Transport Infrastructure and Welfare in Nigeria

Rubaba Ali,
Alvaro Federico Barra,
Claudia N. Berg,
Richard Damania,
John Nash, and
Jason Russ

The World Bank

August 25, 2014

ABSTRACT

Transport infrastructure is deemed to be central to development and consumes a large fraction of the development assistance envelope. Yet there is debate about the economic impact of road projects. This paper proposes an approach to assess the differential development impacts of alternative road construction and prioritize various proposals. Recognizing that there is no perfect measure of economic well-being a variety of outcome metrics are used including: crop revenue, livestock revenue, non-agricultural income, the probability of being multi-dimensionally poor, and local GDP for Nigeria. While our measure of transport is the most accurate possible, it is still endogenous due to the non-random placement of road infrastructure. We address this endogeneity using a novel instrumental variable developed for this paper, termed the natural path. This IV, which measures the time it would take to walk to the nearest market given the terrain and absent any roads, is an improvement over the Euclidean distance IVs typically found in the literature since it more accurately captures what straight-line instruments attempt to estimate, that is, the most logical route connecting two points without taking into account other, bias-causing economic benefits. We find that reducing transportation costs will increase crop revenue, non-agricultural income, wealth index, and local GDP. The probability of being multi-dimensionally poor will decrease while there is no impact on livestock revenue. These findings are robust to relaxing the exclusion restriction, following Conley et al (2012). We demonstrate how to prioritize alternative road programs by comparing the expected development impacts of alternative NEPAD projects.

Table of Contents

List of Tables

1. Introduction

Governments and donors in Sub-Saharan Africa have devoted considerable resources to the construction and rehabilitation of roads. An emphasis on transport infrastructure is also evident in the lending pattern of the World Bank, which commits a larger share of resources to transport infrastructure than education, health and social services combined (World Bank 2007). Total transport commitments in fiscal year 2013 amounted to US$5.9 billion and rural and inter-urban roads remained the largest sub-sector with 60 percent of lending in FY13 (US$3.2 billion) (Transport overview 2013). The rationale behind these investments is self-evident. Roads, while expensive, facilitate the creation of, and the participation in, markets and are deemed to be central to development. Africa has the lowest density of roads in the world, with 204 kilometers of road per 1,000 square km, nearly one fifth the world average, and less than 30% of the next worst region, South Asia. Starting from such a low base, the potential for growth due to improvements in transportation infrastructure is presumed to be especially large in Africa.

However, the existing body of research about the impact of roads on economic well-being remains ambiguous, partially because it is hard to disentangle cause and effect. There is even less evidence on where investments might be the most transformative in creating new opportunities to link producers to markets. Given limited resources, there is a need for selectivity in deciding what investments should occur and where these should be located. This paper aims at tackling these issues by drawing on, and improving upon, the best data available, and by using a somewhat novel approach to overcome some of the technical challenges.

The two key challenges of estimating the impact of road networks on economic activity are well known. First is the difficulty of obtaining data which accurately reflect the conditions of the roads, and the cost of traveling along them. This is always a concern when dealing with road infrastructure—the quality of which is constantly in flux—but it is especially a challenge in Africa where infrastructure assessments are infrequent and rural roads are often unaccounted for. The second challenge is overcoming the endogenous placement of public goods. Roads tend to be built so as to connect major economic activities, e.g. linking cities, markets, mines, or areas of high agricultural productivity. Hence estimates need to take account of reverse causality in looking at the impact of roads: on the one hand economic potential may determine where roads are built, on the other hand, roads may spur greater economic activity. In situations where natural experiments are not feasible, and panel data is unavailable, instrumental variables are the most commonly used technique to correct for these placement effects, which is the approach used in this analysis. While no instrument is perfect, this paper constructs a novel variable, termed the "natural path" (described in the data section), which it is suggested greatly improves the accuracy of the estimates. This instrument is also an improvement over others such as the straight-line instruments frequently used in the literature since it more accurately captures what straight-line instruments attempt to estimate, that is, the most logical route connecting two points without taking into account other, bias-causing economic benefits.

The paper also adds value to the literature in that it attempts to more precisely estimate the actual cost of transporting goods to market by carefully geocoding[1] and digitizing the existing road network, and developing a unique and representative algorithm for optimizing and estimating costs of moving along the network. Taking into account the road classification, quality, type of paving and roughness of the terrain, the measure of transport cost to market that is calculated is perhaps the most accurate possible, given existing information.

The approach used in this paper attempts to provide a more complete picture of the extent to which household welfare and incomes are expected to improve with a given reduction in transport costs. We do so by considering several different outcome variables which we obtain from two household surveys, and a raster dataset on local GDP. The household surveys employed in this paper are the 2010 Living Standards Measurement Study - Integrated Survey on Agriculture (LSMS-ISA) for Nigeria, and the 2008 Nigeria Demographic Health Survey (NDHS). From these surveys, we are able to obtain several welfare indicators, including: revenue from crop production, revenue from livestock sales, non-agricultural income, wealth, and a multi-dimensional poverty indicator A key advantage of these household surveys is that the enumeration areas - the geo-locations -

---

[1] Geocoding is the process of finding associated geographic coordinates from other geographic data, such as street addresses, ZIP codes (postal codes) or physical maps. With geographic coordinates the features can be mapped and entered into GIS.

are recorded (with some randomness for privacy concerns). The raster dataset,[2] which we obtain from Ghosh et al. (2010) gives an estimate of local GDP at a very fine spatial level for the entire land area of Nigeria.

A variety of welfare indicators are used in this paper in recognition of the fact that there are potential data imperfections and that no single measure of welfare can adequately capture all dimensions of economic well-being. These measures also gives us the ability to study the effects of transport costs on both 'flow' measures of welfare, or 'stock' measures, which capture much longer term effects. Measures of income, such as crop revenue, livestock sales, and non-agricultural income are in the former category and will be impacted by idiosyncratic shocks or localized impacts, for instance, a bad harvest due to less than average rainfall, or a sudden illness of the household head. Meanwhile, while improving transportation infrastructure can lead to benefits in the short term, many of the benefits will not become apparent for many years after the improvement, after households and businesses have time to adjust to a new equilibrium. To capture these benefits, we also study 'stock' measures of welfare, including a wealth index available in the NDHS, and a multi-dimensional poverty index (MPI) which we construct, following Alkire and Santos (2010).[3] By looking at these different indicators of welfare, we are able to disaggregate the benefits of a transport cost reduction and obtain insights into the causal pathways to poverty reduction.

Household survey data, though useful, are not representative of all dimensions of spatially heterogeneity across the country due to small sample sizes (small relative only to the entire population of Nigeria, not in a statistical sense). For instance, determining how total income from crop production will increase when a road is improved requires information on crop production data on all of those who use, or will use after improvement, that road, which household surveys cannot provide (as they are not spatially representative).[4] In addition to allowing for cross-regional analysis, nighttime-

[2] A raster is a matrix of cells where each cell contains spatial information, in our case local GDP.

[3] Poverty, when measured in a multi-dimensional fashion which accounts for household capital—both physical and human—captures much more the 'stock', or cumulative welfare effects over time, and is thus a much better measure of welfare over the longer term.

[4] While some income data may be available in a census, censuses typically do not go into such specific details which breakdown the components of household income. Furthermore, even if the Nigerian census did contain the required information, most national censuses do not, and those that do can be difficult for a researcher to obtain from their respective governments. As part of the motivation behind this paper is to

lights data provides baseline data on income and allow for simulations (results in section 6) to forecast benefits of proposed road projects at the regional level, as well as providing an additional result which we use to check the robustness of our household survey elasticities.

The elasticities generated in this paper allow us to forecast the economic impact of the construction of future roads, or the improvement of any portion of the current network. These elasticities are summarized in the table below. Our forecasts allow for heterogeneous benefits depending on current levels of welfare, current transportation costs, and spatially varying transportation cost reductions from new road construction. This will enable decision makers to maximize the efficiency with which they use scarce resources, by prioritizing construction of those roads which would have the biggest impact on economic growth and poverty reduction in the region.

Elasticities

Welfare Indicator	Benefit (from a 10% reduction in transport costs)
Crop Revenue	9.7%
Livestock Revenue	No Benefit
Non-Agricultural Income	4.6%
MPI Poverty Reduction	2.4%
Wealth Index	2.1%
Local GDP	5.3%

2. Literature Review

This paper is related to a vast and rapidly growing literature on the effects of infrastructure on well-being and a continuing debate (among planners, policy makers, and academics) about the role of transport investments in economic growth. The debate has been fostered by limited evidence of a causal relationship and conflicting evidence provided by different studies on the relationship between the two (Gunasekara, Anderson, and Lakshmanan 2007).

develop a toolkit which can be replicated across different countries, using data that is not somewhat ubiquitous would be self-defeating.

Approaches to addressing this issue have varied considerably and evolved over time. Researchers have examined the effects of road infrastructure and transport capital investments on aggregate productivity (usually measured by Gross Domestic Product (GDP) or Personal Income), output elasticity and productivity in developed countries (Aschauer 1989, Lakshman and Anderson 2002, Lakshman and Anderson 2007, Chandra and Thompson 2000, Demetriades and Mamuneas 2000, Annala and Perez 2001, Foster and Araujo 2004, Ihori and Kondo 2001, Lokshin and Yemtsov 2003, Nadiri and Mamuneas 1996, Munnell 1990, Shirley and Winston 2004, and Sturm 2001), and in developing countries (Deichmann et. al 2002, Morrison-Paul et al 2001; Lokshin and Yemtsov 2003; Feltenstein and Ha 1995). The results however remain ambiguous with conflicting evidence of impacts in both developed and developing countries. To a large extent the contradictory evidence and the ensuing debates are a consequence of the identification and reverse causality problems.

One of the challenges in estimating the effect of transportation costs is related to the difficulty of econometrically disentangling cause and effect because of the endogenous placement of transport infrastructure. More specifically, because places that receive better infrastructure are likely to be systematically different from areas that do not, comparing regions with varying endowments of infrastructure will generate biased estimates. A set of recent papers have used rigorous and compelling identification strategies to shed light on the impact of large transport infrastructure improvements (Michaels 2008, Donaldson 2012, Datta 2012, Faber 2012 and Banerjee et al 2012)[5]. The following subsection, 2.1, discusses the estimation strategies used in the recent literature to accurately identify a causal effect of transportation cost reduction and briefly discusses how our estimation strategy builds upon these.

Another limitation in the literature is the absence of rigorous analysis that shows the mechanisms by which transport infrastructure affect well-being. This is perhaps due to the absence of detailed datasets in developing countries. Subsection 2.2 discusses the literature that provides suggestive evidence on how transport cost reductions could affect welfare.

[5] Elhance and Lakshmanan, 1988 and Ford and Poret, 1991 are examples of earlier papers that analyze the impact of aggregate transport investment in Mexico and highway improvement in Sri Lanka.

2.1 Estimation strategies used in the literature

Several approaches have been employed to address the estimation problems arising from the familiar endogeneity of road placement. One approach is to use panel data estimation methods. This is useful if the omitted variables are time-invariant or time varying at a higher administrative level. For example, Dercon et al (2008) use panel data and a dynamic panel model with General Method of Moments instrumental variable estimation method to estimate the impact of transport infrastructure on consumption and poverty rates in rural Ethiopia, while Khandker and Koolwal (2011) focus on per capita expenditure and labor supply in Bangladesh. This approach is an effective and seemingly reliable way to identify the impact of roads. Regrettably, however, panel data on transportation costs with adequate observations over time are rare especially in a developing country context, and not available for our analysis involving Nigeria and other West African countries.

Another approach is to use spatial panel data with natural experiments that exploit the historical context of transport infrastructure. These natural experiments examine scenarios where roads were built for reasons unrelated to the outcome variables under scrutiny. For instance, Donaldson (2012) studies how roads which were meant to aid colonial and military objectives ended up benefiting interregional trade in India. Jedwab and Moradi (2012) investigate how roads built to connect mining facilities in Ghana affected long-term growth and economic development. Likewise Banerjee et al. (2012) examine how transportation infrastructure constructed to connect historical cities and ports during the colonial period in China affected per capita GDP across sectors and regional GDP growth many decades later.

In cases where panel data is available but there are no natural experiments, one needs to establish the exogeneity of transport infrastructure placement. For example, Datta (2012) analyzes the effect of highway improvements in India using panel data and a difference-in-difference strategy on the presumption that improvements in highway quality in non-targeted intermediate areas is exogenous. The assumption is that areas which lie along highways built to connect two cities get better infrastructure merely because of where they happen to be located and not as a consequence of economic or other characteristics. However, because it is possible that priority is given to connecting more important areas between two nodal cities, Faber (2014) used a difference-in-difference set-up along with an instrumental variable estimation strategy to analyze the effect of China's National Trunk Highway System that connected cities on economic conditions in non-targeted peripheral counties. Faber (2014) used two hypothetical spanning tree highway networks, called least cost path spanning tree networks, and Euclidean spanning tree networks as instruments. Both instruments correspond to the question of which routes central planners would have been likely to construct if the sole policy objective had been to connect all targeted destinations on a single network subject to global construction cost minimization. Using similar logic, we create instruments (discussed below) to address the endogeneity of transportation cost.

In the absence of natural experiments and panel data, numerous studies have attempted to capture exogenous variations in transport cost by incorporating exogenous geographic features. For example, Jacoby and Minten (2009) examine the impact of plausibly exogenous variations in transport costs in Madagascar, determined by the terrain of the land which is likely to be exogenous. Using cross-sectional data for Nepal, Shrestha (2012) exploits the fact that constructing a north-south road is cheaper than constructing an east-west road due to the country's topography.

2.2 Literature on mechanisms by which transportation infrastructure affects welfare

Recent papers provide suggestive evidence on how lower transportation costs, by enabling greater access to markets, decrease trade costs and interregional price gaps (Donaldson 2012, Casaburi et al 2013), and affect input and output prices of crops (Khandker et al. 2006, Minten and Kyle 1999). These in turn affect agricultural returns and hence land value (Jacoby 2000, Shrestha 2010, Donaldson 2013). Econometric analysis of household data on the effects of road connectivity on input use, crop output, and household incomes in Madagascar and Ethiopia (Chamberlin et al 2007; Stifel and Minten 2008) suggest that remoteness negatively affects agricultural productivity and incomes at the household level. Not surprisingly the literature also finds that access to good quality roads facilitates economic diversification (Gachassin et al 2010, Fan et al 2000, Mu and van de Walle 2007).

In a novel experimental approach, Gonzalez-Navarro and Quintana-Domequ (2010) report that randomly allocated road pavement in urban areas increases property values and that households in paved areas have collateralized credit as a greater share of mortgages and private bank loans indicating that access to better roads—and hence markets and services—may improve access to credit and increase welfare. Burgess and Donaldson (2012) find that the expansion of railroads made local real incomes less responsive to productivity shocks. This suggests that lowering transportation costs via investments in transportation infrastructure played a key role in raising welfare by enhancing trade integration and lessening the degree to which productivity shocks translated into real income volatility. They also find that mortality rates became significantly less responsive to rainfall shocks where railroads were present indicating that transportation infrastructure has the potential to affect not only living standards but also the health and educational status of households. There is a large body of additional research which uses microeconomic data to examine other impacts, such as on income (Donaldson 2012, Jacoby and Minten 2009), consumption per capita (Khandker et al. 2006), and poverty reduction (Fan et al. 2000; Gibson and Rozelle 2003; Warr 2008).

Using a rigorous estimation strategy, our analysis adds to this literature by analyzing the effect of transportation cost on a variety of measures of well-being in Nigeria, captured by several measures of current income, a wealth index the probability of being multi-dimensionally poor, and local GDP.

3. Data

This paper utilizes several different datasets to analyze the relationship between transportation costs that households incur to access the nearest market (defined as cities with population of at least 100,000) and several different measures of welfare.[6] In order to do so, a very thorough road network was constructed for Nigeria, using several sources

[6] This paper analyzes the combined effect of both large transport infrastructure, such as highways, and rural roads, and thus differ from Michaels, 2008, Donaldson, 2012, Datta, 2012, Faber, 2012 and Banerjee, Duflo and Qian, 2012, which analyze the impact of large transport infrastructures, highways and railways. It also differs from Jacoby and Minten 2008, Dorosh et al 2010, Gibson and Rozelle (2003), Ali (2011), Khandker, Bakht and Koolwal (2011), Mu and van de Walle, 2007 which analyze the impact of smaller rural roads.

of data described in section 3.1 below. To correct for any placement bias inherent in estimating the benefits of transportation expenditures, an IV approach is used, and a novel new instrument was generated for this paper which we refer to as the *Natural Path*, described in section 3.2. Finally, a multitude of welfare indicators are utilized in this paper, and they are described in section 3.3.

3.1 Travel costs to the cheapest market

There is little accurate data on travel costs for much of Africa. The costs of traveling between any two points along a road network depends upon on a number of factors – distance, road conditions, terrain and type of vehicle commonly used. To our knowledge, there exists no definitive dataset which includes the entire road network for any country in Africa, including unpaved, tertiary roads, as well as the road attributes. To incorporate these features requires a digitized network map.

To construct a measure of travel costs to the market in Nigeria we combine road survey data from the Federal Roads Maintenance Agency (FERMA) and World Bank's FADAMA[7] project, with GIS roads network data from Delorme[8]. We used the Delorme dataset for data on both the trunk roads and the rural network of Nigeria, and supplemented this dataset by geocoding data on federal road attributes from FERMA and also with data on rural road attributes from the World Bank's FADAMA project.[9] Using these data and the Highway Development Management Model (HDM-4) programming tool, we accounted for the roughness of the terrain, quality and condition of the road, as well as country level factors (such as the price of fuel, average quality of the fleet, the price of a used truck, and wages)[10] to compute travel costs from households (in the case of household level analysis) or cells (in the case of SPAM data analysis) to markets. The

[7] The FADAMA project is currently in its third stage. It was originally designed to improved utilization of irrigable land, implementing an innovative local development planning (LDP) tool, and building on the success of the community-driven development mechanisms. For more information about the survey and GIS methodology see "Spatial Analysis and GIS Modeling to Promote Private Investment in Agricultural Processing Zones: Nigeria's Staple Crop Processing Zones" presented at the Annual World Bank Conference on Land and Poverty 2013.

[8] Delorme is a company that specializes in mapping and GPS solutions and has the most comprehensive GIS dataset on African roads.

[9] Road segments missing from either of these data sets were deemed to be minor and categorized as tertiary class, unpaved, and in poor condition. Finally, all necessary adjustments were made through consultations with transportation experts familiar with Nigeria's road network in order to arrive at the final road network used in this study.

[10] For more information on the HDM-4 model in general, see http://go.worldbank.org/JGIHXVL460. For more information on how the HDM-4 model was applied to Nigeria, see Appendix VI.

cheapest total travel cost to the “nearest” market is calculated using an iterative cost-minimizing process in which every possible travel path to every available market was calculated, and the least cost one was chosen as the optimal route.[11]

3.2 The natural path instrument

As discussed above, it is well established in the literature that simple OLS regressions will often overestimate the effects of public investments, such as road infrastructure, due to placement endogeneity. In order to eliminate the bias from these placement effects, an instrumental variable (IV) approach is often used. The chosen IV must be correlated with transportation costs, and must only affect the outcome variable (in this case, our welfare variables) through its effect on transportation costs, but cannot affect those outcomes directly. Neither should it be caused by the outcome variable.

To motivate the rationale for our proposed IV, it is instructive to briefly consider the source of the bias and the factors used to determine whether or not to make a public road investment. Benefits from new transportation infrastructure can come in many forms: economic benefits, including the reduction of business costs and enhanced access to markets; social benefits, including improved social cohesion, faster diffusion of information, and better access to schools and hospitals; and political benefits, including the faster deployment of armies to quell unrest or conflict, and the pleasing of certain constituencies, among others.[12] Several factors are important in determining the costs of new road construction or improvement. The most important of these include the length of the road, and the topography of the land over which the road is being built.[13] Flatter land topography beneath the road is more desirable in that it is both cheaper to build the road upon, and to drive on once the road is built.

“Straight line” or Euclidian distance IVs are commonly used in the literature because they are correlated with the costs of building the road, in particular distance, but uncorrelated with its benefits, namely the bias-causing economic benefits. The “straight line” IVs therefore rest on the assumption that, regardless of topography and terrain, two

[11] This was done in ArcGIS using the Network Analysis Extension Closest Facility Tool.
[12] Indeed, many of these benefits spill over into all three categories.
[13] In addition to the standard costs, the opportunity cost of using public funds should also be considered. However, its inclusion in the decision making process has no effect on the choice of instruments.

points which are closer together (have a shorter straight line connecting them), have lower construction and travel costs.

However, while useful, straight line IVs ignore a major determinant of a road's cost—the topography of the land—potentially making them weak instruments. Even if two points are very close together, they may not be easily accessible (e.g. if separated by a canyon or steep terrain), even if a road approximating that straight line existed. Thus, this paper addresses this problem by developing a unique instrument which considers both the distance, and the land topography between a household and the cheapest market.

Specifically this variable, which we refer to as the "natural path", is the route that would be travelled if walking from a given location to the cheapest market, in the absence of roads. The natural path is thus the path that minimizes travel time, given only the geography of the land. It therefore also represents the most cost effective place to construct a road network, if economic benefits were ignored. Moreover in the context of Africa it captures many of the historic trade and caravan routes where head-loading (walking) was the dominant pre-colonial mode of transportation.[14] Therefore, it is plausible to suggest that any endogeneity in the road network from placement decisions (i.e. decisions to place roads in areas which would maximize economic benefits) is captured in the difference between the current road network, and the natural pathway. This instrument is strictly an improvement over "straight-line" instruments as the natural path more accurately represents what straight-line instruments are attempting to estimate, that is, the most cost effective route to connect two points, while excluding any other economic benefits. Details on the GIS algorithm and data used to construct the Natural Path are in Appendix V.

3.3 Welfare indicators

In order to robustly estimate the welfare benefits of reducing transportation costs, we explore several different welfare indicators from three different sources of data, all of which are geolocated. The first data source described here is the 2010 Living Standards Measurement Study - Integrated Survey on Agriculture (LSMS-ISA) for Nigeria.[15] The LSMS-ISA is a national survey on household welfare conducted by the Nigerian Bureau

[14] Animal traction was unavailable due to the high incidence of disease carried by the tsetse fly.
[15] See Kuku-Shittu et al (2013) for a recent example using the Nigeria LSMS-ISA.

of Statistics and the World Bank's Development Research Group (DEC). The panel is a 5,000-household subset of the 22,000-household nationally representative General Household Survey (GHS).[16] LSMS-ISA provides information on total crop revenue, livestock revenue, and non-agricultural income over the past year, at the household level. Another household survey we exploit is the Nigeria Demographics and Health Survey (NDHS) from 2008, which contains an index on household wealth, and various measures of health and educational attainment of household members, which we use to construct a multi-dimensional poverty index. The NDHS is a nationally representative survey of nearly 50,000 Nigerians aged 15-59.[17] Finally, we use the Nigerian portion of the nighttime lights raster dataset (of spatially disaggregated GDP for all of Nigeria) from the lights dataset for the entire world developed by Ghosh et al. (2010).

4. Empirical Framework

Recognizing the need to better understand the pathways through which improved transportation infrastructure affects economic well-being, six outcome variables are used in this paper: (i) crop revenue, (ii) livestock revenue, (iii) non-agricultural income, (iv) household wealth index, (v) the probability of being multi-dimensionally poor, and (vi) local GDP. The first three indicators capture the possibly differential impact of roads in the sub-components of household income. These indicators measure a snap-shot of well-being and can therefore be more volatile due to idiosyncratic temporary shocks. In contrast the wealth and poverty measures are more closely related to expected "permanent income" and as a consequence are impacted less by transitory shocks. They therefore provide better indications of longer term responses to infrastructure.

[16] The LSMS-ISA is part of a $19 million project of the Bill and Malinda Gates Foundation. In Nigeria, the LSMS-ISA data was collected twice over two seasons. The Post-Planting Survey was conducted August-October 2010. This was followed by the Post-Harvest Survey in February-April 2011. Each survey is made up of three integrated questionnaires: household, agriculture, and community. In addition, certain geo-variables are available (including information on agro-ecological zones). Each enumeration area is geo-located, allowing us to merge this data with spatial data from other sources. For the purposes of this analysis, we use the 2010 post-planting survey, mainly focusing on the agriculture questionnaire with a few variables (e.g. labor) from the household survey.

[17] The original purpose of the survey was to inform policy makers on a variety of issues mainly affecting women and children, including fertility preferences, infant and young children feeding practices, nutritional status, and early childhood and maternal mortality. For an explanation of the sampling procedure used in the NDHS, see Appendix IV.

Our main identification strategy is to instrument for cost to market with the Natural Path variable (i.e. time it takes to walk to market along the natural terrain).

To illustrate the approach, consider the following model:

$$lnY_i^k = \beta_0 + \beta_1 lnT_i + X_i'\gamma + \varepsilon_i \tag{1}$$

$$lnT_i = \alpha_0 + \alpha_1 lnN_i + X_i'\theta + \mu_i \tag{2}$$

where Y_i^k denotes the level of outcome k (agricultural revenue, livestock sales, non-agricultural income, wealth index, multi-dimensional poverty and local GDP) indicating welfare of household i in case of LSMS-ISA and NDHS analysis or location i in case of nighttime lights data analysis. T_i is the transport cost to market, X_i is a vector of household controls, and N_i is the natural path variable. For nighttime lights data analysis, the control variables contain geographic level control variables denoted by X_i , which are measured at the center of each raster cell. When analyzing the impact of transportation costs on the probability of being multidimensionally poor, the dependent variable becomes $dMPI_i$ in place of lnY_i, where $dMPI_i$ is a dummy taking the value of one if the household is multidimensionally poor and zero otherwise.

The key parameter of interest is β_1, i.e. the causal impact of the cost of traveling to the cheapest market, on household welfare which is estimated using the customary two-stage approach where the endogenous transport cost variable is instrumented using the natural path variable.

As a robustness check, we calculate a set of Conley Bounds, developed by Conley et al. (2012), for the coefficient of interest, transport cost to market. To illustrate this, let N_i be the IV and rewrite equation (1) as follows:

$$lnY_i = \beta_0 + \beta_1 lnT_i + X_i'\gamma + \delta N_i + \varepsilon_i \tag{3}$$

The traditional IV strategy assumes that $\delta = 0$. Conley et al (2012) allow δ to be close but not actually equal to zero, in other words to allow the IVs to be only plausibly exogenous. By allowing the value of δ to vary, one can test how sensitive the estimates are to different degrees of exogeneity, thereby testing the validity of instruments as well as estimates.

5. Empirical Results

This section presents and discusses the main empirical estimates of the relationship between the cost of transporting goods to market and the chosen measures of household welfare. Overall, the empirical results suggest that improving the quality of roads, thereby lowering cost to market, significantly benefits the rural households—though the impacts appear to depend on the source of income and location.

5.1 LSMS Measures of Household Welfare

We begin by presenting the effects of transport cost on income from different sources, such total revenue from crop sales, livestock sales, and non-agricultural income of the household over the past year, which are the flow measures of household welfare using the LSMS-ISA data for Nigeria.

Crop Revenue

Consistent with prior expectations, the regressions suggest that increased transport costs lower crop revenue on average. Preliminary OLS results are reported in column (1) of Table 1, which shows that that decreasing transport cost by ten percent would increase crop revenue by approximately 8.3 percent.

The coefficients on the controls are consistent with theory. Labor is positively related with crop revenue, while labor squared is negatively related. Land is positively related to crop revenue. Fertilizer and irrigation are both positively related to crop revenue, but fertilizer is not significant. Age and age squared are not statistically significant. Literacy of the household head is positively related to the crop revenue and is significant at the 1% level. The agro-ecological zones (AEZ) control for the agricultural potential of the land and climate. The omitted AEZ category is tropic-warm/semi-arid. The results suggest that tropic warm/sub-humid and tropic cool/sub-humid both have a positive relationship with crop revenue as compared to tropic-warm/semi-arid, while tropic-warm/humid is worse than the omitted category.

While these preliminary OLS results are reassuring in that they conform to prior expectations, they must be treated with caution as they do not take account of the endogeneity of roads. Table 1, column (2) reports the IV estimate where cost to market is instrumented by the natural path. This unbiased estimate of the effect of transport costs is slightly larger than the OLS coefficient, at -0.96. The natural path IV passes the Angrist-Pischke F Test of Weak Identification, with the F statistic far exceeding 10, the rule of thumb.

To check the robustness of the IVs to relaxation of the exclusion restriction, the Conley Bounds are calculated following Conley et al (2012), and reported in Table 7. The 95% confidence interval suggests that the coefficient on the variable of interest remains consistently negative. Taken together with the Angrist-Pischke F statistic and first stage results, these findings suggests that our IV is appropriate.

Livestock Sales

As with crop revenue, we report both our OLS and IV estimates of livestock sales in Table 2. In both cases, we find that the estimated coefficient on cost to market is negative but *not* statistically significant.[18] We report the Conley Bounds for livestock revenue in Table 7. The 95% confidence interval crosses the zero line, further evidence that our estimates are not statistically significant.

Overall, it would seem that there is only a weak relationship between cost to market and the sale of livestock. This might be in part due to the multiple roles of livestock as cultural icons, or a store of value and capital good. In which case sales in a given year are driven by decisions on asset management (e.g., need for revenue to manage temporary household needs for cash—weddings or natural disasters, for example—much more than are crop sales).Alternatively a lack of refrigeration, would have a stronger impact on livestock sales than simply the state of the road.

Non-Agricultural Income

Turning next to the relationship between access to markets and non-agricultural income, economic theory suggests that as transport costs decrease, more opportunities outside of the agricultural sector become available. For example, the commute to and from work may be reduced to a reasonable length by the construction or improvement of

[18] We are clustering at the enumeration area in our regressions. The LSMS and DHS data are sampled within an enumeration area (e.g. village). Since household observations are likely to be correlated within these areas, we cluster our standard errors at the enumeration area level. When looking at the unclustered t-statistics, we find that the coefficient on transport cost is significant at the 10% level.

a road. Also, getting non-agricultural goods to market would be made easier. To investigate this, we regress log non-agricultural income on the log of cost to market, holding constant household characteristics and zone fixed effects. [19]

Table 3 reports the OLS estimates (column 1), and the IV estimates (column 2) for non-agricultural income. The OLS estimates suggest that reducing transport costs by 10 percent increases non-agricultural income by 4.2 percent. After controlling for endogeneity, our IV estimates find a higher increase in income: 4.6 percentage points. The Conley Bounds are reported in Table 7. We find that the estimated coefficient remains consistently negative as we progressively relax the exclusion restriction.

5.2 NDHS Measures of Household Welfare

We now turn to the two measures of household welfare from the NDHS, the wealth index and the multidimensional poverty index.

Nigeria Demographic and Health Survey

The first indicator of household welfare we describe is the "wealth index". This variable is generated by the NDHS surveyors and is available in the DHS data. The wealth index is an estimate of a household's long term standard of living. It is computed using data from the household's ownership of consumer goods; dwelling characteristics; type of drinking water source; toilet facilities; and other characteristics that are related to a household's socio-economic status (NPC 2009). To construct the index, each of these assets are assigned weights (factor scores) generated through principal component analysis, and the resulting asset scores are standardized in relation to a standard normal distribution with a mean of zero and standard deviation of one (Gwatkin et al., 2000). Each household is then assigned a score for each asset, and the scores are summed to arrive at a final number. More detailed information on the wealth index is generated can be obtained from Shea and Johnson (2004).

The second indicator, a multi-dimensional poverty measure, is generated specifically for this study. We follow Alkire and Santos (2010) to calculate the Multi-Dimensional Poverty Index (MPI) for each household. The MPI is a weighted sum of ten indicators of deprivations across three dimensions: education, health, and standard of

[19] Nigeria is divided into six regional zones: north central, north east, north west, south east, south south, and south west.

living. We follow convention and use equal weights for each of the three dimensions and for indicators within dimensions. A household is considered to be multi-dimensionally poor if it is deprived in three of the ten weighted indicators. Table A4 in Appendix II gives more specific details on how this index was constructed.[20]

Household Wealth Index

In order to examine whether a transportation cost reduction allows households to accumulate wealth, and thereby help them to move to a higher welfare equilibrium, we analyze its effect on the wealth index generated by DHS. Table 4 presents the results from regressing this wealth index on transportation costs (both in natural log form).

Column (1) in Table 4 presents the coefficients from OLS estimation, and column (2) presents the coefficients from second stages of the IV estimation. The coefficients of the natural path instrument in the first stage is very highly statistically significant and positively related to transportation cost to the market, as expected. Our results indicate that a 10 percent reduction in transportation cost leads to a statistically significant 2.3 percent increase in the wealth index according to OLS estimation, and a 2.1 percent increase in the wealth index from our IV estimates.[21] Again, our IV passes the Angrist-Pischke F Test of Weak Identification, with the F statistic far exceeding 10. The coefficient estimates of the effect of agricultural potential indicate that, ceteris paribus, as agricultural potential increases, household wealth also increases. Households which are agriculturally involved, and rural households seem to accumulate less wealth. Larger households and households with more males and females in the working age group (15-49 years for females and 15-59 for males) accumulate more wealth, which confirms the intuition that working aged people earn more and are able to save over time to accumulate wealth. Households with more children in the age group 0-5 years accumulate less wealth, which again confirms the intuition that because children require time and resources to be invested in them, households with more children have less wealth. The intuitive and expected results obtained from the regression analysis of the

[20] For robustness, we also calculate an additional MPI using data from the LSMS. Despite differences in the data, results from the LSMS generated MPI are quite similar to that from the NDHS.
[21] As with the regressions using LSMS data, we cluster our standard errors in the DHS regressions at the enumeration area.

factors that affect wealth accumulation lends more credibility to the specification used in the regression analysis.

Multi-dimensional poverty

Next, we consider the impact that reducing the cost to market would have on the probability of a household being multi-dimensionally poor. In Table 5, we report two sets of results: linear probability (OLS and IV) and the marginal effects from maximum likelihood estimation (probit and IV probit). Given that the outcome variable in this case is binary, probit may be more efficient but least squares may be more robust because it does not rely on distributional assumptions. We present both results for robustness, but for space considerations, we only interpret probit models here. (Note, the two sets of estimates are broadly consistent.)

Overall, decreasing a household's transport cost to market decreases their probability of being multi-dimensionally poor. This finding is robust to both linear probability and maximum likelihood estimation. Specifically, using the same controls we used for the wealth index, we find that reducing transport costs to market by 10 percent reduces a household's probability of being multi-dimensionally poor by 2.4 percent.

Our results also indicate, as one might expect, that households that live in rural areas or are agriculturally involved are more likely to be multi-dimensionally poor, and households that live in areas with higher agricultural potential are less likely to be multi-dimensionally poor. Comparing the probit and IV probit marginal effects, we find that the IV probit estimate (0.24) is considerably larger than the probit estimate (0.07) indicating that the Probit model underestimates the effect of transportation costs on multi-dimensional poverty and that the IV estimation approach was important to obtain an unbiased, accurate measure.

5.3 Local GDP

We now turn to our final set of results, which look at the impact of transport costs on local GDP. By using nighttime satellite imagery collected by the National Oceanic and Atmospheric Administration (NOAA), Ghosh et al. (2010) estimate a raster dataset of local economic activity. This dataset spatially disaggregates Nigeria's (among other countries) 2006 GDP into square pixels 30 arc seconds wide (approximately 1km^2), using the fact that brighter lights at night are associated with higher levels of economic activity (see Ghosh et al. 2010 for additional details about how these data were generated). Given the granularity of our control data, we aggregated this data into square cells with sides measuring 5 arc minutes in length (approximately 10km). As control variables, we include total population within each cell,[22] and total population squared, the Euclidean distance to the nearest mining facility,[23] as well as indicators measuring the agro-ecological potential yield of the land for four staple crops[24]—cassava, maize, rice, and yams—and their squared terms (all variables are in natural logs). In addition to these control variables, state fixed-effects are included in the regressions. This specification is tested both for all of Nigeria, and also for only rural areas.

Turning to our results, we measure how overall economic activity is affected by transportation costs. Our dependent variable is local GDP, as estimated using nighttime lights, and the entire country of Nigeria is split into square observation units with sides measuring 5 arc minutes in length (approximately 10km). The results from regressions using local GDP data are presented in Table 6.

Columns (1) and (2) of Table 6 show OLS and IV results when we include all of Nigeria in the sample. The OLS estimate of the coefficient on transportation costs implies that a 10 percent reduction in transport costs increases local GDP by 5.75%. The IV estimate is slightly lower at 5.3%. Both the coefficient on population, and its squared term are significant and positive, implying that there important agglomeration economies to local GDP. The negative and significant coefficient on distance to mine implies that economic activity is, as we should expect, denser around mining facilities. The

[22] Population data is from Landscan and is available here: http://web.ornl.gov/sci/landscan/

[23] Distance to the nearest mining facility is included because mines tend to be areas of great economic activity. In addition to the economic activity at the mine, mines can often generate economic spillovers for industries which service the mine and the mine's workers and their families. Data on mining facilities throughout Nigeria was obtained from the National Minerals Information Center of the USGS. The dataset includes geo-referenced data on all mining facilities, active or closed, between 2006 and 2010. Because of the wide definition of what a mining facility actually is, we selected only a subset of mining facilities available to include in our dataset. Facilities selected were those which involved the extraction of minerals or hydrocarbons from the ground (specifically coal, tin, iron, nitrogen and petroleum), or the processing of hydrocarbons. Mining facilities that were in the USGS dataset but not included in this analysis include facilities like cement plants, or steel mills, which are likely concentrated in large cities or manufacturing areas. We also excluded plants that were labeled as being closed.

[24] Agro-ecological potential data is from GAEZ, a product of FAO. It considers climate and soil conditions to estimate the maximum potential yields in each region for a large number of crops. The data used in this model assumes climactic conditions similar to the 1961-1990 baseline level, and is calculated assuming low input systems.

coefficients on the agricultural potential of various crops are difficult to interpret because they are highly correlated with each other.

Columns (3) and (4) of Table 6 show OLS and IV results when we only include rural areas of Nigeria. Examining the coefficient on transportation costs in the IV regression, we see that the effect of reducing transportation costs is slightly lower when urban areas are omitted; a 10 percent reduction in transportation costs implies a 4.8 percent increase in local GDP. However, a modified t-test shows that the coefficient on transportation costs for the full sample is statistically indistinguishable from that using only rural observations. Similarly, the coefficients on the control variables do not change significantly between columns (2) and (4), with the exception of the coefficient on population becoming insignificant (but the squared term remains significant, leading to the same interpretation). Consistent with all other results, the Conley bounds in Table 7 show that our coefficient on market cost remains negative and within a small range when the exclusion restriction is relaxed.

6. Economic Impact of Alternative Road Investments

In this final section, we use our estimate of the local GDP elasticity of transport costs to simulate the effect of several road infrastructure improvement projects. We analyze several projects which have been proposed by the World Bank, The African Development Bank (ADB) and by the New Partnership for Africa's Development (NEPAD), a planning and coordinating technical body of the Africa Union[25]. However, this methodology could also be applied to study any road improvement or new road construction project within Nigeria.

6.1 The projects

Nigeria has an extensive national road network of more 85,000 km of classified roads (Gwilliam 2011). Both paved and unpaved road network densities are more than twice as high as those for the peer group of resource-rich African countries, although still only half of the levels found in Africa's middle-income countries (Foster and Pushak 2011). According to the Africa Infrastructure Country Diagnostic benchmark study

[25] Several World Bank projects have also been analyzed using the same methodology, both are not included for space considerations.

(Foster and Briceño-Garmendia 2008): if Nigeria wishes to meet its economic and social targets for transportation infrastructure it would need to invest $1.2 billion annually for a 10 year period.

We present below an estimate of the impact of improving the portion of NEPAD's and ADB's Trans African Highway[26] project segments that runs through Nigeria as shown in Figure 1. It is assumed in the simulations that each corridor would be improved from its current quality, to paved and good condition status. The baseline scenario is obtained from FERMA[27] and requires that 20 km need to be paved, while approximately 1,275 km need to be improved from poor to good and 815 km from fair to good condition.

To calculate the change in transportation cost resulting from the improvement of each corridor, we follow the same procedure utilized in section 3.1 to estimate the travel cost to the cheapest market and compare these to current transport costs. The percentage change in transportation costs for each cell, if all three of the corridor improvement projects were completed, is shown in Figure 2.

6.2 Simulation Methodology

For each project, we estimate the increase in local GDP in each grid cell separately, and then aggregate these benefits to arrive at an aggregate total benefit. The increase in local GDP is then summed up amongst all grid cells to arrive at an aggregate value. Formally, this calculation is given by:

$$B_j = \sum_i \eta * \tau_{ij} * y_i, \tag{4}$$

where B_j is the total increase to local GDP due to project j, η is the local GDP elasticity of transportation costs from Table 6 (-0.534), τ_{ij} is the percentage change in transportation costs in cell i due to project j, and y_i is the baseline GDP in cell i, from the local GDP data. This increase in GDP represents an increase in annual GDP over the

---

26 The system of Trans African Highways consists of 9 main corridors with a total length 59 100 km. The concept as originally formulated in the early 1970s, aims at the establishment of a network of all-weather roads of good quality, which would: a) provide as direct routes as possible between the capitals of the continent, b) contribute to the political, economic and social integration and cohesion of Africa and c) ensure road transport facilities between important areas of production and consumption.

27 For more information about the FERMA road survey and GIS methodology see "Spatial Analysis and GIS Modeling to Promote Private Investment in Agricultural Processing Zones: Nigeria's Staple Crop Processing Zones" presented at the Annual World Bank Conference on Land and Poverty 2013

baseline level. These benefits will accrue every year as long as the benefits from reducing transportation costs and baseline GDP levels, both remain constant[28].

The spatial approach also allows us to estimate the total number of Nigerians who would benefit from each road construction project, and to conduct an efficiency analysis, by estimating the benefit per road kilometer improved. Given the inherent uncertainty involved in statistical analysis, we calculate total benefits given our preferred elasticity, -0.534, as well as a range of benefits representing our 95% confidence interval, and other plausible values derived from our Conley bound estimation.

6.3 Simulation Results

We estimate the benefits of each road construction project described above separately, and then the total benefits if all of the projects were completed. Note that the total benefits of all of the construction projects is not equal to the sum of the benefits of each of the projects individually, because there is some overlap between the project locations.

Benefit Point Estimates

We first present results from our preferred point estimation for the benefits of the three NEPAD projects we analyzed. They are given in Table 8. Note that the benefits from these projects are quite large. The North-South Corridor, which is the longest road of the three projects, would result in estimated annual benefits of over $1 billion. Annual benefits from the Northeastern and Southern corridors are significantly lower, at $214 and $288 million, respectively. Nevertheless, these roads are also shorter, potentially implying a lower cost of improvement. If all projects were completed, total estimated annual benefits would be $1.5 billion. Figure 3 shows where exactly the increase in local GDP would occur if all three of the corridor improvement projects were completed.

Turning to the third column of Table 8, we calculate the total benefit per KM of each corridor, which allows us to rank the projects according to their benefit-efficiency; i.e. assuming road improvement costs are uniform and equal across projects, which project gives us the most benefits per unit cost. We see that the North-South Corridor project would return annual benefits of approximately $950,000 per KM improved,

---

[28] This would be a dubious assumption over the long term, but might be a reliable approximation over a short, 3-5 year period.

significant higher than that of the Northeastern and Southern Corridors, which have benefits of $230,000 and $390,000 per KM improved, respectively.

Using Landscan population data, we can also get an approximation of the number of people whose transportation costs to market would decline, as a result of each project. Again, the North-South Corridor has the biggest impact here, benefiting 21.4 million people. The Northeastern and Southern Corridor projects would benefit 13.3 and 9.2 million people, respectively. Figure 4 shows the total population affected if all three corridor improvement projects are completed. By dividing total benefits by the number of people affected, we arrive at estimated benefits per person affected. The North-South Corridor project, in addition to benefiting the largest number of people, also has the biggest benefit per capita, at $41.5 per person benefited.

Benefit Plausible Range

Given that statistical estimates are not perfectly precise, we also present two plausible benefit ranges based on 1) the 95% confidence interval surrounding the estimated local GDP elasticity of transportation costs, and 2) the Conley bounds surrounding this elasticity based on relaxing the exclusion restriction in our 2-stage estimation technique.

The 95% confidence interval around the point estimate of our elasticity is [-0.572 -0.495]. The recalculated benefits for each of the NEPAD projects for these two elasticity bounds are given in Table 9. The range of benefits are quite small, due to the fact that our point estimate is fairly precisely estimated. When all projects are completed, the estimated annual benefits range from $1.405 billion to $1.623 billion. Per capita and per KM benefit could also easily be calculated for this range of benefits. Because the total number of people affected and the length of each road will not change, these value are not shown for brevity.

Finally, we offer a range of annual benefits based on relaxing the exclusion restriction that our natural path IV has no direct impact on local GDP. This range is estimated using Conley bounds, and shown in Table 7. When we allow for our IV to be correlated with local GDP by up to 0.01, or -0.01, we get a range of elasticities equal to [-0.481 -0.586]. Again, we use these elasticities to recalculate total annual estimate benefits for each of the NEPAD transportation improvement projects. These values are given in

Table 10. If all projects were completed, annual benefits would range from $1.365 billion to $1.663 billion. While the range given by the Conley bounds is slightly larger than that from the 95% confidence interval, we still end up with a fairly tight range of benefits.

6.4 Road Prioritization

Thus far, we have calculated the estimated economic benefits of improving three major trunk roads which connect large swaths of Nigeria. These roads are very long, and pass through areas of Nigeria ranging from rural, especially in the North, to very urban, particularly around Lagos. These differences imply that the economic benefits of improving any road will not be homogenous throughout the length of the road; i.e. improving some portions of the road will generate a larger benefit than others. By analyzing small segments of each road separately, we can further prioritize these infrastructure projects not just by which overall project would have the largest impact, but within each project, which segments should be improved with greater urgency.

In order to analyze different segments of the road, we first split up the roads into "marketsheds". We define a marketshed by the land area around each city (with a population of at least 100,000 residents) through which one travels when going to that particular city while minimizing transportation costs. The size and shape of a city's marketshed is therefore going to depend on both the road network around that city, and its proximity to other cities. We partition each road by the marketshed in which it lies. This is the most logical way to partition each road because improving a segment of a road will only benefit those people who reside within the same marketshed as the improved road (by definition, people living within a marketshed other than the one where the improved road resides will not travel on that improved road to reach the market, and it will therefore not affect their transportation costs, or local GDP).

Upon splitting the analysis into different marketsheds, we calculate several different measures as candidates for ranking the priority of improving each road segment. Table 11 shows a list of the marketsheds which contain a portion of one of the three NEPAD roads, and would therefore benefit from its improvement, as well as the measures a policy maker might use to prioritize each segment. Depending on the priorities of the policy maker, several alternative metrics may be the most important. If equitable and shared prosperity are the highest priorities, one might look to maximize the total population affected (column 6), or total population affected per KM of improved road (column 8). In either of these examples, improving roads around the Lagos marketshed would have the largest impact, affecting 8.7 million people, or 251,000 people per KM improved. Total population affected per KM of improved road is displayed visually in Figure 5. If one is looking simply to generate the largest economic benefits, then total GDP increase (column 3) is the variable of interest. In that case, the portion of the road within the Lagos marketshed would again be the most beneficial to improve, as increasing GDP within that marketshed generates an estimated benefit of $681 million. If the goal is to have the largest percentage increase in GDP, then local GDP increase as a share of total marketshed GDP (column 4) is the relevant metric, and the roads falling within Shagamu should take top priority. Finally, if economic efficiency is the most important, then GDP increase per km improved (column 7) should drive decision making. In this instance, we again see that Lagos should be the marketshed which takes top priority. Figure 6 displays visually the benefits per kilometer of each road segment.

7. Discussion and Conclusion

Identifying the causal impact of cost-to-market on the welfare of households is challenged by three potential sources of bias: the endogenous location of (i) households, (ii) markets, and (iii) roads. Failure to take these into account could bias our results (Emran and Hou (2013)). Spatial sorting by households could potentially bias the estimates, if households moved to a particular location on the basis of a variable we did not control for. For example, suppose that households of higher ability were better able to move closer to the market. High ability individuals would also be expected to achieve higher incomes and wealth. In this scenario, the impact of cost to market on welfare would be overestimated.

In the context of Nigerian farmers, who we are considering in this paper, this spatial sorting can be expected to be much less of a concern. Given the lack of a functional land market, it is highly unlikely that farming households would change locations, as moving would require abandoning one’s land. Instead, farming households tend to take over the land that was farmed by their ancestors. In other words, while individuals do migrate (usually to cities) is it rare for the household as a whole to relocate. As such, the location of a farm would have more to do with environmental characteristics, rather than those of the household members themselves. Even so, we do include characteristics of the household head (age and literacy) in our regressions. To capture the environmental characteristics (such as average rainfall and soil quality), that can affect household location choice we include an Agro-Ecological Zone fixed effect.

The markets we focus on in this paper are cities with populations greater than 100,000. Such cities tend to emerge historically in locations of economic potential. Failure to account for this would tend to overestimate the benefits of reducing the cost to market. The inclusion of agro-ecological zone fixed effects, or other regional fixed effects, accounts in part for these sources of economic potential. We checked the robustness of our results controlling for marketshed fixed effects (to account for unobserved heterogeneity in the way markets are formed/located) and find that the effect of reducing transportation cost on wealth index, multidimensional poverty, local GDP are very robust (results will be available upon request).

The above analysis, while by no means exact, represents a robust attempt at estimating the economic impact from several proposed road improvement projects. Although we believe we have used the best possible methods, and the best possible data, several short-comings are acknowledged.

In calculating our estimated local GDP elasticity of transportation cost, we use data which itself is estimated using nighttime lights. This adds an additional level of uncertainty to our estimates, but uncertainty which is unavoidable due to the fact that spatially disaggregated data on actual (non-estimated) GDP is not available. Additionally, even though our elasticity is based on estimated data, it falls within the range of other elasticities we calculated of several income variables, which were obtained using survey data.

Another potential short-coming we acknowledge is the fact that we are using cross-sectional data, which can often make discerning causality very difficult. The instrumental variable technique we employ is one very commonly used in the literature, and we believe our IV is a significant improvement over those used by other very well cited authors. Nevertheless, there is no such thing as a "perfect" instrument. For this reason, we have been careful to present our point estimates along with the respective Conley bounds, which give a range of estimates under the assumption that our instruments are not perfectly exogenous. The ranges given by our Conley bound estimates are relatively small, showing that even if our instrument were to violate the necessary exclusion restriction, our point estimates would not change dramatically.

Finally, it is important to note that benefits from these road projects simulated in section 6 will not all occur immediately, nor all at once. They will likely cascade over time, as people begin learning of the new, lower transportation costs, and adjusting their behavior accordingly. Therefore, these estimates should be considered long term annual benefits.

References

Alkire, S. & Santos, M.E. (2010). Acute multidimensional poverty: A new index for developing countries. OPHI Working paper 38.

Annala, C., & Perez, P. (2001). Convergence of public capital: Investment among the United States 1977–1996. Public Finance and Management, 1(2), 214–229.

Aschauer, D. A. (1989). Is public expenditure productive. Journal of Monetary Economics, 23, 177–200.
http://www.worldbank.org/en/topic/transport/overview#2.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian (2012) "On the Road: Access to Transportation Infrastructure and Economic Growth in China" NBER Working Paper 17897

Burgess, Robin and Dave Donaldson, 2012, "Railroads and the Demise of Famine in Colonial India", working paper

Casaburi, Lorenzo, Glennerster, Rachel and Tavneet Suri "Rural Roads and Intermediated Trade: Regression Discontinuity Evidence from Sierra Leone", MIT Working paper

Chandra, A., & Thompson, E. (2000). Does public infrastructure affect economic activity? Evidence from the rural interstate highway system. Regional Science and Urban Economics, 30(4), 457–490.

Conley, Timothy G., Christian B. Hansen, and Peter E. Rossi (2012) "Plausibly Exogenous", The Review of Economics and Statistics, 94(1): 260-272

Datta, Saugato (2012), "The Impact of Improved Highways on Indian Firms", Journal of Development Economics, 99(1): 46-57.

Deichmann, U., Fay, M., Koo, J., & Lal, S. V. (2002). Economic structure, productivity, and infrastructure quality in southern Mexico. Washington, DC: World Bank.

Demetriades, P., & Mamuneas, T. (2000). Inter-temporal output and employment effects of public infrastructure capital: Evidence from 12 OECD economies. Economic Journal, 110(465), 687–712.

Dercon, Stefan, Daniel O. Gilligan, John Hoddinott, and Tassew Woldehanna (2008) "The Impact of Agricultural Extension and Roads on Poverty and Consumption Growth in Fifteen Ethiopian Villages" IFPRI Discussion Paper 00840

Donaldson, David and Richard Hornbeck, 2013, “RAILROADS AND AMERICAN ECONOMIC GROWTH: A “MARKET ACCESS” APPROACH”, Working paper

Donaldson, David. 2012. Railroads of the Raj: Estimating the impact of transportation infrastructure. Processed, mit.

M. Shahe Emran & Zhaoyang Hou, 2013. "**Access to Markets and Rural Poverty: Evidence from Household Consumption in China**," The Review of Economics and Statistics, MIT Press, vol. 95(2), pages 682-697, May

Faber, Benjamin (2012), “Trade Integration, Market Size, and Industrialization: Evidence From China’s National Truck Highway System”,Working Paper, LSE.

Fan, S., P. Hazell, and S. Thorat (2000). “Government Spending, Growth and Poverty in Rural India,” American Journal of Agricultural Economics 82 (4), 1038{1051.

Foster, V., & Araujo, M. (2004). Does infrastructure reform work for the poor? A case study from Guatemala. Washington, DC: World Bank.

Foster, Vivien, and Cecilia Briceño-Garmendia (2008) “Africa Infrastructure Country Diagnostic.” Overhauling the Engine of Growth: Infrastructure in Africa. September. Washington, DC: World Bank.

Foster, Vivien, and Nataliya Pushak. (2011) “Nigeria's infrastructure: a continental perspective.” Policy Research Working Paper 5686. The World Bank. June, 2011.

Feltenstein, A. and Ha, J. (1995).'The role of infrastructure in Mexican economic reform.' The World Bank Economic Review, vol. 9, pp. 287-304.

Gachassin, Marie, (2010) “Roads Impact on Poverty Reduction: A Case Study of Cameroon”

Ghosh, Tilottama, et al. (2010) "Shedding light on the global distribution of economic activity." The Open Geography Journal 3.1: 148-161.

Gibson, John and Scott Rozelle (2003), “Poverty and Access to Roads in Papua New Guinea”, Economic Development and Cultural Change, 52(1): 159-185.

Gonzalez-Navarro, Marco and Quintana-Domeque, Climent, Roads to Development: Experimental Evidence from Urban Road Pavement (February 23, 2010). Available at SSRN: http://ssrn.com/abstract=1558631 or http://dx.doi.org/10.2139/ssrn.1558631

Gunasekara, K., W.P. Anderson, and T. R. Lakshmanan (2008 November) “Highway Induced Development: Evidence from Sri Lanka”, World Development.

Gwatkin, D.R., S. Rutstein, K. Johnson, R.P. Pande, and A. Wagstaff. 2000. Socio-economic differences in health, nutrition, and population. HNP/Poverty Thematic Group. Washington, D.C.: World Bank.

Gwilliam, Ken (2011) "Africa's Transport Infrastructure: Mainstreaming Maintenance and Management" World Bank.

Ihori, T., & Kondo, H. (2001). The efficiency of disaggregate public capital provision in Japan. Public Finance and Management, 1(2), 161–182.

Jacoby, Hanan G. and Bart Minten (2009) "On Measuring the Benefits of Lower Transportation Costs" Journal of Development Economics 89(1): 28-38

Jedwab, Remi and Alexander Moradi (2012) "Colonial Investments and Long-Term Development in Africa: Evidence from Ghanaian Railroads" Working Paper

Khandker, Shahidur and Gayatri Koolwal (2011), "Estimating the Long-Term Impacts of Rural Roads", Policy ResearchWorking Paper 5867,World Bank.

Khandker, S. R., Z. Bakht, and G. B. Koolwal (2006, April). "The Poverty Impact of Rural Roads : Evidence from Bangladesh". Policy Research Working Paper Series 3875, The World Bank.

Kuku-Shittu, Oluyemisi, Astrid Mathiassen, Amit Wadhwa, Lucy Myles, Akeem Ajibola (2013) "Comprehensive Food Security and Vulnerability Analysis" IFPRI Discussion Paper 01275

Lakshmanan, T.R. and W. Anderson, 2007, "Transport's Role in Regional Integration Processes" Market Access, Trade in Transport Services and Trade Facilitation, Round Table 134. OECDECMT, Paris, pp. 45-71.

Lakshmanan, T.R. and W. Anderson, 2002. A White Paper on "Transportation Infrastructure, Freight Services Sector, and Economic Growth", prepared for the U.S. Department of Transportation, Federal Highway Administration.

Lokshin, M., & Yemtsov, R. (2003). Evaluating the impact of infrastructure rehabilitation projects on household welfare in rural Georgia. Washington, DC: World Bank.

Michaels, G. (2008): The Effect of Trade on the Demand for Skill | Evidence from the Interstate Highway System," Review of Economics and Statistics, 90(4).

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

Mu, R. and D. van de Walle (2007, August). Rural Roads and Poor Area Development in Vietnam. Policy Research Working Paper Series 4340, The World Bank.

Munnell, A. (1990). How does public infrastructure affect regional economic performance. New England Economic Review, (September–October), 11–32.

Nadiri, I., & Mamuneas, T. (1996). Contributions of highway capital to output and productivity growth in the US economy and industries. Washington, DC: Administrative Office of Policy Development, Federal Highway Administration.

National Population Commission (NPC) [Nigeria] and ICF Macro. 2009. Nigeria Demographic and Health Survey 2008. Abuja, Nigeria: National Population Commission and ICF Macro.

Rutstein, Shea O. and Kiersten Johnson. 2004. The DHS Wealth Index. DHS Comparative Reports No. 6. Calverton, Maryland: ORC Macro.

Shreshtha, Slesh A. (2012). Access to the North-South Roads and Farm Profits in Rural Nepal. Working Paper, National University of Singapore

Shirley, W., & Winston, C. (2004). Firm inventory behavior and the returns from highway infrastructure investments. Journal of Urban Economics, 55, 398–415.

Stifel, David, Bart Minten, and Bethlehem Koro (2012) "Economic Benefits and Returns to Rural Feeder Roads: Evidence from a Quasi-Experimental Setting in Ethiopia" IFRPI Policy Research Institute Working Paper

Stifel, David, and Bart Minten. 2008. "Isolation and Agricultural Productivity." Agricultural Economics, 39 (1): 1–15.

Sturm, J. (2001). The impact of public infrastructure capital on the private sector of the Netherlands: An application of the symmetric generalized McFadden Cost Function. PFM. Vol. 1 No. 2

Tobler, Waldo (1993) "Three presentations on geographical analysis and modeling non-isotropic geographic modeling speculations on the geometry of geography global spatial analysis". NATIONAL CENTER FOR GEOGRAPHIC INFORMATION AND ANALYSIS. TECHNICAL REPORT 93-1. February 1993

Transport: Sector Results Profile
Sustainable Transport for All: Helping People to Help Themselves
April 9, 2014

Uchida, H. and Nelson, A. (2009) “Agglomeration Index: Towards a New Measure of Urban Concentration.” Background paper for the World Bank’s World Development Report

Warr, Peter, (2008), “How Road Improvement Reduces Poverty: The Case of Laos,” *Agricultural Economics*, 39, pp. 269-279.

World Bank (2007): Evaluation of World Bank Support to Transportation Infrastructure. Washington DC: World Bank Publications.

Table 1: Crop Revenue

Dependent Variable: ln(Crop Revenue)	**(1) OLS**	**(2) IV**
ln(Cost to Market)	-0.827*** (-5.40)	-0.970*** (-4.83)
hh agricultural labor	0.159** (2.48)	0.170** (2.56)
hh agricultural labor squared	-0.020*** (-3.40)	-0.021*** (-3.43)
land	0.017*** (4.09)	0.018*** (4.19)
fertilizer	0.001 (1.63)	0.001* (1.67)
dummy=1 if irrigates land	0.487 (1.15)	0.497 (1.20)
dummy=1 if tropical warm subhumid	0.346* (1.89)	0.312* (1.72)
dummy=1 if tropical warm humid	-0.022 (-0.08)	-0.072 (-0.26)
dummy=1 if tropical cool humid	1.719*** (7.94)	1.667*** (7.48)
age of hh head	0.014 (0.87)	0.013 (0.77)
age squared	-0.000 (-0.72)	-0.000 (-0.65)
dummy=1 if hh head is literate	0.312*** (2.84)	0.298*** (2.65)
Constant	1.785*** (3.16)	2.090*** (3.31)
First Stage Results IV: ln(Natural Path) Angrist-Pischke Test		 0 .631*** (19.28) 371.66 P=0.0000
Observations	2,600	2,600

Robust t-statistics clustered at the enumeration area in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Data: Nigeria LSMS-ISA 2010

Table 2: Livestock Sales

Dependent Variable: ln(Livestock Sales)	(1) OLS	(2) IV
ln(Cost to Market)	-0.026 (-0.15)	-0.161 (-0.87)
ln(cost of animals)	0.186*** (3.85)	0.186*** (3.88)
hh agricultural labor	0.056 (0.65)	0.061 (0.72)
hh agricultural labor squared	-0.001 (-0.14)	-0.001 (-0.16)
land	0.008** (2.19)	0.009** (2.39)
fertilizer	-0.000 (-0.98)	-0.000 (-0.91)
dummy=1 if irrigates land	0.461 (1.36)	0.466 (1.38)
dummy=1 if tropical warm subhumid	0.355* (1.90)	0.332* (1.77)
dummy=1 if tropical warm humid	-0.134 (-0.47)	-0.195 (-0.68)
dummy=1 if tropical cool humid	-0.245 (-1.23)	-0.280 (-1.42)
age of hh head	0.027 (1.43)	0.027 (1.40)
age squared	-0.000* (-1.72)	-0.000* (-1.72)
dummy=1 if hh head is literate	-0.317** (-2.49)	-0.328*** (-2.63)
Constant	0.116 (0.17)	0.385 (0.56)
First Stage Results IV: ln(Natural Path) Angrist-Pischke Test of Weak Identification		 0.649*** (25.46) 648.20 P=0.0000
Observations	3,297	3,297

Robust t-statistics clustered at the enumeration area in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Data: Nigeria LSMS-ISA 2010

Table 3: Non-Agricultural Income

Dependent Variable: Non-Agricultural Income	**(1) OLS**	**(2) IV**
ln(Cost to Market)	-0.421***	-0.464***
	(-3.35)	(-3.38)
age of hh head	-0.052***	-0.052***
	(-2.62)	(-2.63)
age squared	0.000**	0.000**
	(2.47)	(2.47)
dummy=1 if hh head is literate	0.376***	0.373***
	(3.48)	(3.46)
land	0.003	0.003
	(0.76)	(0.85)
hh labor	0.287***	0.288***
	(3.04)	(3.07)
hh labor squared	-0.015	-0.015*
	(-1.64)	(-1.65)
Dummy=1 if north east	-0.503**	-0.504**
	(-2.27)	(-2.29)
Dummy=1 if north west	-0.266	-0.262
	(-1.23)	(-1.22)
Dummy=1 if south east	-0.236	-0.251
	(-0.99)	(-1.06)
Dummy=1 if south south	0.824***	0.811***
	(3.69)	(3.66)
Dummy=1 if south west	-0.039	-0.064
	(-0.15)	(-0.24)
Total business expenses	0.000***	0.000***
	(28.29)	(28.47)
Constant	5.191***	5.266***
	(9.49)	(9.35)
First Stage Results		
IV: ln(Natural Path)		0.638***
		(19.89)
Angrist-Pischke Test of Weak Identification		395.48
		P=0.0000
Observations	1,355	1,355

Robust t-statistics clustered at the enumeration area in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Data: Nigeria LSMS-ISA 2010

Table 4: Wealth Index

Dependent Variable: ln(Wealth Index)	**(1) OLS**	**(2) IV**
ln(Cost to Market)	-0.228*** (-11.01)	-0.212*** (-7.81)
Agri. Potential	0.027* (1.68)	0.027* (1.66)
Dummy=1 if hh agri. Involved	-0.268*** (-9.19)	-0.275*** (-9.09)
Ln(age of hh. head)	-0.098*** (-2.77)	-0.097*** (-2.74)
Female household head dummy	-0.011 (-0.28)	-0.013 (-0.31)
ln(no. of hh members)	0.061* (1.79)	0.060* (1.76)
ln(no. of females aged 15 to 49 yrs)	0.0921*** (3.67)	0.0927*** (3.70)
ln(of no. of males aged 15 to 59 yrs)	0.0663*** (2.73)	0.0677*** (2.79)
ln(no. of children aged 0 to 5 yrs)	-0.0392** (-2.07)	-0.0384** (-2.03)
Dummy=1 if Rural	-0.482*** (-12.29)	-0.492*** (-11.90)
Dummy=1 if north east	-0.409*** (-6.92)	-0.412*** (-7.01)
Dummy=1 if northwest	-0.160*** (-2.91)	-0.158*** (-2.87)
Dummy=1 if south east	0.264*** (4.67)	0.273*** (4.74)
Dummy=1 if south south	0.383*** (7.04)	0.386*** (7.05)
Dummy=1 if south west	0.232*** (4.20)	0.239*** (4.28)
Constant	12.710*** (107.80)	12.690*** (107.30)
First Stage		
IV: ln(Natural Path)		0.761*** (22.15)
Angrist-Pischke Test of Weak Identification		490.44 P=0.0000
Observations	6,684	6,684

Robust t-statistics clustered at the enumeration area in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Data: Nigeria DHS 2008

Table 5: Multi-dimensional Poverty, NDHS

Dependant Variable: dummy=1 if poor	(1) OLS	(2) IV	(3) Probit	(4) IV Probit
ln(Cost to Market)	0.0846***	0.0650***	0.0725***	0.243***
	(7.085)	(4.216)	(7.29)	(4.43)
Agri. Potential	-0.0173*	-0.0170*	-0.0177*	-0.0706*
	(-1.923)	(-1.880)	(-2.22)	(-2.15)
Dummy=1 if hh agri. Involved	0.139***	0.147***	0.104***	0.457***
	(7.628)	(7.803)	(7.66)	(7.93)
Ln(age of hh. head)	-0.0262	-0.0281	-0.0210	-0.0946
	(-1.047)	(-1.127)	(-0.88)	(-0.95)
Female household head dummy	0.0485	0.0504	0.0249	0.111
	(1.396)	(1.436)	(0.95)	(0.98)
ln(no. of hh members)	0.0387*	0.0402*	0.0364	0.154
	(1.699)	(1.758)	(1.62)	(1.64)
ln(no. of females aged 15 to 49 yrs)	0.110***	0.110***	0.112***	0.461***
	(6.239)	(6.185)	(6.42)	(6.27)
ln(of no. of males aged 15 to 59 yrs)	0.00764	0.00582	-0.000813	-0.00885
	(0.401)	(0.305)	(-0.04)	(-0.11)
ln(no. of children aged 0 to 5 yrs)	0.00977	0.00880	0.0174	0.0684
	(0.733)	(0.655)	(1.28)	(1.21)
Dummy=1 if Rural	0.184***	0.197***	0.166***	0.647***
	(8.318)	(8.416)	(7.96)	(8.51)
Dummy=1 if north east	0.154***	0.158***	0.177***	0.749***
	(6.646)	(6.912)	(7.07)	(7.08)
Dummy=1 if northwest	0.120***	0.118***	0.128***	0.488***
	(4.957)	(4.823)	(4.87)	(4.74)
Dummy=1 if south east	-0.203***	-0.214***	-0.173***	-0.583***
	(-5.903)	(-6.075)	(-5.02)	(-5.27)
Dummy=1 if south south	-0.173***	-0.178***	-0.160***	-0.522***
	(-5.250)	(-5.256)	(-4.98)	(-5.00)
Dummy=1 if south west	-0.175***	-0.184***	-0.144***	-0.485***
	(-5.843)	(-6.037)	(-4.81)	(-5.06)
Constant	0.334***	0.354***		
	(4.215)	(4.504)		
First Stage Results				
IV: ln(Natural Path)		0.761***		0.761***
		(22.15)		(22.17)
Angrist-Pischke Test		490.44		
		P=0.0000		
Observations	6,684	6,684	6,684	6,684

Robust t-statistics clustered at the enumeration area in parentheses, *** p<0.01, ** p<0.05, * p<0.1
Data: Nigeria DHS 2008

Table 6: Local GDP

Dependent Variable: Local GDP	(1) OLS Full Sample	(2) IV Full Sample	(3) OLS Rural Only	(4) IV Rural Only
ln(Cost to Market)	-0.575***	-0.534***	-0.542***	-0.488***
	(-34.13)	(-27.21)	(-30.14)	(-23.16)
ln(Distance to Mine)	-0.0587***	-0.0674***	-0.0557***	-0.0662***
	(-2.91)	(-3.32)	(-2.68)	(-3.17)
ln(Population)	0.140***	0.114***	0.0268	0.0220
	(3.35)	(2.71)	(0.52)	(0.42)
ln(Population)^2	0.0448***	0.0468***	0.0525***	0.0532***
	(18.04)	(18.56)	(15.86)	(15.96)
ln(Cassava potential yield)	0.0278***	0.0247***	0.0252***	0.0217***
	(3.65)	(3.25)	(3.25)	(2.81)
ln(Cassava potential yield)^2	0.00499***	0.00440**	0.00394**	0.00330*
	(2.65)	(2.34)	(2.05)	(1.72)
ln(Yams potential yield)	-0.0174*	-0.0159*	-0.0166*	-0.0147
	(-1.90)	(-1.75)	(-1.80)	(-1.60)
ln(Yams potential yield)^2	-0.00249	-0.00230	-0.00202	-0.00180
	(-1.10)	(-1.02)	(-0.89)	(-0.80)
ln(Maize potential yield)	0.0548***	0.0536***	0.0678***	0.0636***
	(3.87)	(3.56)	(4.77)	(4.22)
ln(Maize potential yield)^2	-0.0074***	-0.0069***	-0.0073***	-0.0063***
	(-3.76)	(-3.35)	(-3.63)	(-3.02)
ln(Rice potential yield)	-0.00984*	-0.00960*	-0.0113*	-0.0117**
	(-1.76)	(-1.71)	(-1.94)	(-2.01)
ln(Rice potential yield)^2	-0.00102	-0.00104	-0.00122	-0.00138
	(-0.80)	(-0.81)	(-0.92)	(-1.04)
Constant	-1.883***	-1.837***	-1.645***	-1.694***
	(-8.84)	(-8.60)	(-6.99)	(-7.19)
First Stage Results				
ln(Natural Path)		0.7342***		0.7342***
		(170.71)		(159.85)
Angrist-Pischke Test		29143.31		25553.44
		P=0.0000		P=0.0000
Observations	10728	10607	9899	9797

*** p<0.01, ** p<0.05, * p<0.1
Data: Nighttime Lights local GDP, Ghosh (2010)

Table 7: Conley Bounds

	Support for possible values of δ	95% Confidence Interval	
	IV: ln(Natural path)	Lower Bound	Upper Bound
ln(Crop revenue)	δ: [-0.0001, 0.0001]	-1.366	-0.575
	δ: [-0.001, 0.001]	-1.367	-0.574
	δ: [-0.01, 0.01]	-1.382	-0.560
ln(Livestock sales)	δ: [-0.0001, 0.0001]	-0.527	0.205
	δ: [-0.001, 0.001]	-0.529	0.206
	δ: [-0.01, 0.01]	-0.543	0.220
ln(Non-agri. income)	δ: [-0.0001, 0.0001]	-0.667	-0.174
	δ: [-0.001, 0.001]	-0.668	-0.173
	δ: [-0.01, 0.01]	-0.679	-0.162
ln(Wealth Index)	δ: [-0.0001, 0.0001]	-0.247	-0.148
	δ: [-0.001, 0.001]	-0.248	-0.147
	δ: [-0.01, 0.01]	-0.259	-0.136
ln(MPI)	δ: [-0.0001, 0.0001]	0.034	0.087
	δ: [-0.001, 0.001]	0.033	0.088
	δ: [-0.01, 0.01]	0.021	0.099
ln(Local GDP) *full sample*	δ: [-0.0001, 0.0001]	-0.572	-0.495
	δ: [-0.001, 0.001]	-0.574	-0.494
	δ: [-0.01, 0.01]	-0.586	-0.481
ln(Local GDP) *rural only*	δ: [-0.0001, 0.0001]	-0.529	-0.446
	δ: [-0.001, 0.001]	-0.531	-0.445
	δ: [-0.01, 0.01]	-0.543	-0.433

Calculated using the code from Conley et al (2012).

Table 8: NEPAD Project Benefits

	GDP Increase (Million USD, 2006 PPP)	Road Length (Kms)	GDP Increase Per KM (Million USD)	Population Affected (Millions)	Per Capita Benefit (USD per person affected)
All Projects	$ 1,515	2,774	$ 0.55	36.5	$ 41.5
North-South Corridor	$ 1,061	1,121	$ 0.95	21.4	$ 49.5
Northeastern Corridor	$ 214	939	$ 0.23	13.3	$ 16.1
Southern Corridor	$ 288	729	$ 0.39	9.2	$ 31.2

Table 9: NEPAD Porject Benefits, 95% confidence interval

	GDP Increase Lower Bound	GDP Increase Point Estimate	GDP Increase Upper Bound
Transport Cost Elasticity	-0.495	-0.534	-0.572
All Projects	$ 1,405	$ 1,515	$ 1,623
North-South Corridor	$ 983	$ 1,061	$ 1,136
Northeastern Corridor	$ 199	$ 214	$230
Southern Corridor	$ 267	$ 288	$ 308

Table 10: NEPAD Project Benefits, Conley Bounds

	GDP Increase Lower Bound	GDP Increase Point Estimate	GDP Increase Upper Bound
Transport Cost Elasticity	-0.481	-0.534	-0.586
All Projects	$ 1,365	$ 1,515	$ 1,663
North-South Corridor	$ 955	$ 1,061	$ 1,164
Northeastern Corridor	$ 193	$ 214	$ 235
Southern Corridor	$ 259	$ 288	$ 316

Table 11: Road Improvement Prioritization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Marketshed	Length of Road Improved, Kms	Total GDP, Millions USD PPP	GDP Increase, Millions USD PPP	Percentage increase in GDP (3/2)	Total Population	Population Effected	GDP Increase per Km improved, Million USD/KM (3/1)	Population Effected per KM of road (6/1)
Abakaliki	201.5	3,095	65.32	2.1%	3,609,315	3,608,627	0.324	17,910
Agbor	75.2	1,973	0.16	0.0%	773,825	33,465	0.002	445
Bauchi	52.1	2,808	0.19	0.0%	2,503,664	135,459	0.003	2,599
Benin	105.0	6,937	3.73	0.1%	1,606,343	158,835	0.036	1,512
Bida	28.4	1,607	0.87	0.1%	1,180,780	84,630	0.031	2,984
Enugu	89.1	7,556	65.33	0.9%	3,682,892	2,705,968	0.733	30,353
Ibadan	65.8	7,832	72.18	0.9%	2,587,179	1,237,726	1.097	18,804
Ijebu Ode	75.3	2,847	32.34	1.1%	796,589	585,154	0.429	7,768
Ilorin	190.0	4,682	58.07	1.2%	2,159,909	660,418	0.306	3,475
Kaduna	1737.0	9,512	52.78	0.6%	2,056,143	632,083	0.304	3,639
Kano	414.0	17,693	95.88	0.5%	13,270,990	7,572,639	0.232	18,293
Katsina	1265.0	5,659	81.63	1.4%	4,627,367	1,625,196	0.645	12,849
Lagos	34.6	42,593	681.60	1.6%	10,642,885	8,714,546	19.691	251,752
Maiduguri	252.9	5,050	73.24	1.5%	4,028,684	4,027,021	0.290	15,924
Minna	140.0	4,617	5.83	0.1%	1,780,793	187,871	0.042	1,342
Ogbomosho	56.0	887	21.32	2.4%	663,362	326,754	0.381	5,835
Okitipupa	73.6	1,217	1.34	0.1%	1,000,196	127,388	0.018	1,731
Onitsha	73.4	7,446	116.75	1.6%	2,778,550	1,935,824	1.592	26,390
Oyo	54.7	844	5.051	0.6%	528,668	203,382	0.092	3,717
Potiskum	267.6	3,852	47.52	1.2%	4,998,330	1,766,429	0.178	6,600
Shagamu	84.2	4,042	100.25	2.5%	653,758	653,758	1.191	7,765
Zaria	114.4	5,089	23.26	0.5%	4,496,220	3,106,335	0.203	27,156

Figure 1: Map of NEPAD road projects

Figure 2: Percentage change in transportation costs

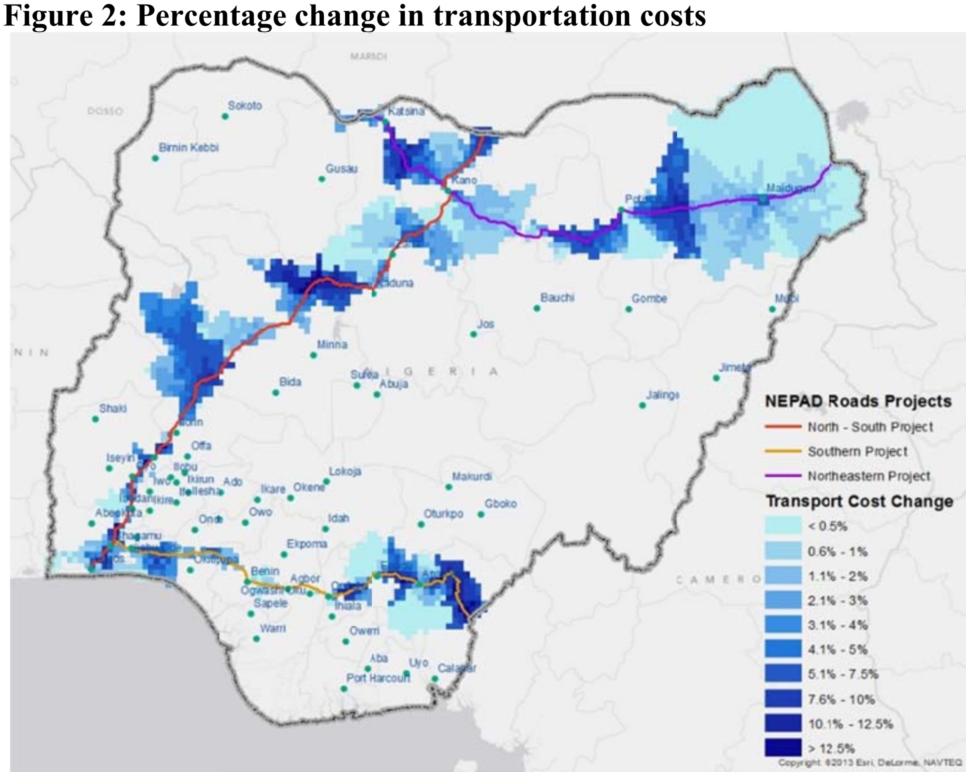


Figure 3: Increase in local GDP

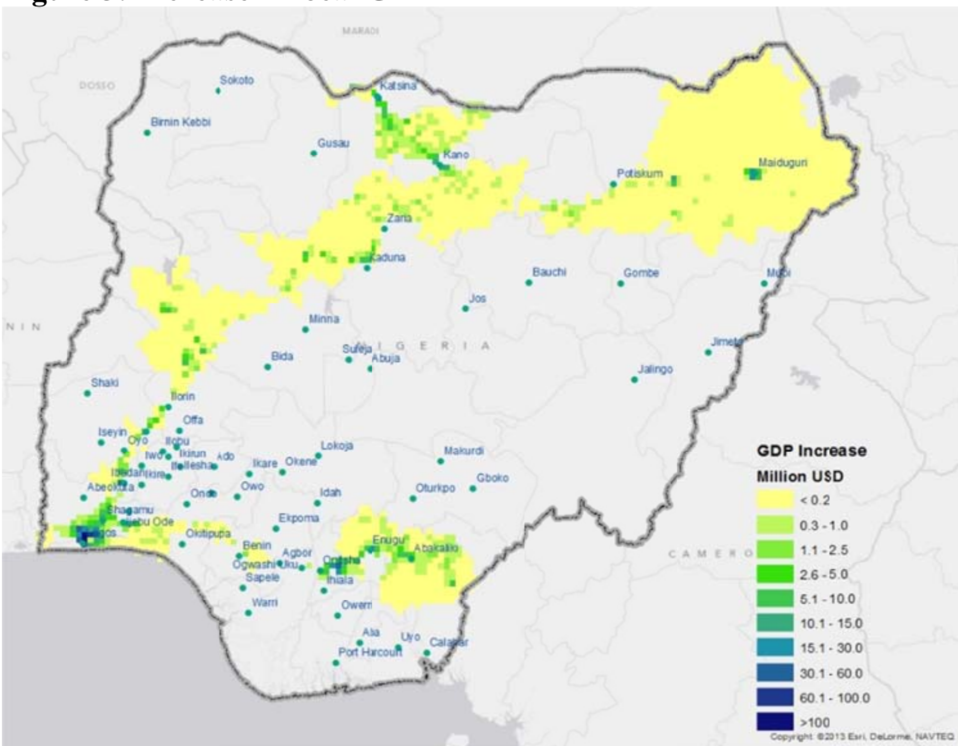


Figure 4: Population affected by Corridor improvement projects

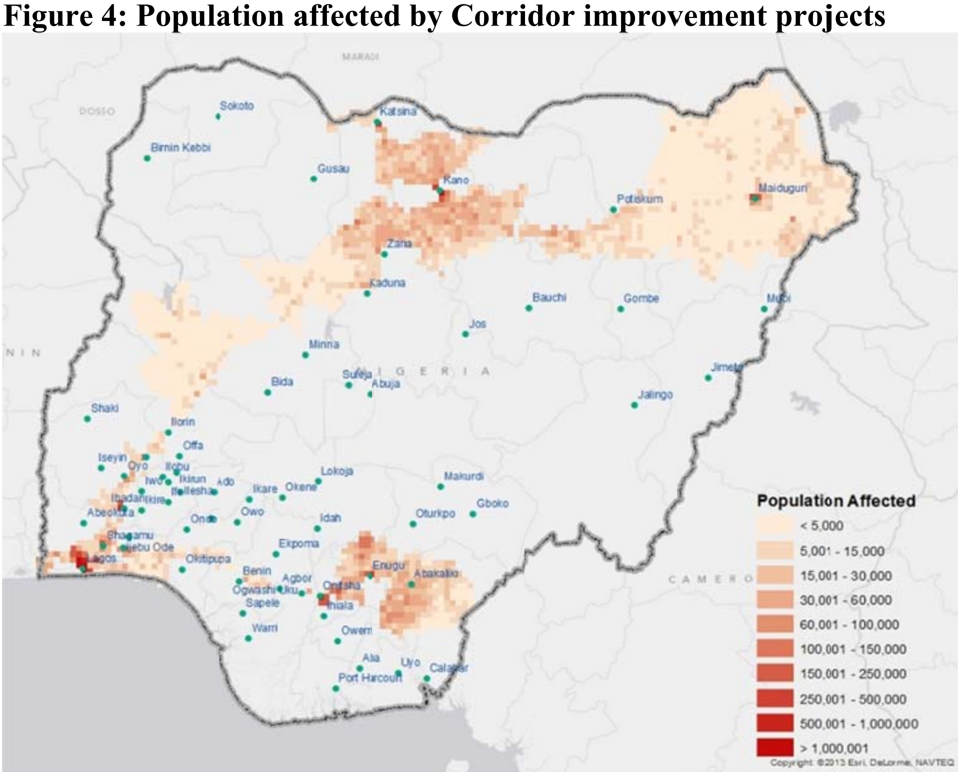


Figure 5: Population Benefited per KM Improved

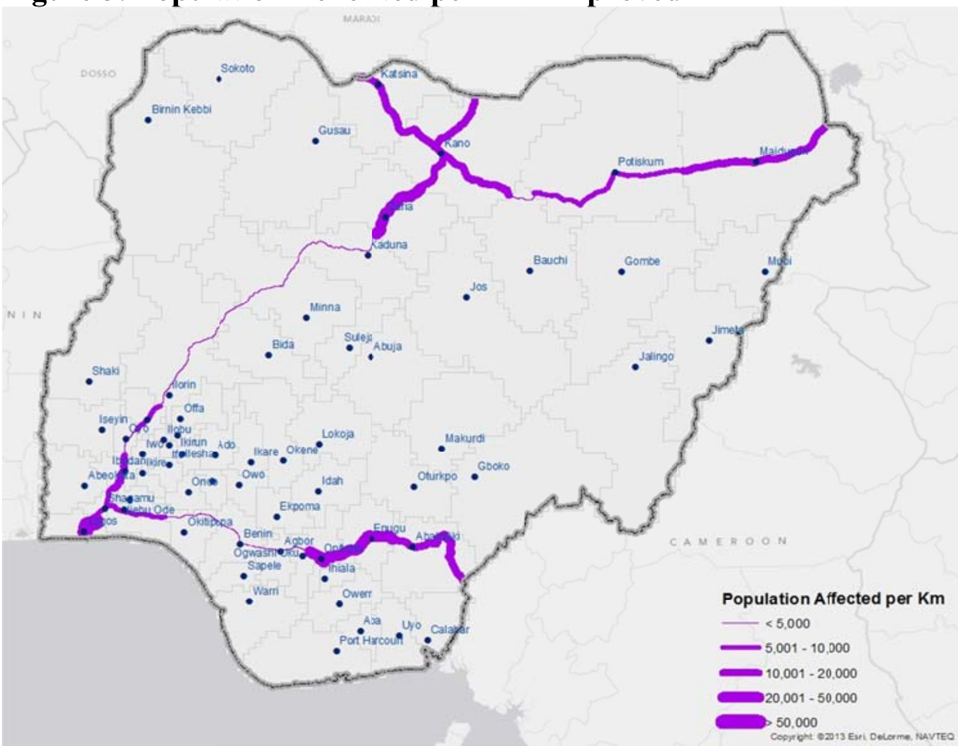
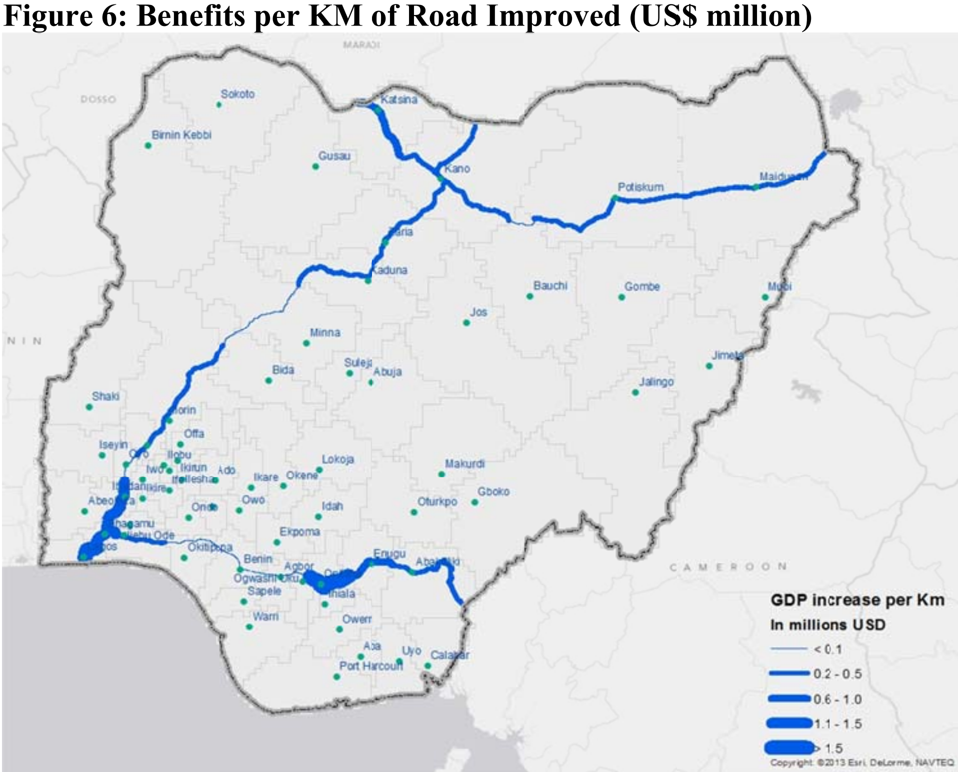


Figure 6: Benefits per KM of Road Improved (US$ million)



Appendix I: Summary Statistics

Table A1: Summary Statistics, LSMS Nigeria

Variable	Obs	Mean	Std. Dev.	Min	Max	Label
Outcome						
crop	2,600	157.84	2,126.75	0.00	105,600.00	Crop revenue (USD)
income	2,600	1,281.07	35,930.45	0.00	1,320,013	Non-agriculture income (USD)
MPI	2,600	0.37	0.15	0.00	0.83	Multi-dimensional Poverty Index: Weighted sum of indicators
animalsale_usd	3,297	48.77	204.33	0.00	3,960.00	Total sales of livestock (USD)
Treatment						
total_cost	2,600	5.10	3.30	0.14	17.36	Cost to market (USD)
IV						
natpath_hrs	2,600	14.06	10.43	0.00	59.42	Natural path to market (hrs)
Controls						
age	2,600	51.40	15.10	15.00	110.00	Age of household head
dliterate	2,600	0.55	0.50	0.00	1.00	Dummy: hh head is literate
land	2,600	9.23	16.92	0.00	265.03	Land (km squared)
labor	2,600	2.92	2.14	0.00	18.00	Number of workers in the house
agrihome	2,600	2.02	2.03	0.00	17.00	Household members working on own plot
total_fertilizer	2,600	11.14	79.91	0.00	2,299.00	Total fertilizer used (km)
dirrigate	2,600	0.95	0.21	0.00	1.00	Dummy: hh irrigates its plot
dwarmsemiarid	2,600	0.31	0.46	0.00	1.00	Dummy: Tropical warm semi-arid
dwarmsubhumid	2,600	0.59	0.49	0.00	1.00	Dummy: tropical warm sub-humid
dwarmhumid	2,600	0.09	0.28	0.00	1.00	Dummy: tropical warm humid
dcoolsubhumid	2,600	0.01	0.11	0.00	1.00	Dummy: tropical cool sub-humid
_Izone_2	2,600	0.17	0.38	0.00	1.00	Dummy: north east
_Izone_3	2,600	0.19	0.40	0.00	1.00	Dummy: north west
_Izone_4	2,600	0.23	0.42	0.00	1.00	Dummy: south east
_Izone_5	2,600	0.14	0.35	0.00	1.00	Dummy: south south
_Izone_6	2,600	0.09	0.29	0.00	1.00	Dummy: south west
totalbiz_costs	2,600	613.69	25,888.87	0	1,320,000	Total business expenses (USD)
animalcosts_usd	3,297	58.90	313.26	0.00	10,312.50	Costs of livestock (USD)

Table A2. Summary Statistics, NDHS

	Mean	Std. dev	Min	Max
Outcomes:				
Wealth index	-12407	98657.2	-145026	305274
Multi-dimensionally poor (dummy)	0.70257	0.457	0	1
Variable of interest:				
Cost to market (USD)	5.652	4.097	0.290	30.301
Instruments:				
Time taken to reach market using natural path (hrs)	15.185	12.417	0	70.794
Controls:				
Agricultural potential (factor of ln agri. potential for cassava, maize and rice)	-0.043	0.966	-1.554	1.211
Dummy: Household agriculturally involved	0.732	0.443	0	1
Age of household head	40.031	11.268	17	99
Sex of household head	1.034	0.182	1	2
No. of household members	6.517	3.079	3	43
No. of female members in households aged 15-49 yrs	1.418	0.775	1	15
No. of female members in households aged 15-59 yrs	1.279	0.678	1	12
No. of children aged 0 to 5 yrs	1.706	0.860	1	9
Dummy=1 if type of residence is rural	0.713	0.452	0	1
Dummy=1 if north east	0.216	0.412	0	1
Dummy=1 if northwest	0.271	0.445	0	1
Dummy=1 if south east	0.080	0.271	0	1
Dummy=1 if south south	0.114	0.318	0	1
Dummy=1 if south west	0.130	0.336	0	1
No. of observations: 6684				

Table A3: Summary Statistics, Local GDP datasets

	Mean	**Std. Dev.**	**Min**	**Max**	**Label**
Local GDP	25.36888	145.3173	0	4469.6	Total income per cell, millions USD (2006 PPP)
Population	13.64	44.45	0	1,639.2	Population per cell (thousands), Landscan 2006
Cassava Potential Yield	833.5371	681.5004	0	2775	Yield Kg/ha, GAEZ FAO
Rice Potential Yield	494.7672	436.5076	0	1792	Yield Kg/ha, GAEZ FAO
Yams Potential Yield	609.166	383.2807	0	1747	Yield Kg/ha, GAEZ FAO
Maize Potential Yield	1209.439	625.5546	0	3556	Yield Kg/ha, GAEZ FAO
No. of Observations: 10015					

Appendix II: NDHS Multi-dimensional Poverty Index

Table A4: NDHS multi-dimensional poverty index components

Dimension	Indicator	Deprived if…	Relative Weight
Education	Highest degree earned	No household member has completed five years of education.	1.67
	Child School Attendance	Household has a school-aged child not attending school	1.67
Health	Child Mortality	Household has had at least one child aged 0-5 years die in the past 5 years.	1.67
	Nutrition	Household has a malnourished woman aged (15-49) or child aged (0-5).	1.67
Standard of Living	Electricity	The household has no electricity.	0.56
	Improved Sanitation	Household does not have improved sanitation.	0.56
	Safe Drinking Water	Household does not have access to improved water source.	0.56
	Flooring	The household has a dirt floor.	0.56
	Cooking fuel	The household uses dirty cooking fuel.	0.56
	Asset Ownership	The household does not own more than one bicycle, motorcycle, radio, fridge, TV, or phone and does not own a car.	0.56

World Heath Organization (WHO) standards were followed in determining what to consider unimproved water sources, inadequate sanitation, and dirty cooking fuel.

A household is considered to be multi-dimensionally poor if its weighted sum of indicators was greater than 3. Note that the weights add up to about 10, the number of indicators (difference due to rounding).

Appendix III: LSMS Multi-dimensional Poverty Index

To test the robustness of our multi-dimensional poverty index indicator, we constructed a second MPI using data from the LSMS. The index is constructed in a similar manner, with three main components each receiving equal weight: education, health and standard of living. Table A5 gives a detailed description of the components of this index.

Table A5: LSMS multi-dimensional poverty index components

Dimension	Indicator	Deprived if…	Relative Weight
Education	Highest degree earned	No household member has completed six years of education, i.e. earned at least the First School Leaving Certificate (FSLC)	1/6
	Child School Attendance	Any school-aged child is not attending school (children 6-16)	1/6
Health	Child Mortality	Any child has died in the family	1/6
	Nutrition	Any household member has gone to sleep hungry during the past week	1/6
Standard of Living	Electricity	The household has no electricity	1/18
	Improved Sanitation	The household does not have a toilet that flushes or a ventilated improved pit, or must share one with other households	1/18
	Safe Drinking Water	Household does not have access or must walk more than 30-minutes round trip to get safe water. (safe water includes: pipe borne water, bore hole/hand pump, well/spring protected, rainwater)	1/18
	Flooring	The household has a straw, dirt, sand, or mud floor	1/18
	Cooking fuel	The household cooks with firewood, coal, grass, or kerosene (as opposed to electricity or gas)	1/18
	Asset Ownership	The household does not own more than one radio, TV, bike, motorbike, or fridge, and does not own any landline, car or other vehicle	1/18

For consistency, similar controls were used in regressions on the LSMS MPI as were used for other LSMS regressions. As with the NDHS MPI, we estimate two linear probability models (OLS and IV), and two maximum likelihood models (probit and IV probit) Table A6, columns (1) report the OLS and IV estimates, respectively, of the effect that log cost to market has on the households' probability of being multi-dimensionally poor. We find that the coefficient on market cost is positive and significant at the one percent level in both regressions. Consistent with a positive bias, the IV estimates is smaller in magnitude, between at 0.073, compared with 0.104 for OLS.

Table A6, column (3) reports the probit marginal effects which are nearly identical to the OLS estimate in column (1). Column (4), which reports the IV probit estimate shows a much larger estimated impact, more than doubled. A ten percent decrease in transport costs decreases the probability of being multidimensional poverty by roughly 2 percent.

As a robustness check of the exogeneity of our IVs, we report the Conley Bounds in Table A7. From the 95% confidence intervals reported, we see as we increase the correlation between the IV and the outcome variable the range of estimated values widens, but remains positive. Taken together with the first stage test results, these test statistics suggest that our IVs have power in explaining the variation in cost to market across the households.

In general, the coefficient on market cost for the MPI constructed using the NDHS is very similar to that for the MPI constructed using the LSMS. The NDHS coefficients on market cost for IV and IV probit are 0.065 and 0.243, respectively. For the LSMS MPI, those coefficients are 0.073 and 0.208. In fact, a modified t-test confirms that there is no statistical difference between these estimates.

Table A6: Multi-dimensionally poor, LSMS

Dependant Variable: dummy=1 if MPI poor	(1) OLS	(2) IV2	(3) Probit	(4) IV Probit
ln(Cost to Market)	0.104***	0.073***	0.114***	0.208***
	(4.68)	(3.02)	(4.47)	(2.77)
hh agricultural labor	0.061***	0.063***	0.065***	0.184***
	(4.40)	(4.59)	(4.43)	(4.64)
hh agricultural labor squared	-0.006***	-0.006***	-0.006***	-0.018***
	(-3.64)	(-3.72)	(-3.71)	(-3.81)
land	-0.001	-0.001	-0.001	-0.002
	(-1.41)	(-1.15)	(-1.57)	(-1.29)
fertilizer	0.000	0.000	0.000	0.000
	(0.71)	(0.74)	(0.68)	(0.67)
dummy=1 if irrigates land	-0.048	-0.048	-0.096	-0.280
	(-1.02)	(-1.07)	(-1.38)	(-1.32)
dummy=1 if tropical warm subhumid	-0.140***	-0.148***	-0.153***	-0.449***
	(-5.10)	(-5.37)	(-4.93)	(-5.04)
dummy=1 if tropical warm humid	-0.170***	-0.181***	-0.193***	-0.535***
	(-4.14)	(-4.42)	(-4.07)	(-4.44)
dummy=1 if tropical cool humid	-0.339***	-0.350***	-0.370***	-0.999***
	(-4.80)	(-5.16)	(-4.93)	(-4.77)
age of hh head	-0.003	-0.003	-0.003	-0.010
	(-0.88)	(-0.99)	(-0.78)	(-0.90)
age squared	0.000	0.000	0.000	0.000
	(1.18)	(1.25)	(1.03)	(1.11)
dummy=1 if hh head is literate	-0.258***	-0.261***	-0.271***	-0.775***
	(-13.76)	(-13.97)	(-14.24)	(-13.34)
Constant	0.740***	0.808***		
	(6.46)	(7.02)		
First Stage Results				
IV: ln(Natural Path)		0.631***		0.631***
		(19.28)		(19.32)
Angrist-Pischke Test		371.71		
		P=0.0000		
Observations	2,600	2,600	2,600	2,600

*** p<0.01, ** p<0.05, * p<0.1

Data: Nigeria LSMS-ISA 2010

Table A7: Conley Bounds, MPI LSMS

	Support for possible values of δ	95% Confidence Interval	
ln(MPI LSMS)	IV: ln(Natural path)	**Lower Bound**	**Upper Bound**
	δ: [-0.0001, 0.0001]	0.025	0.120
	δ: [-0.001, 0.001]	0.024	0.121
	δ: [-0.01, 0.01]	0.009	0.135

Calculated using the code from Conley et al (2012).

Appendix IV: NDHS

Administratively, Nigeria is divided into 36 states and Abuja, the federal capital territory. Each state is subdivided into local government areas (LGAs), and each LGA is divided into localities. In addition to these administrative units, during the 2006 Population Census, each locality was subdivided into convenient areas called census enumeration areas (EAs). The sample frame for this survey was the list of EAs used in that census. The EAs were stratified separately by urban and rural areas. Rural areas are defined as a locality with a population of less than 20,000 constitutes.

The primary sampling unit (PSU), or cluster, for the 2008 NDHS was defined on the basis of enumeration area (EA) from the 2006 census frame. A minimum requirement of 80 households (400 population) for the cluster size was imposed in the design. If the selected EA has a population smaller than this minimum, a supplemental household listing was conducted in the neighboring EA. Although in Nigeria a majority of the population resides in rural areas, the urban areas in some states were over-sampled in order to provide reliable information for the total urban population at the national level. The target of the 2008 NDHS sample was to obtain 36,800 completed interviews. Based on the level of non-responses found in the 2003 Nigeria DHS, to achieve this target, approximately 36,800 households were selected, and all women aged 15-49 were interviewed. A requirement was to reach a minimum of 950 completed interviews per state. In each state, the number of households was distributed proportionately among its urban and rural areas. The selected households were then distributed in 888 clusters in Nigeria, 286 of which were urban area clusters, and 602 were rural area clusters. More details about the sample selection can be obtained in NPC (2009).

Appendix V: Natural Path

To construct the natural pathway instrument, we followed a similar approach that was used for the *Global Map of Accessibility* in the World Bank's World Development Report 2009 Reshaping Economic Geography (Uchida and Nelson 2009). An off-path friction-surface raster was calculated, which is a grid in which each pixel contains the estimated time required to cross that pixel on foot. We assume that all travel is foot based and walking speed is therefore determined by the terrain slope. The slope raster is taken from NASA's Shuttle Radar Topography Mission (SRTM) Digital Elevation Models (DEMs) which has a resolution of 90 meters. The typical velocity of a hiker when walking on uneven or unstable terrain is 1 hour for every 4 kilometers (4 km/hr) and diminishes on steeper terrain. We use a hiking velocity equation[29] (Tobler 1993) to reflect changes in travel speed as a function of trail slope:

$$W = 6e^{-3.5*|S+0.05|}$$

where W is the hiking velocity in km/hr and S is the slope or gradient of the terrain. Finally, we compute the time that it takes to travel from each point in Nigeria to each of our selected markets. The map of Nigeria is divided into a 'fishnet' grid of 10km^2 cells, with approximately 11,000 cells in total. Minimum travel times are calculated using the optimal walking path from the center of each of these 11,000 cells to each of the 65 markets. The algorithm utilizes a node/link cell representation system in which the center of each cell is considered a node and each node is connected to its adjacent nodes by multiple links. Every link has an impedance, which is derived from the time it takes to pass through the cell, according to the natural path friction cost surface, and takes into account the direction of movement through the cell. An ArcGIS/python script was written which creates an optimal path raster for each of the 65 selected cities/markets. This raster defines the optimal path (minimizing walking time), and then records the total time required in each cell. As a result we obtained an origin/destination travel time matrix of more than 11,000 rows (grid cells) and 65 columns (selected markets).

[29] "Three presentations on geographical analysis and modeling non-isotropic geographic modeling speculations on the geometry of geography global spatial analysis". NATIONAL CENTER FOR GEOGRAPHIC INFORMATION AND ANALYSIS. TECHNICAL REPORT 93-1. February 1993

Appendix VI: HDM-4

The Highway Development Management Model (HDM-4) considers several different variables in order to estimate the cost of traveling along each segment of the road network. The data used for the estimates in this paper were collected specifically for Nigeria, to best characterize the transportation conditions one would find there.

In order to estimate the unit cost (in ton per km), the cost of transporting a vehicle with an average weight of 25 tons one kilometer was first estimated. The unit cost per ton-km was derived from the costs per vehicle using a factor of 15 tons per vehicle (average net weight). This factor was obtained based on the assumption of a 30 ton gross vehicle weight, with a 10-ton tare weight and a 75% loading factor.

Characterization of network type and terrain

The road network of Nigeria includes three classes of roads: primary, secondary, and tertiary. Average vehicle speed and width of the main carriage road were used to characterize the differences among network types as follows:

Paved Road Speed (km/hr) by Network & Condition			
Road Condition	Primary 7m	Secondary 6m	Tertiary 5m
Flat	100	80	70
Rolling	80	70	60
Mountainous	60	50	40

Unpaved Road Speed (km/hr) by Network & Condition			
Road Condition	Primary 7m	Secondary 6m	Tertiary 5m
Flat	80	70	60
Rolling	60	50	40
Mountainous	40	30	20

where terrain type is defined using the following concepts and road geometry parameters:

- Flat. Mostly straight and gently undulating
- Rolling. Bendy and gently undulating
- Mountainous. Winding and gently undulating

TERRAIN TYPE	Number			Super-elevation
	Rise & Fall	Rise & Fall	Horizontal Curvature	
	(m/km)	(#)	(deg/km)	(%)
FLAT	10	2	15	2.5
ROLLING	15	2	75	3.0
MOUNTAINOUS	20	3	300	5.0

Characterization of network type and condition

The International Roughness Index IRI (m/km) was used to define the differences in road condition by network as follows:

Paved Road IRI (m/km) by Network & Condition			
Road Condition	Primary 7m	Secondary 6m	Tertiary 5m
Good	2	3	4
Fair	5	6	7
Poor	8	9	10

Unpaved Road IRI (m/km) by Network & Condition			
Road Condition	Primary 7m	Secondary 6m	Tertiary 5m
Good	6	8	10
Fair	12	13	14
Poor	16	18	20

Characterization of vehicle type

A **heavy truck** was defined as the typical vehicle to model freight transport costs. The following key input data was used in the calculation:

FINANCIAL UNIT COSTS	HEAVY TRUCK
Used Vehicle Cost (US$/vehicle)	70,000
New Tire Cost (US$/tire)	800
Fuel Cost (US$/liter)	0.77
Maintenance Labor Cost (US$/hour)	4.73
Crew Cost (US$/hour)	3.15

Finally, using these parameters above, a final cost per ton-km for each road type is estimated ($/ton/km):

Paved FLAT			
Road Condition	**Primary**	**Secondary**	**Tertiary**
Good	0.0526	0.0529	0.0533
Fair	0.0570	0.0583	0.0596
Poor	0.0617	0.0637	0.0986

Paved ROLLING			
Road Condition	**Primary**	**Secondary**	**Tertiary**
Good	0.0533	0.0531	0.0535
Fair	0.0577	0.0586	0.0599
Poor	0.0623	0.0643	0.0996

Paved MOUNTAINOUS			
Road Condition	**Primary**	**Secondary**	**Tertiary**
Good	0.0574	0.0562	0.0584
Fair	0.0620	0.0615	0.0644
Poor	0.0675	0.0676	0.1055

Unpaved FLAT			
Road Condition	**Primary**	**Secondary**	**Tertiary**
Good	0.0629	0.0673	0.0730
Fair	0.0795	0.0831	0.0867
Poor	0.0941	0.1017	0.1091

Unpaved ROLLING			
Road Condition	**Primary**	**Secondary**	**Tertiary**
Good	0.0618	0.0678	0.0752
Fair	0.0801	0.0837	0.0877
Poor	0.0945	0.1021	0.1095

Unpaved MOUNTAINOUS			
Road Condition	**Primary**	**Secondary**	**Tertiary**
Good	0.0651	0.0748	0.0868
Fair	0.0820	0.0884	0.0974
Poor	0.0954	0.1038	0.1130