Isolation and agricultural productivity

David Stifel (Lafayette College)*
Bart Minten (IFPRI)

Abstract: This paper examines the mechanisms that transmit isolation into poverty. In particular, we study the effect of isolation and transport infrastructure on welfare and agricultural productivity in the case of Madagascar. Madagascar is a good case study given the bad shape of its infrastructure and therefore the significant variation in isolation. Based on comprehensive household survey data combined with a census of communes, we discover a strong poverty-isolation relationship. Further we find the inverse relationship between agricultural productivity and isolation to be surprisingly strong. We isolate the following reasons why productivity might decline with isolation: (a) transportation-induced transaction costs, (b) the inverse relationship between plot-size and productivity, (c) increasing price variability and extensification onto less fertile land, and (d) insecurity.

December 2007

* The authors express their gratitude to USAID for providing financial support for this research, to the Institut National de la Statistique (INSTAT) in Madagascar for graciously providing the household survey data, and to Paul Dorosh for making this work possible and for insightful comments. They also want to thank three anonymous referees, Chris Barrett, Dan Gilligan, Jean-Claude Randrianarisoa, John Strauss and participants at seminars held at Cornell University, IFPRI, SUNY Binghamton and the World Bank for valuable comments on an earlier version of this paper, and Lalaina Randrianarison for assistance creating the maps.
Isolation and Agricultural Productivity

I. Introduction:

Geography matters enormously for economic activities and welfare as shown by evidence from cross-country studies (Bloom and Sachs, 1998; Landes, 1998; Gallup et al., 1999) as well as within-country studies (Hentschel et al., 2000; Jalan and Ravallion, 2002; Ravallion and Datt, 2002). The large impact is attributed mainly to differences in market access, access to natural resources, incidences of infectious diseases and effectiveness of governance. For example, it has been argued that Western Europe developed much faster due to its climatic advantages over other parts of the world (Landes, 1998) while the poor economic performance of African economies has largely been blamed on geographic characteristics (Bloom and Sachs, 1998). However, while the effect of geography is widely acknowledged, most of the literature fails to study the types of policies that might be implemented to mitigate adverse geographical conditions.

In this study, we look at this issue in more depth by analyzing the effect of isolation and transport infrastructure on agricultural productivity in Madagascar. Madagascar is a good case study given the bad shape of its infrastructure and therefore the significant variation in isolation. Based on comprehensive data, we discover a strong poverty-isolation relationship. Further we find the inverse relationship between agricultural productivity and isolation to be surprisingly strong. Considering that that eight out of ten people live in rural areas, and that nine out of ten rural poor persons live in farming households (Razafindravonona et al., 2001), understanding the nature of agricultural production and the potential policies that may be employed to improve it is
critical to understanding and addressing rural poverty. The objective of this paper is to do just that by analyzing the specific role of transport infrastructure on agricultural productivity. Although the analysis is limited to this particular relationship, it provides insights into the connection between infrastructure and poverty by extension given the link between agriculture and poverty in Madagascar (Randrianarisoa and Minten, 2001; Razafindravonona et al., 2001; World Bank, 2001; Jacoby and Minten, 2007).

An outline of the remainder of this paper is as follows. In section II, we examine the potential sources of the inverse isolation-productivity relationship. In Section III, we discuss the data and the methodology used to define isolation, while an assessment of the meaning and implications of isolation are described in section IV. We examine the evidence regarding the isolation-productivity relationship for Madagascar in section V, and wrap up with concluding remarks in section VI.

II. Productivity and Isolation

There are various avenues through which isolation can affect agricultural productivity. Those that we address in this study in particular are (1) transportation-induced transaction costs, (2) the inverse relationship between plot-size and productivity, (3) increasing price variability and extensification onto less fertile land, and (4)

---

1 Antle (1983) also examines the relationship between infrastructure and agricultural productivity. But his is a cross-country study of aggregate productivity, whereas our analysis analyzes plot-level productivity.
2 This is much in the spirit of T.W. Shultz’ (1980) opening remarks in his Nobel Prize acceptance speech: “Most of the people in the world are poor, so if we knew the economics of being poor we would know the economics that really matters. Most of the world’s poor people earn their living from agriculture, so if we knew the economics of agriculture we would know much of the economics of being poor.”
3 This issue is not new. For example, Howe and Richards (1985), and Jacoby (2000) examined the extent to which rural road construction – one particular policy intervention – affects rural income inequality, not poverty. This interest has followed, in large part, because of the significance that inequality has on the political constraints related to the allocation of infrastructure investment. While van de Walle (2002) and
insecurity. While other dynamic influences, such as externalities associated with location
effects as discussed in the geographic poverty trap literature (Jalan and Ravallion, 2002;
Deininger and Okidi, 2003; Dercon, 2003; De Vreyer et al., 2003), and induced
innovation (Hayami and Ruttan, 1985), are also likely to be pertinent, a lack of access to
appropriate panel data prevent us from addressing them empirically.

First, transportation-induced transaction costs\(^4\) can generally influence
productivity in two ways. On the one hand, they can alter relative prices in such a
manner that input use is affected. On the other hand, such price differences can also
affect the crop choice. With regard to input use, to the extent that transaction costs drive
a wedge between the value marginal product of the inputs and the market prices of these
inputs (e.g. cost of fertilizer at the factory gate relative to the price paid by the farmer),
input use per unit of land is likely to fall, and consequently so are yields (Jacoby and
Minten, 2007). Indeed, if the bands between the purchase and sale prices due to
transportation costs are wide enough, previously marketed inputs may even become non-
tradables (Sadoulet and de Janvry, 1995). Further, as Minten and Kyle (1999) found in
the former Zaire, traders marketing margins can increase with isolation, implying that
such price bands may even increase at an increasing rate the farther one gets from a
market center.

The presence of transportation-induced transaction costs can also explain the
seemingly inefficient cropping choices of farmers where greater resources are devoted to
low-yielding food crops instead of to cash crops that have higher market returns (see

---

Gibson and Rozelle (2003) study the relationship between road access and poverty, we concentrate
specifically on isolation and its effect on agricultural productivity.
Omamo, 1998, for evidence on Kenya). This is consistent with models that predict farming households shifting out of perishable cash crops (e.g. vegetables and fruits) into heartier and storable crops such as staples and pulses as they get farther from the market centers (von Thünen, 1966).

While the observed inverse relationship between agricultural productivity and isolation is not new, clarifying this link between isolation-related transaction costs associated and agricultural production will shed further light on avenues through which policy interventions intended to alleviate rural poverty may be directed. Indeed, there is the further possibility of not only equity, but efficiency gains to be had from policies designed to reduce transaction costs (Dercon, 2003).

Nonetheless, there are competing explanations as to why agricultural productivity falls with remoteness. This brings us to our second avenue. Given the well-established inverse relationship between plot size and productivity (Carter, 1984; Feder, 1985; Bhalla and Roy, 1988; Barrett, 1996a; Heltberg, 1998), the increase in the median plot size by isolation quintile observed in our data may partially explain the differences in productivity. The policy implications of this particular observation follow from the source of the inverse relationship whether it is due to labor market failures (Carter, 1984), unobserved soil quality differences (Bhalla, 1988; Bhalla and Roy, 1988; Benjamin, 1995), land market failures or financial market failures (Barrett, 1996a). As the literature

---

4 We define transaction costs as any costs that drive a wedge between buyer and seller prices (e.g. farmgate and market prices). As such these costs include transportation costs as well as marketing margins of intermediary traders.

5 For example, Binswanger et al. (1993) found that in India in the 1970s improved roads contributed directly to growth of agricultural output, as well as fertilizer use (see also Ahmed and Hossain, 1990, for Bangladesh), due to “reduced transaction costs of all sorts.”

6 There is evidence of the inverse relationship between plot size and productivity in other studies in Madagascar (Barrett, 1996a; Randrianarisoa, 2001; Randrianarisoa and Minten, 2001).
is far from a consensus as to these sources, the policy implications of this relationship are beyond the scope of this paper.

Third, greater seasonal price variability in more isolated areas (Minten and Randrianarison, 2003) may affect productivity. Agricultural output prices in Madagascar are generally lowest between the months of April and June, a period that corresponds to the harvest when farmers sell much of their crop. Due to liquidity constraints and financial market failures (Barrett, 1996b; Barrett and Dorosh, 1996), many farming households are pressed into using the rice market as a pseudo capital market by selling their rice output after the harvest to pay for inputs, and by buying back all or part of it at higher prices during the lean period (January-March). With rice stored in more central locations, prices in the most remote quintile are lowest during the post-harvest period when rice is sold, and are highest during the lean period when it is bought back. Further, low road quality increases likelihood of interruptions in information and material flows, which tends to increase price variability.  

As higher price risk is equated with higher income risk, poorer and more remote farmers can be expected to accept lower incomes for lower income variance (Rosenzweig and Binswanger, 1993). In the presence of high seasonal price variability, households have the incentive to self insure through agricultural extensification (Barrett, 1999).  

In other words, as an insurance mechanism, these farming households might increase output by expanding production into less fertile land and relying less on inputs and investments to increase land productivity.

---

7 This is consistent with Barrett’s (1996b) finding.
8 Households may also adopt extensification because the costs of expanding production in this manner are lower than the costs with intensification. We thank an anonymous referee for pointing this out.
Finally, physical insecurity also has a potential negative effect on agricultural production. Insecurity in rural Madagascar includes significant theft of standing crop and cattle raiding. The effect is to limit the range over which farmers are willing to cultivate crops and/or to graze their livestock. In so far as this limits the use of more fertile but distant pasture and cultivable soils, the consequence is reduced productivity. In addition, labor productivity\(^9\) and allocation for productive purposes are negatively affected by devoting labor resources to protecting crops and livestock (Freudenberger, 1998). Finally, Minten and Randrianarison (2003) show how wages for agricultural labor are significantly higher in insecure areas as laborers have to be compensated for the riskier living conditions. This leads to relatively less use of hired-in labor and thus lower production.

**III. Data and Measures of Isolation**

Two data sources are used in this analysis: a household survey and a national census of administrative communes. The 2001 *Enquete Prioritaire Aupres des Ménages* (EPM) was a nationally representative integrated household survey of 5,080 households. The data were collected during the months of September, October and November 2001. The sample was selected through a multi-stage sampling technique in which the strata were defined by the *faritany* (province) and *milieu* (rural, secondary urban centers, and primary urban centers), and the primary sampling units (PSU) were *fokontany*.\(^{10}\) Each of the *fokontany* was selected systematically with probability proportional to size (PPS), and

---

\(^9\) Freudenberger (1998) describes the exhausted family members in a Betsileo village whose productivity is diminished because they are forced to watch over their livestock all night.

\(^{10}\) There are 17,433 *fokontany* in Madagascar.
sampling weights defined by the inverse probability of selection to obtain accurate population estimates.

The comprehensive household questionnaire includes sections on education, health, employment, housing, agriculture, non-agricultural enterprises, and household expenditures and assets. The agriculture section is particularly detailed for a nationally representative survey with plot- and crop-level information. For a measure of household well-being, in this analysis we use the estimated household-level consumption aggregate constructed by the Insitut National de la Statistique (INSTAT) and the World Bank (Rakotomahefa et al., 2002).

At the same time that the national household survey was in the field, a commune census was conducted in a collaborative effort between the Ilo program of Cornell University\(^\text{11}\), the national Malagasy agricultural research institute (FOFIFA), and INSTAT. The remoteness of some communes meant that little was known about the spatial distribution of goods and services and economic activity prior to this study. This census gathered information on health centers, educational enrollments, commune budgets, and crime figures from the relevant government offices in the commune. More subjective questions, such as those concerning local prices, transportation, access to various goods and services, major economic activities, and community perceptions of existing conditions were answered for each commune by a focus group representative of the composition of the population of the commune. The survey was conducted at the commune's administrative center. A total of 1385 out of 1392 communes were visited.

\(^{11}\) The Ilo program (http://www.ilo.cornell.edu/) was a USAID-funded program for “Improved Policy Analysis for Economic Decision-Making and Improved Public Information and Dialog.” Ilo (pronounced “ee-lew”) is a Malagasy word that translates roughly as (a) to enlighten and (b) to facilitate the joining of pieces together through lubrication.
Finally, commune-level information from this census was merged with the 163 rural communities that appear in the EPM data.

Rural isolation can be defined in many ways. Distance to urban centers or markets is one commonly used measure (McCabe, 1977; Ahmed and Hossain, 1990; Minten and Kyle, 1999; Jacoby, 2000; Fafchamps and Moser, 2003). In this paper we attempt to capture isolation in Madagascar through two measures: (a) travel time to the nearest primary urban center, and (b) cost of transporting a 50 kg sack of rice to the nearest primary urban center to capture the transaction cost component of isolation.

Travel time to the nearest primary urban center is more precisely the dry season travel time from the commune center to the nearest large city to which commune residents actually travel on a regular basis. This information was collected in the commune census, and was determined in steps. For instance, if multiple forms of travel (e.g. foot, ox-cart, automobile) are necessary, then actual travel time per form of transport was recorded along with waiting time.

The cost of transporting a 50 kg sack of rice was also collected in the commune census, and is the cost of dry season transport between the commune center and the nearest primary urban center to which residents travel on a regular basis. As we shall see, the distance to the urban center and the average travel time per kilometer (interpreted loosely as road quality) each has an independent effect on the cost of transportation. In the econometric estimation in section VI, we use transportation costs as our preferred proxy for isolation because in addition to capturing distance, it also captures road quality and traders’ marketing margins. As such, it provides a framework in which we can
understand the impact of policies designed to reduce these costs, whether it be through road improvements or through price subsidies.

IV. Isolation: Meaning and Implications

Before examining the implications of being isolated, let us characterize isolation itself for households not situated in major urban areas. The disparities in distance to major cities and markets that exist in Madagascar are apparent in Table 1. Using travel time as our measure of isolation, we see that those in the most isolated 20 percent of the rural population must travel 39 hours on average to reach the nearest major city. This is some 39 times longer than it takes for those living in the least isolated quintile of the rural population (see Figure 1 for a map of the communes by travel time). The typical journey for those living in the more isolated areas involves multiple legs starting with an extended walk to a taxi brousse station, and substantial waiting periods.

An additional consideration in the calculus of transaction costs includes the distance from a household’s agricultural plot to the nearest road accessible to animal-drawn carts. The second row of Table 1, illustrates that in addition to road travel time to markets, farmers in more isolated areas must also contend with the basic task of getting the harvest to the road. The average walking time for the most isolated households is over half an hour, whereas plots for those in the least isolated 40 percent of the population are less than 18 minutes from the road.

The cost of transportation relative to local market prices can be viewed as a very rough measure of the degree to which prices must be forgone in order to sell the output in the cities (i.e. transaction costs). The average ratio of the cost of transporting a 50
kilogram sack of rice to the nearest major city during the dry season to the price of 50 kilograms of rice is illustrated in Table 1 by quintile. In the least isolated areas, transportation costs account for approximately three percent of the price of rice, whereas in the most isolated areas these costs take up more than a third of the price. In rural areas in general, the cost of transportation average 17 percent of the price of rice.

A very simple regression model helps to illustrate the correlates of transportation costs (Table 2). The explanatory variables in this model are limited to distance, time in minutes per kilometer traveled, and the number of cattle thefts in the commune. Provincial dummies are also included to capture unobserved heterogeneity at the larger administrative level. The general idea is that costs are a function of the distance, the quality of the road and some risk premium that must be paid for transporting commodities through insecure regions. Although the latter turns out to be insignificant here, distance and road quality (as captured by time per kilometer) are both positive and strongly significant. Since the variables are in logs, the parameter estimates represent estimates of the elasticities. This facilitates a simple exercise such as simulating the partial effect of road improvements or construction that halves the average time it takes to travel one kilometer from 4.55 to 2.28 minutes. Since the road quality elasticity of transportation cost is 0.57, a 50 percent improvement in quality will lead to a 28 percent drop in the cost of transporting rice ($\Delta cost = 0.57 \times \Delta time = 0.57 \times -0.5$), resulting in an average decline of FMG 3,444.

Turning to the implications of remoteness for Malagasy households, we find a strong negative (positive) correlation between household consumption (poverty) and

---

12 These results are qualitatively and statistically no different from another model estimated using Fivondronana (smaller administrative regions) fixed effects.
isolation. Mean household per capita consumption (Table 1) in the most isolated quintile is less than half of that in the least isolated quintile (FMG 638 thousand ($98) versus FMG 1,499 thousand ($230)). Further, the largest gap is between the least isolated and the second quintile where the mean per capita consumption level is FMG 734 thousand ($113). A similar pattern is reflected in the poverty figures, where more than 85 percent of the individuals living in isolated areas are estimated to be poor relative to approximately 55 percent in the least isolated rural areas. 13 Again, the biggest jump is between the first and second quintiles – some 77 percent of the individuals in the second quintile are estimated to be poor. 14

The last row of Table 1 shows the mean shares of food consumption that derive from home production instead of purchases. While autoconsumption is positively correlated with poverty (correlation coefficient of 0.33, see also Razafindravonona et al, 2001), it also represents a measure of market development. For example, without the presence of markets for food, the 53 percent of the individuals in the least isolated quintile 15 who reported no autoconsumption of food would have had to grow their own food. As households are situated in increasingly isolated areas, the share of autoconsumption in total food consumption rises from under 20 percent to over 40 percent suggesting increasingly fragmented or weak markets.

Given the large share of rural households involved primarily in agriculture 16, the next logical step is to examine the relationship between isolation and agricultural

---

13 The poverty line is defined by Rakotomahefa et al. (2002).
14 Similar effects of isolation are found by Jacoby and Minten (2007) in a small region of Madagascar with an enormous transport gradient. They find that a hypothetical road project that would reduce transportation costs by $75/ton would increase the income of rural households by nearly half. One-third of the raise in income is explained by an increase in agricultural income.
15 Note that 37 percent of these individuals are estimated to be poor.
16 83 percent of households report at least one member whose primary employment activity is agriculture.
production. Table 3 shows the surprising degree to which we find that rice, maize and cassava yields differ by extent of isolation. For example, median rice yields drop from above 25 kilograms per are for the first two quintiles to less than 19 kilograms in the most isolated two quintiles. Similar results are found for maize and cassava production where median yields in the most isolated areas are approximately 50 percent below those of the least isolated quintile.

With the competing demands for land and the subsequent higher land values in the least remote areas, it is not entirely surprising that yields should be greater there (Hayami and Ruttan, 1985). Nonetheless, the results observed here are surprising in large part because, while we might expect output to vary substantially and yields to vary moderately in response to changes in transaction costs, it is not clear ex ante that the observed magnitudes of declining productivity associated with increased isolation should be so large. This is an important finding in that it sheds light on an avenue through which policies may be developed to raise the incomes of the poorest sector of the Malagasy economy (Razafindravonona et al., 2001; World Bank 2001).

V. Productivity and Isolation: The Evidence

To better understand the potential policy alternatives, we now review what the data indicate about the four factors that may plausibly contribute to the pattern of declining rice yields.\(^{17}\) As a reminder, these are (a) transportation-induced transaction costs, (b) the inverse relationship between plot-size and productivity, (c) increasing price variability and extensification onto less fertile land, and (d) insecurity. We use three steps

\(^{17}\) We concentrate solely on rice for the remainder of the paper because it is the predominant crop, with 65% of cultivated land allocated to it.
in an attempt to isolate these factors. First, to examine the net effects, we estimate a reduced form plot-level yield function for rice. Second, we estimate separate reduced-form input demand functions to assess the effects of these factors on input use. Finally, we estimate a translog production (yield) function\(^\text{18}\) to study the effect of these inputs on rice production.

Each of these models is estimated using the merged commune-census and household-survey data. The dependent variable in the reduced-form and structural yield models is the log of the yield (kg/are) of principal season rice harvested for each plot. The exogenous explanatory variables in the reduced form yield and input models include plot characteristics, commune land characteristics, weather shocks, household characteristics, measures of institutional constraints in the commune, and province dummies. The explanatory variables in the production function include inputs (land, labor, animal and mechanical traction, and fertilizer) and such shifters as plot characteristics, commune land characteristics, and weather shocks.

Given that the choice variable inputs in the production function are likely correlated with unobserved ability, and as such are correlated with the error term, we estimate the model using instrumental variable (IV) methods. The identifying instruments in the first stage models are the set of exogenous variables that appear in the reduced form input demand models.

The measure of isolation used in these models is the log of the cost of transporting a 50 kilogram sack of rice to the nearest major city, as well as the distance (measured in

\(^{18}\) Given lack of variability in the commune-level input and output price data, as well as the degree to which these prices are imperfect substitutes for the shadow prices of inputs, we prefer the estimation of the primal production (yield) function to the dual profit function. A more detailed description of the model and estimation strategy is provided in the Appendix.
minutes) from the plot to the nearest passable route. One naturally questions whether road placement, and hence isolation, is exogenous. Based on the premise that cities formed around the most fertile land\(^{19}\) (Krugman, 1999; Fujita et al., 2001), one might expect soil/plot characteristics to be at least correlated with isolation. One would also suspect that roads were built in rural areas to access more fertile lands to maximize the availability of agricultural surplus for the urban centers. This is an important issue that hampers the impact of infrastructure analysis in general (Binswanger et al., 1993).

However, we are not terribly concerned about this issue in the case of Madagascar because Berg (1981, 1985, 1988), in a series of papers on the history of the Merina people, lays out convincing evidence that the physical infrastructure was designed not for commercial objectives, but for military ones. Nonetheless, there remains the possibility that road maintenance is correlated with plot characteristics. As such, we include soil characteristics at both the plot and commune level as control variables in the econometric estimates of the determinants of rice yields. We therefore believe that road placement and productivity can cautiously be treated as orthogonal in our estimates.

The output from the three sets of models appears in Tables 4 through 6. Table 4 presents estimates from the reduced form plot-level rice yield models, while the estimates of the separate reduced-form input demand functions appear in Table 5. Finally, the results of the translog production (yield) function are presented in the form of elasticities by remoteness quintile in Table 6.

Three reduced-form rice yield models were estimated (Table 4). In the first, only the isolation measures are included to measure the net effect of remoteness. In this

\(^{19}\) Trading centers and agglomeration effects are clearly important determining factors for the establishment of cities as well.
model, transportation costs have a significantly negative effect on rice yields as expected with an elasticity of -0.17. When soil quality and other control variables are included (Model 2), the elasticity of transportation cost remains significant but falls to -0.07, indicating the importance of controlling for plot characteristics when estimating the effect of isolation on productivity. Finally, plot size, price variability and insecurity variables are introduced in the third model. The fact that the statistically significant negative effect of transportation costs on rice yields falls further to an elasticity of -0.05, suggests that part of the negative isolation-agricultural productivity relationship is explained by these other factors. Not all of it is, however. As such we now turn to each of these factors.

Transportation-Induced Transaction Costs

As discussed in Section II, transaction costs that follow from isolation drive a wedge between farmgate and market prices. We expect this to have two effects on agricultural production. First, input choice and consequently productivity are affected, and second, crop choice will hinge on transportation-induced price changes. With regard to input use, Table 3 shows how fertilizer and pesticide use decline substantially with isolation consistent with our expectations. While overall chemical fertilizer use is very low, with fewer than 12 percent of all rural farming households applying it, it is even more so in the two most isolated quintiles, where three percent use it and on average a tenth of a kilogram is applied per are. This is in sharp contrast to the more than 28 percent of rural farming households in the two least isolated quintiles\(^{20}\) who apply chemical fertilizers.

\(^{20}\) In terms of time to the nearest city, those households within 4.5 hours travel time are in the first two quintiles.
The timing of fertilizer delivery is important for its application, which could also explain the low use observed in isolated areas. Delivery of chemical fertilizers tends to be in bulk, and if deliveries are late due to poor road quality, application is not optimal. Consequently, farmers may opt out of fertilizer use especially given that road access deteriorates during the rainy season which is also the planting season when fertilizer is to be applied.

The differences are even more stark for organic fertilizer. Some 66 percent of farming households in the least isolated quintile apply organic fertilizer depending on the isolation measure used, whereas less than 2 percent in the most isolated quintile use it. In terms of quantities, less than a fifth of a kilogram of organic fertilizer is applied per are in the most isolated areas, while over 7.5 kilograms are applied on average per are in least isolated areas.\(^2\)

With regard to labor inputs, the bulk of which consists of household labor, there appears to be an inverted-U relationship between man-days per are and isolation (Table 3). First, the least remote areas might use relatively more capital than labor in their agricultural activities as agricultural equipment can be acquired at lower costs and as agricultural wages are higher in these areas (Minten and Randrianarison, 2003). Second, the lower labor inputs associated with the least remote areas may also be a result of greater access to non-farm income earning opportunities. Whereas 32 percent of households in the least remote quintile have at least one member involved in a non-farm

\(^2\) The pattern of declining organic fertilizer use does not appear to directly follow from lack of access to excreta. Livestock ownership is generally no less prevalent in more isolated areas. In fact, the average number of heads owned and the average total value increase with isolation. We should note, however, that the largest herds are generally associated with extensive grazing where dung is deposited on pastures, not on fields. Thus, biophysical availability of excreta does not necessarily translate into economic availability. We thank Chris Barrett for making this point.
enterprise, only 13 percent of households in the most remote areas have access to alternative income sources. As remunerative non-farm opportunities are relatively more available in the least remote areas, this sets a higher opportunity wage for the agricultural sector discouraging agricultural extensification (expansion of output at the extensive margin into lower quality soils) and keeping yields higher.

The econometric results are consistent with the hypothesis that transaction costs associated with isolation affect yields through the use (or non-use) of inputs. The reduced-form input demand model estimates (Table 5) indicate that less household labor (column 1) and fertilizer (column 5) are used per are of land as the cost of transportation increase. Further, lack of access to passable routes from the plot has an additional negative effect on fertilizer use.

The input elasticities from the production function estimates (Table 6) suggest that the most important mechanism transmitting the effect of isolation to yields through transaction costs is the use of fertilizer. Interestingly, household labor inputs are not found to have a significant effect (as is the case with animal traction), while fertilizer inputs have a positive and significant rice yield elasticity of 0.05 on average. Further, with the exception of land inputs to which we will return briefly, none of the input elasticities differ across remoteness quintiles. This suggests that any effect of remoteness on the input-yield relationship is due to the quantity of inputs used, not the quality.

The next question is: “Does the choice of crop differ with isolation?” The von Thünen (1966) model suggests that it should. As farmers grow their crops further away

---

22 Note that since current use of fertilizers is quite limited, any increase in application will translate into a large percentage increase.

23 Admittedly, von Thünen (1996) assumes a homogeneous plane so that differences in crop choice are attributable to transport costs alone. Omamo (1998) shows how the presence of transportation-induced
from the cities and market centers, we expect them to shift out of perishable cash crops such as vegetables and into such storable crops as staples and pulses. In terms of cultivated land area devoted to various crops, we do find a decline in vegetables corresponding to increased isolation (Table 7). However, contrary to expectations, we find that staples (of which rice is clearly the most important crop) and pulses also decline. This is due to the fact that in more isolated areas more land is devoted to industrial and export crops such as vanilla, cloves and coffee. This in turn could follow from these crops being grown in the humid regions in the East and to the Northeast, which are isolated. Whether isolation in this region is a consequence of difficulties in maintaining roads under such ecological conditions is difficult to confirm. Note that when we consider all agricultural land except for industrial crops, we do observe that more land is devoted to staple crops the more isolated the area. Although land devoted to pulses decreases slightly, the von Thünen model is generally confirmed for direct food crops – less valuable crops are cultivated by more isolated farming households. As such, crop choice is one plausible transaction costs link between agricultural productivity and isolation.

*Plot-Size and Productivity*

Returning to rice productivity and plot size, both the reduced-form (Table 4) and the production function (Table 6) estimates indicate a negative relationship between land size and rice yields. The magnitudes of the relationship are roughly similar in the two

---

transaction costs can explain the seemingly inefficient cropping choices of farmers in Kenya where greater resources are devoted to low-yielding food crops instead of to cash crops that have higher market returns. See Minten and Kyle (1999) on Zaire, and Fafchamps and Shilpi (2003) on Nepal.
models, with elasticity estimates of -0.21 for the reduced-form model and -0.30 for the production function model. Thus, despite controlling for soil quality and other possible factors, we find further evidence of the inverse relationship between plot size and productivity consistent with other studies in Madagascar (Barrett, 1996a; Randrianarisoa, 2001; Randrianarisoa and Minten, 2001). Interestingly, this relationship also varies statistically across the isolation quintiles with production-function elasticity estimates ranging from -0.12 in the least isolated quintile to -0.30 in the most isolated quintile.

As indicated in Table 3, the median plot sizes are larger in more remote areas. Simple OLS regressions of plot size on dummies for remoteness (cost of transportation) also indicate that plots in the most remote quintile are 114 percent larger than in the least remote quintile on average, and that the increase in plot size by remoteness is monotonic on average.\textsuperscript{26} As such, the increase in plot size by isolation quintile observed in our data appears to explain part of the differences in productivity. Before policy implications can be addressed, however, further research is necessary to understand the source of the inverse relationship (e.g. labor, land and/or financial market failures) and its increasing strength with remoteness.\textsuperscript{27}

\textit{Price Variation and Extensification}

As illustrated in Figure 2, seasonal price variability as measured by the commune census is greater in more isolated areas (Minten and Randrianarison, 2003). The low prices in the April-June period correspond to the harvest when farmers sell much of their

\textsuperscript{25} A simple regression of share of land devoted to staple crops on remoteness controlling for unobserved province effects, confirms that in the two most remote quintiles more land is devoted to staple crops.

\textsuperscript{26} The regression results are available on request from the authors.
crop, while high prices observed from January through March coincide with the lean period. As noted earlier, this is consistent with households using the rice market as a pseudo capital market (Barrett, 1996b; Barrett and Dorosh, 1996). In addition, a simple OLS regression of the standard deviation of commune temporal rice prices on isolation quintile dummies indicates that rice prices are 13 percent more variable in the most remote areas compared to the least remote.

The reduced form model in Table 4 (Model 3) suggests that this greater temporal price variability in more isolated areas likely contributes to lower yields there. The estimated elasticity of -0.23 combined with 13 percent more temporal price variability in more remote areas, indicates that rice yields are 3 percent lower in the most remote quintile compared to the least remote quintile due to factors affected by price variability.

What are these factors? The input demand models (Table 5) suggest that household labor is affected by price variability. Although this is not a formal test, the positive relationship between price variability and household labor allocation is consistent with farming households self-insuring by increasing production through expansion into less fertile land and through less reliance on inputs and investments that increase land productivity (Barrett, 1999). Indeed, this increase in labor household labor into less productive soils may explain the lack of effect of household labor on rice yields in the production functions that appear in Table 6.

27 Because plot sizes are larger in more remote areas, lower yields in more remote areas do not necessarily translate into lower total agricultural production. The agricultural productivity-poverty link is weakened.
28 Note that the decreased use of tractors with price variability is also consistent with less reliance on other inputs, and with extensification into less fertile hillside lands where tractors cannot easily be used.
Insecurity

Physical insecurity as measured by living in a commune designated as a zone rouge, has a negative effect on production.29 This is hardly surprising given the incentive to under-invest resources in agricultural land when the benefits are more uncertain in the presence of rural insecurity. Thus we find that, ceteris paribus, rice yields are 16% lower on average in regions with high insecurity than in relatively safe regions (i.e. coefficient of -0.17 in Table 4). Given Fafchamps and Moser’s (2003) finding using these same data that crime in Madagascar increases with distance from urban centers and decreases with population density (i.e. with increases in isolation),30 insecurity is a likely contributor to lower agricultural productivity in isolated areas. As was the case with price variability, those living in insecure areas (zones rouge) are also likely to employ more household labor (Table 4). But further, they are 9 percent less likely to use fertilizer in rice production. These results are consistent with household allocating household labor resources to protecting crops in areas that are more prone to theft (Freudenberger, 1998) and accepting lower income levels in exchange for lower income variance (Rosenzweig and Binswanger, 1993) by not applying fertilizer.

VI. Concluding Remarks

This paper examines mechanisms that transmit isolation into low agricultural productivity. We study this for the case of Madagascar using unique household survey data combined with a census of administrative communes. Given the importance of

29 The national police divide the country into three zones classified by the degree of insecurity. The red zone (zone rouge) characterizes the most insecure areas.
agriculture to the rural poor, where nine out of ten poor persons is engaged in farming, we concentrate on isolation manifesting itself in the form of high transaction costs such as the cost of transporting agricultural commodities to major market centers.

A surprising result of this analysis is the degree to which crop yields for the three major staple items in Madagascar (rice, maize and cassava) are lower in isolated relative to non-isolated areas, even when we control for soil fertility. In fact, we find that rice yields fall by 33 percent between the least and most isolated quintiles, and that maize and cassava yields are 50 percent lower in the most isolated quintiles. While the slightly lower average market output prices in more isolated areas might lead us to expect lower total output levels, the substantive differences in yields were not expected *ex ante*.

We explore four factors that have the potential to explain the negative productivity-isolation relationship. These include (1) transportation-induced transaction costs, (2) the inverse relationship between plot-size and productivity, (3) increasing price variability and extensification onto less fertile land, and (4) insecurity. Reduced-form and structural rice yield production function estimates as well as input demand function estimates indicate that each of these factors contributes in its own way. First, transaction costs associated with isolation lead households to employ less household labor and to use less fertilizer. Second, larger plot sizes in more isolated areas combine with the confirmed inverse relationship between plot-size and productivity to negatively affect production. Third, the behavior of farming households in the presence of higher price risk in isolated areas is consistent with self-insuring by increasing production through extensification into less fertile lands and through fewer investments that increase land

---

30 Interestingly, they find that the relationship is independent of law enforcement personnel. Although law enforcement personnel help to solve crimes, they do not appear to prevent them (Fafchamps and Moser,
productivity. Finally, we find that households under-invest resources in agricultural land when the benefits are more uncertain in the presence of rural insecurity.

Other dynamic influences, such as externalities associated with location effects as discussed in the geographic poverty trap literature (Jalan and Ravallion, 2002; Deininger and Okidi, 2003; Dercon, 2003; De Vreyer et al., 2003), and induced innovation (Hayami and Ruttan, 1985), are also likely to be pertinent, a lack of access to appropriate panel data prevent us from addressing them empirically.

Finally, in this analysis, we approach the isolation-poverty relationship through but one possible mechanism – agricultural productivity and output – and find a strong link. An additional mechanism that we mention in passing is the possible interactions between non-farm income earning activities and isolation. This is fertile ground for further research.
References:


Ravallion, M., Datt, G., 2002. Why has Economic Growth been more Pro-Poor in some States of India than Others? *Journal of Development Economics* 68, 381-400


Appendix: Model of Rice Production and Input Demand

To analyze the effects of isolation on agricultural productivity, and on rice production in particular, we estimate a rice production function model. We opt to estimate the primal production function, rather than the dual profit function, primarily because the latter is conditioned on prices. And while we employ prices as instruments, we caution that their usefulness as signals is limited. The reasons for this are several. The use of realized *ex post* output prices is complicated by the timing of the input quantity decisions which are made well in advance of the harvest. In fact, these decisions are made conditional on the *ex ante* expected value of these prices. Given the uncertainties in agricultural production and prices, the correspondence between the expected and realized prices is unlikely to be a tight fit. In addition all of the prices available in the data suffer from some degree of aggregation bias (Deaton, 1988; Barrett, 1996a; Barrett, 1997). The commune average prices available for and used in this analysis fail to capture intertemporal and spatial variations in transaction prices within the commune. This uncaptured variation can result from the timing or the volume of sales, inter-linked contracts, differences in quality, and intra-commune isolation.31 Further, the non-separability of household decision-making and the market failures that arise as households selectively opt out of certain markets (often due to high transaction costs), can give rise to household-specific shadow prices that vary considerably from the commune average (de Janvry et al., 1991). As such, we prefer to estimate the production

---

31 We attempt to control for remoteness within the commune with the variable measuring distance from the plot to a passable route. But we note that this captures only part of the effect. In fact, this point highlights a limitation of this analysis in that remoteness is defined at the commune level, not the household level. Nonetheless, this is unlikely to be a first order problem.
function directly and address the endogeneity of input decisions with instrumental variables methods.

We assume a translog functional form of production:

\[
\ln(y) = \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln(x_i) + \sum_{i=1}^{n} \sum_{j=i+1}^{n} \beta_{ij} \ln(x_i) \ln(x_j) + \sum_{i=1}^{n} \beta_i \ln(x_i))^2 + \sum_{k=1}^{m} \delta_k z_k + \mu, \quad (1)
\]

where, \(y\) represents rice output from a plot of land, \(x\) is a vector of variable factor inputs (e.g. land, labor, fertilizer, seed, traction), \(z\) is a vector of productivity shifters (e.g. soil quality, irrigation, plot characteristics, household characteristics, and isolation measures), and \(\mu\) is an error term distributed with zero mean and unknown variance.

In order to estimate the model, a constant value of one is added to all the inputs before the logs are taken because the natural log of zero is undefined. Further, given that the choice variable inputs are likely correlated with unobserved ability, and as such are correlated with the error term, we estimate the model using instrumental variable (IV) methods. Once the model is estimated, the elasticities of the different factor inputs for each plot can be computed as

\[
\varepsilon_{y,x_i} = \alpha_i + \sum_{j<i}^{i-1} \beta_{ji} \ln(x_j) + \sum_{j>i}^{n} \beta_{ij} \ln(x_j) + 2\beta_i \ln(x_i). \quad (2)
\]
The elasticities reported in the tables are averages of the plot specific elasticities for each factor. Note that given that the elasticities are non-linear functions of the inputs, the average of the plot elasticities for a particular factor is not equal to the elasticity evaluated at the means of the inputs. Average elasticities are estimated for the entire sample and for each isolation quintile (where isolation is defined by travel time). Because we estimate the model with IV methods, and then evaluate the elasticities as averages within the samples of interest, we cannot determine the standard errors for the elasticities analytically. As such, the model and the elasticity estimates are bootstrapped (Brownstone and Valletta, 2001). Means of the bootstrapped average elasticities along with their t-statistics are reported in the tables. The instruments used in the first stage pass the usual battery of tests for validity on the base sample. Each of the F-tests for excluded instruments in the first-stage estimates are soundly rejected with p-values less than 0.001, as is the Sargan test of overidentification. The Anderson underidentification test is rejected at the 99 percent level of confidence.

The identifying instruments in the first stage models are output and input prices, household demographics and measures of institutional constraints in the commune. The latter perhaps needs some clarification. Each household in the dataset is asked if they have problems with regard to access to land, animal traction, labor, equipment and credit. They are also asked if they consider financial security to be a problem. Given the potential endogeneity of these variables, we enter as explanatory variables the non-self (household) commune means of the responses (Lanjouw and Ravallion, 1999; Christiaensen and Alderman, 2002) as measures of local institutional constraints.
Further, the estimates are made using ordinary least squares for the entire sample, including those plots on which zero quantities of the inputs are applied. This estimation procedure is adopted instead of a censored regression (Tobit), because we are interested in consistent estimates of the effect of input use on rice production, and consistency of these second-stage estimates does not depend on correctly specifying the functional form of the first-stage estimates. As Angrist and Krueger (2001) note, “using a nonlinear first stage to generate fitted values that are plugged directly into the second-stage equation does not generate consistent estimates unless the nonlinear model happens to be exactly right.”
Figure 1 — Communes by travel time to nearest city

Quintiles of Travel Time

Source: Commune Survey 2001, Programme Ilo, Cornell University
Figure 2: Seasonal Rice Price Variation by Remoteness Quintile

Source: Minten and Randrianarison, 2003
Table 1: Meaning and Implications of Remoteness

<table>
<thead>
<tr>
<th>Averages</th>
<th>Quintiles of Travel Time to Nearest City</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Least Remote</td>
</tr>
<tr>
<td>Meaning</td>
<td></td>
</tr>
<tr>
<td>Travel time (hr) to nearest major city†</td>
<td>0.89</td>
</tr>
<tr>
<td>Travel time (min on foot) from plot to accessible road</td>
<td>17.6</td>
</tr>
<tr>
<td>Ratio of transportation cost to price of 50kg sack of rice†</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Implications

| Percent poor* (headcount ratio) | 53.5 | 76.9 | 84.9 | 85.6 | 85.5 | 77.0 |
| Per capita consumption (FMG) | 1,498,981 | 764,505 | 667,171 | 686,607 | 638,218 | 858,366 |
| Share of autoconsumption in food consumption | 20.1 | 38.3 | 38.5 | 42.0 | 42.4 | 36.0 |

Note: Averages are calculated for households in the 2001 EPM survey.
† Indicates that data is from the 2001 Commune Census. All others are from the 2001 EPM survey.
* Poverty line is defined by INSTAT 2002
Table 2: Models of Cost of Transporting 50kg of Rice to Nearest Major City

*Dry Season Transportation*

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Coeff</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost of transportation (FMG)*</td>
<td>12,131</td>
<td>12,704</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance (km)*</td>
<td>206.1</td>
<td>247.2</td>
<td>0.58</td>
<td>15.17 **</td>
</tr>
<tr>
<td>Time (minutes) per km*</td>
<td>4.55</td>
<td>4.88</td>
<td>0.57</td>
<td>9.15 **</td>
</tr>
<tr>
<td>Number of cattle thefts in commune</td>
<td>51.3</td>
<td>177.8</td>
<td>0.00</td>
<td>0.82</td>
</tr>
<tr>
<td>Province Dummies</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antananarivo</td>
<td>0.26</td>
<td>0.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fianarantsoa</td>
<td>0.24</td>
<td>0.43</td>
<td>0.48</td>
<td>3.76 **</td>
</tr>
<tr>
<td>Toamasina</td>
<td>0.20</td>
<td>0.40</td>
<td>0.74</td>
<td>5.43 **</td>
</tr>
<tr>
<td>Mahajanga</td>
<td>0.12</td>
<td>0.32</td>
<td>0.30</td>
<td>1.94 +</td>
</tr>
<tr>
<td>Toliara</td>
<td>0.12</td>
<td>0.32</td>
<td>0.46</td>
<td>2.71 **</td>
</tr>
<tr>
<td>Antsiranana</td>
<td>0.06</td>
<td>0.23</td>
<td>0.54</td>
<td>2.25 *</td>
</tr>
<tr>
<td>Constant</td>
<td>5.08</td>
<td>29.50</td>
<td>0.814</td>
<td>29.50 **</td>
</tr>
</tbody>
</table>

R² = 0.814

No. observations = 163

* Logs in the regression

Data: 2001 Commune Census
Table 3: Agricultural Yields, Inputs and Prices by Degrees of Remoteness

<table>
<thead>
<tr>
<th>Quintiles of Travel Time to Nearest City</th>
<th>Least Remote</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Most Remote</th>
<th>All Rural</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Median Yields (kilograms per are)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice</td>
<td>28.0</td>
<td>28.5</td>
<td>15.8</td>
<td>18.6</td>
<td>19.0</td>
<td>22.8</td>
</tr>
<tr>
<td>Maize</td>
<td>16.7</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
<td>8.0</td>
<td>10.0</td>
</tr>
<tr>
<td>Cassava</td>
<td>50.0</td>
<td>32.0</td>
<td>33.0</td>
<td>24.0</td>
<td>25.0</td>
<td>30.0</td>
</tr>
<tr>
<td><strong>Input Use</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage of households using…</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemical fertilizers</td>
<td>28.1</td>
<td>25.4</td>
<td>4.9</td>
<td>0.9</td>
<td>3.8</td>
<td>11.7</td>
</tr>
<tr>
<td>Organic fertilizers</td>
<td>66.2</td>
<td>54.8</td>
<td>11.9</td>
<td>17.9</td>
<td>1.5</td>
<td>28.0</td>
</tr>
<tr>
<td>Pesticides/herbicides</td>
<td>18.8</td>
<td>15.6</td>
<td>3.3</td>
<td>2.9</td>
<td>1.9</td>
<td>7.8</td>
</tr>
<tr>
<td><strong>Average quantity (kg/are) of…</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemical fertilizers</td>
<td>0.28</td>
<td>0.48</td>
<td>0.10</td>
<td>0.01</td>
<td>0.10</td>
<td>0.21</td>
</tr>
<tr>
<td>Organic fertilizers</td>
<td>7.59</td>
<td>3.49</td>
<td>0.45</td>
<td>0.13</td>
<td>0.19</td>
<td>2.54</td>
</tr>
<tr>
<td>Pesticides/herbicides</td>
<td>1,013</td>
<td>369</td>
<td>36</td>
<td>27</td>
<td>62</td>
<td>319</td>
</tr>
<tr>
<td><strong>Average value (FMG/are) of…</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pesticides/herbicides</td>
<td>1,013</td>
<td>369</td>
<td>36</td>
<td>27</td>
<td>62</td>
<td>319</td>
</tr>
<tr>
<td><strong>Labor (average number of man-days per are)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All crops</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>28.1</td>
<td>29.3</td>
<td>51.9</td>
<td>45.5</td>
<td>38.1</td>
<td>38.0</td>
</tr>
<tr>
<td>Household</td>
<td>22.6</td>
<td>22.4</td>
<td>41.0</td>
<td>36.5</td>
<td>33.1</td>
<td>30.6</td>
</tr>
<tr>
<td>Reciprocal</td>
<td>2.6</td>
<td>3.5</td>
<td>6.4</td>
<td>5.1</td>
<td>4.0</td>
<td>4.3</td>
</tr>
<tr>
<td>Hired</td>
<td>2.9</td>
<td>3.4</td>
<td>4.5</td>
<td>3.9</td>
<td>1.0</td>
<td>3.2</td>
</tr>
<tr>
<td><strong>Rice (principal season)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>33.1</td>
<td>52.2</td>
<td>79.9</td>
<td>76.2</td>
<td>58.3</td>
<td>59.0</td>
</tr>
<tr>
<td>Household</td>
<td>24.3</td>
<td>39.1</td>
<td>65.2</td>
<td>62.9</td>
<td>48.6</td>
<td>47.0</td>
</tr>
<tr>
<td>Reciprocal</td>
<td>2.9</td>
<td>7.8</td>
<td>9.6</td>
<td>7.5</td>
<td>7.6</td>
<td>7.1</td>
</tr>
<tr>
<td>Hired</td>
<td>5.9</td>
<td>5.3</td>
<td>5.1</td>
<td>5.7</td>
<td>2.1</td>
<td>4.9</td>
</tr>
<tr>
<td><strong>Median Land Area (are/plot)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>20</td>
<td>30</td>
<td>48</td>
<td>30</td>
<td>24</td>
</tr>
<tr>
<td><strong>Prices</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daily wage rate (male)</td>
<td>5,909</td>
<td>5,851</td>
<td>5,914</td>
<td>5,777</td>
<td>10,079</td>
<td>6,667</td>
</tr>
<tr>
<td>Price of paddy (kg)</td>
<td>1,472</td>
<td>1,179</td>
<td>1,323</td>
<td>1,269</td>
<td>1,282</td>
<td>1,311</td>
</tr>
</tbody>
</table>

Data: 2001 Commune Census for remotenes; 2001 EPM for remainder.
### Table 4: Reduced form estimates of determinants of rice yields in Madagascar

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Coeff</th>
<th>t-stat</th>
<th>Coeff</th>
<th>t-stat</th>
<th>Coeff</th>
<th>t-stat</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost of transporting rice (log)</td>
<td>8.78</td>
<td>1.25</td>
<td>-0.169</td>
<td>-10.17 ***</td>
<td>-0.066</td>
<td>-3.28 ***</td>
<td>-0.047</td>
<td>-2.49 **</td>
</tr>
<tr>
<td>Dist plot to passable route (minutes walk)</td>
<td>20.10</td>
<td>22.65</td>
<td>-0.005</td>
<td>-6.70 ***</td>
<td>-0.002</td>
<td>-2.59 ***</td>
<td>-0.001</td>
<td>-1.99 **</td>
</tr>
<tr>
<td>Plot size (log are)</td>
<td>769.18</td>
<td>626.35</td>
<td>0.0002</td>
<td>0.70</td>
<td>0.00004</td>
<td>0.52</td>
<td>-0.003</td>
<td>4.67 ***</td>
</tr>
<tr>
<td>Average precipitation</td>
<td>1809.51</td>
<td>551.77</td>
<td>-0.00004</td>
<td>0.52</td>
<td>-0.00003</td>
<td>0.52</td>
<td>-0.003</td>
<td>4.67 ***</td>
</tr>
<tr>
<td>Average temperature</td>
<td>187.31</td>
<td>33.08</td>
<td>0.005</td>
<td>1.13</td>
<td>0.009</td>
<td>2.11 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std. dev. elevation</td>
<td>119.68</td>
<td>93.43</td>
<td>-0.007</td>
<td>2.14 **</td>
<td>-0.013</td>
<td>4.33 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std. dev. precipitation</td>
<td>74.04</td>
<td>82.30</td>
<td>-0.001</td>
<td>1.91 *</td>
<td>-0.001</td>
<td>1.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std. dev. temperature</td>
<td>6.62</td>
<td>4.38</td>
<td>0.173</td>
<td>2.55 **</td>
<td>0.274</td>
<td>4.53 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flood</td>
<td>0.413</td>
<td>0.49</td>
<td>-0.137</td>
<td>3.00 ***</td>
<td>-0.097</td>
<td>2.37 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drought</td>
<td>0.369</td>
<td>0.48</td>
<td>-0.085</td>
<td>2.22 **</td>
<td>-0.027</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irrigated field</td>
<td>0.746</td>
<td>0.44</td>
<td>0.111</td>
<td>2.68 ***</td>
<td>0.102</td>
<td>2.77 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commune population density</td>
<td>105.40</td>
<td>266.68</td>
<td>0.0005</td>
<td>5.19 ***</td>
<td>0.0003</td>
<td>3.30 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Model 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost of transporting rice (log)</td>
<td>8.78</td>
<td>1.25</td>
<td>-0.169</td>
<td>-10.17 ***</td>
<td>-0.066</td>
<td>-3.28 ***</td>
<td>-0.047</td>
<td>-2.49 **</td>
</tr>
<tr>
<td>Dist plot to passable route (minutes walk)</td>
<td>20.10</td>
<td>22.65</td>
<td>-0.005</td>
<td>-6.70 ***</td>
<td>-0.002</td>
<td>-2.59 ***</td>
<td>-0.001</td>
<td>-1.99 **</td>
</tr>
<tr>
<td>Plot size (log are)</td>
<td>769.18</td>
<td>626.35</td>
<td>0.0002</td>
<td>0.70</td>
<td>0.00004</td>
<td>0.52</td>
<td>-0.003</td>
<td>4.67 ***</td>
</tr>
<tr>
<td>Average precipitation</td>
<td>1809.51</td>
<td>551.77</td>
<td>-0.00004</td>
<td>0.52</td>
<td>-0.00003</td>
<td>0.52</td>
<td>-0.003</td>
<td>4.67 ***</td>
</tr>
<tr>
<td>Average temperature</td>
<td>187.31</td>
<td>33.08</td>
<td>0.005</td>
<td>1.13</td>
<td>0.009</td>
<td>2.11 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std. dev. elevation</td>
<td>119.68</td>
<td>93.43</td>
<td>-0.007</td>
<td>2.14 **</td>
<td>-0.013</td>
<td>4.33 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std. dev. precipitation</td>
<td>74.04</td>
<td>82.30</td>
<td>-0.001</td>
<td>1.91 *</td>
<td>-0.001</td>
<td>1.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std. dev. temperature</td>
<td>6.62</td>
<td>4.38</td>
<td>0.173</td>
<td>2.55 **</td>
<td>0.274</td>
<td>4.53 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flood</td>
<td>0.413</td>
<td>0.49</td>
<td>-0.137</td>
<td>3.00 ***</td>
<td>-0.097</td>
<td>2.37 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drought</td>
<td>0.369</td>
<td>0.48</td>
<td>-0.085</td>
<td>2.22 **</td>
<td>-0.027</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irrigated field</td>
<td>0.746</td>
<td>0.44</td>
<td>0.111</td>
<td>2.68 ***</td>
<td>0.102</td>
<td>2.77 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commune population density</td>
<td>105.40</td>
<td>266.68</td>
<td>0.0005</td>
<td>5.19 ***</td>
<td>0.0003</td>
<td>3.30 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Model 3</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost of transporting rice (log)</td>
<td>8.78</td>
<td>1.25</td>
<td>-0.169</td>
<td>-10.17 ***</td>
<td>-0.066</td>
<td>-3.28 ***</td>
<td>-0.047</td>
<td>-2.49 **</td>
</tr>
<tr>
<td>Dist plot to passable route (minutes walk)</td>
<td>20.10</td>
<td>22.65</td>
<td>-0.005</td>
<td>-6.70 ***</td>
<td>-0.002</td>
<td>-2.59 ***</td>
<td>-0.001</td>
<td>-1.99 **</td>
</tr>
<tr>
<td>Plot size (log are)</td>
<td>769.18</td>
<td>626.35</td>
<td>0.0002</td>
<td>0.70</td>
<td>0.00004</td>
<td>0.52</td>
<td>-0.003</td>
<td>4.67 ***</td>
</tr>
<tr>
<td>Average precipitation</td>
<td>1809.51</td>
<td>551.77</td>
<td>-0.00004</td>
<td>0.52</td>
<td>-0.00003</td>
<td>0.52</td>
<td>-0.003</td>
<td>4.67 ***</td>
</tr>
<tr>
<td>Average temperature</td>
<td>187.31</td>
<td>33.08</td>
<td>0.005</td>
<td>1.13</td>
<td>0.009</td>
<td>2.11 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std. dev. elevation</td>
<td>119.68</td>
<td>93.43</td>
<td>-0.007</td>
<td>2.14 **</td>
<td>-0.013</td>
<td>4.33 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std. dev. precipitation</td>
<td>74.04</td>
<td>82.30</td>
<td>-0.001</td>
<td>1.91 *</td>
<td>-0.001</td>
<td>1.19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Std. dev. temperature</td>
<td>6.62</td>
<td>4.38</td>
<td>0.173</td>
<td>2.55 **</td>
<td>0.274</td>
<td>4.53 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flood</td>
<td>0.413</td>
<td>0.49</td>
<td>-0.137</td>
<td>3.00 ***</td>
<td>-0.097</td>
<td>2.37 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drought</td>
<td>0.369</td>
<td>0.48</td>
<td>-0.085</td>
<td>2.22 **</td>
<td>-0.027</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irrigated field</td>
<td>0.746</td>
<td>0.44</td>
<td>0.111</td>
<td>2.68 ***</td>
<td>0.102</td>
<td>2.77 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Commune population density</td>
<td>105.40</td>
<td>266.68</td>
<td>0.0005</td>
<td>5.19 ***</td>
<td>0.0003</td>
<td>3.30 ***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Dependent variable is kilograms of rice produced per are for each plot.
<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>Tobit</th>
<th>Tobit</th>
<th>Probit</th>
<th>Probit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Household</td>
<td>Hired</td>
<td>Animal Traction</td>
<td>Tractor</td>
<td>Tractor</td>
</tr>
<tr>
<td></td>
<td>Coeff</td>
<td>Coeff</td>
<td>Coeff</td>
<td>Effect</td>
<td>Effect</td>
</tr>
<tr>
<td>Log cost of transporting rice</td>
<td>-0.401</td>
<td>-3.33 ***</td>
<td>-0.729</td>
<td>0.43 ***</td>
<td>0.002</td>
</tr>
<tr>
<td>Dist plot to passable route (minutes walk)</td>
<td>-0.001</td>
<td>-3.33 ***</td>
<td>-0.005</td>
<td>-0.10</td>
<td>0.001</td>
</tr>
<tr>
<td>Log Std. dev. temporal prices in commune</td>
<td>0.046</td>
<td>1.96 **</td>
<td>-0.034</td>
<td>0.30</td>
<td>-0.002</td>
</tr>
<tr>
<td>Designated &quot;zone rouge&quot; *</td>
<td>0.618</td>
<td>2.33 **</td>
<td>0.174</td>
<td>1.33</td>
<td>-0.007</td>
</tr>
<tr>
<td></td>
<td>0.191</td>
<td>-0.97</td>
<td>0.086</td>
<td>0.40</td>
<td>0.121</td>
</tr>
<tr>
<td></td>
<td>0.507</td>
<td>-1.67</td>
<td>0.180</td>
<td>1.08</td>
<td>-0.645</td>
</tr>
<tr>
<td></td>
<td>0.314</td>
<td>1.23</td>
<td>-0.012</td>
<td>-0.09</td>
<td>-1.603</td>
</tr>
<tr>
<td></td>
<td>-0.045</td>
<td>-0.22</td>
<td>-0.057</td>
<td>-0.58</td>
<td>-0.260</td>
</tr>
<tr>
<td></td>
<td>0.002</td>
<td>0.01</td>
<td>-0.016</td>
<td>-0.13</td>
<td>-0.090</td>
</tr>
<tr>
<td>Majority district soil sediment</td>
<td>1.199</td>
<td>3.45 ***</td>
<td>-0.167</td>
<td>-0.94</td>
<td>-1.396</td>
</tr>
<tr>
<td>Majority district soil volcanic</td>
<td>0.809</td>
<td>2.73 ***</td>
<td>-0.010</td>
<td>-0.07</td>
<td>-2.780</td>
</tr>
<tr>
<td>Majority district soil alluvial</td>
<td>0.332</td>
<td>2.08</td>
<td>-0.116</td>
<td>-0.61</td>
<td>-0.218</td>
</tr>
<tr>
<td>Average elevation</td>
<td>0.003</td>
<td>2.15 **</td>
<td>0.001</td>
<td>1.89</td>
<td>0.002</td>
</tr>
<tr>
<td>Average precipitation</td>
<td>0.002</td>
<td>6.90 ***</td>
<td>0.000</td>
<td>0.38</td>
<td>-0.002</td>
</tr>
<tr>
<td>Average temperature</td>
<td>0.015</td>
<td>0.67</td>
<td>0.032</td>
<td>2.75 ***</td>
<td>0.031</td>
</tr>
<tr>
<td>Std. dev. elevation</td>
<td>0.083</td>
<td>5.22 ***</td>
<td>0.029</td>
<td>3.74</td>
<td>-0.063</td>
</tr>
<tr>
<td>Std. dev. precipitation</td>
<td>-0.007</td>
<td>3.19</td>
<td>-0.001</td>
<td>-0.90</td>
<td>-0.004</td>
</tr>
<tr>
<td>Std. dev. temperature</td>
<td>-1.560</td>
<td>4.70 ***</td>
<td>-0.587</td>
<td>-3.62 ***</td>
<td>1.204</td>
</tr>
<tr>
<td>Flood</td>
<td>-0.240</td>
<td>-1.08</td>
<td>0.087</td>
<td>0.79</td>
<td>-0.304</td>
</tr>
<tr>
<td>Drought</td>
<td>-1.018</td>
<td>-5.28 ***</td>
<td>-0.281</td>
<td>2.97</td>
<td>-0.274</td>
</tr>
<tr>
<td>Irrigated field</td>
<td>0.373</td>
<td>1.87 *</td>
<td>0.094</td>
<td>0.97</td>
<td>-0.094</td>
</tr>
<tr>
<td>Commune population density</td>
<td>0.004</td>
<td>9.53 ***</td>
<td>0.001</td>
<td>6.69</td>
<td>0.000</td>
</tr>
<tr>
<td>Problem: Access to land (non-self cluster mean)</td>
<td>-1.304</td>
<td>-3.34 ***</td>
<td>0.427</td>
<td>2.19 **</td>
<td>1.314</td>
</tr>
<tr>
<td>Problem: Access to zebu (non-self cluster mean)</td>
<td>-2.802</td>
<td>-5.44 ***</td>
<td>-0.229</td>
<td>-0.88</td>
<td>1.562</td>
</tr>
<tr>
<td>Problem: Access to labor (non-self cluster mean)</td>
<td>2.992</td>
<td>5.47 ***</td>
<td>-0.586</td>
<td>-2.28 ***</td>
<td>0.546</td>
</tr>
<tr>
<td>Problem: Access to equipment (non-self cluster mean)</td>
<td>0.479</td>
<td>0.91</td>
<td>0.493</td>
<td>1.89 **</td>
<td>-0.749</td>
</tr>
<tr>
<td>Problem: Access to credit (non-self cluster mean)</td>
<td>0.396</td>
<td>0.87</td>
<td>-0.429</td>
<td>2.09 **</td>
<td>-1.534</td>
</tr>
<tr>
<td>Agr. is principal activity</td>
<td>0.040</td>
<td>0.13</td>
<td>0.027</td>
<td>0.19</td>
<td>-0.901</td>
</tr>
<tr>
<td>HH owns 1 to 5 head of livestock</td>
<td>-0.354</td>
<td>-1.76 *</td>
<td>0.265</td>
<td>2.68 ***</td>
<td>1.276</td>
</tr>
<tr>
<td>HH owns more than 5 head of livestock</td>
<td>-0.168</td>
<td>-0.58</td>
<td>0.046</td>
<td>0.31</td>
<td>1.386</td>
</tr>
<tr>
<td>Female head</td>
<td>-0.019</td>
<td>-0.08</td>
<td>0.383</td>
<td>3.42 ***</td>
<td>0.009</td>
</tr>
<tr>
<td>Age of head</td>
<td>-0.003</td>
<td>-0.54</td>
<td>-0.005</td>
<td>-3.63 ***</td>
<td>0.000</td>
</tr>
<tr>
<td>Number of kids 6-14 in HH</td>
<td>0.046</td>
<td>0.73</td>
<td>-0.014</td>
<td>-0.45</td>
<td>0.057</td>
</tr>
<tr>
<td>Number of adult women in HH</td>
<td>-0.256</td>
<td>-2.14 **</td>
<td>-0.158</td>
<td>2.66 ***</td>
<td>0.343</td>
</tr>
<tr>
<td>Number of adult men in HH</td>
<td>0.068</td>
<td>0.83</td>
<td>-0.085</td>
<td>1.58</td>
<td>-0.249</td>
</tr>
<tr>
<td>No. adults w/ primary educ</td>
<td>0.036</td>
<td>0.43</td>
<td>0.019</td>
<td>0.46</td>
<td>0.008</td>
</tr>
<tr>
<td>No. adults w/ secondary educ</td>
<td>0.028</td>
<td>-2.37 **</td>
<td>0.094</td>
<td>0.92</td>
<td>-0.145</td>
</tr>
<tr>
<td>No. adults w/ tertiary educ</td>
<td>-1.106</td>
<td>2.09 **</td>
<td>0.090</td>
<td>4.02 ***</td>
<td>-0.101</td>
</tr>
<tr>
<td>Fianarantsoa</td>
<td>0.239</td>
<td>0.53</td>
<td>-0.571</td>
<td>2.61 ***</td>
<td>-1.957</td>
</tr>
<tr>
<td>Taonantsika</td>
<td>0.988</td>
<td>2.02 **</td>
<td>-1.258</td>
<td>5.24</td>
<td>-1.721</td>
</tr>
<tr>
<td>Mahajanga</td>
<td>1.443</td>
<td>2.32 **</td>
<td>-1.189</td>
<td>-3.89 ***</td>
<td>-1.894</td>
</tr>
<tr>
<td>Tolara</td>
<td>1.471</td>
<td>2.12 **</td>
<td>-1.361</td>
<td>3.98 ***</td>
<td>0.433</td>
</tr>
<tr>
<td>Ankaramana</td>
<td>0.659</td>
<td>0.95</td>
<td>-1.915</td>
<td>6.65 ***</td>
<td>-0.702</td>
</tr>
<tr>
<td>Constant</td>
<td>-4.126</td>
<td>-0.75</td>
<td>-5.749</td>
<td>-2.11 **</td>
<td>-7.421</td>
</tr>
<tr>
<td>Number of left-censored observations</td>
<td>508</td>
<td>1065</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.161</td>
<td>0.052</td>
<td>0.073</td>
<td>0.31</td>
<td>0.408</td>
</tr>
<tr>
<td>Number of observations</td>
<td>1,939</td>
<td>1,939</td>
<td>1,939</td>
<td>1,939</td>
<td>1,939</td>
</tr>
</tbody>
</table>
Table 6: Input Elasticities and Elasticities of Productivity Shifters in Rice Yield Production Function Estimates

<table>
<thead>
<tr>
<th></th>
<th>Full Sample</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Least Remote</td>
<td>Quintile 2</td>
<td>Quintile 3</td>
<td>Quintile 4</td>
<td>Most Remote</td>
<td>Difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elasticity</td>
<td>t-stat</td>
<td>Elasticity</td>
<td>t-stat</td>
<td>Elasticity</td>
<td>t-stat</td>
<td>Elasticity</td>
<td>t-stat</td>
</tr>
<tr>
<td>Inputs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Area under cultivation (are)</td>
<td>-0.210</td>
<td>-9.53 ***</td>
<td>-0.122</td>
<td>-5.99 ***</td>
<td>-0.203</td>
<td>-8.22 ***</td>
<td>-0.236</td>
<td>-8.96 ***</td>
<td>-0.225</td>
</tr>
<tr>
<td>Household labor (days)</td>
<td>0.003</td>
<td>0.18</td>
<td>-0.020</td>
<td>-0.96</td>
<td>0.011</td>
<td>0.56</td>
<td>0.005</td>
<td>0.25</td>
<td>0.016</td>
</tr>
<tr>
<td>Hired/reciprocal labor (days)</td>
<td>0.085</td>
<td>4.72 ***</td>
<td>0.094</td>
<td>4.59 ***</td>
<td>0.071</td>
<td>3.72 ***</td>
<td>0.097</td>
<td>4.89 ***</td>
<td>0.093</td>
</tr>
<tr>
<td>Animal traction (hours)</td>
<td>0.021</td>
<td>1.13</td>
<td>-0.006</td>
<td>-0.35</td>
<td>0.022</td>
<td>1.06</td>
<td>0.035</td>
<td>1.88 *</td>
<td>0.024</td>
</tr>
<tr>
<td>Tractor use (hours)</td>
<td>0.102</td>
<td>3.33 ***</td>
<td>0.082</td>
<td>2.43 **</td>
<td>0.095</td>
<td>3.59 ***</td>
<td>0.126</td>
<td>3.77 ***</td>
<td>0.097</td>
</tr>
<tr>
<td>Fertilizer (kg)</td>
<td>0.052</td>
<td>2.77 ***</td>
<td>0.047</td>
<td>2.88 ***</td>
<td>0.061</td>
<td>3.43 ***</td>
<td>0.051</td>
<td>2.61 ***</td>
<td>0.056</td>
</tr>
<tr>
<td>Productivity Shifters</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hill-side</td>
<td>-0.003</td>
<td>-0.37</td>
<td>-0.003</td>
<td>-0.39</td>
<td>-0.004</td>
<td>-0.37</td>
<td>-0.004</td>
<td>-0.37</td>
<td>-0.003</td>
</tr>
<tr>
<td>Terraced</td>
<td>0.001</td>
<td>0.30</td>
<td>0.002</td>
<td>0.29</td>
<td>0.001</td>
<td>0.31</td>
<td>0.000</td>
<td>0.29</td>
<td>0.000</td>
</tr>
<tr>
<td>Mod. to steep slope</td>
<td>0.002</td>
<td>0.31</td>
<td>0.002</td>
<td>0.31</td>
<td>0.002</td>
<td>0.30</td>
<td>0.002</td>
<td>0.32</td>
<td>0.002</td>
</tr>
<tr>
<td>Soil - claylike</td>
<td>0.006</td>
<td>0.46</td>
<td>0.005</td>
<td>0.47</td>
<td>0.007</td>
<td>0.46</td>
<td>0.005</td>
<td>0.45</td>
<td>0.008</td>
</tr>
<tr>
<td>Soil - muddy</td>
<td>-0.018</td>
<td>-2.05 *</td>
<td>-0.015</td>
<td>-1.98 *</td>
<td>-0.014</td>
<td>-2.03 *</td>
<td>-0.025</td>
<td>-2.04 *</td>
<td>-0.015</td>
</tr>
<tr>
<td>Majority of commune soil is sediment</td>
<td>0.006</td>
<td>0.94</td>
<td>0.003</td>
<td>0.93</td>
<td>0.008</td>
<td>0.92</td>
<td>0.010</td>
<td>0.94</td>
<td>0.005</td>
</tr>
<tr>
<td>Majority of commune soil is volcanic</td>
<td>-0.007</td>
<td>-1.15</td>
<td>-0.013</td>
<td>-1.13</td>
<td>-0.004</td>
<td>-1.09</td>
<td>-0.004</td>
<td>-1.11</td>
<td>-0.009</td>
</tr>
<tr>
<td>Majority of commune soil is alluvial</td>
<td>-0.001</td>
<td>-0.36</td>
<td>-0.002</td>
<td>-0.34</td>
<td>-0.000</td>
<td>-0.40</td>
<td>-0.001</td>
<td>-0.39</td>
<td>-0.001</td>
</tr>
<tr>
<td>Average elevation in commune</td>
<td>-0.220</td>
<td>-1.59</td>
<td>-0.331</td>
<td>-1.59</td>
<td>-0.292</td>
<td>-1.58</td>
<td>-0.156</td>
<td>-1.58</td>
<td>-0.127</td>
</tr>
<tr>
<td>Average precipitation in commune</td>
<td>-0.277</td>
<td>-3.22 **</td>
<td>-0.234</td>
<td>-3.21 **</td>
<td>-0.253</td>
<td>-3.22 **</td>
<td>-0.305</td>
<td>-3.22 **</td>
<td>-0.313</td>
</tr>
<tr>
<td>Average temperature in commute</td>
<td>-1.050</td>
<td>-1.58</td>
<td>-0.935</td>
<td>-1.58</td>
<td>-0.978</td>
<td>-1.58</td>
<td>-1.108</td>
<td>-1.58</td>
<td>-1.144</td>
</tr>
<tr>
<td>Std. Deviation of elevation in commune</td>
<td>-0.433</td>
<td>-1.51</td>
<td>-0.488</td>
<td>-1.51</td>
<td>-0.297</td>
<td>-1.51</td>
<td>-0.458</td>
<td>-1.50</td>
<td>-0.411</td>
</tr>
<tr>
<td>Std. Deviation of precipitation in commune</td>
<td>0.046</td>
<td>1.52</td>
<td>0.033</td>
<td>1.51</td>
<td>0.031</td>
<td>1.50</td>
<td>0.064</td>
<td>1.53</td>
<td>0.042</td>
</tr>
<tr>
<td>Std. Deviation of temperature in commune</td>
<td>0.425</td>
<td>1.55</td>
<td>0.477</td>
<td>1.55</td>
<td>0.293</td>
<td>1.54</td>
<td>0.451</td>
<td>1.54</td>
<td>0.414</td>
</tr>
<tr>
<td>Irrigated field**</td>
<td>0.093</td>
<td>3.32 **</td>
<td>0.080</td>
<td>2.38 *</td>
<td>0.110</td>
<td>3.51 **</td>
<td>0.102</td>
<td>3.69 **</td>
<td>0.065</td>
</tr>
<tr>
<td>Flood</td>
<td>-0.015</td>
<td>-0.96</td>
<td>-0.008</td>
<td>-0.95</td>
<td>-0.010</td>
<td>-0.94</td>
<td>-0.024</td>
<td>-0.96</td>
<td>-0.019</td>
</tr>
<tr>
<td>Drought**</td>
<td>0.008</td>
<td>0.66</td>
<td>0.006</td>
<td>0.90</td>
<td>0.012</td>
<td>0.76</td>
<td>0.009</td>
<td>0.54</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Adjusted \( R^2 \) from first replication | 0.459 |
No. observations | 1,939 | 376 | 514 | 569 | 243 | 238 |

++ Elasticity calculations and t-statistics include the interaction term
Note: 1,500 bootstrap replications
Table 7: Average Share of Household Cultivated Land Devoted to Various Crops, by Remoteness Quintiles

<table>
<thead>
<tr>
<th>Quintiles of Transport Cost</th>
<th>Staple Crops</th>
<th>Industrial and Export Crops</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Rice</td>
</tr>
<tr>
<td><strong>All Crops</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Least Remote</td>
<td>83.2</td>
<td>50.9</td>
</tr>
<tr>
<td>2</td>
<td>82.2</td>
<td>50.8</td>
</tr>
<tr>
<td>3</td>
<td>86.5</td>
<td>60.0</td>
</tr>
<tr>
<td>4</td>
<td>84.3</td>
<td>49.2</td>
</tr>
<tr>
<td>Most Remote</td>
<td>79.6</td>
<td>52.9</td>
</tr>
<tr>
<td>Total Rural</td>
<td>83.2</td>
<td>53.1</td>
</tr>
<tr>
<td><strong>Just Non-Industrial Domestic Food Crops</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Least Remote</td>
<td>86.3</td>
<td>53.5</td>
</tr>
<tr>
<td>2</td>
<td>87.0</td>
<td>53.8</td>
</tr>
<tr>
<td>3</td>
<td>94.0</td>
<td>65.6</td>
</tr>
<tr>
<td>4</td>
<td>93.8</td>
<td>56.3</td>
</tr>
<tr>
<td>Most Remote</td>
<td>94.6</td>
<td>64.6</td>
</tr>
<tr>
<td>Total Rural</td>
<td>91.5</td>
<td>59.3</td>
</tr>
</tbody>
</table>
**Data Appendix**

The following are Stata programs used to creates Tables 1 - 7.

```

*****************************************************************
**  Table 1 of Stifel and Minten (2008).**  Stata file to examine the meaning and implications
**  of remoteness using the 2001 commune census and the
**  2001 EPM
**
*****************************************************************

version 10.0
clear
set memory 15m
set more off

*-------------*
*-- Meaning --*
*-------------*

use data\cc_epm01_remoteness_index.dta, clear
    keep codcom idmen poids pop time_cup tquint costt
    sort idmen
    save data\temp.dta, replace

sum time_cup, detail
* Drop extreme values *
drop if time >= 99

tabstat time_cup [aw=poids], stat(mean) by(tquint)
```

```
use data\agricula1.dta, clear
rename slallq3 t_plot_road
keep idmen t_plot_road
sort idmen
merge idmen using data\temp.dta
keep if _m == 3

tabstat t_plot_road [aw=poid], stat(mean) by(tquint)

use data\c.dta, clear
gen p_rice = c3c5 * 50
keep codcom p_rice
label variable p_rice "Price of 50kg of milled local rice"
sort codcom
save data\temp.dta, replace

use data\cc_epm01_remoteness_index.dta, clear
keep codcom idmen poids hhmemb pop time* rdec-tremote costt
sort codcom
merge codcom using data\temp.dta
keep if _m==3
drop _m
gen cost_price = costtran / p_rice
sum cost_pr, detail
* Recode to 1, the observations for which price < transport cost *
replace cost_pr = 1 if cost_pr>1

tabstat cost_pr [aw=poid], stat(mean) by(tquint)

*------------------*
*-- Implications --*
*------------------*
use data\cc_epm01 remoteness_index.dta, clear
  keep idmen tquint
  sort idmen
save data\temp.dta, replace

use data\total_exp.dta, clear
  keep idmen pcpoids poids tot_pc_pa autocon autocons food_tot
  sort idmen
merge idmen using data\temp.dta
  tab _m
  keep if _m == 3
  drop _m
  gen byte poor = 100*(tot_pc_pa<=1051425.1)
  gen autoshr = 100* (autocon + autocons) / food_tot
  gen byte total = 1

  table tquint [aw=pcpoid], c(mean poor mean tot_pc_pa mean autoshr)
  table total [aw=pcpoid], c(mean poor mean tot_pc_pa mean autoshr)
*-----------------------*

*--------------------------------------------------------*

*******************************************************************************
* Table 2 of Stifel and Minten (2008).
*******************************************************************************
* Stata program to run a regression of the cost of
* transporting 50kg of rice to the nearest CUP (primary
* city) on
* *
* - distance (km) to nearest CUP
* - distance (hours) to nearest CUP
* - insecurity in the commune
* - Faritany dummies
*******************************************************************************
*--------------------------------------------------------*

*------------------*
use data\cc_epi01_remoteness_index.dta, clear
qui tab s0qb8, gen(far)
keep idmen codcom* poids hhmemb pop costtran time_cup we* far* dist_cup czebu zeb_theft-veh_theft_3

gend tkm = 60 * time / dist
gen lncost = ln(costt)
gen lncost = ln(wet_costt)
gen lndist = ln(dist)
gen lntkm = ln(tkm)
gen dist2 = dist^2
gen tkm2 = tkm^2
gen zdpoid = sum(poids), by(codcom_epm)

*-- Keep only one observation per commune --*

bysort codcom_epm: gen ind = _n
drop ind if ind > 1

drop lncost if lncost < 5
drop lncost if lncost == .

*--------------------*
*-- Estimate Model --*
*--------------------*

sum costt dist tkm zeb_theft far1-far6 [aw=zdpoid]
regress lncost lndist lntkm zeb_theft far2-far6 [aw=zdpoid], robust

*------------------------------------------------------------------------*

* Table 3 of Stifel and Minten (2008).
* Stata file to examine the variation in agric yields,*
* input use and prices across remoteness quintiles*
* using the 2001 commune census and the 2001 EPM*

*-----------------------------------------------------------*

*--- Yields ---*

use data\epm01_prod_fn_data.dta, clear
  gen y_rice  = yield if code == 101
  gen y_maize = yield if code == 201
  gen y_cass  = yield if code == 302

  tabstat y_rice  [aw=poid], stat(median) by(tquint)
  tabstat y_maize [aw=poid], stat(median) by(tquint)
  tabstat y_cass  [aw=poid], stat(median) by(tquint)

*--- Input Use ---*

** Fertilizer **

use data\cc_epm01_remoteness_index.dta, clear
  keep idmen poids hhmemb pop rindex time* rdec-tremote pnpk pthb codcom
  sort idmen
  save data\temp.dta, replace
  use data\agricb.dta, clear
  rename s1lb1q3 superficie
  keep idmen codetany s1lb3* superficie
  for any 4 5 7a 7b 7e 7d 8: rename s1lb3qX qX
  drop s1l*
  gen byte chemfert = 100 * (q4 >0 | q5>0)
  gen byte orgfert  = 100 * (q7a>0)
  gen byte pest     = 100 * (q8 >0)
egen qchemfert = rsum(q4 q5)
replace qchemfert = qchemfert / superficie

gen qorgfert = q7b * q7a / superficie
replace qorgfert = q7a / superficie if q7b == 4
recode qorgfert .=0

gen vpest = q8 / superficie
recode vpest .=0

qui for var qche-qor: sum X, detail \\ replace X = r(p99) if X>r(p99)

sort idmen
merge idmen using data\temp.dta

tab _m
keep if _m == 3
drop _m

replace vpest = 150000 if vpest > 150000

* Average quantity/value... *

tabstat qchem [aw=poid], stat(mean) by(tquint)
tabstat qorg [aw=poid], stat(mean) by(tquint)
tabstat vpest [aw=poid], stat(mean) by(tquint)

collapse (sum) qchemfert qorgfert vpest, by(idmen tquint poid)

gen use_chem = 100*(qchem > 0)
gen use_org = 100*(qorg > 0)
gen use_pest = 100*(vpest > 0)

* Percent of HH using... *

tabstat use_chem [aw=poid], stat(mean) by(tquint)
tabstat use_org [aw=poid], stat(mean) by(tquint)
tabstat use_pest [aw=poid], stat(mean) by(tquint)

***********
** Labor **
***********

use data\cc_epm01_remoteness_index.dta, clear
keep idmen poids tquint
sort idmen
save data\temp.dta, replace
use d:\data\madagascar\epm2001\data\agricb.dta, clear
rename s11b2q1  lab_hh
rename s11b2q2a  lab_ass
rename s11b2q2b  lab_hlp
rename s11b1q7b  code
keep idmen  code  lab*
sort idmen
merge idmen using data\temp.dta
tab  _m
keep if  _m == 3
drop  _m
drop if  code == 3
recode lab_hh .=0
egen lab_tot = rsum(lab_hh lab_a lab_hl)
gen byte total = 1
    * All crops... *
    table  tqu [aw=poid],  c(mean  lab_tot  mean  lab_hh  mean  lab_ass  mean  lab_hlp)  f(%5.2f)
table  tot [aw=poid],  c(mean  lab_tot  mean  lab_hh  mean  lab_ass  mean  lab_hlp)  f(%5.2f)
    * Rice... *
    table  tqu [aw=poid] if  code==101,  c(mean  lab_tot  mean  lab_hh  mean  lab_ass  mean  lab_hlp)  f(%5.2f)
table  tot [aw=poid] if  code==101,  c(mean  lab_tot  mean  lab_hh  mean  lab_ass  mean  lab_hlp)  f(%5.2f)

*--------*
*-- Prices --*
*--------*
use d:\dstifel\ifpri_s02\data\cc_epm01 remoteness_index.dta, clear
rename s0qb8  far
rename s0qb9  fiv
keep codcom*  far  fiv  pop  poids  wagem_mu  ppaddy_mu  tquint
    ** Replace missing with fivondronana mean **
preserve
dropdup codcom
collapse (p50) fvwagem=wagem fvppaddy=ppaddy [aw=pop], by(far fiv) fast
compress
sort far fiv
save data\temp.dta, replace
restore
sort far fiv
merge far fiv using data\temp.dta
tab _m
drop _m
replace wagem = fvwagem if wagem == .
replace ppadd = fvppaddy if ppadd == .
tabstat wagem_mu [aw=poid], stat(mean) by(tquint)
tabstat ppaddy_m [aw=poid], stat(mean) by(tquint)

*-----------------------------------------------------------------------*
* Table 4 of Stifel and Minten (2008). Stata programs to estimate reduced form models of rice yields in Madagascar. These models include remoteness variables as explanatory variables.
* The general idea is to isolate the effects of remoteness through transaction costs from those through insecurity, price risk, and plot size. In other words three reduced-form models are estimated...
* 1. yield = f(cost)
* 2. yield = f(cost, other stuff)
* 3. yield = f(cost, insecurity, price risk, plot size, other stuff)
*-----------------------------------------------------------------------*
clear
set memory 50m
use data\epm_2001_rice.dta

*---------------------------------------------------------------*
*-- Define output and inputs in per are units --*
*-- i.e. "output" is really "yield" --*
*---------------------------------------------------------------*

order output cult_area hhlabor labor semenc animtr tractor fertilizer

for var output hhlabor-fertilizer: replace X = X/cult_area
    sum output-fertilizer [aw=poid], sep(0)
bysort tquint: sum output-fertilizer [aw=poid], sep(0)

qui for var output time_cup densite: gen lnX = ln(X)
rename lntime_cup lntime

for var cult_area hhlabor labor semenc animtr tractor fertilizer: count if X == 0 \ count if X > 0
#delimit ;

order output
    cult_area
    std_pprice
    zone_rouge dist_route
    hhlabor labor semenc animtr tractor fertilizer
    riz_hill terrace slopevar soil_clay soil_mud
dsoil_sediment dsoil_vol-dsoil_alluvial
    mean_elev mean_prec mean_temp
    std_elev std_prec std_temp
    flood drought
    irrigation
    densite
    p_acc_land p_acc_plzreb p_acc_lab p_acc_equip p_acc_cred
    agr_prin_head lvstkl_5 lvstk6_
    headfem headage
    kids6_14 female male
    prim-psec clinic fady2-fady4
    wagef wagem prseed pthb przebu prtrct pnpk prur prorgfe
    agext title;
replace std_pprice = exp(std_pprice)
drop if miss==1

qui reg lncost agext cult_area-dist_route riz_hill-psec far2-far6 [aw=poid]
keep if e(sample)

xtile cost_quint = lncost [aw=poid], nq(5)
tabstat output [aw=poid], stat(mean median) by(cost_quint)
tab cost_quint [aw=poid], gen(cost)

replace cult_area = ln(cult_area)

summ lncost dist_route cult_area-zone_rouge riz_hill-psec far2-far6 [aw=poid], sep(0)

*-------------------------*
*-- Estimate the models --*
*-------------------------*

reg lnoutput lncost dist_route
reg lnoutput lncost dist_route riz_hill-psec far2-far6

qui test riz_hill, notest
qui test terrace , accum notest
qui test slopevar, accum notest
qui test soil_cla, accum notest
qui test soil_mud, accum notest
qui test dsoil_se, accum notest
qui test dsoil_vo, accum notest
qui test dsoil_al, accum notest
qui test mean_ele, accum notest
qui test mean_pre, accum notest
qui test mean_tem, accum notest
qui test std_elev, accum notest
qui test std_prec, accum notest
test std_temp, accum

*-- Full model (with plot size) --*

reg lnoutput lnecost dist_route cult_area-zone_rouge riz_hill-psec far2-far6

qui test riz_hill, notest
qui test terrace, accum notest
qui test slopevar, accum notest
qui test soil_cla, accum notest
qui test soil_mud, accum notest
qui test dsoil_se, accum notest
qui test dsoil_vo, accum notest
qui test dsoil_al, accum notest
qui test mean_ele, accum notest
qui test mean_pre, accum notest
qui test mean_tem, accum notest
qui test std_elev, accum notest
qui test std_prec, accum notest
test std_temp, accum

test lncost, notest
test cult_area, accum notest
test std_pprice, accum notest
test zone_rouge, accum

test cult_area, notest
test std_pprice, accum notest
test zone_rouge, accum

*-----------------------------------------------------------------------*
*-----------------------------------------------------------------------*
*-----------------------------------------------------------------------*
*-----------------------------------------------------------------------*
*-----------------------------------------------------------------------*
**  Table 5 of Stifel and Minten (2008).**
* Stata program to estimate reduced form input demand models for
  rice production in Madagascar using the 2001 EPM and Commune
  Census.
*-----------------------------------------------------------------------*
use data\epm_2001_rice.dta

order output cult_area hhlabor labor semenc animtr tractor fertilizer

for var output hhlabor-fertilizer: replace X = X/cult_area
    sum output-fertilizer [aw=poid], sep(0)
bysort tquint: sum output-fertilizer [aw=poid], sep(0)

qui for var output time_cup: gen lnX = ln(X)
rename lntime_cup lntime

for var cult_area hhlabor labor semenc animtr tractor fertilizer: count if X == 0 \ count if X > 0

#delimit ;
order output
cult_area hhlabor labor semenc animtr tractor fertilizer
riz_hill terrace slopevar soil_clay soil_mud
dsoil_sediment-dsoil_alluvial
mean_elev mean_prec mean_temp
std_elev std_prec std_temp
floöd drought
irrigation
densite p_acc_land p_acc_plzeb p_acc_lab p_accEquip p_acc_cred
agr_prin_head lvstkl_5 lvstk6_
headfem headage
kids6_14 female male
prim-psec
wagf wagem prseed pthb przebu prtrct pnpk prur prorgfe
agext title;
#delimit cr;

gen use_tract = tractor > 0
gen use_fert = fertilizer > 0

global remote lncost dist_route std_pprice zone_rouge
global rhs riz_hill-dsoil_sediment dsoil_vol-psec far2-far6
summ lncost, detail
drop if lncost < 5

*-------------------------*
*-- Estimate the models --*
*-------------------------*

*-- Check for outliers --*
summ    hhlabor, detail
replace hhlabor = . if hhlabor < 0.01 | hhlabor > 30

summ    labor, detail
replace labor = . if labor > 20

summ    animtr, detail
replace animtr = . if animtr >= 50

*-- Estimate models --*
regress hhlabor   $remote $rhs
tobit   labor     $remote $rhs, ll
tobit   animtr    $remote $rhs, ll
dprobit use_tract $remote $rhs
dprobit use_fert  $remote $rhs

*  ***************************************************************
*                         Table 6 of Stifel and Minten (2008).
*                         Stata program to estimate rice yield production functions
*                         using EPM 2001 data and 2001 Commune Census
*                         (transportation costs, remoteness)
*                         - Translog production technology
*                         - Calculates elasticities of shifters across
* travel-time quintiles
*
*
******************************************************************************
clear
set seed 100
set memory 20m
set matsize 500

use data\epm_2001_rice.dta, replace
  sort codcom
  merge codcom using data\soil.dta
  tab _m
  drop if _m == 2
  drop _m
  sort codcom
  merge codcom using data\climate.dta
  tab _m
  list codcom if _m==1
  drop if _m==2
  drop _m

sum output cult_area yield
keep if pt_yield ~= .
replace cult_area = pt_area
replace output = pt_area * pt_yield

capture drop slopevar
capture drop terrace
gen slopevar = mod==1 | steep==1
replace riz_hill = 1 if topo > 1
gen terrace = topo == 3
replace riz_hill = 0 if terrace == 1

qui tab codcom, gen(com)                      // 150 communes

*--------------------------------------------------------------------------*
*-- Define output and inputs in per are units --*
*--------------------------------------------------------------------------*
--- i.e. "output" is really "yield"  --*
*-------------------------------------------------------------*

for var output hhlabor-fertilizer: replace X = X/cult_area
    sum output-fertilizer [aw=poid], sep(0)
bysort tquint: sum output-fertilizer [aw=poid], sep(0)

qui for var output time_cup: gen lnX = ln(X)
rename lntime_cup lntime

*-------------------------------------------------------------*
*-- Instead of adding 1 to all values, --*                *-- add 1/10th of minimum positive --*          *-- to the zero values only --*               *-------------------------------------------------------------*

for var cult_area hhlabor labor semenc animtr tractor fertilizer: sum X if X > 0
*-- Add 0.01 to the zero values --*

for var cult_area hhlabor labor semenc animtr tractor fertilizer: replace X = 0.01 if X == 0

#delimit ;
order codcom
output
cult_area hhlabor labor semenc animtr tractor fertilizer
riz_hill terrace slopevar soil_clay soil_mud
dsoIl_sediment dsoil_volcanic-dsoil_alluvial
mean_elev mean_prec mean_temp
std_elev std_prec std_temp
irrigration drought flOod dr_irr title
zone_rouge dist_route plant_late flood
ag_r_prin_head lvstk1_5 lvstk6_
headfem headage
prim-psec clinic fady2-fady4
kids6_14 female male
   p_acc_land-p_acc_cred credit_max agext;
#delimit cr;
*-- Rename the inputs to facilitate estimation and post-estimation commands --*

```
#delimit ;
for var cult_area hhlabor labor semenc animtr tractor fertilizer \ num 1/7: rename X inputY \
gen inY = ln(inputY);
#delimit cr;
local  i = 1
local  p = 1
while `i' <= 7 {       
local  j = `i'
while `j' <= 7 {         
qui gen in`i'`j' = in`i' * in`j'
local j = `j' + 1
local p = `p' + 1       }
local i = `i' + 1
}
```

*---------------------------------------------------*
*-- ivreg to test for validity of the instruments --*
*---------------------------------------------------*

```
ivreg2 lnoutput  in1 in11-in77 riz_hill-dr_irr (in2-in7 = title dist_route-male), first

** Note: The results that appear in the paper are the means of the
** bootstrapped estimates. So they differ from those above.

*-----------------------------*
*-- Define the bootstrap program --*
*-----------------------------*

capture program drop elast
program define elast

*-- First-stage instrumenting equations --*
```
** Inputs **

```stata
#delimit ;
qui for num 2/7:              regress inX riz_hill-soil_mud irrigation-title dist_route-male com2-com150 \
    predict uhatX, resid;
#delimit cr;

*-- Estimate the model --*

if `1' == 1 {
    reg lnoutput in1-in77 riz_hill-dr_irr uhat2-uhat7
    matrix a = e(b)
    matrix betas = a[1,"riz_hill".."dr_irr"]
    matrix b = a["riz_hill".."dr_irr","riz_hill".."dr_irr"]
    matrix a = vecdiag(b)
    matrix b = diag(a)
    matsqrt b a
    matrix b = syminv(a)
    matrix tstat = b * betas
    matrix scoeff = betas', tstat
    scalar R2 = e(r2)
}       qui reg lnoutput in1-in77 riz_hill-dr_irr uhat2-uhat7

*-- Inputs - Elasticities --*

local  i = 1
while `i' <= 7 {
    qui gen e`i' = _b[ln`i''] + (2*_b[ln`i'``i'']*ln`i'
local  j = `i' + 1
while `j' <= 7 {
    qui replace e`i' = e`i' + (_b[ln`i'``j']*ln`j')
    local j = `j' + 1
}      local k = `i' - 1
while `k' >= 1 {
```
qui replace e`i' = e`i' + (_b[in`k'`i']*in`k')
local k = `k' - 1
}
local i = `i' + 1
}

** By remoteness quintile **
local i = 1
while `i' <= 5 {
    qui for num 1/7: gen q`i'_eX = eX if time`i' ==1
    local i = `i' + 1
}

*** Inputs - Marginal Products ***
local i = 1
while `i' <= 7 {
    qui gen mp`i' = e`i' * output / input`i'
    local i = `i' + 1
}

** By remoteness quintile **
local i = 1
while `i' <= 5 {
    qui for num 1/7: gen q`i'_mpX = mpX if time`i' ==1
    local i = `i' + 1
}

*** Shifters - elasticities ***
#delimit ;
qui for var riz_hill-std_temp flood:
gen X_e = _b[X] * X;
#delimit cr;

qui gen  irrig_e = (_b[irrigation] + _b[dr_irr]*drought) *irrigation
qui gen drought_e = (_b[drought] + _b[dr_irr]*irrigation)*drought
order riz_hill_e-std_temp_e irrig_e flood_e drought_e

** By remoteness quintile **

    local  i = 1
    while `i' <= 5 {        qui for var riz_hill_e-drought_e: gen q`i'_X = X if time`i' ==1        local i = `i' + 1      }

*-- Find the average elasticities --*

    collapse (mean) e1-q5_mp7 riz_hill_e-drought_e q1_riz_hill_e-q5_drought_e [aw=poid], fast

    keep lnoutput in* riz_hill-agext com* lntime lndist time_km* output poid time*
    drop if time_km >0.5

*---------------------------*

*-- Bootstrap full sample --*

*---------------------------*

*-- First pass through the data --*

    preserve
    elast 1
    qui save data\temp.dta, replace
    restore

*-- Remaining passes through the data --*

    local i   = 2
    local max = 200
    while `i' <= `max' {
        if `i'/10 == int(`i'/10) {        display `i'/10
            display `i'      }
    }
preserve
bsample
elast 0
append using data\temp.dta
qui save data\temp.dta, replace
restore
local i = `i' + 1
}

use data\temp.dta, clear
qui sum el
scalar mu = r(mean)
scalar tstat = r(mean) / r(sd)
matrix output = ( mu, tstat )
matrix colnames output = "Mean" "tstat"
matrix rownames output = "el"
capture program drop dispres
program define dispres
    qui sum `1'
    scalar mu = r(mean)
    scalar tstat = r(mean) / r(sd)
    matrix out = ( mu, tstat )
    matrix rownames out = "`1""
    matrix output= output out
end
qui for var e2-q5_drought_e: dispres X
drop _all
svmat output, names(col)
matrix list output
matrix list scoeff
display R2
*___________________________________________________________________________________*
Table 7 of Stifel and Minten (2008).

Stata file to examine the relationship between remoteness and share of land dedicated to various crops using the 2001 commune census and the 2001 EPM.

```stata
#delimit cr;
*-----------------------*
*-- Load commune data --*
*-----------------------*

use data/cc_epm01_remoteness_index.dta, clear
    rename s0qb8 far
    keep idmen poids far hhmemb pop rindex time_cup rdec-cremote
    sort idmen
    save data/temp.dta, replace

*--------------------------*
*-- Load EPM Region data --*
*--------------------------*

use data/total_exp.dta, clear
    rename s0qb8 far
    rename s0qb12 milieu
    keep idmen-idfivo far milieu
    sort idmen
    merge idmen using data/temp.dta
    tab _m
    keep if _m == 3
    drop _m
    sort idmen
    save data/temp.dta, replace

*------------------------*
*-- Load EPM Section B --*
*------------------------*
```

use data\agricb.dta, clear  
rename s11blq3  area  
rename s11blq7b code  

gen byte rice  = code < 200  
gen byte cass  = code > 300 & code < 305  
gen byte staple  = code < 400  
gen byte pulses  = code >=400 & code < 500  
gen byte veggie  = code >=500 & code < 600  
gen byte fruit  = code >=600 & code < 700  
gen byte index  = code >=700  
for var rice-index: replace X = X * area  
collapse (sum) area rice-index, by(idmen) fast  
for var rice-index: gen sh_X = 100*X/area  

** For all non-industrial & export crops **  
gen area2 = area - index  
for var rice-fruit: gen sh2_X = 100*X/area2  

sort idmen  
merge idmen using data\temp.dta  
tab _m  
keep if _m == 3  
drop _m  
gen byte total = 1  

*---------------------------------*  
* Shares by Remoteness Quintile -*  
*---------------------------------*  

table tquint [aw=poid], c(mean sh_st mean sh_rice mean sh_cass) format(%5.2f) 
table tquint [aw=poid], c(mean sh_pu mean sh_ve mean sh_fr mean sh_in) format(%5.2f) 

table total [aw=poid], c(mean sh_st mean sh_rice mean sh_cass) format(%5.2f) 
table total [aw=poid], c(mean sh_pu mean sh_ve mean sh_fr mean sh_in) format(%5.2f)
*-- Non-industrial & export crops... --*

```
table tquint [aw=poid], c(mean sh2_st mean sh2_rice mean sh2_cass) format(%5.2f)
table tquint [aw=poid], c(mean sh2_pu mean sh2_ve mean sh2_fr) format(%5.2f)

table total [aw=poid], c(mean sh2_st mean sh2_rice mean sh2_cass) format(%5.2f)
table total [aw=poid], c(mean sh2_pu mean sh2_ve mean sh2_fr) format(%5.2f)
```