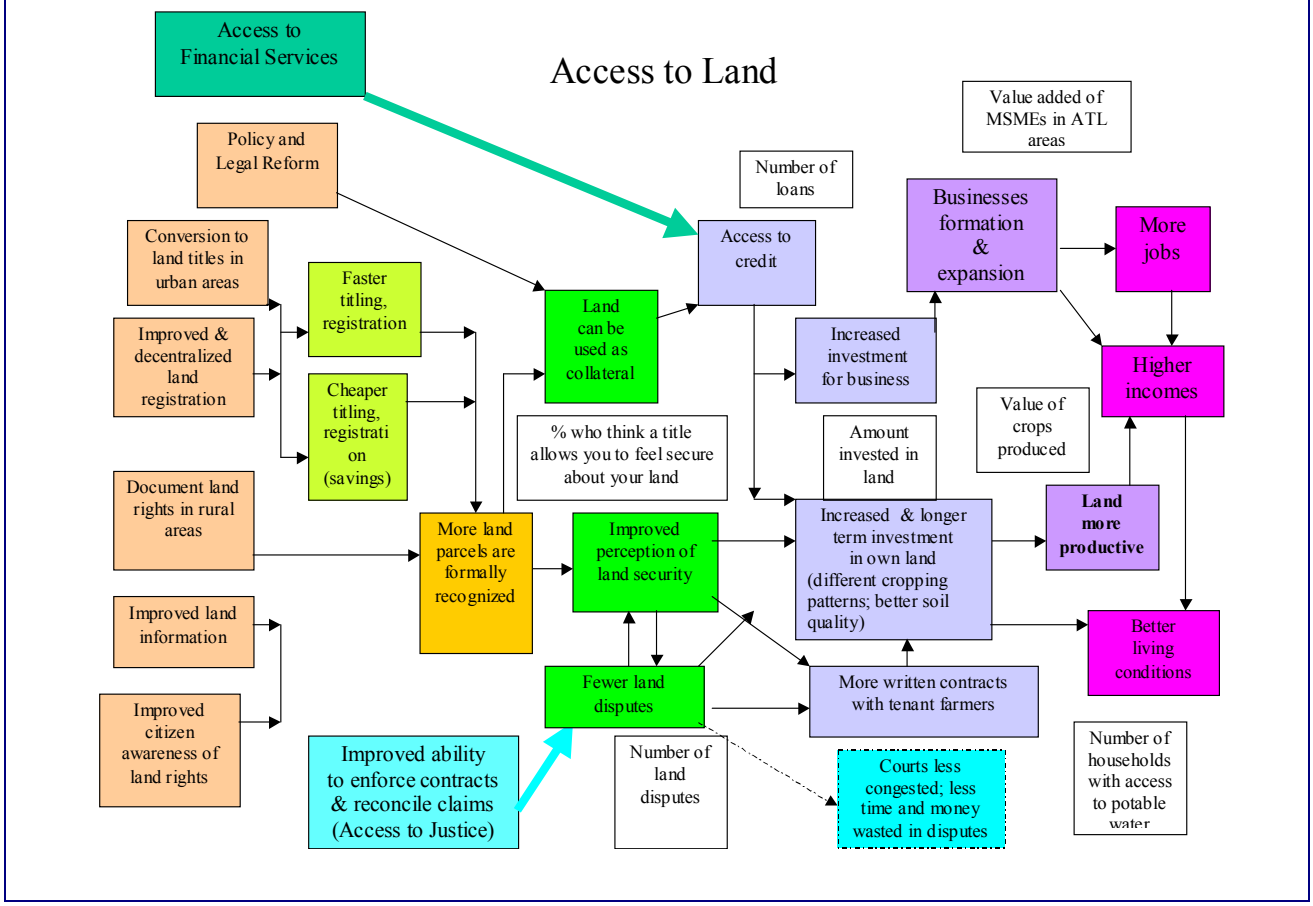
The overall expected result of making the land more productive (through investment) is higher household incomes.

The chart below illustrates the anticipated effects of the activities in the Access to Land project (depicted to the left of the chart), as greater formal recognition of land parcels leads to improved security and increased investment in land and in other enterprises (moving toward the right). The white boxes contain the principal outcomes that this evaluation will seek to assess.

[6] Deininger et al., "Tenure Security and Land-Related Investment: Evidence from Ethiopia," World Bank Policy Research Paper 2991, Washington, DC: World Bank, 2003.

[7] The design for evaluation of the Access to Finance project is presented separately from this report. Recommendations on how to link the evaluation of the two programs where they intersect are presented in Section 6 below.

Figure C.1 — ATL Outcome Sequence



C.2 Methodological Challenges

Recent discussions with MCC with respect to the ATL component indicate that an experimental design with village level randomization can be used for rural areas.[8] Based on discussions with the project implementer, ProCGN/GTZ, randomization will be possible within communes, but the objectives of the ATL project will result in the communes being placed in the treatment queue based upon the *likelihood* that they will benefit from the program. Thus, communes which have a greater likelihood of success, will receive the ATL program earlier in the implementation process. The design will need to take into account both this characteristic of the pipeline of villages into the rural ATL component and the fact that the treatment will not be applied to all villages in the impact evaluation at the same time, but over a period of approximately three years. We will look to take

[8] It is unclear yet whether a randomized approach can also be used urban areas. For the present, we assume that such an approach will also be used in urban areas, but will later describe an alternative approach if this assumption should prove to be incorrect.

advantage of the features of the new EMICoV national household survey as much as possible to address these concerns.

Rural ATL Component Village Selection

The rural activities of the ATL component of the Compact are planned to take place in 300 villages that will be selected from 42 communes in 9 departments according to criteria to be developed by MCA-Benin and approved by MCC.

MCC, MCA-Benin, and the implementing agencies (ProCGN/GTZ) have agreed upon an implementation schedule that prioritizes the EMICoV sample villages; this will allow the EMICoV data to be used as part of the ATL evaluation. MCC and MCA-Benin will work closely with the implementing agencies to determine the order in which the communes will be selected. MCC has indicated that the selection of communes will not be random, but rather will prioritize communes where key indicators suggest the program would be most beneficial.

A second key challenge for the rural ATL evaluation will be to obtain the appropriate data needed both for measuring the outcomes that are of interest (mainly household income/productivity and investment/credit information), but also to include data on possible confounding factors outside of the ATL project that will affect these measures. This is particularly important for the rural ATL component, as factors such as soil quality, rainfall, and overall accessibility (including access to markets) have a significant effect on household outcomes (with or without the ATL activities). The use of spatial data on topography, road networks, soil fertility and rainfall variation, assembled in a GIS, combined with EMICoV data (which will also contain geo-located observations), will allow for the explicit and direct integration of these variables into the impact analysis.[9]

The GIS will also be used in coordination with the MCC Benin land cadastre mapping component to bring into the analysis highly accurate and detailed digital maps of political boundaries, road networks, and when possible land ownership boundaries. These data will be useful not only for the explicit consideration of measures such as accessibility of individual villages or parcels to markets, the relative proximity of assets such as good soils or high rainfall which could increase land values (and incentives for land investment), but also for display and monitoring purposes. On-going discussions with MCC Benin land cadastre program implementers are working to ensure the coordination and integration of the Benin evaluation GIS with the MCC Benin digital cadastre system, providing synergies for both.

[9] A detailed review of the use of GIS in the evaluation and the construction of GIS indices is given in Annex A to this report.

Finally, data quality and availability are always a concern in any evaluation. The value of the EMICoV survey in meeting evaluation objectives in terms of providing representative data through its random sample must be balanced against the quality and timing of the data provided by the survey. Where the EMICoV survey is unable to provide the needed data for the evaluation, consideration must be given to alternative data sources or the collection of primary data.

C.3 Role of GIS in the Access to Land Evaluation

C.3.1 Derived Variables and GIS Access Indices

Each village that will be potentially sampled possesses different geographic, biophysical and market accessibility conditions. For example, some villages will be located on soils that are inherently richer and more agriculturally productive than the soils of other villages, some villages will receive greater annual rainfall amounts or lower intra- or inter-annual rainfall variation, and some villages will possess better accessibility to markets, infrastructure, government or health services by virtue of location and quality of road networks. Villages with superior or favorable bio-geophysical conditions or with lower access and transport costs will benefit from these unequal exogenous factors. Thus, the variation in these conditions should be considered when analyzing outcomes.

The data necessary to consider these varying exogenous factors can be obtained through compiling maps of sufficient quality on topography, soil quality, rainfall variation, climate variation, and transport networks (including road, rail or river network transport speeds and quality) from Benin government agencies or from non-government organizations operating in Benin. In some cases, country-wide maps (such as topography or land-cover) can be obtained freely on the internet. These maps can then be digitized and merged into a Geographic Information System (GIS) to calculate for each village the approximate relative soil, climate, topographic and transport network accessibility measures. These measures can then aid as controls in the selection of villages for sampling.

Without this information, this important source of variation in outcomes would not be considered. Furthermore, including the information would not necessarily require a survey, since we believe the data can be obtained from the internet or from Benin government agencies, and thus may be obtained for relatively low cost. Annex A to this report describes the available data in detail.

The spatial analysis and the digital geodatabase would also allow for the consideration and controlling of "neighborhood" effects, including the presence of spatial autocorrelation in village outcomes. This effect has been shown to be significant in numerous social science studies (often as or more important than temporal autocorrelation) but cannot be considered in a non-spatial analysis.

Finally, this spatial analysis could produce important insights: maps can often reveal interesting spatial and/or time trends quickly that might not be apparent in looking at tabular data. A GIS created for the evaluation can provide a powerful tool to assist in the evaluation and assessment of all program time-sequence and spatial elements, as the GIS acts as a central database to store all data, both spatial/geographic data but also all (non-spatial) tabular data such as survey responses, socioeconomic statistics, etc. Any socioeconomic or tabular data on villages, communes, respondents, etc., can be linked in the GIS database to geo-locations for these units, and then displayed spatially. Time-changes or time-sequences in program elements or outcomes can also be displayed on a series of highly accurate maps, and all data exported and imported in and out of standard statistical software for more rigorous statistical, panel or time-series analysis. In addition, geo-location of and spatial visualization of the data will often reveal trends and patterns that might not otherwise be apparent, and suggest possible solutions to achieve adequate country-wide spatial distribution. Annex A to the report provides detail about how a GIS can be used to enhance the evaluation.

Access Indices and GIS

For calculating access indices and other spatial analysis relating to the impact of the road, the GIS should contain data such as:

- Existing road network spatial configurations, with approximate road speeds and conditions (determined from road maps, satellite imagery, and/or survey/census information)
- Elevation data
- Data on land quality (such as soil type, rainfall)
- Stream, river, lake and reservoirs
- Other infrastructure (power networks, railways, water access systems)
- Location of other features that are important for the calculation of welfare gains to the communities (such as hospitals, major market centers, government centers, health clinics)
- Data on the location and timing of construction of major and feeder roads, to judge when their "impact" began

Once the GIS is assembled, access indices can be constructed; example approaches include:

- *For communities that are directly connected to existing major road networks:* Travel time from each community to the road calculating travel along connecting feeder roads, considering approximate road speed/road quality and other factors such as topography or traffic congestion.
- *For communities that are not connected directly to the major roads or are bordering a road:* A cost surface would be constructed for the relevant area, based on a combination of all GIS layers that provide data on factors that increase the cost/difficulty of travel (for example, steep slopes or unimproved roads). The GIS would then calculate the pathway from each unconnected community to the road, following the path of lowest travel cost, to determine the accessibility index value for that community

C.3.2 Coordination with MCC Benin Land Cadastre Program. The NORC evaluation team is currently in on-going communication with MCC personnel regarding the development of the MCC Benin cadastre program, which includes programmatic elements such as the collection of high-resolution digital spatial data to support cadastre mapping. Continuing coordination with MCC during the development of the cadastre program design should facilitate mutual synergistic support between the NORC evaluation Benin GIS and the MCC Benin cadastre program. For example, GIS physiographic data imported, cleaned and collected by NORC from CENETAL can be exported directly into the MCC cadastre GIS, and also digital spatial data on administrative or land plot boundaries, or digital airphotos, created for the cadastre can be imported directly into the NORC GIS. (See Annex A for more details).

C.4 Addressing the Challenges

C.4.1 Methodology

The agreement of the project implementers to use the villages in the EMICoV sample allows for a design that takes advantage of the randomized representative sample that underlies the EMICoV survey.

The ATL implementing agencies will be selecting the order of communes (and the corresponding villages) in which they will work according to subjective programmatic concerns. In each commune, the ATL implementers will carry out a rapid diagnostic in each of the villages in the commune that are part of the EMICoV sample.[10] The result of the diagnostic will be a decision by the ATL implementing agencies as to whether the village possesses the characteristics that would allow the land reform activities to succeed; these villages will then be categorized as "eligible" villages. From this pool of eligible villages, half will be assigned to the "treatment' group of villages where the ATL land reform activities will occur and half will be assigned to the "control" group of villages where no ATL land reform activities will take place.

[10] The EMICoV sample is actually mapped to *zones dénombremens* (ZDs) which do not necessarily correspond to villages on a 1:1 basis. Thus, the villages assessed by the ATL implementing agencies will have to be chosen so as to ensure that each ZD in the EMICoV sample is covered by assessed villages. In some cases, a ZD only covers part of a village; in this case we would expect that the entire village will be assessed. Similarly, in cases where more than one village is part of an EMICoV sample ZD, we would expect that all of those villages that are part of the ZD will be assessed. Thus, it may be the case that there are more villages being assessed in a commune than EMICoV ZDs in that commune.

The ATL implementing agencies expect, given resource limitations, that they will be able to carry out the project activities in approximately 20 villages in the 2007, 60 villages the following year, and 100 villages in subsequent years. Based on data provided by INSAE about the EMICoV sample, there appear to be 341 rural villages in the communes to be covered by the ATL project. If we assume (arbitrarily, for illustrative purposes) that approximately 20% of the rural villages do not qualify as being eligible for the project, then approximately 270 EMICoV villages would be in the universe of eligible villages. Thus, it would take approximately three years for the ATL implementing agencies to assess and begin implementation in the 135 villages in the treatment group, as shown in Table 1 below.

Table C.1 – Sequencing of ATL Rural Villages

Cohort	*Treatment Villages*	*Control Villages*	*Total Villages*
2007	20	20	40
2008	60	60	120
2009	55	55	110
TOTAL	135	135	270

Communes will be assigned to three yearly cohorts according to subjective programmatic concerns (i.e., they will not be assigned to cohorts at random). One cohort will begin the program in 2007; the other cohorts will begin in 2008 and 2009. Within each commune, eligible villages will be assigned to the treatment and control groups.[11] However, any one commune will not have a large enough sample size to permit inferences to that specific commune. Results need to be combined over communes within a cohort. If there is a large enough sample, and there likely will be in the 2008 and 2009 cohorts, we have some chance of making inferences to the population of communes in the specific cohort. To make inferences to the population of all communes, we need to make the regression adjustments for the observed differences between the cohorts and the population.

[11] While it would be statistically better to have assignment of eligible villages to the treatment and control groups done randomly, this is likely to be at odds with the programmatic desire to begin ATL activities as soon as a village has been declared eligible. Given the balanced sizes of the treatment and control groups, systematic assignment (i.e., first village assessed as eligible is assigned to treatment group, second eligible village to control group, third eligible village to treatment group, and so on) is our recommendation. This approach would fit better with programmatic needs and, we conjecture, based on our experience elsewhere, can be done with very little impact to the statistical validity of the evaluation. This conjecture is, however, just a working hypothesis; we will calibrate it further as the evaluation proceeds.

At this point, the GIS-derived variables describing the physio-geographic (soil, rainfall, land cover, topographic, percent of land that is highly fertile, percent of land that is not prime buildable, etc.) and measures of accessibility (to markets, ports, infrastructure) will improve the estimation accuracy of these regressions, because they will consider these important factors that have powerful long-term impacts on overall land value in Benin. For each of these variables, spatial data on soil, rainfall, road networks, etc. can be used to calculated *aggregated* measures for each commune (using robust geographical statistical measures designed for such spatial aggregation). These aggregated commune variables when included in the regression adjustments, will improve the ability to consider comparisons between the cohorts and the population.

As we will describe below in Section 6.2, a baseline measurement will be made at the time of treatment, and then follow-up measurements will be made at regular intervals thereafter. Preliminary project impact will be obtained as the difference between the follow-up and baseline measurements. This preliminary program impact will reflect different time periods for each of the cohorts. This disparity of periods would be of little concern if each cohort would be broadly representative of the population (i.e, if each cohort was a random sample of all communes). However, such representativeness will surely not be the case—the cohorts are likely to be very different. Because of the approach taken by the ATL project implementers, the first cohort will be more "promising" than the second, and the second more "promising" than the third. To compile a composite evaluation of the program's impact, combining all cohorts into one assessment, we will need to make adjustments to the preliminary program impacts for the ways in which the cohorts are different.

This can be accomplished by regressing the preliminary impacts on the observable characteristics of the communes. Again, the GIS-derived variables can be included in the regressions, and because they will introduce important data describing each commune that otherwise might not be available (and specifically this will be data that has a direct impact on long-term land values and investment), they will improve the estimation accuracy in the regressions of the differences between communes. We will report on program impacts after controlling for the observable differences between communes.

In carrying out the analysis of the data, the use of randomization gives us a powerful tool to account for the causes of variation in outcomes between villages from sources other that the rural ATL activities. Nonetheless, it is also important to develop as much information as possible about the covariation of outcomes with other factors in order to minimize the variance of the measure of the impacts associated with the ATL project. Thus, blocking or stratifying the sample where feasible will be done in the analysis to achieve this result.[12] The analysis will be carried out by regressions of a matrix of the outcome measures, Y, on a matrix of observed covariates, X, which will also include GIS-derived phyiogeographic variables and dummy variables for treatment and cohort. This will allow us to assess whether the treatment effect is statistically significant after controlling for all of the other observable variables.

C.4.2 Data Sources

EMICoV. We will rely primarily upon data from the EMICoV national household survey being collected by INSAE beginning in 2006. Several of the EMICoV survey modules include items necessary to conduct the impact evaluation, including those on changes in income and investment that are expected to result from the ATL activity. The *Foncier* module includes questions on land disputes and perceived security of tenure that may also be useful in analyzing the impact of the interventions. The *Microfinance* module includes data on investments in property and other areas. In meetings in Benin with INSAE and ProGCN, we determined that a number of additional questions could be added to the several EMICoV modules for Phase 4 and subsequent rounds. The literature also suggests that qualitative indicators of household perceptions of confidence in their tenure and its security are good short term proxy indicators of longer-term success (in terms of increased investment and incomes). Several EMICoV modules include items which measure confidence in institutions and overall security, but we anticipate adding questions to the survey to focus on this aspect.

Significant resources have been invested in EMICoV by both MCC and MCA-Benin, so it seems we should use it where feasible. This suggests that the likely data collection strategy will be to identify

[12] It is possible that events may occur over the course of the evaluation that undermine the assumption of randomization. In such a case, having sufficient information available to interpret the data as an observational study, through the use of propensity score matching, for example, to match treatment villages with similar comparison villages, will be invaluable. Again, GIS-derived physiogeographic and accessibility variables will be extracted for all villages, and these variables will improve the accuracy and power of the propensity score matching, thus improving the overall accuracy of the impact evaluation.

any limitations in the EMICoV modules (from the perspective of the impact evaluation) and add a few critical items to existing modules that would be administered beginning in Phase 4.

As discussed above, the timing of EMICoV baseline data collection will not be consistent across villages because ATL implementation will occur on a flow basis. Since villages will only begin the program once the diagnostic and randomization process are concluded, the timing of the EMICoV data collection will differ across villages. This different timing will have implications for data analysis. We recommend the development and implementation of a short survey conducted in each village just before they begin ATL implementation. The survey would be a subset of EMICoV items, but also potentially include a small number of critical items that are not presently in the current EMICoV modules. While this would add data collection costs to the evaluation, we estimate the short survey would only place a maximum of 30-60 minutes burden upon respondents, compared to the nearly 10 hours of survey time that the EMICoV requires in a household. In addition, since these households have been enumerated and cooperated on the EMICoV, and interviewers have been trained, the overall cost of the survey, while not insignificant, is greatly reduced. One possibility is to administer the survey as one step in the land demarcation process in the treatment villages.

GIS and Spatial Data Sources. For a detailed description of our assessment of Benin GIS data sources, please refer to Annex A. However, in sum, our field visit to Benin in August 2006 sought to determine if a sufficient quantity of high-accuracy digital spatial data existed and could be obtained. Based on our investigation, and the initial gathering of initial GIS data from CENETEL in Benin, our assessment is that sufficient data exists and will be obtainable to complete the methodology as described above. Key to the analysis will be data on village geo-locations, but these are being obtained through the use of Geographic Positioning System (GPS) units by INSAE, and will thus be highly accurate.

Meetings and discussions conducted at CENETEL indicate that they possess GIS data at the national level for Benin that includes digital road network data, as well as data on vegetation and land cover, hydrology, topography, soils and rainfall. Graphical outputs of some of the CENETAL GIS data is shown in Annex A. CENETAL also has an extensive collection of paper maps for Benin for environmental variables, and as described in Annex A these paper maps can be converted relatively inexpensively to digital GIS data for analysis. CENETAL also possesses an archive of both digital

aerial photographs and satellite images for Benin, including Landsat, Spot and Quickbird satellite imagery. This imagery can be used to extract GIS digital data for road networks, village (or even household) locations, and land cover. Evaluation team members are experts in the use of both GIS and in the processing of satellite imagery and aerial photographs for land cover extraction. In addition to CENETAL, IGN also has an archive of GIS data.

In addition to spatial data at Benin agencies, extensive GIS and satellite image data for Benin is available for free from numerous online sources (see Annex A for more details). Finally, there exists an active private market for GIS data (highly accurate geo-spatial data is needed constantly by engineering firms, highway contractors, large-scale irrigation construction, urban and water planning, etc.) and extensive and accurate GIS data for Benin is available from private sector contractors if necessary[13]. These sources collectively should be sufficient for the analysis.

6.3 Evaluation Timing and Sequencing

ProGCN/GTZ and the implementing agencies' diagnostic assessment to determine eligible villages from villages in the EMICoV sample is expected to begin in early 2007. The commencement of ATL activity is denoted by *T* in the table.[14] Based on our assumptions outlined above, it is estimated that 20 EMICoV villages will begin the program in 2007, 60 in 2008, and as many as another 55 in 2009. There will be equal numbers of control villages. Table 2 below shows the sequencing of the various rounds of EMICoV data collections. Data collections for each cohort are denoted by Y_iX_i, where *i* signifies the data collection period. *Y* denotes the data, or outcome measures, to be collected in either the EMICoV interviews or in other data collection. Collection of the covariates, (*X*, or exogenous variables) for final modeling must also be done. If the covariates are not time varying (e.g., many of the GIS variables, such as soil quality), then we would collect them in a one-time operation. Otherwise, some covariates may be collected each year as part of the EMICoV survey. The covariates are needed, of course, for both treatment and control villages.

[13] Such as the French company GeoConcept (based in Paris) that sells GIS data for Benin (see http://www.geoconcept.com/index-en.php3).

[14] If villages are randomly assigned to treatment and control ATL activity can only begin once an entire commune has received the diagnostic and all villages in the ATL pool have been assigned.

Table 2 — ATL Data Collection Sequencing				
	Time (in Months) from Baseline			
Cohorts	***0-12***	***13-24***	***25-36***	***37+***
2007				
Treatment	Y_1X_1, *T*	Y_2X_2	Y_3X_3	Y_4X_4
Control	Y_1X_1	Y_2X_2	Y_3X_3	Y_4X_4
2008				
Treatment		Y_2X_2, *T*	Y_3X_3	Y_4X_4
Control		Y_2X_2	Y_3X_3	Y_4X_4
2009				
Treatment			Y_3X_3, *T*	Y_4X_4
Control			Y_3X_3	Y_4X_4

The initial round of the EMICoV survey will provide pre-treatment baseline data on all the treatment and control villages. Additional rounds of data collection occur 1, 2, and 3 years post intervention (assuming the final round can be completed prior to the end of the task order to allow adequate time for analysis). As noted above, ideally, these post-intervention data collections could be fulfilled by the annual EMICoV updates which INASE is planning. Since it is not feasible that the start of ATL implementation coincide with the administration of EMICoV updates, the supplementary data collection at the time of entry of each ATL village is needed to provide data that will allow us to control for the timing differences of each village's entry in the treatment group where the EMICoV updates have not been carried out close enough to the time of entry into the treatment group.

Valid Estimation Techniques for Use with Randomized Design

- The before-and-after estimator compares differences in the means of the outcome variables between the treatment or control group during the periods after and before the implementation of the treatment.
- The single (or cross-sectional) difference estimator compares differences in the means of the outcome variables between the treatment and control groups during the periods after the implementation of the program
- The double differences estimator measures program impact by comparing differences in the means of the outcome between the treatment and control groups in post-treatment rounds with the differences in the means of the outcome between the treatment and control group in the pre-program round.

Once the baseline and later rounds of data are collected together with the mini-survey, the analysis can proceed using the various estimation techniques that are potentially valid for a randomized design (see box at right). Randomized assignment to treatment implies that the distribution of all the variables for treatments and control samples should be equal prior to the administration of the program. This can be checked by comparing the means of key variables in control and treatment

localities and by comparing the means of the same variables with household level data.[15] If the hypothesis that the means are equal is rejected, then there may be some bias present and double difference estimators may be preferred (assuming adequate baseline and post-treatment data is available). Double differences also require larger sample sizes to achieve the same power.

C.4.4 Possible Special Considerations for Urban ATL Activities

The implementing agencies will implement the urban ATL activities in three urban centers—Cotonou, Parakou, and Porto Novo—where formal land titles will be issued for urban properties. MCA-Benin and the implementing agencies will conduct department level diagnostics to determine the queue for communes in these urban areas. Communes in urban areas, like rural ones, will be placed in the pipeline based upon perceived likelihood that the commune is well-suited to the program. While there are differences between the program implementation in rural villages and urban *quartiers*, we assume that MCA-Benin, ProGCN/GTZ and the other implementing agencies will adhere to a similar randomization approach in both urban and rural areas. In urban areas, all *quartiers* in the EMICoV sample within a commune will receive the diagnostic and based upon their interest and diagnostic results, eligible *quartiers* will be placed in either the treatment or control group. As with the rural villages, we assume that control *quartiers* will remain in the control group until the conclusion of the evaluation.

However, this approach assumes that it will be possible to map the selected quartiers to EMICoV urban ZDs in a manner that is consistent with the programmatic requirements (which have not yet been as well-defined as those for the rural ATL activities). The Benin evaluation GIS may aid in this consideration, by allowing for the production of maps of multiple geographic scales showing the overlap between ZD boundaries and *quartiers*. This cartographic analysis could assist in the evaluation of whether or not it will be feasible to map the *quartiers* to the ZDs in a useful manner, or it may aid in suggesting alternative approaches.

If a similar approach to that used in rural areas is not feasible and randomization on the EMICoV sample urban ZDs is not possible, then the design for urban ATL activities must be adapted to reflect its non-experimental nature. If selection of the treatment *quartiers* were based purely on observable

[15] This would include spatial, GIS-derived variables, such as soil and market accessibility. Means for these variables can be calculated by the GIS just as for any other (socioeconomic) variable at any of a number of specified spatial levels of aggregation (village, ZD, commune, nationally, etc.) and then exported to a statistical processor for regression analysis.

characteristics, then a propensity-score matching (PSM) method would remove the selection bias due to differences between *quartiers* that were and were not affected by the urban ATL activity. The propensity score measures the probability that a *quartier* is selected for the activity as a function of that *quartier*'s observed pre-activity characteristics. Here exogenous GIS-derived variables including accessibility could improve the power of the PSM process by providing a rich source of additional observables for the *quartier*. While environmental variables such as soil fertility and topography may not be as significant for urban areas, nonetheless key variables such as relative accessibility to markets or roads can be calculated by the GIS for each spatial urban area (combined with road network data, which is often of superior resolution in urban areas), and these variables could be used in the propensity score process to further increase the accuracy of the treatment and comparison zone selection.

Propensity Score Matching

Propensity score matching (PSM) is useful when the aim of matching between control and treatment groups is to find the closest comparison group from a sample of communities not receiving treatment to the sample of communities receiving treatment. "Closest" is measured in terms of observable characteristics.

The main steps in matching based on propensity scores are as follows.

- First, obtain a representative sample of eligible treatment and non-treatment communities; the larger the sample of eligible comparison communities the better, to facilitate matching.
- Second, pool the two samples and estimate a probit or logit model of participation as a function of all available variables that are likely to determine participation[16].
- Third, create the predicted values of the probability of participation from the estimated regression; these are the propensity scores (one for each sampled community).
- Fourth, exclude non-treatment communities in the sample if they have a propensity score that is outside the range (typically too low) found for the treatment sample.
- Finally, for each community in the treatment sample, find the observation in the non-treatment sample that has the closest propensity score, as measured by the absolute difference in scores. This is called the "nearest neighbor."

More precise estimates can be obtained by comparing the mean of multiple nearest neighbors for each treatment observation.

If treatment and comparison *quartiers* have the same propensity scores and all characteristics relevant to assignment of treatment are captured in the propensity score (i.e., the relevant characteristics are all observable), then the difference in their outcomes yields an unbiased estimate of the intervention's impact.

However, some unobserved characteristics of the *quartier* that correlate with, say, investment outcomes might also correlate with investment placement, which can introduce bias in the estimation

[16] These variables should include ones describing physiographic conditions of impacted communities, such as soil quality, rainfall, elevation, etc., as well as measures of relative accessibility to markets, as these are likely to influence economic output and/or land investment. The GIS will allow these variables to be calculated and included for all locations.

of investment impact. As long as the pre-investment differences between the control and treated *quartiers* are the result of unobservable characteristics omitted from the propensity score that do not change over time in their impact on outcomes, then the double difference method will correct for the possible bias. Thus, matching using PSM removes the selection bias due to the observed differences between the treated and comparison communities. Double differences corrects for possible bias due to the differences in time-invariant unobserved characteristics between the two groups. The impact of the investment is the change in the outcome indicators between matched *quartiers* from the treatment and comparison groups.

C.5. Other Evaluation Design Issues

Other MCC Project Activities. Some of the villages selected for the creation of rural land plans will also be subject to other MCC interventions:

- Four rural communes will be selected for an activity that will provide parcel-based cadastre registration, where we understand that the more formal land titles (*titre foncier*) will be issued.
- Villages that will be affected by Access to Credit and Access to Justice activities.

We do not recommend trying to obtain sufficient data to try to discern the marginal effect of each additional intervention, as this would greatly increase the sample size required. Rather, we suggest including a dummy variable in the analysis for each program to see if there is a statistically significant effect on aggregate outcomes where multiple programs are in effect. Of course, to do so, we will need to ensure that the sample where the other programs are active is sufficiently large to make statistically valid conclusions.

Collected Data Quality. During the month of August, NORC staff accompanied INSAE's data collection team to a number of rural and urban sites. NORC witnessed the data effort and was satisfied that INSAE and the team are following accepted standards of field collection. INSAE has secured well educated staff and implemented a comprehensive training program that prepares them for the field interviewing. It will also be important to assess other elements in the data collection

process including the receipting, transfer, and data entry of the questionnaires to ensure that the highest quality is maintained at each of these steps.

INSAE has indicated that it will perform the necessary reliability checks of data collected in each round. This analysis and conclusions should be provided to NORC at regular intervals so that we can evaluate the quality of the EMICoV data on an on-going basis during the evaluation.

During the August field visit, it became clear that some key modules and questions were being refused by a number of respondents. We could not determine in a short visit how widespread this problem was. Critical items on informal sector activity and investment are potentially sensitive but critical to determining if resources are being invested. It appeared that this might be a greater problem in urban ZDs than in rural ones. Since respondents may not feel secure in the confidential nature of their respondents, their fear that reporting this information could result in low response rates for these key items; jeopardizing key elements of the impact evaluation. It will be important for interviewers to gain cooperation of respondents and convince them of the confidentiality of their responses. NORC can provide assistance to INSAE to help address this problem.

EMICoV Follow-up Surveys. INSAE has yet to determine which modules will be administered in subsequent rounds. It is essential to the success of the ATL evaluation that key questions in the *foncier*, microfinance, qualitative and informal sector modules be asked during each round of the survey. INSAE has suggested that several of these modules may be skipped in subsequent rounds. If this occurs, critical data for the evaluation may be lacking. Agreement must be reached on this issue soon in order to determine how much of the data needs for the evaluation can be derived from future EMICoV surveys and how much data will need to be collected directly as part of the evaluation.

Randomization. In the discussion above, we observed that it was not strictly necessary that the assignment of eligible villages to the treatment and control groups be random; it could be carried out in a systematic manner. However, if MCC, MCA-Benin, ProGCN/GTZ and any other implementing agencies decide to carry out random assignment to treatment and control groups, we recommend that NORC be responsible for making the random assignment.

The assignment of villages to the control group will result in the control villages not receiving the ATL program sooner than the conclusion of the evaluation. There is a risk that these villages will become impatient with the delay in implementation of a program that they both demonstrated an interest and for which they qualified. While MCA-Benin and ProCGN/GTZ have provided assurances that the sue of randomization will not provoke a negative reaction from villages assigned to the control group, we remain concerned about this possibility and recommend that the strategy for how this approach will be communicated to potential beneficiaries be carefully discussed and agreed upon by all parties concerned before the village assessments begin.

C.6 Implementation Schedule and Next Steps

While this design report provides a roadmap for the next steps for the evaluation, as we note above, there still remain several key issues to be addressed with MCC, MCA-Benin, and the ATL project implementers. In summary, these include:

- Agreement with INSAE on the scope and timing of EMICoV follow-up surveys, any amendments to the current EMICoV modules, and means for assuring data quality..
- Understanding of the operational details of the urban ATL component and assessment of whether the proposed methodology for the rural ATL component can be applied in the same way or needs adjustment.
- Obtaining needed GIS data from government or other sources (see Annex A).
- Consensus on the communication strategy to potential beneficiaries concerning randomization.

Given the difficulties faced to date in obtaining follow-up information from Benin since the August field visit, we would recommend scheduling a follow-up field visit for the resolution of these remaining issues as early as possible in 2007.

Table D.3 — ATL Evaluation Implementation Timetable

Task No.	*Activity*	*Period*
Finalize Design		
1	Resolution of remaining issues from Design Report (Benin field visit)	Q1/2007
Assess & Develop Data Sources		
2	Provide additional questions for EMICoV survey modules for Phase 4	Q1/2007

Table D.3 — ATL Evaluation Implementation Timetable		
Task No.	***Activity***	***Period***
3	Identify and obtain GIS data	Q1/207
4	Develop supplementary short survey to be administered just before ATL commences	Q1/2007
5	Review and assess EMICoV survey data for Phase 1	Q2/2007
Data Collection		
6	Work with MCA-Benin and ProGCN to develop randomization procedures and treatment/control assignment protocols.	Q1/2007
7	If necessary, provide technical assistance	
8	Monitor village assessment and assignment of villages to treatment and control groups	Q2/2007 – Q4/2009
9	Where necessary, collect data at entry for assessed villages through short survey	Q2/2007 – Q4/2009
10	Complete development of GIS database to support evaluation	Q3/2007
11	Annual updates of EMICoV Survey	2008, 2009
12	Review EMICoV data from each annual cohort of ATL villages (treatment and control)	Q4/2007, Q4/2008, Q4/2009
Impact Evaluation Analysis		
13	Baseline data/analysis report	Q4/2007
14	Interim impact analysis reports	Q4/2008, Q4/2009
15	Final report on impact evaluation	Q3/2010

D. ACCESS TO FINANCIAL SERVICES COMPONENT

D.1 Evaluation Goals

This evaluation aims to assess the impact of the AFS component on the principal objectives of the program. First, poverty reduction, which is the ultimate aim of the program, will in this case be measured through the level of wages and number of jobs created in "treated" MSMEs, because of the assumed important role of small and micro businesses in particular in offering increased economic opportunities, more jobs, and higher wages for low income populations. Further, the evaluation will look at the steps leading to those outcomes, in specifically three areas:

- Profits and wages of MSMEs (micro-, small- and medium- enterprises) benefiting from the capacity-building activity
- Value of new financial services offered by financial institutions.
- Number and value of loans taken out by MSMEs

A fourth impact, changes in average household income in the MCC Access to Land (ATL) component areas, will also be assessed, but this will be done based primarily on data that will be developed under the ATL impact evaluation.[17]

The focus of this evaluation will be the impact of the Financial Innovation and Expansion Challenge Facility, a funding mechanism that will provide grants to institutions to support institutional strengthening, innovation, and the provision of business development services. The presumptive causal chain at work here is as follows:

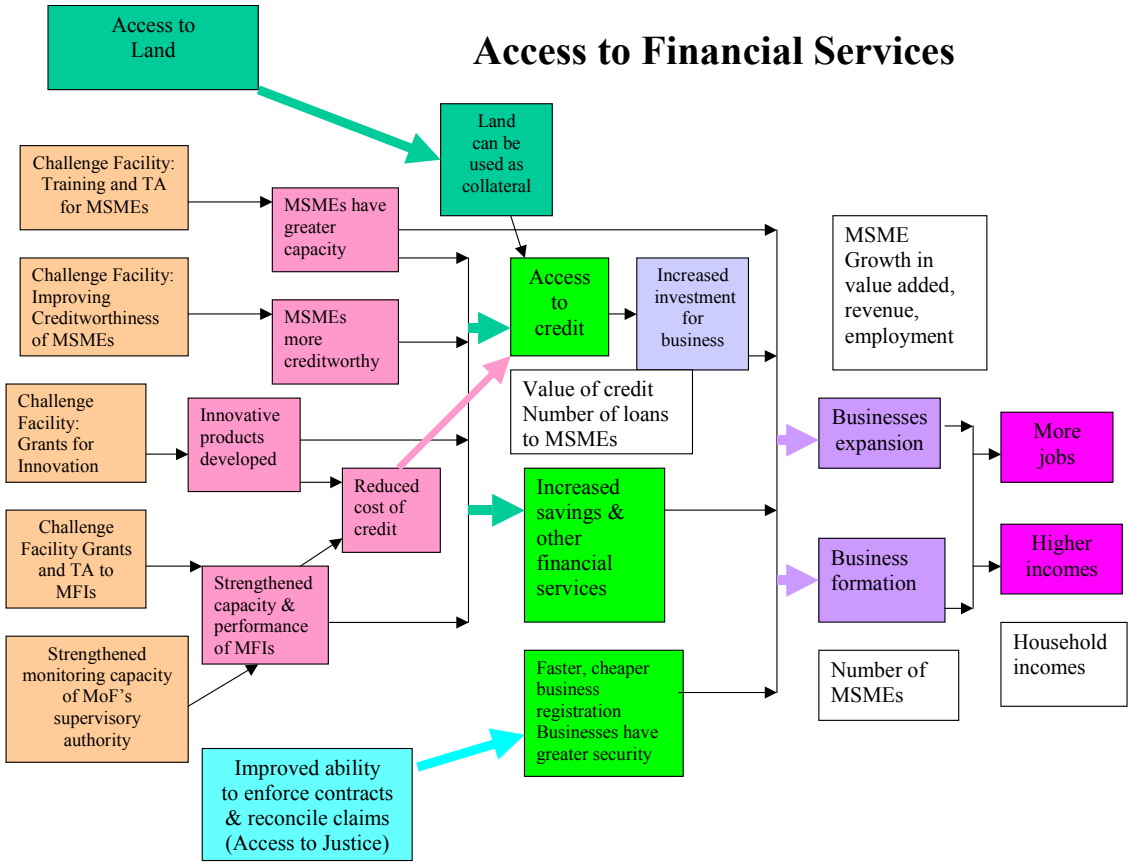
- MCA grants encourage more innovative/widespread lending or services by banks and/or microfinance institutions (MFIs) to MSMEs or more effective/widespread technical

[17] The effect of the AFS project on the households in the ATL project areas is discussed in design subreport for the ATL project. The proposed method would use dummy variables in multivariate regression models to discern if there is a statistically significant effect of the Access to Financial Services program on aggregate household outcomes, but will not attempt to measure the marginal effects of the Access to Financial Services intervention. One concern that will have to wait until the details of the AFS grant-funded activities are know is how to determine which ATL rural villages and urban *quartiers* are affected by an AFS activity—if an activity is household specific, we will have to identify ATL sample households that are affected; if the activity is regional (such as a marketing campaign), we will have to determine which ATL communities are affected.

assistance to MSMEs by MFIs or capacity-building institutions (CBIs) such as cooperatives, associations, or business development services (BDS) providers.

- MSMEs use improved access to credit or increased capacity to increase their output and/or productivity
- Higher capacity/more productive MSMEs increase profits and/or hire more workers or pay higher wages.

Figure D.1 — AFS Outcome Sequence



The outcome sequence chart appended below (Figure D.1 above) illustrates the linkages between project activities (on the left) and the desired intermediate and end outcomes. The key impacts to be measured in terms of MSMEs and households are shown in white boxes. The relationship of the AFS component to two other MCA program components – Access to Land and Access to Justice – is depicted through the boxes in the upper left and lower left corners, as they contribute through enhancing access to credit and strengthen contractual relationships for MSMEs. The chart illustrates that while the primary activities of the project (shown to the left in the chart) – specifically the grants

issued by the Challenge Facility – take place with banks, MFIs, CBIs, or BDS providers, the key impacts to be measured are those that affect the end beneficiaries, i.e., MSMEs and households. The distinction between these two groups of players will affect the design of the evaluation.

D.2 Methodological Challenges

The details of the specific interventions that will be funded by Challenge Facility will not be known with precision until the second half of 2007 when grants are awarded for specific projects. Therefore, the discussion here is based on a generic model of an intervention (see box at right) in order to frame the issues; the design may need to be revised once more details concerning the Challenge Facility activities are known.

Generic Model of AFS Intervention

As the AFS component will not become operational until the second half of 2007, we have made some assumptions about what the interventions will look like based on information gathered so far. The AFS Challenge Facility is expected to support grants to:

- Financial institutions seeking to expand the scope and scale of their services to MSMEs in previously underserved regions and sectors or introducing innovative technologies that lead to economies of scale and reduced operating costs and risks;
- Microfinance institutions (MFIs) seeking to improve internal controls, transparency and management;
- MFIs assisting borrowers' financial literacy, business management skills and ability to complete credit applications; and
- Rural networks and organizations improving the creditworthiness of their MSME members.

We assume that each grantee will have a current roster of customers, which may include MSMEs. The effect of the grant will be to:

- Increase the number of MSMEs served by the grantee, and/or
- Change the type of scope and scale of services (including availability of credit) delivered to MSMEs.

Thus, the intervention supported by the grant may affect both MSMEs that were clients of the grantee before the grant and MSMEs served by the grantee after the grant. It is expected that the intervention supported by the grant will either make credit more available to MSMEs -- either by increasing the volume of credit to MSMEs or by making MSMEs more creditworthy – or will increase their access to other financial services.

D.2.1 Randomization

The structure of the Challenge Facility, which solicits grant applications from banks, MFIs, and business development service (BDS) providers to improve their outreach to and effectiveness with MSMEs, implies that randomization is not likely to be possible, either in terms of dividing grant applicants into treatment and control groups or among the MSMEs that are the expected direct beneficiaries of the grants. We look at each of these in turn.

Grant Applicants. The grant application process will select grantees based on the quality of their applications. The use of quality based selection means that grantees will be selected with bias. However, if the number of grant applicants with

quality scores that were acceptably high for receiving a grant were large enough, then grants could be awarded using randomization (perhaps stratifying between the highest scoring grants and those with lower scores to allow a greater proportion of high-scoring applicants to receive grants—this is the approach being employed in the evaluation of the MCC Agricultural Development Activity (ADA) in Georgia).

However, the Challenge Facility is accepting grant applications from several different types of organizations—banks, MFIs, and BDS providers—and the nature of the activities from each of those groups is also likely to vary widely among applicants. Under such circumstances, the number of similar type applicants proposing to carry out similar activities will be small (given the small number of grantees).[18] If analysis is to be carried out comparing grantees and non-grantees, then the small number of potential treatment (grantee) and control (non-grantee) organizations will not support statistically powerful analysis comparing similar activities (see box above for a general discussion of power calculations and the footnote below for an illustrative example of a power calculation that might apply in this case).[19]

---

[18] The number of grants to be made has not yet been fully defined: The impact evaluation task order estimates 40 grants will be made; the Compact M&E plan estimates 50 grants; and interviews in Benin suggest as many as 100 grants. With this number of grantees, it is quite likely that the majority of the target organizations will receive a grant at some point over the course of the activity. Indeed, during the August field visit to Benin, it was stated that all or almost all MFIs active in Benin are likely to be selected as grantees by the Challenge Facility at some point.

[19] To translate this into something practical, suppose the proportion of the MSMEs with access to credit is $p = 0.25$ or 25% and we want to be able to detect a difference of $\delta = 0.10$. Once we know the value of the proportion p we can immediately obtain that the standard deviation is

$$\sigma = \{p(1-p)\}^{1/2}$$

Now with $p = 0.25$, we obtain $\sigma = 0.4$ approximately, hence the ratio

$$\delta/\sigma = 0.10/0.4 = 1/4 = 25\%$$

Now in the case of independent selections of n_1 treatments and n_2 controls and no contamination or mixing, the standard error is expressed in general as

$$\sigma\{1/n_1 + 1/n_2\}^{1/2}.$$

If we assume that $n_1 = n_2$, then we can determine either α, the probability of choosing the alternative when the null hypothesis is *true*, or β, the chance of choosing the null hypothesis when the null hypothesis is *false*. Usually we want to maximize $(1-\beta)$, or the power for a given type I or α level. If we assume $n_1 = n_2 = 20$ for a set of grantees with similar activities, then the resultant power calculations $(1-\beta)$ for various levels of α are as follows:

$\alpha = 5\%$	0.1965
$\alpha = 10\%$	0.3117
$\alpha = 20\%$	0.4796

Thus, at such small sample sizes, achieving even minimally acceptable power requires much higher levels of α than the usual 5%.

Power Calculations and Sample Size

Sample size determination requires information about the variability in the outcome being measured. For example, consider a simple hypothesis test that an intervention significantly increases an outcome. A simple comparison of means is used to test the hypothesis that the treatment had no effect versus the alternative hypothesis that the treatment made a material improvement. Assuming that the design is balanced with n units receiving the treatment and n units in the control group (the most efficient design), the sample size required depends on three main things:

- The probability of rejecting a true hypothesis (Type I error), denoted as α,
- The probability of accepting a false null hypothesis (Type II error) when a particular value of the alternative hypothesis is true, commonly denoted as β, and
- The population standard deviation, σ.

The power of the test is $1-\beta$, the probability of accepting the alternative hypothesis that the treatment improved the outcome, when in fact the treatment resulted in an average improvement of a particular size. One wants to determine a sample size that will provide high power to detect a "material" difference.

One also needs to know the underlying standard deviation. Because there is often no initial information regarding the standard deviation, the sample size is determined based on detecting an increase in the population mean of κ standard deviations, that is $\delta = \kappa\sigma$. The table below show sample sizes to achieve three different levels of power, $1-\beta$, and κ, for a fixed α=0.05 size test. The sample size shown is for the number of units in each group, treatment and control.

Power $1-\beta$	$\kappa = \delta / \sigma$		
	0.25	0.50	1.00
95%	350	90	25
90%	275	70	20
85%	230	60	15

For example, if one wanted to detect a treatment difference of at least one standard deviation (κ=1) with 95% probability, then one would need a sample of 25 control and 25 treatment units.

Caution: The illustration above uses the normal distribution and is good for large sample sizes or when σ is known. When σ is unknown the power calculation must use the t-distribution and the sample sizes would

MSMEs. The second potential aspect of the AFS activity upon which randomization could occur is the placement of MSMEs within or outside the grant-funded activity. In the case in which a completely new activity were being funded by the grant, then a mechanism could be envisaged whereby MSMEs that wished to participate in the program could be randomly assigned to treatment or control groups. (This assumes that the grantee organization would be willing to allow random selection of its potential customers, a non-trivial assumption for an organization whose objectives are likely to be focused on maximizing the success of its operations.)

However, since the expected grant applicants are organizations that are already working with MSMEs, it is likely that the grantees will already have MSMEs as existing customers. This will make the creation of randomly-assigned treatment and control groups of MSMEs impossible, as the

treatment group for the grantee will already be populated by the organization's existing MSME customers.

In short, we can divide the possibilities into two cases based on the key characteristics of the grantee projects:

- Randomization is possible if the grantee is able to – and willing to – offer the proposed "treatment" (improved service or product) to MSMEs selected randomly. For example.
  - — Training offered to applicant MSMEs, with a large client base to choose from
  - — Product that grantee is planning to introduce to a relatively limited number of MSMEs.
- Randomization is not possible if a proposed project is expected to benefit all MSMEs (for example all customers of one grantee) or benefits will accrue to MSMEs in a way that is not predictable. For example,
  - — Improved service that will benefit all grantee's current clients
  - — New product that grantee will want to make available as widely as possible

In the event that it is possible to randomize selection of participants for one or more grant projects, the evaluation would require of surveys of both treatment and control groups at the start of the project and at the end of the evaluation period, with possible annual updates. (Section D.3.1 on methodology below describes this process in greater detail.)

We provide now an alternative approach in the event that randomization is not possible.

D.2.2 Alternative to Randomization

If randomization is not possible, then a quasi-experimental approach will be needed to compare those MSMEs that are affected by the grant-funded activity—the treatment group—with similar MSMEs that are not—the comparison group. The typical quasi-experimental approach would be to use propensity score matching (PSM) to match the treatment group MSMEs to comparison group MSMEs. To do this matching, the characteristics of the MSMEs (or a representative sample of these) receiving the treatment have to be identified so that MSMEs not receiving treatment (*i.e.*, not on the existing or newly-sought customer list of the grantee) with characteristics that would result in the same propensity of being selected can be identified as the comparison group (see box) so as to

eliminate the bias implicit in the selection process. There are two timing problems that have to be addressed:

- Because the treatment firms will not be known until the grantees are selected by the Challenge Facility, the comparison group cannot be identified until the treatment group is known. This implies that sufficient data on possible comparison MSMEs are available to serve as comparisons, or that a comparison group has to be developed for each treatment group.
- If one of the objectives of the grants is to expand the target population of MSMEs with which a grantee interacts, then the treatment group for that grantee will change over time. This implies that additional members will have to be added to the treatment group over time (and the comparison group will have to be augmented as well). One scenario is that it is possible that members of the comparison / control group will actually become part of the treatment group over the course of the program, and so they will have to be replaced via the same matching approach.

Propensity Score Matching

Propensity score matching (PSM) is useful when the aim of matching between control and treatment groups is to find the closest comparison group from a sample of MSMEs not receiving treatment to the sample of MSMEs receiving treatment. "Closest" is measured in terms of observable characteristics.

The main steps in matching based on propensity scores are as follows.

- First, obtain a representative sample of eligible treatment and non-treatment MSMEs; the larger the sample of eligible comparison MSMEs the better, to facilitate matching.
- Second, pool the two samples and estimate a probit or logit model of participation as a function of all available variables that are likely to determine participation[20].
- Third, create the predicted values of the probability of participation from the estimated regression; these are the propensity scores (one for each sampled MSME).
- Fourth, exclude non-treatment MSMEs in the sample if they have a propensity score that is outside the range (typically too low) found for the treatment sample.
- Finally, for each MSME in the treatment sample, find the observation in the non-treatment sample that has the closest propensity score, as measured by the absolute difference in scores. This is called the "nearest neighbor."

More precise estimates can be obtained by comparing the mean of multiple nearest neighbors for each treatment observation.

[20] These variables should include ones describing physiographic conditions of MSMEs' location, such as measures of relative accessibility to markets, financial centers, transportation networks, and services such as water and power, as these are likely to influence economic output and/or investment. The GIS will allow these variables to be calculated and included for all locations (see Annex A of the ATL Design Subreport).

D.2.3 Data

The main challenge to be addressed at this moment (at least until more is known about the specifics of the organizations and activities to be funded by the Challenge Facility grants) is developing a suitable source of data about appropriately matched MSMEs that can be compared to the treatment MSMEs. The data requirements can be summarized as follows:

- Baseline data is available that allows for the development of a comparison group of MSMEs matched on the treatment group using PSM. This requires both adequate data on the characteristics of the MSMEs—likely those related to size (employment or turnover), sector or industry, and location (with respect to markets and financial centers)—as well as sufficient observations from which to make high quality matches.
- The baseline data needs to be of adequate quality to allow statistical inferences to be drawn with confidence. This implies both that the survey and data collection be of high quality, but also that the AFS grant applicants and MSME respondents must be willing to provide the data (i.e., adequately low rates of refusal and/or missing data).

A key factor, of course, is understanding the universe of enterprises (formal and informal), and we have yet to find much data to provide a comprehensive picture.[21] One of the most valuable potential contributions of the EMICoV and DDI surveys would be the information they might provide of the overall scope of enterprises in Benin.

D.3. Addressing the Challenges

D.3.1 Methodology

Given the small number of grantees and the likely wide variation of their interventions, the results of the impact evaluation are unlikely to be generalizable to similar programs in a statistically powerful way. Thus, the results will only be relevant for the particular grantee program for which the

[21] According to INSAE, for example, about 15,000 enterprises registered between 2000 and 2004, of which about 9,700 were in *commerce* (which presumably covers retail, wholesale and international trade). Some sectors are very small; banking and insurance had only 50 registered enterprises in 2004. On average, about 3,000 new registrations were filed annually during 2000-2004. However, it is unclear if all of these registrations are currently active, and these numbers obviously do not provide any information on the number or distribution of informal enterprises.

treatment and comparison group are selected. Identification of the individual projects to be evaluated will therefore be an important decision to be made by MCC once the grantees are selected.

The methodology described below is based on the assumed generic model of AFS intervention described in Section D.2 above. While we believe this generic model suffices for advancing the design discussion for the AFS project, the design will need to be revisited for each type of grantee activity for which an impact evaluation is desired by MCC, as the specifics of each activity may require some adjustment to the general approach described below. We examine the approach for assessing the impact on MSMEs first, and then discuss how to obtain the desired evaluation data on AFS grantees.

MSMEs. For each organization selected as a grantee under the AFS Challenge Facility whose activity is to be subject to an impact evaluation, the following steps would be taken at the start of the evaluation period:

- Obtain a listing of existing MSME customers who will be affected by the planned activity.[22] The listing should provide basic data on the customers to allow them to be contacted for a baseline survey. If one of the objectives of the grant will be to increase the number of MSME customers, then the baseline sample will need to be updated periodically with the new customers added by the grantee as a result of the grant activity, as they come online.
- Given the type of activity and expected impacts, determine the size of the sample of MSME customers to be surveyed as the treatment group. Sample size will be determined based on the expected changes (variance) in outcome measures to be detected, the statistical power of the results required by MCC, and the number of MSMEs in the relevant universe of MSMEs being examined.
- Carry out the baseline survey of the sampled treatment MSMEs. Data on numbers of jobs by type of worker and occupation, wages, sex of head of business, revenue, and profit would be collected.[23]

[22] In meetings with MFIs during the August field visit to Benin, there seems to be a general willingness to provide this information if it is so required.

[23] The baseline and subsequent follow-up surveys will be structured so as to allow comparability to the data source from which the comparison group will be drawn. The possible data sources for the comparison group are described in Section 5.2.

- Obtain covariate information in addition to MSME data gathered through the baseline survey (such as GIS-based information about access to markets and financial services).
- Using PSM, match MSMEs from the treatment group to comparison MSMEs from an appropriate data source (see Section D.3.2 below).

Once the treatment and comparison groups have been identified and baseline data gathered, follow-up surveys will be carried out on a periodic basis (see Section D.3.3 below).

If selection of a treatment MSME were based purely on observable characteristics, then using PSM would remove the selection bias due to differences between MSMEs that were and were not affected by the grantee activity. The propensity score measures the probability that a MSME would take part in the grantee activity as a function of that MSME's observed pre-treatment characteristics. If treatment and comparison MSMEs have the same propensity scores and all characteristics relevant to assignment of treatment are captured in the propensity score (i.e., the relevant characteristics are all observable), then the difference in their outcomes yields an unbiased estimate of the intervention's impact.

However, some unobserved characteristics of the MSME that correlate with intervention outcomes might also correlate with the MSME taking part in the grantee activity, which can introduce bias in the estimation of the activity impact. As long as the pre-intervention differences between the comparison and treatment MSMEs are the result of unobservable characteristics omitted from the propensity score that do not change over time in their impact on outcomes, then the double difference method will correct for the possible bias.

Double-Difference Method of Analysis

The standard approach to calculating double differences with respect to an intervention is based on the two situations faced by MSMEs: those that take part in the treatment and those that do not. The first difference is the comparison of average values for the outcome variables in the MSMEs in the comparison group and the same variables for the MSMEs which take part in the treatment. The second difference is between the pre-treatment and post-treatment situations. The steps to be taken can be summarized as follows:

- Undertake a baseline survey before the treatment is applied, covering MSMEs affected by the treatment and a comparison group of similar MSMEs.
- After the treatment is completed, undertake one or more follow-up surveys. These should be highly comparable to the baseline survey, both in terms of the questionnaire and the sampled observations (ideally the same sampled observations as the baseline survey).
- Calculate the mean difference between the pre- and post-treatment values of the outcome indicators for each of the treatment and comparison groups.
- Calculate the difference between these two mean differences to obtain the estimate of the impact of the program.

Thus, matching using PSM removes the selection bias due to the observed differences between the treated and comparison communities. Double differences corrects for possible bias due to the differences in time-invariant unobserved characteristics between the two groups. The impact of the activity is the change in the outcome indicators between matched MSMEs from the treatment and comparison groups.

Grantees. The evaluation objective also seeks to measure the changes in the practices of the grantees due to the Challenge Facility grants. Because the nature of each grant program is likely to vary significantly from other grants, we believe that the most effective way to collect such data is to make it part of the grant requirements. Once the nature of the grant programs is established, a set of indicator data can be developed (for example, value of new loans, use of collateral, use of new technology, use of training of MSMEs in financial management practices) about which grantees can be required to report. It is likely that one set of data can be defined to be comparable across all grantees, while additional specific sets of data may be necessary for each different type of grant.[24]

D.3.2 Data Sources

As described above, we recommend obtaining data from MSMEs in the treatment group through surveys administered to a random sample of that group. However, since carrying out the impact evaluation is dependent on matching MSMEs in the treatment group, the primary data requirement is for a source of data on MSMEs from which a comparison group can be matched. Ideally, the data source will also contain the needed information for measuring the desired evaluation outcomes. Three possible data sources that can be used for creating the comparison group have been identified:

- EMICoV informal sector module;
- Survey of enterprises carried out by the *Direction du Développement Industriel* (Directorate of Industrial Development, or DDI) at the Ministry of Industry; or
- Baseline survey of MSMEs that MCA-Benin is planning to execute to gather data to help design the Challenge Facility.

[24] The data on grantees described here is very similar to the data listed in MCA-Benin's M&E plan, so consideration should be given to having this data collected as part of MCA-Benin's M&E activity, rather than under the impact evaluation contract.

EMICoV Survey. The EMICoV informal sector module (as well as the employment module that precedes it) has a screening question that asks if the respondent either works independently or owns his or her own business and if the response is positive, the module is administered. It is unclear to what extent the EMICoV survey module on informal sector activity will capture the universe of registered MSMEs, because it was just beginning to be administered in August of 2006, but potentially, EMICoV could yield a fairly broad source of data on the informal sector. We will need to assess how well the module is capturing informal sector activity.[25] The EMICoV survey has the advantage of regular update surveys (expected to take place on an annual basis) being planned in the future. However, the content and exact timing of these update surveys has not yet been determined by INSAE.

DDI Survey of Enterprises. This survey was last carried out in 2000 with the expectation that it would be administered regularly, every two to three years. It has been postponed since then for budgetary reasons, but is now scheduled for 2007. The DDI survey in the past has used as its sampling frame a census also carried out by DDI in which interviewers walk through neighborhoods identifying businesses and asking basic questions on registration, address, sector, level of capital. If respondents are not willing to participate, the interviewer at least notes their identification. In 2007, DDI has decided to combine the survey with its regular survey on the economic climate and at the same time update its full census of enterprises in all the major cities and towns of Benin.

The survey at present includes questions regarding firms' production, sales, employment, and marketing. DDI expressed willingness to collaborate and if useful to add questions to their questionnaire, for example, on investor confidence in the courts or on access to credit. The timing of this survey as now scheduled is excellent to serve as a baseline—it is now scheduled to begin in March of 2007, with results available starting in July. There are some indications that the department carrying out the survey may require a fair amount of technical assistance to ensure the quality of the design and implementation of the survey, which will be reflected in implementation plans for the evaluation.

[25] One particular concern is how many observations the informal sector module will yield. If the number of households that respond positively to the screening question is small, then the usefulness of the EMICoV survey as a source of potential matches for the comparison group is greatly reduced because the matching will be poorer than it would likely be with a larger (and more varied) number of observations.

MCA-Benin Baseline Survey of MSMEs. Another potential data source is the baseline survey of MSMEs that MCA-Benin is planning to execute to gather data for design of the Challenge Facility. The draft Terms of Reference (TOR) for the baseline survey requests a representative sample of 3,000 business establishments, but does not provide any guidance on how such a sample is to be developed. If the sample for the MSME baseline survey is developed in a sound manner and obtains the appropriate data for the evaluation of the AFS project, it could serve as a source from which PSM could be done to identify comparable firms to match to the treatment firms. However, until we can gather more information on the structure of the population of MSMEs in Benin, we cannot say yet whether a sample of 3,000 is sufficient, given that it is likely that the sample will need to be stratified by industry, enterprise size, and (probably, given the apparent concentration of business enterprises near the coast) region. As described, if the survey were to be carried out only once, then unlike the EMICoV and DDI surveys, it doesn't provide for future updates that would provide outcome data on comparison group MSMEs.[26]

Data Source Compatibility. In fact, it may be possible to use some or all of the three sources combined—the EMICoV survey, the DDI survey, and the MCA-Benin baseline survey of MSMEs—as sources of data for the comparison group if it can be determined that there is little or no overlap between the data sets or if a way can be found to deal with the overlap of the three data sets.[27] Further information about each of the data sources is needed to determine how feasible such an approach would be.

Assuming an approach that allows the use of more than one data source is feasible, the data from each source should be as consistent as possible. The EMICoV and DDI surveys both collect information on employment, wages and salaries, hours worked, production, inputs, and investments, but there are differences in how each survey obtains the data and at what level of aggregation.[28]

[26] There is some possibility of a follow-on to this survey to be carried out in Year 3, but that is not definite.

[27] For example, the EMICoV survey distinguishes between registered and unregistered enterprises. If the DDI survey is limited to registered enterprises only, then it may be possible to use the EMICoV survey as a data source for unregistered enterprises only, assuming that the protocol for selecting households that will receive the informal sector module is appropriate (i.e., the screening questions used are effective). Initial results from the EMICoV survey in December 2006 will provide more information about the extent to which EMICoV will be able to produce a large enough data set of informal enterprises for the purposes of providing information on both the treatment and comparison groups.

[28] For example, the EMICoV collects labor data by individual employee, but does not use the same classification of employees (i.e., specialists/ managers, skilled workers, administrative personnel) as the DDI survey.

Neither survey attempts to elicit direct estimates of profitability, but seem to have sufficient production and cost data to make estimates of profitability. The DDI survey does not seek any information on credit; the EMICoV survey has quite detailed questions on the sources, terms, and uses of financing. Both the EMICoV and DDI surveys offer the opportunity for making adjustments to the survey instruments in the coming months, so this opportunity needs to be explored. The table below illustrates some of the differences between survey coverage of key topics, which will be addressed in our additions to the two surveys.[29]

Table D.1 — Comparison of Survey Questions by Subject

Subject	EMICoV – Informal Sector	DDI Questionnaire
Employees	Number, divided into wage vs. salary; age, sex, ethnicity; recruitment, type of contract, salary, benefits	Number, divided into specialized/managers; skilled and laborers; administrative personnel – and then by permanent/ temporary; resident/ foreigner.
Hours Worked	Hours worked the previous month by each employee	Total hours worked by skilled vs. unskilled laborers for each of 3 half-years
Production	Amount & price transformed, sold, value over the last month	Amount produced / sold by product, by month, over the last two years
Investment	Categorized by type of investment (land, building, or machine)	Categorized by type (land, bldg, machine, R&D, training) and condition (new, secondhand, renovation, other)
Competition/ Marketing	Exports, competition Marketing (how do you seek clients)	Questions about competition, fiscal policies, participation in fairs, exports
Location	Captures SMEs without own site; Rent or own	No questions about tenure; only includes sited enterprises
Date of Establishment	Asked	No questions
Clients	Principal clients by type (public sector, private, size)	No questions
Regulation, Corruption	Questions about bribes, taxes, fines	No questions
Loans	Source, amount use, contract, repayment, duration; Micro-credit	No questions
Registration	Attitude to registration	No questions
Court System	Asks about confidence in the judiciary	No questions

[29] An alternative approach would be to carry out a two stage sampling approach as INSAE has done for the EMICoV survey. (We do not recommend this approach because it is likely to be more expensive than the other alternatives discussed above, but present it here for the sake of completeness.) A random sample of ZDs could be drawn in the first stage, and then an enumeration of business establishments in each selected ZD could be developed, from which a random sample could be drawn. If the sample were designed to be representative at the department, regional, or national level, the number of observations required would be significantly less than the approximately 18,000 in the EMICoV sample (which is designed to be representative at the commune level). The manner in which the INSAE enumeration is done, which lists all locations in the ZD and then excludes non-household locations, may provide data to allow such an approach (although it is unclear if the non-household enumeration data is readily accessible and coded in a consistent or adequately detailed manner to be used as the basis for drawing a sample of business establishments).

GIS. As with the Access to Land (ATL) project, consideration of geographic conditions and especially variation in village accessibility to markets, cities or locations providing credit could greatly increase the accuracy of the analysis. This will be possible using the GIS that will be constructed to support the Access to Land project (see Section C above and Annex A). The same geo-database will be used for all three Benin MCC components being evaluated (Access to Land, Access to Financial Services, and Access to Justice), and thus additional marginal costs will be low while potential synergies could be high. Most of the data that will be entered into the GIS database will be entered once to support all three MCC Benin projects. These data will include geo-locations of EMICoV and business survey respondents, geo-locations of villages/towns, ZDs, and communes, data on digital road networks (including approximate road quality/speed) allowing calculation of access measures for respondents or villages to cities, markets, financial centers, courthouses, etc. In addition, Benin-wide spatial distribution of exogenous geographic variables such as rainfall levels, soil fertility, topography, river/stream and hydrology data, etc., will also be entered into the GIS (from Benin government and other sources, see Annex A), allowing precise calculation of village point estimates for these variables, as well as ZD, commune or *quartier*-wide aggregated estimates for any variables.

For example, the spatial analysis could be used to accurately measure approximate geographic/travel distance to credit providers locations (calculated as a measure of "accessibility" along existing road networks, controlling for road travel quality/speed, topography, land cover—details are again provided in the Annex A. Data on road networks and geo-locations of villages (collected already for the access to land analysis) could be used to create accessibility indices for distance to credit provider locations (such as city/town markets or government district centers providing credit). This could be an important control on evaluating access to credit: for example, if a MSME producer has to walk or travel a short distance to talk to a credit provider officer, he may be much more inclined to purchase credit than if he has to walk or travel a long, costly distance[30]. Thus, some locations may possess exogenous access advantages that should be considered.

[30] Variation in travel cost per farmer or household to the nearest credit provision facility has been shown in the development literature to be quite significant, and pose a significant barrier to providing credit to poor or rural populations in many developing countries. See Felkner and Townsend (2006), for example.

D.3.3 Evaluation Timing and Sequencing

Apparently the Challenge Facility is now scheduled to be established in thethird quarter of 2007, with the first call for applications to take place in September. It will be important to have the baseline survey (or surveys) carried out before that process begins. That now seems feasible for each of the proposed data sources:

- Data collection for the EMICoV survey (informal sector module) is already underway and even the fourth "passage" will be completed by November of 2007, which is likely to precede the earliest grants.
- Data collection for the DDI survey is now scheduled for mid 2007, and may even be carried out earlier.
- The MCA-Benin survey of MSMEs is scheduled to be carried out in early 2007.
- MFIs and other grantees will be asked to provide baseline information on their customers and from their own records at the time of grant application. Whatever information cannot be provided by the grantees will be sought via surveys of beneficiary MSMEs.

It appears that the timing of the Challenge Facility's start-up and schedule for grant applications and awards is on track with the anticipated timing outlined in work plan (i.e., with the first call for applications to be issued in September of 2007). It will be important to be sure of the timing in order to ensure the desired sequencing.

Assuming that annual update surveys of enterprises(as are planned for EMICoV) will be possible, Table 2 outlines the proposed approach to data collection for the AFS evaluation. The symbol T designates the point in time of the intervention; in this case, the Challenge Facility grant award to grantees to be evaluated. Each Treatment Group consists of a representative sample of MSMEs associated with the grantee being evaluated.[31] The comparison group is identified using PSM. Data collections for the grantee are denoted by Y_iX_i, where i signifies the data collection period; period 0 is the baseline. Y denotes the data, or outcome measures, to be collected in either the EMICoV

[31] Since the number of grantees is small and will be spread over three different kinds of grants, there is no possibility of drawing a statistically meaningful sample of the grantees. MCC and MCA-Benin will need to provide guidance as to how many grantees will be selected for evaluation.

interviews or in other data collection. Collection of the covariates (X, or exogenous variables) for final modeling must also be done.[32]

Table D.2 — AFS Data Collection Sequencing for Typical Grantee				
	Time (in Months) from Baseline			
Cohorts	*Baseline*	*0-12*	*13-24*	*25-26*
Grantee A				
Treatment	Y_0X_0, T	Y_1X_1	Y_2X_2	Y_3X_3
Comparison	Y_0X_0	Y_1X_1	Y_2X_2	Y_3X_3

The initial round of the EMICoV and DDI surveys (and/or other data sources) will provide pre-treatment baseline data on the treatment and comparison MSMEs. Additional rounds of data collection occur annually in the 3 years post intervention (assuming the final round can be completed prior to the end of the task order to allow adequate time for analysis). Ideally, these post-intervention data collections could be fulfilled by the annual EMICoV updates which INASE is planning and by subsequent rounds of the DDI survey.[33] Since the matching of treatment and comparison MSMEs will not be possible until the AFS grant-funded activity is defined and the sample of treatment MSMEs drawn, data collection from the treatment MSMEs will have to be undertaken at the start of the activity. In some instances—where the baseline data is out of date, fore example, data collection at the start of the activity may also be required for the comparison group MSMEs.

D.4 Other Evaluation Design Issues

Other Activities by MCC, Donors, or Government. There may be some MSMEs in the treatment or comparison group that will be subject to influences that will affect the outcomes being measured by the evaluation. In the case of other MCC activities, as with the ATL project evaluation, we do not recommend trying to obtain sufficient data to try to discern the marginal effect of each additional intervention, as this would greatly increase the sample size required. Rather, we suggest including a

[32] If the covariates are not time varying (e.g., locational data), then we would collect them in a one-time operation. Otherwise, some covariates may be collected each year as part of the update survey. The covariates are needed, of course, for both treatment and comparison MSMEs.

[33] The DDI survey has been discussed as a bi-annual survey. If annual updates of DDI data were desired for impact evaluation purposes, then the "off-year" updates could focus only on the MSMEs in the comparison groups.

dummy variable in the analysis for each program to see if there is a statistically significant effect on aggregate outcomes where multiple programs are in effect.

In other cases, actions by other donors or the government may cause the outcomes for some MSMEs in the sample to be significantly affected. For example, if a new road is to be constructed that will drastically change the market access of a group of MSMEs during the course of the evaluation, then a decision will need to be made as to whether those affected MSMEs should be dropped from the sample or whether the change in market access can be adequately modeled to control for the change. In any event, it will be important that we stay informed about such actions and explicitly make decisions with MCC and MCA-Benin about how to handle such circumstances as they occur.

Data Quality. During the month of August, NORC staff accompanied INSAE's data collection team to a number of rural and urban sites. NORC witnessed the data effort and was satisfied that INSAE and the team are following accepted standards of field collection. INSAE has secured well educated staff and implemented a comprehensive training program that prepares them for the field interviewing. It will also be important to assess other elements in the data collection process including the receipting, transfer, and data entry of the questionnaires to ensure that the highest quality is maintained at each of these steps.

INSAE has indicated that it will perform the necessary reliability checks of data collected in each round. This analysis and conclusions should be provided to NORC at regular intervals so that we can evaluate the quality of the EMICoV data on an on-going basis during the evaluation.

During the August field visit, it became clear that some key modules and questions were being refused by a number of respondents. We could not determine in a short visit how widespread this problem was. Critical items on informal sector activity and investment are potentially sensitive but critical to determining if resources are being invested. It appeared that this might be a greater problem in urban ZDs than in rural ones. Since respondents may not feel secure in the confidential nature of their responses, their fear could result in low response rates for these key items, jeopardizing key elements of the impact evaluation. It will be important for interviewers to gain cooperation of respondents and convince them of the confidentiality of their responses. NORC can provide assistance to INSAE to help address this problem.

NORC will work closely with INSAE field staff to expand their understanding of confidentiality and provide them with a comprehensive list of Frequently Asked Questions (FAQ) that include items related to confidentiality; NORC will also work with the staff to learn and practice providing this information to respondents. It is recommended that the interviewers spend time working in groups, just as they did when translating the questionnaires, to ensure that they can provide both consistent and appropriate translations for these assurances.

While the DDI at the Ministry of Industry has voiced its interest in collaborating with the evaluation, and their survey seems to be aimed at collecting the kinds of data needed for this evaluation, we have not yet had any opportunity to assess the quality of the data (or underlying data collection procedures and processing) for the DDI survey. This is obviously a high priority task in the coming months if the DDI survey is going to be used as a data source for matching the treatment and comparison groups and for baseline and outcomes data for the comparison group.

Schedule of Follow-up EMICoV and DDI Surveys. INSAE has yet to determine which modules will be administered in subsequent rounds. For the AFS project, the overlap of data requirements with the ATL project is quite large, so the key questions in the *foncier*, microfinance, qualitative and informal sector modules required for the ATL impact evaluation update surveys may also meet the AFS data needs.[34] INSAE has suggested that several of these modules may be skipped in subsequent rounds. If this occurs, critical data for the evaluation may be lacking. Agreement must be reached on this issue soon in order to determine how much of the data needs for the evaluation can be derived from future EMICoV surveys and how much data will need to be collected directly as part of the evaluation.

Similarly, discussion is needed with the DDI as to the timing of the future rounds of the DDI enterprise survey. In particular, we will need to discuss whether the DDI survey will be updated annually and if the timing of those updates could be coordinated with the EMICoV updates.

[34] Even if the AFS impact evaluation requires some additional questions beyond those needed for ATL, the fixed costs of fielding the survey make the marginal costs of a few additional AFS questions very small.

Choosing Evaluation Targets Carefully. As there will be as many as 40 to 100 grants awarded by the AFS activity and the grants are likely to cover projects of various different types, it is our assumption that not all of the grant activities will be evaluated and that some selection of evaluation targets will be done. It may be possible to group some of these grants into a single evaluation,[35]

No matter what approach is used to obtain a comparison group to match against the treatment group to be selected from Challenge Facility grantee customer lists, a fundamental issue remains: Until more is known about the specifics of the Challenge Facility grants and grantees (*i.e.*, until we can understand the characteristics of the treatment group), it is difficult to specify in advance what is the appropriate comparison group. Thus, if we rely on data from EMICoV and/or the DDI enterprise survey for our comparison group, it cannot be known with certainty that these data sets will have sufficient observations for enterprises that can be considered as a good comparison group to the treatment group. However, we can increase the odds for success by avoiding grants that focus on sectors or types of enterprises that have few MSMEs.

Additionally, grant-funded activities that have treatments and expected outcomes that are clearly and consistently defined and measured are likely to produce higher quality evaluation results.

Communication with Stakeholders. The MFIs have exhibited interest both in the project and the evaluation, and have expressed themselves to be willing to provide data as needed. Nevertheless, MFI representatives expressed an interest in seeing the design of the project in greater detail before fully endorsing it. Another MFI representative observed that the impacts of microfinance are much debated and cannot be taken for granted. MCA-Benin has provided information that a series of outreach meetings are now underway with potential grantees. We wish to stress that MCC, MCA-Benin, and NORC will need to work closely with the AFS implementing organization to ensure that the community of AFS potential grant applicants is consulted closely throughout the evaluation to ensure that any concerns that are raised that might jeopardize the cooperation needed from participating MSMEs are addressed promptly an effectively.

[35] For planning purposes, guidance is needed from MCC as to how many AFS grant activities are to be evaluated.

D.5 Implementation Schedule and Next Steps

While this design report provides a roadmap for the next steps for the evaluation, as we note above, there still remain several key issues to be addressed with MCC, MCA-Benin, and potential partners in the AFS evaluation. In summary, these include:

- Further discussion with DDI on the suitability of the DDI survey of enterprises for use in the evaluation. These discussions will include the coverage and structure of the sample used for the survey and the content of the survey itself. Our working assumption going into these discussions is that the most desirable outcome is the use of both the DDI and EMICoV surveys as sources of baseline data and matching for MSME comparison groups to treatment groups of MSMEs sampled from customer lists of AFS grantees.
- Agreement with both INSAE and DDI on the scope and timing of follow-up EMICoV and DDI enterprise surveys, respectively, amendments to the current EMICoV modules and DDI survey to achieve maximum comparability of data across the surveys, and appropriate means for assuring data quality.
- Work with DDI to determine types of technical assistance that will be useful in designing and carrying out the enterprise survey.
- Developing a better understanding of the details of the AFS Challenge Facility and its operation, first through the results of the MCA-Benin baseline survey of MSMEs and then as those results are translated into the approach that the AFS implementing organization will be taking with respect to soliciting grant applications (particularly in terms of any guidance or preference it will give for certain types of activities).
- Obtaining needed GIS data from government or other sources.[36]
- Consensus on the communication strategy to potential grantees and their customers concerning the evaluation, our approach, and the need for cooperation from customer MSMEs.

As there are still a number of areas that would benefit from on-site discussion, we would recommend scheduling a follow-up field visit as early as possible in 2007 for the resolution of these remaining issues.

[36] This is also being done as part of the ATL design development.

Table D.3 — AFS Evaluation Implementation Timetable		
Task No.	***Activity***	***Period***
Finalize Design		
1	Resolution of remaining issues from Design Report (Benin field visit)	Q1/2007
Assess Data Sources		
2	Review analysis of EMICoV first passage	Q1/2007
3	Review sampling approach and frame for DDI census	Q1/2007
Baseline Data Collection		
4	Provide input for MCA-Benin baseline survey of enterprises	Q1/2007
5	Provide input for questions to add to 4th passage of EMICoV	Q1/2007
6	Provide input for questionnaire and sampling of DDI survey	Q1/2007
7	Review findings from study of demand and explanation of financial and non-financial services for MSMEs	Q2/2007
8	Provide design input for grant application data requirements for Challenge Facility	Q2/2007
9	Work with potential grant applicants to review / refine data forms for their customers	Q3/2007
Baseline Analysis		
10	Carry out data quality audits EMICoV and DDI surveys; prepare data for use in PSM for comparison groups and as baseline data source	Q4/2007
11	Entry of first AFS grantees into evaluation	Q4/2007
12	Draw treatment group sample	Q4/2007
13	Select comparison group via PSM	Q4/2007
14	Prepare report on baseline analysis	Q1/2008
Impact Evaluation Analysis		
15	Periodic updates in each of the sample surveys	Q2-4/2008, 2009
14	Interim impact analysis reports	Q4/2008, Q4/2009
15	Final report on impact evaluation	Q3/2010

E. ACCESS TO JUSTICE COMPONENT

E.1 Evaluation Goals

The impact evaluation seeks to measure the effect of the ATJ project on the following:

- Average time required to enforce a contract
- Investor confidence in the judicial system
- MSME growth (number of businesses and value added)
- Effect of construction of new courts of first instance (TPIs) on dispute resolution (in terms of improving access to the courts and reducing the time to resolve disputes).

The ATJ project interventions are based on the following premises:

- A more efficient and expanded network of business registration centers (*Guichets Uniques*) will reduce the average time for enterprise registration. Reduced time to register a business will in turn lead to growth in the number of MSMEs and their value added.
- Increased use of alternative dispute resolution (ADR) will decrease the cost of resolving commercial disputes and thereby lead to a reduction in the average time required to enforce a contract.
- The construction of new courthouses will lead to a reduction in the average time it takes to resolve a dispute and will provide greater access to the Beninese, who will in turn have greater confidence in the judicial system as it will reduce the costs to resolve disputes and reduce travel costs.
- Effective training for court judges and dissemination of court decisions, laws and other legal information will create greater transparency in the judicial system and greater certainty in commercial transactions. Greater transparency and certainty in commercial transactions will lead to increased investor confidence in the judicial system.
- In turn, both the increase in judicial confidence and the reduction of time and cost to enforce a contract, will also lead to MSME expansion, which will create more jobs and alleviate poverty.

E.2 Methodological Challenges

The ATJ project does not appear to lend itself to an impact evaluation approach given the nature of the interventions:

- Interventions to improve the services of the courts are either general improvements to the system (such as the training of judges, improving access to information, or increasing the availability of legal aid) that are difficult to attribute to specific locations or populations, or where there are localized effects (as in the case of the new courthouses to be constructed) it will be difficult to identify comparison areas against which to measure impact, given the selection criteria enumerated in the Compact and the fact that other donors are also constructing new courthouses.
- Development of a system of alternative dispute resolution (ADR) that will be available nationally (although located in Cotonou) where none existed before does not allow for the division of the target population into treatment and control groups.
- Improvement and expansion of the *Guichet Unique* is also planned to occur on a few sites that are intended to serve large areas that will make it difficult to construct comparison groups that are comparable and yet unaffected by the intervention.

Based on the above, we do not believe it is appropriate to attempt to carry out an impact evaluation using experimental or usual quasi-experimental approaches, as the nature of the interventions make it unlikely that valid control/comparison groups can be defined and identified. In such circumstances, a quasi-experimental design using reflexive comparison can be used. This type of design is particularly useful in evaluations of full-coverage interventions such as nationwide policies and programs in which the entire population participates and there is no scope for a control group. The counterfactual is constructed on the basis of the situation of program participants before the program. Thus, program participants are compared to themselves before and after the intervention and function as both treatment and comparison group.

The major drawback with reflexive comparisons is that the situation of program participants before and after the intervention may change owing to myriad reasons independent of the program. In the case of the ATJ project, for example, participants in the justices system may see improved outcomes

in terms of the time required to complete cases. While this improvement may be due to the ATJ project interventions—better information systems, training for judges, new courthouses—it may also be due to the fact that the economy is performing better and less cases are being pursued. Thus, comparisons of the "before" and "after" circumstances must be carefully done so as to distinguish between effects of the project and other external factors, so as not to confound the two.

However, based on the evaluation designs presented for the Access to Land (ATL) and Access to Financial Services (AFS) projects and other data that may be available, there are measures that can be developed that will allow such analysis of the effects of the ATJ activities. These measures fall into two categories:

- "Primary" measures that will require collection of primary data specifically for the purpose of the ATJ project (such as the time taken to settle contract, property, and other similar civil disputes).
- "Secondary" measures that can be derived from existing data sources or from data collection recommended for the ATL and AFS impact evaluations; these include data on confidence in the judicial system and growth of MSMEs.

E.3 Addressing the Challenges

E.3.1 Primary Measures

The primary measures we recommend developing relate to the time to settle contractual and other similar civil disputes that affect the conduct of economic activity by businesses and households. Based on the proposed ATJ project activities, the analysis should examine the throughput and whether the number of cases change, if the composition of cases (*i.e.* the types of claims) change, and how the filing location and location of the parties in the cases change in response to the construction of new courthouses (including both those constructed by MCC and other donors).

While the description of data collection in this section focuses on court records, we note that as the ADR system develops, comparable data should be collected for those cases as well, so that the relationship between changes in the judicial caseload can be compared to those in the ADR system.

Data collection procedures for ADR records will be developed as the new records systems are established.

The existing data source for this measure will be the court records currently archived (for cases that have been resolved) and those that are filed in the office of the court clerks (active cases). Future cases beginning with criminal in 2007 and civil in 2008, will have their data stored in a computerized case management system that is being developed by the EU. The proposed analysis will compare data from three categories of cases:

- Cases completed in the last 5 years (i.e., completion dates during the period 2002-2006;
- Cases still active at the end of 2006; and
- Cases filed after 2006.

Data Required for Case Completion Analysis

The following is a preliminary listing of the data required for the analysis of the time required to complete judicial cases. The data list below assumes the existence of a GIS (as described in the ATL evaluation design) that will allow spatial analysis linking the location of the parties involved and the site of the case proceedings with access indices and other spatial measures.

- Type of case; only civil or commercial cases concerning business or economic matters (e.g., contracts, property, labor)[37]
- Date case was filed
- Date of first court hearing
- Date case was closed
- Date payment/restitution required by judgment was made
- Address of plaintiff and of defendant
- Location of the courthouse
- Identifier for judge presiding
- Decision and monetary value (if any)
- Whether the case was appealed and, if appealed, date of appeal decision and whether decision was reversed or modified.

The paper case files stored in the archive and in the office of the court clerks do not seem to be filed in any systematic manner, so we do not have an exact estimate of the number of files in each class. The files are classified (by the color of the file folder) between civil, commercial and criminal cases; appealed files are also color-coded separately. For the purposes of developing an approach, we have conservatively estimated that there are about 15,000 cases in each of the archived and active categories.

To carry out the proposed analysis of the time required to complete each case, we must not only collect data on the date the case was filed and the date a judgment was entered, but also on covariate data that might influence how long cases take to adjudicate (such as the type of case—contract, land/property, labor, and other economic/business-related civil cases—and value of the dispute—

[37] We need guidance from MCC and MCA-Benin whether this list should be more inclusive, so as to include, for example, tort cases that have economic aspects, but are not strictly related to commercial interests. We have assumed that non-economic civil cases (such as those falling under family law) will not be included in the evaluation.

assuming the value is a proxy for the complexity of the case) and locational data (the location of the courthouse where the case was heard and addresses of the plaintiff and defendant) that will allow analysis of any possible linkages between the construction of new courthouses and the creation of ADR centers.

Given the number of cases involved and the fact that the cases are apparently not filed or organized in any systematic manner, we recommend a two-stage process to develop a sample of the paper files upon which to base the analysis.[38] In the first stage, we would select a proportion (we are assuming 50% at this point) of the eligible files and collect basic information on each case—type of case, starting date (and ending dates for completed cases), and some locational information.[39] Once we have this basic data for eligible cases, we can then use the data to develop estimates of the variance in the time to complete cases, assuming that the time required for the completed cases represents the upper range of the estimate for the rest of the cases and that this assumption is not contradicted by the time-in-

Grouping of Cases for Data Collection

The data collection for the eligible cases will be organized into three groups:

- ***Group 1 – Cases completed within the last 5 years.*** In this group, we will be looking at cases that have been resolved. These cases are filed in a room of the *Greffe* (Archives). The cases are organized by the type of case (as indicated by the color of the file cover—red for criminal, green for commercial, yellow for civil, and pink for appealed). Once the first round basic case data has been collected, the cases will be organized by start date (and possibly blocked or stratified on other covariates) for sampling. Once the second round sample has been drawn, detailed case data will be added to the database for analysis.
- ***Group 2 – Cases still in progress.*** In this group we will look at cases that are still in process by the end of 2006. Cases that are still in process are stored in a room with the court clerks and do not seem to be organized in any manner aside from the color of their file cover (red for criminal, green for commercial and yellow for civil). We will follow the same two-stage process as outlined for the Group 1. For the cases sampled at the second stage, data will have to be collected during the evaluation period on cases that have been resolved. It still needs to be determined who can access the data for cases in progress and enter it into the database system.
- ***Group 3 – Future cases.*** Beginning in January 2007, criminal case data will be entered into a case management data system under an EU-funded project. Civil case data will be entered into the system by the EU beginning January 2008. Due to the late processing for civil cases, further discussion with MCA-Benin is needed to see if contractors could enter these cases into the database at an earlier time. Access to the data and its compatibility with data to be extracted from the paper files has been verified with the EU and a follow-up comparison in late 2007 will be needed.

[38] We are assuming that compatible data will be available to us for analysis once the EU case management system is operational; this assumption needs to be verified.

[39] Some basic case data is recorded on the cover of the file; other data needs to be mined form within the file itself. Our approach in the first stage would be to try to limit the data being collected to that available on the file cover. This approach is based on our preliminary estimate of there being approximately 15,000 eligible cases in each of the archived and active categories for the period 2001-2006. If, on closer examination of the numbers of eligible cases is significantly lower, then this preliminary data gathering could be modified so as to become a census of eligible cases rather than a first sample if such a modification could be accommodated with the available budget resources. (MCC has indicated that up to $10,000 can be made available for the creation of this database.)

progress measures for the current cases. With these variance estimates, the statistical power of various sample sizes can be calculated and a sample size selected - for a second round of data collection to obtain the remaining detailed case information - that fits best with the resources available.[40]

With a database populated with the detailed data extracted from the sampled paper files and compatible data from the planned EU case management system and which is linked to appropriate GIS data, we can carry out analysis of outcomes along the following lines:

- Changes in the number of cases filed
- Changes in the type of cases filed
- Changes in the time required to complete cases
- Changes in the location of plaintiffs and defendants from which cases are filed
- Changes in the location at which cases are filed (as new courthouses are constructed)
- Changes in the satisfactory resolution of cases

By linking the case data to the GIS, we can analyze and disaggregate the analysis geographically (see box).

As we noted above, the nature of the ATJ project does not allow for the definition of treatment and control/comparison groups, so the above analysis

Geocoding of Case Data

We recommend geocoding the detailed case data to be collected. Using the GIS proposed to be developed for the ATL impact evaluation, we anticipate that geo-coding of claims should be possible by collecting data on the village or town of plaintiffs and defendants, and then linking that data to the geo-locations of Benin towns and villages (which will have been entered into the GIS database with accurate data obtained from Benin government agencies as well as extracted from satellite images and aerial photographs).

An example of the type of analysis that use of the GIS would allow is the effect of building new courthouses. Areas with new courthouses may witness both a surge in the number of claims and eventually reduced time to process claims. The GIS could be used to calculate measures of access to new courthouses for each respondent (using the techniques described in Annex A), and this data could be quantitatively compared against the dates of the filing of claims, to control for geographic access of respondents to courthouses. This would provide a quantifiable approximate measure of the impact of new courthouses in encouraging claim filing.

Data on the processing time of claims for all courthouses or per specific courthouse could be compared with data on respondent travel-time access, etc. The GIS would allow the spatial display of locations of all courthouses as well as the spatial display in the variation of access difficulty for respondents, as well as the spatial display of the location of the filing of new claims and the variation in the timing of processing of claims (by respondents or by courthouses).

Visual analysis of this data could reveal interesting trends and patterns that might not otherwise be discernible (as is often the case with spatial display of data) and spatial statistical measures can be applied to quantify perceived trends/patterns.

[40] Under typical assumptions (normal distribution, $\alpha = 5\%$, and detection of a change in the population mean equal to 0.25 standard deviations), a sample size of 350 is required to achieve statistical power of 95%. (See Section 4 of the AFS design subreport for a more detailed discussion.)

will be limited in its ability to attribute causality to the ATJ activities. However, we will attempt, in our analysis to control for as many factors as possible in order to draw as strong a relationship as possible between the results and the ATJ project.

E.3.2 Secondary Measures

Other measures of the effect of the ATJ project can be developed using data that is either already being collected (such as the EMICoV survey) or which can be collected through coordination with other planned data collection efforts (the planned survey of enterprises carried out by the *Direction du Développement Industriel* (Directorate of Industrial Development, or DDI) at the Ministry of Industry and the judicial system satisfaction survey carried out by the Ministry of Justice). Measures of confidence in the judicial systems and growth of MSMEs are discussed in turn.

Confidence in the Judicial System. The EMICoV survey module on governance has questions that ask respondents about their confidence in the justice system, so a ready source of data on perceptions of households and informal businesses associated with those households is already available. The DDI survey of enterprises does not currently include any items relating to the justice system. If comparable questions to those included in the EMICoV survey could be added to the DDI survey, then the combination of the two surveys would provide data on this topic from households and both formal and informal enterprises. The satisfaction survey is carried out by the Ministry of Justice every two years and asks the respondents about their opinion of the judicial system. This survey has a random sample size of 2827 individuals in the 12 Administrative Departments in Benin. Of those interviewed, 91.5% of respondents are selected from households and 8.5% are selected individuals who have used court services. Clarification from the Ministry of Justice is still outstanding regarding how the respondents are selected and what type of questions are included in the survey.

The responses to these questions can be used to estimate the percent of households and business enterprises at various scale levels of confidence (the EMICoV survey uses a four-point scale) to construct a confidence index. If respondent data in all three surveys is geocoded[41] (which we believe is the case for the EMICoV survey and could be incorporated into the design of the DDI survey of establishments as well as the Ministry of Justice satisfaction survey), variation of the

[41] Inexpensive (approximately $100) hand-held Global Positioning System (GPS) receivers carried by interviewers would allow almost instant recording of extremely precise geo-locations with an accuracy of less than 10 meters.

confidence index with proximity to new courthouses could be examined to see if improved access through new courthouses leads to changes in the confidence index. Access index values for each respondent capturing the variation in respondent access to the new courthouses, measured as travel-time along existing road networks will be calculated using the GIS (see description in Annex A of the ATL design subreport for details on calculation of access indices).

These relative access values per respondent can then be regressed onto the confidence indices. This regression model can be examined rigorously by including other controls as covariates, such as wealth, education, income, etc., or exogenous geographic properties such as inherent soil fertility of occupied land, rainfall amounts for that respondent, proximity to water sources, etc. (all of which would be available through the GIS database). Including these other controls as well as rigorous econometrics to account for other potential biases can quantify the degree of confidence in the judicial system per respondent as a function of travel-time/accessibility to new courthouses.

The GIS is also a powerful display tool, and can be used to produce highly accurate maps of respondent spatial distributions, respondent geo-locations color-coded by variation in measured confidence indices, spatial displays of the variation in confidence by proximity to new courthouse locations. In addition, spatial statistical summaries can be generated and can provide useful insights into the relationship between judicial confidence and new courthouse impact, such as judicial confidence by zones of courthouse proximity (for all new courthouses), or the "catchment" or approximate spatial extent of an increase in judicial confidence for a particular courthouse, etc.

Growth of MSMEs. The tracking of growth of MSMEs, in terms of their numbers, employment, revenue, profit, and similar indicators is also desired to be examined with respect to the ATJ project.

In terms of numbers of MSMEs, data on this measure are available from a number of sources:

- ***EMICoV survey.*** The survey, currently being conducted and scheduled for future updates, includes a module that obtains data on both formal and informal enterprises associated with sampled households.
- ***DDI Survey of Establishments.*** In 2000, DDI conducted a survey on the formal and informal business sector and plans to conduct another survey in beginning 2007. While we need to obtain further details from DDI as to how this survey will be conducted, we believe it

includes an attempt to enumerate all business establishments as part of the sampling process for the survey itself, which collects information on revenues, employment, labor and input costs, and views on the business environment. DDI has stated that they intend to control and update their data by entering newly installed businesses, although the timing of this update is unclear. If the updating and survey were done periodically, then the DDI survey would represent a strong data source on the growth of MSMEs.

- ***Guichet Unique.*** The *Guichet Unique* can also provide data on the growth of MSMEs as new businesses register each year. Additionally, limited data on their growth can be obtained through the information they are required to file to renew their business card if they plan on buying any goods from suppliers, importing other goods or exporting their products. We may also be able to use these data to determine how many registered businesses cease operation each year (although the failure to renew the business card does not make it certain that an establishment has closed).

Again, we can use the availability of GIS data to analyze the data spatially to see if there are any patterns that emerge in terms of MSME growth that relate to the construction of new courthouses. However, it must be noted that it will not be possible to make any causal inference between the ATJ project and changes in MSMEs; at best, we will only be able to describe the covariation between the two.

E.3.3 Data Sources

As described above, we recommend obtaining data for the ATJ evaluation both through primary data collection and from secondary data sources. We examine the issues related to each data source below.

Judicial Case Data. We have only had the opportunity to make a preliminary visit to the Archive and the office of the court clerks, so additional investigation is required to determine more precisely the content of the files and the consistency of the data they contain. Similarly, the compatibility of the data in the paper files and that which will be contained in the EU case management system also must be verified before they commence their civil case database processing in early 2008 (although we have been assured by the EU that their system will be comprehensive and will contain the data required for the analysis).

Another issue to be resolved with the data on cases currently in process is how this data will be updated as cases are completed, specifically whether we should use a "flow" or "batch" approach (i.e., whether the data on cases should be updated as they are resolved, or whether the data on cases currently in progress will be updated periodically or at the end of the evaluation period). Doing the updating on a limited periodic basis would imply lower costs, but also requires that we will be able to quickly locate specific cases once they have been moved to the Archive (a condition that does not now seem to obtain).

Additionally, issues related to access to the data and confidentiality have not yet been solved and these must be resolved in a manner that satisfies requirements of the Court and IRB guidelines. The Ministry of Justice has informed us that for cases still in process, we will not be able to have access to those files. However, the question remains if someone with access to these files has the authority to enter them into the database system.

EMICoV Survey. The EMICoV informal sector module (as well as the employment module that precedes it) has a screening question that asks if the respondent either works independently or owns his or her own business and if the response is positive, the module is administered. It is unclear to what extent the EMICoV survey module on informal sector activity will capture the universe of registered MSMEs, because it was just beginning to be administered in August of 2006, but potentially, EMICoV could yield a fairly broad source of data on the informal sector. We will need to assess how well the module is capturing informal sector activity.[42] The EMICoV survey has the advantage of regular update surveys (expected to take place on an annual basis) being planned in the future. However, the content and exact timing of these update surveys has not yet been determined by INSAE.

DDI Survey of Enterprises. This survey was last carried out in 2000 with the expectation that it would be administered regularly, every two to three years. It has been postponed since then for budgetary reasons, but is now scheduled for 2007. The survey at present includes questions

[42] One particular concern is how many observations the informal sector module will yield. If the number of households that respond positively to the screening question is small, then the usefulness of the EMICoV survey as a source of potential matches for the comparison group is greatly reduced because the matching will be poorer than it would likely be with a larger (and more varied) number of observations.

regarding firms' production, sales, employment, and marketing. DDI expressed willingness to collaborate and if useful to add questions to their questionnaire, for example, on investor confidence in the courts or on access to credit. The timing of this survey may be a little late to serve as a baseline—it is now scheduled for mid to late 2007—but it has been suggested that DDI may be open to accelerating the process.

The DDI survey uses as its sampling frame a census also carried out by DDI in which interviewers walk through neighborhoods identifying businesses and asking basic questions on registration, address, sector, level of capital. If respondents are not willing to participate, the interviewer at least notes their identification.[43]

The Ministry of Justice Satisfaction Survey. This survey was last carried out in 2005 and will continue to be administered by the Ministry of Justice every two years. Currently, a quality control test is being conducted to verify the reliability of data that was collected in 2005. Continued discussion with the Ministry of Justice is needed to determine how the sample is chosen, the questions asked and the timing of the next survey.

Guichet Unique. Similar issues obtain with the *Guichet Unique* as with the judicial case data. Additional research is required to determine more precisely the nature and content of the data and an assessment of its quality with respect to consistency of reporting and accuracy. As with the judicial case data, issues related to access to the data and confidentiality have not yet been discussed with the *Guichet Unique* and these must be resolved in a manner that satisfies requirements of the *Guichet Unique* and IRB guidelines.

GIS. As with the Access to Land (ATL) project, consideration of geographic conditions and especially variation in village accessibility to markets, cities or locations providing credit could greatly increase the accuracy of the analysis. This will be possible using the GIS that is proposed to be constructed to support the Access to Land project (see Annex A). The same geo-database will be used for all three Benin MCC components being evaluated (ATL, AFS, ATJ), and thus additional marginal costs will be low while potential synergies could be high. Most of the data that will be

[43] We still have an outstanding question to DDI on whether the sampling frame is a comprehensive census, or a census based on randomly selected localities, and about the size of the sample.

entered into the GIS database will be entered once to support all three MCC Benin projects. These data will include geo-locations of EMICoV and business survey respondents, geo-locations of villages/towns, ZDs, and communes, data on digital road networks (including approximate road quality/speed) allowing calculation of access measures for respondents or villages to cities, markets, financial centers, courthouses, etc. In addition, Benin-wide spatial distribution of exogenous geographic variables such as rainfall levels, soil fertility, topography, river/stream and hydrology data, etc., will also be entered into the GIS (from Benin government and other sources, see Annex A), allowing precise calculation of village point estimates for these variables, as well as ZD, commune or *quartier*-wide aggregated estimates for any variables.

E.3.3 Evaluation Timing and Sequencing

Assuming we follow the approach outlined above, the timing of most activities for the ATJ project will be driven by the schedule for the collection and availability of data from other sources, some of which (such as the EU case management system, the EMICoV survey, the DDI survey of enterprises and the Ministry of Justice Satisfaction survey) will be decided primarily by other entities, and others (such as the case data in paper files and the *Guichet Unique*) which are presently available, but for which access and terms of use need to be negotiated.

Data from the first round of the EMICoV survey should begin to become available in the first quarter of 2007, The DDI survey of enterprises is currently planned for mid-2007, which implies that data will become available in late 2007. We believe that the next Judicial System Satisfaction survey results could be available by the last quarter of 2007. Assuming agreement can be reached with the Court and the *Guichet Unique* in the coming quarter over the use of their data, we estimate that the data from these sources could be available by the last quarter of 2007. In early 2008 data from civil cases will be entered into the EU system and should begin to become available in the first quarter of 2008. This would allow preliminary analysis to take place at the end of 2007 and early 2008. Depending on the availability of data updates from the EMICoV and DDI surveys, an additional round of analysis could be completed in 2009 before final results were developed at the end of the task order in 2010.

E.4 Other Evaluation Design Issues

Data Quality. During the month of August, NORC staff accompanied INSAE's data collection team to a number of rural and urban sites. NORC witnessed the data effort and was satisfied that INSAE and the team are following accepted standards of field collection. INSAE has secured well educated staff and implemented a comprehensive training program that prepares them for the field interviewing. It will also be important to assess other elements in the data collection process including the receipting, transfer, and data entry of the questionnaires to ensure that the highest quality is maintained at each of these steps.

INSAE has indicated that it will perform the necessary reliability checks of data collected in each round. This analysis and conclusions should be provided to NORC at regular intervals so that we can evaluate the quality of the EMICoV data on an on-going basis during the evaluation.

During the August field visit, it became clear that some key modules and questions were being refused by a number of respondents. We could not determine in a short visit how widespread this problem was. Specific items on informal sector activity and investment are potentially sensitive but critical to determining if resources are being invested. It appeared that this might be a greater problem in urban ZDs than in rural ones. Since respondents may not feel secure in the confidential nature of their responses, we fear that this could result in low response rates for these key items; jeopardizing key elements of the impact evaluation. It will be important for interviewers to gain cooperation of respondents and convince them of the confidentiality of their responses. NORC can provide assistance to INSAE to help address this problem.

While the DDI at the Ministry of Industry has voiced its interest in collaborating with the evaluation, and their survey seems to be appropriate, we have not yet had any opportunity to assess the quality of the data (or underlying data collection procedures and processing) for the DDI survey. The DDI survey instrument that we have reviewed (which appears to date from 2005 and so may not be the questionnaire planned for use in 2007) does not contain any questions regarding enterprises experience with/views of the justice system, so questions relevant to the ATJ project evaluation would need to be developed for inclusion in the 2007 survey.

Additionally, the Ministry of Justice survey would be crucial for the evaluation, however, some questions continue to remain unanswered and the quality of the data will also have to be assessed. Similarly, we have also not yet had an opportunity to look in detail at the data in the paper case files, the EU case management system (as planned),and the *Guichet Unique* to verify if the assumptions we have made above concerning their use in the evaluation are accurate.

Schedule of Follow-up EMICoV, DDI and Ministry of Justice Surveys. INSAE has yet to determine which modules will be administered in subsequent rounds. INSAE has suggested that several of these modules may be skipped in subsequent rounds. If this occurs, critical data for the ATJ evaluation may be lacking. Agreement must be reached on this issue soon in order to determine how much of the data needs for the evaluation can be derived from future EMICoV surveys and how much data will need to be collected directly as part of the evaluation.

Similarly, discussion is needed with DDI as to the timing of the future rounds of the DDI enterprise survey. In particular, we will need to discuss whether the DDI survey will be updated annually and if the timing of those updates could be coordinated with the EMICoV updates. We would also need to discuss the DDI's willingness to include questions related to satisfaction with the justice system in the survey of enterprises.

Follow-up is also needed with the Ministry of Justice to determine the timing of the next satisfaction survey, how respondents are selected and to obtain a copy of the survey instrument to see what questions are asked.

E.5 Implementation Schedule and Next Steps

While this design report provides a roadmap for the next steps for the evaluation, as we note above, there still remain several key issues to be addressed with MCC, MCA-Benin, and potential partners in the ATJ evaluation. In summary, these include:

- Further discussion with DDI on the suitability of the DDI survey of enterprises for use in the evaluation. These discussions will include the coverage and structure of the sample used for the survey and the content of the survey itself.

- Agreement with both INSAE and DDI on the scope and timing of follow-up EMICoV and DDI enterprise surveys, respectively, amendments to the current EMICoV modules and DDI survey to achieve maximum comparability of data across the surveys, and appropriate means for assuring data quality.
- Developing a better understanding of the details of the data available in the paper case files and EU case management system in the Court and data available through the Guichet Unique (particularly with respect to the annual filings by registered enterprises).
- Continued discussion with the Ministry of Justice regarding cases in process and who has access to enter these files into the database and additional questions concerning the Satisfaction Survey on the judicial system.
- Obtaining needed GIS data from government or other sources.[44]

Given the difficulties faced to date in obtaining follow-up information from Benin since the August field visit, we would recommend scheduling a follow-up field visit for the resolution of these remaining issues.

Table D.3 — ATJ Evaluation Implementation Timetable

Task No.	*Activity*	*Period*
Finalize Design		
1	Resolution of remaining issues from Design Report (Benin field visit)	Q1/2007
Assess Data Sources		
2	Review analysis of EMICoV first passage	Q1/2007
3	Review sampling approach and frame for DDI census	Q1/2007
4	Review sampling approach and frame for Ministry of Justice survey	Q1/2007
5	Reach agreement on timing of follow-up EMICoV and DDI surveys	Q1/2007
6	Review data to be collected by EU court information system and ADR case information system	Q1/2007
Data Collection		
7	Provide input for questions to add to 4th passage of EMICoV	Q1/2007
8	Provide input for questionnaire and sampling of DDI survey	Q1/2007
9	Data collection for DDI survey of business establishments	Q2/2007-Q3/2007
10	Data collection with archived and current paper case files	Q2/2007-Q4/2007
11	Obtain initial data from ADR case information system	Q3/2007-Q4/2007
12	Obtain initial data from EU court information system	Q1/2008-Q2/2008
13	Updates of EMICoV, DDI, Ministry of Justice, ADR/courts case data	2008, 2009

[44] This is also being done as part of the ATL design development.

Table D.3 — ATJ Evaluation Implementation Timetable		
Task No.	***Activity***	***Period***
12	Review data from each updated survey	Q4/2007, Q4/2008, Q4/2009
Impact Evaluation Analysis		
15	Baseline data/analysis report	Q4/2007
14	Interim impact analysis reports	Q4/2008, Q4/2009
15	Final report on impact evaluation	Q3/2010

ANNEX A

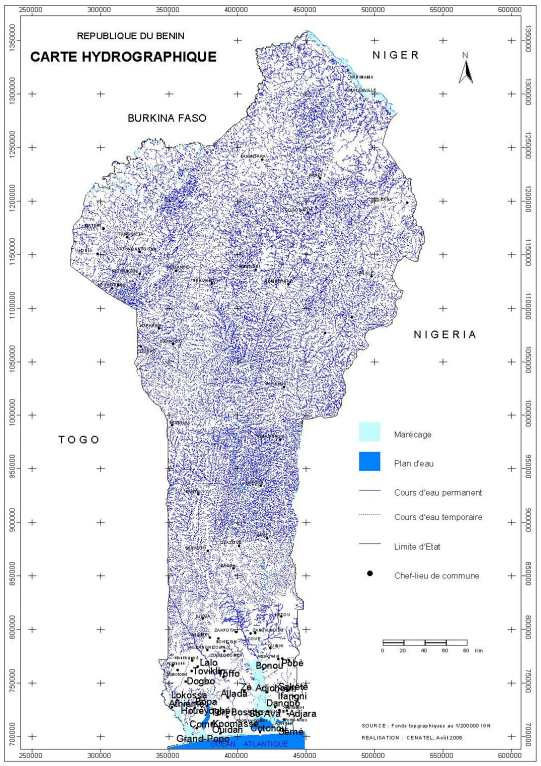
CONSIDERATION OF EXOGENOUS GEOGRAPHIC AND ACCESSIBILITY CONDITIONS USING GIS

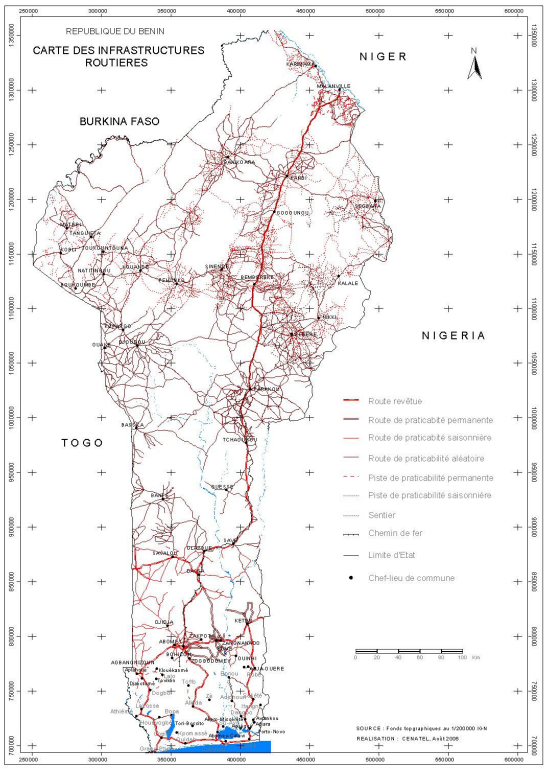
Unlike most socioeconomic data, much of the GIS physiographic data that will be used for the Benin impact evaluation has the advantage that it can be collected in one time step. Presumably data on soil fertility and topography, for example, is not varying over time, or not varying significantly within the scope of this analysis to be significant. In addition a single map on topography from CENETAL (see below) covers the entire country, and thus the spatial extent of the data can also be collected in one time step. The result is that, in comparison to survey data (for example), GIS data acquisition costs and effort are relatively quite low, but with repeated marginal benefits for all analyses.

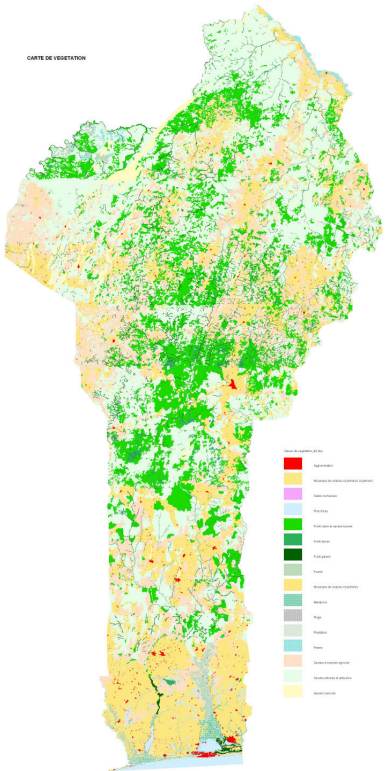
Geographic, Environmental and Road Network Data Sources in Benin

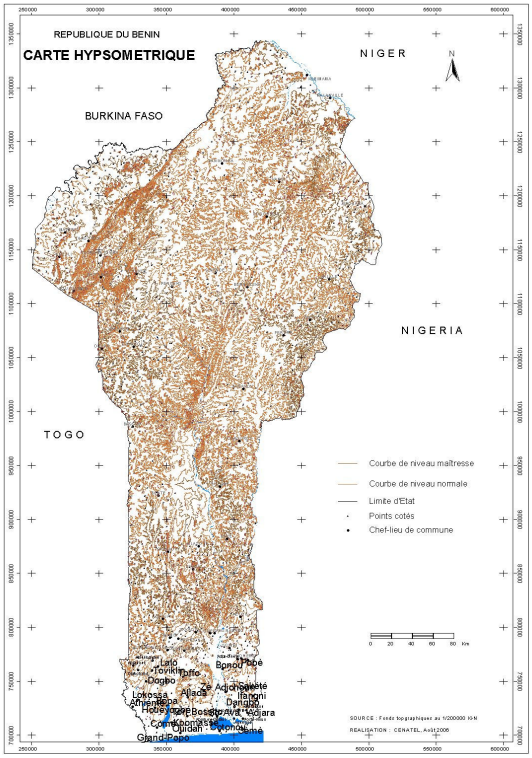
The use of GIS-calculated exogenous geographic and access variables for villages and communes depends directly on the ability to obtain sufficiently accurate digital geographic data for Benin and the study areas. A key goal of the evaluation trip to Benin in August, 2006, was to determine if a sufficient quantity of high-accuracy digital spatial data existed and could be obtained[45]. Based on our investigation, and the initial gathering of initial GIS data from CENETEL in Benin, our assessment is that sufficient data exists and will be obtainable to complete the methodology as described above. Key to the analysis will be data on village geo-locations, but these are being obtained through the use of Geographic Positioning System (GPS) units by INSAE, and will thus be highly accurate. Meetings and discussions conducted at CENETEL indicate that they possess GIS data at the national level for Benin that includes digital road network data, as well as data on vegetation and land cover, hydrology, topography, soils and rainfall. Graphical outputs of some of the CENETAL GIS data is shown below, including road network data (crucial for the calculation of village access indices), hydrology, land cover and topography.

[45] The accuracy and detail of available GIS data for most developing countries world-wide is improving dramatically. This is likely due to the extremely rapid evolution and development of the digital geo-spatial mapping industries world-wide, driven both by market demand and the spread of powerful and lower-cost technologies (including remote sensing instruments for data capture, and powerful desktop GIS software for data preparation and analysis).


REPUBLIQUE DU BENIN
CARTE HYDROGRAPHIQUE
NIGER
BURKINA FASO
NIGERIA
TOGO
Marécage
Plan d'eau
Cours d'eau permanent
Cours d'eau temporaire
Limite d'Etat
Chef-lieu de commune


REPUBLIQUE DU BENIN
CARTE DES INFRASTRUCTURES ROUTIERES
NIGER
BURKINA FASO
NIGERIA
TOGO
Route revêtue
Route de praticabilité permanente
Route de praticabilité saisonnière
Route de praticabilité aléatoire
Piste de praticabilité permanente
Piste de praticabilité saisonnière
Sentier
Chemin de fer
Limite d'Etat
Chef-lieu de commune


CARTE DE VEGETATION

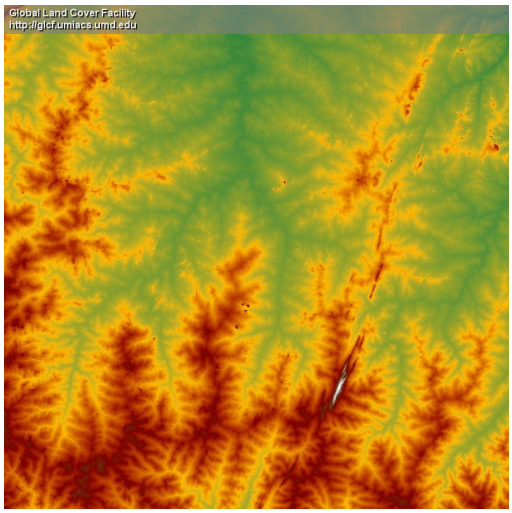

REPUBLIQUE DU BENIN
CARTE HYPSOMETRIQUE
NIGER
BURKINA FASO
NIGERIA
TOGO
Courbe de niveau maitresse
Courbe de niveau normale
Limite d'Etat
Points cotés
Chef-lieu de commune

CENETAL also has an extensive collection of paper maps for Benin for environmental variables. In the case that key data necessary for deriving village, household or commune exogenous geographic variables or access index values exists on paper maps but is not in digital GIS format, the data can be relatively inexpensively converted to digital format by private sector firms in the US or Africa that specialized in that conversion. Once the data is in digital format, it can be highly processed for accurate analysis. For example, topographic contour lines such as those on the map above ("Carte Hypsometrique") can be interpolated into a continuous elevation surface model (DEM or Digital Elevation Model) using GIS conversion algorithms. This DEM can in turn be used to improve the accuracy of the calculation of access indices to markets for villages, but including topographic variation in the calculation of approximate travel times.

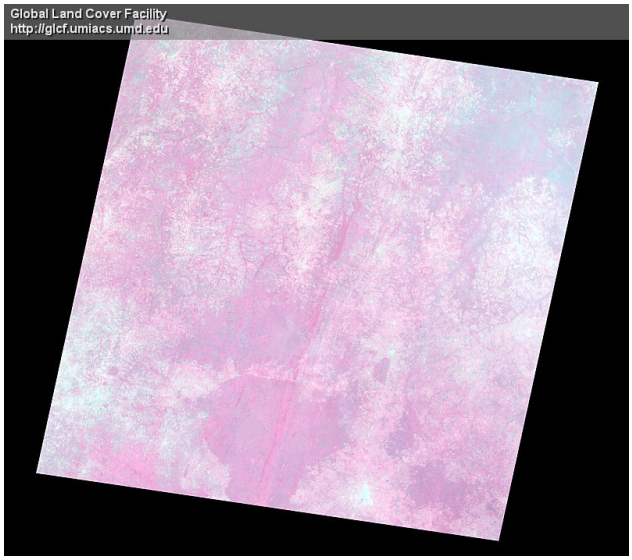
During visits in August, 2006, NORC also determined that CENETAL possesses an archive of both digital aerial photographs and satellite images for Benin, including Landsat, Spot and Quickbird satellite imagery. This imagery can be used to extract GIS digital data for road networks, village (or even household) locations, and land cover, and provide a strong cross-section in terms of both spatial and spectral resolution (Quickbird features powerful spatial resolution, allowing for the delineation of individual buildings on the ground, while Landsat and Spot provide powerful spectral resolution, allowing for the extraction of land cover across wide areas). NORC team members are experts in the use of both GIS and in the processing of satellite imagery and aerial photographs for land cover extraction. In addition to CENETAL, IGN also has an archive of GIS data.

In addition to spatial data at Benin agencies, extensive GIS and satellite image data for Benin is available for free from numerous online sources[46], including highly detailed and accurate elevation models, road network data and satellite images covering the entire country. Below is an example of a digital elevation model obtained online (from the Global Land Cover Facility) as well as a graphic showing its "footprint" for Benin (elevation pixels are color coded by elevation value):

[46] Such as the Global Land Cover Facility (GLCF) at the University of Maryland (see http://glcf.umiacs.umd.edu/index.shtml), the US Geological Survey (see http://edc.usgs.gov/geodata/) and the Geography Network (see http://www.geographynetwork.com/).



Also, satellite imagery is available for the whole country from online archives. Here is a preview and then a graphic showing a Landsat Thematic Mapper image and "footprint":



Finally, there exists an active private market for GIS data (highly accurate geo-spatial data is needed constantly by engineering firms, highway contractors, large-scale irrigation construction, urban and water planning, etc.) and extensive and accurate GIS data for Benin is available from private sector contractors if necessary[47]. These sources collectively should be sufficient for the analysis.

[47] Such as the French company GeoConcept (based in Paris) that sells GIS data for Benin (see http://www.geoconcept.com/index-en.php3).

Coordination and Data from MCC Benin Cadastre Program

The NORC evaluation team is currently in on-going communication with MCC personnel regarding the development of the MCC Benin cadastre program, which includes programmatic elements such as the collection of high-resolution digital spatial data to support cadastre mapping. Continuing coordination with MCC during the development of the cadastre program design should facilitate mutual synergistic support between the NORC evaluation Benin GIS and the MCC Benin cadastre program. For example, GIS physiographic data imported, cleaned and collected by NORC from CENETAL can be exported directly into the MCC cadastre GIS, and also digital spatial data on administrative or land plot boundaries, or digital airphotos, created for the cadastre can be imported directly into the GIS.

Calculation of Accessibility Conditions Using GIS

Spatial economic theory as far back as the famous Von Thunen (1826) land rent model has been based on the assumption that spatial access to markets, controlled by transportation costs, is crucial for economic development. In principle, improved access to consumer markets (including inter-industry buyers and suppliers) will increase the demand for a firm's products, thereby providing the incentive to increase scale and invest in cost-reducing technologies. Workers and firms would benefit from gaining access to an agglomeration as they could expect higher wages and to have access to a larger set of employers. Furthermore, access to markets or economic city/town agglomerations can determine if a household is able to afford the cost of shipping products for sale, earning potentially higher wages in agglomeration centers, or gaining access to information spillovers or technology advances, further reducing costs.

There is a rich body of literature on the benefits to firms from gaining improved access or proximity to other firms in the same industry (Henderson, 1974 and 1988; Carlino, 1978; Selting et al.,1994).

In general, “access” to markets is determined by the household’s or village’s true cost of traveling to or accessing market centers. This could include the cost of transporting goods for sale, transporting (back to the village) key inputs for production or consumption, or the cost of transporting people for migratory or more permanent employment. Thus, effective access to urban markets also depends on

the willingness and ability to afford transport costs, and these in turn are directly a function of road quality as well as actual measured road distance, topography, climate, rivers or any other potentially inhibiting (and thus more costly) exogenous geo-physical barriers.

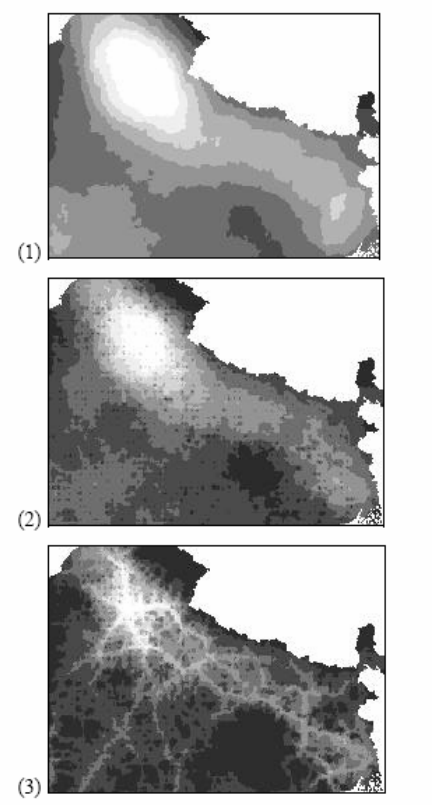
The classic gravity model which is commonly used in the analysis of trade between regions and countries states that the interaction between two places is proportional to the size of the two places as measured by population, employment or some other index of social or economic activity, and inversely proportional to some measure of separation such as distance. Following Hansen (1959)

$$I_i = \sum_j \frac{S_j}{d_{ij}^b}$$

where I is the "classical" accessibility indicator estimated for location i (for example, a village), S is a size indicator at a market destination j (for example, population, purchasing power or employment), and d is a measure of distance (or more generally, *friction*) between origin i and destination j, while b describes how increasing distance reduces the expected level of interaction. Empirical research suggests that simple inverse distance weighting describes a more rapid decline of interaction with increasing distance than is often observed in the real world (Weibull, 1976), and thus a negative exponential function is often used.

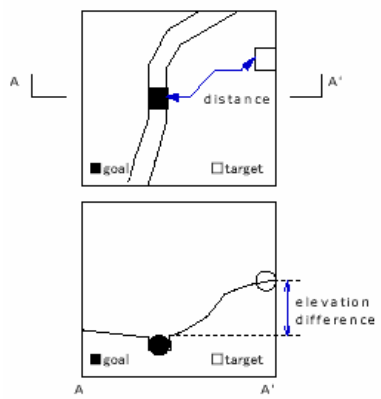
There are several options for developing accessibility indicators depending on the choice of distance variables used in the computation. These include: (a) indicators based on "straight-line" or Euclidean distance; (b) indicators incorporating topography; (c) indicators incorporating the availability of transport networks; (d) indicators incorporating the quality of transport networks; and (e) movement across a "cost surface". A better alternative is to use actual measured distance along road networks as the basis of the inverse weighting parameter and to incorporate information on the quality of different transportation links. Feasible travel speed and thus travel times will vary depending on each type of network link. A place located near a national highway will be more accessible than one on a rural, secondary road. The choice of the friction parameter of the access measure will therefore strongly influence the shape of the catchment area for a given point—i.e., the area that can be reached within a given travel time. This, in turn, determines the size of potential market demand as measured by the population within the catchment area.

The figure below illustrates these points using an accessibility surface for the Northern Indian Gangetic plain using three measures of market access: (1) based on Euclidean distance, (2) network distance, and (3) network travel time. It is clear that indicators based on (1) and (2) overestimate potential market area. The variation in infrastructure quality (3) between regions leads to a more realistic representation of the structure of market areas. Thus, incorporating the quality of the transport network is important in assessing the potential market integration.



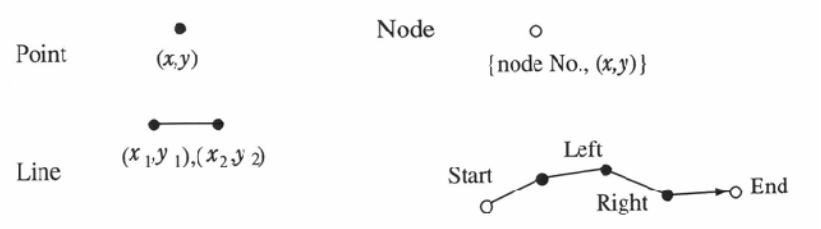
In studies related to agglomeration economies and economic geography (e.g., Hanson, 1998), the distance measure of choice is usually the straight-line (Euclidean, or "as-the-crow-flies") distance, which has the advantage of computational simplicity. However, this assumption of uniform (isotropic) plane is clearly unrealistic, particularly in countries where topography and sparse transport networks of uneven quality greatly affect the effort required to move between different parts of the country. Such an access index takes no account of the fact that hills and mountains greatly reduce travel times and greatly increase travel costs. Nor does it take into account the fact that people and goods move along road networks – not across a uniform plane. If data on topography is obtained (either from contour lines digitized from paper maps, or from spot samples taken on the ground by surveyors, or from airborne or satellite instruments) it can be converted using GIS algorithms to a continuous elevation

surface. In that case, distance *across topography* can be calculated: the GIS calculates Euclidean distance, but then further calculates the actual distance on the ground considering topographic variation. This is partially illustrated by this graphic:

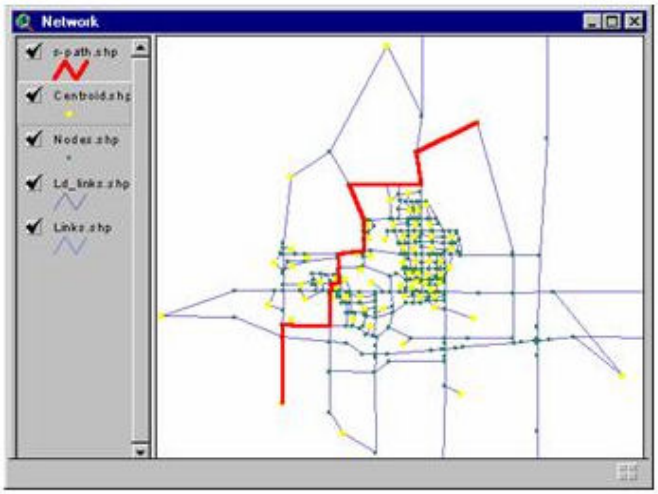


Where distance is measured both across the two-dimensional x,y surface, but also across the topographic z surface, calculating actual distance traveled (again in meters, kilometers, etc.).

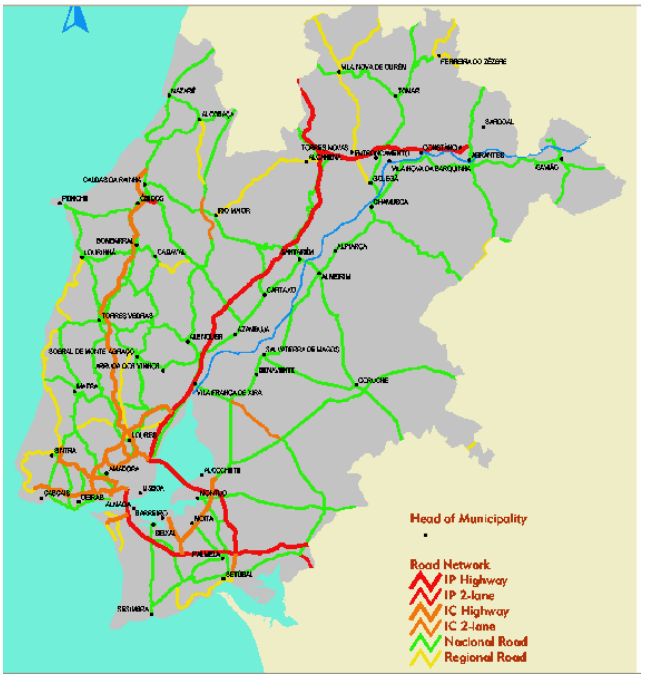
A far better alternative, however, is to use actual measured distance along existing road networks, considering the fact that goods and people move predominantly along infrastructure networks. This can be accomplished by obtaining an accurate digital GIS road network. Such a road network has all roads digitized into GIS digital "vector" objects. That is, rather than simply a graphical image of the roads, the road network is actually made up of many individual line segments, connected to each other at the end points, which are called "nodes". Each individual road segment in the larger network is "seen" by the GIS as an individual digital object. The GIS can calculate the exact length, direction and curvature of each line segment (just as it can for polygonal objects). These graphics illustrates the underlying road network structure, which is technically referred to as "vector line topological structure":



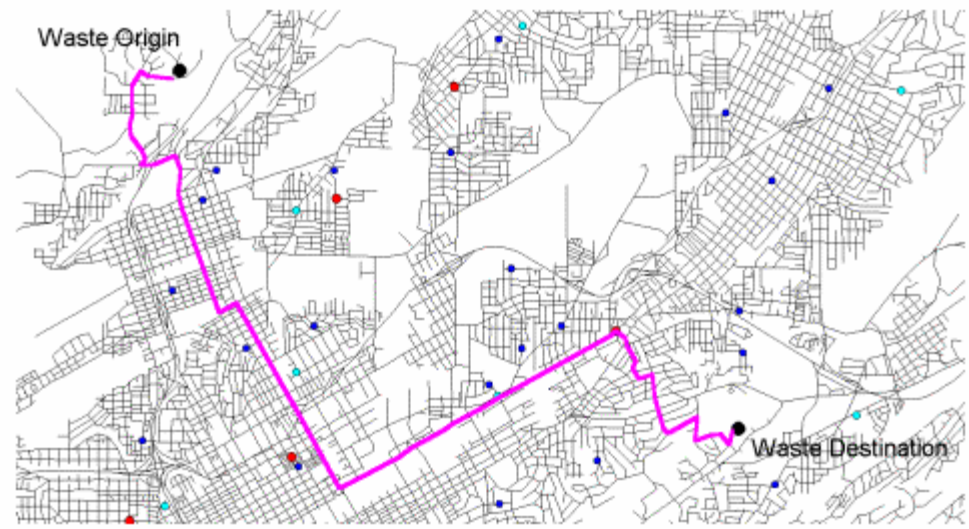
The GIS keeps tract of the exact geo-location of each node connecting linear road segments, as well as the exact curve of each line segment. Thus the GIS can calculate precisely the exact distance along each segment in any desired unit (such as meters, kilometers, etc.). Thus, using advanced algorithms, the GIS can calculate travel distance *through* the road network from any node to any other node. Other algorithms will pick a "shortest path" through the network to get from node X to node Y, minimizing travel distance, as in this graphic:



Here the GIS has simply found the shortest path through the road network assuming that all road network segments are equal in terms of road quality or road speed. However, data on road quality or road speed of each individual road segment is often available, and can be entered into the GIS database and attached to each road segment (in fact any amount of information on road segments – or any other object in the GIS – can be entered into the database, such as data on road segment names, date of paving, cost per segment, number bridges per segment, etc – all of this information is kept track of in the GIS database). If data on road quality is available, then approximate road speeds can be estimated. Typically, road maps categorize roads into categories as in this map from Portugal:



For example, if a road is categorized as "one-lane paved", then an approximate road speed of 45 miles per hour could be assigned to all road segments with that categorization. Once categories of roads are assigned approximate road speeds, then travel times through the road network *considering road speed/road quality* can be calculated. This is a simple calculation: road length divided by road speed. For example, if the road segment is 50 kilometers long, and the road speed of that segment is 25 kilometers per hour, then the travel time (or travel "cost" if the definition of cost here is time) would be 50/25 or 2 hours of travel time. Often, this results in a different "least cost" or "least time" (if the "cost" is in terms of speed) pathway than the minimum distance pathway along all road networks. For example, it may be quicker in terms of time/cost to drive onto a highway and then exit to get to a destination than to travel along intermediate roads even though they provide a more direct link. Thus, the pathway of minimum distance may not always be the same as the pathway of minimum time or cost. This graphic shows the *fastest* route through a network from one destination to another, rather than the minimum distance route:



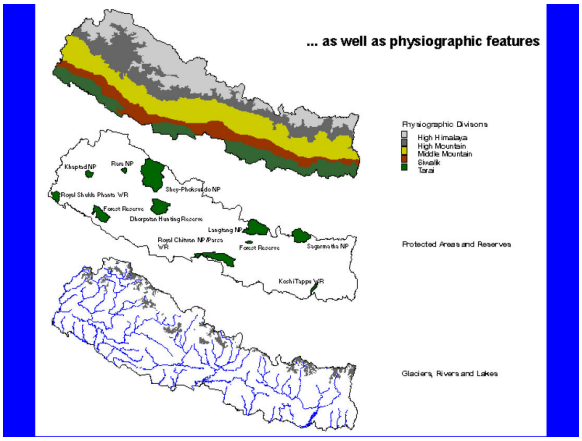
Topographic information could be combined with road network speed information, so that the road network segments are weighted by elevation or slope. For example, one might burn less gas or put less stress on a truck (lower "cost") to drive around a mountain than across it, even though the minimum distance pathway is across the mountain. In this case, the path of "accessibility" would likely be around the mountain.

Calculation of Accessibility Conditions Using GIS

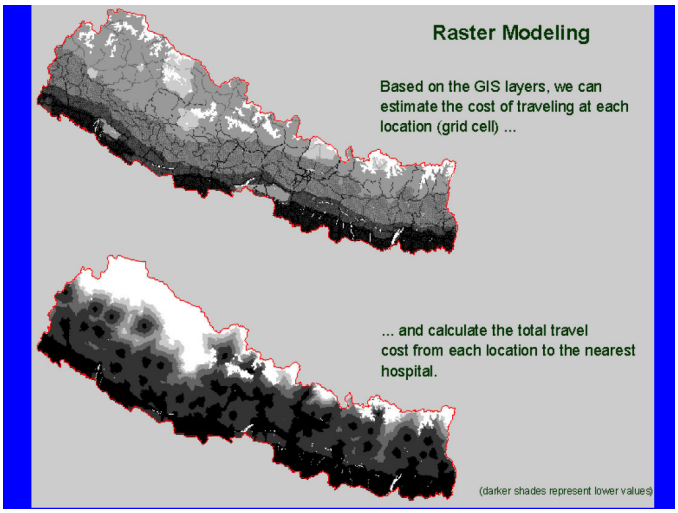
Calculating accessibility or distance through a road network is sufficient if one is only interested in locations that are on that road networks. However, what about locations for which accessibility must be calculated that are not located on the road network? Or, what about calculating travel time/cost across areas for which there are no road networks? There are a variety of approaches to this, but it is usually accomplished by creating a GIS *cost surface*.

The first step in the creation of a cost surface is to convert all the GIS data into what is known as raster format data, which means a regular grid of pixels. All input layers, therefore, such as elevation, roads, village locations, landcover, etc., are converted to grids of regular pixels of equal size, with each pixel having a value that corresponds to the range of values in its respective input layer. For example, the elevation raster would have pixel values reflecting the elevation of that pixel. The landcover raster would have pixel values

corresponding to the specific landcover that pixel occupies in space[48]. All the individual input layer rasters would then be combined algebraically to create a single output cost surface. In this example in Nepal several individual input layers describing features that could influence travel time are rasterized:



And then combined into a single cost surface (darker shades represent lower travel costs):



The cost surface can be created reporting to many different cost units (distance, time, currency, or any other unit) and by many different methods, and it establishes the impedance for crossing each *individual* cell. If we then want to incorporate into the cost surface road network travel time estimates from our road network (including data on road quality or road speed), then that network would also be rasterized with the approximate road speed for each road segment mapped to the

[48] There are various routines for "resampling" or interpolating values from the input layer to the output raster, and their respective measures of error have been carefully delineated in the geography literature.

corresponding (spatially overlying) grid cell. Then, the averaged road speed would be used to calculate the cell crossing time by using the following equation:

$$CCT = \frac{P \times 60}{TS \times 1000}$$

where:

CCT - Cell Crossing Time (minutes)
P - Pixel Size
TS - Traveling Speed (Km/h)

For example if one is traveling in a 2-lane highway with average road speed of 80 kilometers per hour through pixels 100 meters across, the result is the following:

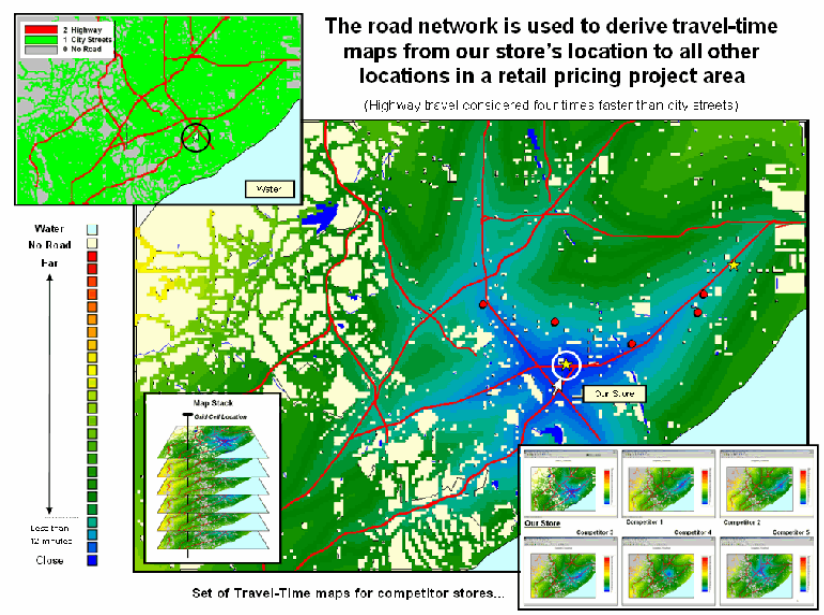
$$CCT = \frac{P \times 60}{TS \times 1000} = \frac{100 \times 60}{80 \times 1000} = \frac{6}{80} = 0.075$$

or .075 minutes (4.5 seconds to cross 100 meters). The following table shows the resulting pixel travel times for a variety of potential road speeds:

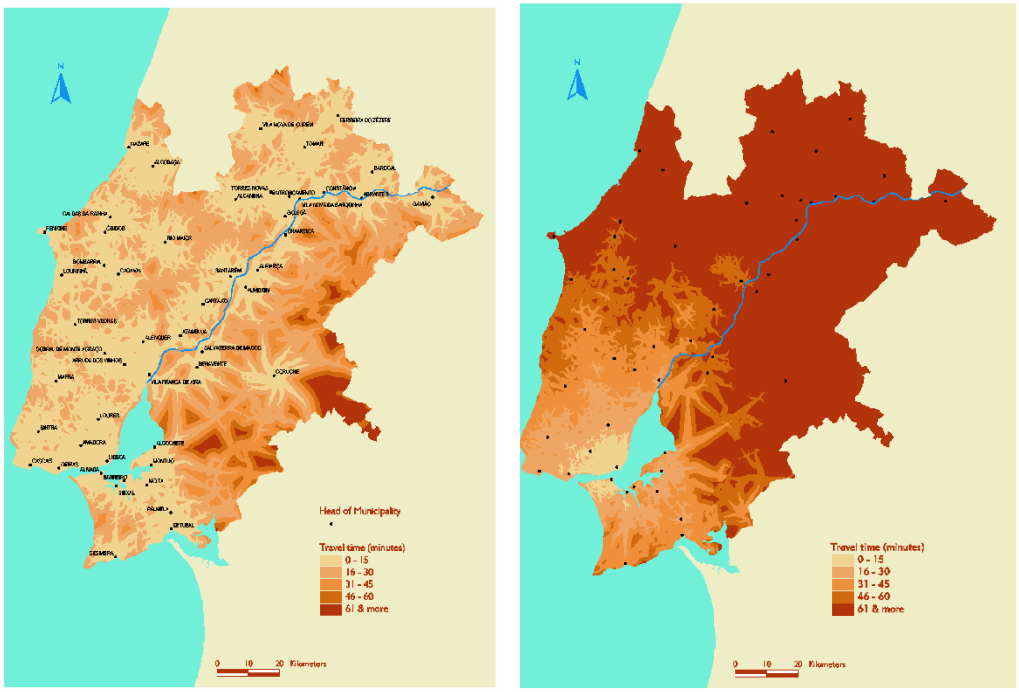
Road Category	**Average Speed**	**Cell Crossing Time** (minutes)
IP highway	110	0.0545
IP 2 lane	80	0.0750
IC highway	110	0.0545
IC 2 lane	70	0.0857
National Road	60	0.1000
Regional Road	55	0.1091
Municipal Road (former national)	50	0.1200
Municipal Road	50	0.1200

For land outside the road infrastructures, pixel crossing speeds are estimated based on the combination of input layers going into the overall construction of the cost surface. For example, a base "walking time" of 6 kilometers per hour might be established, but be altered depending steepness of slope, elevation, landcover, etc. Other considerations are important, such as the fact

that highways can typically only be entered or exited at specific points, whereas dirt or smaller paved roads could theoretically be entered at any point. Specific GIS algorithms exist to model these situations. Feasible travel speed and thus travel times will vary depending on each type of network link. A place located near a national highway will be more accessible than one on a rural, secondary road. The graphic below illustrates an example of an accessibility model run across a cost surface utilizing road network information:



Using the measures described above, accessibility indices can be calculated for all villages, communities or indeed for all points in space. For example, the figures below display access computed to the nearest city or town (on the left) and then access to one major city (on the right):



Integration of Physiogeographic Data to Weight Cost Surface Calculations

While measuring distance along road networks incorporating data on varying road quality or varying road speed is a far superior method than measuring access "as the crow flies" (Euclidian distance) or even along road networks without considering road quality, the accuracy of the computed access indices can be further enhanced by incorporating weights that reflect further variable that impede travel, adding travel cost and time. For example, topography (as well as slope angle) is an extremely important variable that could dramatically alter travel times and costs, but might not be considered at all if only road network distance and road quality were considered. While digital data on a road network might indicate that a particular stretch of road was paved at high quality, with an official speed limit of 80 kilometers per hour, nonetheless in reality that stretch might involve movement up and down steep hills, in effect slowing travel time and increasing travel cost beyond what is measured simply by the road network data. Furthermore, a flat stretch of road in a low-lying area that rarely encounters debilitating weather such as snowstorms might overall be much easier (and cheaper) to travel than a similar flat stretch of identical road quality located at high elevations. On the latter, travel may frequently be inhibited by severe snow or ice, thus dramatically increasing travel costs.

Other important physiographic factors can affect actual travel costs and times, including land cover, climate, rainfall amounts, and the presence of lakes, rivers, streams and glaciers, which may periodically overflow, or swell during certain times of the year. Furthermore, a road network map may not indicate that certain areas are restricted because they are protected – either for conservation or military purposes, for example – and thus travel through them is impractical. In that case, the road network will need to be digitally altered to reflect the actual travel routes.

By the same token, certain physiographic factors can provide exogenous drivers of village economic productivity, such as inherently fertile soils which would result in higher agricultural productivity, or favorable rainfall patterns or climate, etc. Villages located in areas with good access to clean water, or with less intimidating (and costly) topography for villagers to drive and navigate, might have an inherent (exogenous) advantage over other villages with very similar socioeconomic measures. Consideration of all these variables will be crucial in our PSM process, as villages will need to be "matched" as accurately as possible to measure changes in "before" and "after" economic levels. In their absence, for example, an economic increase in one village over another might be falsely attributed to superior program benefits, rather than to superior soil fertility which may be the true driver. Or the reverse could occur, blunting the effective measurement of true, positive program benefits.

Ignoring such physiographic conditions for villages could also ignore another key element of "accessibility": the fact that market access may be more valuable for some communities than for others. For example, a community with inherently poor soil fertility may benefit more from access to a fertilizer market than a community with inherently rich soils, but having the same level of access as measured by road distance, quality and even topography.

Consequently, the gathering and obtaining of data on physiographic conditions and the integration of this data into the GIS is necessary both to enhance the calculation of village access indices (to delineate the gradations of "treatment"), but also to provide important variables characterizing villages for the PSM process. These variables are not available in the EMICOB socio-economic survey and census data, but are or will be available from large scale GIS datasets (many of which have been already obtained – see below).

Once these data are assembled in the GIS, along with geo-locations of impacted communities (such as villages), then the GIS can quickly "map" to each community variables describing the respective physiographic conditions for each. Also, these data inputs can be used to weight the road network segments, as well as the areas of land leading to the nearest road network (in the case of villages that have no road network connection, if these exist). The graphic below depicts the creation of a GIS travel cost surface (which can be used to calculate accessibility) after incorporation of physiographic data on topography, rivers, streams, glaciers, and protected areas.

These variables can greatly enhance the PSM models and comparisons, because they describe exogenous conditions unique to each village. Inclusion of these variables in the propensity score models will dramatically alter the score for an individual village, and thus greatly improve the overall accuracy of the PSM process.

ANNEX B

PROPOSED REVISED TASK ORDER BUDGET