UNIVERSITÉ DE FRIBOURG
UNIVERSITÄT FREIBURG

WORKING PAPERS SES

Better roads, better off? Evidence on improving roads in Tanzania

Christelle Dumas
and
Ximena Játiva

N. 518
IX.2020

Faculté des sciences économiques et sociales et du management
Wirtschafts- und sozialwissenschaftliche Fakultät

Better roads, better off? Evidence on improving roads in Tanzania *

Christelle Dumas[a], Ximena Játiva[b]

[a] *Corresponding author. University of Fribourg, Bd de Perolles 90, 1700 Fribourg, Switzerland. Email: christelle.dumas@unifr.ch*

[b] *University of Fribourg, Bd de Perolles 90, 1700 Fribourg, Switzerland. Email: ximena.jativa@unifr.ch*

Abstract

Spatial isolation is considered as one of the main determinants of poverty. Therefore, many transport investments are undertaken with a stated objective of poverty reduction. In our paper, we evaluate a Tanzanian program that rehabilitated 2500km of major roads between 2008 and 2013. We deal with endogenous placement issues with a household fixed-effect strategy combined with a propensity score matching. Contrary to most studies, we find damaging effects of the road on the rural population: the price of the main product (rice) decreases, they reduce rice production and reallocate labor away from farm but opportunities of off-farm work are scarce. This results in depressed wages and households declare a lower satisfaction. This is consistent with a situation where rural households face an increased competition due to lower transportation costs.

Keywords: Roads; Poverty; Rural households; Africa

JEL Codes : O12; O13; J43; O15; O18.

1. Introduction

Eighty percent of the extreme poor and seventy five percent of the moderate poor live in rural areas (Castaneda et al., 2016). Therefore, isolation is considered as one of the main contributors to poverty. According to the World Bank's Rural Access Index, one billion rural dwellers do not have reliable transport access and only 34 percent of the rural Sub-Saharan African population live within two kilometers of an all-season road. Spatial isolation can impose serious constraints on agricultural production, access to health, education and work opportunities. As a consequence, infrastructure investment has been considered for a long time as a key aspect of development policy.[1] However, isolation also limits the competition from external producers and therefore might allow the subsistence of non-competitive rural farmers. The construction

*We would like to thank Stefan Bauernschuster, Lorenzo Casaburi, Denis Cogneau, Martin Huber, Sebastian Krautheim, Julien Labonne, Sylvie Lambert, Karen Macours, Ferdinand Rauch, Dominique van de Walle, Oliver Vanden Eynde, Joachim De Weerdt for helpful comments and suggestions. We also thank seminar participants at the University of Fribourg, Paris School of Economics, LICOS, Passau, Heidelberg, Cergy-Pontoise, Graduate Institute and participants of the following conferences: Swiss Development Economics Network, German Development and Economics Policy (Berlin), EEA-ESEM (Manchester), ESPE (Bath), Nordic Development Economics (Copenhague), DIAL (Paris), NOVAfrica (Lisbon), NCID (Madrid).

[1] The World Bank Transport Business Strategy for 2008-2012 stated: "One of the best ways to promote rural development is to ensure good accessibility to growing and competitive urban markets." (The World Bank, 2008, p. 3)

of transport infrastructure could well have damaging impacts on small-scale farm producers, at least in the short run on vulnerable populations. This paper evaluates the consequences of a road rehabilitation program in Tanzania on rural households and covers a wide array of possible adjustments: production and consumption decisions, market participation, labor choices, types of activities, assets, welfare, migration and prices. We find that rural households actually suffer from the program.

A growing literature seeks to identify the consequences of improved access to transportation infrastructure, be they railroads, highways or rural roads. Most of the literature concludes to positive effects of a better access. Better connectivity to other locations has been shown to reduce transport costs (Khandker et al., 2009; Donaldson, 2018), lower prices and increase the availability of non-local goods (Aggarwal, 2018; Khandker et al., 2009), increase the use of productivity-enhancing inputs such as fertilizers or hybrid seeds (Shamdasani, 2016; Shrestha, 2019), increase agricultural production (Khandker et al., 2009), reallocate labor from agriculture to urban activities (Asher and Novosad, 2018), facilitate agricultural specialization and market-orientation of farm activity (Qin and Zhang, 2016; Shamdasani, 2016), and increase income (Donaldson, 2018; Qin and Zhang, 2016). Only a few find no or very limited effects (Banerjee et al., 2020). These results are consistent with the combination of two previous strands of literature: the analysis of agricultural households in presence of market imperfections (Singh et al., 1986; de Janvry et al., 1991) and the quantification of transaction costs (Renkow et al., 2004; Jacoby and Minten, 2009; Minten and Kyle, 1999; Harrison et al., 2014). Additional effects identified by the infrastructure literature are the following: younger children seem to spend more time enrolled in school (Aggarwal, 2018; Adukia et al., 2017; Khandker et al., 2009), local markets develop (Mu and van de Walle, 2011), and in the case of major roads or highways, firms are more likely to export (Volpe Martincus et al., 2017; Storeygard, 2016) and manufacturing activity is stimulated (Ghani et al., 2016). In addition, the few papers interested in identifying heterogeneous impacts of roads identify stronger (positive) effects for the poor population (Khandker et al., 2009; Lokshin and Yemtsov, 2005).

However, from a theoretical point of view, the reduction in transaction costs modifies the whole vector of relatice prices. Given that consumers differ in their consumption bundle and in their sector of activity, the whole population is not affected homogeneously (Porto, 2006). In particular, rural households may well suffer from increased price competition instead of benefiting from greater market access. If the sectoral reallocation was high, then households could easily cope with such changes happening in their economic environment but a host of evidence shows reallocation is particularly slow in developing countries.[2]

In this paper, we evaluate the effects of a program that by 2013 upgraded 2'500km of major roads in Tanzania on: agricultural decisions, market participation for labor and products, prices

[2]See for instance, Revenga (1997); Harrison and Hanson (1999); Attanasio et al. (2004); Currie and Harrison (1997); Topalova (2004). Congruent with this conclusion is the evidence, in some countries, that spatial mobility is limited (Munshi and Rosenzweig, 2016).

and wages, migration, consumption and individuals' life satisfaction. We combine a household panel data with a GIS dataset listing changes in road infrastructure in the country since 2008. This allows us to identify the effect of roads controlling for household fixed effects. In order to guarantee that the within household change would be similar among treated and non-treated households in absence of the treatment, we build a comparison group of the treated households by propensity score matching. We then assess in greater detail changes in employment using two rounds of the Labor Force Survey and a similar methodology. This allows us to show that rural households, instead of reaching new markets, actually face a decrease in the price of rice and have to reduce their farm activity, increase off-farm work and see their life satisfaction decrease. We also provide evidence that, consistent with the increase in market labor supply, wages are depressed, so that the whole population is affected. This result is very different from the ones obtained previously in the literature and is in our view extremely important. If there are negative consequences of roads on poor rural households, even if they are transitory, then we need to be aware of them in order to design safety nets that would help these households to smooth transition to their new economic environment. We also provide a series of robustness tests to assess the validity of our identification strategy. Among others, we show that the common trend assumptions are not rejected; we run placebo tests with roads that were identified as potentially rehabilitated but were not improved in the same timeframe; and we use alternative control and treatment definitions.

Our paper is a valuable addition to the existing roads literature for the following reasons. First, the existing research is either on the impact of rural roads, i.e. roads that reach villages, or of railroads and has focused mainly on Asia or the United States. Little has been done on major roads, to the exception of Storeygard (2016), Ghani et al. (2016), Shrestha (2019) and few papers have focused on Sub-Saharan Africa (Casaburi et al., 2013; Burgess et al., 2015; Gachassin et al., 2010).[3] There are few reasons to expect similar consequences from the improvement of rural roads and major roads. For instance, rural roads could well improve access to other villages and therefore improve the spatial integration of the labor market, while a major road would fail to have such consequences, simply because accessing the major road is too time-consuming. Covertly, main roads might trigger a larger reduction in the price of goods than rural roads. Insofar as the Tanzanian government decided to fight poverty using the instrument of major roads, it is important to assess whether this type of roads also fosters economic growth in rural areas. Second, compared to most papers in the transportation literature, we acknowledge the variety of adjustments households might implement in reaction to their new environment by using a wide range of outcomes rather than focusing on one dimension. Our results show this approach is meaningful: indeed, if we had focused on labor market participation, for which we find an increase, it would have been tempting to conclude to an improvement of the households' situation, while this is actually not the case. The ability to follow several indicators gives us the opportunity to provide a consistent story of the adjustments we observe. Last, to our knowledge,

[3]These papers address questions related to prices in rural markets, ethnic favoritism and poverty reduction.

this is the first paper that empirically identifies a negative effect of roads on some segment of the population.

The article is structured as follows. Section 2.2 describes the road improvements in Tanzania during the period 2008-2013 and section 3 presents the data. Section 4 describes the identification strategy and section 5 provides the results. Section 6 concludes the paper.

2. Context

2.1. Agriculture in Rural Tanzania

Over two-thirds of Tanzanians live in rural areas where agriculture serves as their main livelihood activity and accounts for 70% of their income. Livestock keeping and poultry production is the second largest source of income. Agriculture is the cornerstone of the Tanzanian economy, it employs 74% of the country's labor force and contributes to 31% of the GDP and 20% of the total annual export earnings, where the main export crops are cashew, tobacco, sugar, coffee and cotton. The vast majority of agricultural output (75%) is produced by small family farmers, producing mostly for home consumption and selling the excess in local markets. Tanzanian smallholders sell on average only 35% of their agricultural production. The main food crops are maize, cassava, rice, sorghum and bananas. Local and informal markets are the main selling and buying channel for smallholders for both agricultural inputs (99%) and outputs (98%) (Government of Tanzania, 2011; FAO, 2018, 2020).

Table 1: Main crops in Tanzania

	Harvest (kg)	Harvest (TSh)	Consumption (kg)
Maize	458.8	87776.1	351.9
Rice	136.1	40593.8	138.6
Beans	61.7	25517.4	89.6
Millet	66.2	12149.7	10.9
Cotton	40.7	18636.6	
Groundnut	41.0	12430.9	29.9
Sweet potato	70.2	8172.1	112.6
Cassava	9.1	2084.3	100.8

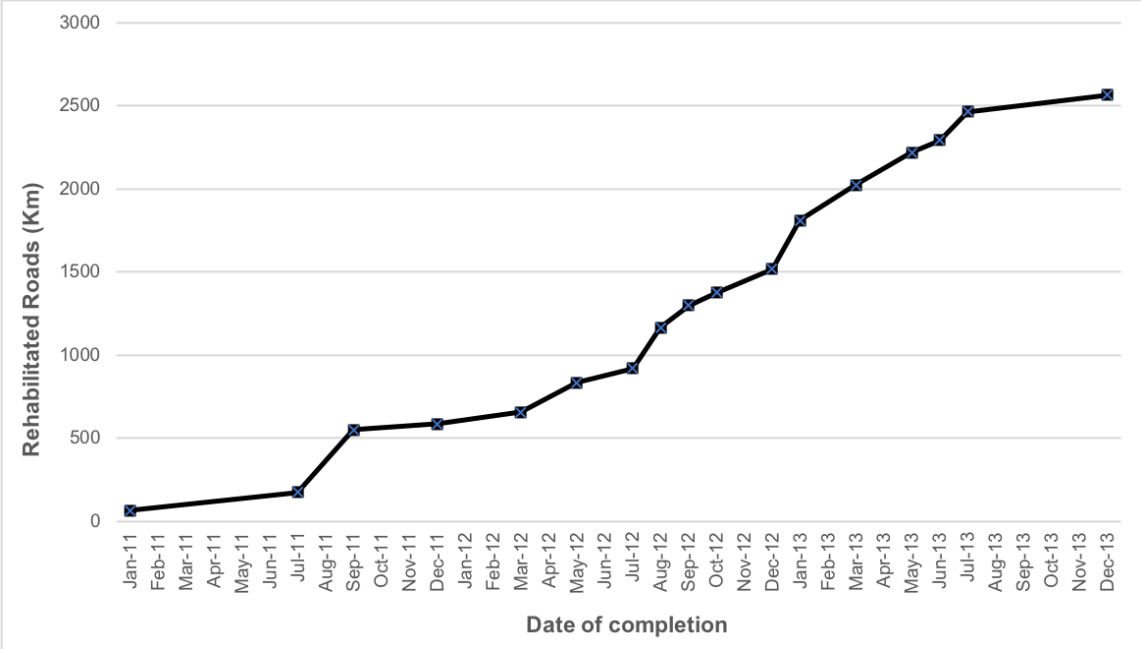
Note: This table is obtained from the LSMS 2008/09 and provides the harvest for the main crops, both in kilograms and in value. The measure for the consumption is obtained on a 7-days recall and then inflated to estimate the yearly consumption.

Table 1 provides descriptive statistics on agricultural production and consumption of rural households in Tanzania. This table is obtained from our dataset, described later. We see that maize is the most important crop, both in value and in kilograms. A large share but not all the production is consumed. The second crop is rice, and the consumption is of the same order of magnitude as production. All the other crops, even cash crops as cotton and groundnut,

generate comparatively less value than the cereals. According to Dercon (1998), rice has a higher return than maize but is also riskier. It is also a more appreciated good by households.

2.2. The road construction program

Figure 1: Roads Rehabilitated (km) 2008-2013

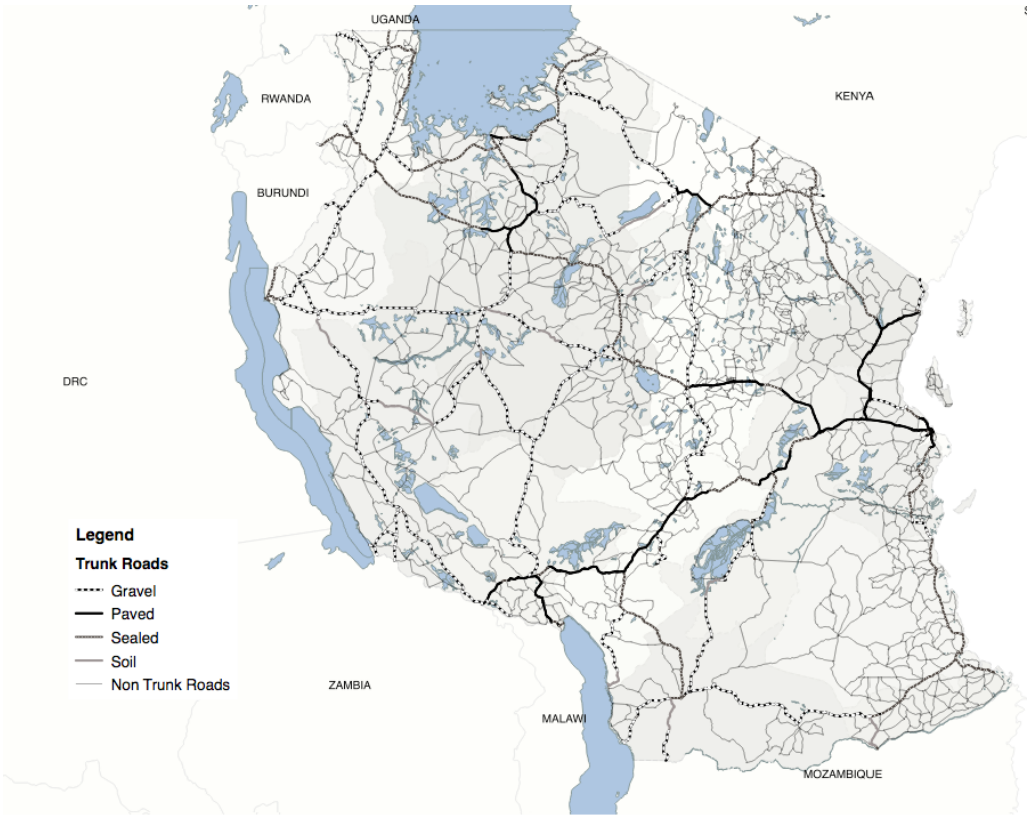


The road construction program takes place in a context where the Tanzanian infrastructure is considered as extremely poor: the country ranked 118th out of 134 economies in the infrastructure dimension of the World Economic Forum's Global Competitiveness Index in 2008-2009 (Schwab and Porter, 2008). This is due to both a low road density (96.5 meters per square kilometer compared to 296.95 in Kenya and Uganda) and to a poor condition of the road network: only 36.63% of the road network was paved or sealed and was classified as in "good or fair condition" (Government of Tanzania, 2008).

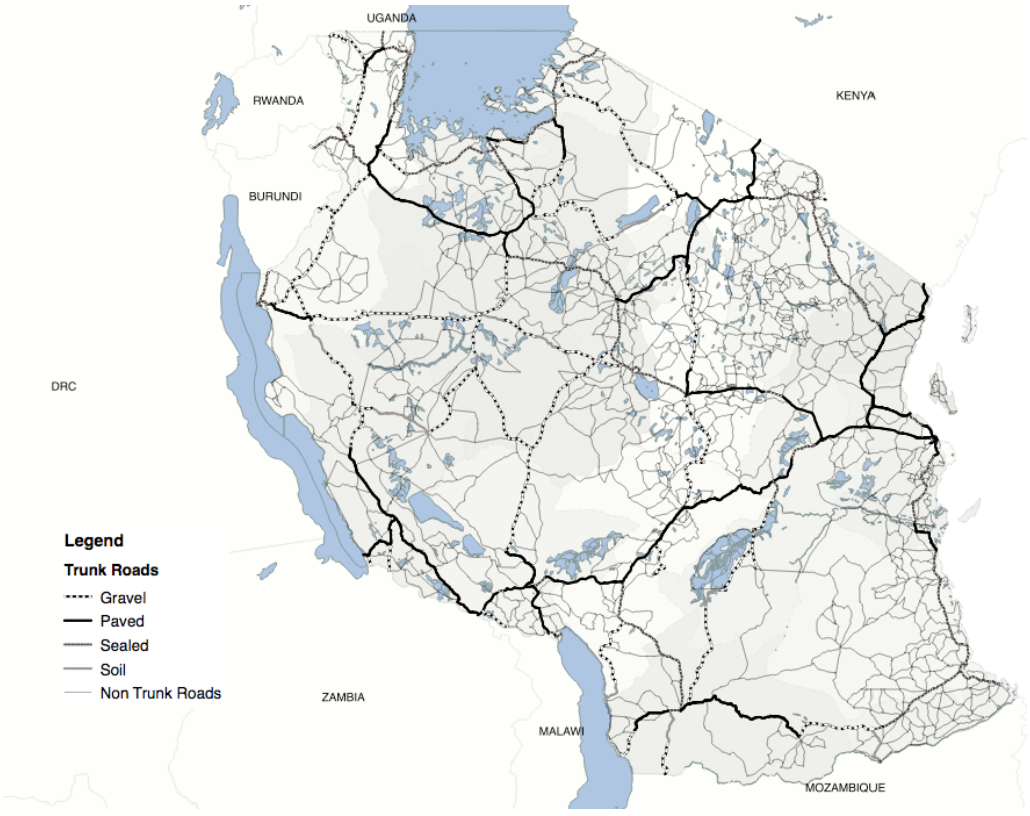
The 2005 National Strategy for Growth and the Reduction of Poverty in Tanzania identified the poor condition of the rural road network as one of the major impediments to development: insufficient and poor quality infrastructure hampers access to labor and input markets, hinders agricultural productivity, and generates significant losses for the rural poor. Therefore, the plan emphasized that investing in adequate road infrastructure had the potential to boost the rural economy, promote growth and bring down the level of rural poverty.

In this light, the Government of Tanzania adopted two important measures. First, the government launched the "10 Year Transport Sector Investment Programme", which identified nine road transport corridors (about 10'300 km of roads) that were key to enhance integration of regions within the country, as well as better connect the country to its neighbors. During the ten years of the program, all roads in the nine development corridors that were not in good condition (6'000 kms) were planned to be upgraded to paved standards. The first phase was scheduled to be implemented between 2007/2008 and 2012/2013. The Tanzanian National

Figure 2: Road Rehabilitation 2008-2013



(a) Road Network 2008



(b) Road Network 2013

Roads Agency (TANROADS) received in 2007 the responsibility to improve the Trunk and Regional Road Network.[4] The infrastructure works therefore started after 2008. The second phase was scheduled to start in 2012-13 and end in 2016-17. However, a review of recent data on public works show that it has actually been postponed.[5]

Three policy goals had been set: promote economic growth, improve economic wellbeing of the rural poor by enhancing production and market exchanges of both subsistence and cash crops, and take advantage of the geographic location of the country vis-à-vis of its landlocked neighboring countries.[6] The guidelines to determine the allocation of funds and prioritization of road projects were therefore the following: fully funded projects related to maintenance of roads in bad condition were given first priority. Second, road segments that would enhance exports and key food crop production had to be given preference. Third, all regional centers should be linked with paved roads, and district headquarters should be connected with all-weather roads of at least gravel standards. Last, the projects should open up all major population areas to modern economy and trade. We will proxy these guidelines by a set of geo-referenced variables.

Figure 1 shows that by the end of 2013, the road network showed significant improvements: 2'564 km of roads were paved. Figure 2 shows the state of the road network before (panel 2a) and after (panel 2b) the infrastructure works. Prior to the rehabilitation program the majority of the road network was either graveled (49.7%) or sealed (30.5%). Only 11% of the total main network was paved and in good condition, whereas at the end of 2013 this percentage rose to 29.4%. As a result, Tanzania moved from the 118th to the 90th place of the infrastructure dimension of the Global Competitiveness Index in 2016-2017 (Schwab and Sala-i Martin, 2016).

The objective of this paper is to evaluate the consequences of this major roads' improvement program on the reduction of poverty and improvement of opportunities for Tanzanian rural households.

3. Data

We mostly use three sources of data: the detailed geolocated information on roads obtained from TANROADS, a household panel survey and two rounds of the Labor Force Survey.

3.1. TANROADS Data

The Tanzanian Roads Agency has collected a geo-referenced dataset on road condition and rehabilitation, which we were able to access. It covers the entire road network and contains primary and secondary roads, as well as information about their length, surface condition (paved, sealed, gravel and soil type), infrastructure projects, type of infrastructure project and date of

[4]This network is estimated to be 33'891 km of the 86'472 km total classified road network in Mainland Tanzania (Government of Tanzania, 2007). The remaining network of Urban, District, and Feeder Roads is managed by the Local Government Authorities.

[5]Tanzania National Roads Agency, http://tanroads.go.tz

[6]Tanzania is an important transit gateway for six landlocked countries of Southern and Central Africa: Malawi, Zambia, DRC, Burundi, Rwanda, and Uganda.

completion. For the purpose of this study we are mostly interested in primary roads for which rehabilitation works started and were completed between 2008-2013.

3.2. LSMS-ISA Data

The LSMS-ISA (Living Standards Measurement Study - Integrated Surveys on Agriculture) data for Tanzania is a panel of three rounds of a nationally representative household sample, collected by the Tanzanian National Bureau of Statistics. The survey is made of 3'265 households in the first wave, clustered in 410 enumeration areas across mainland Tanzania. 258 of these enumeration areas are rural and included in the analysis. The three first rounds were collected in 2008-2009, 2010-2011 and 2012-2013.[7]

These data are particularly suitable for our study. They have very detailed information on the household, agricultural and community dimensions and they are geo-located. The wide variety of variables collected in the survey is important for the assessment of the consequences of the road rehabilitation since it may affect substantially the economic environment of the households.

Table 2 summarizes our key variables, by reference period on which they are collected. Regarding the agricultural activities, we use the information for the previous long rainy season (roughly the period covering January to August). The survey collects all data related to inputs (market inputs, labor) and outputs on each of the household plots. We aggregate this information at the household-year level. Information pertaining to market labor supply and household consumption are collected for a shorter reference period (typically for the week before the survey). Prices for food and standard goods are collected at the village level at the time of the survey. All variables used in the study are described in Appendix A.1. The difference in recall durations will impose constraints on the estimation.

3.3. LFS Data

We also use the 2006 and 2014 Labor Force Surveys (LFS) to further evaluate the effects of the program on labor market participation. These data will allow us to check the consistency of the results with the LSMS, as well as provide details on the employment changes due to the road rehabilitation. The LFS is a nationally representative survey that gathered data on 16'445 households in 326 wards in 2006 and 11'473 in 363 wards in 2014. Each village contains on average 28 households and approximately 52% of the sample is located in rural areas. We are able to geo-locate these villages at the ward level.[8]

[7]The first round was conducted from October 2008 to September 2009, the second from October 2010 to September 2011 and the third from October 2012 to November 2013. The last round, collected in 2014-15, draws 3'352 new households clustered in 419 enumeration areas, so as to ensure the representativity of the sample. Roughly a third of the original sample is re-interviewed but the data for this sub-sample is not yet released and in any case, would likely be too small to achieve identification. Several key variables are still missing from the release of the new dataset, which prevents us from extending our analysis to the fourth round, even as a new cross-section.

[8]To do so, we use ward level boundary polygons from the 2002 and 2012 Tanzania Population Census obtained from the Tanzania Bureau of Statistics. A ward is the smallest administrative unit above the village in Tanzania. We define centroids for all wards to obtain their geographic coordinates using Q-GIS.

Table 2: Outcome Variables in the LSMS, by recall period

	Labor Market	**Product Market**	**Welfare**
Household	*Annual information* Off-farm labor Hired-in labor Household on-farm labor *Last 7 days* Household wage labor	*Annual information* Harvest (kg, TSh)+ Sales (kg, TSh)+ *Last 7 days* Consumption(kg)+	*Annual information* Migration *Last 7 days* Subjective Welfare Durable assets Productive assets
Community		*Last 7 days* Prices	

+ By crop: rice, maize

3.4. Additional Data

We collected a set of geographic measures from multiple sources, in order to proxy for the main drivers of the road program. Table A1 presents the list of variables created as well as their corresponding sources. We geo-referenced major cities with more than 20'000 habitants according to the 2002 Tanzania Population Census, District Headquarters and Major Border Crossings and we obtained geo-referenced 2010 population density and IFPRI standardized Agroecological zones, both at 100mx100m resolution.

We also build variables such as the labor market activity and education level in the LFS 2006 at the ward/district level that we then merge with the other datasets.[9] For example, to assess labor market activity, we compute the share of households in the ward who have at least one member providing wage work.

Finally, the data were also matched with rainfall data,[10] to control for productivity shocks. We check that the treatment does not correlate with shocks. Given that there is no such correlation, and that the estimates are the same with and without control for shocks, we simply provide the estimates without this additional control.

4. Estimation Strategy

We start describing the identification strategy for the LSMS panel and then explain how we adapt it to the LFS data.

4.1. Defining the treatment: timing

In 2008, the program had not started and we therefore have a pre-treatment period. A small number of roads were rehabilitated before the second round of the panel and a much

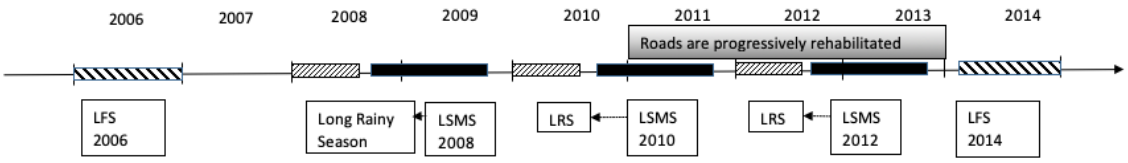
[9] We merge information at the ward level when the same ward can be found in the LFS 2006, if not we use the district-average information.

[10] Rainfall estimates are given by squares of roughly 10km x 10km. These data were obtained from `http://iridl.ldeo.columbia.edu/SOURCES/.NOAA/.NCEP/.CPC/.FEWS/.Africa/.DAILY/.ARC2/.daily/.est_prcp/datafiles.html`. We compute the deviation from the mean in the area, divided by the standard deviation in the area using years 2001-2013. We construct two variables using rainfall from January to December of the current and previous year of the LSMS wave.

larger number of roads were rehabilitated before the third round of the panel.

However, the duration of the survey for each round exceeded one year and therefore the actual treatment depends on the exact date of interview of the household. The treatment is defined as such: for outcomes that are measured on a yearly basis, the road has to be finished at least one year before the interview; and for outcomes that are measured on a lower-period basis, the road has to be finished at least a month before the interview. Figure 3 shows the timeline of the road rehabilitation program, the dates where the data collection took place and the dates of each long rainy season (LRS). Some additional roads were rehabilitated after 2014, but we only use them for placebo tests. It is clear from the figure that no household will be considered as treated in the 2010-2011 LSMS data collection for agricultural outcomes and that some (actually only a limited number of them) may be considered as treated for short-recall period outcomes.

Figure 3: Timeline of the surveys



4.2. Estimation in the panel: Household fixed effect model with propensity score matching

The second difficult point in defining the treatment is to clarify who is treated and who is not when a road is rehabilitated. There is probably a continuous effect depending on the distance of the household to the road. However, households living further away from existing roads might substantially differ from the ones who are better connected. For this reason, we will implement three complementary strategies. First, we will select control households among the ones who live at a similar distance from a (non-rehabilitated) road than the treated households. Second, we will control for household fixed-effects in order to avoid bias arising from remaining latent differences between treated and non-treated households. The strategy is therefore extremely close to a difference-in-difference strategy except that we control for any unobserved household characteristics that are not time-varying. However, this does not guarantee that the evolution of these households would have been similar, in absence of the road rehabilitation. So, third, we implement a matching strategy to guarantee the comparability of treated and control households. The identifying assumption is that this control group would have had a similar evolution than the treated group in absence of the program (Heckman and Navarro-Lozano, 2004; Smith and Todd, 2005; Caliendo and Kopeinig, 2008). Compared to a control strategy, PSM does not impose linear functional form restrictions in the estimation of the conditional expectation of the outcome variables. In addition, even if the PSM cannot

take into account unobserved characteristics, this is a mild concern in our case given the use of household fixed effects.

Implementing a matching requires to have a 0/1 treatment and to use a distance threshold for defining the treatment. However, we can perform the exercise for many thresholds and will do so for all consecutive kilometers between 20km and 50km. We cannot use a lower threshold because the estimation would rely on a too low number of treated households. Given that for yearly outcomes, the number of treated households is reduced, we can only provide an assessment of the road program for treatments starting at 30km until 50km. In the main analysis, we use the as-the-crow-flies distance to define the treatment. However, we provide robustness tests and show that a distance defined in time to reach the main road gives similar conclusions.

4.2.1. Specification

Our specification is the following:

$$Y_{ivt} = \alpha + \delta_{1d} 1_{it}(t = 2012) Road_{ivtd} + \delta_{0d} 1_{it}(t = 2010) Road_{ivtd} + \theta Z_{ivt} + \phi_i + \beta_t + \epsilon_{ivt} \quad (1)$$

where Y_{ivt} is an outcome variable for household i in village v at time t. The outcome variables are listed in section 3. $Road_{ivtd}$ is a dummy indicating whether the household is located within a radius of d km of a treated road, d taking values from 20 to 50. Figure A1 shows the composition of the treated group, by distance to the road, for each of these thresholds d. β_t are the time fixed effects for each round of the survey, 2008, 2010 and 2012. Z_{ivt} is a vector of time varying household, farm and village characteristics. ϕ_i contains household fixed effects and ϵ_{ivt} is the error term (the standard errors are clustered at the village level). $1_{it}(t = 2012)$ takes the value one when the household is surveyed in 2012. As already pointed out, all households are pre-treatment in 2008, only a few of them are already treated in 2010 and the majority of treated households are observed in 2012. The parameter δ_{1d} is the effect of living less than d kilometers away from a road, once it has been rehabilitated and evaluated in 2012. We chose to discard the few households who are treated in 2010 (472 observations),[11] so that the parameter δ_{0d} provides a test of the common trend assumption. Indeed, if treated and control households do not satisfy this criteria, then δ_{0d} will be significantly different from 0. As a result, our evaluation of the program mostly compares 2008 and 2012.[12]

4.2.2. Matching in the panel

We now turn to a discussion of the matching in the panel. First, given that we use a panel, we only need to match households based on their 2008 characteristics. As previously

[11] More precisely, we exclude the households for which the short-period recall outcomes might be impacted by the road rehabilitation as early as 2010. Our results are robust to including them and assigning them as treated in the 2010 round.

[12] We have also tried a different specification which keeps all the observations and where the treatment is a dummy indicating the post intervention wave, *Post*, interacted with our treatment dummies, and the results are similar.

mentioned, there is little attrition in this panel. We construct a propensity score to receive a road rehabilitation for each household since this allows us to use individual characteristics as predictors in the logit regression.

First, we select households based on their distance to the nearest major road. Indeed, we want to compare households who benefited from an improvement in their nearest road to those who did not, but had the same proximity to a road. We thus create "buffers" around roads that define which households might act as a control. By definition, treated households have to be in the buffer. For instance, if we assess the effect of living less than 30km away from a rehabilitated road, the households that are used in the control group also live less than 30km from a main road. This set of potential controls being defined, we do a standard matching, based on observable characteristics. We also tried to simply use this distance to the nearest major road as a covariate in the matching but the common trend assumption was more often rejected and we decided to impose this constraint.

For the matching, we use the main selection criteria for the improvement of roads in Tanzania. We include the following variables: quality of the closest road prior 2008 (gravel, paved and sealed[13]), "Distance to the Nearest Border Crossing", "Distance to the Nearest Population Center with 20'000 inhabitants", "Distance to the Nearest District Headquarter" and "Distance to the Nearest Market" to capture the socio-economic importance of the area. In addition, we take into account village characteristics, including population density, agroecological zones and labor market availability. Finally, we include household and farm characteristics to find the best comparable households and to control for potential confounders influencing the outcome of interest. Table A4 in the Appendix presents summary statistics of the variables used to predict the propensity score.

We use an Epanechnikov kernel matching procedure based on the estimated propensity scores. We choose this non-parametric matching estimation because it uses more information to construct the counterfactual (Caliendo and Kopeinig, 2008). Compared to specification (1), it simply consists in running the household fixed effects regression with the kernel weights obtained from the logit, so that the control group is statistically comparable to the treatment group. We report two specifications for each outcome variable: the household fixed-effects with matching, and the household fixed-effects without matching but restrained on the common support sample. We do not report the pure household fixed-effects model (without any type of matching) since the common trend assumptions were more often rejected with this specification.

4.3. Estimation in the repeated cross-section: Difference-in-Difference with propensity score matching

The two rounds of the Labor Force Survey do not allow us to control for household fixed effects; however, we can still compare the evolution of treated and untreated households in a diff-in-diff strategy. We can control for constant characteristics at the group-level (the characteristics

[13]Soil is the reference quality type variable.

associated to living close to a road that will be rehabilitated). The main difficulty compared to the panel consists in guaranteeing that the composition of the groups does not change too much between the two rounds. Using again a matching approach strengthens the validity of the identifying assumption. Here we need to perform two matches: one to create a control group for the pre-intervention wave in 2006, and another to create a control group for the post-intervention wave in 2014. The treatment is defined by the distance to a rehabilitated road, as previously.

The specification we implement in the LFS is the following:

$$Y_{ivt} = \alpha + \delta_d 1_{it}(t = 2014) Road_{ivtd} + \theta Z_{ivt} + \phi Road_{ivtd} + \beta_t + \epsilon_{ivt} \quad (2)$$

We use a similar set of controls Z_{ivt} as in the panel. Tables A5 and A6 present summary statistics of the variables used to predict the propensity score, as well as both waves propensity score estimations.

5. Results

5.1. Matching results

Table A7 presents the results of the propensity score estimation. We start by commenting the matching performed for the panel. We compute a propensity score for each household using a logit regression based on 45 baseline characteristics. This matching is performed for each distance threshold but Table A7 only reports the results for the 20km threshold, as an example. We find that geo-referenced controls, road and village characteristics are the best predictors for the treatment. Several results point to the fact that the program guidelines were followed. First, households living closer to a gravel major road are more likely to have that road rehabilitated compared to soil roads, followed by paved and sealed roads. As stated in section 2.2, only roads in bad condition were selected to be improved and infrastructure works related to maintenance were given first priority. These results indicate that indeed roads that were in bad condition but did not demand large investments were selected first. Further, the closer to a border post, the higher the probability of having a road rehabilitated, in line with the government's objective of taking advantage of their location vis-à-vis of its bordering countries. Figure 1 shows that major roads connecting the country to its neighbors were prioritized. Likewise, households closer to a major population center are more likely to be treated. However, conditional on these characteristics, the closer to a major road and to a market, the less likely the household is treated. A few characteristics that may indicate higher living standards seem to have a negative effect on the probability of treatment, namely, durable and productive assets, a higher share of households with at least a member wage employed and the presence of a bank branch within 10km. For each matching, we manage to find a suitable control group. We give as examples the common support assessment for the 20km and the 30km matching procedures (Figures A2 and A3). Figure A4 displays the spatial location of treated and control households on the common support, as well as the observations that are not used in the analysis for the same treatment.

We carried out a balancing test of the difference in the means of the covariates between the control and the treated groups after the matching. Table A8 shows the quality of that matching. We achieve a perfect balancing on the baseline characteristics[14] and we find no significant differences between treated and control groups based on observables. We postpone the discussion on the balancing test results for the outcomes of interest.

For the repeated cross-section, we follow a similar methodology. Table A6 provides the results for the matching. We use a total of 26 regressors, and the same set of geo-coded variables as for the panel; but given that we can only geolocate households at the ward level, we loose some precision in the matching. However, we achieved an almost perfect balancing of the initial observed characteristics in both waves and a perfect matching for all outcome variables measured in 2006 (Table A9). Finally, Figures A5, A6 and A7 present the results for the common support at 20 and 30kms respectively and show that this requirement is obtained both for 2006 and 2014.

5.2. LSMS results

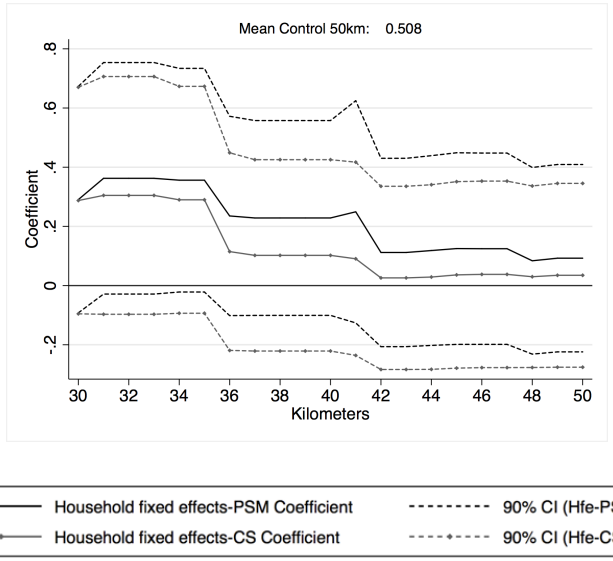
Migration. We start by assessing whether roads triggered more migration. This could be a potential threat to our identification strategy if better roads induce selective attrition in the panel. In addition, it is interesting to assess whether better roads induce changes on how the subsistence means are obtained. However, traveling along a road of poor quality is unlikely to be a substantial share of the migration costs. Therefore, we expect that it has little effect except for temporary migration. The next figures display the estimates for each treatment cut-off (ranging from living less than 20km away from the road to living less than 50km away from the road). The provided estimates are the household fixed effects with propensity score matching and the household fixed effects without propensity score matching but on the common support. The two confidence intervals are at the 90% level. Figure 4 provides the effect of the treatment on the number of household members away at least one month in the previous year. The average of this variable in the control sample of households living less than 50 kilometers away from a road is 0.508. The figure shows that there is little evidence that better roads induce migration. The figure A8a provides the test of the common trend assumption (i.e. the δ_{0d} coefficients). It shows that the common trend assumption might be violated in the instance of the migration variable: households living less than 40 kilometers away from the road display a higher likelihood to migrate and this likelihood has increased as early as 2010. From now on and for the sake of brevity, we will only refer to the tests of the common trend assumption when they seem to be rejected but the full set of results is available in the Appendix.

Welfare. We now turn to a general and subjective welfare measure to assess whether households valued their access to a better road. The subjective welfare measure is based on a 7-point ladder question asked to household members. The exact wording of this question is "*How satisfied or*

[14]Results for all the other treatments for both the panel and the cross-section case are available from authors upon request.

dissatisfied would you say you are with your life?". Obviously, the use of a subjective welfare question has limitations. However, in a setting where prices may change, it is extremely difficult to use aggregate consumption measures to make a welfare statement. In particular, we would need the full set of prices households face, which is not available in our dataset (Atkin et al., 2018). Therefore, we rely on this subjective welfare statement, which has also the advantage of encompassing other potential changes than just the ones associated to pure consumption. The outcome variable used in this section corresponds to the average subjective welfare among all respondents in the household.[15] Figure 5 shows that households who obtained a better road declare a *lower* life satisfaction. The decrease is actually quite substantial, since from an average of 3.7 it decreases by 0.5 (-13.5%) for households living less than 20 kilometers away from the road. The decrease is significantly different from zero for almost all treatments. However, the fact that it becomes non significantly different from zero for a treatment larger than 45 kilometers suggests that the households living further away from the road might actually have a no effect or a positive effect, which compensates for the loss for the households closer to the road. We will now try to understand where this reduction in life satisfaction might come from. One possibility is that roads trigger changes in the price system, which affects real income.

Figure 4: Migration



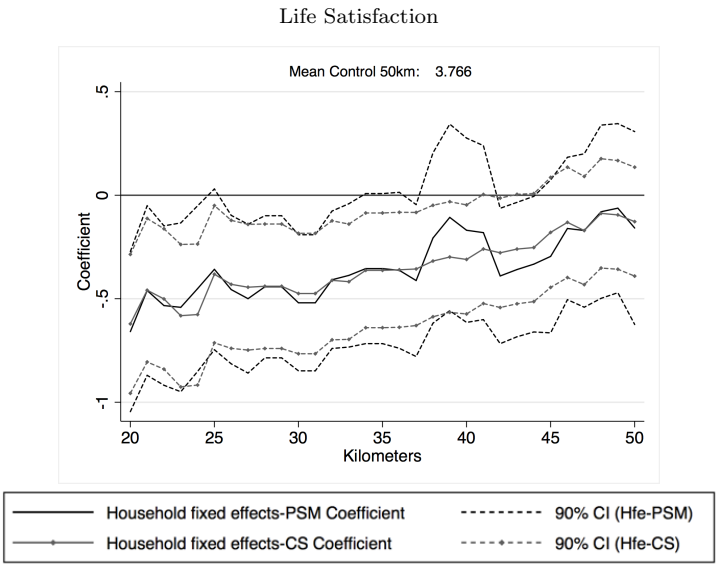
Note: The figure shows estimates from both a household fixed effects strategy with propensity score matching and from a household fixed effects strategy without propensity score but on the common support sample estimated separately for each treatment cut-off. Confidence intervals are at the 90% level.

Prices. We start by reviewing changes in the price of goods. We only have one observation per village for prices and only the price of available goods is reported. Therefore we need to focus on commonly exchanged goods. Rice and maize are by far the most traded products and they are the products for which we have enough observations to do the evaluation. Figure 6a

[15] As an alternative measure, we use the subjective welfare of the head of the household. Results are similar.

displays a clear decline in the price of rice, for all treatment thresholds, and the decrease is very substantial compared to the mean. The point estimates for maize also tend to be negative (Figure 6b) but the coefficients are much less often significant. A decrease in prices would have a negative impact on welfare for net producers and a positive impact on welfare for net consumers. Insofar as rice and maize are commonly produced by Tanzanian farmers, this decrease in rice price could explain the lower satisfaction. However, if transaction costs are lower, then other imported goods might also be cheaper (though it might not be automatically the case due to general equilibrium effects). We therefore check whether prices are generally lower, which presumably would limit the negative effect on welfare for producers. Figures 6c and 6d show that the price of sugar and of kerosene are not negatively impacted by road improvement. Therefore, the decrease in the price of produced goods might be a good explanation for the lower declared welfare.

Figure 5: Subjective Welfare



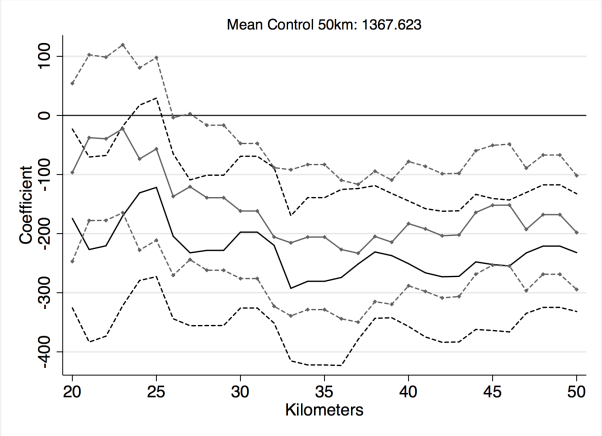
Note: The figure shows estimates from both a household fixed effects strategy with propensity score matching and from a household fixed effects strategy without propensity score but on the common support sample estimated separately for each treatment cut-off. Confidence intervals are at the 90% level.

Agricultural outcomes and consumption. In order to confirm this interpretation, we check the effect of roads on harvest, sales, land area and consumption for rice and maize. Figure 7 shows that the harvest for rice tends to be lower (not significant at the 10% level), that the sales and land area devoted to the crop are significantly lower for all treatments and that the consumption does not change. All the variables are expressed in kilograms and not in value and therefore are not mechanical effects of the reduction in price. Because the rice price is lower, households sell a lower share of their harvest, which might itself be somewhat lower. However, they fail to take advantage of rice price decrease because they do not consume more rice. We do not find any significant effect on the maize sector (see Figure 8), except that household consumption decreases for treated households, when the cut-off is sufficiently high. The negative effects of a rice price decrease are only expected if most of the treated households are actually net rice

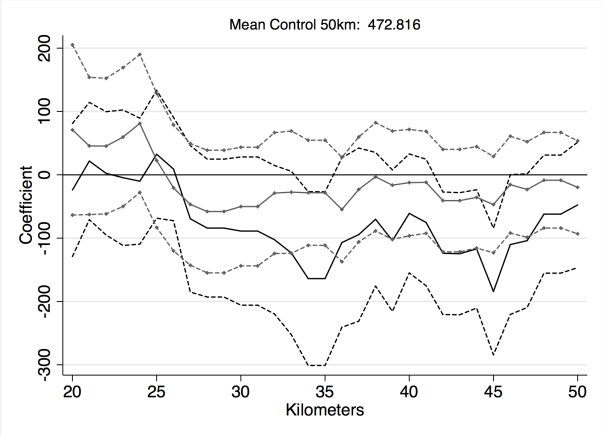
producers. Table A8 shows that, despite the matching, there is actually an imbalance between the treatment and the control group in terms of crops: the treated households are clearly in a rice area, while the control households are in a maize area. This could explain why we observe an impact on the rice sector but not so much in the maize sector.[16] Last, given the price drop and the reduction in harvested area in rice, we should observe an agricultural income reduction, except if the households manage to substitute rice with other high return crops. The evidence is inconclusive (see Figure 10d below) since we do not observe any significant change, maybe due to very imprecise income measurements.

Figure 6: Market Prices (TSh)

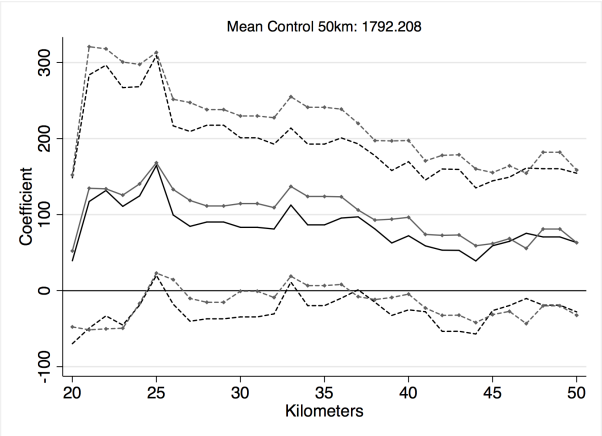
(a) Rice



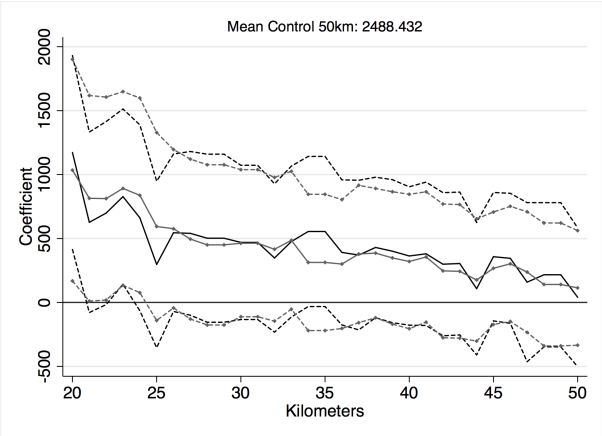
(b) Maize

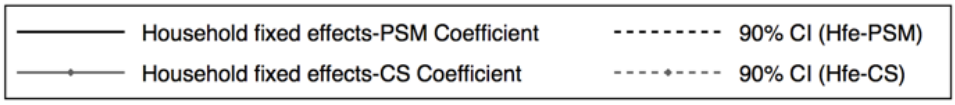


(c) Sugar



(d) Kerosene



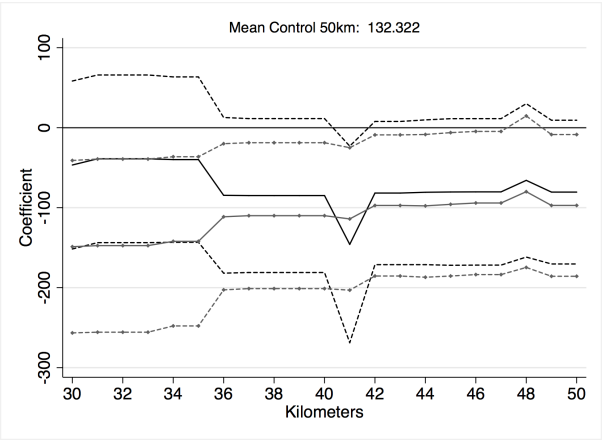


Note: The figure shows estimates from both a household fixed effects strategy with propensity score matching and from a household fixed effects strategy without propensity score but on the common support sample estimated separately for each treatment cut-off. Confidence intervals are at the 90% level. Prices are measured at the village level.

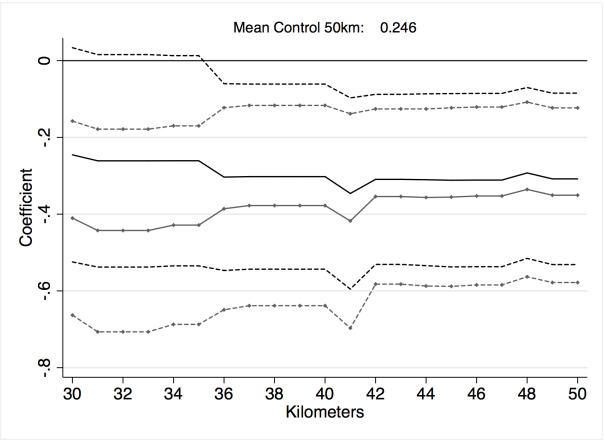
[16]However, when implementing a matching that takes into account the household crops in 2008, the results remain the same.

Figure 7: Harvest, Land Area, Sales and Consumption: Rice

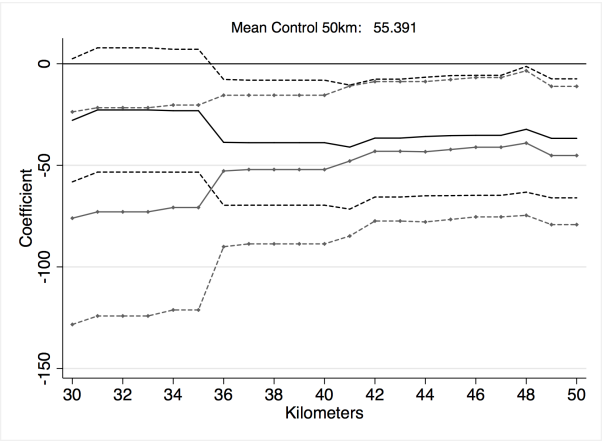
(a) Harvest: rice (kg)



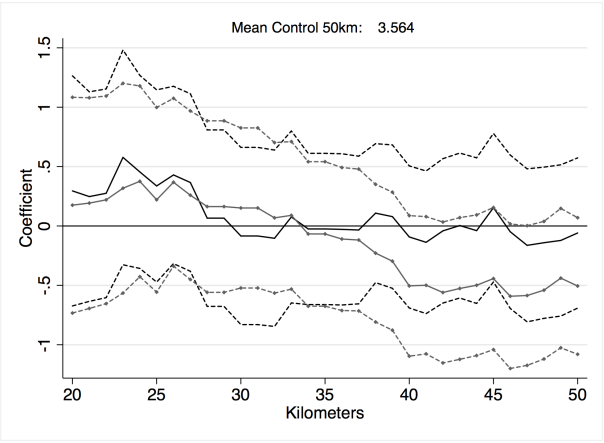
(b) Land area: rice (acres)

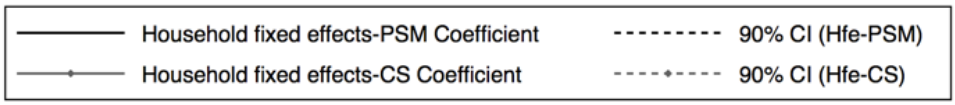


(c) Sales: rice (kg)



(d) Consumption: rice (kg)





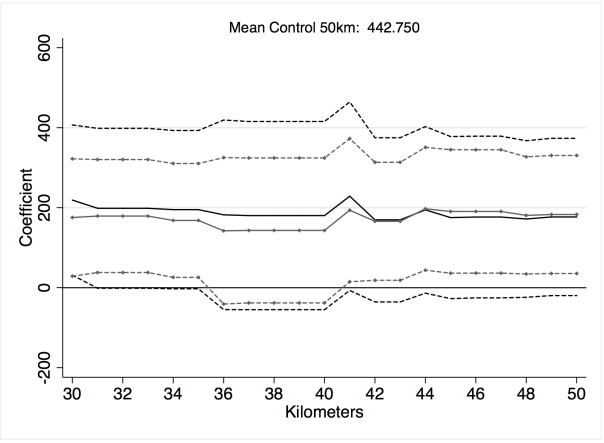
Note: The figure shows estimates from both a household fixed effects strategy with propensity score matching and from a household fixed effects strategy without propensity score but on the common support sample estimated separately for each treatment cut-off. Confidence intervals are at the 90% level.

Assets. Figure 9 assesses the impact of better roads on durable ownership and on productive assets ownership. Both variables are built with a principal component analysis run on a series of durable goods ownership (consumption goods for the first, productive assets for the second). The effect on durable assets is large and negative: treated households deplete assets. However, the common trend assumption is rejected for this variable (Figure A8p), even though the margin of significance is lower than for the δ_{1d} coefficients. The effect of roads on productive assets is also negative but marginally significant. We will now try understand better how households adjust their activity.

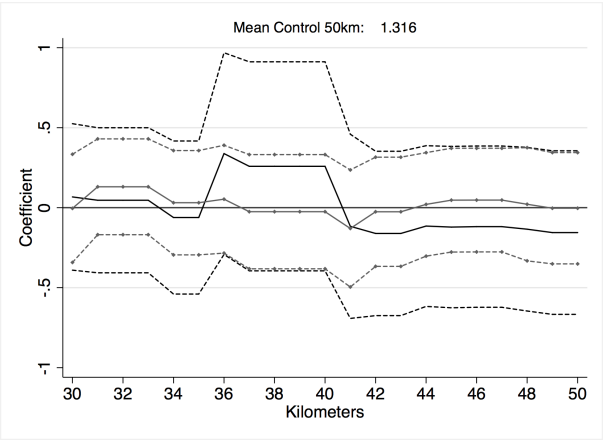
Labor market. To do so, we turn to the information available in the panel on the labor market. One of the expected advantages of a better road is a better access to markets and therefore a

Figure 8: Harvest, Land Area, Sales and Consumption: Maize

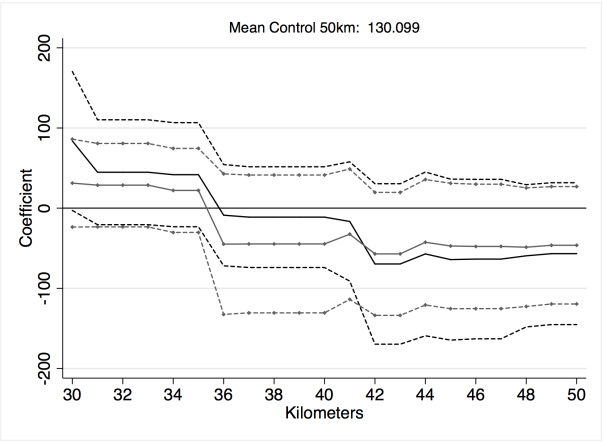
(a) Harvest: maize (kg)



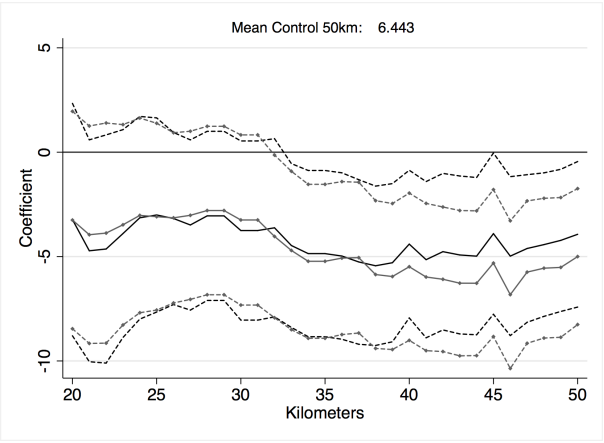
(b) Land area: maize (acres)

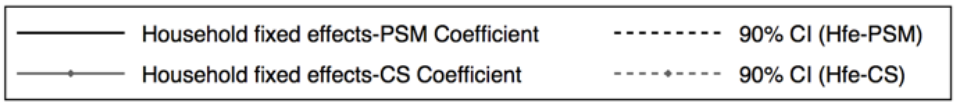


(c) Sales: maize (kg)



(d) Consumption: maize (kg)





Note: The figure shows estimates from both a household fixed effects strategy with propensity score matching and from a household fixed effects strategy without propensity score but on the common support sample estimated separately for each treatment cut-off. Confidence intervals are at the 90% level.

better access to farm inputs and to labor demand. Figure 10a identifies that treatment above 36 kilometers leads to an increase in wage labor (at least one household member works for a wage) and we also find in one specification that these households tend to reduce their on-farm labor (Figure 10c). A better road does not increase the use of farm inputs (Figure 10b for hired labor). However, given the product price decrease, farming becomes less attractive and this could be sufficient to counterbalance any improvement in market access. The analysis on the Labor Force Survey will help us be more specific on the changes in the labor market.

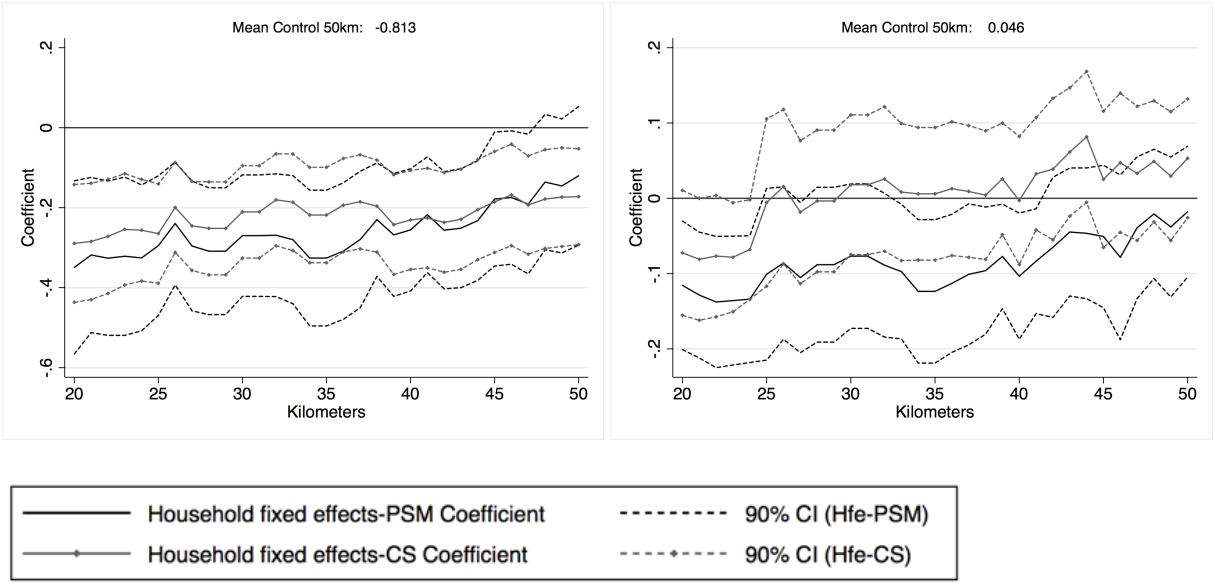
5.3. LFS results

Economic activities. We first assess whether the results are consistent between the two surveys, despite the fact that the samples are different. Figure 11a shows that treated households living

Figure 9: Assets (score)

(a) Durable Assets

(b) Productive Assets



Note: The figure shows estimates from both a household fixed effects strategy with propensity score matching and from a household fixed effects strategy without propensity score but on the common support sample estimated separately for each treatment cut-off. Confidence intervals are at the 90% level.

less than 40 kilometers away from the road reduce their economic activity: roughly 0.4 household members less have a current activity (against 2.4 on average in the control group).[17] Figure 11b confirms that treated households (but living far away from the road) increase wage work. Not only the result is the same than in the panel, but the cut-off is also the same. With the LFS, we also detect a small increase in the number of household members earning a wage for treatments below 30 kilometers. The reduction in the number of individuals with a current activity (-.4) is largely due to fewer individuals working on-farm (-.5) for households sufficiently close to the road and hardly compensated by self-employment (+.15 for the PSM specification) and wage work as previously mentioned. There are small effects on the likelihood to provide unpaid work but the prevalence of this type of work remains small. All in all, it seems that households reduce their economic activity, in particular on-farm, with only a very partial ability to compensate in self-employment and wage work.

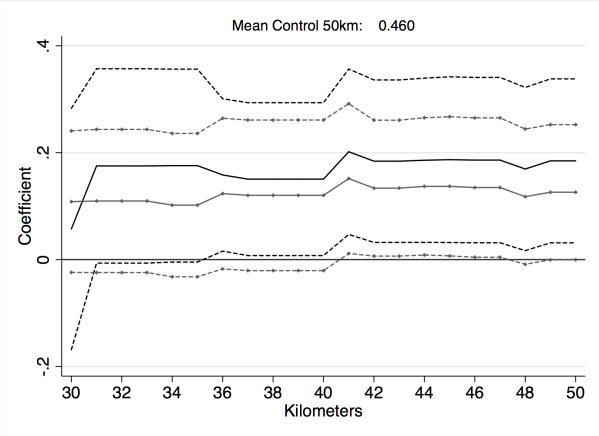
Wages. The LFS provide information on the payments obtained in the last month. Assuming that work hours in the last week reflect the usual activity of individuals, we can compute the hourly wage. If labor supply globally increases without an equivalent increase in labor demand, then wages should go down if the labor market is flexible but spatially limited. Figure 12a confirms that the wages tend to be depressed in villages that become better connected. This explains why, despite an increase in market economic activity, the earnings from wage work and self-employment do not increase (Figure 12b). Given the extent of reduction in the farming

[17]The number of household members is not impacted by the treatment. It is also a control variable in the matching.

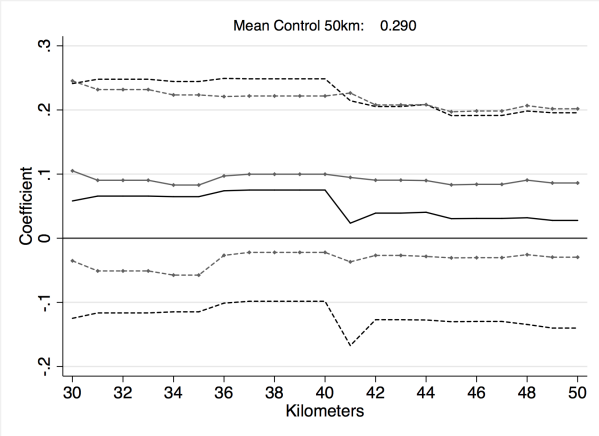
activity, the productivity in self-employment and wage work would have to be much larger than in farming to allow households to keep the income level they had before the reduction in prices.[18] This does not seem to be the case and provides additional evidence that households living in rural areas are actually negatively impacted by the road improvement.

Figure 10: Labor Market

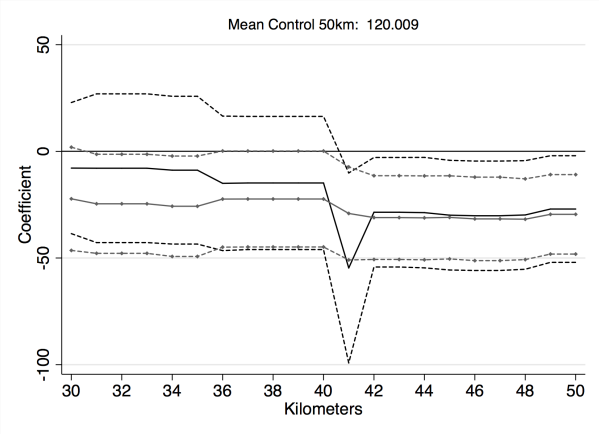
(a) Wage labor



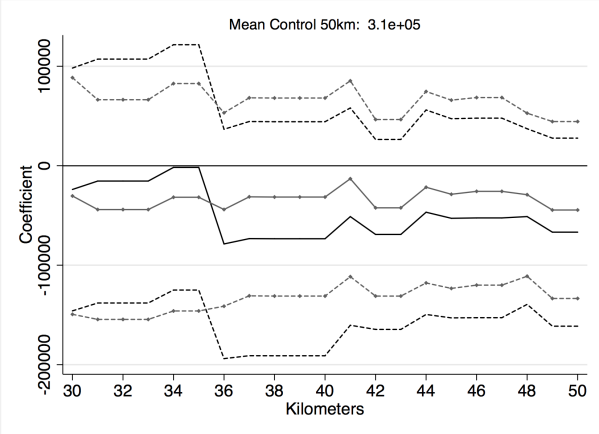
(b) Hired labor

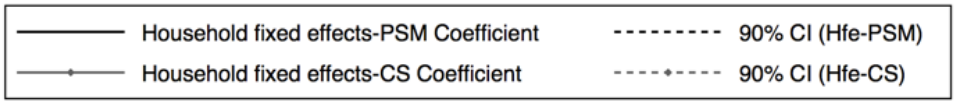


(c) N. days agricultural labor by household members



(d) Agricultural nominal income





Note: The figure shows estimates from both a household fixed effects strategy with propensity score matching and from a household fixed effects strategy without propensity score but on the common support sample estimated separately for each treatment cut-off. Confidence intervals are at the 90% level.

Sectors. We finally provide information on the new sector of activity of individuals. Consistently, they are less employed in agriculture. The only sectors where we observe an increase are: hospitality industry (e.g. restaurants), construction and transport. This seems consistent with expected effects of a road improvement and that the area may benefit from more activity

[18] Income from farming activity is not collected in the LFS.

in these sectors. However, again, the new opportunities from these sectors for the households seem to be fairly low, given that they allow hiring between +.01 and +.04 household members more each.

5.4. Results with simple differences-in-differences

In this section, we assess whether our results hold in a more standard diff-in-diff framework. We make two changes to our identification. First, we use as a control group all the households who live less than d kilometers away from a non-rehabilitated road, without weighting them with a propensity score matching procedure. Second, instead of cumulating all households living less than d kilometers away, with d increasing from 20 to 50, we simply build two groups of treated households: those living less than 30 kilometers away and those living 31 to 50 kilometers away. We expect the second group to be less affected than the first.[19] The results are provided in Tables A10 and A11. With this specification, we find no effect when households are more than 30 kilometers away from the road. The main effects when the households live less than 30kms away are confirmed with strong negative impacts on welfare, rice price, sales and harvest, as well as farm and economic activities. This is consistent with the conclusion that the effects tend to fade out when households live further away from the rehabilitated road. Importantly, it also shows that our results do not hinge on the propensity score matching.

5.5. Other possible effects of roads

In this section, we provide tests of other possible effects of roads. We do so in a more compact way, by simply providing estimates for treatment being: "less than 20kms away from a rehabilitated road", "less than 30kms away" and "less than 40kms away".

One possible threat to our identification could come from any other complementary policy that happened in response to the road infrastructure program. If, for instance, local governments decided to upgrade or construct roads connecting to the treated major roads, we would be evaluating the joint effect of those policies. To address this concern, we test at the community level whether the treatment at the different thresholds had any impact on the implementation of community projects in the following areas: road construction, market construction, irrigation schemes, water supply and grain storage. We find no evidence that this is the case (Table A12).

The reduction in household welfare could also be due to changes in security associated to the road (Ridgeway and MacDonald, 2017; Agnew, 2018). We therefore check whether there is a change in declared security and, if anything, find that theft is reduced with the program but the coefficients are insignificant.[20] This result, therefore, does not explain the reduction in well-being.

We also check whether a better market connection improves agricultural practices. We do not find any effect on access to credit, nor use of pesticides and inorganic fertilizer. We do find

[19] We would have liked to build a third group with households living more than 50kms away, which would have provided a placebo test, but there are not enough observations in this category.

[20] The exact question is how much was stolen from the household in the last 12 months.

an increase in use of organic fertilizer, but this variable does not pass the common trend test. We also do not detect any change in the number of cattle heads owned by households.

Last, we also want to see if households have access to a greater variety of goods (Gunning et al., 2018; Aggarwal, 2018). We use as an outcome variable the number of goods for which prices are collected in the market and do not find any change. We also check whether the number of sectors that hire labor has increased and this is not the case either. The benefit due to improved market access seems therefore hardly visible.

5.6. Heterogeneity of the effect

We now want to assess whether the effects of the road rehabilitation program are heterogeneous with respect to household characteristics, in particular with respect to being a cereal producer.[21] Our expectation would be that cereals producers should be more affected than others. However, households producing cereals are also wealthier and presumably better equipped to face shocks such as the price decrease.[22] In addition, general equilibrium effects taking place at the village level could blur the results. Indeed, the labor supply increase has been shown to depress wages, which will in turn affect differentially households depending on their labor supply and demand choices. As a result, the following test is not a definitive test of the mechanism at stake. The objective is also to determine who is more adversely impacted in the villages.[23]

We start by providing an assessment of the effect of the program, allowing for heterogeneity based on whether the household produced maize or rice in 2008 and based on whether it had a low or large land endowment in 2008. This is simply done by including an interaction term between the treatment and the land/cereals status of the household in the main regression. Given that households may adjust their choices on very different margins, we first use the life satisfaction variable, that reflects overall well-being in the household. Table 3 shows that the program led to a decrease in satisfaction specifically for cereals producers for households living close to the new road, but that the negative effect is not different by land ownership status. The significance of the main effect is reduced for households living further away from the road but the coefficients are still negative.

We then move on to analyze changes for other outcome variables. We find that only households producing cereals reduce their sales in rice and maize and that the difference to the non-producers is significant, which is expected. We also find that cereals producers significantly reduce their labor demand compared to the non cash crop producers, which is consistent with a reduction in the rice returns. However, we also find that non cereals producers actually increase their labor demand. This has to be interpreted jointly with the result that non cereals

[21] Given that we do not have many rice producers in the control group, we use rice and maize producers, called here "cereals producers" as the group of interest.

[22] 62% of households without maize nor rice crops have a land area below the median, and 60% of households with cereals have a land area above the median.

[23] We restrict this analysis to the LSMS data since it allows us to use household characteristics that were collected in 2008, and therefore not impacted by the treatment, which is not doable in the LFS, for lack of panel data.

producers also increase their wage labor supply. This group of households therefore participate more actively in the labor market, which is predicted by the theory. Interestingly, the group of cereals producers have a different behaviour and do not increase neither their demand nor their supply. This is consistent with the result obtained in Jayachandran (2006) for India: wealthier households are protected from shocks because wages adjust. Here, because the wages decrease, cereals producers even though they have a lower rice return, also face lower labor costs and do not change their demand.

5.7. Placebo tests

We now turn to placebo checks in order to assess the validity of our estimation strategy. First, we provide an additional test of the common trend assumption, in a much longer run than the one that was provided using the intermediate round of the LSMS data. To do so, we use the DHS data that were collected in 2003, 2007 and 2010 and build a durable assets index with a principal component analysis. We find that the areas that were treated were not diverging from the control areas prior to the treatment on the basis of this wealth index (Table A13).[24]

Second, we can use the roads that have been rehabilitated after our time window (meaning after 2013) as placebo roads, in order to assess whether selected roads have specificities that could invalidate the common trend assumption. More precisely, given that most major roads will end up being rehabilitated, we use the information on whether roads had undergone a feasibility study by 2014 as a placebo treatment and compare them to roads that were neither rehabilitated in 2012 nor had undergone a feasibility study by 2014. The results are provided in Table A14.[25] Most of the coefficients are non significant, as required for a placebo test. The coefficients that are significantly different from 0 (for instance, rice sales) are actually of the opposite sign than the one obtained in our main results, and therefore cannot explain what we find.

5.8. Using alternative samples

We also want to assess the robustness of our results to changes in the sample. First, our estimates might be biased if treated villages have spillover effects on the control villages. One option consists in using a control group that is further away from the treated areas. We do so by restricting control villages to regions that do not receive any road. This set of estimates has to be treated with caution since the control villages are actually less similar to the treated villages since they belong to more remote regions. The balancing test is more often non satisfactory than when using the control sample located in the treated region.[26] This being said, with this

[24]The DHS data include geolocation of households. We build the distance to the main road and define treated households as the ones who are close to a rehabilitated road, which is then interacted with year dummies. This provides the test of the common trend assumption.

[25]For the LSMS data, we implement the same identification strategy as before but for the LFS data, we are unable to have a propensity score matching for lack of common support and therefore provide only the diff-in-diff results.

[26]This result is not included in the Appendix but is available upon request.

alternative sample of control villages, we still find that the subjective welfare decreases, and that the rice price sharply decreases (Table A15). We also find household production adjustments to this change in the rice price. However, the effects on labor supply obtained from the LFS are very imprecisely estimated and sometimes at odds with the previous results. This comes from the unability of the propensity score matching procedure to obtain a good matching with this alternative control sample.[27]

Second, we also want to exclude the villages that are close to nodes of the road network because there might be endogeneity in the selection of the nodes that will be treated (Banerjee et al., 2020). We therefore run the same regressions but excluding villages that are located

Table 3: Heterogenous Effects (LSMS-ISA)

Variable		**less than 20km**	**less than 30km**	**less than 40km**
Subjective Welfare	Road	0.0844	−0.4083	−0.5032
		(0.371)	(0.327)	(0.323)
	Road*cereals	−1.0151**	−0.2052	0.4325
		(0.462)	(0.385)	(0.497)
	Road	−0.7832**	−0.7240**	−0.2228
		(0.356)	(0.289)	(0.419)
	Road*big land owner	0.0904	0.2933	−0.0618
		(0.388)	(0.344)	(0.480)
Wage labor	Road		0.3671**	0.3153***
			(0.174)	(0.120)
	Road*cereals		−0.5535**	−0.3517*
			(0.219)	(0.191)
Hire labor	Road		0.3118	0.3689**
			(0.198)	(0.141)
	Road*cereals		−0.2066	−0.3735**
			(0.229)	(0.157)
Sales maize (kg)	Road		76.7906	110.3818*
			(50.688)	(57.048)
	Road*cereals		12.8782	−239.1726**
			(76.637)	(96.121)
Sales rice (kg)	Road		3.5439	−12.0408
			(4.308)	(9.811)
	Road*cereals		−49.5079	−44.5196
			(31.247)	(31.733)

Note: The reported treatment effects are the result of a household fixed effects strategy with propensity score matching. *Road* corresponds to the parameter δ_{1d} of equation 1. *Cereals* indicates that a household produced maize or rice on the last long raining season. Robust standard errors clustered at the village level in parenthesis.
* p< 0.10, ** p<0.05, *** p<0.01, **** p<0.001.

less than 5 kilometers away from one of the cities with more than 5000 inhabitants. Given the number of such cities, we cannot exclude a larger perimeter from the city. Table A16 shows that we confirm the effect on welfare, on the rice price and on the type of activities undertaken by the household member. In addition, we find a negative effect on productive assets.

[27]The result is not included in the Appendix but is available upon request.

5.9. Alternative treatment definitions

We now want to assess two alternative treatment definitions. First, our paper identifies a negative effect of being close to a road due to the increased competition one face. However, it is not clear what is the market in which such competition takes place. Indeed, our treatment could proxy for another treatment: being close to a city (i.e. a market) which benefits from a better road connection. If this is the case, then some villages may not be considered as treated if we consider how close they are from the new road, while they would be treated because they are close to a city, which is treated. Those villages would particularly suffer from the road program since they would have the increased competition without having a better connection.

In order to gain insight on this question, we define a new treatment and run similar regressions as before. The new treatment is to be located less than d kilometers away from a city, which is treated in the sense that the programme improved the connectivity to the rest of the country or to neighbouring countries (Figure A9).

We provide a subsample of the results. Globally, the effects go in the same direction but they are lower and often loose significance (see for instance, the effect on life satisfaction, rice price and rice area in Table A17). Our previous treatment variable seems therefore to have a stronger predictive power than the distance to a treated city. We can therefore conclude that the proximity to the road triggers the reduction in well-being rather than the proximity to a treated city. This suggests that the main market in which the households face increased competition are local rather than located in major cities.

Second, in all our analysis we have used the as-the-crow-flies distance to the road. This has the great advantage of being computed without additional assumptions. However, if individuals have to use specific secondary roads to reach the major road, this distance measure may not reflect the true transportation cost between the village and the road. If this is the case then our treatment variable has some measurement error and our estimates may be biased. We therefore use an alternative treatment definition based on time needed to reach the main road. For doing so, we use the network of secondary roads of the open source routing service OpenStreetMap (OSM).[28] We use this additional data because it records more secondary roads than our primary data. However, we do not use travel time obtained directly from OpenStreetMap since its computation relies on real-time data. Instead, we use the Stata command osrmtime (Huber and Rust, 2016).[29] Figure A10 displays the secondary roads in these data. Table A18 shows that the effects on the rice price, rice harvest and rice sale are confirmed with this alternative treatment. We also find a reduction in the number of household members who have an economic activity (and work on farm). The main difference is that we do not identify anymore an effect on

[28]OSM maps were obtained from http://download.geofabrik.de in February 2020. OSM road data for 2008 cannot be retrieved due to data quality issues.

[29]The command osrmtime computes offline the optimal travel time between two points using their geographic coordinates and speed profiles that depend on the mean of transportation and the road segment classification. For instance, the default speed used for a car in a secondary road is 55 km/h. All our requests use "car" as a mean of transport and the default speed profile.

welfare. We find it reassuring that most of the effects are similar with this alternative treatment variable. However, there are various reasons why we prefer the treatment variable based on the as-the-crow-flies distance to the road. First, it does not require additional assumptions on travel time along the secondary roads. Second, the secondary roads were collected in 2020 and could have changed compared to our study period, and in any case, it is not completely clear that they reflect the actual way people connect to the road.

6. Conclusion

We evaluate the effects of road infrastructure improvements on market participation of labor and products, prices and welfare of rural households in Tanzania. Our empirical method consists on combining household fixed-effects and difference-in-difference with a propensity score matching. We provide tests of common support and balancing between the treatment and the control group that are satisfactory. The tests also do not reject the common trend assumption between the two groups in the vast majority of the outcome variables.

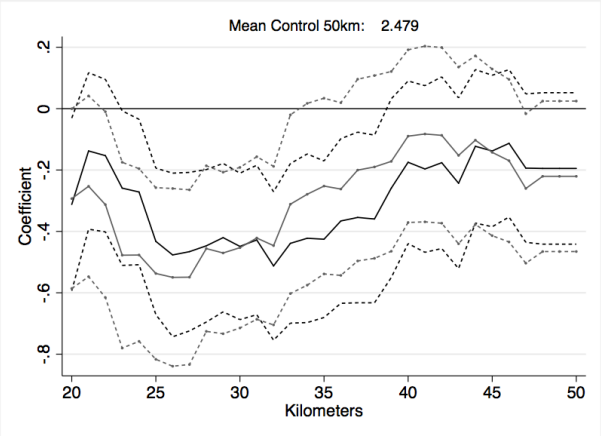
Contrary to most studies, we find damaging effects of the road improvements on the rural population: we observe a reduction in life satisfaction, a reallocation of labor away from farm without many additional opportunities of work, that we trace back to a decrease in the price of the main product (rice). This is consistent with predictions obtained from trade models where rural households, with lower transaction costs, actually face competition from lower price goods. We also find that the increase in market labor supply depressed wages. In addition, the heterogeneity analysis shows that large landowners and cereals producres are actually insured against the shock associated to the program, as found in Jayachandran (2006). By comparison, the individuals with fewer assets bear a large share of the burden. Not only are rural households adversely impacted by the program, but the general equilibrium effects propagates the shock to the poorer segment of the population.

Our conclusion raises the question of the specificity of our results to the Tanzanian case. First, as recalled in the introduction, there are actually only few contributions which evaluate the impact of these type of roads. Often, rural roads (i.e. better connecting villages to neighboring cities) are evaluated. Articles on major roads (Storeygard, 2016; Ghani et al., 2016; Shrestha, 2019) use as outcomes exports and manufacturing and therefore say little on effects on the rural population. Second, the road rehabilitation program fulfilled several objectives: not only poverty reduction, but also improvements in the connectivity with the neighboring countries. These countries are also producers of rice and maize and Tanzania is a natural export country for them. In our study, we are unable to tell where the competition comes from, because we do not have information on trade flows. One possibility is that the better connectivity with the neighboring countries triggered higher imports from them and therefore fiercer competition for the Tanzanian producers. We acknowledge that, had the choice of roads been different, the welfare effect on rural households could have been different as well. However, it is not innocuous that the actual placement of roads was as such. Under cover of poverty reduction,

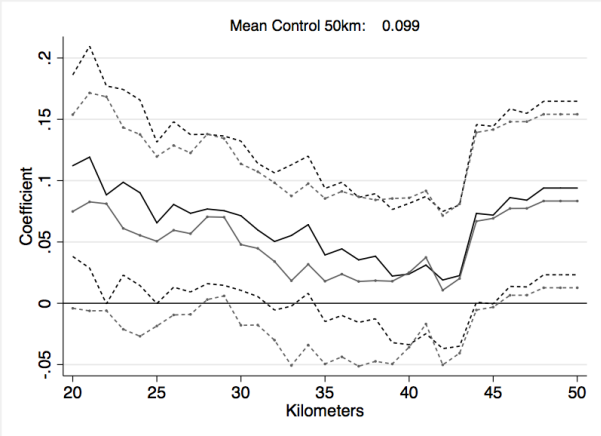
the governments may actually implement programs which simply aim at fostering economic growth. This paper does not address the question of the impact of roads on urban areas. However, preliminary estimations show that urban households (in the panel) actually benefit from the road improvement: they have a higher life satisfaction and a lower exposure to price shocks. It would be interesting as well to document the effect of roads in Tanzania on the farm households living in the neighboring countries. This is left for further research.

Figure 11: Labor Market
(N. household members)

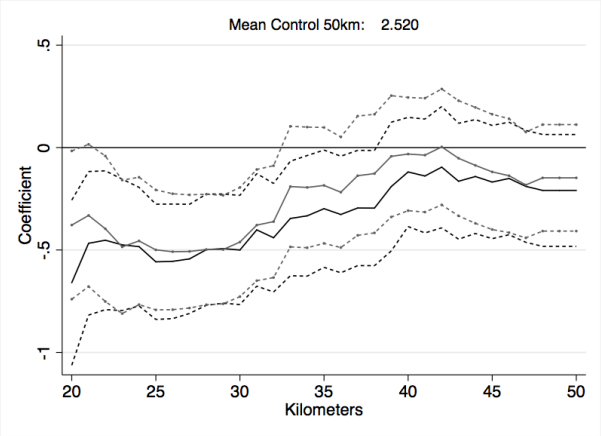
(a) With a current activity



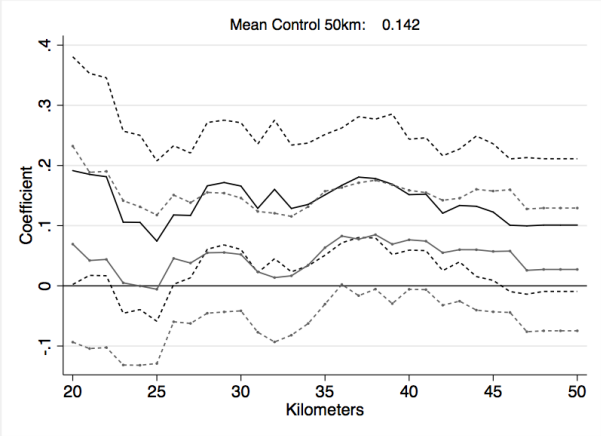
(b) Earn a wage



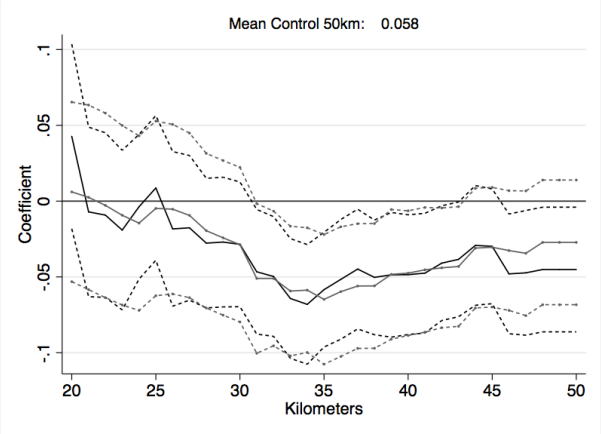
(c) Work on farm

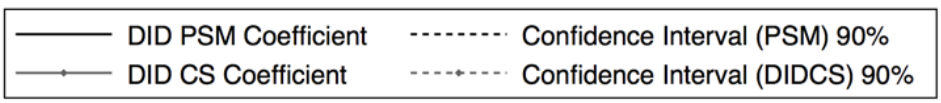


(d) Self-employed



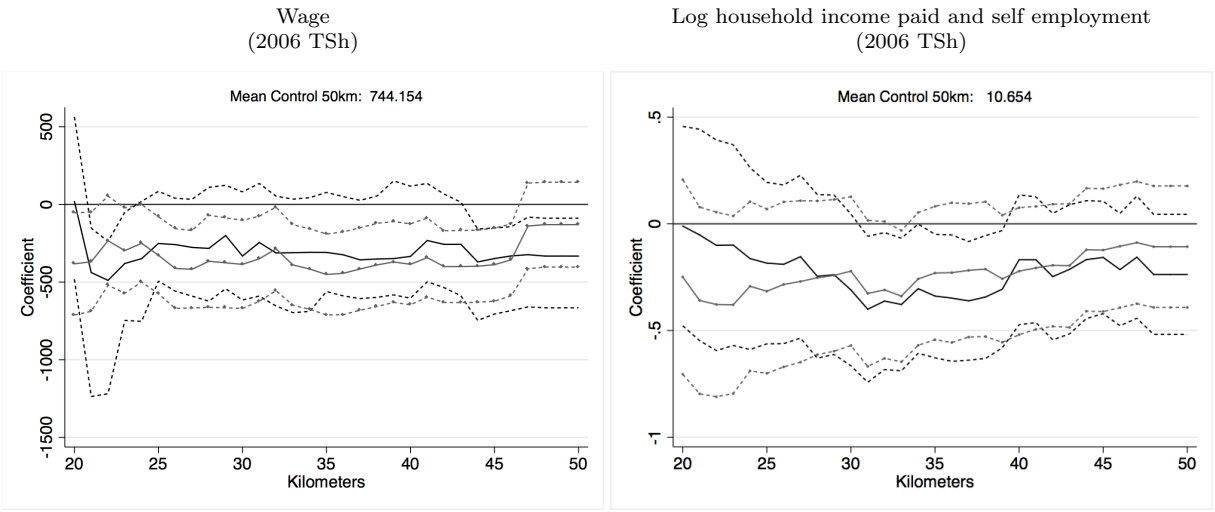
(e) Work unpaid





Note: The figure shows estimates from both a household fixed effects strategy with propensity score matching and from a household fixed effects strategy without propensity score but on the common support sample estimated separately for each treatment cut-off. Confidence intervals are at the 90% level.

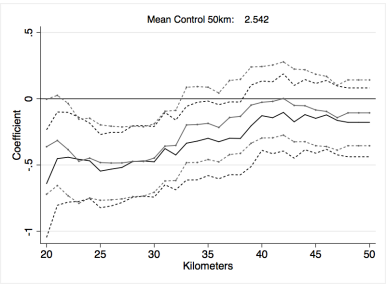
Figure 12: Wage and income



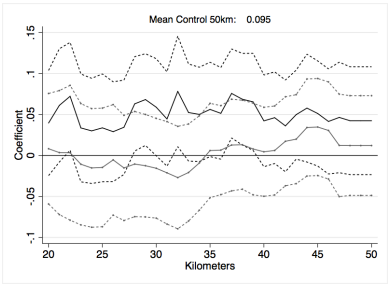
Note: The figure shows estimates from both a household fixed effects strategy with propensity score matching and from a household fixed effects strategy without propensity score but on the common support sample estimated separately for each treatment cut-off. Confidence intervals are at the 90% level.

Figure 13: Sector of activity (ISIC)
(N. household members)

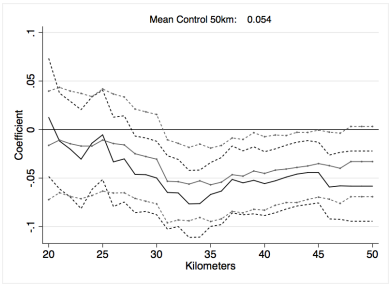
(a) Agriculture



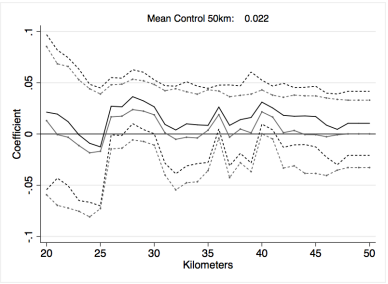
(b) Retail



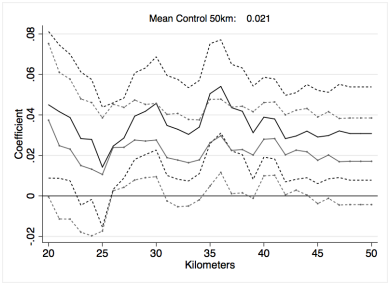
(c) Domestic personnel



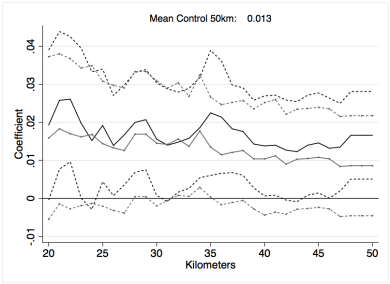
(d) Manufacturing



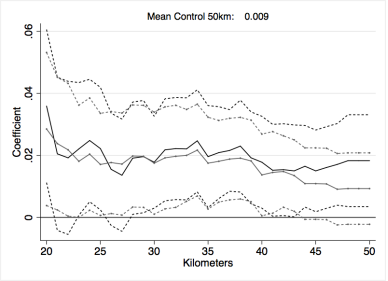
(e) Hospitality industry

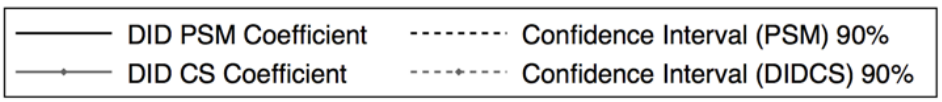


(f) Construction



(g) Transport





Note: The figure shows estimates from both a household fixed effects strategy with propensity score matching and from a household fixed effects strategy without propensity score but on the common support sample estimated separately for each treatment cut-off. Confidence intervals are at the 90% level.

References

Adukia, A., Asher, S., and Novosad, P. (2017). Educational Investment Responses to Economic Opportunity: Evidence from Indian Road Construction. MPRA Paper 80194, University Library of Munich, Germany.

Aggarwal, S. (2018). Do rural roads create pathways out of poverty? Evidence from India. *Journal of Development Economics*, 133(C):375–395.

Agnew, K. (2018). Crime highways: the effect of motorway expansion on burglary rates.

Asher, S. E. and Novosad, P. M. (2018). Rural roads and local economic development. Policy Research Working Paper Series 8466, The World Bank.

Atkin, D., Faber, B., and Gonzalez-Navarro, M. (2018). Retail globalization and household welfare: Evidence from mexico. *Journal of Political Economy*, 126(1):1–73.

Attanasio, O., Goldberg, P. K., and Pavcnik, N. (2004). Trade reforms and wage inequality in Colombia. *Journal of Development Economics*, 74(2):331–366.

Banerjee, A., Duflo, E., and Qian, N. (2020). On the road: Access to transportation infrastructure and economic growth in China. *Journal of Development Economics*, 145:102442.

Burgess, R., Jedwab, R., Miguel, E., Morjaria, A., and Padró i Miquel, G. (2015). The Value of Democracy: Evidence from Road Building in Kenya. *American Economic Review*, 105(6):1817–1851.

Caliendo, M. and Kopeinig, S. (2008). Some Practical Guidance for the Implementation of Propensity Score Matching. *Journal of Economic Surveys*, 22(1):31–72.

Casaburi, L., Glennerster, R., and Suri, T. (2013). Rural roads and intermediated trade: Regression discontinuity evidence from Sierra Leone.

Castaneda, A., Doan, D., Newhouse, D., Nguyen, M. C., Uematsu, H., and Azevedo, J. P. (2016). Who are the poor in the developing world ? Policy Research Working Paper Series 7844, The World Bank.

Currie, J. and Harrison, A. E. (1997). Sharing the Costs: The Impact of Trade Reform on Capital and Labor in Morocco. *Journal of Labor Economics*, 15(3):44–71.

de Janvry, A., Fafchamps, M., and Sadoulet, E. (1991). Peasant household behaviour with missing markets: Some paradoxes explained. *Economic Journal*, 101(409):1400–1417.

Dercon, S. (1998). Wealth, risk and activity choice: cattle in western tanzania. *Journal of Development Economics*, 55(1):1–42.

Donaldson, D. (2018). Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *American Economic Review*, 108(4-5):899–934.

FAO (2018). Small Family Farms Country Factsheet. *Food and Agriculture Organization of the United Nations (FAO) Country Factsheets*, pages 1–2.

FAO (2020). FAOSTAT Database.

Gachassin, M., Najman, B., and Raballand, G. (2010). The Impact Of Roads On Poverty Reduction : A Case Study Of Cameroon. *The World Bank Policy Research Working Paper*, (5209):1–39.

Ghani, E., Goswami, A. G., and Kerr, W. R. (2016). Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing. *Economic Journal*, 126(591):317–357.

Government of Tanzania (2007). The Roads Act, Act No.13 of 2007. pages 1–239.

Government of Tanzania (2008). 10 Year Transport Sector Investment Programme (TSIP). pages 1–119.

Government of Tanzania (2011). The Tanzania Five Year Development Plan 2011/2012-2015/2016. Technical report, President's Office, Planning Commission.

Gunning, J. W., Krishnan, P., and Mengistu, A. T. (2018). Fading Choice: Transport Costs and Variety in Consumer Goods. CEPR Discussion Papers 12910, C.E.P.R. Discussion Papers.

Harrison, A. and Hanson, G. (1999). Who gains from trade reform? Some remaining puzzles. *Journal of Development Economics*, 59(1):125–154.

Harrison, A. E., Lin, J. Y., and Xu, L. C. (2014). Explaining Africa's (Dis)advantage. *World Development*, 63(C):59–77.

Heckman, J. and Navarro-Lozano, S. (2004). Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models. *Review of Economics and Statistics*, 86(1):30–57.

Huber, S. and Rust, C. (2016). osrmtime: Calculate Travel Time and Distance with OpenStreetMap Data Using the Open Source Routing Machine (OSRM). *The Stata Journal*, 16(2):1–8.

Jacoby, H. G. and Minten, B. (2009). On measuring the benefits of lower transport costs. *Journal of Development Economics*, 89(1):28–38.

Jayachandran, S. (2006). Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of Political Economy*, 114(3):538–575.

Khandker, S. R., Bakht, Z., and Koolwal, G. B. (2009). The Poverty Impact of Rural Roads: Evidence from Bangladesh. *Economic Development and Cultural Change*, 57(4):685–722.

Lokshin, M. and Yemtsov, R. (2005). Has Rural Infrastructure Rehabilitation in Georgia Helped the Poor? *World Bank Economic Review*, 19(2):311–333.

Minten, B. and Kyle, S. (1999). The effect of distance and road quality on food collection, marketing margins, and traders' wages: evidence from the former Zaire. *Journal of Development Economics*, 60(2):467–495.

Mu, R. and van de Walle, D. (2011). Rural Roads and Local Market Development in Vietnam. *Journal of Development Studies*, 47(5):709–734.

Munshi, K. and Rosenzweig, M. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, 106(1):46–98.

Porto, G. G. (2006). Using survey data to assess the distributional effects of trade policy. *Journal of International Economics*, 70(1):140–160.

Qin, Y. and Zhang, X. (2016). The Road to Specialization in Agricultural Production: Evidence from Rural China. *World Development*, 77(C):1–16.

Renkow, M., Hallstrom, D. G., and Karanja, D. D. (2004). Rural infrastructure, transactions costs and market participation in Kenya. *Journal of Development Economics*, 73(1):349–367.

Revenga, A. (1997). Employment and Wage Effects of Trade Liberalization: The Case of Mexican Manufacturing. *Journal of Labor Economics*, 15(3):20–43.

Ridgeway, G. and MacDonald, J. M. (2017). Effect of rail transit on crime: A study of los angeles from 1988 to 2014. *Journal of Quantitative Criminology*, 33:277–291.

Schwab, K. and Porter, M. E., editors (2008). *The Global Competitiveness Report 2008–2009*. World Economic Forum, Geneva.

Schwab, K. and Sala-i Martin, X., editors (2016). *The Global Competitiveness Report 2016–2017*. World Economic Forum, Geneva.

Shamdasani, Y. (2016). Rural road infrastructure and agricultural production: Evidence from india.

Shrestha, S. A. (2019). Roads, Participation in Markets, and Benefits to Agricultural Households: Evidence from the Topography-based Highway Network in Nepal. *Economic Development and Cultural Change*, forthcoming.

Singh, I., Squire, L., and Strauss, J. (1986). *Agricultural Household Models: Extension, Application and Policy*. Johns Hopkins Univ. Press, Baltimore, MD.

Smith, J. A. and Todd, P. E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics*, 125(1-2):305–353.

Storeygard, A. (2016). Farther on down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa. *Review of Economic Studies*, 83(3):1263–1295.

The World Bank (2008). Safe, clean, and affordable...transport for development : the World Bank Group's transport business strategy for 2008-2012. Technical report.

Topalova, P. (2004). Factor immobility and regional impacts of trade liberalization: Evidence on poverty and inequality form india.

Volpe Martincus, C., Carballo, J., and Cusolito, A. (2017). Roads, exports and employment: Evidence from a developing country. *Journal of Development Economics*, 125(C):21–39.

Appendix A. Appendix

Appendix A.1. Variable definition

Outcome variables in the LSMS:

Life satisfaction: Average self-reported life satisfaction among all respondents in the household on a scale from 1 to 7, where 1 corresponds to very dissatisfied and 7 to very satisfied.

Consumption rice (kg) last 7d: Household consumption of rice (kg) over the last week.

Consumption maize (kg) last 7d: Household consumption of maize (kg) over the last week.

Prices (TSh): Village level prices (TSh) of rice husked, maize grain, sugar and kerosene.

Wage labor: Dummy informing if at least one member of the household engaged in wage labor over the last year.

Hire labor: Dummy informing if the household hired labor over the last year.

N. days agri. labor by hhm: Number of farming days in the last rainy season performed by household members.

Harvest rice (kg): Harvest of rice (kg) over the last rainy season.

Sales rice (kg): Sales of rice (kg) over the last rainy season.

Harvest maize (kg): Harvest of maize (kg) over the last rainy season.

Sales maize (kg): Sales of maize (kg) over the last rainy season.

Log PC real annual consumption: Logarithm of the per capita annual household consumption. The price deflator is computed by the World Bank (see the data documentation).

N. members away at least 1m last 12m: Number of household members that lived away for at least one month over the last year.

Outcome variables in the LFS:

N. household members have current act.: Number of household members that had an economic activity over the last week.

N. household members earned wage: Number of household members that earned a wage over the last year.

N. household members self employed: Number of household members self-employed over the last year.

N. household members work unpaid: Number of household members that worked unpaid over the last year.

N. household members work on farm: Number of household members that worked on farm over the last year.

Isic: work in agriculture: Number of household members that worked in agriculture over the last year.

Isic: work in manufacturing: Number of household members that worked in manufacturing over the last year.

Isic: work in construction: Number of household members that worked in construction over the last year.

Isic: work in retail: Number of household members that worked in retail over the last year.
Isic: work in transport: Number of household members that worked in transport over the last year.
Isic: work in hospitality industry: Number of household members that worked in hotels and restaurants over the last year.
Isic: work as domestic personnel: Number of household members that worked as domestic personnel the last year.
Log Hh income paid and self employment (2006 TSh): Logarithm of the total household income paid and self employment (TSh) adjusted to 2006 values.

Control variables:
Head Male: Dummy variable equal to 1 if the household's head is male.
Head Age: Age in years of the household head.
Fraction hh members 0-6 years old: Share of household members aged 0 to 6 years old.
Fraction hh members 7-12 years old : Share of household members aged 7 to 12 years old.
Fraction hh members 13-18 years old : Share of household members aged 13 to 18 years old.
Fraction male hh members >18 years old : Share of household members who are male and older than 18 years old.
Fraction female hh members >18 years old: Share of household members who are female and older than 18 years old.
Household size: Number of people living in the household.
More than primary Head: Dummy equal to 1 if the household head has attained at least primary education.
N. adult members with > primary: Number of adults aged 18 and older in the household that attained at least primary education.
Durable Assets (score): Principal component analysis (PCA) asset index constructed on the basis of 30 household durable assets, i.e. radio, telephone, refrigerator, beds, etc.
Productive Assets (score): Principal component analysis (PCA) asset index constructed on the basis of 20 household productive assets, i.e. carts, tractor, hoes, etc.
Land title: Dummy variables equal to 1 if the household holds a land title for at least one of their plots.
Land rights: Dummy variable equal to 1 if the household has the right to sell or to use as collateral at least one of their plots.
Used org fertilizer: Dummy variable equal to 1 if the household uses organic fertilizer in at least one plot.
Used inorg fertilizer: Dummy variable equal to 1 if the household uses inorganic fertilizer in at least one plot.
Used pesticides: Dummy variable equal to 1 if the household uses pesticides in at least one

plot.

Erosion: Dummy variable equal to 1 if the household has had an erosion issue for at least one plot.

Irrigation: Dummy variable equal to 1 if the household has at least one irrigated plot.

Tropic-warm/humid agroecological zone: Dummy variable equal to 1 if the climate is Tropical-warm and humid and 0 if the climate is Tropical-cool and humid.

Quality of land: Total surface area of land (in acres) of good, average and bad quality.

2008 Distance (Km) to Nearest:

Major Road: Distance (km) to the nearest major road.

Pop Center with >20'000: Distance (km) to the nearest population center with more than 20,000 inhabitants.

Border Crossing: Distance (km) to the nearest border crossing.

District Headquarters: Distance (km) to the nearest district headquarters.

Quality of closest road: Dummy variables equal to 1 if the closest major road is graveled, paved, sealed or in soil condition.

Average hh size in the ward: Average household size in the ward.

Population Density: 2010 population density (people per km^2) in the ward.

Share hhs with a member wage employed: Share of household members in the ward with at least a member engaged in wage employment based on the labor force survey sample.

Share illiterate people >15 years old : Share of illiterate people 15 years old and older in the ward based on the labor force survey sample.

N. farmer cooperative groups: Number of farmer cooperative groups/association in the village.

Services available (less 10km): Dummy variables indicating if there is a bank branch, a savings and credit cooperative organization (SACCOs), a daily market and a weekly market available within a radius of 10km.

Services available (less 2km): Dummy variables informing if there is a primary school, a clinic, a shop and a market available within a radius of 2km.

Table A1: Geovariables Description

Variable	Source	Data type	Resolution	Description
2008 Distance (km) to Nearest				
Major Road	TANROADS	Shapefile	N.A.	Euclidean distance computed using QGIS
Population Center with $>$ 20'000	City population	Text List	N.A.	Georeferenced using QGIS
Border Crossing	Tracks4Africa	Text List	N.A.	Georeferenced using QGIS
District Headquarters	Statoids	Text List	N.A.	Georeferenced using QGIS
Quality of the closest Road	TANROADS	Shapefile	N.A.	Euclidean distance computed using QGIS
2010 Population Density (people per km^2)	WorldPop	Raster	0.008333 dd	2010 Population Density Range (people per km^2), with national total adjusted to match UN population division estimates
Agroecological Zone	IFPRI	Raster	0.008333 dd	Agro-ecological zones created using IFPRI standardized AEZ data based on elevation, climatology

Figure A1: Treatment Distribution by distance to a treated road

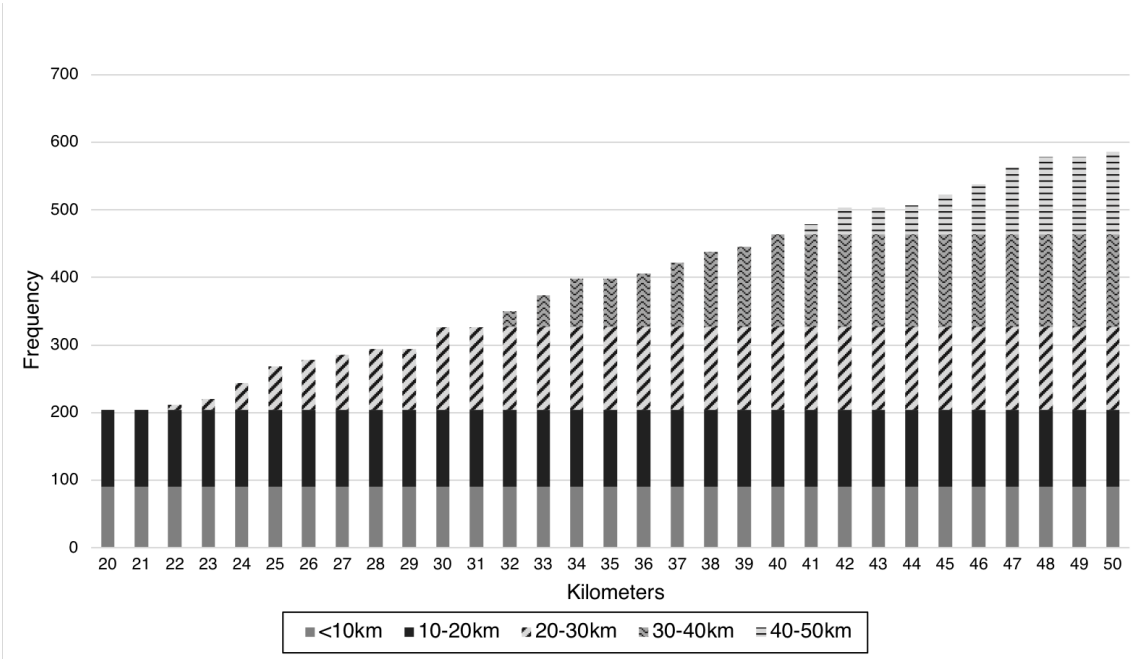


Table A2: LSMS-ISA Observations (3 rounds pooled)

	Household fixed effects-CS	
	Treated	Control
Outcome 7d		
20km treatment	676	2003
50km treatment	2223	2352
Outcome last year		
30km treatment	370	834
50km treatment	1237	1754
	Village fixed effects-CS	
	Treated	Control
Prices		
20km treatment	58	148
50km treatment	182	236

Table A3: LFS Observations (by round)

	Diff-in-diff-CS	
	Treated	Control
2006		
20km treatment	1055	2651
50km treatment	2967	3088
2014		
20km treatment	327	737
50km treatment	1047	996

Table A4: LSMS-ISA Summary Statistics

	Mean	(SD)		Mean	(SD)
Treatment			***Land Quality***		
Road rehabilitation 20km	0.117	(0.004)	Good quality*area	3.476	(0.238)
Road Rehabilitation 30km	0.189	(0.005)	Avg quality*area	2.313	(0.074)
Road Rehabilitation 50km	0.344	(0.006)	Bad quality*area	0.230	(0.020)
Household Characteristics			Soil quality missing	0.067	(0.003)
Head Age	48.184	(0.195)	***2008 Distance (KMs) to Nearest***		
Head Male	0.760	(0.005)	Major Road (GIS)	19.182	(0.273)
Fraction hh members 0-6 years old	0.205	(0.002)	Pop Center with 20'000	56.997	(0.495)
Fraction hh members 7-12 years old	0.147	(0.002)	Market	79.352	(0.644)
Fraction hh members 13-18 years old	0.140	(0.002)	Border Crossing	163.421	(1.336)
Fraction male hh members >18 years old	0.234	(0.002)	District Headquarters	34.408	(0.345)
Household size	5.872	(0.042)	***Quality of closest road***		
Education			Gravel	0.281	(0.006)
More than primary Head	0.474	(0.006)	Paved	0.207	(0.005)
Education's head missing	0.015	(0.002)	Sealed	0.482	(0.006)
N. adult members with > primary	1.290	(0.015)	Soil	0.030	(0.002)
Asset ownership (score)			***Village Characteristics***		
Durable Assets	−1.057	(0.020)	Average hh size in the ward	6.073	(0.025)
Productive Assets	0.095	(0.012)	Population Density	3.595	(0.032)
Farm's Characteristics			Population density missing	0.011	(0.001)
Land title	0.104	(0.004)	Share hhs with a member wage employed	0.141	(0.001)
Land rights	0.719	(0.006)	Share illiterate people >15 years old	0.302	(0.002)
Land rights missing	0.061	(0.003)	***Services available(less 10km)***		
Used org fertilizer	0.196	(0.005)	Bank	0.222	(0.005)
Used inorg fertilizer	0.111	(0.004)	SACCOs	0.512	(0.006)
Used pesticides	0.155	(0.004)	Market (daily)	0.385	(0.006)
Inputs missing	0.008	(0.001)	Market (weekly)	0.468	(0.006)
Erosion	0.222	(0.005)			
Irrigation	0.028	(0.002)			
Erosion/irrigation missing	0.068	(0.003)			
Tropic-warm/humid agroecological zone	0.549	(0.006)			

Note: Values at baseline.

Table A5: LFS Summary Statistics

	2006		2014	
	Mean	(SD)	Mean	(SD)
Treatment				
Road rehabilitation 20km	0.147	(0.035)	0.100	(0.030)
Road Rehabilitation 30km	0.216	(0.041)	0.271	(0.047)
Road Rehabilitation 50km	0.390	(0.049)	0.456	(0.055)
Household Characteristics				
Head Age	46.219	(0.385)	45.877	(0.510)
Head Male	0.740	(0.009)	0.744	(0.016)
Fraction hh members 0-6 years old	0.210	(0.005)	0.197	(0.006)
Fraction hh members 7-12 years old	0.135	(0.003)	0.141	(0.005)
Fraction hh members 13-18 years old	0.107	(0.002)	0.105	(0.004)
Fraction male hh members >18 years old	0.255	(0.004)	0.270	(0.008)
Fraction female hh members >18 years old	0.293	(0.004)	0.287	(0.007)
Household size	4.716	(0.101)	4.715	(0.122)
Education				
More than primary Head	0.493	(0.012)	0.548	(0.020)
Education's head missing	0.000	(omitted)	0.008	(0.002)
N. adult members with > primary	1.329	(0.037)	1.629	(0.071)
2008 Distance (KMs) to Nearest+				
Major Road (GIS)	20.570	(1.944)	22.049	(1.599)
Pop Center with >20'000	52.894	(3.266)	54.472	(3.583)
Border Crossing	194.920	(13.802)	180.682	(13.745)
District Headquarters	27.709	(1.718)	32.735	(1.704)
Quality of closest road+				
Gravel	0.275	(0.044)	0.354	(0.052)
Paved	0.177	(0.038)	0.153	(0.037)
Seated	0.509	(0.050)	0.493	(0.055)
Soil	0.039	(0.019)	0.000	(omitted)
Village Characteristics+				
Average hh size in the ward	5.946	(0.140)	5.956	(0.160)
Population Density	1.116	(0.180)	0.777	(0.071)
Share hhs with a member wage employed	0.147	(0.010)	0.146	(0.008)
Share illiterate people >15 years old	0.273	(0.012)	0.298	(0.014)
Tropic-warm/humid agroecological zone	0.481	(0.050)	0.419	(0.053)
Services available(less 2km)+				
Primary School	0.171	(0.005)	0.947	(0.011)
Clinic	0.114	(0.007)	0.664	(0.035)
Shop	0.178	(0.006)	0.936	(0.013)
Market	0.108	(0.008)	0.582	(0.036)

Note: Means and standard deviations computed using sample design weights. +2006 value

Table A6: LFS - Propensity Score Matching: Logit 20km

	2006		2014	
Household Characteristics				
Head Age	-0.0142	(0.014)	0.029	(0.031)
Head Age^2	0.0001	(0.000)	-0.0004	(0.000)
Head Male	-0.1131	(0.205)	0.1866	(0.342)
Fraction hh members 0-6 years old	0.2365	(0.859)	-0.0533	(0.461)
Fraction hh members 7-12 years old	0.2062	(0.702)	0.0266	(0.544)
Fraction hh members 13-18 years old	0.3251	(0.546)	-0.0877	(0.623)
Fraction male hh members >18 years old	0.4995	(0.387)	0.0302	(0.585)
Log. Household size	0.1629	(0.243)	0.0974	(0.167)
Education				
More than primary Head	0.0315	(0.110)	-0.2707	(0.217)
Education's head missing	0	(.)	-0.2814	(0.416)
N. adult members with > primary	-0.1228	(0.074)	-0.082	(0.101)
2008 Distance (KMs) to Nearest+				
Major Road	1.39E-06	(0.000)	-0.0002*	(0.000)
Pop Center with >20'000	-0.0293	(0.041)	0.0107	(0.038)
Border Crossing	-0.0036	(0.005)	0.0028	(0.008)
District Headquarters	-0.0253	(0.047)	-0.0759	(0.054)
Quality of closest road+				
Gravel	1.2289	(1.139)	3.7725*	(2.068)
Paved	-0.9823	(2.291)	2.7459*	(1.420)
Village Characteristics+				
Average hh size in the ward	-0.5713	(0.876)	-0.8228	(0.530)
Population Density	-1.6585	(1.857)	-1.0118	(0.743)
Share hhs with a member wage employed	-4.8053	(8.376)	10.0594**	(4.314)
Share illiterate people >15 years old	2.4564	(5.921)	-8.0934*	(4.081)
Tropic-warm/humid agroecological zone	2.0353**	(0.881)	0.6393	(1.141)
Services available(less 2km)+				
Primary School	-5.4388	(10.683)	17.1844	(26.026)
Clinic	-13.3242	(13.601)	0.8406	(1.902)
Shop	25.4096	(21.245)	9.5332	(15.896)
Market	-11.7854*	(6.435)	-1.2597	(2.629)
Constant	4.0781	(11.784)	-21.2738	(17.697)
Observations	4573		1415	

Note: The reported coefficients and standard errors are the result of a logit regression estimated separately for each wave. Sample design adjusted standard errors are presented in parentheses. + 2006 value. * p< 0.10, ** p<0.05, *** p<0.01, **** p<0.001

Table A7: LSMS-ISA Propensity Score Matching: Logit 20km

Household Characteristics			***Land Quality***		
Head Age	0.0157	(0.036)	Good quality*area	0.0068**	(0.003)
Head Age^2	−0.0002	(0.000)	Avg quality*area	0.0322*	(0.017)
Head Male	0.6300**	(0.299)	Bad quality*area	0.0563*	(0.031)
Fraction hh members 0-6 years old	0.5453	(0.972)	Soil quality missing	0.1356	(0.458)
Fraction hh members 7-12 years old	0.5078	(0.943)	***Quality of closest road***		
Fraction hh members 13-18 years old	1.4847*	(0.866)	Gravel	15.2924****	(0.368)
Fraction male hh members >18 years old	−0.2300	(0.733)	Paved	14.6676****	(0.419)
Log. Household size	−0.2234	(0.319)	Sealed	13.3548****	(0.469)
Education			***2008 Distance (KMs) to Nearest***		
More than primary Head	0.1289	(0.302)	Border Crossing	−0.0076****	(0.001)
Education's head missing	−1.6443**	(0.830)	Pop Center with >20'000	−0.0274****	(0.004)
N. adult members with > primary	−0.085	(0.163)	District Headquarters	0.0024	(0.004)
Asset ownership (score)			Major Road	0.0487***	(0.019)
Durable Assets	−0.1623*	(0.086)	Market	0.0065***	(0.002)
Productive Assets	−0.3861*	(0.213)	***Village Characteristics***		
Farm's Characteristics			Average hh size in the ward	0.0351	(0.062)
Land title	0.4395	(0.350)	Population Density	−0.0430	(0.043)
Land rights	0.2453	(0.261)	Population density missing	0.0459	(0.590)
Land rights missing	−0.0351	(0.410)	Share hhs with a member wage employed	−4.6657****	(1.185)
Used org fertilizer	0.3318	(0.265)	Share illiterate people >15 years old	1.9836***	(0.764)
Used inorg fertilizer	0.3977	(0.277)	***Services available(less 10km)***		
Used pesticides	−0.1366	(0.286)	Bank	−0.4224*	(0.251)
Erosion	−0.3751*	(0.222)	SACCOs	0.3555*	(0.210)
Irrigation	−0.9154*	(0.490)	Market (daily)	−0.3371	(0.225)
Tropic-warm/humid agroecological zone	0.4480**	(0.214)	Market (weekly)	0.0249	(0.223)
Constant	−15.8821****	(1.254)			
Pseudo R2	0.238				
Observations	1022				

Note: The reported coefficients and standard errors are the result of a logit regression estimated at baseline. Heteroscedasticity corrected and clustered standard errors are presented in parentheses. * p< 0.10, ** p<0.05, *** p<0.01, **** p<0.001

Figure A2: LSMS-ISA Common Support Treatment 20km

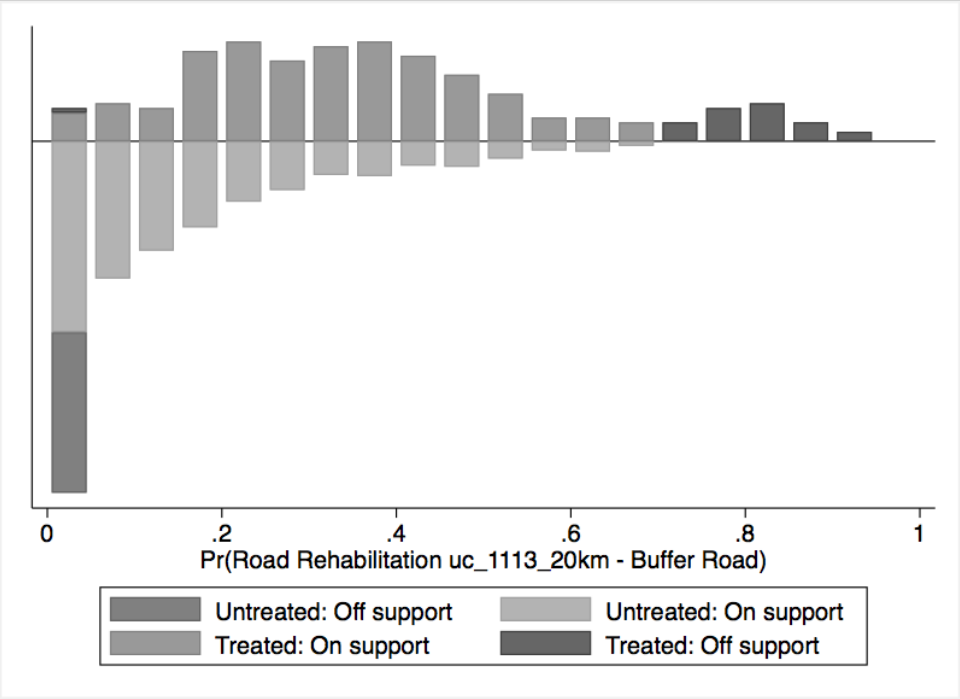


Figure A3: LSMS-ISA Common Support Treatment 30km

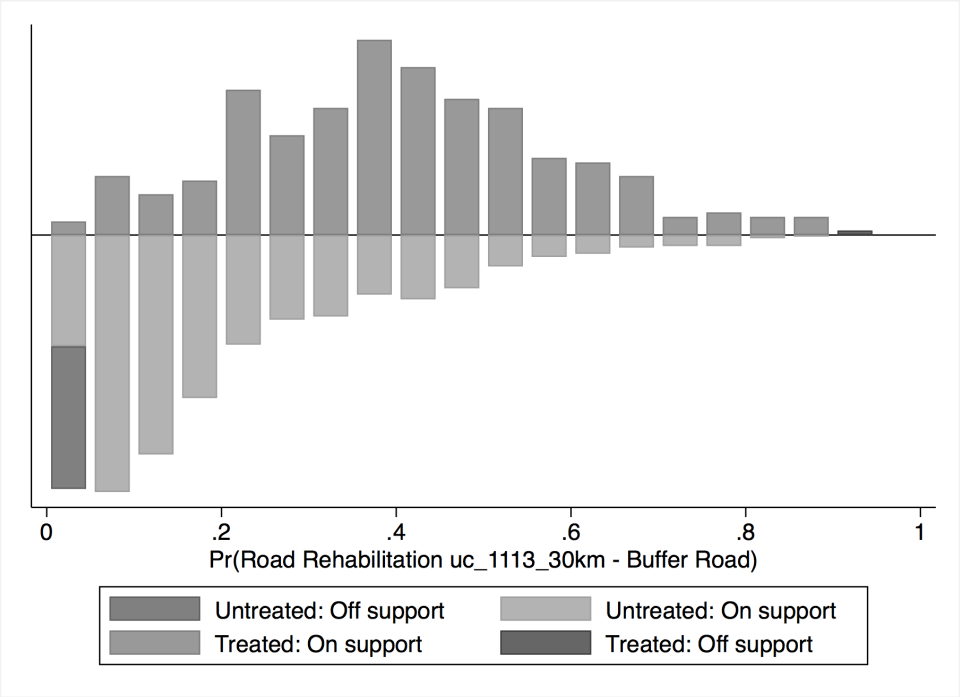


Figure A4: LSMS-ISA Common Support Treatment 20km

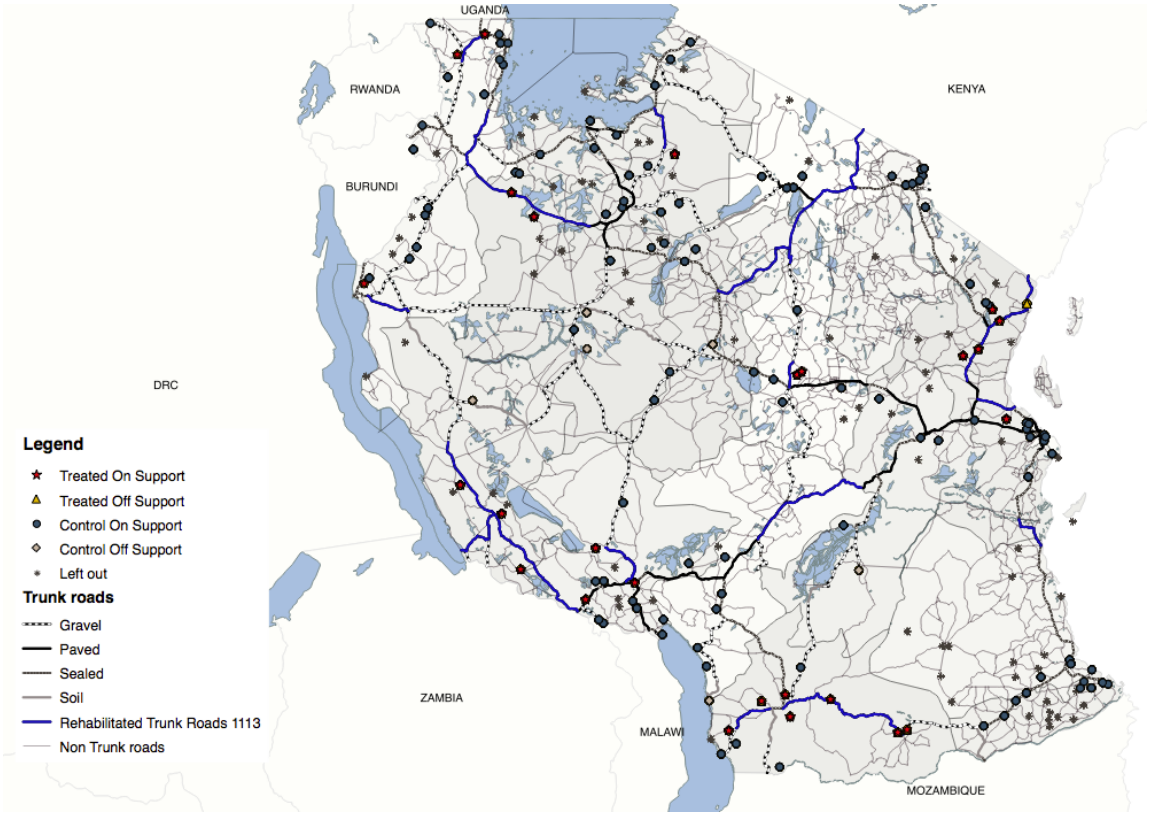


Table A8: LSMS-ISA Balancing Test

Variable	Mean Treated	Mean Control	Diff.	
	Panel A			
Household Characteristics				
Head Male	0.789	0.781	0.008	(0.028)
Head Age	47.760	48.005	−0.245	(1.109)
Head Age^2	2536.926	2574.646	−37.721	(116.940)
Fraction hh members 0-6 years old	0.204	0.206	−0.002	(0.013)
Fraction hh members 7-12 years old	0.154	0.152	0.002	(0.010)
Fraction hh members 13-18 years old	0.139	0.135	0.004	(0.011)
Fraction male hh members >18 years old	0.238	0.238	0.000	(0.014)
Household size	5.286	5.203	0.083	(0.180)
Education				
More than primary Head	0.497	0.513	−0.016	(0.034)
Education's head missing	0.006	0.004	0.002	(0.005)
N. adult members with > primary	1.143	1.206	−0.063	(0.073)
Asset ownership (score)				
Durable Assets	−1.305	−1.260	−0.045	(0.083)
Productive Assets	−0.040	−0.038	−0.001	(0.029)
Farm Characteristics				
Land title	0.103	0.104	−0.001	(0.021)
Land rights	0.743	0.740	0.003	(0.030)
Land rights missing	0.046	0.041	0.005	(0.014)
Used org fertilizer	0.229	0.231	−0.003	(0.029)
Used inorg fertilizer	0.211	0.188	0.024	(0.027)
Used pesticides	0.149	0.143	0.005	(0.024)
Erosion	0.246	0.250	−0.004	(0.030)
Irrigation	0.029	0.017	0.012	(0.010)
Erosion/Irrigation missing	0.063	0.065	−0.002	(0.017)
Tropic-warm/humid agroecological zone	0.469	0.462	0.006	(0.034)
Land Quality				
Good quality*area	3.292	3.350	−0.058	(1.464)
Avg quality*area	2.377	2.504	−0.127	(0.352)
Bad quality*area	0.245	0.256	−0.011	(0.130)
Soil quality missing	0.063	0.065	−0.002	(0.017)
2008 Distance (Km) to Nearest				
Major Road	8.558	8.543	0.015	(0.433)
Pop Center with >20'000	45.574	43.667	1.908	(2.042)
Market	75.599	75.371	0.228	(3.185)
Border Crossing	134.439	134.037	0.402	(6.639)
District Headquarters	32.211	30.172	2.039	(1.844)
Quality of closest road				
Gravel	0.406	0.431	−0.026	(0.034)
Paved	0.269	0.208	0.061**	(0.029)
Sealed	0.326	0.361	−0.035	(0.032)

Table A8: LSMS-ISA Balancing Test (continued)

Variable	Mean Treated	Mean Control	Diff.	
Village Characteristics				
Average hh size in the ward	5.653	5.634	0.019	(0.098)
Population Density	3.171	3.215	−0.044	(0.153)
Population density missing	0.023	0.022	0.001	(0.010)
Sh. hhs with a member wage employed	0.110	0.110	0.000	(0.006)
Sh. illiterate people >15 years old	0.299	0.294	0.005	(0.009)
Services available (less 10km)				
Bank	0.189	0.212	−0.023	(0.027)
SACCOs	0.583	0.565	0.018	(0.034)
Market (daily)	0.291	0.285	0.006	(0.031)
Market (weekly)	0.514	0.548	−0.034	(0.034)
	Panel B			
Outcome Variables				
Subjective Welfare: Life	3.895	3.688	0.206	(0.127)
Durable Assets (score)	−1.305	−1.260	−0.045	(0.083)
Productive Assets (score)	−0.040	−0.038	−0.001	(0.029)
Consumption rice (kg) 7d	1.617	2.970	−1.354****	(0.328)
Consumption maize (kg) 7d	9.967	8.234	1.734	(1.343)
Outcome Variables last year+				
Wage labor	0.348	0.295	0.053	(0.047)
Hire labor	0.379	0.395	−0.016	(0.049)
N. days agri. labor by hhm	102.136	110.590	−8.453	(9.907)
Harvest rice (kg)	192.576	9.663	182.913****	(43.691)
Sales rice (kg)	137.258	4.476	132.781***	(41.382)
Harvest maize (kg)	293.485	402.741	−109.256**	(46.718)
Sales maize (kg)	34.015	114.464	−80.448****	(22.264)
Log PC real annual consumption	13.050	13.027	0.023	(0.057)
N. members away at least 1m last 12m	0.167	0.260	−0.093	(0.057)

Note: The reported coefficients and standard errors are from OLS regression models using PSM weights on the common support sample and performed separately for each variable. Test using baseline observations within 20km of any major road. Standard errors are reported in parenthesis. * p< 0.10, ** p<0.05, *** p<0.01.
+Test at 30km

Table A9: LFS Balancing Test

Variable	Mean Treated	Mean Control	Diff.	
Panel A: 2006 characteristics				
Household Characteristics				
Head Male	0.743	0.723	0.020	(0.031)
Head Age	46.211	44.786	1.426	(1.284)
Head Age^2	2406.234	2252.429	153.805	(127.237)
Fraction hh members 0-6 years old	0.208	0.243	−0.035*	(0.019)
Fraction hh members 7-12 years old	0.132	0.137	−0.005	(0.013)
Fraction hh members 13-18 years old	0.108	0.102	0.006	(0.008)
Fraction male hh members >18 years old	0.267	0.238	0.028*	(0.017)
Fraction female hh members >18 years old	0.285	0.280	0.006	(0.012)
Household size	4.500	4.836	−0.336	(0.303)
Education				
More than primary Head	0.460	0.437	0.023	(0.044)
Education's head missing	0.000	0.000	0.000	(.)
N. adult members with > primary	1.145	1.132	0.013	(0.125)
2008 Distance (Km) to Nearest+				
Major Road	7.307	7.477	−0.170	(1.884)
Pop Center with >20'000	38.272	54.189	−15.918	(11.048)
Border Crossing	173.804	184.530	−10.726	(42.717)
District Headquarters	23.860	24.753	−0.893	(4.650)
Quality of closest road+				
Gravel	0.419	0.493	−0.074	(0.177)
Paved	0.166	−0.015	0.181	(0.116)
Sealed	0.415	0.500	−0.085	(0.195)
Soil	−0.001	0.021	−0.022	(0.019)
Village Characteristics+				
Average hh size in the ward	5.638	5.938	−0.300	(0.373)
Population Density	0.638	0.622	0.016	(0.138)
Share hhs with a member wage employed	0.149	0.119	0.031	(0.028)
Share illiterate people >15 years old	0.270	0.317	−0.048	(0.045)
Tropic-warm/humid agroecological zone	0.600	0.324	0.275	(0.169)
Services available (less 2km)+				
Primary school	0.175	0.157	0.018	(0.015)
Clinic	0.105	0.091	0.014	(0.023)
Shop	0.196	0.160	0.0357**	(0.017)
Market	0.093	0.081	0.012	(0.019)

Table A9: LFS Balancing Test (continued)

Variable	Mean Treated	Mean Control	Diff.	
Panel B: 2014 characteristics				
Household Characteristics				
Head Male	0.798	0.720	0.078	(0.086)
Head Age	42.749	48.277	−5.528*	(2.981)
Head Age^2	2023.860	2625.183	−601.323*	(326.210)
Fraction hh members 0-6 years old	0.204	0.165	0.039**	(0.019)
Fraction hh members 7-12 years old	0.147	0.152	−0.005	(0.015)
Fraction hh members 13-18 years old	0.103	0.128	−0.025	(0.017)
Fraction male hh members >18 years old	0.282	0.272	0.011	(0.038)
Fraction female hh members >18 years old	0.263	0.283	−0.020	(0.033)
Household size	4.558	4.299	0.259	(0.250)
Education				
More than primary Head	0.582	0.529	0.053	(0.070)
Education's head missing	0.012	0.006	0.007	(0.008)
N. adult members with > primary	1.603	1.529	0.074	(0.176)
2008 Distance (Km) to Nearest+				
Major Road	9.037	10.660	−1.624	(2.157)
Pop Center with >20.000	45.422	45.410	0.013	(11.506)
Border Crossing	137.271	167.211	−29.940	(49.786)
District Headquarters	24.435	24.449	−0.014	(4.217)
Quality of closest road+				
Gravel	0.560	0.621	−0.060	(0.205)
Paved	0.270	0.178	0.091	(0.149)
Sealed	0.170	0.201	−0.031	(0.150)
Soil	0.000	0.000	0.000	(.)
Village Characteristics+				
Average hh size in the ward	5.661	5.274	0.387	(0.301)
Population Density	0.936	0.726	0.209	(0.273)
Share hhs with a member wage employed	0.198	0.180	0.018	(0.035)
Share illiterate people >15 years old	0.255	0.255	0.001	(0.039)
Tropic-warm/humid agroecological zone	0.410	0.273	0.137	(0.159)
Services available (less 2km)+				
Primary school	0.976	0.983	−0.007	(0.012)
Clinic	0.687	0.790	−0.103	(0.114)
Shop	0.971	0.978	−0.007	(0.021)
Market	0.565	0.741	−0.175*	(0.101)

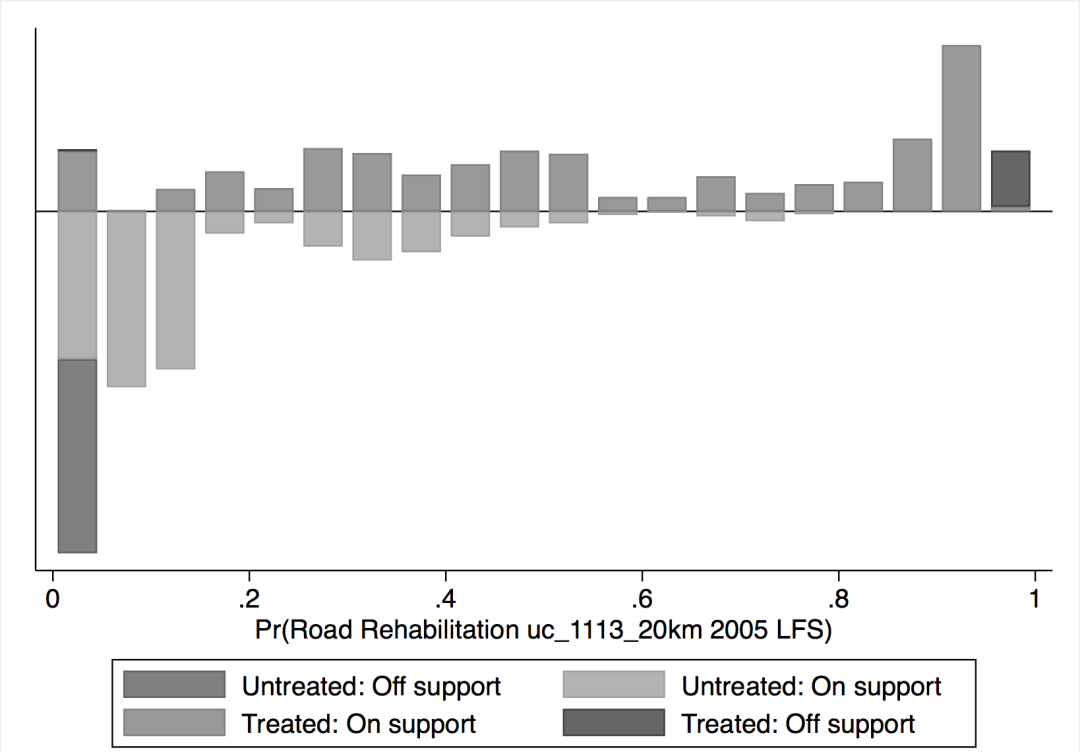
Table A9: LFS Balancing Test (continued)

Variable	Mean Treated	Mean Control	Diff.	
Panel C: 2006 characteristics				
Outcome Variables				
N. household members:				
Have current act.	2.476	2.341	0.134	(0.213)
Earned wage	0.072	0.077	−0.006	(0.023)
Self employed	0.168	0.102	0.066	(0.076)
Work unpaid	0.100	0.104	−0.005	(0.037)
Work on farm	2.417	2.411	0.006	(0.201)
Isic: work in agriculture	2.450	2.440	0.010	(0.196)
Isic: work in manufacturing	0.078	0.037	0.041	(0.044)
Isic: work in construction	0.006	0.009	−0.003	(0.005)
Isic: work in retail	0.053	0.056	−0.003	(0.021)
Isic: work in transport	0.003	0.007	−0.005	(0.004)
Isic: work in hospitality industry	0.030	0.018	0.012	(0.016)
Isic: work as domestic personnel	0.096	0.105	−0.010	(0.037)
Log Hh income paid & self employment (2006 TSh)	10.288	10.141	0.147	(0.234)

The reported coefficients and standard errors are from OLS regression models using PSM weights on the common support sample and estimated separately for each variable. Test using observations within 20km of any major road. Sample design adjusted standard errors are reported in parenthesis. * p< 0.10, ** p<0.05, *** p<0.01.
+ 2006 value.

Figure A5: LFS Common Support Treatment 20km

(a) Common Support 2006



(b) Common Support 2014

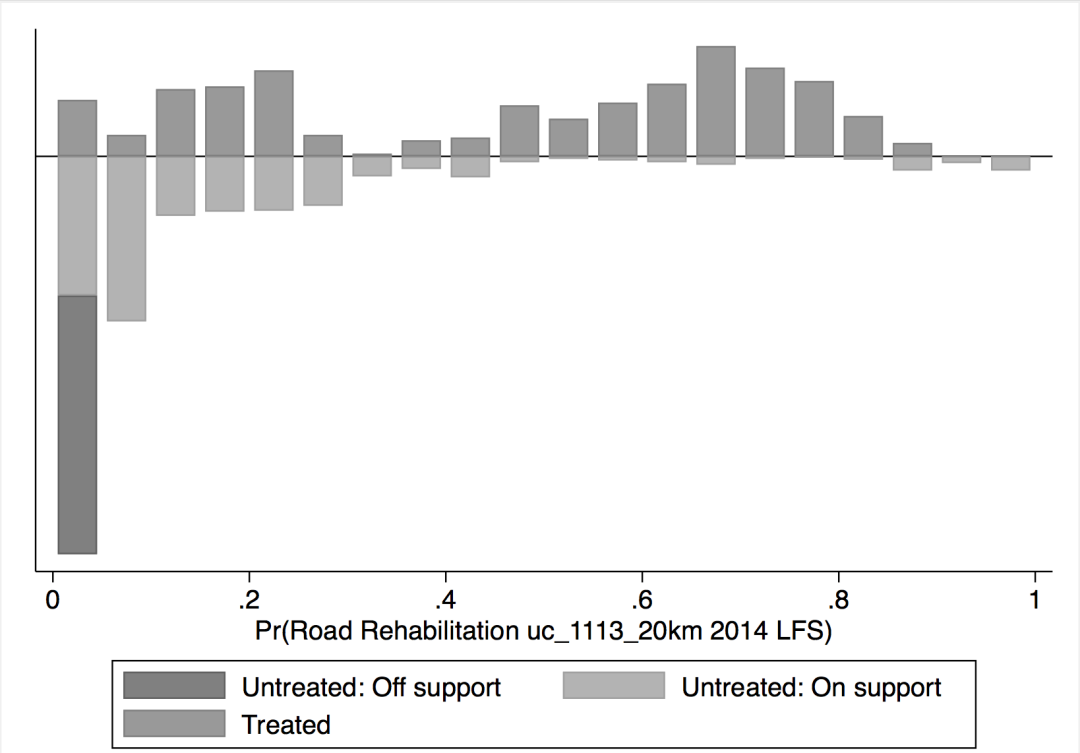
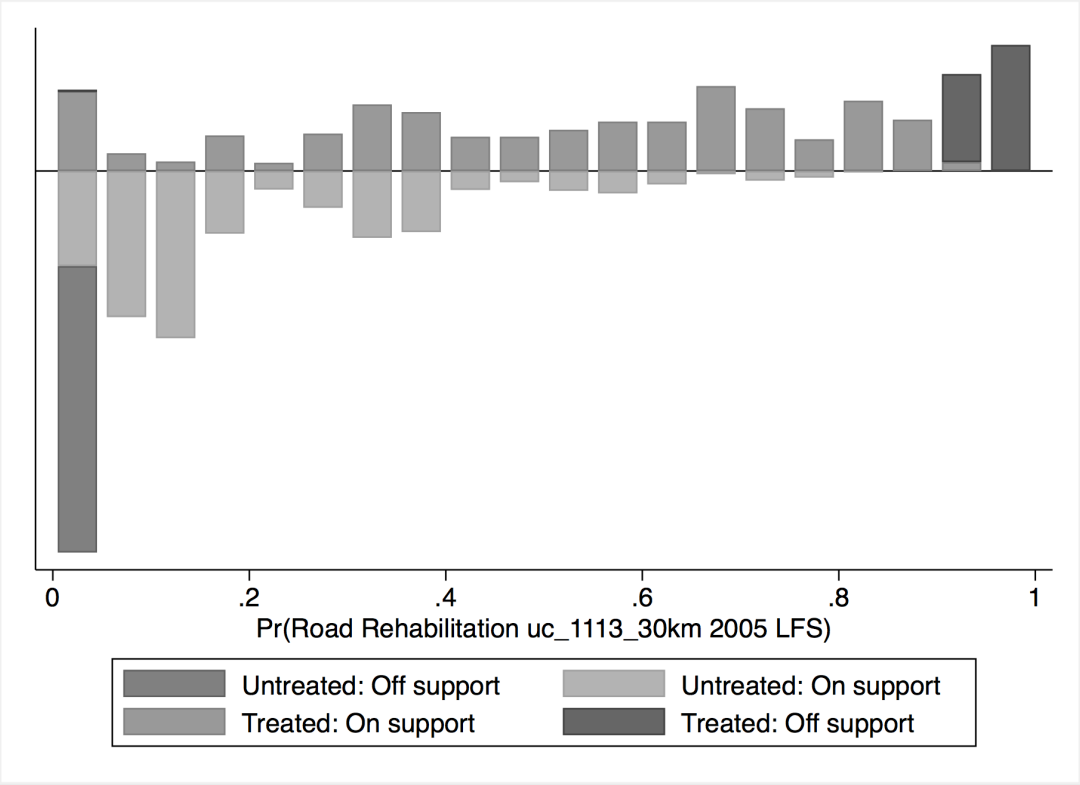


Figure A6: LFS Common Support Treatment 30km

(a) Common Support 2006



(b) Common Support 2014

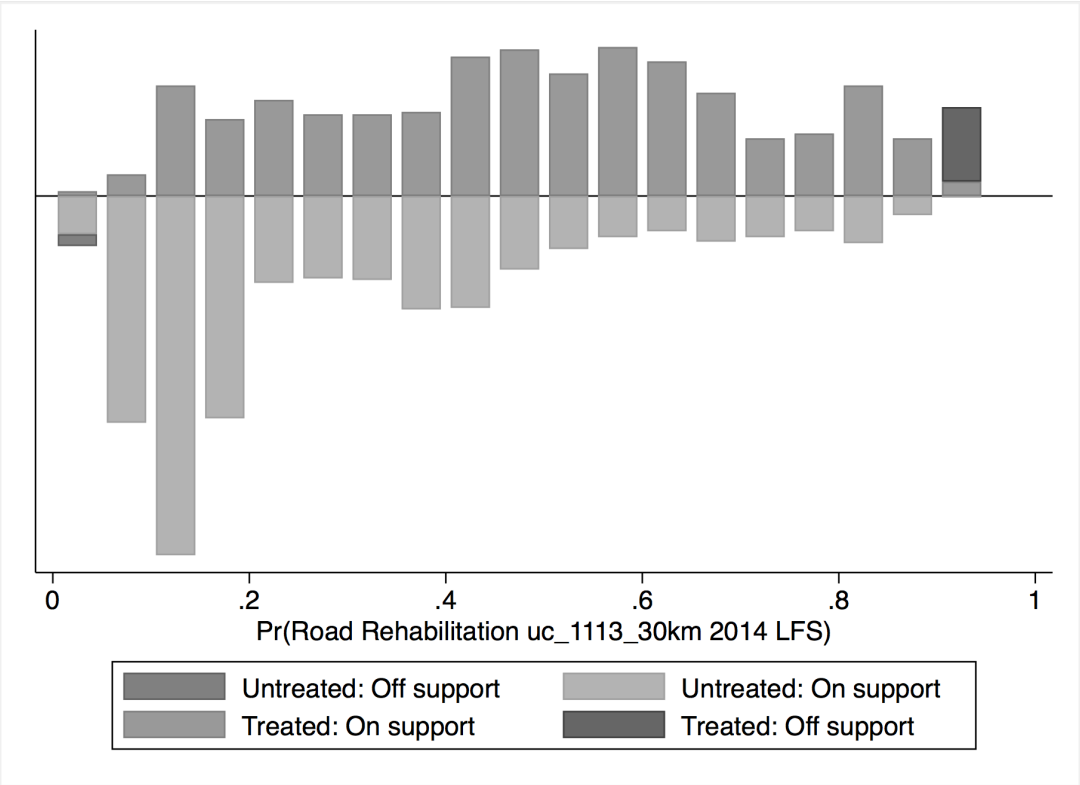
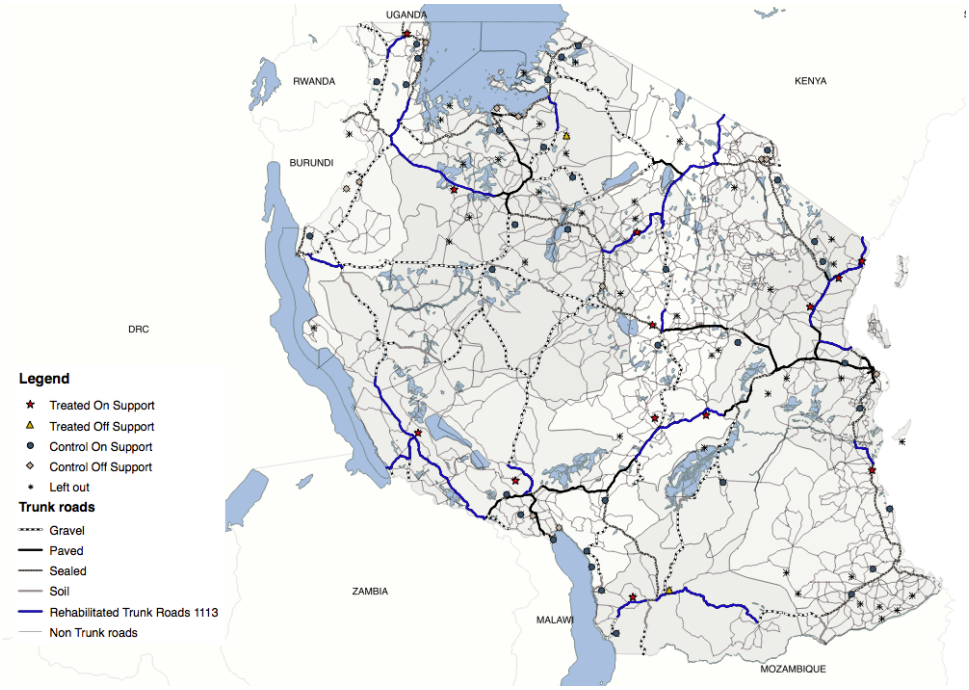


Figure A7: LFS Common Support Treatment 20km

(a) Common Support 2006



(b) Common Support 2014

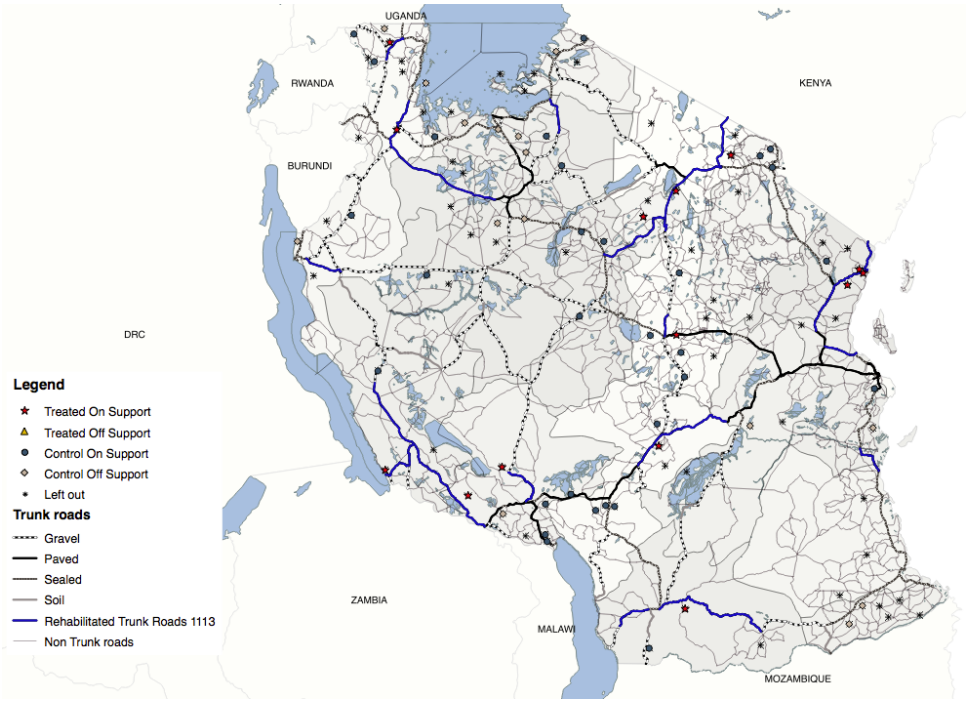
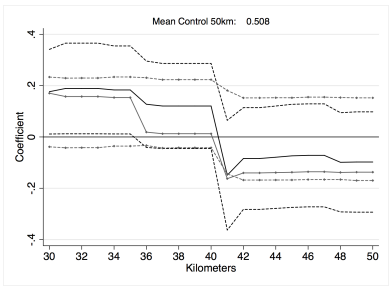
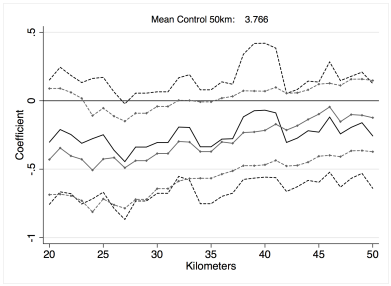


Figure A8: Common Trend

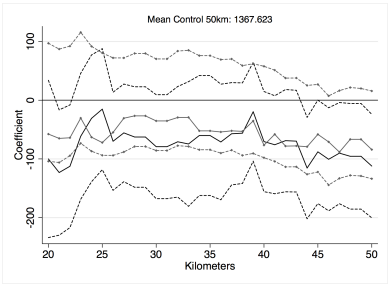
(a) N. members away at least 1 month last 12m



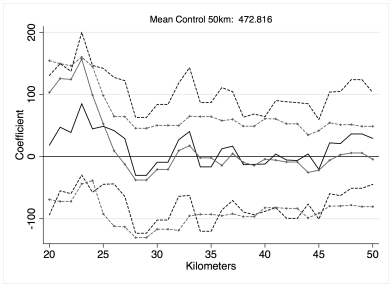
(b) Life Satisfaction



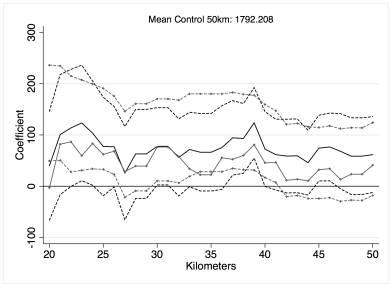
(c) Price rice husked (TSh)



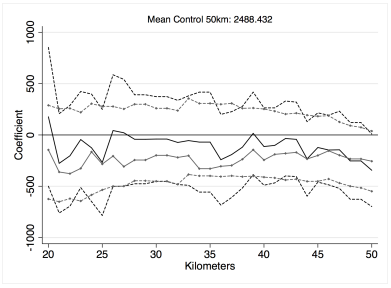
(d) Price maize grain (TSh)



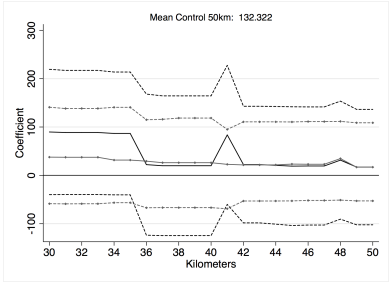
(e) Price sugar (TSh)



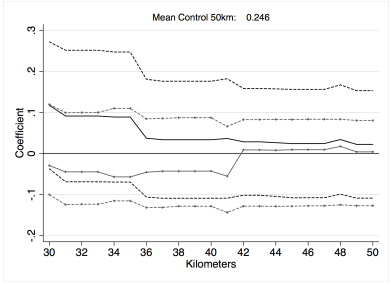
(f) Price kerosene (TSh)



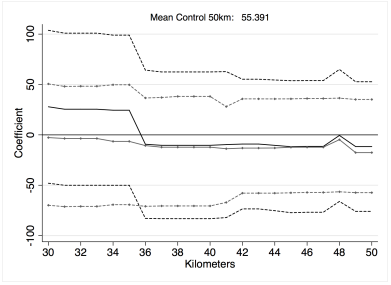
(g) Harvest rice (kg)



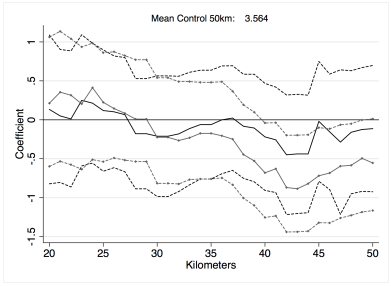
(h) Land area: rice (acres)



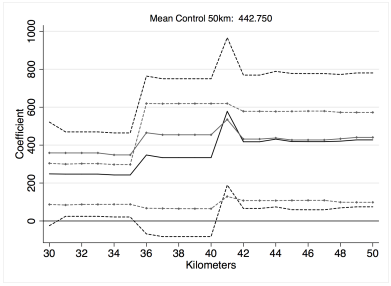
(i) Sales rice (kg)



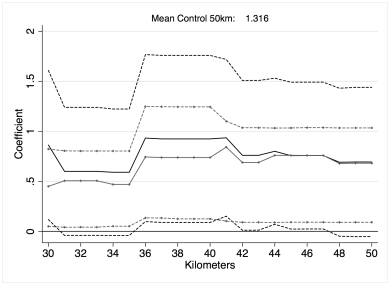
(j) Consumption rice (kg)



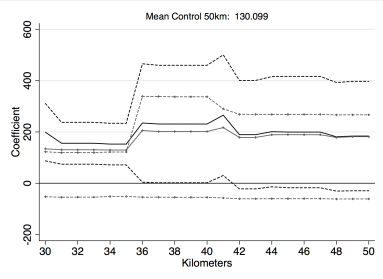
(k) Harvest maize (kg)



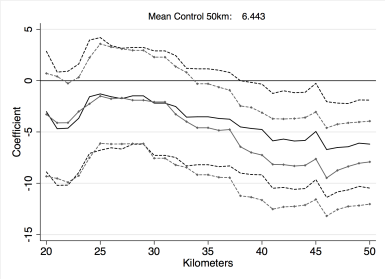
(l) Land area: maize (acres)



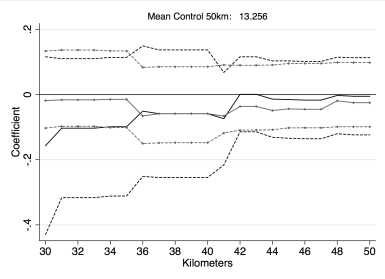
(m) Sales maize (kg)



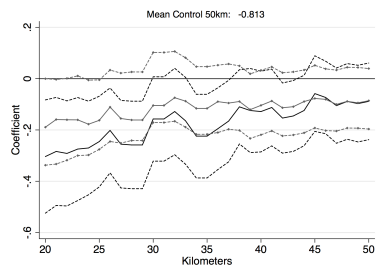
(n) Consumption maize (kg)



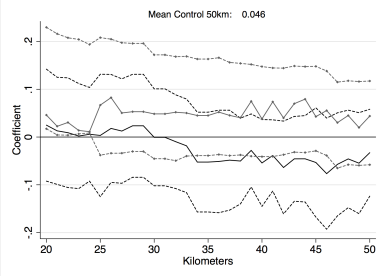
(o) Log PC real annual consumption



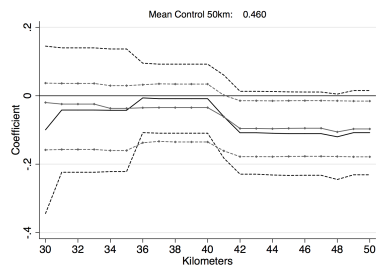
(p) Durable Assets (score)



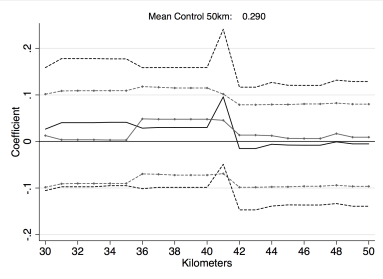
(q) Productive Assets (score)



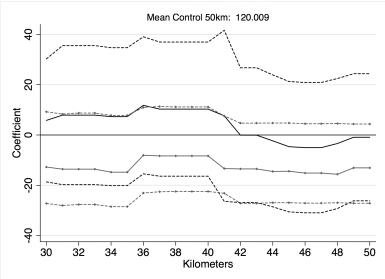
(r) Wage labor



(s) Hired labor



(t) N. days agricultural labor by household members



Note: The figure shows estimates from the interaction term between the treatment and the 2010 wave from both a household fixed effects strategy with propensity score matching and from a household fixed effects strategy without propensity score but on the common support sample estimated separately for each treatment cut-off. Confidence intervals are at the 90% level.

Table A10: Treatment effects (LSMS-ISA)

	less than 30km			30-50km		
	Coef	Mean Control	Obs.	Coef	Mean Control	Obs.
Subjective Welfare: Life	−0.5236***	3.728	4871	0.5377	3.917	657
	(0.175)			(0.479)		
Price rice husked (TSh)	−187.9508***	1360.510	476	208.9884	1434.091	60
	(66.793)			(165.065)		
Supply labor	−0.0059	0.450	4958	0.2179	0.440	580
	(0.073)			(0.253)		
Hire labor	0.0314	0.298	4958	0.0885	0.295	580
	(0.066)			(0.223)		
Sales rice (kg)	−53.4764**	64.869	4491	123.4664*	41.788	524
	(20.659)			(63.137)		
Harvest rice (kg)	−101.6859*	149.267	4524	205.4054	108.374	529
	(54.772)			(123.255)		
Harvest area rice (acres)	−0.3071**	0.277	4524	−0.0153	0.274	529
	(0.142)			(0.202)		

Note: The reported treatment effects are the result of a household fixed effects strategy without propensity score matching. We report results of the parameter δ_{1d} of equation 1 for each outcome of interest. Robust standard errors clustered at the village level in parenthesis. * p< 0.10, ** p<0.05, *** p<0.01.

Table A11: Treatment effects (LFS)

	less than 30km			30-50km		
	Coef	Mean Control	Obs.	Coef	Mean Control	Obs.
N. hhm have current act.	−0.4714***	2.469	7759	−0.0526	2.456	1667
	(0.158)			(0.379)		
N. hhm earned wage	0.0189	0.120	7625	0.0968	0.029	1640
	(0.041)			(0.059)		
N. hhm work on farm	−0.3454**	2.463	7625	−0.0082	2.521	1640
	(0.166)			(0.415)		
Log Hh income paid &	−0.1969	104021.710	3164	0.1548	87449.886	496
self employment 2005 TSh	(0.193)			(0.333)		

Note: The reported treatment effects are the result of a difference-in-difference strategy, without propensity score matching. We report results of the parameter δ_{1d} of equation 1 for each outcome of interest. Robust standard errors clustered at the village level in parenthesis. * p< 0.10, ** p<0.05, *** p<0.01. The mean of the variable "Log Hh income paid & self employment" is expressed in TSh.

Table A12: Additional outcome variables

Variable	less than 20km		less than 30km		less than 40km	
Community level projects						
Road construction			0.1468	(0.371)	0.2807	(0.228)
Market construction			−0.2	(0.204)	−0.1426	(0.135)
Irrigation schemes			0.099	(0.115)	−0.1417	(0.136)
Water supply			0	(0.292)	−0.0871	(0.191)
Grain Storage constr.			0	(.)	0.0207	(0.022)
Health center/dispensary constr.			−0.4	(0.249)	−0.2956	(0.265)
Health center/dispensary maintenance			−0.101	(0.319)	−0.0704	(0.154)
Other possible effects						
Theft losses 12m			−13840.775	(11537.279)	−4473.6273	(9464.544)
Had a credit 12m			−0.0357	(0.075)	−0.0515	(0.068)
Used pesticides			0.0516	(0.064)	0.0164	(0.057)
Used org fertilizer			0.3369**	(0.132)	0.2603***	(0.093)
Used inorg fertilizer			0.0224	(0.044)	0.0121	(0.027)
N. livestock currently owned	0.4469	(0.411)	0.1802	(0.401)	−0.552	(0.392)
N. goods with prices	1.3454	(3.044)	−1.4064	(2.441)	−1.8362	(2.529)
N. wage labor sectors cluster	−0.6168	(0.425)	−0.2162	(0.347)	−0.0256	(0.322)

Note: The reported treatment effects are the result of a household fixed effects strategy with propensity score matching. We report results of the parameter δ_{1d} of equation 1 for each outcome of interest. Robust standard errors clustered at the village level in parenthesis.
* p< 0.10, ** p<0.05, *** p<0.01, **** p<0.001.

Table A13: Test of the common trend assumption with the wealth index from the DHS

	less than 20km		**less than 30km**		**less than 40km**	
Road	0.0649	(0.111)	0.0338	(0.093)	0.0439	(0.073)
Road*2003	−0.0727	(0.113)	−0.0581	(0.097)	−0.0962	(0.087)
Road*2007	−0.1559	(0.125)	−0.1016	(0.098)	−0.073	(0.085)
Road*2010	−0.0089	(0.104)	0.0167	(0.093)	0.0038	(0.064)
Observations	10.503		12.683		14.019	
Mean Y Control	−0.3784		−0.4009		−0.4173	

Note: Robust standard errors clustered at the village level in parenthesis.
* p< 0.10, ** p<0.05, *** p<0.01, **** p<0.001.

Table A14: Placebo Test

Variable	less than 20km		less than 30km		less than 40km	
Subjective Welfare: Life	0.3278	(0.324)	0.4467*	(0.264)	0.3715	(0.312)
Durable Assets (score)	0.0428	(0.137)	−0.0495	(0.131)	0.0047	(0.136)
Productive Assets (score)	0.0161	(0.042)	0.0184	(0.039)	−0.0088	(0.046)
Consumption rice (kg) last 7d	−0.8789	(1.193)	−1.1156	(1.103)	−1.2917	(0.901)
Consumption maize (kg) last 7d	10.2044***	(3.243)	4.6897**	(2.031)	6.9498****	(1.979)
Supply labor	0.007	(0.071)	0.0828	(0.071)	0.0415	(0.065)
Hire labor	0.0897	(0.086)	0.0162	(0.053)	0.04	(0.060)
N. days agri. labor by hhm	9.3365	(37.820)	−9.9176	(21.782)	2.2501	(21.056)
Sales maize (kg)	2.9122	(39.408)	−48.6289	(44.050)	−5.3838	(61.195)
Sales rice (kg)	102.1606**	(42.962)	126.9193***	(43.459)	87.9155***	(31.355)
Harvest maize (kg)	125.253	(135.333)	−41.1142	(136.821)	16.7109	(134.862)
Harvest rice (kg)	49.1632	(53.903)	42.1732	(60.939)	11.7345	(45.574)
Harvest area maize (acres)	0.3429	(0.380)	0.0282	(0.300)	−0.0601	(0.265)
Harvest area rice (acres)	−0.1289	(0.155)	−0.0293	(0.178)	−0.1782	(0.242)
N. hhm have current act.	0.7101	(0.661)	0.0999	(0.239)	0.4095*	(0.225)
N. hhm earned wage	0.009	(0.083)	−0.0176	(0.045)	−0.0576	(0.043)
N. hhm self employed	−0.2531**	(0.107)	0.1792**	(0.080)	0.0486	(0.058)
N. hhm work on farm	1.2631***	(0.453)	−0.3542	(0.247)	0.152	(0.168)
N. hhm work unpaid	0.1085*	(0.059)	0.0499	(0.046)	0.063	(0.044)

Note: The reported treatment effects are the result of a household fixed effects strategy with propensity score matching. Robust standard errors clustered at the village level in parenthesis. Treatment is defined based on roads that have been rehabilitated after 2013. Treated observations prior 2013 are excluded. * p< 0.10, ** p<0.05, *** p<0.01, **** p<0.001.

Table A15: Effect using control villages from non-treated regions

Variable	**less than 20km**		**less than 30km**		**less than 40km**	
Subjective Welfare: Life	−0.9829***	(0.309)	−0.8382****	(0.228)	−0.5485*	(0.302)
Durable Assets (score)	−0.4637**	(0.182)	−0.2794***	(0.088)	−0.1985	(0.140)
Productive Assets (score)	−0.1402	(0.090)	−0.0218	(0.062)	−0.0556	(0.050)
Consumption rice (kg) last 7d	0.4231	(0.857)	0.6152	(0.652)	1.6417*	(0.891)
Consumption maize (kg) last 7d	−3.914	(4.211)	−9.2691***	(2.910)	−6.5105***	(2.126)
Price rice husked (TSh)	−195.2169*	(114.356)	−289.6204***	(100.288)	−204.4189**	(92.423)
Price maize grain (TSh)	−56.792	(157.367)	−177.5162*	(89.441)	−96.4121*	(55.659)
Supply labor			0.0772	(0.197)	0.1368	(0.101)
Hire labor lrs			−0.0019	(0.191)	−0.0187	(0.118)
N. days agri. labor by hhm			−25.2012	(27.614)	−35.4668	(24.869)
Sales maize (kg)			17.9096	(49.773)	−43.5988	(91.293)
Sales rice (kg)			−31.8329	(20.883)	−53.3199*	(28.531)
Harvest maize (kg)			171.9127	(139.394)	154.6012	(168.536)
Harvest rice (kg)			−52.4941	(70.546)	−140.7710*	(81.807)
Harvest area maize (acres)			−0.2305	(0.382)	−0.0264	(0.352)
Harvest area rice (acres)			−0.2906	(0.188)	−0.4109**	(0.185)
N. hhm have current act.	−0.1717	(0.200)	0.1982	(0.153)	0.3398**	(0.152)
N. hhm earned wage	0.0298	(0.050)	0.025	(0.050)	−0.0993*	(0.059)
N. hhm self employed	0.0138	(0.131)	−0.0118	(0.083)	0.0273	(0.095)
N. hhm work on farm	−0.2087	(0.259)	0.1941	(0.189)	0.4675**	(0.193)
N. hhm work unpaid	−0.0192	(0.042)	−0.0343	(0.035)	−0.0107	(0.031)

Note: The reported treatment effects are the result of a household fixed effects strategy with propensity score matching using control unit from regions that did not received a road. Robust standard errors clustered at the village level in parenthesis.
* p< 0.10, ** p<0.05, *** p<0.01, **** p<0.001.

Table A16: Effects excluding villages close to cities

Variable	**less than 20km**		**less than 30km**		**less than 40km**	
Subjective Welfare: Life	−0.5171*	(0.263)	−0.3880*	(0.202)	−0.2561	(0.266)
Durable Assets (score)	−0.4353***	(0.146)	−0.2795***	(0.100)	−0.2794***	(0.099)
Productive Assets (score)	−0.1207**	(0.047)	−0.0676	(0.057)	−0.1054*	(0.058)
Consumption rice (kg) last 7d	0.2702	(0.592)	0.1391	(0.453)	−0.0415	(0.361)
Consumption maize (kg) last 7d	−2.7404	(3.232)	−4.1287	(2.689)	−5.0755**	(2.500)
Price rice husked (TSh)	−71.3772	(94.988)	−244.3996***	(84.193)	−329.7070****	(69.804)
Price maize grain (TSh)	−99.2251	(74.585)	−93.0238	(61.628)	−87.9145	(55.551)
Supply labor			0.1774	(0.118)	0.1907*	(0.105)
Hire labor lrs			0.0827	(0.115)	0.1467	(0.120)
N. days agri. labor by hhm			−3.0379	(20.072)	−12.6901	(19.572)
Sales maize (kg)			47.2982	(37.489)	34.0171	(51.834)
Sales rice (kg)			−27.4538	(19.427)	−29.8342	(19.644)
Harvest maize (kg)			214.8563*	(114.853)	279.0319	(176.346)
Harvest rice (kg)			−43.3263	(65.858)	−69.6804	(66.361)
Harvest area maize (acres)			0.237	(0.317)	0.3381	(0.309)
Harvest area rice (acres)			−0.2334	(0.179)	−0.3236*	(0.176)
N. hhm have current act.	−0.1753	(0.140)	−0.3782***	(0.133)	−0.1283	(0.158)
N. hhm earned wage	0.06	(0.046)	0.0428	(0.036)	0.0342	(0.032)
N. hhm self employed	0.1284	(0.104)	0.0605	(0.070)	0.1246*	(0.066)
N. hhm work on farm	−0.3474*	(0.203)	−0.3458**	(0.159)	−0.0794	(0.170)
N. hhm work unpaid	0.0097	(0.030)	−0.0182	(0.028)	−0.0516**	(0.025)

Note: The reported treatment effects are the result of a household fixed effects strategy with propensity score matching excluding observations within 5 km of a major city (> 5000 inhabitants). Robust standard errors clustered at the village level in parenthesis.
* p< 0.10, ** p<0.05, *** p<0.01, **** p<0.001.

Figure A9: Treated Major Cities (Population > 5000)

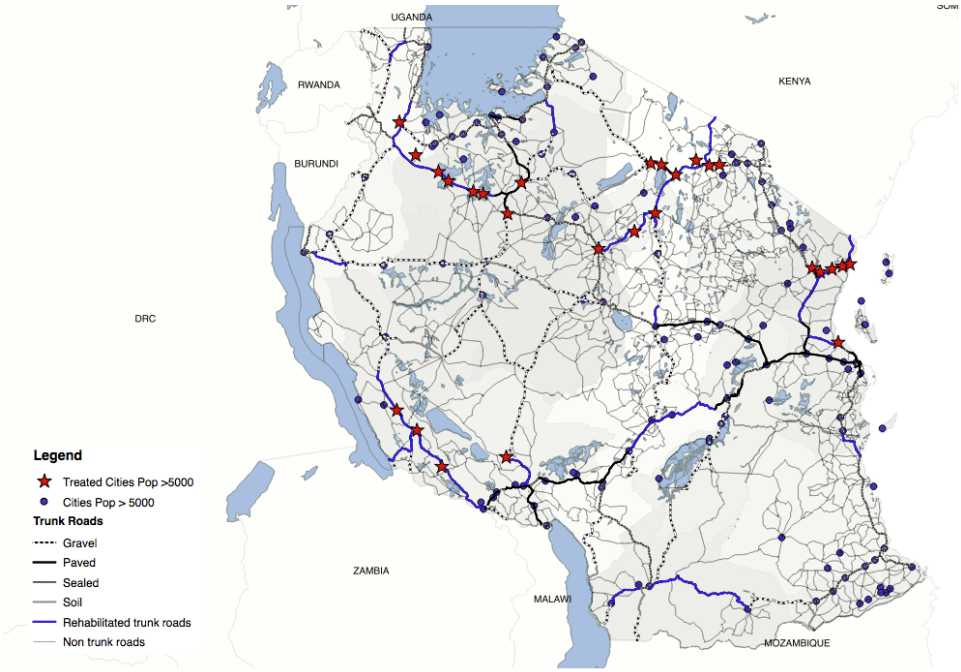


Table A17: Effect using a treatment defined based on distance to a major city (> 5000 inhabitants)

Variable	less than 20km		less than 30km		less than 40km	
Subjective Welfare: Life	−0.384	(0.350)	−0.2791	(0.329)	−0.3538	(0.248)
Durable Assets (score)	−0.135	(0.354)	−0.1111	(0.223)	−0.3604***	(0.130)
Productive Assets (score)	−0.0449	(0.053)	−0.0781	(0.057)	−0.0739	(0.053)
Consumption rice (kg) last 7d	0.2613	(0.756)	−0.3002	(0.925)	−0.0985	(0.687)
Consumption maize (kg) last 7d	−5.3867	(4.691)	−6.6655	(4.326)	−4.4789	(3.688)
Price rice husked (TSh)	−59.193	(132.048)	−228.2154**	(96.846)	−32.092	(97.400)
Price maize grain (TSh)	126.2699*	(65.092)	97.6178**	(44.933)	65.412	(50.059)
Supply labor			−0.0164	(0.139)	0.0517	(0.123)
Hire labor lrs			−0.1011	(0.068)	−0.0954	(0.119)
N. days agri. labor by hhm			−48.9717**	(20.801)	−45.0515	(28.355)
Sales maize (kg)			−24.615	(35.460)	1.1818	(26.789)
Sales rice (kg)			−26.7566**	(12.311)	−11.8992*	(6.383)
Harvest maize (kg)			27.9655	(65.139)	100.9359	(120.452)
Harvest rice (kg)			39.9666	(61.988)	19.0239	(36.444)
Harvest area maize (acres)			−0.3582	(0.322)	−0.3318	(0.295)
Harvest area rice (acres)			−0.0704***	(0.026)	−0.0351***	(0.013)
N. hhm have current act.	0.0745	(0.269)	−0.3544**	(0.172)	−0.1908	(0.176)
N. hhm earned wage	0.129	(0.114)	−0.0467	(0.072)	−0.0202	(0.051)
N. hhm self employed	0.9167****	(0.194)	0.1878	(0.152)	0.0363	(0.125)
N. hhm work on farm	−0.9269**	(0.403)	−0.4475	(0.273)	−0.0506	(0.289)
N. hhm work unpaid	0.1061	(0.069)	0.0705**	(0.031)	−0.0135	(0.047)

Note: The reported treatment effects are the result of a household fixed effects strategy with propensity score matching. Treatment is defined based on roads the distance to a major city. Robust standard errors clustered at the village level in parenthesis.
* p< 0.10, ** p<0.05, *** p<0.01, **** p<0.001.

Figure A10: OpenStreetMap Road Network

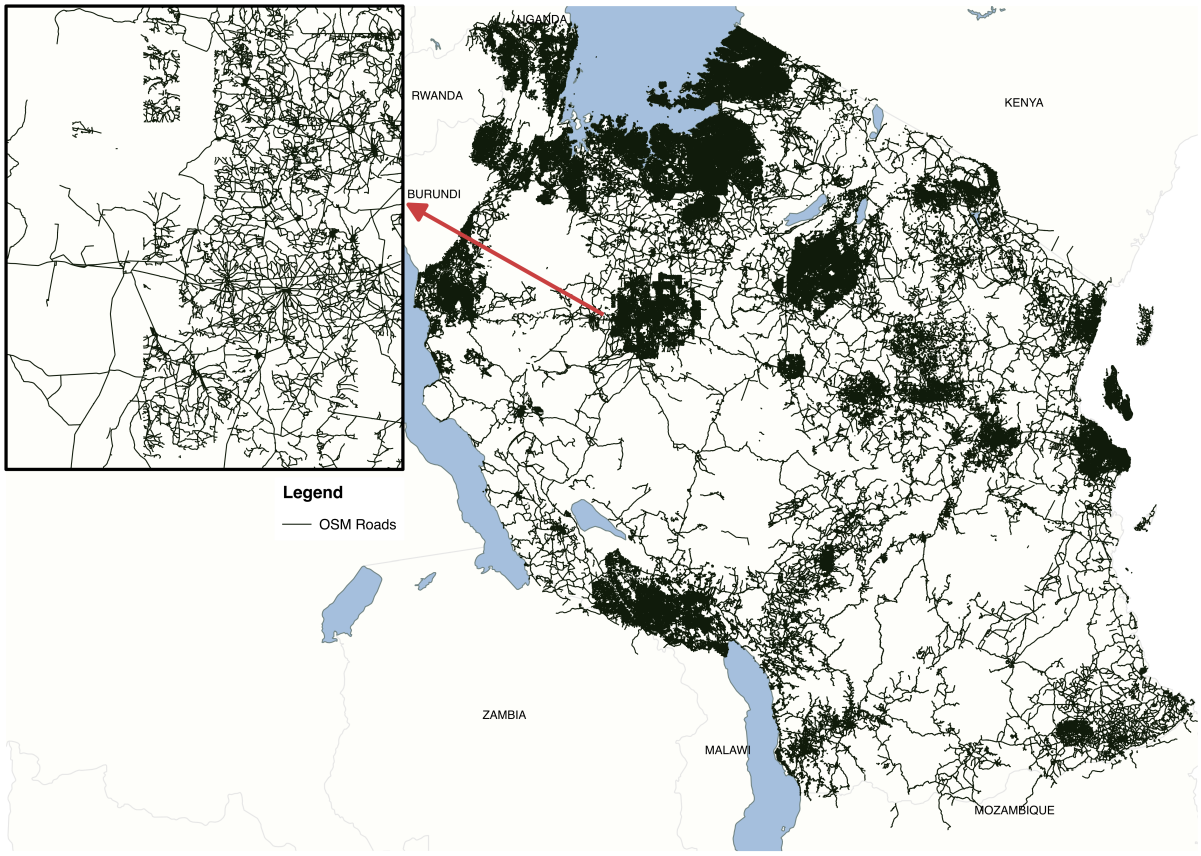


Table A18: Effect using a treatment defined based on travel time to the road

Variable	less than 60min		less than 90min		less than 120min	
Subjective Welfare: Life	−0.0561	(0.266)	−0.2055	(0.202)	0.118	(0.232)
Durable Assets (score)	−0.2143**	(0.106)	−0.1893*	(0.099)	−0.1907	(0.150)
Productive Assets (score)	−0.1177	(0.074)	−0.0548	(0.053)	−0.0725	(0.055)
Consumption rice (kg) last 7d	0.1426	(0.454)	0.4494	(0.470)	−0.1172	(0.449)
Consumption maize (kg) last 7d	−3.1343	(2.776)	−2.2552	(2.230)	−2.4059	(2.131)
Price rice husked (TSh)	−152.8886*	(78.588)	−27.9923	(76.390)	−51.5281	(89.001)
Price maize grain (TSh)	−101.5277	(63.257)	−51.9976	(50.561)	−30.8937	(55.437)
Supply labor	−0.0468	(0.108)	0.1258*	(0.072)	0.103	(0.072)
Hire labor lrs	−0.0242	(0.069)	0.0954	(0.081)	0.1203*	(0.069)
N. days agri. labor by hhm	−31.8951	(21.858)	−19.8579	(19.662)	−20.0056	(17.278)
Sales maize (kg)	−38.3374	(49.514)	−40.1783	(43.302)	−41.0949	(39.128)
Sales rice (kg)	−21.7010**	(8.386)	−46.2372*	(24.556)	−37.1845	(22.630)
Harvest maize (kg)	−6.1147	(98.851)	166.0026	(114.099)	169.5168	(105.859)
Harvest rice (kg)	−33.0462	(41.610)	−96.1480**	(45.332)	−73.8773	(48.826)
Harvest area maize (acres)	−0.4306	(0.366)	0.0257	(0.249)	0.0144	(0.206)
Harvest area rice (acres)	−0.1177**	(0.051)	−0.3075**	(0.123)	−0.2755**	(0.139)
N. hhm have current act.	−0.3335**	(0.166)	−0.1379	(0.178)	0.1551	(0.159)
N. hhm earned wage	0.0703	(0.044)	0.0496	(0.033)	0.0456	(0.032)
N. hhm self employed	0.0594	(0.090)	0.0466	(0.081)	0.0174	(0.058)
N. hhm work on farm	−0.3537*	(0.193)	−0.1311	(0.179)	−0.0605	(0.174)
N. hhm work unpaid	−0.0513*	(0.030)	0.0023	(0.029)	0.0021	(0.029)

Note: The reported treatment effects are the result of a household fixed effects strategy with propensity score matching. Treatment is defined based on the travel time distance to a treated road. Robust standard errors clustered at the village level in parenthesis. * p< 0.10, ** p<0.05, *** p<0.01, **** p<0.001.

Working Paper 518 September 2020

Authors
Christelle DUMAS
University of Fribourg, Bd de Perolles 90, 1700 Fribourg, Switzerland. Email: christelle.dumas@unifr.ch

Ximena JATIVA
University of Fribourg, Bd de Pérolles 90, 1700 Fribourg, Switzerland. Email: ximena.jativa@unifr.ch

Abstract
Spatial isolation is considered as one of the main determinants of poverty. Therefore, many transport investments are undertaken with a stated objective of poverty reduction. In our paper, we evaluate a Tanzanian program that rehabilitated 2500km of major roads between 2008 and 2013. We deal with endogenous placement issues with a household fixed-effect strategy combined with a propensity score matching. Contrary to most studies, we find damaging effects of the road on the rural population: the price of the main product (rice) decreases, they reduce rice production and reallocate labor away from farm but opportunities of o_-farm work are scarce. This results in depressed wages and households declare a lower satisfaction. This is consistent with a situation where rural households face an increased competition due to lower transportation costs.

Citation proposal
Christelle Dumas, Ximena Játiva. 2020. «Better roads, better off? Evidence on improving roads in Tanzania». Working Papers SES 518, Faculty of Economics and Social Sciences, University of Fribourg (Switzerland)

Jel Classification
O12; O13; J43; O15; O18

Keywords
Roads; Poverty; Rural households; Africa.

Working Papers SES collection

Last published
511 Ducret R., Isakov D.: The Korea discount and chaebols; 2019
512 Wu J., Morschett D.: M&A by Chinese POEs in Developed Countries. Acquiring and Bundling Strategic Assets; 2020
513 Wallimann H., Imhof D., Huber M.: A Machine Learning Approach for Flagging Incomplete Bid-rigging Cartels; 2020
514 Huber M., Lafférs L.: Bounds on direct and indirect effects under treatment/mediator endogeneity and outcome attrition; 2020
515 Farbmacher H., Huber M., Langen H., Spindler M.: Causal mediation analysis with double machine learning; 2020
516 Fernandes A., Huber M., Vaccaro G.: Gender Differences in Wage Expectations; 2020
517 Deschênes S., Dumas C., Lambert S.: Household Resources and Individual Strategies; 2020

Catalogue and download links
http://www.unifr.ch/ses/wp
http://doc.rero.ch/collection/WORKING_PAPERS_SES

Publisher
Université de Fribourg, Suisse, Faculté des sciences économiques et sociales et du management
Universität Freiburg, Schweiz, Wirtschafts- und sozialwissenschaftliche Fakultät
University of Fribourg, Switzerland, Faculty of Management, Economics and Social Sciences

Bd de Pérolles 90
CH-1700 Fribourg
Tél.: +41 (0) 26 300 82 00
decanat-ses@unifr.ch
www.unifr.ch/ses