Working paper

Rural Roads and Intermediated Trade

Regression Discontinuity Evidence from Sierra Leone

Lorenzo Casaburi
Rachel Glennerster
Tavneet Suri

December 2012

IGC

International
Growth Centre

DIRECTED BY
LSE · UNIVERSITY OF OXFORD

FUNDED BY
UKaid
from the British people

Rural Roads and Intermediated Trade:
Regression Discontinuity Evidence from Sierra Leone[*]

Lorenzo Casaburi, Rachel Glennerster and Tavneet Suri[†]

December 31, 2012

Abstract

There is little evidence on the impacts of rural road infrastructure investments in developing economies. This paper contributes to this literature in two ways. First, using a road-level regression discontinuity identification strategy based on the specific design of a rural road rehabilitation program in Sierra Leone, we study the impacts of rural road infrastructure improvements on prices. We show that the improved roads led to a reduction in local crop prices in markets. We find that this effect is stronger in markets that are more distant from major urban centers and in those located in less productive areas. Second, we theoretically derive equilibrium prices in agricultural markets under alternative structures of trader competition. We show that our empirical findings are consistent with a search cost framework à la Mortensen, but inconsistent with other models, such as Bertrand competition, bilateral bargaining, and Cournot oligopsony.

1 Introduction

In spite of the central relevance of infrastructure in the development policy debate and in foreign aid, there has been little rigorous research on the role of local infrastructure investments in developing countries, especially in the context of very poor economies. Over the last two decades, governments and donors in Sub-Saharan Africa have devoted considerable resources to rural road construction and rehabilitation, of which a large share was used to upgrade feeder roads that link up small localities with each other or to larger roads.[1] The rationale behind these investments is that good feeder roads, while expensive, are central to the integration of markets, primarily because, by reducing transport costs, they improve the access of farmers to markets for their crops. The existing body of research on the impact of rural feeder roads is plagued by identification problems, and usually does not provide compelling causal evidence on the impact of infrastructure improvement (van de Walle (2009)).

In addition, the empirical literature on rural roads, for the most part, has not explicitly dealt with the role of the underlying market structure in shaping the response to these infrastructure investments. Agricultural markets throughout Sub-Saharan Africa are characterized by high levels of fragmentation and poor transport infrastructure, with intermediaries playing a central role (Fafchamps et al. (2004)). Different models of trader competition and intermediation generate different predictions about the price response to an improvement in rural road quality and how this response varies with market characteristics. The empirical analysis of these price responses can therefore provide a test for competing theoretical frameworks of the market structure in agricultural trade. Understanding which of these theoretical frameworks best explains the nature of competition amongst these agricultural intermediaries can shed light on the impact of other supply chain interventions, such as subsidies to different agents of the value chain or exporting promotion policies.

In this paper, we make two contributions. First, we provide empirical evidence on the impact of rural road rehabilitation on transport costs and rural market prices. We focus on a specific program in Sierra Leone where the road selection algorithm allows us to use a Regression Discontinuity Design (RDD), and we use novel data collection strategies to measure changes in prices and transport costs. Sierra Leone is an economy with low population density where feeder roads could potentially have large impacts. Rural markets in the economy are not well integrated, and farmers and traders both travel long distances to engage in trade, Transaction costs are extremely high but transport costs are only a part of these transaction costs. These characteristics are common to many rural African economies (Fafchamps and Hill (2008), Fafchamps et al (2004)) and other developing countries (World Bank, 2009). As feeder roads are commonly

---

[1]Since 2000, major feeder roads rehabilitation programs have been implemented in Cameroon, the DRC, Ghana, Mozambique, Sierra Leone, and South Sudan, among other countries. Carruthers et al. (2010) document that it would cost about 2% of African GDP every year for ten years to reach some reasonable targets on improved transportation.

advocated as a key policy intervention in these contexts, our first objective is to understand their impact. Well identified measures of feeder road impacts are particularly important given the high costs of effectively connecting such dispersed populations.

As a second contribution, we use our empirical results to test between alternative models of competition in agricultural markets. We focus on Bertrand oligopsony, bilateral bargaining, Cournot oligopsony (with and without endogenous entry), and a basic Mortensen (2003) search model. For each of these frameworks, we derive the equilibrium price in rural markets and provide comparative statics on the price effects of a change in rural road transport costs. We then show how this price effect varies by market characteristics, such as the distance to the main urban centers and the agricultural productivity in the surrounding areas. To the extent that the four classes of models deliver different predictions on these comparative statics, we can use our empirical results to test between these models. It is an open question as to what the relevant model of market structure for traders in agriculture is in developing economies. As we describe below, understanding market structure is important to understanding what policies are most effective at improving market access and, more broadly, welfare for farmers in Sub Saharan Africa.

Our empirical analysis focuses on a feeder road rehabilitation program in Sierra Leone that was financed by the European Union (EU) and implemented in four districts that cover 27% of the country's area and 30% of its population.[2] Sierra Leone's infrastructure is extremely poor: in 2002, at the end of a decade-long civil war, only 8% of the country's 11,300 km of roads were paved. Agriculture is the country's largest employer, with 64% of households farming, of which 87% produce rice, the main staple. The infrastructure network is generally described as inadequate to support well-integrated agricultural markets.[3]

The project targeted local dirt feeder roads, measuring 20 km on average, linking local markets to villages or to a more important road. They did not connect major cities in different parts of the country together. To structure the use of its funds, the EU created a priority ranking for each of 47 eligible roads based on an index of quantitative economic data. The highest-ranked roads were chosen in order until at least as close to a total of 150 km of roads had been assigned to each district. A total of 31 roads were ultimately selected. This method of fund allocation allows us to use a road-level RDD to study the impacts of the feeder road improvements.

Our empirical results can be summarized as follows. First, the rehabilitation program did improve road quality on the roads selected. Using data on transport fares and GIS video stream data collected on the roads, we show that transport costs fell significantly. Second, from trader surveys and monthly price surveys, we find that an improvement in rural road quality leads to a reduction in the prices of rice and cassava (the two main staples produced domestically) in

---

[2]Sierra Leone National Census (2004).

[3]See the Agricultural Household Tracking Survey (AHTS) Report (2011) and Pushak and Foster (2011).

rural markets along the rehabilitated roads.[4] Third, we find that this price reduction is stronger in markets that are farther away from main urban centers and is weaker in markets that are located in more productive areas.

Which of the standard classes of models best explains these findings? We look across a set of simple models which assume the city is a small open economy, traders are homogeneous and their transport costs are linear. In these models, rural roads can potentially affect transport costs for both buyers traveling from urban areas (primarily traders) and sellers (farmers) since both travel to rural marketplaces. In our analysis, we refer to the impact of the rural road rehabilitation on urban traders as the "demand effect", and the effect on producers as the "supply effect", with the former driving up prices in rural markets, and the latter driving them down. The relative importance of these two effects varies across the theoretical frameworks we present. In particular, both the Bertrand framework and the Cournot model with endogenous entry predict that reduced transport costs should be associated unambiguously with *increases* in prices in rural markets (i.e. the demand effect always dominates). However, this is inconsistent with our empirical findings - we find that improvements in road quality that reduce transport costs on average *reduce* prices. On the other hand, the remaining frameworks (bilateral bargaining, Cournot with an exogenous number of traders, and the search frictions framework) predict a role for both demand and supply effects.

Second, the models have different predictions regarding the heterogeneity in the price effect of reduced rural transports costs along two dimensions: (i) the distance between the rural markets and major urban areas, and (ii) farmer productivity. Both the bilateral bargaining model[5] and the Cournot oligopsony case with an exogenous number of traders predict that market characteristics have no impacts on these price effects, which is inconsistent with our results. On the other hand, a basic Mortensen search framework predicts that the interactions of the price effect with distance and productivity are non-zero because the magnitude of both the demand and the supply effect varies with these two market characteristics. Specifically, for the markets that are further away, or the markets in areas with lower productivity, the relevance of search costs in determining equilibrium prices is higher. In turn, this implies that the negative price effects of the improved roads should be stronger for these sets of markets. Overall, our empirical results are consistent only with a search framework and inconsistent with other models.

These results have important policy-implications. First, the impact of roads projects varies by the characteristics of the location where roads are placed. Features such as productivity

---

[4]Rural markets in our sample are medium-sized marketplaces in which several crops are traded (rice, cassava, palm oil in particular). In Section 3 we describe in detail our sampling strategy for markets and traders within markets.

[5]The bilateral bargaining model represents the case where a given producer is locked into a relationship with one particular trader. It can also be interpreted as a model with search costs approaching infinity, so that producers have no outside option, outside of the existing relationship.

and linkages with urban consumers and traders can affect not just the magnitude of rural market price responses, but also their sign. This in turn affects the optimal road placement a social planner would choose under different objective functions. As in Suri (2011), where the returns to agricultural technologies are heterogeneous and that some of this heterogeneity arises from differences in access to the rural road infrastructure network, returns to different road projects will vary sharply with underlying farmer heterogeneity. Second, the finding that agricultural intermediary markets in this setting are best characterized by a framework with search frictions has implications for other policies designed to influence agents in the agricultural value chain. These policies include price subsidies, agricultural export promotion interventions, credit-provision policies targeting traders, and more major road projects. Third, the empirical evidence in support of the presence for search frictions suggests a possible complementarity between hard infrastructure projects and other interventions aimed at reducing search costs, such as the introduction of marketing boards or the extension of mobile phone coverage in rural areas (for example, mobile phone penetration rates are only 36% in Sierra Leone (World Development Indicators, 2012)).

This paper's findings are consistent with recent empirical evidence on search frictions (see, for example, Jensen (2007), Aker (2010) and Goyal (2010)). While these studies focus on the price impact of a reduction in information frictions, we show that the response to improved transport infrastructure also depends on the presence of search frictions. From this standpoint, our approach is similar to that of Allen (2012) who shows that, in the presence of search costs, the elasticity of trade flows to destination prices depends on supply parameters, such as producer heterogeneity. In addition, our work draws from three other strands of literature. First, a small but influential set of papers has used rigorous identification strategies to shed light on the impact of large transport infrastructure improvements (examples include Michaels (2008), Donaldson (2012), Datta (2012), Faber (2012) and Banerjee, Duflo and Qian (2012)). We look at rural roads and not highways - van de Walle (2009) provides a good review of the literature on rural roads.[6] Second, a recent growing literature uses micro-data to estimate the degree of price pass-through internationally and domestically.[7] Third, our theory focus is motivated by the recent emphasis on intermediated trade (for example, Bardhan, Mookherjee, and Tsumagari (2012), Antràs and Costinot (2011)). In particular, we rely heavily on Chau, Goto and Kanbur (2009) in our theoretical setup and in the way we model search frictions.

The rest of the paper is structured as follows. In Section 2, we outline the theoretical frameworks and derive the comparative statics of interest. Section 3 describes some background on Sierra Leone, the EU road rehabilitation program and the data sources. Section 4 outlines the

---

[6]Examples include Jacoby and Minten (2008), Dorosh et al. (2010), Gibson and Rozelle (2003), Ali (2011), Khandker, Bakht and Koolwal (2009), Khandker and Koolwal (2011), and Mu and van de Walle (2007).

[7]See, for example, Broda and Weinstein (2008), Burstein and Jaimovich (2009), Gopinath et al. (2011), Borraz et al (2012) and Atkin and Donaldson (2012).

empirical strategy and tests the validity of the RDD for our setting. Section 5 presents the results of the empirical analysis, including a comparison of our results with the theoretical predictions of the different models of competition and results from a variety of robustness checks. Section 6 concludes.

2 Theoretical Frameworks

In this section, we present four different theories of trade intermediation between agricultural producers and traders. Traders play a central role in the rural economies of Sub-Saharan African by channeling produce between crop-producing areas in the countryside and the consumers living in urban areas (Fafchamps et al. (2004)). We focus on rural markets where traders purchase and sell a number of different crops. Traders are mostly small scale, traveling from market to market and then transporting crops to the capital city. We describe these traders in more detail below.

We focus on four broad classes of theoretical models, representing different market structures: Bertrand competition, bilateral Nash bargaining, Cournot oligopsony (both with and without endogenous entry) and search frictions à la Mortensen (2003). The Bertrand case and the search frictions case, as well as the theoretical setup, are based on Chau, Goto and Kanbur (2009). We adapt their framework in three directions to fit our research questions. First, we explicitly model the role of rural roads in linking producers and traders to rural market places. Second, we introduce a productivity parameter that varies across villages. Third, we model two additional cases: a specific case of bilateral bargaining between traders and producers and the case of Cournot competition (with and without endogenous entry).

The primary goal is to derive key comparative statics to be tested in our empirical framework. We focus specifically on the price effects of an improvement in rural road infrastructure, and on the heterogeneity in this effect across market characteristics.

2.1 Setup

We model the transactions that occur between traders and rural agricultural producers in rural markets to which both types of agents travel. Throughout, we use the term producers to refer to either farmers or "aggregators", where aggregators refer to larger farmers who aggregate produce from other farmers in a village to bring it to the market to sell. The theoretical models we present below all follow through whether it is farmers themselves that come to the market or aggregators. We make two important simplifying assumptions here. The first is that traders are homogeneous - this seems to be a reasonable interpretation of our data. For example, 97% of traders are male, 88% are from two main ethnic groups, 65% have no education, 86% are

married, 96% report that they started trading on their own, 62% own a mobile phone, 96% own a radio and only 19 traders of local rice own their own mode of transport (motorbike, car or truck - unfortunately the survey did not distinguish between these). The second assumption we make is that is that utility is linear: both farmers and traders maximize their (expected) profits.

Around each market, there is an exogenous number of producers, N, who produce σ units of a certain crop. The opportunity cost of each unit (for instance the utility from consuming each unit) is c. In order to reach the local market, rural producers use rural roads to travel distance α. The unit transport cost on the rural road, τ, is therefore the inverse of a measure of road quality. When road quality improves, the unit transport cost on the road decreases.

City traders travel to the local market to purchase the crop which they can then sell in the urban areas at an exogenous price p^*. The exogeneity of the city price p^* is based on the intuition that the city receives the crop from many different markets across the country. Each market is thus a small open economy and changes to the transport costs between a given village and the urban areas are assumed not to affect the city price or prices in other rural markets. In order to reach the market, the traders bear two types of transport costs: first, a "major road" transport cost, x, and second, a rural road cost $\beta\tau$. The "net city price" is thus $p^* - x - \beta\tau$. Throughout, we assume this form of linear transport costs as well as a separation between the major road and the rural road transport costs for the trader. The linearity seems a natural assumption for the transport costs since the main form of transport for traders is renting motorbikes, which price a fare per distance travelled. The travel along these roads (major and rural) is reasonably competitive. Very few traders actually own a mode of transport - in our trader data only 19 traders owned any form of transport. There is therefore little scope for a (discrete) investment that simultaneously affects transport costs on both the main road and the rural road. In addition, the improvement in the rural roads due to the EU program causes a reduction in transport costs but not a big enough reduction to allow traders to invest in their own mode of transport. Finally, this assumption is consistent with Fafchamps and Hill (2008) who find that there are no economies of scale at this level of traders.

We denote the (endogenous) sale price in the rural market by p. In addition to city traders, we assume that farmers can sell to local consumers (non-farming rural households) at cost-recovery price $p^0 = c + \alpha\tau$. Throughout the paper, we assume there are gains from trade, that is: $p^* - x - \beta\tau > c + \alpha\tau$. The assumption that there are gains to trade implies that there is no extensive margin effect of changes in τ on traders or producers entering the market.

Below, we discuss equilibrium prices (or price distributions) under four alternative market structures. For each of these, we are interested in three comparative statics:

1. The equilibrium price change in response to a change in rural road transport costs: $\frac{\partial p}{\partial \tau}$

2. Heterogeneity in this price response by the distance travelled on the major road, x: $\frac{\partial^2 p}{\partial \tau \partial x}$

3. Heterogeneity in this price response by market-level productivity, σ: $\frac{\partial^2 p}{\partial \tau \partial \sigma}$

2.2 Bertrand Oligopsony

Bertrand competition is characterized by free entry of city traders into each village and perfect information. This implies that all the producers are matched to city traders and that competition drives up the equilibrium price, p^B to equalize the net city price:

$$p^B = p^* - x - \beta\tau \tag{1}$$

It is easy to see that producers appropriate all the gains from trade. Looking at the predictions of the Bertrand model with regards to the three comparative statics of interest, we obtain:

1. $\frac{\partial p^B}{\partial \tau} = -\beta < 0$: rural road transport costs enter the equilibrium price only via their effect on traders' transport costs.

2. $\frac{\partial^2 p^B}{\partial \tau \partial x} = 0$: as the trader cost function is separable between the rural road and the major road costs, the price response to rural road costs is not affected by the major road transport costs.

3. $\frac{\partial^2 p^B}{\partial \tau \partial \sigma} = 0$: productivity, or any other supply characteristic, does not affect the equilibrium price, and hence does not affect the price response to a reduction in rural transport costs.

2.3 Bilateral Bargaining with Lock In

We now present a simple model of prices under a specific model of bilateral bargaining in which each producer can only trade with one specific trader, with an outside option limited to the cost-recovery price $c + \alpha\tau$. In this scenario, the transactions equilibrium price, p^N, is assumed to be the generalized Nash-Bargaining solution:

$$\begin{aligned} p^N &= \arg\max_p \{[\sigma(p^* - x - \beta\tau - p)]^\gamma * [\sigma(p - c - \alpha\tau)]^{1-\gamma}\} \\ &= \gamma(c + \alpha\tau) + (1 - \gamma)(p^* - x - \beta\tau), \end{aligned} \tag{2}$$

where $\sigma(p^* - x - \beta\tau - p)$ is the trader's surplus, $\sigma(p - c - \alpha\tau)$ is the producer's surplus, and γ is the bargaining weight for the trader. The Bertrand outcome is a special case of the Nash Bargaining solution for $\gamma = 0$. In addition, the bilateral bargaining setup implies that, in the presence of a positive surplus from trade, all the producers are matched to traders.

The model delivers the following predictions with regards to our three key comparative statics:

1. $\frac{\partial p^N}{\partial \tau} = \gamma\alpha - (1-\gamma)\beta \gtrless 0$: with positive α and β, rural road transport costs now enter prices through their effect both on producers' and on traders' costs. The former drives down prices (what we term the supply effect), while the latter raises prices (what we term the demand effect).

2. $\frac{\partial^2 p^N}{\partial \tau \partial x} = 0$, as in the Bertrand case.

3. $\frac{\partial^2 p^N}{\partial \tau \partial \sigma} = 0$: the amount transacted, σ, does not affect how the price responds to a change in τ, even though σ does enter the equilibrium price.

2.4 Cournot Oligopsony

The third framework we present is one of Cournot competition. Here, we introduce an additional assumption to generate an upward sloping supply curve. We assume that the parameter α (the distance travelled by the producer on the rural feeder road) is a random variable uniformly distributed over $[\alpha^* - z, \alpha^* + z]$. Producers sell in the market only if the equilibrium price p is larger than their reservation price: $p > c + \alpha\tau$. This extensive margin on selling generates an upward sloping supply curve as only a share $F(\frac{p-c}{\tau}) = \frac{(p-c)/\tau - (\alpha^* - z)}{2z}$ of producers enters the market, where $F(\cdot)$ is the cdf of the uniform distribution.[8]

We study two versions of the Cournot model. First, we look at the case with an exogenous number of traders operating in the market, M. Second, we endogenize the number of entrants using a free entry condition, where we assume entrants have a fixed cost of entry K.

We focus on symmetric equilibria. With an exogenous number of traders, each trader chooses his optimal quantity given other traders' quantities which enter the profit functions through the (inverse) supply curve. After imposing symmetry, the equilibrium price, p^C is

$$p^C = \frac{c + \tau(\alpha^* - z) + M(p^* - x - \beta\tau)}{1 + M} \tag{3}$$

The Cournot model with exogenous entry delivers the following results with regards to our comparative statics of interest:

1. $\frac{\partial p^C}{\partial \tau} = \frac{\alpha^* - z - M\beta}{1 + M} \gtrless 0$: both the supply and the demand effect of a change in τ enter the derivative of price with respect to transport costs.

2. $\frac{\partial^2 p^C}{\partial \tau \partial x} = 0$: while x shifts the net city price, it does not interact with τ in determining the equilibrium price.

3. $\frac{\partial^2 p^C}{\partial \tau \partial \sigma} = 0$: σ does not enter the solution for the equilibrium price, p^C.

---

[8]We also studied the case where farmers are heterogeneous in the opportunity cost, c, instead of being heterogeneous in the parameter α. Since all the comparative statics for this case collapse to the Bertrand case, we do not report it here.

We now extend the basic Cournot model above to allow for endogenous trader entry. By equating individual profits with the cost of entry, we find the equilibrium number of traders operating in the economy and then derive the corresponding equilibrium price, p^{CE}:

$$p^{CE} = p^* - x - \beta\tau - \sqrt{\frac{2\tau K}{Nz\sigma}}, \tag{4}$$

The comparative statics of interest are now:

1. $\frac{\partial p^{CE}}{\partial \tau} = -\beta - \sqrt{\frac{K}{2\tau Nz\sigma}} < 0$: τ affects p^{CE} through its impact on both demand and supply. The demand effect operates though β, as before. However, the reduction in aggregate supply generates a more than proportional reduction in the number of traders entering the economy. As a result, the price reduction in response to an increase in τ is stronger than in the benchmark Bertrand case. Intuitively, an increase in τ has two effects. First, it reduces the "net city price", $p^* - x - \beta\tau$. All else equal, this leads traders to reduce p. Second, it reduces the elasticity of supply, $Q = \frac{\sigma N(p-c)}{\tau}$. All else equal, this induces traders to *increase* the mark down, thereby contributing further to a decrease in price.

2. $\frac{\partial^2 p^{CE}}{\partial \tau \partial x} = 0$: terms that only enter the net city price do not interact with transport costs in determining the equilibrium price.

3. $\frac{\partial^2 p^{CE}}{\partial \tau \partial \sigma} = \frac{\sqrt{K}}{2\sigma\sqrt{2rNz\sigma}} > 0$: by increasing aggregate supply, σ reduces the role of fixed costs in determining the equilibrium price. An economy with a higher productivity will be closer to the perfectly competitive case relative to one with a lower σ. Thus, an increase in σ implies a smaller in absolute value (i.e. closer to $-\beta$) price response to an increase in τ since the impact of τ on the elasticity of supply is lower.

Summarizing these results, the Cournot model with exogenous entry cannot be distinguished from the bilateral bargaining model based solely on the sign of the comparative statics. On the other hand, the version with endogenous entry delivers a unique set of predictions relative to the previous models.

2.5 Search Frictions

We now introduce search frictions in the model economy. Empirically, we think search costs may arise from two potential mechanisms. A first potential mechanism for search costs is that these markets are difficult to reach. Farmers and traders often travel on motorbikes to these markets and although transport is available, the timing is uncertain. This generates some search frictions in that farmers and traders may not end up in a given rural market at the same time. There may therefore be waiting costs as well as uncertainty of what will be available in the

market by the time the trader gets there. This is the type of search cost considered in Fafchamps and Hill (2008).

Second, farmers develop relationships with traders over time. There is a lot of evidence for such contracts both in Sierra Leone (see Casaburi and Reed (2012) for an example) as well as in other developing countries (see Deb and Suri (2012)). These relationships operate as a search friction in the sense that there are switching costs for farmers to shift across traders and this can generate market power for the traders. This is one explanation given by Fafchamps and Minten (2011) to explain why a price information intervention did not have impacts on prices in India. These relational contracts may exist to enforce quality standards (see Bardhan, Mookherjee and Tsumagari (2012)) or due to the existence of trade credit as in the case of Casaburi and Reed (2012). The existence of such relationships between farmers and traders would generate search frictions of the sort we model here.

To construct our model, we closely follow Chau, Goto, and Kanbur (2009) and we refer readers to their paper for detailed derivations of the results. Their paper is in turn based on the static search framework derived by Mortensen (2003), which provides the key results. These results can also be derived in more elaborate dynamic search models (for example, Burdett and Mortensen (1998)). We choose a static framework so that we can easily introduce search frictions in the previously described setup.

The interactions between traders and producers now occur in three stages. In the first stage, traders decide whether to enter a certain market. If they enter, they incur an entry cost, K, which can be thought of as including bargaining time, travel fixed costs, waiting time and uncertainty. In the second stage, traders who entered the market choose a price offer. Due to the search frictions, only one random producer receives the offer. In the third stage, producers take up the best offer among those received. For a large enough market, the distribution of offers received by each producer follows a Poisson with mean $\lambda \equiv M/N$, where M is the endogenous number of traders entering the market:

$$Pr(z; \lambda) = e^{-\lambda} \frac{\lambda^z}{z!} \tag{5}$$

Thus, the cumulative distribution of the highest price offer received by each producer, or equivalently, from the trader's perspective, the probability of outbidding the second-best offer is:

$$H(p) = \sum_{z=0}^{\infty} Pr(z; \lambda) F(p)^z = \exp(1 - F(p)) \tag{6}$$

where $F(p)$ is the endogenous cdf of price offers made by the traders.

The trader's expected profit maximization implies that no symmetric pure strategy equilibrium exists. Rather, traders follow a mixed strategy where prices are in the support $[p_l, p_h]$, given

by:
$$[c + \alpha\tau, (1 - e^{-\lambda})(p^* - x - \beta\tau) + e^{-\lambda}(c + \alpha\tau)] \tag{7}$$

and are drawn according to distribution:
$$F(p) = \frac{1}{\lambda} \ln\left(\frac{p^* - x - \beta\tau - c - \alpha\tau}{p^* - x - \beta\tau - p}\right) \tag{8}$$

Each price in the support leads to an equal expected profit. The free entry condition can therefore be written as $E[\pi(p_l)] = K$, which allows us to solve for the equilibrium ratio of traders to producers, λ:
$$\lambda^* = \ln\left(\frac{p^* - x - \beta\tau - c - \alpha\tau}{K/\sigma}\right) \tag{9}$$

We substitute λ^* into the optimal trader's price offer distribution, $F(p)$, in (8) and solve for both the expected price for farmers conditional on receiving at least one offer from traders and for the unconditional expected price, p^S. The latter can be written as:
$$p^S = p^* - x - \beta\tau - \frac{K}{\sigma}\left[1 + \ln(p^* - x - \beta\tau - c - \alpha\tau) - \ln\left(\frac{K}{\sigma}\right)\right] \tag{10}$$

We focus on the comparative statics from this expected unconditional price equilibrium. The model delivers the following predictions with respect to our three comparative statics:

1. $\frac{\partial p^S}{\partial \tau} = -\beta\left(1 - \frac{(K/\sigma)}{p^* - x - \beta r - c - \alpha\tau}\right) + \frac{(K/\sigma)\alpha}{p^* - x - \beta\tau - c - \alpha\tau} \gtrless 0$, where the first term represents the demand effect, and the second the supply effect, of a change in transport costs. The magnitude of the demand effect is lower in absolute value than the demand effect in the Bertrand case: traders' market power induced by search frictions implies an imperfect pass-through of traders' cost shocks.

2. $\frac{\partial^2 p^S}{\partial \tau \partial x} = \frac{K(\alpha+\beta)}{\sigma(p^* - x - \beta\tau - c - \alpha\tau)^2} > 0$: in locations that are farther away for the city, a lower net price induces lower entry, more monopsony power and stronger deviations from the Bertrand benchmark. Thus, if x is higher, the (negative) price effect of a higher transport cost for traders is weaker and the (positive) price effect of a higher transport cost for producers is stronger.

3. $\frac{\partial^2 p^S}{\partial \tau \partial \sigma} = -\frac{K(\alpha+\beta)}{\sigma^2(p^* - x - \beta\tau - c - \alpha\tau)} < 0$: intuitively, an increase in σ lowers the "real" entry cost K/σ, bringing the economy closer to the benchmark competitive Bertrand case. A higher σ therefore moves the equilibrium both toward a stronger (negative) demand effect and toward a weaker (positive) supply effect of an increase in τ.

2.6 Summary of Theoretical Results

Table 1 summarizes the results from the models presented in this section. With regard to our comparative statics of interest, the models can be differentiated easily in two ways. First, some models predict that a decrease in τ will unambiguously lead to higher equilibrium prices, while others predict an ambiguous effect. This is due to the fact that a change in τ can induce either a demand effect alone, or both a demand and a supply effect. Second, the Cournot model with exogenous entry and the search friction frameworks are the only models predicting that the magnitude of the price response will depend on the market-specific features, x and σ. In section 5, we compare these predictions to our empirical results.

3 Background and Data Sources

3.1 The EU Rural Feeder Roads Rehabilitation program

3.1.1 Background and Implementation

The EU feeder roads rehabilitation program was designed to contribute to the reconstruction of Sierra Leone's infrastructure in the aftermath of a destructive civil war (1991-2002). In 2008 the government of Sierra Leone acknowledged that the development of a road network was a strategic priority and set a goal of rehabilitating 2,055km of feeder roads within five years (Republic of Sierra Leone (2008)). Beyond Sierra Leone, distance and access to markets have been shown to be severe constraints for agricultural growth and investment by farmers in Sub-Saharan Africa (see Suri (2011) and World Bank (2008) for a review).[9]

The EU's program was implemented over 2009 to 2011 at a total cost of EUR 9.5 million (USD 13 million) - approximately EUR 16,000 per kilometer of rehabilitated road. The program targeted four districts in three different provinces: Kambia and Port Loko (Northern Province), Kenema (Eastern Province) and Pujehun (Southern Province). The roads in the rehabilitation program primarily connected small towns and markets to one another or to a major road, rather than linking big cities and regions to one other. Because these feeder roads do not dramatically alter access to population centers, they are unlikely to modify migration patterns in the areas surrounding the rehabilitated roads. The roads are also unlikely to affect agricultural technology adoption because improved technologies such as fertilizer and improved seeds are rarely available even in urban centers, as reflected in extremely low baseline rates of technology adoption.[10]

---

[9]The cost of agricultural inputs is extremely high in these economies, partly because of high transport costs due to poor infrastructure. In 2005, for example, transport costs represented approximately 25% of the farm-gate price of urea fertilizer in Sub-Saharan African countries. The ratio of wholesale fertilizer prices plus transport, insurance, and freight costs to wholesale prices ranges from 1.55 in Malawi to 1.75 in Nigeria, versus 1.2 in the United States (Gregory and Bumb (2006)).

[10]Roughly 5% of farmers use fertilizer nationally (AHTS (2011)) and much of this use is centered around the capital

Rehabilitation work began in mid-2009 and ended by early 2011 for all the roads. In the final construction progress report, dated August 2011, 25 of the 31 roads selected for rehabilitation were described as fully rehabilitated and 6 roads as partially rehabilitated.[11]

3.1.2 Program Design

The rehabilitation program was designed in a way that allows for a road-level Regression Discontinuity Design (RDD) analysis. In 2003, field investigations with local stakeholders led to the identification of a base list of 47 rural roads eligible for rehabilitation, totaling about 800 km. The task of prioritizing 600 kilometers of roads for rehabilitation from amongst this larger sample was given to an external consultant, Edward Davies & Associates (EDA). EDA (2004) provides extensive details of the prioritization process that was used to decide which of these roads would be rehabilitated. The roads were ranked according to a score, which was a weighted average of five components:

i) **Economic Production per kilometer**, defined on the basis of survey measures of the volume of crops produced, income from economic activities and mode of transportation;

ii) **Population per kilometer** within the area of influence of each road;

iii) **Road Assessment**, a 1 to 5 score measuring the pre-existing condition of the selected roads, based on seven parameters (culverts, bridges, drainage, pavement surface, vertical alignment, horizontal alignment, and riding quality);

iv) **Social Value**, a 1 to 5 score, based on the number of schools, health centers, wells and toilets in the catchment area of the road;

v) **Length**.

The data on these components was compiled by EDA. Each component of the score was normalized by its district-level maximum value. The final weighted sum of each of the five components had weights of 0.4, 0.2, 0.2, 0.1 and 0.1 for components (i) to (v) above, respectively.[12] The decision rule was that, in each district, roads would be rehabilitated starting with the highest-priority one (based on the score), following in order of decreasing score until as close to 150 km of road as possible was allocated to be rehabilitated in each district. Since th roads could not be split to get exactly 150 km, this means that in some districts slightly over 150 km

Freetown and the national agricultural research station. In separate work, Suri (2011) found that deficient infrastructure and differential access to good infrastructure were among the reasons explaining low technology adoptions rates in Kenya. In our context, the small size of the roads considered makes it reasonable to assume that the supply of agricultural technology from the main urban centers to the countryside did not dramatically increase as a result of the program.

[11]The average completion rate for the six roads partially rehabilitated was 64%. This percentage does not reflect the fraction of the total length of the road that had been rehabilitated. Instead, it illustrates what fraction of the planned improvements were completed on the road i.e. the average completion across all specifications of the work - in many cases, improvements were implemented along the entire length of the road, but not all the planned improvements were fully completed.

[12]Appendix 1 provides the exact formula used to compute the score.

of road was rehabilitated, and in some districts just under 150 km was rehabilitated. Following this rule, 30 roads out of the eligible 47 should have been selected across the four districts. Figure 1 presents maps of the priority roads in the initial list across the four districts, as well as the connections between these roads and the major roads in the country.

After collecting data from the EU, EDA, and the Sierra Leone Roads Authority (SLRA) to verify that the RDD had been followed, we found evidence of two potential manipulations that occurred around the cutoff. In the district of Kenema, the final list of roads rehabilitated covered 184 km, while selecting one less road would have brought the total rehabilitated in this district to 153 km, closer to the approximate desired cutoff of 150 km. One extra road was therefore rehabilitated that was not supposed to be. We treat this as a fuzzy RDD where the extra treated road is allocated to the control group. In Kambia District, two roads exactly at the cutoff had the same score. Only one of the two roads was rehabilitated, giving the appropriate total for rehabilitation, but no rules were established for how roads should be selected in the case of a tie. In this case, we drop the two roads from our baseline specifications, though we check the robustness of our results to their inclusion. We discuss the implications of these issues for our identification strategy in more detail in Section 4 below.

3.2 Data

This section provides an overview of the data used in our empirical framework, ordered chronologically by collection date.

3.2.1 Baseline Roads Data (2003-2004)

To construct the scoring index, EDA collected data around all 47 eligible roads before any rehabilitation began. We use this data to check for baseline differences between the roads above and below the cutoff in our RDD framework. We include all five components of the score, as well as a number of road characteristics, such as the number of bridges, palmlog bridges[13] and culverts,[14] data for which was provided by the EU. Table 2 shows summary statistics on the baseline condition of the roads considered for rehabilitation. The average length of the roads was 21 km. In this table we also report the average rescaled score, which is the value of the index created by the EU, rescaled so that zero is the selection cutoff score between the first road below and the first road above in each district.

---

[13]A palmlog bridge is a makeshift bridge made from logs of palm trees. Regular cars are usually unable to cross such bridges, but four-wheel drive vehicles and motorbikes often can, although at times with considerable risk. During the rainy season, palmlog bridges get flooded, making the wood rot and sometimes causing the entire bridge to collapse.

[14]A culvert is a pipe or a drain that lets water flow below a road. Culverts are much smaller than bridges and are often used to create crossings over small waterways or as a way to improve drainage in locations where rainy seasons are extremely intense (as they are in Sierra Leone).

3.2.2 National Population Census (2004)

To test our RDD assumptions, we use data from the last population census conducted in Sierra Leone, which was completed in 2004. From the census data, our variables of interest are the fraction of households involved in crop farming, the fraction of individuals over ten years old involved in crop farming, the literacy rate and the years of education completed by household heads.

3.2.3 Agricultural Households Tracking Survey (2010)

To identify areas of high productivity or high surplus for different crops, we use measures of agricultural production at the road level, computed from the Agricultural Households Tracking Survey (AHTS). The AHTS was a nationally representative agricultural survey conducted in March-April 2010 with a sample of 8,803 farming households. Since the AHTS was administered potentially concurrently with the implementation of the EU program, we do not use it as baseline data. Instead we only use the household-level data to compute measures of local production around the markets for rice and cassava. We also use measures of the fractions of households selling rice and cassava. We compute chiefdom-level averages, as discussed in more detail below.

3.2.4 Endline Roads Data (2011)

We designed novel GIS data collection strategies to verify whether the roads had indeed been rehabilitated, and to collect objective measures of transport costs. These variables also provide evidence for the first stage of our analysis.

First, we drove a sport utility vehicle with a geo-referenced camera secured to the hood along each road. This exercise was completed for 46 out of the 47 eligible roads in November 2011.[15] The camera recorded the GPS position and speed of the vehicle every second, as well as collected a continuous stream of video along all roads. In the analysis presented below, we use average speed on the roads traveled as one of our first-stage variables of interest.

In another effort to gather first-stage evidence of the impact of the roads, we collected data on transport fares along all 47 roads. We negotiated with motorbike taxi riders (locally known as *okadas*) for a route fare on every road,[16] and actually travelled with the *okada* on a random subset of roads. *Okadas* are the most commonly available type of public transportation available in rural Sierra Leone, for both people and goods, and unlike buses they can travel on most roads

---

[15]One road was missed as part of this exercise because it could not be located. The road is 2.6 km long and was not selected for rehabilitation. Example videos are posted at `http://www.mit.edu/~tavneet/research.html.`.

[16]This exercise was designed so that the surveyor bargaining with the taxi was the residual claimant.

all year round. These data give us a measure of actual transport costs and freight rates on the project roads.

We show summary statistics on these first stage variables in Table 2. The average speed on the roads was about 21 km per hour and the motorbike fare was an average of 16,000 Leones per road ($3.72). Since the roads are of different lengths we report the fare per km travelled, which is about 900 Leones ($0.21) per km.

3.2.5 Trader Surveys (2011-2012)

To understand the effects of these road improvements on prices, we use data from two waves of a nationwide survey targeting rice traders in rural markets. The first wave, conducted in February/March 2011, targeted all the markets located within five kilometers of the 35 roads that were within five positions of the rehabilitation threshold in each of the four districts. In addition, random sampling of the remaining markets located in the rest of the country led to a sample of 54 markets located within 11 kilometers of the 47 roads. The second wave, conducted in January/February 2012, included the universe of markets (82) located within 11 kilometers of any of the 47 roads, including the 54 markets sampled for the first wave.[17] Within each market, we first listed all traders of local rice, then randomly sampled two traders to participate in the survey (or surveyed the unique trader when only one trader was present).

Table 2 shows summary statistics from these trader surveys. Sale prices from traders most recent sales transaction are presented for local rice, but not for cassava, which was not covered by this particular survey. We report prices in cups, as this is the standard unit across all our price data.[18] The trader surveys were conducted during the harvest season (February/March), a period when local rice is plentiful and the price of local rice is lower than its annual average.

The GPS coordinates of the markets are used to compute as-the-crow-flies distances between rural markets and larger towns or urban centers in Sierra Leone. We define a large town to be either one of the six largest towns in the country based on the 2004 census (these towns were Freetown, Bo, Kenema, Makeni, Koidu, Waterloo and Lunsar, with a median population across these towns of 81,000 individuals) or any district headquarter town across all the 13 districts in the country. This gives us a total of 14 unique large towns, however, our markets only match to a subset of nine of those towns.[19]

---

[17]The threshold of 11km was defined before the second wave of the survey to maximize the number of markets surveyed under tight survey budget constraints. Within our sampling frame of markets, 54 were located within 5 kilometers of the nearest road, 78 were located within 10 kilometers, and 86 within 11 kilometers. Out of the 86 markets visited, 4 had no rice trader at the time of the survey, leading to the final number of 82 markets.

[18]There are roughly four cups of local rice per kilogram. Two varieties of local rice are sold by traders - parboiled and milled. These varieties are nearly identical to each other. We averaged these prices to obtain one local rice price.

[19]Some of the large towns are not in the districts in which we have our RDD, but are still the closest big town for a market in our sample. An example is the town of Kailahun, which is in Kailahun District but is very close to some of the markets in our sample as can be seen from Figure 1.

3.2.6 High Frequency Price Surveys (2011-2012)

Finally, between May 2011 and July 2012, we targeted the markets included in the trader survey sample to collect monthly price data via phone surveys. We think of the prices collected as the average price in the market for that month. The respondents were individuals identified as focal points for the markets during the data collection for the trader surveys. The number of markets for which we sought monthly pricing data increased after the tenth wave of calls to include the 28 additional markets targeted by the second wave of the trader survey. We focus on two main crops in our analysis of the high frequency data: local rice and a common type of processed cassava known locally as *gari*.[20] Table 2 shows summary statistics on the prices of these two crops, again reported in cups. The price of a cup of gari is about 340 Leones ($0.08) and that of a cup of rice about 1,000 Leones ($0.23).

4 Empirical Strategy

4.1 Regression Discontinuity Design

The design of the EU program allows us to use an RDD[21] to identify the causal effect of rural road rehabilitation on transport costs and rural market prices.[22] We have a small number of clusters for the RDD, though a large number of observations, especially when using the high frequency price survey data. This RDD is thus closer in spirit to the geographic or boundary RDD studies (for example, see Dell (2010)).

In what follows, we describe the identification problem using the standard Rubin (1974) potential outcomes framework. Define $y_i(1)$ as the potential outcome for road i when the matched road is rehabilitated by the EU program and $y_i(0)$ when it is not rehabilitated. Then the (sub) population parameter of interest is $\Delta = E[y_i(1) - y_i(0)]$, the average treatment effect (ATE). The basic identification problem arises because the econometrician only observes the realized outcome based on the actual rehabilitation status T_i of the road:

$$y_i = T_i \cdot y_i(1) + (1 - T_i) \cdot y_i(0) \tag{11}$$

---

[20]Gari is a form of processed cassava similar in aspect to kuskus or bulghur, common throughout most of West Africa. Gari is obtained from peeling and grating cassava tubers often using manual equipment.

[21]We refer the reader to Imbens and Lemieux (2007) and Lee and Lemieux (2010) for extensive reviews of the RDD.

[22]Our main outcomes of interest are measured either at the road level (transport costs) or at the market level (market prices). In addition, since we do not have baseline (pre-intervention) data for market prices, we use instead road level data collected from household surveys in Enumeration Area (EAs) around the project roads during the 2004 Census and Chiefdom level data collected from an agricultural household survey for the heterogeneous treatment effects analysis. For the sake of clarity, in this section we only refer to our outcomes as being road-level or market-level. However, the reader should bear in mind that some of the variables used for baseline checks were actually collected at the household level and appropriately aggregated.

An OLS framework would lead to an inconsistent estimate of Δ in the presence of unobservable covariates that are correlated both with the outcome and with the rehabilitation status. With multiple rounds of data, a difference-in-difference estimator would be biased if changes in the rehabilitation status are correlated with changes in other unobservable variables.[23]

Our RDD relies on the comparison between roads "just above" and roads "just below" the rehabilitation cutoff. In the case of perfect correspondence between the rehabilitation plan and the actual rehabilitation status, this empirical strategy would identify the average treatment effect (ATE) around the cutoff. This ATE is local in the sense that results around the treatment effect at the cutoff cannot be generalized to other points in the domain of the forcing variable. In the context of this study, we argue that our empirical design identifies the marginal effect of a program expansion, which is a policy relevant local average treatment effect.

We define S_i^N as the road score, normalized to zero at the mid-point between the first road below the cutoff and the first road above the cutoff.[24] Our empirical analysis starts with a graphical approach. For each of the variables of interest, we plot the bin level means of the outcome and include a linear fit of all the underlying data. We restrict the graphs to be within our chosen main specification bandwidth (see below) so that the graphs are easily comparable to the parametric results. For the parametric estimation, we restrict our attention to roads that are "close" to the cutoff, i.e. the sample of roads whose score is within h points from the cutoff ($S^N \in [-h, h]$). We use one main bandwidth as our preferred specification, $h = 0.15$, which delivers a subsample of 31 roads. This bandwidth of $h = 0.15$ is close to the Imbens and Kalyanaraman (2009) optimal bandwidth across a range of our outcomes at the road-level. As Lee and Lemieux (2010) mention, optimal bandwidth algorithms may suggest nadwidths that are larger than the rule of thumbs used by many researchers. Since we have a small sample size, we choose the optimal IK bandwidth as our main specification. We also report results in our main tables for two additional bandwidths, ($h = 0.075$ and $h = 0.3$) which span half and double the preferred bandwidth, and deliver subsamples of 18 and 38 roads, respectively. For our first stage variables, where our outcomes are at the road level, we do not report results for the $h = 0.075$ bandiwdth as these specifications only have 18 observations and are therefore not particularly meaningful.

For the preferred bandwidth as well as the two alternatives, we run a local-linear regression of the form:

$$y_{ik} = \gamma_0 + \gamma T_i + \gamma_R T_i * S_i^N + \gamma_L (1 - T_i) * S_i^N + \delta_k + \epsilon_{ik} \tag{12}$$

---

[23]Studies using a differences-in-differences approach include Ali (2011), Bakht, Khandker and Koolwal (2009), Khandker and Koolwal (2011), Bell and van Dillen (2012), and Mu and Van de Walle (2007)).

[24]Our results are similar if we normalize the cutoff to be at the first treated road in each district. These results are available upon request.

where y_{ik} is a road-level outcome, i denotes the road and k the district. We control for district fixed effects in the road level specifications and we present heteroskedasticity robust standard errors.

For market-level outcomes (prices), the regression model is:

$$y_{ijkt} = \gamma_0 + \gamma T_i + \gamma_R T_i * S_i^N + \gamma_L (1 - T_i) * S_i^N + \delta_k + \mu_t + \epsilon_{ijkt} \tag{13}$$

where i denotes the road that the market is close to, j is the market, k is the district (we control for both the district the road is in as well as the district the market is in if different) and t is the survey round. For the trader survey, our regressions are at the trader level and we have multiple traders per market.

In addition to the local-linear regressions, we use the full sample of 47 roads and include a third-order polynomial approximation in S_i^N. The road-level regression model becomes:

$$y_{ik} = \gamma_0 + \gamma T_i + \sum_{j=1}^{3} \gamma_{jR} T_i * S_i^N + \sum_{j=1}^{3} \gamma_{jL} (1 - T_i) * S_i^N + \delta_k + \epsilon_{ik} \tag{14}$$

The market-level model presents analogous changes relative to equation (13).

When running the market-level regressions, we adjust our standard errors for two way clustering, allowing for clustering at the road level as well as at the market level. This is important since some markets match to multiple roads. There are 82 unique markets in our surveys. On average, we have about 6 markets matching to each road. Approximately 48% of markets match to just one road. The others match to two or more roads, with the average being about 2.8 roads per market. We also weight the regressions by the inverse of the distance between the road and the market in question. For road-level regressions, we present heteroskedacity robust (instead of clustered) standard errors since the level of observation is at the road level.

Finally, we look not just at the treatment effects of the improved roads but also the heterogeneity in these effects along three specific dimensions: distance between the market and the closest town, harvest of the relevant crop (rice or cassava) and the fraction of households selling the relevant crop as described in detail below. For the heterogeneity analysis we use a dummy variable for whether the market is above the median value for the relevant variable. Below, we therefore report results from the specifications described above as well as specifications where we control for these heterogeneity variables, in particular we control for dummies indicating whether the log mean harvest of rice and cassava per household and the fraction of farming households selling those crops in the area are above or below their median value at the chiefdom level and whether the distance to the closest urban center is greater than the median distance in the sample.

4.2 Testing the Validity of the Identification Strategy

The above RDD relies on the assumption that there is no manipulation of the theoretical rehabilitation status around the threshold. One potential challenge to identification therefore comes from the following two compliance issues briefly discussed in Section 3.1.2:

1. In Kenema district, there was one case where the rehabilitation assignment plan and the actual rehabilitation status were not the same - an extra road was treated. This implies that we cannot estimate the average treatment effect of the actual rehabilitation status. Instead, we present the (local) Intent-To-Treat (ITT) estimator identifying the effect of the rehabilitation plan around the cutoff.[25]

2. In Kambia district, the same score was assigned to two roads, but only one road was rehabilitated and there was no clear rule for a tie break. To complicate matters, these two roads are exactly at the cutoff in the rescaled score (i.e. across the whole sample, they are the two roads closest to the cutoff on either side) giving them the highest weight in a classic RDD. Since we lack a clear ranking protocol for this specific case, we present results after dropping these two roads from the sample.[26]

We adopt two standard strategies to test for the presence of manipulation. First, we test for discontinuity in the density of the forcing variable (McCrary (2008)). Second, we inspect whether baseline covariates show discontinuities around the score cuttoff.

Figure 2 presents the results of the McCrary test.[27] Although we do see a jump in Figure 2 at the cutoff in the density point estimate, this jump is well within the confidence interval. In addition, if anything, the jump is downward while any manipulation would predict a jump upward. Finally, given the very low number of observations at the extreme values or tails, the density is not well estimated at these tails. Our density test is, therefore, likely to be a low powered one.

The analysis of covariate balance at the cutoff relies on three data sources. For a subset of variables (the most relevant ones), we also show graphical results. Table 3a focuses on variables collected by the EU as part of their baseline assessment of the roads and directly entering the scoring index. We look at just the population per km graphically - see Figure 3a. Both the figure and Table 3a show no significant discontinuity at the cutoff (only one coefficient of the twenty is significant). Table 3b presents other baseline characteristics of the roads, including variables

---

[25]LATE results are available on request.

[26]This is akin to the recent donut RDD (see, for example, Barreca et al. (2011)). Notice that we include these roads in one of the robustness checks later in Section 5.2.3.

[27]See McCrary (2008) for a detailed description of the test. For our test we use a binsize of 0.025 and a bandwidth of 0.15. Our estimate of discontinuity at the cutoff is -0.325 with a standard error of 0.684.

collected by EDA during the prioritization process but that do not enter the score, and the agricultural and education variables collected during the 2004 Census. Unfortunately, the census has limited agricultural information but we present results for the available variables - the fraction of households involved in agriculture, the fraction of individuals involved in agriculture, literacy and education of the household head. We keep EAs within 2 km of the roads and we aggregate the data to the road level for these tests. We also show results for the fraction of individuals in farming and the education of the household head graphically in Figures 3b and 3c. Neither the figures nor the regressions show evidence of discontinuities at the cutoff - in Table 3b, none of the 28 coefficients are significant.

In Table 3c, we focus on the five variables used for the heterogeneity analysis. This includes four variables collected during the AHTS in early 2010 as proxies of σ: the average rice and cassava harvests per household and the fractions of households selling rice and cassava.[28] Although these variables were not collected at baseline in the strict sense, the harvest variables can be considered predetermined as the relevant farmer decisions (such as acreage planted and input use) were made the prior year, which was before the road rehabilitation. Because harvest data may be noisy, we also look at the extensive margin sale choice as a proxy for σ. Since the majority of the population relies on subsistence farming for the staple crops, the share of producers who enter the market economy is a reasonable measure of the average surplus of rice relative to the consumption needs.

We use chiefdom-level averages of these variables for the chiefdom that a given market is in. Chiefdom-level data are less subject to endogeneity concerns since the roads considered typically affect a small fraction of the population of each chiefdom.[29] The specification for the agricultural variables is therefore at the chiefdom-road level for the set of chiefdoms that the markets belong to. Table 3c also includes a check for the distance between markets and the closest town (as a proxy for x in the heterogeneity analysis). This specification is at the market-road pair level. For the four agricultural variables and the distance, we find no systematic evidence of a discontinuity at the cutoff. For three of the five variables, there is one specification that shows an effect that is both economically and statistically significant (for example, in the tight bandwidth linear specification, the fraction of households selling cassava is more than 50% lower in the treatment group than in the control group). The polynomial specification is responsible for two of the three problematic coefficients. Due to the small sample size in these cells, we think that the polynomial is not a good approximation to the data as it overfits and the magnitude of the coefficients is probably overestimated. Graphically, Figure 3d shows results

---

[28]For this table, we use the continuous version of these variables. This is unlike the heterogeneity analysis, where, in the main specification, we use dummies for whether the value of the variable is above or below the median, as described later in the paper.

[29]Chiefdoms are the third-level administrative unit in Sierra Leone, coming after provinces and districts. There are 149 chiefdoms throughout the country.

for the fraction of households selling rice with little evidence of a discontinuity.

5 Results

We now discuss our empirical results. We first present estimates of the effects of the road improvements on transport costs and speed, and then for each of the three comparative statics for rural market prices that we derived in Section 2. The comparative statics presented in the theory section focus on a change in crop prices with respect to a change in τ, the rural transport costs. The empirical results are for an *improvement* in road quality that *reduces* rural transport costs.

5.1 Speed and Transport Costs

Figures 4a and 4b present a graphical analysis of the impact of the program on our two primary measures of transport costs: travel speed and (log) travel fares per km. As mentioned above, for the graphs, we show the data within the main bandwidth, $h = 0.15$ - we show bin level means for the data and a linear fit of all the data. Table 4 shows the results from the parametric regressions on these variables according to the specifications outlined in Section 4. However, the results from the $h = 0.075$ bandwidth for these road level outcomes are not meaningful as the sample size is too small and given the number of controls for the RDD specification, we are left with too few degrees of freedom. Looking at the impact of road quality improvements on speed, the bin means show a jump at the discontinuity (see Figure 4a), but the flexible polynomial specification is driven up at the left of the discontinuity. This is due to the fact that, as discussed earlier, the last road on the left of the cutoff was in fact rehabilitated even though it should not have been, based on the pre-determined decision rule. The parametric analysis in Table 4 confirms this intuition: the wide ($h = 0.3$) bandwidth local linear specifications point at an increase in speed across the cutoff. The results are also robust to the inclusion of controls we later use for the heterogeneity analysis. The polynomial is less robust, but is likely to be overfitting the data. The discontinuity in transport fares is visible (though noisy) in Figure 4b and in Table 4: an improvement in road quality has a negative, significant impact on fares. In the local linear regression with the optimal bandwidth ($h = 0.15$), the road improvements lead to a 59% reduction in transport costs per km. This result is robust to the inclusion of the controls we use for the heterogeneity analysis in later sections of the paper.

5.2 Rural Market Prices

5.2.1 Basic Price Effects

We now look at the effect of these improvements on crop prices along the roads. First, we relate our empirical approach to the theoretical framework above. The impact of rural road rehabilitation corresponds to the first comparative static of interest we presented in Section 2, $\frac{\partial p^S}{\partial \tau}$, in Table 1. Since the theoretical models derive predictions with respect to transport costs (not road quality), a negative $\frac{\partial p^S}{\partial \tau}$ implies that improvements in road quality have a positive effect on prices, and vice versa. Hence, if our empirical results are consistent with any given model, the sign of our regression coefficients should be the *opposite* of the signs of the theoretical predictions.

Our price results focus on the two main staple crops produced domestically, rice and cassava (traded in the form of *gari*). In our data, rice and cassava are available 85% and 96% of the time, respectively (considering market-round level observations), and both crops are sold in all markets at some point during the year. Before looking at prices, we first analyzed the effects of the road improvement program on crop availability in markets. The RDD treatment effect on availability of local rice is -0.10 with a standard error of 0.07. For cassava, the effect is -0.04 with a standard error of 0.06 (detailed results available upon request). We interpret this as evidence that market availability on the extensive margin was not different across treatment and control roads, although we unfortunately do not have data on quantities traded that would allow us to further test this result.

Figure 5 and Table 5 (columns (1) and (2)) show that road rehabilitation leads to a substantial reduction in the sale price of rice in the trader data. The graphical analysis shows a discontinuity in the price of rice at the cutoff. The price decrease is significant across all specifications (columns (1) and (2) of Table 5). In the rice price data from the high frequency surveys, evidence for a treatment effect in the form of a price reduction is much weaker (see Figure 6 and Table 5, columns (3)-(4)). While still negative in the tight bandwidth, the point estimates in these regressions are much smaller and mostly not significant. The results for rice from the high frequency data are different from those in the trader survey. One probable reason for this is the timing being different for the two surveys. The trader survey is conducted just after harvest, which the high frequency surveys are collected every month. We do not have enough sample size to look at just the harvest period in the high frequency data as we only have one year (hence one harvest season) of data. For cassava there is no specific harvest season. Cassava can remain in the ground for multiple seasons and can be harvested any time and is usually only harvested for immediate consumption or sale (it has to be processed immediately it is harvested).

Figure 7 and Table 5 (columns (5) and (6)) cover price effects for *gari*, based on the monthly price surveys only. The trader surveys did not cover cassava. Table 5 shows a large price drop

in cassava. In the optimal bandwidth specification (see the top panel in column (5)), the road rehabilitation reduces prices of cassava by 17.8% (significant at standard levels). This result is robust to different bandwidths, a polynomial specification, and to the inclusion of the heterogeneity controls described above. The effect on cassava prices is clearly larger than the effect on rice prices. We expect this to be the case as cassava is usually processed at the market site rather than on the farm,[30] and raw cassava is more expensive to transport than unprocessed (husk or threshed) rice.[31] Hence, cassava prices should be more responsive to changes in road infrastructure.

In summary, Table 5 provides evidence of price reductions for both rice and cassava in response to the road improvements. We interpret this as evidence that the road improvements have primarily facilitated access to rural markets for producers, as opposed to reducing transport costs on rural roads for traders. More formally, using the notation from our theoretical setup, these results imply that, in our sample, $\alpha > \beta$, which is consistent with the qualitative evidence that, on average, city traders' transport costs are primarily determined by the quality of the major roads, as opposed to rural road infrastructure. In our earlier terminology, the supply effect prevails over the demand effect. Most importantly, the stronger relative importance of the supply effect (versus the demand effect) is inconsistent with a basic model of Bertrand competition. In the Bertrand framework, a change in transport costs affects the equilibrium price only through β, the costs for the trader. Similarly, the Cournot case with endogenous entry predicts that a reduction in transport costs would unambiguously increase prices. These predictions are not borne out in the empirics.

5.2.2 Heterogenous Price Effects

We now focus on the heterogeneity in the price responses to road rehabilitation. As described in Section 2, the sign of the heterogeneity in these effects can shed light on the underlying market structure and allow us to empirically test between the different classes of models we presented above. Using insights from the theory, our analysis of heterogeneous effects focuses on two market characteristics: the distance between the rural market and the closest urban center (the variable x in the theory) and crop-specific productivity around the markets (the variable σ in the theory). We proxy x by the Euclidean distance between the market and the closest large town, as described above. In order to proxy for agricultural productivity, we use 2010 household survey

---

[30]In the AHTS (2011), only 2% of households and 8% of villages reported being equipped with at least one cassava grater; 0.05% of households and 8% of villages were equipped with a mechanized rice mill.

[31]Since we do not have accurate data on the transport cost of unprocessed cassava, we justify this by the fact that cassava tubers (raw cassava) are more voluminous (and hence more expensive, holding volume constant) to transport than most other crops. In addition the processing rate to transform raw cassava into *gari* is low: 1 unit of unprocessed cassava yields 0.36 units of *gari* based on the AHTS (2011) data, while the typical milling rate for rice is 0.5. In other words, producers must transport larger volumes of raw cassava than rice to the marketplace to sell the same volume of the processed crop.

data to compute a chiefdom-level measure of 2010 harvest per farming household for both rice and cassava and the fraction of farmers selling the crop, as described above.

Before presenting the results, we describe our empirical framework in some detail. First, we work with dummy variables that indicate whether a certain market is above or below the median value of each of our heterogeneity variables. Using these binary indicators allows us to present the results in an intuitive way and to quantify the economic magnitude of the interactions easily.

In Table 6 (columns (1) and (2)), we present our heterogeneity analysis of price effects for rice and cassava based on the distance between markets and the nearest large town. The price reductions induced by road quality improvements are always stronger (i.e., more negative) for markets that are further away from large towns. The sign of the interaction term of treatment and the dummy indicating the market is far from urban centers is negative and statistically significant in all four specifications for rice, and all but the wide-bandwidth local linear regression for cassava. By contrast, the sign of the treatment coefficient is always positive, and significant in all cases for rice, but in no cases for cassava. Based on this hypothesis, we also report one-sided p-values on the treatment interaction coefficient in each panel, all eight of which are significant at the 10% level (and seven are significant at the 1% level).

Next, we examine the heterogeneous impacts on crop prices by the chiefdom-level average household harvest, one of our proxies for productivity (see Table 6, columns (3) and (4)). For rice, we find that the treatment effect is negative in low-productivity areas, and in fact positive in the high-productivity areas, indicating a stronger demand effect in these areas. With respect to rice, the coefficient of treatment interacted with high-productivity areas is statistically significant in the optimal and tight bandwidths, but not significant in the other two specifications. We find similar results for cassava, with a strongly negative treatment effect in low-productivity markets (significant in the tight bandwidth and polynomial specifications) and a coefficient that is positive and statistically different from zero in two specifications in the high-harvest areas. The one-sided p values are significant in four of the eight specifications.

In columns (5) and (6), we present evidence of treatment effect heterogeneity by our second proxy of farmer productivity, the share of farmers selling any of their harvest. The price reduction induced by the improvements in road quality is significantly weaker in markets with higher density of sellers. For rice, in markets with a high density of sellers, the price effects are positive. For both crops, the difference in the percentage effect is economically sizable, although it is larger for cassava. The significance of the interaction term, as well as the one-sided p-values (computed for the relevant coefficient sign) in each panel, indicate that the effects across low and high seller density markets are significantly different from each other for all specifications for cassava, and all except the wide-bandwidth specification for rice.

Summarizing these results, we find evidence that market-specific characteristics, both on

the demand side (distance from major market centers) and on the supply side (productivity), significantly affect the magnitude and the direction of the price effect of improvements in rural road infrastructure. Turning again to the models presented in Section 2, these findings are inconsistent with both the bilateral bargaining framework as well as the Cournot model with an exogenous number of traders.[32] In the bilateral bargaining setting, the productivity parameter affects the size of the surplus but does not affect the optimal solution to the bargaining problem. In addition, the distance between the rural market and the urban center, x, affects the outside option of the city-traders but does not affect the impact of rural road quality on the bargaining solution. The Cournot model with an exogenous number of traders also predicts that neither x nor σ affect the price derivative with respect to rural transport costs. For the case of x, it shifts the net city price but does not interact with rural transport costs in determining prices. Similarly, σ does not enter the equilibrium price.

Our empirical findings are therefore consistent with a model with search frictions, as formalized by Chau, Goto and Kanbur (2009). In this framework, a reduction in rural road transport costs has both demand and supply effects on prices. It also predicts that more isolated markets have lower entry and hence more imperfect pass-through and that markets with higher producer productivity have lower real search frictions and hence stronger demand effects. We find evidence of both forms of heterogeneity and in the directions predicted by a search framework.

5.2.3 Robustness Checks

Given our small sample size and empirical design, we report three different sets of robustness checks on our main results. More specifically, we look at three sets of checks: alternative specifications, alternative samples, and a small sample correction. We report all robustness check results for the main price effects in Table 7 (both with and without the heterogeneity controls) and for the heterogeneous effects in Table 8.

In Panel A of Table 7 and Table 8, we present results from three additional specifications: (i) one that controls for second-order polynomials in the score (separately on both sides of the cutoff), (ii) one that uses a triangle weight in the local linear regression (i.e., a weight that places less emphasis on observations further from the cutoff), and (iii) one that uses continuous heterogeneity variables (reported only in Table 8) instead of the 'greater than median' dummies we use in Table 6. As Panel A in each of Tables 7 and 8 shows, our main results are robust to these alternative specifications. In some cases we lose precision in our estimates, but the coefficients do not change significantly.

With respect to alternative samples, we present results from four additional specifications:

---

[32]The Bertrand model and the Cournot model with an endogenous number of traders were ruled out on the basis of our main price effects in Table 5.

(i) one that includes the two roads at the cutoff in Kambia District[33] (ii) one that restricts the sample to just the closest market to each road and therefore has no weighting, (iii) one that restricts the sample to markets within 5km of each road and does not weight these results by the market-distance to the road, (iv) one that excludes markets that are in more urban areas, and (v) one that excludes markets that match to more than four roads. Once again, looking at Panel B in Tables 7 and 8, we see our results are robust across all these alternative samples though we lose power especially in the specifications with low sample sizes (such as the specification with just one market per road).

Finally, Panel C in each of Tables 7 and 8 includes a robustness check to address our small sample size and low number of clusters. Only 47 roads were considered for rehabilitation, but since we have market-level data, our results are generally two-way clustered by market and road. We are not worried about the number of markets as we have many more markets than roads. Our concern would rather be on the road-level clustering. From the applied literature (see Cameron, Gelbach and Miller (2008) and Angrist and Pischke (2008)), it seems that 47 clusters is not small enough to be a concern, as is the case with the local linear regressions with our preferred bandwidth where we have a total of 31 roads. However, as a check, we implement a technique that has been proposed for cases with small numbers of cluster, the wild bootstrap (Cameron, Gelbach and Miller (2008)). Since our concern is clustering at the road-level, as a preliminary step, we show that one-way clustering by road does not change our results for the preferred bandwidth specification (see Panel C of Tables 7 and 8). We then report the 95% confidence interval from a wild bootstrap routine also based on clustering by road only. The wild bootstrap confidence bands support our main results. We retain significance in both Tables 7 and 8 as per our main results.

6 Conclusion

Road infrastructure projects represent a large share of current foreign aid and government expenditure in Sub-Saharan Africa. Promoters of these investments argue that enhancing market access is a prerequisite for agricultural development and economic growth. As such, the assessment of the impact of local improvements in rural road quality has immediate policy rel-

---

[33]In the documents given to us by the EU, these two roads were right at the cutoff with one treated and one not. Their ranks were identical as their listed scores were tied at two decimal places - the level of precision used to decide rankings. Given there was no rule for a tie, to allocate these two roads to treatment or control, we computed the score from the underlying data and allocate the road with the higher score to the treatment (this road has a higher score only at the third decimal place, while the EU reports only considered scores up to two decimal places). Using this allocation rule, the road we allocate to the treatment group was, in fact, not rehabilitated, but the control road was rehabilitated. In the paper, due to the lack of a clear description of how the rule operated in case of ties (and hence of how these two roads were allocated to treatment and control), we drop the two roads from the sample we use in the main specifications.

evance. In addition, any rigorous estimates of this impact should focus not only on the overall effect of the infrastructure improvements, but also on the variation in this impact across markets, in particular across the proximity of the market to urban centers and across productivity.

In this paper, we make two contributions relative to the existing literature. First, we use the specific design of a rural road rehabilitation program in Sierra Leone to estimate the causal impact of road quality on rural market prices of rice and cassava, the two most important staple crops in the country. We also estimate how this response varies with specific market characteristics. Second, we use these empirical findings to test between alternative stylized frameworks of intermediated trade which describe the interaction between rural producers, intermediaries and final consumers. Specifically, we model responses to a change in rural road transport costs under the following four settings: (i) Bertrand traders' oligopsony, which leads to perfect competition outcomes; (ii) bilateral bargaining, which models the case of rural monopsonies; (iii) Cournot oligopsony, both with and without an endogenous number of traders; and (iv) a basic search framework à la Mortensen (2003). We show that these frameworks deliver different predictions on how prices respond to a reduction in rural transport costs. The predictions for how these price effects vary also differ. Comparing these theoretical predictions to our empirical findings, we find support for the model with search frictions but not for the other frameworks of trader competition.

The evidence on search frictions is consistent with a recent literature on the effects of information on prices (see Jensen (2007), Aker (2010), and Goyal (2010)). However, relative to this literature, we show that search frictions that determine the nature of traders' competition also have an effect on how the local economy responds to improvements in infrastructure. More generally, we point out that the structure of the market for agricultural commodities will affect the direction and the magnitude of price response to such investments, thus affecting the size of the gains for different types of agents. We believe our results could inform the decisions of policymakers considering similar infrastructure improvements as well as a wide range of other policies that would impact agricultural trade, such as price subsidies, agricultural export promotion interventions, credit-provision policies targeting traders, and more major road projects. In addition, finding evidence consistent with the presence of search frictions implies that there may be strong complementarities between roads (or other infrastructure projects) and policy interventions that may reduce search frictions, such as the expansion of mobile phone coverage.

Possible extensions of this research would look at the longer-term effects of the rural roads improvements. Our study focused on the short-term impacts of these improvements since our surveys were conducted between one and two years after the completion of the rehabilitation program. Given the characteristics of these roads, i.e. that they are small dirt roads, we potentially expect the infrastructure to degrade over time and hence the impact of the rehabilitation to be more limited as this degradation occurs.

References

[1] Aker, Jenny (2010), "Information From Markets Near and Far: The Impact of Mobile Phones on Grain Markets in Niger", *American Economic Journal: Applied Economics*, 2(1): 46-59.

[2] Ali, Rubaba (2011), "Impact of Rural Road Improvement on High Yield Variety Technology Adoption: Evidence From Bangladesh", Working Paper, University of Maryland.

[3] Allen, Treb (2012), "Information Frictions in Trade", Working Paper, Princeton University and Northwestern University.

[4] Antràs, Pol and Arnaud Costinot, "Intermediated Trade", *Quarterly Journal of Economics*, 126(3): 1319-1374.

[5] Atkin, David and Dave Donaldson (2012), "Who's Getting Globalized? Intranational Trade Costs and World Price Pass-Through", Working Paper, Stanford University.

[6] Banerjee, Abhijit, Esther Duflo, and Nancy Qian (2012), "On the Road: Access to Transportation Infrastructure and Economic Growth in China", NBER Working Paper No. 17897.

[7] Bardhan, Pranab, Dilip Mookherjee, and Masatoshi Tsumagari (2012), "Middlemen Margins And Globalization," Working Paper, Boston University.

[8] Barreca, Alan, Melanie Guldi, Jason Lindo, and Glen Waddell (2011), "Saving Babies? Revisiting the Effect of Very Low Birth Weight Classification", *Quarterly Journal of Economics*, 126(4): 2117-2123.

[9] Bell, Clive, and Susanne van Dillen (2012), "How Does India's Rural Roads Program Affect the Grassroots? Findings from a Survey in Upland Orissa", Working Paper, University of Heidelberg.

[10] Borraz, Fernando, Alberto Cavallo, Roberto Rigobon, and Leandro Zipitria (2012), "Distance and Political Boundaries: Estimating Border Effects under Inequality Constraints", NBER Working Paper No. 18122.

[11] Broda, Christian and David E. Weinstein (2008), "Understanding International Price Differences Using Barcode Data," NBER Working Paper No. 14017.

[12] Burstein, Ariel and Nir Jaimovich (2009), "Understanding Movements in Aggregate and Product-Level Real Exchange Rates", Working Paper, UCLA.

[13] Cameron, A. Colin, Jonah Gelbach, and Douglas Miller (2008), "Bootstrap-Based Improvements for Inference with Clustered Errors", *Review of Economics and Statistics*, 90(3): 414-427.

[14] Carruthers, Robin, Ranga Rajan Krishnamani, and Siobhan Murray (2009), "Improving Connectivity: Investing in Transport Infrastructure in Sub-Saharan Africa", Africa Infrastructure Country Diagnostic Background Paper 7 (Phase II), World Bank.

[15] Casaburi, Lorenzo, and Tristan Reed (2012), "Incomplete Pass-Through and Interlinked Transactions: Experimental Evidence from Sierra Leone", Working Paper, Harvard University.

[16] Chau, Nancy H., Hideaki Goto, and Ravi Kanbur (2009), "Middlemen, Non-Profits, and Poverty", Working Paper, Cornell University.

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

[18] Deb, Rahul, and Tavneet Suri (2012), "Endogenous Emergence of Credit Markets: Contracting in Response to a New Technology in Ghana", Forthcoming, *Journal of Development Economics*.

[19] Donaldson, Dave (2012), "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure", Working Paper, MIT.

[20] Dorosh, Paul, Hyoung-Gun Wang, Liang You, and Emily Schmidt (2010), "Crop Production and Road Connectivity in Sub-Saharan Africa", Policy Research Working Paper 5385, World Bank.

[21] EDA (2004), "Identification and Prioritization of Rural Roads in Kambia, Port Loko, Kenema and Pujehun Districts. Final Report, Volume 1".

[22] Faber, Benjamin (2012), "Trade Integration, Market Size, and Industrialization: Evidence From China's National Truck Highway System", Working Paper, The London School of Economics.

[23] Fafchamps, Marcel, and Ruth Vargas Hill (2008), "Price Transmission and Trader Entry in Domestic Commodity Markets", Working Paper, University of Oxford.

[24] Fafchamps, Marcel, Eleni Gabre-Madhin, and Bart Minten (2004), "Increasing Returns and Market Efficiency in Agricultural Trade", *Journal of Development Economics*, 78(2): 406-442.

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

[26] Gopinath, Gita, Pierre-Olivier Gourinchas, Chang Tai-Hsieh, and Nicholas Li (2011), "International Prices, Costs, and Markup Differences", *American Economic Review*, 101(6): 2450-2486.

[27] Goyal, Aparajita (2010), "Information, Direct Access to Farmers, and Rural Market Performance in Central India," *American Economic Journal: Applied Economics*, 2(3): 2245.

[28] Gregory, D. I. and B. L. Bumb (2006), "Factors affecting Supply of Fertilizer in Sub-Saharan Africa", Agriculture and Rural Development Discussion Paper 24, Washington, D.C.: World Bank.

[29] Imbens, Guido and Karthik Kalyanaraman (2009), "Optimal Bandwidth Choice for the Regression Discontinuity Estimator", NBER Working Paper No. 14726.

[30] Imbens, Guido and Thomas Lemieux (2007), "Regression Discontinuity Designs: A Guide to Practice", *Journal of Econometrics*, 142(2): 615-635.

[31] Jacoby, Hanan and Bart Minten (2008), "On Measuring the Benefits of Lower Transport Costs", *Journal of Development Economics*, 89(1): 28-38.

[32] Jensen, Robert (2007), "The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector", *Quarterly Journal of Economics*, 122(3): 879-924.

[33] Khandker, Shahidur, Zaid Bakht, and Gayatri Koolwal (2009), "The Poverty Impact of Rural Roads: Evidence From Bangladesh", *Economic Development and Cultural Change*, 57(4): 685-722.

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

[35] Lee, David and Thomas Lemieux (2010), "Regression Discontinuity Designs in Economics", *Journal of Economic Literature*, 48(2): 281-355.

[36] McCrary, Justin (2008), "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test", *Journal of Econometrics*, 142(2): 698-714.

[37] Michaels, Guy (2008), "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System", *Review of Economics and Statistics*, 90(4): 683-701.

[38] Mortensen, Dale (2003), *Wage Dispersion: Why Are Similar Workers Paid Differently?*, MIT Press, Cambridge.

[39] Mu, Ren and Dominique van de Walle (2007), "Rural Roads and Local Market Developments in Vietnam", Policy Research Working Paper 4340, World Bank.

[40] Pushak, Nataliya and Vivien Foster (2011), "Sierra Leone's Infrastructure: A Continental Perspective", *Africa Infrastructure Country Diagnostic Country Report*, Washington, D.C.: The World Bank.

[41] Republic of Sierra Leone (2008), *An Agenda for Change*, Second Poverty Reduction Strategy (PRSP II), 2008-2012, `http://unipsil.unmissions.org/portals/unipsil/media/publications/agenda_for_change.pdf`.

[42] Rubin, Donald (1974), "Estimating Causal Effects of Treatments in Randomized and Non-randomized Studies", *Journal of Educational Psychology*, 66(5), 688-701.

[43] Statistics Sierra Leone (2004), "2004 Population and Housing Census".

[44] Statistics Sierra Leone (2011), "Agricultural Households Tracking Survey Report".

[45] Suri, Tavneet (2011), "Selection and Comparative Advantage in Technology Adoption", *Econometrica*, 79(1): 159-209.

[46] Van de Walle, Dominique (2009), "Impact Evaluation of Rural Road Projects", *Journal of Development Effectiveness*, 1(1): 15-36.

[47] World Bank (2012), *World Development Indicators Online*, `http://data.worldbank.org/data-catalog/world-development-indicators`.

[48] World Bank (2009), *World Development Report 2009: Reshaping Economic Geography*, `http://siteresources.worldbank.org/INTWDRS/Resources/477365-1327525347307/8392086-1327528510568/WDR09_bookweb_1.pdf`.

[49] World Bank (2008), *World Development Report 2008: Agriculture for Development*, `http://siteresources.worldbank.org/INTWDR2008/Resources/WDR_00_book.pdf`.

7 Appendix 1: Formula for Score

The score of road i in district k is given by:

$$score_i = 0.4 * \frac{ep_i}{ep\_max_k} + 0.2 * \frac{pd_i}{pd\_max_k} + 0.2 * \frac{ra_i}{ra\_max_k} + 0.1 * \frac{sv_i}{sv\_max_k} + 0.1 * \frac{l_i}{l\_max_k} \quad (15)$$

where:

ep_i is economic productivity per kilometer of road i,

ep_max_k is the maximum economic productivity in district k,

pd_i is population density around road i,

pd_max_k is the maximum population density in dsitrict k,

ra_i is the baseline road condition assessment of road i,

ra_max_k is the maximum baseline road condition in district k,

sv_i is the social value indicator of road i,

sv_max_k is the maximum social value in district k,

l_i is the length of road i,

l_max_k is the maximum road length in district k.

Figure 1: District Maps of Sierra Leone

Legend
- Markets in sample
- Paved Highways
- Roads rehabilitated
- Roads not rehabilitated

KENEMA

PUJEHUN

Legend
- Markets in sample
- Paved Highways
- Roads rehabilitated
- Roads not rehabilitated

KAMBIA

PORT LOKO

Waterloo
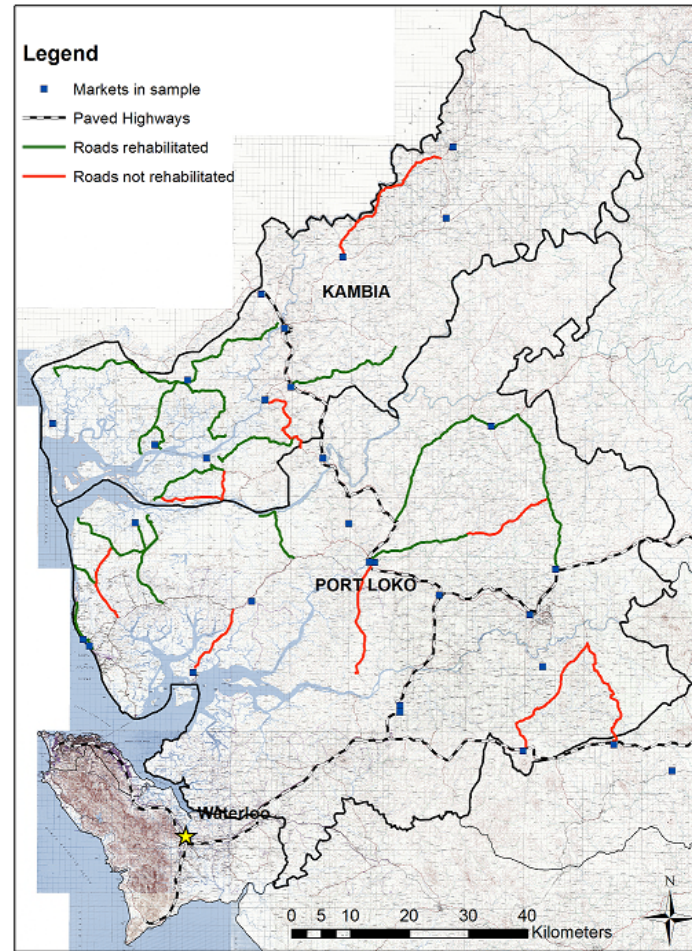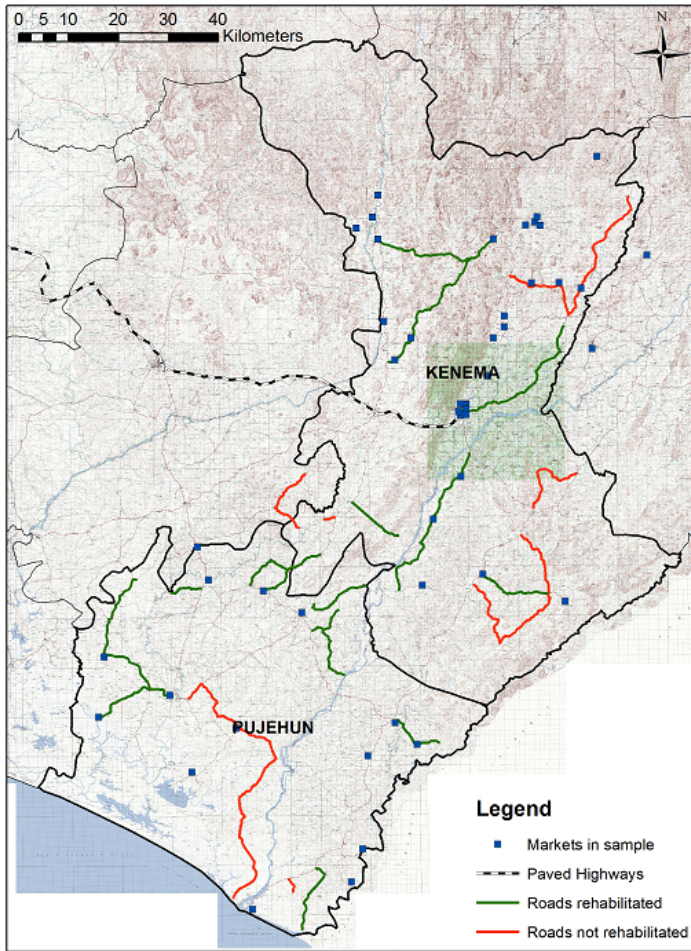
Figure 2: McCrary (2008) Density Test
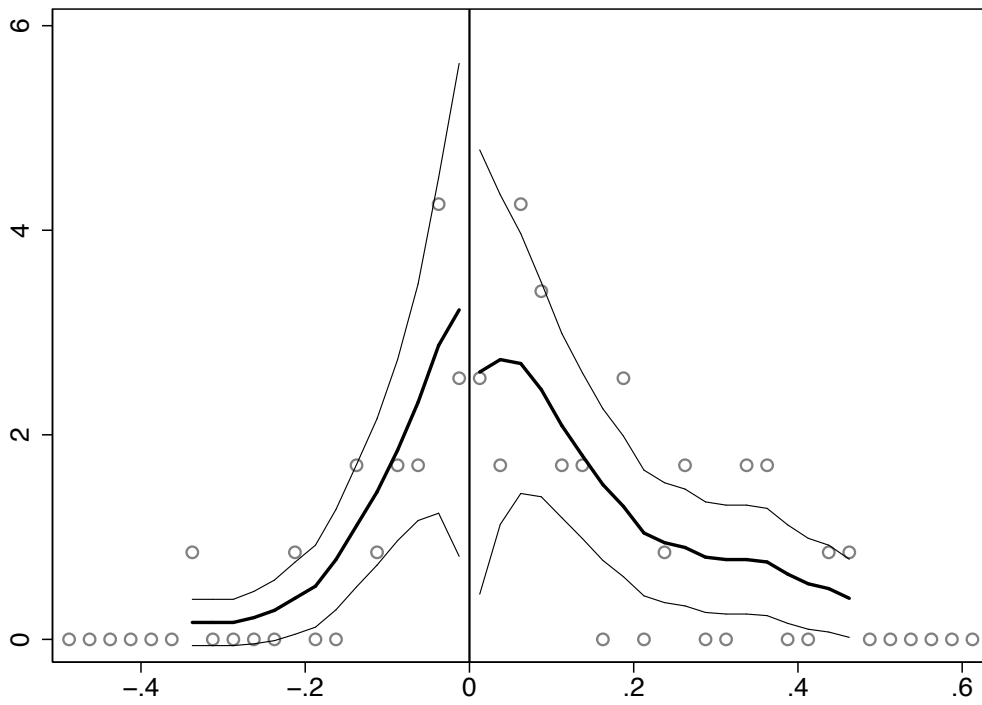
Figure 3a: Baseline: Population per km
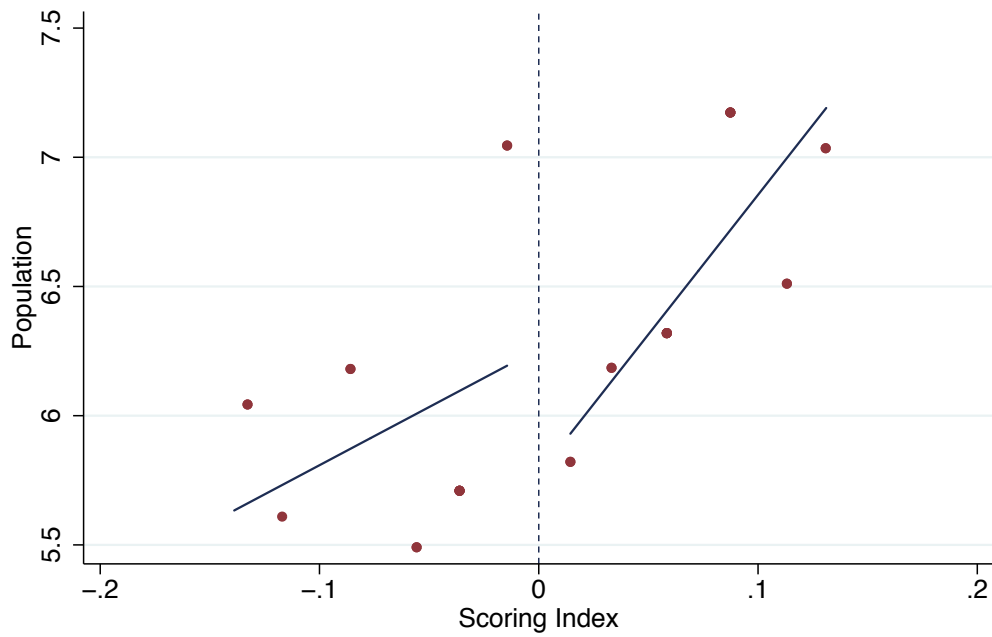
The solid line is a linear fit of the data at the road level.

The scatter represents bin–level means of the outcome (bin size is .025).

Figure 3b: Baseline: Fraction of Individuals in Crop Farming

The solid line is a linear fit of the data at the road level.
The scatter represents bin–level means of the outcome (bin size is .025).
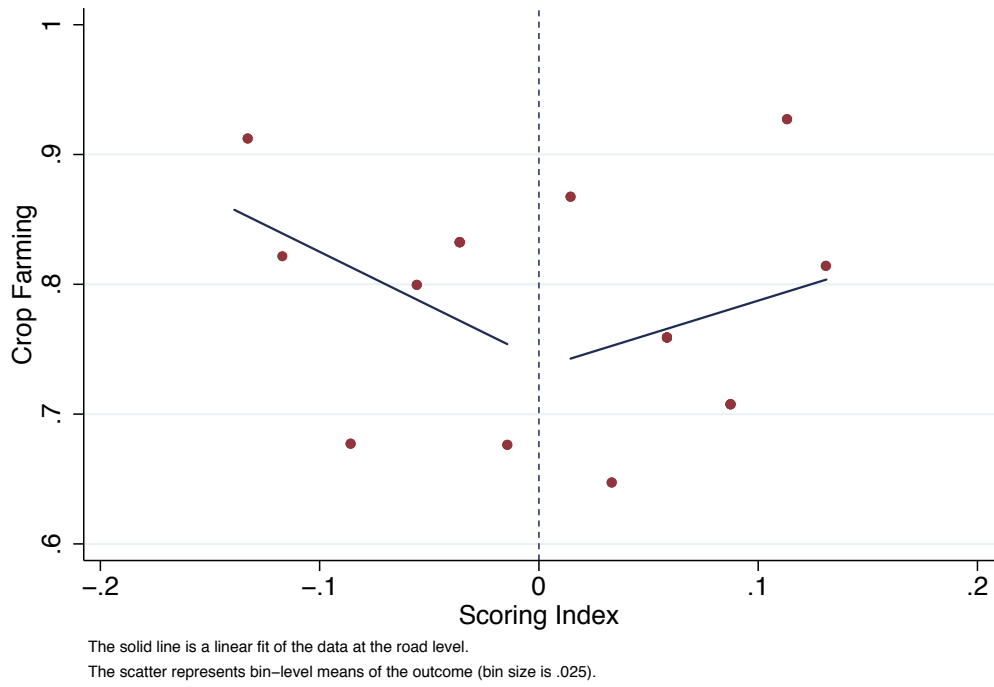
Figure 3c: Baseline: Education of HH Head)

The solid line is a linear fit of the data at the road level.
The scatter represents bin–level means of the outcome (bin size is .025).
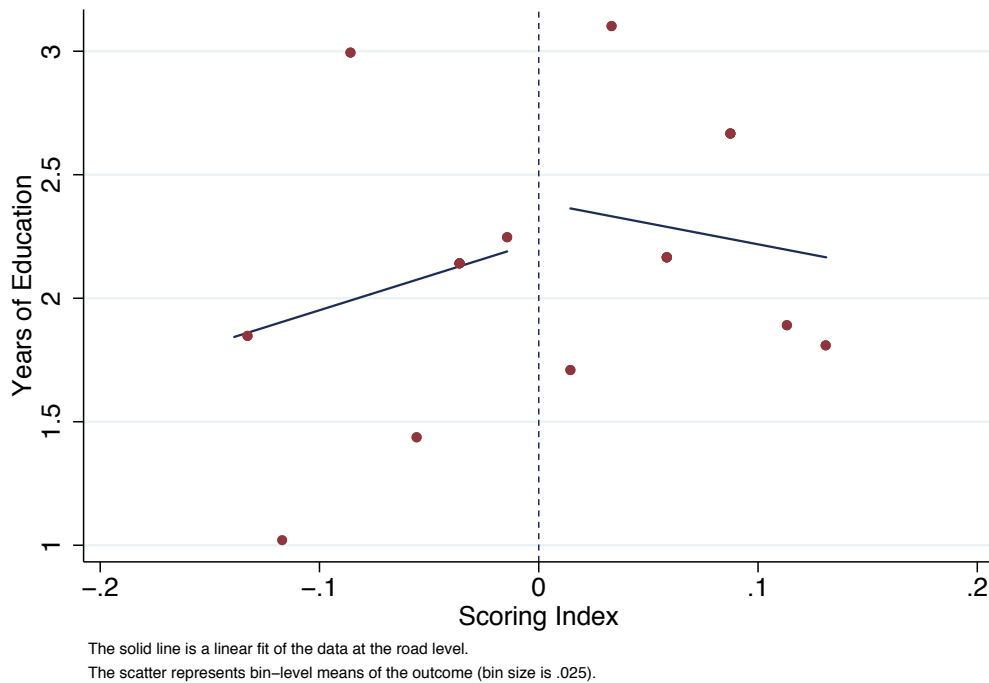
Figure 3d: Baseline: Fraction of HHs Selling Rice

The solid line is a linear fit of the data at the road level.
The scatter represents bin–level means of the outcome (bin size is .025).
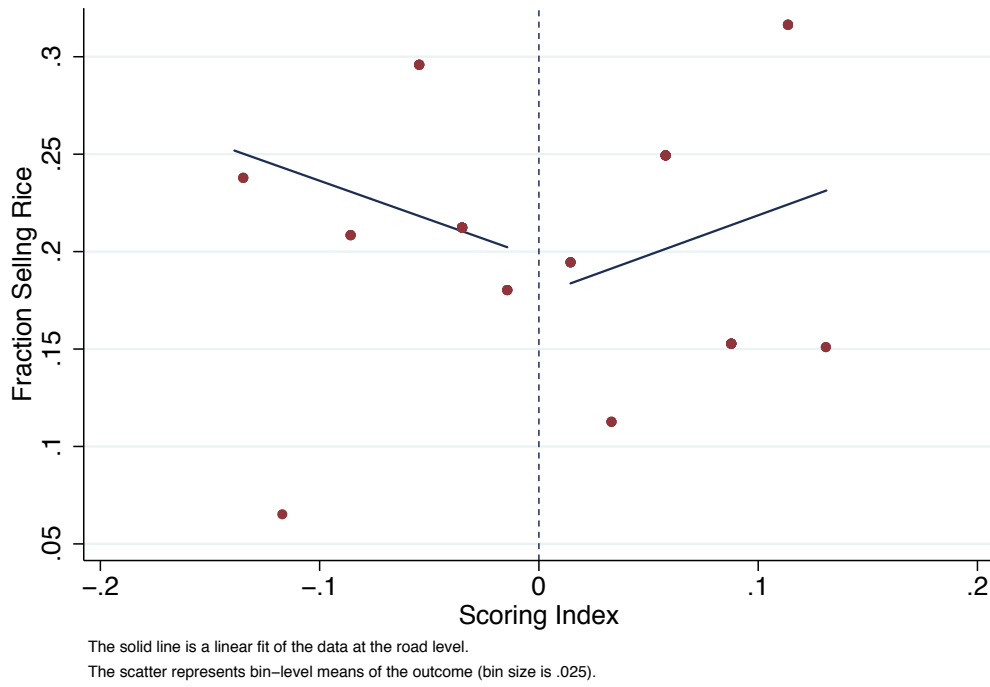
Figure 4a: First Stage: Speed (km/h)

The solid line is a linear fit of the data at the road level.
The scatter represents bin–level means of the outcome (bin size is .025).
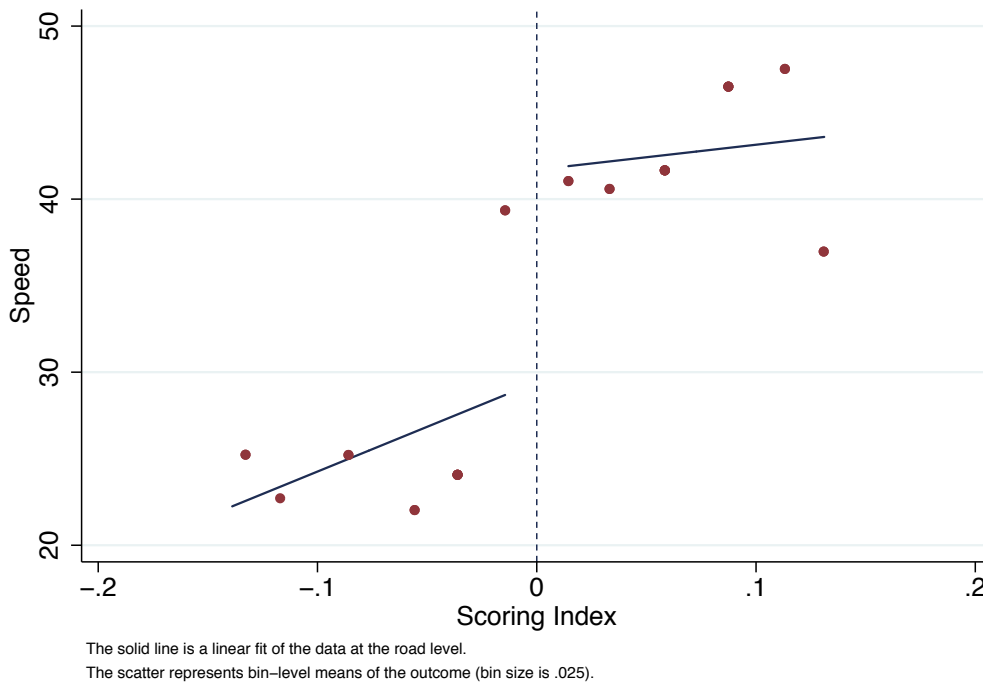
Figure 4b: First Stage: Log Motorbike Fare (per km)
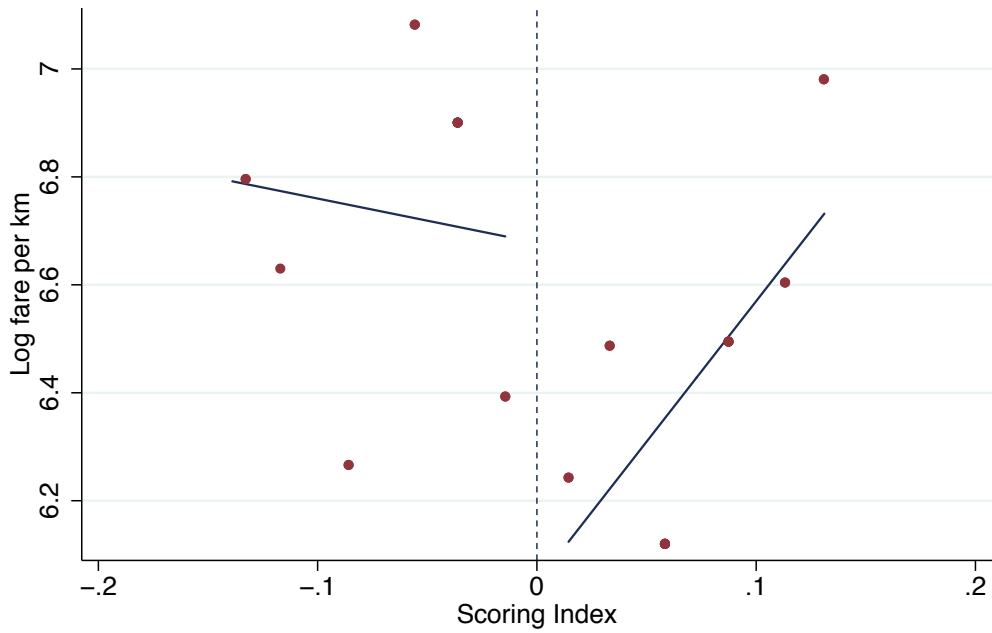
The solid line is a linear fit of the data at the road level.
The scatter represents bin–level means of the outcome (bin size is .025).

Figure 5: Trader Survey: Log Price of Local Rice
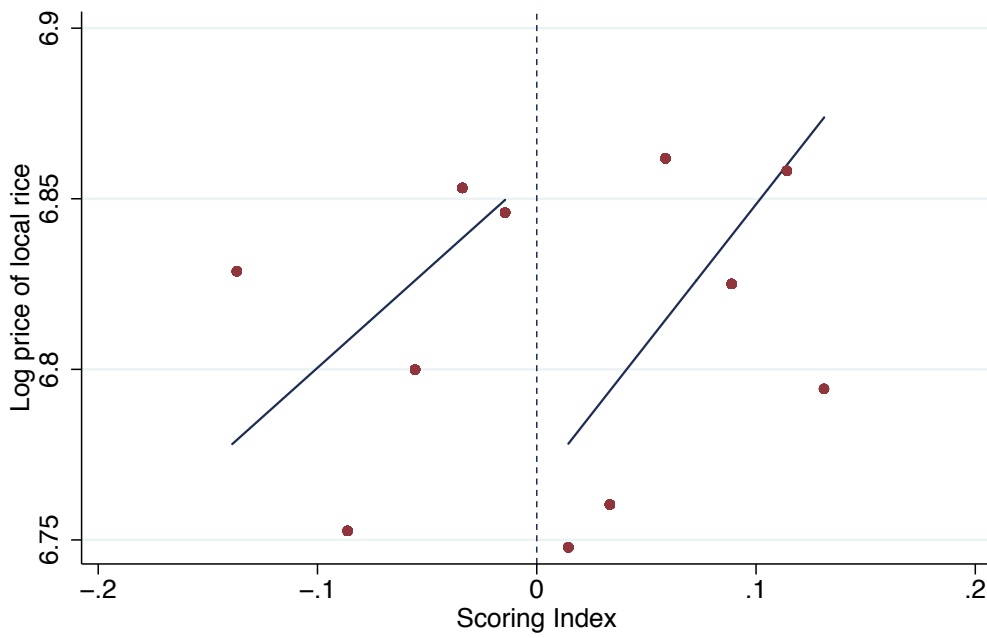
The solid line is a linear fit of the data at the road–market level.
The scatter represents bin–level means of the outcome (bin size is .025).

Figure 6: High Frequency Price Survey: Log Price of Local Rice

The solid line is a linear fit of the data at the road–market level.
The scatter represents bin–level means of the outcome (bin size is .025).
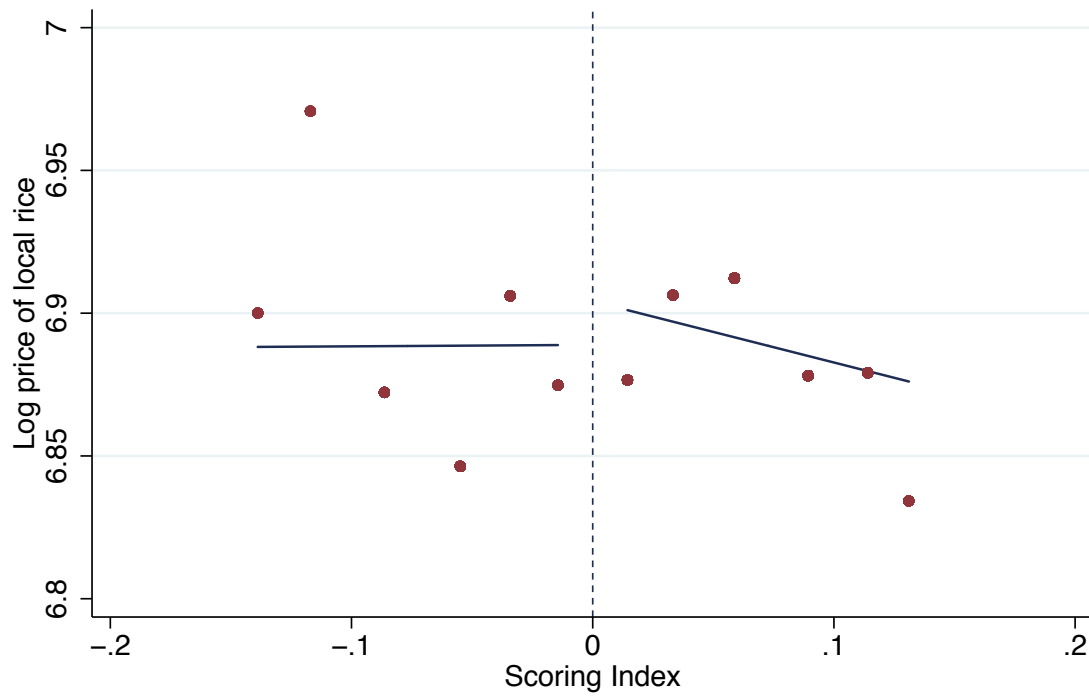
Figure 7: High Frequency Price Survey: Log Price of Cassava (Gari)

The solid line is a linear fit of the data at the road–market level.
The scatter represents bin–level means of the outcome (bin size is .025).
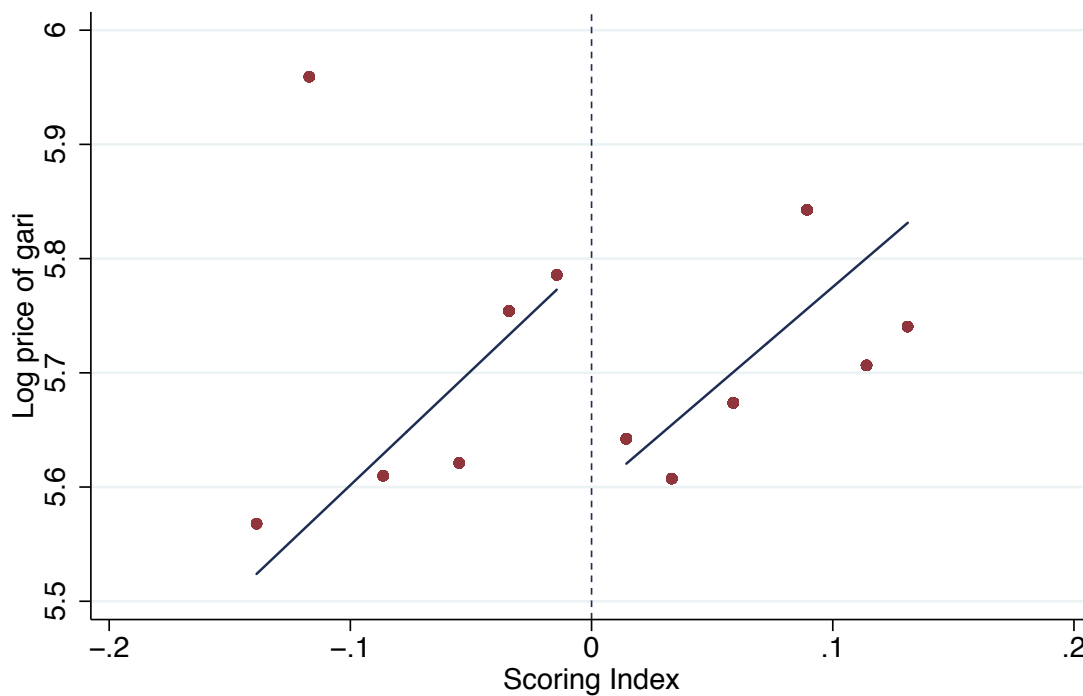
Table 1: Theory Comparative Statics and Empirical Results

Comp. Statics	Bertrand	Bilateral Bargaining	Cournot Exogenous M	Cournot Endogenous M	Search Frictions	Empirical Results	Regression Coefficient ($-\frac{\partial p}{\partial \tau}$)
$\frac{\partial p}{\partial \tau}$	< 0	$\gtrless 0$	$\gtrless 0$	< 0	$\gtrless 0$	> 0	Treat < 0
$\frac{\partial^2 p}{\partial \tau \partial x}$	$= 0$	$=0$	$=0$	$= 0$	> 0	> 0	Treat*Distance< 0
$\frac{\partial^2 p}{\partial \tau \partial \sigma}$	$= 0$	$= 0$	$= 0$	> 0	< 0	< 0	Treat*Harvest per hh> 0; Treat*% Selling> 0

Table 2: Summary Statistics

	Mean	SD	Obs
EA-Level 2004 Census Data			
Proportion farming (all individuals over 10 years)	0.68	0.36	1373
Proportion farming (all households)	0.78	0.32	1373
Proportion literate (all individuals over 12 years)	0.31	0.19	1374
Years of education (HH heads)	2.86	2.26	1369
Roads Baseline Data			
Log population per km	6.53	0.93	45
Log economic production per km	18.60	1.01	45
Road assessment	17.24	3.81	45
Social value	9.56	2.15	45
Number of bridges (excluding palm logs)	3.21	4.07	45
Number of palm log crossings	2.80	5.01	45
Number of culverts	18.29	17.05	45
Road length (km)	20.61	14.72	45
Score rescaled at midpoint	0.08	0.17	45
Chiefdom-Level AHTS Data			
Proportion of HHs selling rice	0.20	0.13	34
Proportion of HHs selling cassava	0.37	0.20	34
Log HH rice harvest (kg)	6.15	0.51	34
Log HH cassava harvest (kg)	7.18	0.53	34
Roads Endline: First Stage			
Total travel fare (one way)	16100	12085	45
Travel fare per km (one way)	900	558	45
Average speed (kph)	36.26	10.42	44
Trader Data			
Local rice purchase price	774.84	90.08	149
Local rice sale price	931.14	105.46	118
Market-Level Price Survey Data			
Price of cassava/gari	339.71	99.92	787
Price of local rice	996.51	143.06	764
Cassava/gari available (proportion of mkts)	0.95	0.21	918
Local rice available (proportion of mkts)	0.84	0.37	918
Number of traders per market (cassava/gari)	9.37	11.31	918
Number of traders per market (local rice)	9.05	9.90	918
Log distance (km) to nearest large town	2.60	1.40	78

Note: All prices are in SLL (Sierra Leonean Leones).

The relevant exchange rate was roughly $1=4300 SLL.

Total travel fare is for a motorbike ride along the entire distance of the road.

Outliers (top 1% of observations) are removed from price data.

All prices are in cups unless otherwise indicated.

Conversion rates are approximately 4 (5) cups of rice (cassava) per kg.

Table 3a: Pre-Rehabilitation Analysis: Score Components

	(1) Population per km	(2) Production per km	(3) Road Assessment	(4) Social Value	(5) Road Length
Preferred LLR (h=.15)					
Treatment	-0.537	0.342	-0.072	-1.008	15.694
	[0.409]	[0.328]	[2.264]	[1.599]	[17.433]
Mean for Control	5.977	17.863	17.143	8.857	23.199
Observations	31	31	31	31	31
LLR (h=.075)					
Treatment	-1.578***	0.006	2.897	-0.229	11.099
	[0.430]	[0.553]	[3.310]	[3.038]	[28.657]
Mean for Control	5.958	17.681	18.667	9.444	22.943
Observations	18	18	18	18	18
LLR (h=.3)					
Treatment	-0.213	0.232	1.641	-0.457	4.988
	[0.314]	[0.223]	[1.644]	[1.137]	[10.770]
Mean for Control	5.967	17.902	17.067	8.733	22.306
Observations	38	38	38	38	38
3rd Order Polynomial					
Treatment	-0.789	0.085	0.862	-1.065	24.238
	[0.530]	[0.488]	[2.929]	[2.356]	[23.118]
Mean for Control	5.935	17.889	16.750	8.750	21.074
Observations	45	45	45	45	45

Note: Road assessment is a 1-5 score assigned by EDA that measures the pre-existing condition of the selected road based on seven parameters (culverts, bridges, drainage, pavement surface, vertical/ horizontal alignment, and riding quality).

Social value is a 1-5 score assigned by EDA based on the conditions of schools, health centers, wells and toilets in the catchment area of the road.

LLR is short for local linear regression.

All specifications include district fixed effects.

Heteroskedasticity-robust standard errors reported in brackets.

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

Table 3b: Pre-Rehabilitation Analysis: Road Characteristics

	(1) Number of Bridges	(2) Number of Palmlog Bridges	(3) Number of Culverts	(4) Prop of HHs Crop Farming	(5) Prop of Indiv Crop Farming	(6) Literate HH Head Dummy	(7) HH Head Years of Education
Preferred LLR (h=.15)							
Treatment	2.801	-2.242	4.647	-0.036	0.003	0.040	0.284
	[5.263]	[4.466]	[11.778]	[0.097]	[0.125]	[0.059]	[0.582]
Mean for Control	3.857	4.552	15.430	0.880	0.794	0.239	2.056
Observations	31	31	31	31	31	31	31
LLR (h=.075)							
Treatment	0.578	-5.051	-9.375	0.100	0.190	-0.059	-0.249
	[11.257]	[4.421]	[9.420]	[0.136]	[0.172]	[0.067]	[0.694]
Mean for Control	4.778	3.525	14.336	0.890	0.790	0.240	2.008
Observations	18	18	18	18	18	18	18
LLR (h=.3)							
Treatment	0.130	-3.645	2.613	-0.003	0.001	0.033	0.212
	[3.499]	[3.427]	[8.724]	[0.070]	[0.090]	[0.039]	[0.398]
Mean for Control	3.800	4.248	15.735	0.886	0.804	0.238	2.035
Observations	38	38	38	38	38	38	38
3rd Order Polynomial							
Treatment	2.287	3.863	-0.263	-0.065	-0.001	0.038	0.640
	[7.704]	[6.487]	[15.785]	[0.149]	[0.185]	[0.090]	[0.880]
Mean for Control	3.563	3.983	15.001	0.890	0.810	0.237	2.021
Observations	45	45	45	45	45	45	45

Note: Columns (4)-(7) based on 2004 Census data.

Note: A culvert is a pipe or drain allowing water to flow under a road.

LLR is short for local linear regression.

All specifications include district fixed effects.

Heteroskedasticity-robust standard errors reported in brackets.

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

Table 3c: Pre-Rehabilitation Analysis: Farming Activity and Harvests

	(1) Log Market Distance to Big Town	(2) Chiefdom Frac of HHs Selling Rice	(3) Chiefdom Frac of HHs Selling Cassava	(4) Log Rice Harvest	(5) Log Cassava Harvest
Preferred LLR (h=.15)					
Treatment	-0.518	-0.031	-0.063	-0.005	0.166
	[0.341]	[0.027]	[0.061]	[0.168]	[0.112]
Mean for Control	9.707	0.218	0.389	6.197	7.307
Observations	104	65	65	65	65
LLR (h=.075)					
Treatment	-0.670	0.007	-0.256***	-0.197	0.060
	[0.623]	[0.054]	[0.067]	[0.219]	[0.280]
Mean for Control	9.857	0.223	0.405	6.202	7.253
Observations	64	41	41	41	41
LLR (h=.3)					
Treatment	-0.702*	-0.027	-0.016	-0.040	0.081
	[0.365]	[0.021]	[0.045]	[0.157]	[0.103]
Mean for Control	9.707	0.218	0.389	6.197	7.307
Observations	129	77	77	77	77
3rd Order Polynomial					
Treatment	-2.292**	-0.018	-0.165	-0.727**	-0.314
	[1.058]	[0.036]	[0.168]	[0.285]	[0.322]
Mean for Control	9.707	0.218	0.389	6.197	7.307
Observations	156	91	91	91	91

Note: All data is from the 2010 AHTS.

Distance data (km) is at the market-road level. Harvest data (kg) is at the chiefdom-road level.

LLR is short for local linear regression.

All specifications include road-district and market-district fixed effects.

Standard errors two-way clustered by road and market reported in brackets in column (1).

Standard errors two-way clustered by road and chiefdom reported in brackets in columns (2)-(5).

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

Table 4: First Stage Analysis (Transport Fares and Road Speed)

	Average Speed (kph)		Log Fare/km	
	(1)	(2)	(3)	(4)
Preferred LLR (h=.15)				
Treatment	12.085**	12.769**	-0.594**	-0.610***
	[5.345]	[5.282]	[0.236]	[0.203]
Mean for Control	26.196	26.196	6.729	6.729
Heterogeneity Controls		X		X
Observations	31	31	31	31
LLR (h=.3)				
Treatment	16.122***	16.350***	-0.149	-0.178
	[4.881]	[4.704]	[0.214]	[0.215]
Mean for Control	26.732	26.732	6.789	6.789
Heterogeneity Controls		X		X
Observations	38	37	38	37
3rd Order Polynomial				
Treatment	-4.274	-3.674	-0.836**	-0.980**
	[8.509]	[8.498]	[0.325]	[0.419]
Mean for Control	26.732	26.732	6.805	6.805
Heterogeneity Controls		X		X
Observations	44	43	45	44

Note: Fares are one-way motorbike fares in SLL (Sierra Leonean Leones).

The relevant exchange rate was roughly $1=4300 SLL.

Average speed measured in kilometers per hour along the entire road.

Heterogeneity controls include cassava/rice sales indicators/harvests
measured for EAs within 2km of the roads.

LLR is short for local linear regression.

All specifications include district fixed effects.

Heteroskedasticity-robust standard errors reported in brackets.

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

Table 5: Effects on Prices from Trader and High Frequency Surveys

	Log Local Rice Price		Log Local Rice Price		Log Cassava Price	
	(1)	(2)	(3)	(4)	(5)	(6)
Preferred LLR (h=.15)						
Treatment	-0.116**	-0.105***	0.006	0.015	-0.178**	-0.126**
	[0.058]	[0.029]	[0.029]	[0.017]	[0.086]	[0.056]
Mean for Control	6.831	6.831	6.889	6.889	5.703	5.703
Heterogeneity Controls		X		X		X
Observations	190	188	896	883	918	906
LLR (h=.075)						
Treatment	-0.178***	-0.174***	-0.030	-0.022	-0.227*	-0.237***
	[0.066]	[0.047]	[0.048]	[0.031]	[0.125]	[0.072]
Mean for Control	6.843	6.843	6.887	6.887	5.728	5.728
Heterogeneity Controls		X		X		X
Observations	124	122	599	586	595	583
LLR (h=.3)						
Treatment	-0.074*	-0.069**	0.005	0.003	-0.159*	-0.093*
	[0.045]	[0.031]	[0.021]	[0.022]	[0.089]	[0.050]
Mean for Control	6.831	6.831	6.889	6.889	5.703	5.703
Heterogeneity Controls		X		X		X
Observations	237	233	1099	1073	1112	1088
3rd Order Polynomial						
Treatment	-0.165***	-0.168***	-0.000	-0.053	-0.403**	-0.338***
	[0.058]	[0.044]	[0.055]	[0.057]	[0.170]	[0.068]
Mean for Control	6.831	6.831	6.889	6.889	5.703	5.703
Heterogeneity Controls		X		X		X
Observations	298	292	1314	1275	1338	1302

Note: Trader data (columns (1)-(2)) is annual; high frequency price data (columns (3)-(6)) is monthly.

All prices in SLL (Sierra Leonean Leones) per cup.

Heterogeneity controls include log distance to the nearest major town, and rice and cassava sales fractions and harvest amount medians.

LLR is short for local linear regression.

All specifications include round, road-district, and market-district fixed effects.

Standard errors two-way clustered by road and market reported in brackets.

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

Table 6: Heterogeneity by Access to Towns, Agricultural Production and Density of Rice Sellers Near Roads

	Distance		Harvest		Seller Density	
	(1)	(2)	(3)	(4)	(5)	(6)
	Log Rice Price	Log Cassava Price	Log Rice Price	Log Cassava Price	Log Rice Price	Log Cassava Price
Preferred LLR (h=.15)						
Treatment	0.085***	0.002	-0.022	-0.115	-0.040	-0.275**
	[0.024]	[0.033]	[0.036]	[0.071]	[0.040]	[0.108]
Treat * Above Median	-0.162***	-0.345***	0.065*	0.064	0.095**	0.236**
	[0.047]	[0.104]	[0.039]	[0.080]	[0.039]	[0.113]
One-sided p-value	0.000	0.000	0.048	0.212	0.008	0.018
Mean for Control	6.889	5.703	6.889	5.703	6.889	5.703
Observations	896	918	883	906	883	906
LLR (h=.075)						
Treatment	0.107***	0.051	-0.138**	-0.404**	-0.121*	-0.476***
	[0.040]	[0.068]	[0.070]	[0.178]	[0.072]	[0.115]
Treat * Above Median	-0.210***	-0.535***	0.178**	0.410**	0.176**	0.490***
	[0.073]	[0.129]	[0.076]	[0.188]	[0.078]	[0.132]
One-sided p-value	0.002	0.000	0.009	0.015	0.012	0.000
Mean for Control	6.887	5.728	6.887	5.728	6.887	5.728
Observations	599	595	586	583	586	583
LLR (h=.3)						
Treatment	0.072***	0.002	-0.008	-0.254	-0.007	-0.278***
	[0.022]	[0.023]	[0.036]	[0.185]	[0.034]	[0.102]
Treat * Above Median	-0.128***	-0.186	0.019	0.219	0.040	0.308***
	[0.037]	[0.136]	[0.040]	[0.191]	[0.037]	[0.117]
One-sided p-value	0.000	0.086	0.318	0.126	0.136	0.004
Mean for Control	6.889	5.703	6.889	5.703	6.889	5.703
Observations	1099	1112	1073	1088	1073	1088
3rd Order Polynomial						
Treatment	0.135**	0.107	-0.053	-0.362***	-0.138***	-0.584***
	[0.055]	[0.075]	[0.065]	[0.095]	[0.033]	[0.101]
Treat * Above Median	-0.242***	-0.552***	0.065	0.420***	0.205***	0.556***
	[0.092]	[0.089]	[0.063]	[0.134]	[0.049]	[0.123]
One-sided p-value	0.004	0.000	0.151	0.001	0.000	0.000
Mean for Control	6.889	5.703	6.889	5.703	6.889	5.703
Observations	1314	1338	1275	1302	1275	1302

Note: All data is from high frequency price surveys.

All prices in SLL (Sierra Leonean Leones) per cup.

LLR is short for local linear regression.

All specifications include round, road-district, and market-district fixed effects.

One sided p-values are calculated on the interaction term (>0 for distance, <0 for harvest and seller density).

Heteroskedasticity-robust standard errors for all narrow bandwidth (h=.075) regressions. Standard errors two-way clustered by road and market for cassava price regressions, except the h=.075 specification.

For rice price (distance heterogeneity) regressions, we use heteroskedasticity-robust standard errors as there are too few clusters to allow for two-way clustering. Standard one-way clustering at either the road level or market level produces standard errors that are uniformly smaller than those reported.

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

Table 7: Robustness Checks: Main Effects on Prices

	Log Local Rice Price		Log Local Rice Price		Log Cassava Price	
	(1) Basic	(2) Het Controls	(3) Basic	(4) Het Controls	(5) Basic	(6) Het Controls
Panel A: Alternative Specifications						
Paper Spec: 2nd Poly						
Treatment	-0.140**	-0.120***	0.005	-0.020	-0.218	-0.113
	[0.066]	[0.044]	[0.036]	[0.039]	[0.135]	[0.074]
Observations	298	292	1314	1275	1338	1302
Triangle Weight						
Treatment	-0.082	-0.074***	0.027	0.033**	-0.079	-0.077**
	[0.050]	[0.028]	[0.032]	[0.016]	[0.067]	[0.036]
Observations	190	188	896	883	918	906
Panel B: Alternative Samples						
All Roads						
Treatment	-0.100**	-0.073**	0.015	0.024	-0.169***	-0.095*
	[0.048]	[0.031]	[0.023]	[0.016]	[0.063]	[0.050]
Observations	203	201	970	957	995	983
Closest Market (unw)						
Treatment	-0.123*	-0.179**	0.030	0.027	-0.179**	-0.161***
	[0.074]	[0.073]	[0.026]	[0.027]	[0.077]	[0.061]
Observations	56	56	282	282	284	284
Markets w/in 5km (unw)						
Treatment	-0.085	-0.063*	0.032	0.018	-0.122*	-0.080
	[0.061]	[0.033]	[0.028]	[0.027]	[0.073]	[0.061]
Observations	82	82	438	438	459	459
Drop Town Markets						
Treatment	-0.131**	-0.127***	-0.002	-0.000	-0.198**	-0.172***
	[0.056]	[0.022]	[0.030]	[0.018]	[0.091]	[0.052]
Observations	163	161	733	720	753	741
Drop Multimatch (>4) Markets						
Treatment	-0.131**	-0.127***	0.001	0.010	-0.168*	-0.132**
	[0.056]	[0.022]	[0.029]	[0.016]	[0.092]	[0.058]
Observations	163	161	778	765	780	768
Panel C: Small Sample Corrections						
One-way Cluster (Road)						
Treatment	-0.116**	-0.105***	0.006	0.015	-0.178**	-0.126**
	[0.051]	[0.027]	[0.028]	[0.018]	[0.080]	[0.053]
Wild Bootstrap Coeff 95% CI	-0.21, -0.04	-0.15, -0.06	-0.05, 0.06	-0.02, 0.05	-0.33, -0.03	-0.23, -0.02
Observations	190	188	896	883	918	906

Note: All specifications are local linear regressions at the preferred (h=.15) bandwidth unless otherwise indicated.
Wild bootstrap coefficient confidence intervals based on 400 iterations.

Table 8: Robustness Checks: Heterogeneity

	Distance		Harvest		Seller Density	
	(1) Log Rice Price	(2) Log Cassava Price	(3) Log Rice Price	(4) Log Cassava Price	(5) Log Rice Price	(6) Log Cassava Price
Panel A: Alternative Specifications						
Paper Spec: 2nd Poly						
Treatment	0.087***	0.061*	-0.054	-0.422	-0.064	-0.438***
	[0.033]	[0.036]	[0.049]	[0.273]	[0.052]	[0.100]
Treat * Above Median	-0.156***	-0.369**	0.088	0.466*	0.101*	0.564***
	[0.057]	[0.149]	[0.056]	[0.277]	[0.060]	[0.126]
One-sided p-value	0.003	0.007	0.057	0.046	0.047	0.000
Observations	1314	1338	1275	1302	1275	1302
Triangle Weight						
Treatment	0.089***	0.018	-0.012	-0.042	-0.036	-0.180**
	[0.019]	[0.023]	[0.028]	[0.043]	[0.038]	[0.086]
Treat * Above Median	-0.160***	-0.266***	0.073**	0.019	0.101**	0.173*
	[0.042]	[0.086]	[0.031]	[0.045]	[0.039]	[0.089]
One-sided p-value	0.000	0.001	0.010	0.335	0.005	0.025
Observations	896	918	883	906	883	906
Continuous Het Vars						
Treatment	0.183	1.009**	0.290	-0.982***	-0.023	-0.390**
	[0.124]	[0.497]	[0.221]	[0.306]	[0.056]	[0.178]
Treatment Interaction	-0.019	-0.121**	-0.047	0.123***	0.085	0.505*
	[0.014]	[0.056]	[0.035]	[0.035]	[0.165]	[0.287]
One-sided p-value	0.087	0.015	0.912	0.000	0.302	0.039
Mean of Het Variable	9.390	9.373	6.207	7.244	0.230	0.408
Observations	896	918	883	906	883	906
Panel B: Alternative Samples						
All Roads						
Treatment	0.070***	-0.038	-0.003	-0.075	-0.031	-0.277***
	[0.017]	[0.029]	[0.041]	[0.059]	[0.042]	[0.105]
Treat * Above Median	-0.148***	-0.304***	0.023	0.007	0.080*	0.189*
	[0.044]	[0.099]	[0.045]	[0.070]	[0.042]	[0.108]
One-sided p-value	0.000	0.001	0.307	0.460	0.030	0.041
Observations	970	995	957	983	957	983
Closest Market (unw)						
Treatment	0.066**	-0.075**	0.017	-0.242	0.020	-0.245**
	[0.033]	[0.038]	[0.045]	[0.315]	[0.062]	[0.103]
Treat * Above Median	-0.096	-0.252**	0.035	0.139	0.023	0.181*
	[0.059]	[0.121]	[0.063]	[0.322]	[0.065]	[0.109]
One-sided p-value	0.053	0.019	0.290	0.333	0.360	0.049
Observations	282	284	282	284	282	284
Markets w/in 5km (unw)						
Treatment	0.080***	0.056**	-0.068***	-0.139**	-0.046	-0.265***
	[0.029]	[0.025]	[0.018]	[0.069]	[0.029]	[0.103]
Treat * Above Median	-0.183***	-0.369***	0.154***	0.126*	0.135***	0.220**
	[0.050]	[0.118]	[0.018]	[0.070]	[0.031]	[0.105]
One-sided p-value	0.000	0.001	0.000	0.036	0.000	0.018
Observations	438	459	438	459	438	459

Drop Town Markets

Treatment	0.088***	0.009	-0.028	-0.196***	-0.042	-0.312***
	[0.031]	[0.039]	[0.038]	[0.071]	[0.042]	[0.100]
Treat * Above Median	-0.165***	-0.352***	0.054	0.136*	0.084**	0.273***
	[0.051]	[0.108]	[0.036]	[0.076]	[0.039]	[0.105]
One-sided p-value	0.001	0.001	0.066	0.037	0.016	0.005
Observations	733	753	720	741	720	741

Drop Multimatch (>4) Markets

Treatment	0.085***	0.002	-0.019	-0.115	-0.038	-0.279**
	[0.024]	[0.033]	[0.037]	[0.070]	[0.040]	[0.111]
Treat * Above Median	-0.177***	-0.431***	0.053	0.088	0.082**	0.261**
	[0.054]	[0.066]	[0.046]	[0.078]	[0.041]	[0.114]
One-sided p-value	0.001	0.000	0.124	0.131	0.022	0.011
Observations	778	780	765	768	765	768

Panel C: Small Sample Corrections

One-way Cluster (Road)

Treatment	0.085***	0.002	-0.022	-0.115	-0.040	-0.275***
	[0.024]	[0.036]	[0.034]	[0.072]	[0.040]	[0.095]
Treat * Above Median	-0.162***	-0.345***	0.065*	0.064	0.095**	0.236**
	[0.047]	[0.088]	[0.038]	[0.086]	[0.042]	[0.102]
One-sided p-value	0.000	0.000	0.045	0.228	0.012	0.010
Wild Bootstrap Coeff 95% CI (Lower)	-0.21, -0.12	-0.51, -0.19	-0.01, 0.14	-0.11, 0.23	0.02, 0.17	0.05, 0.41
Observations	896	918	883	906	883	906

Note: All specifications are local linear regressions at the preferred (h=.15) bandwidth unless otherwise indicated.

  Two-way clustered standard errors except rice price/distance heterogeneity regressions and one-way cluster spec.

  All one-sided p-values are for the interaction term (>0 for distance, <0 for harvest and seller density).

  Wild bootstrap interaction coefficient confidence intervals based on 400 iterations.

The International Growth Centre
(IGC) aims to promote sustainable
growth in developing countries
by providing demand-led policy
advice based on frontier research.

Find out more about
our work on our website
www.theigc.org

For media or communications
enquiries, please contact
mail@theigc.org

Subscribe to our newsletter
and topic updates
www.theigc.org/newsletter

Follow us on Twitter
@the_igc

Contact us
International Growth Centre,
London School of Economic
and Political Science,
Houghton Street,
London WC2A 2AE

IGC

**International
Growth Centre**