WPS7272

Policy Research Working Paper 7272

Agricultural Technology Choice and Transport

Rubaba Ali
A. Federico Barra
Claudia N. Berg
Richard Damania
John D. Nash
Jason Russ

WORLD BANK GROUP
Agriculture Global Practice Group
May 2015

Abstract

This paper addresses an old and recurring theme in development economics: the slow adoption of new technologies by farmers in many developing countries. The paper explores a somewhat novel link to explain this puzzle—the link between market access and the incentives to adopt a new technology when there are non-convexities. The paper develops a theoretical model to guide the empirical analysis, which uses spatially disaggregated agricultural production data from Spatial Production Allocation Model and Living Standards Measurement Study survey data for Nigeria. The model is used to estimate the impact of transport costs on crop production, the adoption of modern technologies, and the differential impact on returns of modern versus traditional farmers. To overcome the limitation of data availability on travel costs for much of Africa, road survey data are combined with geographic information road network data to generate the most thorough and accurate road network available. With these data and the Highway Development Management Model, minimum travel costs from each location to the market are computed. Consistent with the theory, analysis finds that transportation costs are critical in determining technology choices, with a greater responsiveness among farmers who adopt modern technologies, and at times a perverse (negative) response to lower transport costs among those who employ more traditional techniques. In sum, the paper presents compelling evidence that the constraints to the adoption of modern technologies and access to markets are interconnected, and so should be targeted jointly.

This paper is a product of the Agriculture Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at rdamania@worldbank.org.

Agricultural Technology Choice and Transport

Rubaba Ali[1]*
A. Federico Barra**
Claudia N. Berg***
Richard Damania**
John D. Nash**
Jason Russ***

JEL Codes: O1, I3, Q1, L9

Keywords: Agriculture, Technology choice, Transport, Nigeria, Africa

[1] Authors in alphabetical order. * University of Maryland; ** World Bank; *** George Washington University

Agricultural Technology Choice and Transport

1. Introduction

Yields across many parts of Sub-Saharan Africa (SSA) have stagnated and even declined in some countries, despite the availability of new agricultural technologies, improvements in transportation infrastructure and policy reforms aimed at market liberalization. A comparison of Africa's performance with that in South America and three subregions of Asia over the course of two decades indicates that Africa showed lower total factor productivity (TFP) growth overall than its counterparts in the 1990s, then fell even further behind in the 2000s, magnifying the TFP gap (World Bank 2013, based on Fuglie et al 2012). As an example, cereal yields in SSA grew at an average annual rate of 0.7 percent between 1980 and 2000, compared to other developing regions which saw growth ranging from 1.2 to 2.3 percent (Slootmaker 2013). Similarly the average yield of tuber crops in SSA is the lowest in the world at around 8 tons per hectare (Pinstrup-Anderson et al 1997). Since the 1980s, fertilizer intensity in SSA has grown on average by 0.93 percent annually compared to 5 percent in South Asia (Slootmaker *op cit*). Furthermore, Gollin *et al* (2005) note that by 2000, modern maize varieties only accounted for an estimated 17 percent of total maize production in Africa, compared to 90 percent in East Asia and South East Asia and 57 percent in Latin America. With a population that is set to double by 2040 at current trends, these are worrying signs for a region that aspires to generate food surpluses.

Numerous explanations have been offered for the low productivity of agriculture in Africa. Early work focused on the need for price reforms to sharpen producer incentives and promote greater competition and efficiency (World Bank 2012). Other studies highlight the role of transactions costs. If there are impediments to market participation, price incentives will be dampened and may even be rendered ineffectual (Schiff and Montenegro 1997). For a recent study of agricultural intensification in Africa, see Binswanger-Mzhize and Savastano (2014). Another strand of literature studies the reasons for the slow adoption of modern technologies and advances numerous explanations including learning impediments, credit constraints, risk, and farmer heterogeneity.[2]

[2] For an early and concise summary see Feder, Just and Zilberman (1985).

Undoubtedly, many factors play a role in depressing the productivity of African agriculture. Our paper makes a modest contribution to this extensive literature by recognizing that agricultural transformation involves interlinked processes. It explores why a significant fraction of farmers in developing countries may not adopt more modern technologies despite the potential for these to increase yields. It is hypothesized that there are non-convexities in switching to alternative technologies. These may arise from a variety of factors, including fixed costs (pecuniary or non-pecuniary) of adopting a new technology or process, indivisibilities (such as the size of a machine), non-linear prices, or increasing returns to scale, among other factors. It is suggested that barriers to the adoption of newer technologies occur from the interaction between these non-convexities and access to markets. In what follows, we focus on this hitherto neglected link.

To frame the analysis, we outline a minimalist model of technology adoption with transport costs. It is assumed that there are fixed costs, or minimum threshold costs, to adopting more modern agricultural inputs (for instance, a minimum rental or purchase price of tractors, harvesters, and planters; learning costs, etc.). Analogous to the big-push literature, a switch from traditional techniques to more modern farming occurs if the gain in payoffs from switching exceeds the fixed costs of the new technology. Unlike the big-push models, or the technology adoption literature, the outcome does not necessarily depend on the decisions of other farmers, demand externalities, risk, or market imperfections. The fixed costs (which can be interpreted broadly) create a hurdle that households must overcome in order to adopt the more productive technology. Transport costs influence the returns to technology adoption and thereby create heterogeneity of responses. Constraints on technology adoption can impede entry into markets and lock farmers into traditional, low input modes of production, while variations in transport costs generate differences in returns.

We examine the relationship between transport cost and technology adoption, and the output of users of modern technology relative to their traditional counterparts in greater detail both theoretically and empirically. Our model generates two testable hypotheses. First, that adoption of new technologies will be more pervasive where transport costs are lower. And second, that reductions in transport costs will have a larger impact on the revenue of farmers already utilizing more modern farming techniques than on the revenue of those still using traditional technologies. The policy implication is that there is a need to target both issues –

lowering transport costs, and removing other barriers farmers face to adopting improved technologies – to trigger transformational improvements in production in developing countries.

The empirical analysis is based on Nigeria. To motivate the issue empirically, we begin by presenting preliminary evidence using spatially disaggregated agricultural production data from IFPRI's Spatial Production Allocation Model (SPAM). The SPAM data set is particularly useful for our purposes as it distinguishes clearly how much of each crop is produced under different input levels and technologies (high input/technology versus low input/technology). Consistent with the theoretical model, these indicative results suggest that farmers' responses to changes in transport costs are heterogeneous and vary systematically with production technology decisions and crop type. Decreasing transport costs increases the production of crops using high input (more modern) production systems, but has no effect or even a seemingly perverse (opposite) effect on production of crops using low input systems. The SPAM data, while useful, are limited and cannot explain when and why farmers switch technology[3] and the precise mechanisms by which transportation costs affect technology choice, and thus, production levels.

To explore these links in further detail, a richer data set is used from the Nigeria Living Standards Measurement Study: Integrated Survey on Agriculture (LSMS-ISA), which contains household-level data on production revenues and technology choices as well as household characteristics (age and literacy of the household head, amongst others). In evaluating the effect of transport costs on agriculture empirically, we consider three aspects. Households face a discrete decision on technology choice—to produce using traditional input levels, or to pay a fixed cost (or cost with a minimum threshold) and use more modern techniques. They also face two continuous decisions on how much to produce and their level of market participation (i.e. what share of crops should be consumed at home versus sold at the market). Consistent with the theory, the results suggest that a) farmers with lower transport costs exhibit a greater proclivity to adopt modern techniques, and b) the revenue of farmers using modern techniques is also more responsive to changes in transport costs.

These findings advance the literature on the role of transport costs in promoting agricultural transitions. Transport infrastructure is costly and despite significant investments, has

[3] For instance, the spatially disaggregated data cannot distinguish whether production levels rise because farmers are switching from low input to high input production, or if this is a consequence of high input farmers producing more output.

often failed to induce a transition to modernization. Likewise, despite the availability of modern farming methods, adoption rates in Africa have been below desired levels. Our analysis suggests the two outcomes are connected and need to be tackled simultaneously. Adoption of modern technologies (and by implication, the production of a surplus) is a prerequisite for enhancing output and promoting market participation. Lower transport costs may assist in stimulating market participation but may not be sufficient, especially among the most marginal farmers. That is, modest changes in transport costs may not in itself stimulate large numbers of farmers to make the technological leap, as there are high barriers to modernization. Hence there is a need for broader policies that consider the bundling of incentives – policies focused on improving technology availability need to consider constraints on connectivity and vice-versa.

The remainder of this paper is organized as follows. Section 2 provides a brief survey of the literature. Section 3 outlines the theoretical model that guides the empirical work. Section 4 discusses two of the novel data sets used herein to estimate transportation costs and their effects. Section 5 presents the empirical analysis using the SPAM data and Section 6 discusses the LSMS analysis. Section 7 concludes.

2. Literature Survey

This section provides a brief review of three prominent though often unconnected strands of literature that seek to explain the slow pace of agricultural transformation, especially in Sub-Saharan Africa. First, that of technology adoption in agriculture; second the big-push literature; and finally a somewhat separate body of literature on the commercialization of farming. The purpose is not to provide a comprehensive review of these burgeoning areas of research, but merely to highlight some of the more prominent contributions.

There is a vast and varied literature analyzing factors that affect technology adoption in the United States (Griliches 1957), and in developing countries (Besley and Case 1993, Foster and Rosenzweig 1995, Conley and Udry 2008, Moser and Barrett 2006, Croppenstedt, Demeke and Meschi 2003). The constraining factors identified in the literature are lack of knowledge about new technology (Besley and Case 1993, Foster and Rosenzweig 1995, Conley and Udry 2008), capital market imperfections (Croppenstedt, Demeke and Meschi 2003) and heterogeneity among farmers (Suri 2011). However in recent years it is risk that has gained the most traction as an explanation for the slow adoption of technologies (Dercon and Christiansen 2011, Lamb

2003). Binding wealth or liquidity constraints combined with uncertain fluctuations in output, can induce the adoption of known and safer technologies. The findings from the literature suggest the need for reducing risk exposure through the provision of (perhaps subsidized) credit or insurance mechanisms, and improvement in access to information related to modern technology through better extension service provision. The resulting policy suggestions have thus targeted largely standard prescriptions.

There is also an extensive literature focused on big-push models and coordination failures. Rosenstein-Rodan (1943) was among the first to examine basic coordination failures. Murphy, Shleifer, and Vishny (1989) formalized the approach, presenting a theoretical model depicting the conditions under which there would be multiple stable equilibria – including an inferior equilibrium, also known as a poverty or underdevelopment trap.[4] In this context it may be necessary for a "big push", such as a large investment, to drive the economy to the superior equilibrium.[5] The implication is that investment in infrastructure will push local economies from traditional sectors to modern industries and thus induce the adoption of superior technologies. An empirical example of the big push literature is Emran and Shilpi (2002), who present evidence of spillovers of sales decisions at the household-level, stemming from increasing returns to scale in marketing.

Finally there is a rich and related strand of literature addressing the determinants of small farmer participation in markets, with a special focus on Africa. The literature begins with the presumption that transactions costs – broadly defined – impede decisions to enter markets. Thus, policies such as the construction of roads and other infrastructure that lower these transaction costs will promote agricultural commercialization (Goetz 1992, Heltberg and Tarp 2001, Key *et al* 2000). Several studies seek to estimate the size of the unobserved transaction costs. As an example, Renkow *et al* (2004) find that transaction costs are the equivalent of a 15 percent *ad valorem* tax in Kenya. Consistent with these results, Abrar *et al* (2004) find that integration into

[4] Under *pecuniary externalities* (where one farmer's actions affect his neighbors' costs) and *complementarities* (where one farmer's actions influenced his neighbors' incentives to take similar actions), there may be two stable equilibria.

[5] For example, Murphy et al (1989) present a model of investment in infrastructure whereby industrialization will only occur if a railroad is built (the logic is easily extended to transport infrastructure more generally). Cottage industries do not need a railroad and can locate anywhere, while modern sectors in contract must cluster around the road.

markets in northern Ethiopia is low and as a result there is little responsiveness to price incentives. There is, however, a dissenting view. Holloway and Lapar (2007) and Lapar, Holloway and Ehui (2003) find that market access is not correlated to market participation once a correction is made for spatial autocorrelation. Takeshima and Winter-Nelson (2011) study the decision making process for market participation and conclude that market participation and production decisions are made sequentially by farmers in Benin. This result confirms earlier work by Bellemare and Barrett (2006). Most of the studies find that wealthier households have a higher probability of selling in markets – though reverse causality cannot be ruled out. While there is much discussion of market access costs in this literature, to our knowledge none of this work considers the problems that arise with the endogeneity of infrastructure – especially roads.[6]

This paper seeks to augment this literature by suggesting that the constraints to technology adoption (modernization of agriculture) and market participation are likely closely linked due to non-convexities in technology adoption. We explore the nature of the links and suggest how policies that reduce transportation costs impact technology adoption and facilitate market participation. Section 3 provides the theoretical motivation for the empirical analysis that follows.

3. The Model

This section outlines a minimalist model which describes the manner in which responses might diverge between different types of farmers. We distinguish between traditional farmers who utilize a lower productivity technology and those who adopt improved technology that generates higher payoffs and yields. In contrast to the existing literature the focus is not on differences in factor endowments, risk, or imperfections in capital or labor markets that might lead to differential responses. Instead, we present a model with identical endowments where differences emerge simply from non-convexities – specifically the need to cover the fixed costs of accessing a more productive technology. Such non-convexities may emerge either from the technology per se (such as if there is a minimum size or other indivisibility), non-linear pricing

[6] A central obstacle to establishing causality is the need to take account of the fact that road placement is not random. Governments build roads in places with the greatest economic potential, or the greatest need, so causal links between road access and economic outcomes may be obscured by policy decisions and estimates will be biased upwards.

(such as when a machine is rented for a minimum amount of time), or information and learning costs (such as learning how to drive a tractor, or grow a new crop) and other forms of inertia.[7] Analogous to the big-push literature – a switch from traditional to more modern farming occurs if the payoffs from switching exceed the fixed costs of adopting the new technology. Unlike the big-push models the outcome does not necessarily depend on the decisions of other farmers, demand factors, or market failures. This theoretical framework fits into a larger family of models of technology choices developed by Mundlak (1988) and later applied by Mundlak et al (1999, 2012) and Larson and León (2006). Essentially, farmers face different circumstances (different roads, markets, climates, etc.) and so choose different technology choices so as to maximize their profit. By making technology choices, farmers move between production functions as well as along them.

For brevity a stylized version of the model which underlies the empirical work is outlined, based on specific functional forms that yield closed form solutions. Generalizations without these functional forms are straightforward.

There are two types of farmers in the model, traditional farmers who use a less productive technology and those with an improved technology such as access to (better or any) machinery. Use of the improved technology requires payment of a fixed cost (F) which enhances the productivity of farming. As noted earlier, F could be interpreted broadly to represent a variety of impediments to adoption – a threshold price on the rental of machinery, learning costs, technological lumpiness and so on. In all other respects the farmers are identical. As discussed above, the fixed costs may be financial or non-pecuniary, and may arise for a number of reasons that have been frequently documented in the literature.

There is only one period with two stages. Production decisions are made sequentially. In the first stage each farmer independently decides whether to pay the fixed costs and adopt the productivity enhancing technology, or remain with the traditional technology. Having made this technology decision, in the second stage each farmer determines how much to produce and how much of this output to sell in a market (or conversely to consume domestically). For simplicity

[7] Even renting space in a truck to carry goods to market represents a lumpy cost (Emran and Shilpi 2002). Hayami and Kawagoe (1993) note that the transportation cost in Indonesia is Rp. 5/kg or higher for a load up to 200kg by pony wagon, but that the cost declines to Rp. 2.5/kg if a two-ton load is carried by a small truck.

we start with the case of two farmers indexed *i* (more modern) and *j* (traditional). By backward induction the final stage is solved first.

Modern farmers

The utility function of the modern farmer is simply given by:

$$U_i = B_i^{\beta}(\delta_i y_i)^{\alpha} \tag{3.1}$$

where B_i is the quantity of goods bought from the market for consumption by farmer *i*, δ_i is the proportion of goods produced, y_i, that are consumed domestically (i.e. the amount sold to markets is (1 - δ_i)y_i). Equation (3.1) is maximized subject to the following constraints. The production function is $y_i = W_i^{\gamma}$, where $1 > \gamma > 0$ and the input *W* is supplied under competitive markets at price *v*. The budget constraint is:

$$(1-\delta_i)(p-t)W_i^{\gamma} = (p_B + t_B)B_i + vW + F \tag{3.2}$$

Where *t* and t_B are transport costs for goods *y* and *B* respectively; and *p* and p_B are the given market prices of farm output and consumption goods respectively. *F* is the fixed cost for using technology $y_i = W_i^{\gamma}$ and *v* is the variable cost. By equation (3.2) sales of goods produced ($(1-\delta_i)(p-t)W_i^{\gamma}$) must equal total expenditures; i.e. the sum of spending for consumption goods $\left((p_B + t_B)B_i\right)$, the purchase of the input (vW) and the fixed cost for using the improved technology (*F*). Maximizing (3.1) subject to (3.2) yields the following first-order-conditions.

$$\frac{dL}{dB} = \beta B^{\beta-1}\delta_i^{\alpha}W_i^{\varepsilon} - \lambda(p_B + t_B) = 0 \tag{3.3}$$

where, $\varepsilon = \alpha\gamma$

$$\frac{dL}{dW} = \varepsilon B^{\beta}\delta_i^{\alpha}W_i^{\varepsilon-1} + \lambda\left(\varepsilon(p-t)(1-\delta_i)W_i^{\gamma-1} - v\right) = 0 \tag{3.4}$$

$$\frac{dL}{d\delta_i} = \alpha B^{\beta}\delta_i^{\alpha-1}W_i^{\varepsilon} - \lambda\left((p-t)W_i^{\gamma}\right) = 0 \tag{3.5}$$

$$\frac{dL}{d\lambda} = (1-\delta_i)(p-t)W_i^{\gamma} - (p_B + t_B)B_i - vW - F = 0 \tag{3.6}$$

Which can be solved for the endogenous variables:

$$W_i = \left(\frac{\gamma(p-t)}{v}\right)^{\frac{1}{1-\gamma}} \tag{3.7}$$

$$\delta_i = \frac{\varepsilon(p_B - t_B)}{\beta}\left(\frac{1}{(p-t)}\left(\frac{v}{\gamma}\right)^{\gamma}\right)^{\frac{1}{1-\gamma}} \tag{3.8}$$

$$B_i = \frac{1-F-\left(\frac{\gamma(p-t)}{v^{\gamma}}\right)^{\frac{1}{1-\gamma}}}{p_{B+}t_B} - \frac{\varepsilon}{\beta} \tag{3.9}$$

Substituting these into (3.1) defines the indirect utility function:

$$U_i^* = \left(\frac{1-F-\left(\frac{\gamma(p-t)}{v^{\gamma}}\right)^{\frac{1}{1-\gamma}}}{p_{B+}t_B}\right)^{\beta}\left(\frac{\varepsilon(p_B+t_B)}{\beta}\right)^{\alpha} \tag{3.10}$$

Traditional Farmer

The farmer with the traditional technology has identical preferences. The only difference is that they pay no fixed costs and thus utilize a less efficient technology $y_j = W_j^{\eta}$, with $0 < \eta < \gamma < 1$. The traditional farmer's maximization problem is given by:

$$Max\ U_j = B_j^{\beta}(\delta_j y_j)^{\alpha} \tag{3.11}$$

Subject to:

$$(1-\delta_j)(p-t)W_j^{\eta} = (p_B + t_B)B_j + vW_j \tag{3.12}$$

Which by an analogous procedure yields the indirect utility function:

$$U_j^* = \left(\frac{1-\left(\frac{\eta(p-t)}{v^{\eta}}\right)^{\frac{1}{1-\eta}}}{p_{B+}t_B}\right)^{\beta}\left(\frac{\varepsilon(p_B+t_B)}{\beta}\right)^{\alpha} \tag{3.13}$$

Technology Choice

Given these production and utility levels, in stage one farmers will (or will not) switch from subsistence farming to modern farming if $U_i^* > (<) U_j^*$. Substituting from (3.10) and (3.13) and rearranging we obtain:

$$U_i^* - U_j^* < 0 \text{ if } \Psi < F \tag{3.14}$$

where $\Psi \equiv \left(\frac{\gamma(p-t)}{v^\gamma}\right)^{\frac{1}{1-\gamma}} - \left(\frac{\eta(p-t)}{v^\eta}\right)^{\frac{1}{1-\eta}}$

Suppose next that farmers are situated at a range of locations with differing transport costs $t_h \in [t_0, \ldots, \tilde{t}, \ldots T]$, where $t_0 < \tilde{t} < T$. Define $\tilde{t}$ such that for some $\tilde{F} > 0$, $\tilde{\Psi} = \left(\frac{\gamma(p-\tilde{t})}{v^\gamma}\right)^{\frac{1}{1-\gamma}} - \left(\frac{\eta(p-\tilde{t})}{v^\eta}\right)^{\frac{1}{1-\eta}} = \tilde{F}$. Observe that at $\tilde{\Psi}$ farmers are indifferent between the technologies. We obtain the following results:

Result 1: At locations $t_h < \tilde{t}$, farmers will find it profitable to switch to the improved technology, whereas at locations with $t_h > \tilde{t}$, farmers will remain with the old technology.

Proof: Note that $\frac{d\Psi}{dt} = -\frac{1}{1-\gamma}\left(\frac{\gamma(p-t)}{v^\gamma}\right)^{\frac{1}{1-\gamma}} + \frac{1}{1-\eta}\left(\frac{\eta(p-t)}{v^\eta}\right)^{\frac{1}{1-\eta}} < 0$ since $\eta < \gamma$ by construction. Since Ψ is continuous in *t*, then from the Intermediate Value Theorem it follows that for locations where $t_h < \tilde{t}$, then $\Psi > \tilde{F}$ and these farmers switch to the improved technology, whereas on locations with $t_h > \tilde{t}$, then $\Psi < \tilde{F}$ and it pays to remain with the old technology. ■

Let $E_k = \frac{\partial y_k}{\partial t}\frac{t}{y_k}$, ($k = i, j$) be the output elasticity of demand with respect to transport costs of goods sold in markets.

Results 2: The outputs of farmers who adopt more modern, improved technologies is more responsive to changes in transport costs than that of farmers who utilize the old technology (i.e. $|E_i| > |E_j|$).

Proof: Note that since $y_i = \left(\frac{\gamma(p-t)}{v}\right)^{\frac{\gamma}{1-\gamma}}$ then $E_i = \frac{-\gamma t}{(p-t)(1-\gamma)}$ and by analogy $E_j = \frac{-\eta t}{(p-t)(1-\eta)}$. Since by assumption $\gamma > \eta$, then it follows that $|E_i| > |E_j|$. ■

The analysis therefore suggests two hypotheses that we test in subsequent sections. First, the model suggests that adoption of new technologies will be more pervasive where transport costs are lower (Result 1). Second, the model predicts that reductions in transport costs will have

a larger impact on the output of farmers utilizing modern farming techniques (Results 2). Sections 5 and 6 seek to test these hypotheses using a variety of publicly available data sets.

4. Data on Transportation Costs and Instruments

To estimate the effect of transportation costs on technology choice and crop production requires data on the complete road network for Nigeria, as well as the associated costs of moving between any two points. To our knowledge, no complete, definitive data set exists which accurately includes the entire road network for Nigeria. Recognizing the importance of accurately estimating transportation costs, several sources of data were utilized to construct the transportation network that is used in this study. The road network vector data was obtained from Delorme[8], which appears to be the only available data set that includes both trunk roads, as well as the rural road network. The costs of travel are calculated taking account of road attributes (including road class: primary, secondary or tertiary; paving status; and road quality: good, fair, and poor) that were obtained from two sources: Nigeria's Federal Roads Maintenance Agency (FERMA), and a specialized road survey that was conducted using the World Bank's FADAMA program.[9] This information was digitized and merged into the Delorme data set.[10] Next, costs of traveling along each segment of the road were computed using the Highway Development Management Model (HDM-4), which is a standard application used by road engineers that considers the road attributes, as well as roughness of the terrain, and country level factors (for example, the price of fuel, the cost of purchasing and maintaining a heavy truck, and wages), in order to calculate the cost, in USD, of transporting one ton of goods along each segment of the road network.[11] With this information, the cost of traveling from each SPAM cell (Section 5) or each LSMS household (Section 6) to the cheapest market was calculated, where a market is defined as a city with at least 100,000 people. This was done using an iterative algorithm in which the cost of traveling to every market along every possible route is calculated for each location. The cheapest market/route combination is then selected.

[8] Delorme is a company which specializes in GPS devises and has compiled a very thorough network of geo-referenced roads across Africa

[9] 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

[10] Roads that are in the Delorme data set but do not appear in FERMA or our survey data are characterized as being of the lowest quality (tertiary class, unpaved, and poor condition).

[11] For specific details on what goes into the HDM-4 model, and on the final estimated costs, see Appendix I

Endogeneity bias is a familiar problem when estimating the impact of roads. Roads are non-randomly placed and tend to be built to connect areas with higher economic potential. Hence, estimates need to take these placement effects into account when looking at the impact of roads: there could be higher levels of economic activity in an area because of a higher density of roads, or there might be a higher density of roads because there was originally greater economic potential.

In order to correct for the endogeneity of road placement this analysis uses a novel instrumental variable (IV) for transport costs which we refer to as the *natural path*. The natural path variable measures the time it would take to walk from the area of production (for SPAM this is the center of each gridded cell in Nigeria, and for the LSMS it is the geo-referenced location of each enumeration area) to the cheapest market as described above, in the absence of any transportation infrastructure. Under the assumption that impediments to pedestrians correlate highly with impediments to road construction, the natural path variable is the best estimate for the ideal placement of a road in a cost-minimization only framework (i.e. positioning the path of a road to be most cost efficient, without making deviations from this path to connect areas with higher economic potential). The natural path variable is therefore completely exogenous with respect to the economic benefits of the transportation network. It also correlates highly with transportation costs because points which require a shorter amount of time to walk between generally also have lower transportation costs (both because they are usually closer together, and because they have less rough terrain between them). This is a similar approach to that of Faber (2014), who uses a hypothetical least cost path.

This variable represents a significant improvement over the Euclidean Distance, or "straight-line", IVs typically used in the literature, as it better captures what the straight-line is trying to measure; namely, the most logical path to the market absent any confounding, bias-causing factors. Straight-line IVs, while they do capture the shortest path "as the crow flies" between any two points, do not consider the fact that the underlying terrain may be impassable, or especially costly to travel through. They are therefore not as accurate an approximation of the cheapest route to construct a road between two points as the natural path, and therefore may be a less efficient instrument. See Appendix II for more information on how this variable was constructed.

5. SPAM Estimates

In order to motivate our empirical analysis, we first use spatial data to determine the impact of transportation costs on crop production under different input systems. We use a data set on crop production, SPAM, which disaggregates crop production statistics into 10km x 10km cells throughout the entire country of Nigeria. This data set is useful for our purposes since it clearly distinguishes between production of various crops using different input systems (labeled low input and high input, with the former referred to as "subsistence/low input" and the latter referred to as "irrigated", although it encompasses mechanized and all input intensive forms of agriculture).[12] These categories map into conventional definitions of traditional and modern production. Examining the effects of transport costs on low and high input systems separately provides an indication of the possibly differential impacts of transport costs on agricultural output (as suggested by Result 2 from our empirical model).

5.1 SPAM Model and Data

The data used in this analysis are spatially organized into a gridded framework. The total land area of Nigeria is split into pixels of 5 arc minute x 5 arc minute (approximately 10km x 10km), which line up with the SPAM cells described above. Each pixel is a unique observation. An agricultural production function was estimated for four crops: yams, rice, cassava, and maize. These crops were chosen because they are the most widely grown in Nigeria, and combined they represent more than 60% of the total agricultural production value in Nigeria in 2011 according to FAO.[13] Cassava and maize are typically grown using low inputs, for home consumption (in fact, in Nigeria, neither of these crops have high input or irrigated production according to SPAM), while yams and rice are more commonly marketed and are grown under both low input and high input crop management regimes. The agricultural production function takes the form of:

$$\ln(Y_{ik}^{j}) = \beta_{0k}^{j} + \beta_{1k}^{j}\ln(T_i) + X_{ik}^{j}{}'\delta^{j} + \varepsilon_{ik}^{j}; \tag{5.1}$$

where Y_{ik}^{j} is total production of crop *j* in pixel *i* using input system *k*, where *k* can be traditional (subsistence, as coded by SPAM) or modern (irrigated) inputs. T_i is transportation costs from cell

[12] For instance, subsistence input systems are defined as "rainfed crop production which uses traditional varieties and mainly manual labor without (or with little) application of nutrients or chemicals for pest and disease control", and irrigated input systems refers to production in which "the crop area [is] equipped with either full or partial control irrigation…[normally using]…high level of inputs such as modern varieties and fertilizer as well as advanced management such as soil/water conservation measures.

[13] See FAOSTAT: http://faostat.fao.org/site/339/default.aspx

i to the cheapest market, with a market being defined as a city of at least 100,000 residents, and X_{ik}^{j} is a vector of control variables. To correct for potential bias due to the non-random placement of roads, Equation 5.1 is estimated using two-stage least squares, where the natural path variable described in Section 4 is used as the instrument for transportation costs.

Several spatial data sets were utilized or generated to control for confounding variables, including population, agro-ecological production potential, and distance to mining facilities. Population data come from Landscan[14] (2009), which uses satellite imagery analysis to disaggregate census data into a gridded network. Agro-ecological potential data are from GAEZ, a product of FAO, which considers climate and soil conditions to estimate the maximum potential yields in each pixel, for each crop.[15] Euclidean distance to the nearest mining facility is calculated using data from the National Minerals Information Center of the USGS, and is included to account for the fact that mining facilities often have high concentrations of workers and their families, making them very high demand centers.[16] Table 5.1 provides summary statistics for the variables used in this analysis. Because this analysis is focused on agricultural production, which mainly occurs in rural areas, urban areas were removed from the data set.[17] In addition, pixels falling entirely in water, and pixels in which there is zero agricultural potential according to GAEZ for the particular crop being regressed are also omitted from the analysis. All other pixels are included.

5.2 SPAM Results

The results from the SPAM analysis for yam and rice are displayed in Tables 5.2 and 5.3, respectively. Only traditional level production results are given for cassava and maize in Tables

[14] LandScan population data available here: http://web.ornl.gov/sci/landscan/landscan_data_avail.shtml

[15] The data used in this model assumes climactic conditions similar to the 1961-1990 baseline level. GAEZ also differentiates potential yields by input system. For our traditional input equation, we use potential yields under low inputs, and for our modern input equation, we use potential yields under high inputs.

[16] We considered only a subset of mining facilities available in the raw data. 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 data set 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 only considered plants which were considered active between 2006 and 2010.

[17] The methodology for determining which pixels of the Landscan data set are urban areas went as follows. Nigeria's urbanization rate as defined by the Central Intelligence Agency World Factbook was 49.6% in 2011 (see: https://www.cia.gov/library/publications/the-world-factbook/fields/2212.html). The total population in the Landscan data set is approximately 136 million, implying an urban population of 67 million. The pixels with the largest number of people according to Landscan are marked as being urban pixels until the total number of people living in these marked pixels equals 67 million. These marked pixel are then omitted from the regressions.

5.4 and 5.5, respectively, as there is no modern input production of these crops in Nigeria, according to SPAM. In each table, regressions using ordinary least squares and two-stage least squares are summarized.

Turning first to the variable of interest, the cost of traveling to market, column (2) in Tables 5.2 and 5.3 show that as the cost to market increases, irrigated production of yams and rice decline with elasticities of -0.31 and -0.53, respectively. This is as we would expect—lower transport costs leads to greater, high-input production of these two crops. However, the results are strikingly different when these crops are produced under low input production systems. Notably, transportation costs have no statistically significant effect on low input production of yams and cassava, and low input production of rice and maize actually increase when transport costs rise. In the context of our model, there are two reasons why this relationship might hold. The first reason is that low input farmers do not generate sufficient marketable surplus to benefit from cheaper transportation to markets. This finding is consistent with Result 2 from the theoretical model in Section 3 and is a plausible explanation for the results observed for cassava and yams. The results for rice and maize, however, require a different explanation since not only does reducing transport costs not lead to increased low input production of these two crops, it actually leads to decreased levels. This could be explained by Result 1: that when transportation costs decline, farmers find it profitable to switch to more modern inputs. This leads to a reduction in low-input production in the aggregate statistics, while at the same time aggregate production of these crops actually increases- some production is simply shifted from low input to high input agriculture.

Briefly turning to the control variables, population has a positive, but diminishing effect on production of the four crops, under both input systems, with the exception of irrigated rice where coefficients on both the linear and squared terms are insignificant. This is what one would expect—when holding other inputs constant, additional labor will lead to more production, but will exhibit diminishing marginal returns as land becomes crowded. Further, crop production tends to decline with distance to a mining facility, as was originally hypothesized. Finally, all four crops exhibit a non-linear relationship between agricultural potential (land and climate suitability) and actual production. High input production of yams and rice display a convex relationship between the two, implying increasing returns to crop suitability. Conversely low input yams, rice, maize and cassava, show a concave relationship, implying diminishing returns

to crop suitability. This relationship is perhaps not surprising. Low input producers are not able to take advantage of better climate and soil conditions as much as high input producers might be.

The empirical model presented here provides preliminary evidence of the results predicted by our theoretical model. However, due to the fact that the data set is aggregated at the geographic level, rather than at the household, it is not possible to assess directly whether variations in transport costs induce shifts in technology choice. The next section tests for this and the other predictions of our empirical model, using household level survey data from the LSMS-ISA.

6. LSMS Analysis

Next, we extend the analysis by focusing on the household level using LSMS-ISA data. This rich survey data enables us to explore the underlying mechanism behind increased production and to further test the hypotheses suggested by the theoretical model in Section 3. We test the impact of transport costs on the probability of adopting modern technology (Result 1), and how the impact of transport costs varies with technology choices (Result 2).

6.1 LSMS-ISA Model and Data

To test the above hypotheses, a two-step treatment effects model is employed, as summarized in equations (6.1) and (6.2):

$$dM_i = I(\alpha_0 + \alpha_1 \ln(T_i) + \alpha_2 \theta_i + X_i^{'} \delta + v_i > 0), \quad (6.1)$$

$$\ln(R_i) = \beta_0 + \beta_1 \ln(T_i) + X_i^{'} \gamma + \rho \lambda_i + u_i, \text{ if } dM_i = 1. \quad (6.2)$$

Equation 6.1 is the treatment equation suggesting that the choice of modern technology ($dM_i = 1$) depends on transport costs (T_i), an exclusion restriction (θ_i), and a vector of control variables (X_i).[18] Equation 6.2 investigates how transport costs (T_i) influence crop revenue (R_i) for those farmers that adopt more modern techniques (i.e. the subsample for which $dM_i = 1$). Equation 6.2 obviously omits the exclusion restriction used in Equation 6.1 (which is described below) and uses as controls the other variables in 6.1 as well as a set of additional controls. These include whether the plot is irrigated, the number of household members engaged in agricultural labor,

[18] The control variables we include in the selection equation are total land of the household, age and age squared of the household head, a dummy indicating whether the head is literate, amount of fertilizer purchased, and distance to the nearest mine.

and fixed effects indicating the agro-ecological zone. (There are four agro-ecological zones in Nigeria: tropic-warm/semi-arid, tropic-warm/sub-humid, tropic-warm/humid, and tropic-cool/sub-humid.) To account for possible selection bias, the Inverse Mills Ratio (λ_i) enters equation 6.2 and is discussed further below.

Machinery usage is used as a proxy for the adoption of more modern technologies (dM_i). The choice of this indicator is guided by pragmatism and data availability. While information is also available on other input indicators such as high yielding varieties (HYV) of seeds, pesticides, herbicides, and fertilizers their use patterns are known to be distorted by the widespread distribution of free or subsidized inputs throughout Nigeria, which would potentially bias the estimates. Irrigation is also not a suitable indicator in this context because it is more a function of geography and where a farmer is located, rather than technology decisions. Machinery, in contrast, must be either bought or rented and is typically associated with a greater investment in inputs.

The parameters of interest are α_1 and β_1 which indicate the causal effect of transport costs on the probability of becoming mechanized (Result 1) and on crop revenue of mechanized farmers (Result 2), respectively. As noted earlier, the endogeneity of roads complicates the identification of these parameters, as does the potential selection bias in equation 6.2.

To address both the endogeneity and potential sample selection, we implement a two-stage estimation process. First, equation 6.1 is estimated using IV-Probit, followed by the estimation of 6.2 by two-stage least squares. The endogeneity issue is overcome in both equations by instrumenting for transport costs with the natural path variable. In equation 6.2 the Inverse Mills Ratio accounts for possible sample selection bias.[19] Two alternative exclusion restrictions (θ_i) are tested. The first is non-agricultural income of the household. There are typically indivisibilities in the purchase or rental of machinery (such as a minimum cost). As such, farmers with greater endowments of non-agricultural income would be better positioned to pay for these. Non-agricultural income should therefore positively influence the likelihood of machinery adoption, but do not necessarily have a direct effect on actual crop revenue once selection is taken into account. Further, we control for the amount of other inputs, such as

[19] Given that the endogenous variable, transport costs, appears in the selection equation (6.1), we proxy for it with the natural path variable. We then estimate the modified equation by Probit, and use the fitted values to calculate the inverse mills ratio (λ_i).

fertilizer purchased, to account for the possibility that income may be spent on other agricultural inputs besides machinery.

The second exclusion restriction tested is neighborhood effects of machinery use in the area. As is customary neighborhood effects are computed by calculating the percentage of households using machinery within the enumeration area, leaving out the household under consideration. It is well established in the literature that neighborhood effects have a strong influence on a household's adoption of new technologies. For example, Conley and Udry (2008) find strong evidence that pineapple farmers in Ghana adjust their fertilizer use on the basis of their neighbors' experiences. Thus, we expect the neighborhood effects of machinery use to positively influence the likelihood of adopting machinery, but not to impact the revenue from crop production. Because richer areas (those with higher non-agricultural income) are more likely to use machinery, non-agricultural income and average machinery use within an enumeration area are highly correlated. We therefore include them one at a time, and report both results as a robustness check.

6.2 LSMS Results

Table 6.2 presents the IV-Probit estimates of equation 6.1, using two alternative specifications. We test the two exclusion restrictions described above: in column (1) we include the natural log of non-agricultural income and in column (2) we include the neighborhood effects of machinery use instead, with transport costs instrumented with the natural path IV. In both cases, we see that the exclusion restrictions are positive and significant at at least the 5% level, giving support that they are both valid exclusion restrictions. In column (1), we see that a 10% increase in transportation costs reduces the probability that a farmer uses mechanized farming techniques by 1.8%. Similarly, estimates in column (2) indicate that a 10% increase in transportation costs reduces that same probability by 2.5% (which, according to a difference in means test is indistinguishable from the coefficient in the former model). In both cases, the coefficient is significant at the 1% level. This is clear evidence in support of Result 1 – that reducing transportation costs induces a "technological switch" from non-mechanized, to mechanized farming.

Next, we explore the results from the second stage of the two-step Heckman. Table 6.3 reports the IV-estimates of the augmented outcome equation, which includes the Inverse Mills

Ratio (IMR) as an additional control. Column (1) includes IMR_1, which was calculated using the log non-agricultural income, and column (2) includes IMR_2, which is computed using machinery-use neighborhood effects.[20] Control variables include those from the selection equation in addition to whether the plot is irrigated, agricultural labor, as well as climate (i.e. agro-ecological zone fixed effects).

In both cases, the results indicate that cost to market has a significantly negative impact on the crop revenue of mechanized farmers. We find that decreasing transport costs by 10% would increase crop revenue by between 22 and 25%, respectively. While these elasticities seem quite high, it is important to note that they reflect the benefit from reducing transport costs for those farmers who are most impacted by transportation costs—mechanized, and more commercialized farmers. They also represent gross benefits, and are not net of input costs. Note also that the coefficients on both alternate IMRs are not significant, suggesting that selection bias is not present.

As a robustness check, we estimate an alternative specification in which we omit the variables from the two-stage least squares specification that were not present in the selection equation (namely, whether the plot was irrigated, labor, and regional fixed effects). These results are reported in Table 6.4. Here too, we find that transport costs exert a significantly negative effect on the crop revenues of mechanized farmers, of the same magnitude as in Table 6.3. Further, we again find that the IMRs are not significant.

Given the absence of evidence of selection bias, IV estimates are presented which do not control for the IMR in Table 6.5. Column (1) estimates the model for the full sample, column (2) focuses on the subsample of mechanized farmers, and column (3) the traditional farmers. There is striking evidence of a heterogeneity of impacts that vary with technology adoption. As predicted by our theoretical model, the coefficient on transport costs is significantly lower for non-mechanized farmers, at -1.4 (Table 6.4, column 3) than it is for mechanized farmers, -2.4 (Table 6.4, column 2). This is clear evidence of Result 2: reducing transport costs benefits farmers using modern inputs significantly more than traditional-input farmers.

In sum, we find compelling evidence to support our hypotheses that reducing transport costs will entice more farmers to switch to more mechanized forms of agriculture, and that the

[20] Recall that both non-agricultural income and machinery neighborhood effects are excluded from the outcome equation.

mechanized farmers are more sensitive to variations on transport costs. These findings are robust to alternative specifications and to controlling for sources of endogeneity and selection biases.

7. Conclusions

While it has long been recognized in the literature that reduction of transport costs will induce greater market participation and increase welfare, we take a fresh look at the underlying mechanisms.

Using two sources of data—SPAM and LSMS-ISA—we analyze the impact of transportation costs on the adoption of modern inputs and on its potentially differential effect on modern versus traditional farmers. We present robust evidence that a reduction of transport costs will increase the adoption of modern technologies. Further, we demonstrate that transport costs have a greater impact on modern as compared to traditional farmers.

From the SPAM analysis, we find that when transport costs are reduced, crop production under a high-input regime is increased, while crops produced under a low-input regime are either unaffected or may see a decline in output. It is likely that this decline in low-input production is explained by the switch towards more modern agriculture. From the LSMS, we find evidence supporting our hypothesis that reducing transport costs increases the likelihood of modern technology use. Further, we find that all farmers (both modern and traditional) see an increase in crop revenue when transport costs decline, but modern farmers' revenue increases by a significantly greater margin. These findings are consistent with the implications of our theoretical model.

In sum, we present compelling evidence that the constraints to the adoption of modern technologies and access to markets are interconnected and so they should be targeted jointly. For example, reducing transport costs to the market may not be enough to push the local economy towards a more favorable equilibrium. It may be necessary to also expand access to credit, which would enable farmers to cover the fixed costs involved in modernizing. By the same token, expanding credit by itself may not be enough of a push either, if market access is insufficient or too costly. In order to increase yields in SSA, there is a need for policies which bundle these interventions, focusing on improving both technology availability and connectivity.

References

Abrar, S., O., Morrissey and T., Rayner, (2004), Crop Level Supply Response by Agro-Climatic Region in Ethiopia, Journal of Agricultural Economics, 55-2:289-311.

Bellemare, Mark F. and Christopher B. Barrett (2006) “An Ordered Tobit Model of Market Participation: Evidence from Kenya and Ethiopia” Amer. J. Agr. Econ. 88(2) (May 2006): 324–337

Besley, Timothy and Anne Case (1993) “Modeling Technology Adoption in Developing Countries”, American Economic Review, Vol. 83, No. 2, (May 1993) pp. 396-402.

Binswanger-Mkhize, Hans and Sara Savastano (2014) “Agricultural Intensification: The Status in Six African Countries” World Bank Policy Research Working Paper 7116

Conley, Timothy G. and Christopher R. Udry (2008) “Learning About a New Technology: Pineapple in Ghana” Working Paper.

Croppenstedt, Andre, Mulat Demeke, and Meloria M. Meschi (2003) “Technology Adoption in the Presence of Constraints: the Case of Fertilizer Demand in Ethiopia” Review of Development Economics, Vol. 7 No. 1, pp. 58–70

Dercon, Stefan & Christiaensen, Luc (2011) “Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia,” Journal of Development Economics, Elsevier, vol. 96(2), pages 159-173, November.

Emran, Shahe and Forhad Shilpi (2002) “Marketing Externalities and Market Development” *World Bank Policy Research Working Paper*, No. 2839

Faber, Benjamin. "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System*." The Review of Economic Studies (2014): rdu010.

Feder, Gershon, Richard E. Just, and David Zilberman (1985) “Adoption of Agricultural Innovations in Developing Countries: A Survey.” Economic Development and Cultural Change 33:255-299.

Foster, Andrew D. and Mark R. Rosenzweig (1995) “Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture” The Journal of Political Economy, Vol. 103, No. 6 (Dec. 1995), pp. 1176-1209

Fuglie, Keith O., Sun Ling Wang, and V. Eldon Ball (eds.) (2012) Productivity Growth in Agriculture: An International Perspective, by, CAB International, Oxfordshire, UK, January.

Goetz, Stephan J. (1992) “A Selectivity Model of Houshold Food Marketing Behavior in Sub-Saharan Africa”, American Journal of Agricultural Economics 74(2), pp. 444-452.

Gollin, Douglas, Michael Morris, and Derek Byerlee (2005) “Technology Adoption in Intensive Post-Green Revolution Systems” American Journal of Agricultural Economics, Vol. 87, No. 5, Proceedings Issue (Dec., 2005), pp. 1310-1316

Griliches, Zvi (1957) “An Exploration in the Economics of Technological Change” Econometrica, Vol. 25, No. 4 (Oct., 1957), pp. 501-522.

Hayami, Yujiro and Toshihiko Kawagoe (1993) *The Agrarian Origins of Commerce and Industry: A study of Peasent Marketing in Indonesia*, St. Martin’s Press.

Heckman, James J. (1976) "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models". Annals of Economic and Social Measurement, 5(4): 475–492.

Heltberg, Rasmus and Finn Tarp (2001) “Agricultural Supply Response and Poverty in Mozambique” WIDER Discussion Paper No. 2001/114

Holloway, Garth and Ma. Lucila A. Lapar (2003) “How Big is Your Neighbourhood? Spatial Implications of Market Participation Among Filipino Smallholders” Journal of Agricultural Economics, Vol. 58, No. 1, 2007, 37–60

Key, Nigel, Elisabeth Sadoulet, and Alain de Janvry (2000) “Transaction Costs and Agricultural Household Response.” American Journal of Agricultural Economics, 82 (2), 245-259.

Lamb, Russel L., 2003, “Fertiliser Use, Risk and Off-Farm Labor Markets in the Semi-Arid Tropics of India,” American Journal of Agricultural Economics, 85-2: pp. 359-371.

Moser, Christine M. and Christopher B. Barrett, (2006) “The complex dynamics of smallholder technology adoption: the case of SRI in Madagascar” Agricultural Economics, International Association of Agricultural Economists, vol. 35(3), pages 373-388, November.

Mundlak, Yair (1988) “Endogenous Technology and the Measurement of Productivity” Agricultural Productivity: Measurement and Explanation (ed. Susan M. Dapalbo and John M. Antle), Washington, D.C., IFRPI

Mundlak, Yair, Rita Butzer, and Donald Larson (2012) “Heterogeneous technology and panel data: the case of the agricultural production function,” Journal of Development Economics 99 (1).

Mundlak, Yair, Donald Larson, and Rita Butzer (1999) “Rethinking within and between regressions: The case of agricultural production functions,” Annales d’Economie et Statistique, pp. 55-56

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny (1989) “Industrialization and the Big Push,” *Journal of Political Economy*, vol. 97, no. 5, pp. 1003-1026.

Pinstrup-Andersen, Per, Rajul Pandya-Lorch, and Mark W. Rosegrant. "The world food situation: Recent developments, emerging issues, and long-term prospects." (1997).

Renkow, M., D. Hallstrom, D. Karanja (2004) “Rural infrastructure, transactions cost and market participation in Kenya.” Journal of Development Economics 73: 349–367.

Rosenstein-Rodan, Paul (1943) “Problems of Industrialization of Eastern and Southeastern Europe.” *Economic Journal* 53 (June-September), pp. 202-211.

Slootmaker, Chris 2013 “Technology Adoption, Risk, and Intrahousehold Bargaining in Subsistence Agriculture” Working Paper, October 2013.

Schiff, Maurice & Montenegro, Claudio E, 1997. “Aggregate Agricultural Supply Response in Developing Countries: A Survey of Selected Issues,” Economic Development and Cultural Change, University of Chicago Press, vol. 45(2), pages 393-410, January.

Suri, Tavneet (2011), “Selection and Comparative Advantage in Technology Adoption”, Econometrica, 79 (1), pp. 159-209, January 2011.

Takeshima, Hiroyuki and Alex Winter-Nelson (2011). “Sales Location of Semi-Subsistence Cassava Farmers in Benin.” Agricultural Economics 53: 655-670.

Lapar, Ma. Lucila A., Garth Holloway, and Simeon Ehui (2003) “How Big is your Neighborhood? Spatial Implications of Market Participation by Smallholder Livestock Producers?” Working Paper.

Larson, Donald F. and Mauricio León. "How endowments, accumulations, and choice determine the geography of agricultural productivity in Ecuador." The World Bank Economic Review 20(3), (2006).

World Bank (2012) Agriculture, Policy Reform and Poverty in Sub Saharan Africa, World Bank Discussion Paper 280.

World Bank (2013) Growing Africa: Unlocking the Potential of Agribusiness. Washington, DC: World Bank.

Table 5.1 SPAM Summary Statistics

Variable	Obs	Mean	St. Dev.	Min	Max	Label
Subsistence Production						
Yams	10015	1072.397	2256.246	0	29988.5	Production, Tons (SPAM 2006)
Rice	10015	93.25695	136.4678	0	1019.2	Production, Tons (SPAM 2006)
Cassava	10015	3419.833	8387.33	0	143613.1	Production, Tons (SPAM 2006)
Maize	10015	491.527	854.9417	0	10103.8	Production, Tons (SPAM 2006)
Irrigated Production						
Yams	10015	1115.142	7199.12	0	179121	Production, Tons (SPAM 2006)
Rice	10015	196.0147	1249.861	0	31679.5	Production, Tons (SPAM 2006)
Low Input Potential Yield						
Yams	10015	609.166	383.2807	0	1747	Yield (Kg/Ha)
Rice	10015	494.7672	436.5076	0	1792	Yield (Kg/Ha)
Cassava	10015	833.5371	681.5004	0	2775	Yield (Kg/Ha)
Maize	10015	1209.439	625.5546	0	3556	Yield (Kg/Ha)
High Input Potential Yield						
Yams	10015	3253.258	1780.383	0	7028	Yield (Kg/Ha)
Rice	10015	1646.77	1532.992	0	6389	Yield (Kg/Ha)
Other Controls						
Population	10015	13639.3	44454.81	0	1.639MM	Population (Landscan 2009)
Market Transport Cost	10015	7.989746	4.76409	0.1	38.2	Cost to Market (US$)
Natural Path	10015	20.98574	13.42766	0	81.02	Time to Cheapest Market (hrs)
Distance to mine	10003	189.7956	110.7933	2.9	499.4	Euclidean Distance (kms)

Table 5.2: Yams

Dependent Variable: Ln(Yams Production)	**Modern Inputs**		**Traditional Inputs**	
	(1)	(2)	(3)	(4)
	OLS	2SLS	OLS	2SLS
Ln(Market Transport Cost)	-0.23**	-0.31***	-0.03	0.08
	(-2.47)	(-2.81)	(-0.26)	(0.64)
Ln(Population)	0.63**	0.65**	0.92***	0.84**
	(2.18)	(2.26)	(2.65)	(2.42)
Ln(Population)^2	-0.04*	-0.04**	-0.04*	-0.03
	(-1.95)	(-2.10)	(-1.93)	(-1.54)
Ln(Distance to mine)	-0.16	-0.15	-0.38***	-0.35***
	(-1.53)	(-1.42)	(-3.02)	(-2.80)
Ln(High Input Potential)	-2.84***	-2.75***		
	(-5.38)	(-5.16)		
Ln(High Input Potential)^2	0.25***	0.24***		
	(6.38)	(6.12)		
Ln(Low Input Potential)			0.48*	0.54**
			(1.83)	(2.07)
Ln(Low Input Potential)^2			-0.0004	-0.008
			(-0.02)	(-0.29)
Observations	8,783	8,684	8,747	8,649

T-statistics in parentheses
* significant at 10% level
** significant at 5% level
*** significant at 1% level

Table 5.3: Rice

Dependent Variable: Ln(Rice Production)	**Modern Inputs**		**Traditional Inputs**	
	(1)	(2)	(3)	(4)
	OLS	2SLS	OLS	2SLS
Ln(Market Transport Cost)	-0.36***	-0.53***	0.35***	0.52***
	(-3.18)	(-3.96)	(2.98)	(3.79)
Ln(Population)	0.57*	0.60*	0.07	-0.05
	(1.79)	(1.87)	(0.22)	(-0.16)
Ln(Population)^2	-0.04*	-0.04**	0.01	0.02
	(-1.86)	(-2.06)	(0.52)	(1.15)
Ln(Distance to mine)	-0.20	-0.18	-0.58***	-0.51***
	(-1.64)	(-1.43)	(-4.56)	(-4.04)
Ln(High Input Potential)	-3.24***	-3.25***		
	(-10.07)	(-10.05)		
Ln(High Input Potential)^2	0.35***	0.35***		
	(13.40)	(13.37)		
Ln(Low Input Potential)			2.70***	2.79***
			(8.76)	(9.03)
Ln(Low Input Potential)^2			-0.11***	-0.12***
			(-3.74)	(-3.94)
Observations	6,548	6,449	6,778	6,679

T-statistics in parentheses
* significant at 10% level
** significant at 5% level
*** significant at 1% level

Table 5.4: Cassava

Dependent Variable: Ln(Cassava Production)	**Traditional Inputs**	
	(1)	(2)
	OLS	2SLS
Ln(Market Transport Cost)	-0.32**	-0.24
	(-2.02)	(-1.29)
Ln(Population)	1.12**	1.01**
	(2.42)	(2.17)
Ln(Population)^2	-0.06**	-0.05*
	(-2.09)	(-1.75)
Ln(Distance to mine)	-0.75***	-0.69***
	(-4.34)	(-3.94)
Ln(Low Input Potential)	1.63***	1.73***
	(5.15)	(5.46)
Ln(Low Input Potential)^2	-0.02	-0.03
	(-0.96)	(-1.25)
Observations	7,533	7,435

T-statistics in parentheses
* significant at 10% level
** significant at 5% level
*** significant at 1% level

Table 5.5: Maize

	Traditional Inputs	
	(1)	(2)

Dependent Variable: Ln(Maize Production)	OLS	2SLS
Ln(Market Transport Cost)	0.89***	1.21***
	(6.27)	(7.31)
Ln(Population)	1.35***	1.26***
	(3.09)	(2.87)
Ln(Population)^2	-0.07**	-0.06**
	(-2.48)	(-2.07)
Ln(Distance to mine)	-1.86***	-1.87***
	(-11.95)	(-11.86)
Ln(Low Input Potential)	1.91***	2.34***
	(2.87)	(3.42)
Ln(Low Input Potential)^2	0.01	-0.02
	(0.26)	(-0.30)
Observations	9,242	9,144

T-statistics in parentheses
* significant at 10% level
** significant at 5% level
*** significant at 1% level

Table 6.1 Summary Statistics

	Full Sample		**Mechanized Only**		**Traditional Only**	
Variable	**Obs.**	**Mean**	**Obs.**	**Mean**	**Obs.**	**Mean**
Crop Revenue (USD)	3,022	115.43	764	218.12	2,255	80.66
Cost to Market (USD)	2,607	5.10	624	5.49	1,983	4.97
Non-Agricultural Income	2,986	461.38	474	225.74	1,134	309.77
Neighborhood Machinery use	3,068	0.25	770	0.70	2,298	0.10
Dummy=1 if mechanized	3,076	0.25	773	1	2,303	0

Land of household (km^2)	3,087	10.68	773	16.33	2,303	8.81
Age of household head	4,978	49.56	773	49.85	2,295	50.89
Dummy=1 if literate	4,982	0.63	773	0.53	2,300	0.53
Dummy=1 if irriagates land	3,085	0.04	773	0.06	2,302	0.03
Fertilizer purchased (kg)	3,034	1.28	764	1.45	2,269	1.22
Household Agricultural Labor	4,991	1.31	773	2.22	2,301	1.99
Distance to mine (km)	2,607	139.95	624	155.33	2,255	80.66

Data: Nigeria LSMS-ISA

Table 6.2 IV-PROBIT Estimates

Dependent Variable Dummy=1 if uses machinery	**(1)**	**(2)**
ln(Cost to Market)	-0.18*** (-2.33)	-0.25*** (-3.28)
ln(Non-Agricultural Income)	0.04** (1.98)	
ln(Average Machine Use in Village)		4.81*** (31.86)
Land	0.02*** (5.19)	0.01*** (2.43)
Age of Household Head	-0.01 (-0.89)	-0.02 (-1.54)
Age Squared	0.00 (1.14)	0.00 (1.69)
Dummy=1 if Head is literate	0.01 (0.12)	0.08 (0.97)
Fertilizer purchased (kg)	-0.000 (-0.17)	0.001 (0.44)
Distance to Mine	0.12** (2.12)	0.22*** (3.51)
First Stage Estimates ln(Natural Path)	0.64*** (48.63)	0.63*** (51.95)
Observations	1,354	1,354

Note: Robust t-statistics in parenthesis,
* significant at 10% level,
** significant at 5% level,
*** significant at 1% level

Table 6.3 IV Estimates with Heckman Selection Correction

Dependent variable: ln(Crop Revenue)	**(1)** **IMR_1** **(non-agri. income)**	**(2)** **IMR_2** **(neighborhood effects)**
ln(Cost to Market)	-2.20*** (-3.43)	-2.46*** (-6.54)
Land	0.053* (1.73)	0.048*** (4.47)
Age of head	0.181** (2.14)	0.094 (1.51)
Age Squared	-0.001** (-1.96)	-0.001 (-1.28)
Dummy=1 if head is literate	1.595*** (2.92)	0.614 (1.44)
Fertilizer purchased	-0.016 (-1.17)	-0.012 (-0.97)
ln(Distance to Mine)	-0.360 (-0.62)	-0.620 (-1.61)
Dummy=1 if plot is Irrigated	-0.696 (-0.72)	0.049 (0.06)
Household agri. labor	0.386 (1.37)	0.471** (2.14)
Labor squared	-0.049** (-2.00)	-0.048*** (-2.59)
Dummy=1 if AEZ is warm/subhumid	-0.817 (-1.25)	-0.828 (-1.49)
Dummy=1 if AEZ is warm/humid	1.687 (0.47)	-0.629 (-0.17)
Dummy=1 if AEZ is cool/subhumid	2.283** (1.98)	1.670 (1.58)
IMR_1	-0.413 (-0.12)	
IMR_2		-0.193 (-0.56)
Constant	3.806 (0.67)	7.701*** (2.66)
First Stage Results		
ln(Natural Path)	0.699*** (42.01)	0.687*** (67.08)
Angrist-Pischke F test	1,764.72 P=0.0000	4,499.78 P=0.0000
Observations	393	621

Note: Robust t-statistics in parenthesis,
* significant at 10% level,
** significant at 5% level,
*** significant at 1% level

Table 6.4 Robustness Check: Alternative Specification

Dependent variable: ln(crop revenue)	**(1)** **IMR_1** **(non-agri. income)**	**(2)** **IMR_2** **(neighborhood effects)**
ln(Cost to Market)	-2.172*** (-3.56)	-2.247*** (-6.37)
Land	0.054* (1.74)	0.050*** (4.67)
Age of head	0.166* (1.95)	0.088 (1.40)
Age Squared	-0.001* (-1.84)	-0.001 (-1.22)
Dummy=1 if head is literate	1.587*** (2.89)	0.666 (1.58)
Fertilizer purchased	-0.010 (-0.91)	-0.008 (-0.71)
ln(Distance to Mine)	-0.534 (-0.92)	-0.613 (-1.62)
IMR_1	-0.367 (-0.11)	
IMR_2		-0.177 (-0.56)
Constant	4.891 (0.86)	7.461*** (2.78)
First Stage Results ln(Natural Path) Angrist-Pischke F test	 0.712*** (42.27) 1,786.40 P=0.0000	 0.691*** (69.51) 4,831.50 P=0.0000
Observations	393	621

Note: Robust t-statistics in parenthesis,
* significant at 10% level,
** significant at 5% level,
*** significant at 1% level

Table 6.5 IV Estimates, Heterogeneous Effects

Dependent variable: ln(crop revenue)	(1) Full	(2) Modern	(3) Traditional
ln(Cost to Market)	-1.774***	-2.441***	-1.422***
	(-8.44)	(-6.57)	(-5.47)
Land	0.031***	0.049***	0.018**
	(4.38)	(4.66)	(2.37)
Age of head	0.033	0.094	0.015
	(0.98)	(1.51)	(0.39)
Age Squared	-0.000	-0.001	-0.000
	(-0.76)	(-1.29)	(-0.29)
Dummy=1 if head is literate	0.618***	0.634	0.533***
	(3.34)	(1.49)	(2.62)
Fertilizer purchased	-0.005	-0.012	-0.000
	(-0.62)	(-0.98)	(-0.02)
ln(Distance to Mine)	-0.351**	-0.609	-0.280
	(-2.09)	(-1.60)	(-1.45)
Dummy=1 if plot is Irrigated	0.781	0.046	1.189*
	(1.55)	(0.05)	(1.90)
Household agri. labor	0.322***	0.472**	0.337***
	(3.22)	(2.17)	(2.77)
Labor squared	-0.041***	-0.048***	-0.047***
	(-4.07)	(-2.59)	(-3.29)
Dummy=1 if AEZ is warm/subhumid	0.664***	-0.737	0.960***
	(2.81)	(-1.45)	(3.39)
Dummy=1 if AEZ is warm/humid	-0.201	-0.677	0.279
	(-0.51)	(-0.18)	(0.67)
Dummy=1 if AEZ is cool/subhumid	3.362***	1.830*	1.504
	(3.89)	(1.83)	(0.49)
Constant	5.491***	7.393***	4.843***
	(4.15)	(2.65)	(3.23)
First Stage Results			
ln(Natural Path)	0.632***	0.685***	0.607
	(53.06)	(67.94)	(37.68)
Angrist-Pischke F test	2814.98	4616.49	1420.08
	P=0.0000	P=0.0000	P=0.0000
Observations	2,598	624	1,974

Note: Robust t-statistics in parenthesis,
* significant at 10% level,
** significant at 5% level,
*** significant at 1% level

Appendix I: 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 used in this paper was 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 ton 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

Appendix II: Natural Pathway

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[21] (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).

[21] "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