THE GRADUATE INSTITUTE | GENEVA

INSTITUT DE HAUTES ETUDES
INTERNATIONALES ET DU DÉVELOPPEMENT

GRADUATE INSTITUTE OF INTERNATIONAL
AND DEVELOPMENT STUDIES

Graduate Institute of
International and Development Studies Working Paper
No: 12/2012

The Linkage between Outcome Differences in Cotton Production and Rural Roads Improvements: A Matching Approach

Christian K.M. Kingombe
Overseas Development Institute / Graduate Institute of International Studies

Abstract

This paper tests the linkage between a binary treatment (rural road improvement project) and a continuous outcome (cotton productivity) in Zambia's agro-based Eastern Province as measured by repeated cross-sections of farm-level data from the Zambian post-harvest survey (PHS). We use this PHS dataset, which covers the period from 1996/1997 to 2001/2002 across two phases, the pre-treatment phase (1996/1998) and the treatment phase when the Eastern Province Feeder Road Project (EPFRP) was being implemented (1998/2002). The identification strategy relies on the implementing of matching estimators for all three treatment parameters: Average Treatment Effect (ATE); Treatment on the Treated (TT) and Treatment on the Untreated (TUT), which is crucial in terms of policy relevance (Arcand, 2012). Matching ensures a sub-set of non-project areas that best represents the counterfactual and is done at the same geographic level of aggregation (van de Walle, 2009). Since treatment participation is not by random assignment we use the propensity score as a method to reduce the bias in the estimation of these treatment effects with observational PHS data sets in order to reduce the dimensionality of the matching problem. We find the ATT estimation results are not the same when implementing various matching using 'the logarithm of (cotton) yield' compared to using 'cotton productivity' as variable. In the latter case the following matching methods all have negative difference between treated and controls: 1-to-1 propensity score matching; k-nearest neighbours matching; radius matching; and 'spline-smoothing'. However, the Kernel matching has positive difference between treated and controls for the 'productivity' variable: Finally, some of the local linear regression and the Mahalanobis matching specifications yields positive difference between treated and controls for the 'logyield' variable, but not for the 'productivity' variable and not for all specifications either. Through our robustness checks of the Matching Assumption and Sensitivity of Estimates we find that the matching doesn't reduce the starting unbalancing. The comparison of the simulated ATT and the baseline ATT tells us that the latter is robust. We conclude that the application of various non-parametric matching methods didn't enable us to identify a robust linkage, most likely due to the PHS data source and the evaluation design. Future rigorous rural roads impact evaluation requires panel (with pre-intervention) data for project and appropriate non-project areas, which allows for an evaluation design that combines a double difference (DID) with controls for initial conditions either through propensity score matching, regression controls or an IV (van de Walle, 2009). Regression discontinuity designs would offer an alternative method for impact evaluation (ADB, 2011; see Arcand, 2012).

The Linkage between Outcome Differences in Cotton Production and Rural Roads Improvements in Zambia's Eastern Province: A Matching Approach

Christian K.M. Kingombe[1]

Abstract

This paper tests the linkage between a binary treatment (rural road improvement project) and a continuous outcome (cotton productivity) in Zambia's agro-based Eastern Province as measured by repeated cross-sections of farm-level data from the Zambian post-harvest survey (PHS). We use this PHS dataset, which covers the period from 1996/1997 to 2001/2002 across two phases, the pre-treatment phase (1996/1998) and the treatment phase when the Eastern Province Feeder Road Project (EPFRP) was being implemented (1998/2002).

The identification strategy relies on the implementing of matching estimators for all three treatment parameters: Average Treatment Effect (ATE); Treatment on the Treated (TT) and Treatment on the Untreated (TUT), which is crucial in terms of policy relevance (Arcand, 2012). Matching ensures a sub-set of non-project areas that best represents the counterfactual and is done at the same geographic level of aggregation (van de Walle, 2009). Since treatment participation is not by random assignment we use the propensity score as a method to reduce the bias in the estimation of these treatment effects with observational PHS data sets in order to reduce the dimensionality of the matching problem.

We find the ATT estimation results are not the same when implementing various matching using 'the logarithm of (cotton) yield' compared to using 'cotton productivity' as variable.

In the latter case the following matching methods all have negative difference between treated and controls: 1-to-1 propensity score matching; k-nearest neighbours matching; radius matching; and 'spline-smoothing'. However, the Kernel matching has positive difference between treated and controls for the 'productivity' variable: Finally, some of the local linear regression and the Mahalanobis matching specifications yields positive difference between treated and controls for the 'logyield' variable, but not for the 'productivity' variable and not for all specifications either.

Through our robustness checks of the Matching Assumption and Sensitivity of Estimates we find that the matching doesn't reduce the starting unbalancing. The comparison of the simulated ATT and the baseline ATT tells us that the latter is robust. We conclude that the application of various non-parametric matching methods didn't enable us to identify a robust linkage, most likely due to the PHS data source and the evaluation design.

Future rigorous rural roads impact evaluation requires panel (with pre-intervention) data for project and appropriate non-project areas, which allows for an evaluation design that combines a double difference (DID) with controls for initial conditions either through propensity score matching, regression controls or an IV (van de Walle, 2009). Regression discontinuity designs would offer an alternative method for impact evaluation (ADB, 2011; see Arcand 2012).

Key words: Average Treatment Effects; Average Treatment on the Treated; Matching Methods; Poor rural area development project; Impact evaluation of cotton productivity; Africa; Zambia (Eastern Province).

JEL-codes: C2; C83; D2; O12; O13; Q12; R3.

[1] **Acknowledgement:** We would like to thank Jones Govereh and Ballard Zulu from the Zambia Food Security Research Project (FSRP) for providing us with cleaned versions of the Post-Harvest Surveys 1997-2002 collected by Zambia's CSO. The paper benefited from discussions with Colin Thirtle, Bhavani Shankar, Peter Hazell, Jonathan Kydd and Salvatore di Falco. All errors are our own.
Correspondence: Senior Research Officer, Trade, Investment and Growth Programme, Overseas Development Institute, 111 Westminster Bridge Road, London SE1 7JD and Visiting Research Fellow, Development Studies, Graduate Institute of International and Development Studies (IHEID), (c.kingombe@odi.org.uk or Christian.kingombe@graduateinstitute.ch).

1. Introduction

This paper aims to establish *the structural relationships* between rural transport infrastructure (RTI) development and rural growth in the short-to medium long-term. The potential linkages between rural roads improvements and transformative rural development as measured by agricultural productivity are tested by using a matching approach.

Concerning the importance of infrastructure as an instrument of economic development and, potentially, poverty reduction, the Commission for Africa in particular argues that investing in assets such as rural roads, and a transport network, in addition to health and education, can lead to growth and job creation, helping Africa make progress towards the Millennium Development Goals (MDGs).

Despite the fact that rural infrastructure has become a major development priority (World Bank, 1994, Commission for Africa, 2005; Foster and Briceño-Garmendia, 2010; G20, 2010), yet little is known about the size and especially the distribution of benefits from such investments in Least Developed Countries (LDCs). *Roads* are particularly important forms of rural infrastructure, providing cheap access to both markets for agricultural output and for modern inputs. Given limited policy instruments for reaching the remote rural poor, *road building* at first glance seems desirable on distributional grounds (Jacoby, 2002).

In fact, road investment constitutes a major portfolio of public investment in rural areas, reinforcing the notion that rural income and productivity growth depend critically on roads and other public investments (Khandker et al., 2006, Van de Walle, 2002, Howe, 2001). Given Zambia‘s developmental challenges, especially the high poverty levels, there is a real need for stepping up efforts aimed at *strengthening and broadening the growth process*. Therefore *two critical areas* where public spending (on development, e.g. in the form of Aid for Trade) should be focused if growth is to be accelerated and broadened are:

(a) Strengthening the relevant *economic and social infrastructure*; and
(b) Enhancing *agriculture and rural development*.

Although reforms have led to promising signs of agricultural growth in recent years in Zambia, the persistence of poverty suggests that there remain significant constraints to poor Zambian households’ participation in this growth and wealth creation process. One of the key constraints is market access created by poor rural infrastructure such that around 40 percent of agricultural households are still engaged solely in subsistence agriculture (Thurlow and Wobst, 2005).

We use data from the Zambian Post-Harvey Surveys (PHS) covering all the districts of Eastern Province in the period from 1996/1997 to 2001/2002 (CSO, 2002, 2000a, 1997), allowing us to measure the short-term and medium-term gains from an United National Capital Development Fund (UNCDF) and United Nations Development Programme (UNDP) funded, ILO-executed, feeder road project covering five districts in Eastern province (Chadiza; Chipata; Lundazi; Katete; and Petauke districts) (see **Map**

A1-A2), that is the Eastern Province Feeder Road Project (EPFRP), which was implemented during this period (see **Tables A3-A4**).

The objective of this paper is to quantify the direct and indirect rural transport infrastructure investment impacts of the EPFRP. Although, the estimation of supply responses has proved difficult in the preceding literature, we will nevertheless explore the impacts on the production of the main cash crop in Zambia's Eastern Province. The aim is to estimate whether *the differential cotton yield* generated by increased market agricultural activities mainly is due to the EPFRP treatment.[2]

In other words the paper addresses a hypothesis test proposed in the following statistical terms: *The mean response in cotton productivity growth to labour-based investment in rural roads within the treatment areas is the same as the mean response in the control areas.*

The following section presents the background and setting. Section 3 presents the framework. Section 4 describes our PHS data, while section 5 presents our empirical results. Finally, section 6 summarises our conclusions.

2. Background and Setting

Growth in agricultural production in Sub-Sahara Africa (SSA) over the last 30 years has been disappointing. Rates of productivity growth have been slower than in other regions. In SSA very low rates of growth in the 1970s were followed by increases in the 1980s and 1990s, but *per capita growth* has been very low or negative over much of the period: SSA is the only region with agriculture growing at a rate below overall population growth from 1965 to 1998, and at a lower rate than growth in the agricultural labour force from 1980 to 1998 (Kydd et al., 2004).

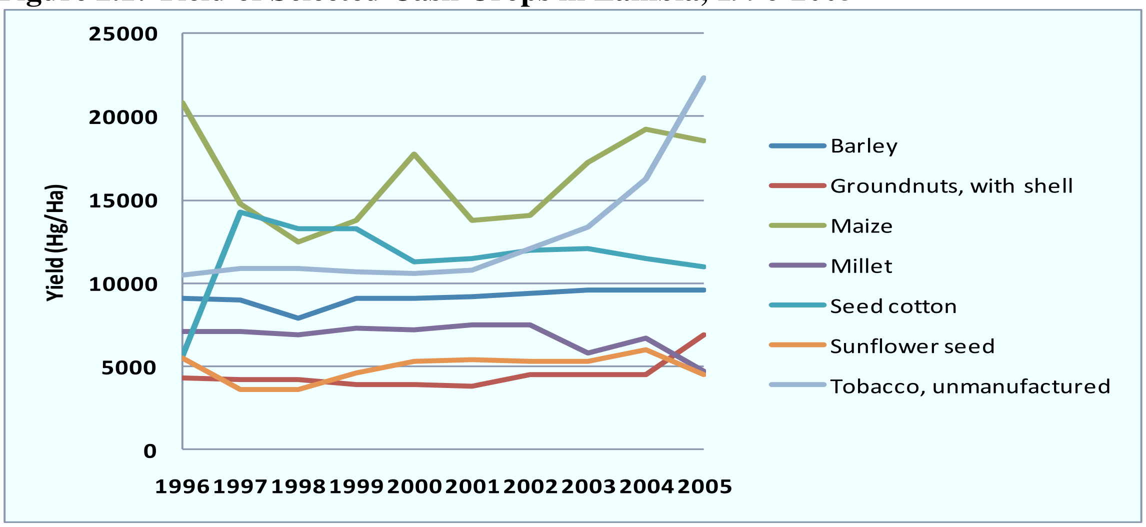
Overall 47 percent of Zambia's land area is defined as agricultural land. However, from 1995 to 2002 of Zambia's 5,260,000 hectares of arable land the percentage share under cereal production fluctuated between 10 and 15 percent. The irrigated land as a percentage of cropland only increased slightly from 1.33% in 1994/1995 to 2.95% in 2002/2003 despite the plentiful water supply from rivers and lakes.[3] Moreover, although Zambia's agricultural value added percentage share of GDP incrementally grew from 1969 to 1977, the positive trend was reversed from 1978 until 1989, where it in 1989 for the first time since independence exceeded 20%. However, it wasn't until the period 1998 to 2008 that agriculture's value added share consistently exceeded 20% of GDP (WDI, 2010) due to the fact that agriculture has been one of the faster growing sectors of the Zambian economy (FAO, 2009).

[2] Only a total of 34,329 worker days were generated in Mambwe by Rehabilitation works which is less than 20% of the average workers days of the catchment districts. Moreover no workers days were created by Maintenance Road Works, therefore Mambwe is categorised as a control district.

[3] One estimate shows that Zambia's water potential could enable it to irrigate up to 500,000 hectares of land. Currently only 13 percent of this potential is utilized, mainly by medium- and large-scale farmers. However, the small scale farmers remain the key players of the local Eastern economy (Lungu, 2006).

This positive trend in agriculture is confirmed by national crop production (tonnes) data, which shows a slight upward trend between 1996 and 2003 for e.g.: Barley; cassava; groundnuts; seed cotton; and tobacco, whereas maize; millet and sunflower seed had decreased. **Figure 2.1** shows **the changing levels of yield (Hg/Ha)** for the main food and cash crops in Zambia. *Maize yield* fell dramatically both in absolute terms and relative to other crops in the latter part of the 1990s after which the maize yield incrementally converged towards its earlier level. The fluctuations were driven both by shifting area size devoted to harvesting maize as well as production levels. The *yield of seed cotton* almost experienced a reversed trend, in the sense that the yield increased significantly towards 1997/98 after which it gradually declined until 2005, although without entirely reaching the low level at the outset. As seen from **figure 2.1** there was a wealth of diverging growth experiences amongst the other non-maize crops, some of which such as *groundnuts and tobacco* have performed well over the decade, whereas the yield of *millet and sunflower seed* declined. However, despite its declining importance the more-drought susceptible crops *maize* has remained one of the dominant staple crops in Zambia together with *cassava*.

Figure 2.1: Yield of Selected Cash Crops in Zambia, 1996-2005



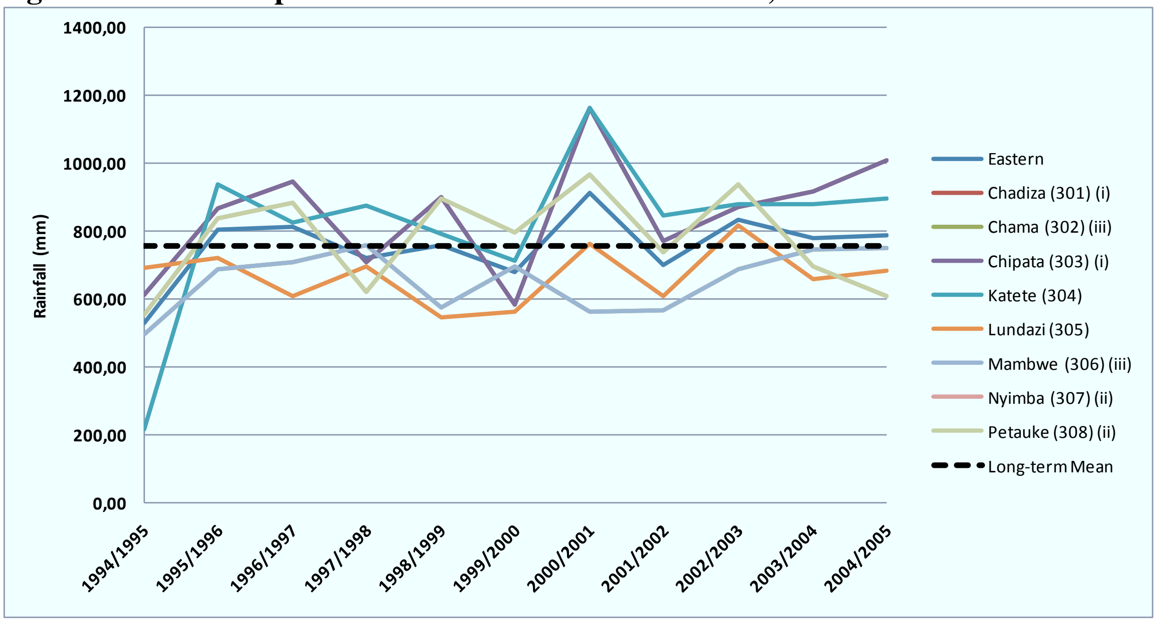
Source: Author's calculation based on Food and Agriculture Organization (FAOSTAT, 2009).
Note: This figure does not show floricultural production, which has been one of the fastest growing cash crops in recent years (World Bank, 2004).

Zambia's **Eastern Province** covers an area of 69,106 square kilometres and has 8 districts namely Chadiza, Chama, Chipata, Katete, Lundazi, Nyimba, Petauke and Mambwe. In 2000 Eastern Province had a population of 1,300,973 (**Table A1**). Of this population, 49.4 per cent were male and 50.6 per cent were female. Eastern Province was growing at an average annual population growth rate of 2.6 per cent (CSO, 2001). Eastern Province economy is agro-based and depends entirely on the soil with maize, cotton and tobacco being the major cash crops most of which are intended for the export market.[4]

[4] The Zambia-Malawi-Mozambique Growth Triangle (ZMM-GT) project incorporates fruit and vegetable growers, paprika growers and various agro-forestry programs (Patel, 2006; see Kingombe, 2012b).

However, there has been considerable volatility in agricultural growth driven by high variations in rainfall (see **figure 2.2;** and **table A2**) and the low share of irrigated land. Crop production was negatively affected by the severe 1992 and 1995 draught. Both short-term fluctuations in rainfall as well as the long-term effects of climate change have made rural farm households vulnerable to successive periods of famine (Kingombe, 2012a).

Figure 2.2: Rainfall pattern in Eastern Province Zambia, 1994-2005



Notes: 1994 and 2002 were modest drought years in Zambia.
Source: Authors' based on Zambia Meteorological Service data.

Apart from changes in the level of crop production, there have also been substantial changes in its **composition**. Much of this has been driven by the agricultural policies that were implemented by the MMD government (Smale & Jayne, 2002; Pletcher, 2000).

3. Framework

The contribution to the literature of this paper is the attempt to identify the impact of a rural transport infrastructure programme on local economic development using a matching approach. This linkage can be expressed using concepts such as economic expansion measures, e.g. district output or value added. Or other economic development measures such as cash crop productivity.[5] In other words, transport infrastructure improvements which influence travel behaviour and transport markets must eventually be transferred into these measurable economic benefits, which also include improved factor productivity, increased demand for inputs, and greater demand for consumer goods. Banister & Berechman (2000) argue that the degree to which infrastructure improvements will affect economic development is not independent of the level and performance of the in-place capital infrastructure.

Moreover, the impact of a transportation project on a regional economy varies depending on the phase of the project, because the interrelationships are not instantaneous and, in general, require considerable periods of time to transpire. Transportation spending for maintenance and rehabilitation of feeder roads affects current economic activity but also represents an investment in future growth. The main reasons for this are the long period necessary for investment implementation (1998-2001) as well as the time needed for the demand side adjustment (**table A3**).[6] The longer-term effect fosters economic growth that contributes to the expansion of a regional economy.[7]

Underlying these *time lags* are market imperfections including incomplete information concerning infrastructure development, uncertainty regarding the behaviour of public authorities and private entities, high transaction costs emanating from imperfect land market and general market externalities (see e.g. Dorward et al., 1998; Kydd et al., 2003). All of these make the transformation of transport improvements into economic benefits highly *time dependent*. The overall result is *a dynamic process* whose evolution depends on the initial conditions of local transport and activity systems and on the local transport and economic policies (Banister & Berechman, 2000).[8]

On this background we want to *evaluate the causal effect* of the binary treatment (EPFRP) on a continuous 'logarithm of cotton productivity' outcome Y experienced by units in the population of interest. For our unit of observation individual i, i = 1,. . ., N, with all units exchangeable,[9] let (Y_{0i}, Y_{1i}) denote the two potential outcomes, i.e.:

---

[5] In our context **agricultural productivity** is defined as output per hectare (Kg/Ha).

[6] As the effects of a transport project reverberate through the economy, increasing income levels, consumer spending, etc., government coffers will increase, allowing for an expansion and / or improvement of public services.

[7] Cost related indirect economic benefits of transportation investment do not materialize *instantaneously* because they involve *long-term* business and household location decisions. In fact, a prevalent view is that economic effects are realized after lags between 4 and 7 years in the case of highway developments.

[8] There is an alleged complementarity between transport and telecommunication technologies. The ability to use telecommunications (e.g. Agricultural Extension Services through radio programmes or providing agriculture market price information more recently through short message service (SMS)) may affect travel needs of the agricultural extension service officers.

[9] *The unit of analysis* for measuring benefits is at a level below the project area or PSU. We look at outcomes for rural households / farms within the project area, recognizing that certain units may benefit more than others (van de Walle, 2009).

Y_{1i} the outcome of unit i if i were exposed to the treatment: $D_i = 1$.
Y_{0i} the outcome of unit i if i were not exposed to the treatment: $D_i = 0$, where
$D_i \in \{0, 1\}$ indicator of the treatment by some social programme (e.g. Aid for Trade intervention) actually received at the level of individual (i).
$Y_i = Y_{01} + D_i(Y_{1i} - Y_{0i})$ the actually observed outcome of unit i.
X the set of pre-treatment characteristics.
$\tau_i = Y_{1i} - Y_{0i}$ the causal (treatment) effect for a single unit i.

The fundamental problem of causal inference' is that it is impossible to observe the individual treatment effect. It is impossible to make causal inference without making generally untestable assumptions (Sianesi, 2001; Abadie et al., 2001; Dehejia and Wahba, 2002).

Under some assumptions in **tables 5.2a-b** we estimate the causal estimand of interest, namely '*the average treatment effect (ATE)'* of the sampled Eastern Province agricultural household population (**Table A5a**):[10]

- Average Treatment Effect (ATE) = $\frac{1}{N}\sum_{i-1}^{N}(Y_{1i} - Y_{0i})$ = $E(y_1 - y_0)$;[11]
- Average Treatment Effect on the Untreated (ATU) = $E(y_1 - y_0 \mid D = 0)$;
- Average Treatment Effect for the sub-population of the Treated (ATT) = $E(y_1 - y_0 \mid D = 1)$.[12]

The primary treatment effect of interest in our non-experimental settings is the expected treatment effect for the treated population (ATT); hence:

(3.1a) $\tau|_{D=1} = E(Y_{1i} - Y_{0i} \mid D = 1) = E(Y_{1i} \mid D = 1) - E(Y_{0i} \mid D = 1) = \frac{1}{N_1}\sum_{i \mid D_i = 1}(Y_{1i} - Y_{0i})$,

(3.1b) $\tau|_{D=0} = E(Y_{1i} - Y_{0i} \mid D = 0) = E(Y_{1i} \mid D = 0) - E(Y_{0i} \mid D = 0) = \frac{1}{N_0}\sum_{i \mid D_i = 0}(Y_{1i} - Y_{0i})$, where

$N_1 = \sum_i D_i$ and $N_0 = \sum_i (1 - D_i)$ are the number of treated and control units respectively (Sianesi, 2001; Abadie et al., 2001).

Table 3.1 shows that the basic issue is a problem of missing information. *The problem of unobservability* is summarized by the fact that we can estimate $E(Y_{1i} \mid D = 1)$ but not $E(Y_{0i} \mid D = 1)$.

Table 3.1: The Problem of Missing Data

	Y_1	Y_0
D = 1	$Y_1 \mid D = 1$: Observed	$Y_0 \mid D = 1$: Unobserved
D = 0	$Y_1 \mid D = 0$: Unobserved	$Y_0 \mid D = 0$: Observed

Source: Arcand, 2012.

[10] Whether one is interested in the average treatment effect in the population (PATE) or the sample (SATE) does not affect the choice of estimator: the sample matching estimator will estimate both. However, in general the variance for SATE is smaller than for the PATE (Abadie et al., 2001; cf. Imbens, 2002, 2003).
[11] Heckman (1997) notes that ATE might not be of relevance to policy makers because it includes the effect on persons for whom the programme was never intended (Grilli and Rampichini, 2011).
[12] The parameter of interest in most evaluation studies (ibid.).

Thus, we need to construct the unobserved *the counterfactual mean* $E(Y_{0i}| D = 1)$, the outcome participants would have experienced, on average, had they not participated, by choosing a proper substitute for it to estimate ATT. The difference,

$$(3.2) \qquad \tau^e = E(Y_{1i} | D = 1) - E(Y_{0i} | D = 1),$$

can be estimated, but it is potentially a biased estimator of the difference in the outcomes with and without treatment, τ. Intuitively, if Y_{0i} for the treated and comparison units systematically differ, then in observing only Y_{i0} for the comparison group we do not correctly estimate Y_{i0} for the treated group. Such *bias* is of paramount concern in non-experimental studies.[13] The role of *randomization* is to prevent this (Deheija and Wahba, 2002).[14]

Thus, in our observational study of the EPFRP's impact on cotton productivity (*logyield*), by definition there are no experimental controls. Therefore, there is no direct counterpart of the **ATE**. In other words, the *counterfactual is not identified*. As a substitute we may obtain data from a set of *potential comparison units* that are not necessarily drawn from the same population as the treated units, but for whom the observable characteristics, **x**, match those of the treated units up to some selected degree of closeness (see **tables A5a-h**).

Van de Walle(2009) argues that *road projects* typically select road links or segments, not geographic areas. However, these road segments are not independent of their project areas; by selecting a road segment one automatically selects a project area (see **table A3**). Selection of a road segment can thus be treated as the (implicit) selection of a project area. **The comparison units** must then be selected from within *the sub-set of the non-project areas* that appear to best represent the counterfactual of what would have happened in the project areas in the absence of the project. *Matching* to ensure a sub-set of non-project areas that best represents the counterfactual should be done at the same geographic level of aggregation (e.g. local government area or community level) used in defining the PSUs according to van de Walle(2009).

The average outcome for the *untreated matched group* identifies the *mean counterfactual outcome* for the treated group in the absence of the treatment. This approach solves *the evaluation problem* by assuming that selection is unrelated to the untreated outcome, conditional on **x** (Cameron & Triverdi, 2005).

Propensity Score Matching

In this kind of evaluation problems, data often do not come from randomized trials but from (non-randomized) observational studies. Hence, Rosenbaum and Rubin (1983, 1985) suggest the use of ***the propensity score*** — the probability of receiving treatment conditional on covariates (**x**)[15] — as a

[13] There are three sources of bias in any piece of empirical work: (1) *"Garden variety" endogeneity* in which, for example, common unobservables determine both treatment status and outcomes; (2) the decision to implement or participate in the intervention (D) is based in part on what people expect to gain from it (b); (3) The impact of the intervention (β) is correlated with unobservables that determine the outcome (ε). Most methods deal with the first source of bias, because it is much harder to deal with the other two (Arcand, 2012).

[14] In *a non-experimental setting*, the treatment and comparison samples are either drawn from distinct groups or are nonrandom samples from a common population. In contrast, in *a randomized experiment*, the treatment and control samples are randomly drawn from the same population, and thus the treatment effect for the treated group is identical to the treatment effect for the untreated group (Dehejia and Wahba, 2002).

[15] The propensity score is a possible balancing score b(X), i.e. functions of the relevant observed covariates X such that the conditional distribution of X given b(X) is independent of assignment into treatment (Grilli and Rampichini, 2011).

method to reduce the bias in the estimation of treatment effects with observational data sets in order to reduce the dimensionality of the matching problem, by allowing us to condition on a scalar variable rather than in a general n-space (Dehejia and Wehba, 2002; Grilli and Rampichini, 2011).[16]

Thus, when treatment participation is not by random assignment but depends stochastically on a vector of observable variables **x**, as in our *observational PHS data*, then the concept of **propensity scores** is useful.[17] This is a conditional probability measure of treatment participation given **x** and is denoted **p(x)** (i.e. the probability of unit i having been assigned to treatment), where

(3.3a) $p(x) = \Pr\{D = 1 \mid X = x\} = E(D_i \mid X_i)$,

The individual assignment possibilities (i.e. propensity scores) as a function of unit i's value of covariates, p_i, are strictly between zero and one,

(3.3b) $0 < p_i < 1$

An exogeneity assumption that plays an important role in treatment evaluation is the *balancing condition* of the estimated propensity score (PS), which states that

(3.4) $D \perp x \mid p(x)$.

We can investigate whether, approximately, Eq.(3.4), by stratifying the sample into subsamples (blocks) with similar value of p(x), and then testing independence of D_i and x_i within each resulting stratum. For each covariate, we test whether the means for the treated and for the controls are statistically different in all blocks. If one covariate is not balanced in one block, we split the block and test again within each finer block. If one covariate is not balanced in all blocks, modify the specification of the propensity score adding more interaction and higher order terms and then test again (Grilli and Rampichini, 2011).

Type of Matching Estimators

Matching on the propensity score is essentially a weighting scheme, which determines what *weights* are placed on comparison units when computing the estimated treatment effect:

(3.5) $$\hat{\tau}\,|_{D=1} = \frac{1}{|N|}\sum_{i\in N}\left(Y_i - \frac{1}{|J_i|}\sum_{j\in J_i} Y_j\right),$$

where **N** is the treatment group, **|N|** the number of units in the treatment group, $\mathbf{J_i}$ is the set of comparison units matched to treatment unit **i** (see Heckman et al., 1998), and $\mathbf{|J_i|}$ is the number of comparison units in $\mathbf{J_{1i}}$. Expectations are replaced by sample means, and we condition on $\mathbf{p(X_i)}$ by *matching* each treatment unit **i** to a set of comparison units, $\mathbf{J_i}$, with a similar propensity score. Our **matching strategy objective** is to *match treated units to comparison units whose propensity scores are sufficiently close* to consider the conditioning on *p*(*Xi*) in the following proposition:

(3.6) $\tau|_{D=1} = E_{p(X)}[(\tau|_{D=1,\, p(X)}) | D_i = 1]$,

to be approximately valid (Dehejia and Wehba, 2002).

[16] As the number of variables increases, the number of cells increases exponentially, increasing the difficulty of finding exact matches for each of the treated units.

[17] Estimate *the propensity score* on the X's e.g. via probit or logit.

Three issues arise in **implementing matching:** (i) whether or not to match with replacement, (ii) how many comparison units to match to each treated unit, and (iii) finally which matching method to choose (ibid).

The unit level treatment effect is $\mathbf{Y_{1i} – Y_{0i}}$. However, only one of the potential outcomes $\mathbf{Y_{1i}}$ or $\mathbf{Y_{0i}}$ is observed for each individual and the other is unobserved or missing (**table 3.1**). ***The matching estimators*** we consider impute the missing potential outcome by using average outcomes for individuals with "similar" values for the covariates. Pair to each treated individual **i** some group of 'comparable' non-treated individuals and then associate to the outcome of the treated individual i, $\mathbf{y_i}$, the (weighted) outcomes of his 'neighbours' **j** in the comparison group:

(3.7) $$\hat{y}_i = \sum_{j \in C^0(p_i)} w_{ij} y_j$$

Where:

$\mathbf{C^0(p_i)}$ is the set of neighbours of treated i in the control group $\mathbf{w_{ij}} \in [0, 1]$ with $\sum_{j \in C^0(p_i)} w_{ij} = 1$

is the weight on control **j** in forming a comparison with treated **i**.

The first step in PSM is the estimation of the propensity score:[18] this affects the large sample distribution of propensity score matching estimators.[19] However, an estimate of the propensity score is not enough to estimate the ATT.

Several matching methods have been proposed in the literature. The most widely used are: Nearest-Neighbor Matching (with or without within caliper; with or without replacement)[20]; Radius Matching; Kernel Matching; Stratification Matching; and one-to-one matching is also common as well as k-Nearest neighbours; local linear regression, and Mahalanobis matching (Grilli and Rampichini, 2011).

We associate to the outcome $\mathbf{y_i}$ of treated unit i a *'matched' outcome* given by the outcome of the most observably similar control unit ('traditional matching estimators') ***one-to-one matching:***

(3.8) $$C^0(p_i) = \left\{ j : |p_i - p_j| = \min_{k \in \{D=0\}} \{ |p_i - p_k| \} \right\}, \quad w_{ik} = 1(k=j).$$

A weighted average of the outcomes of more (possibly all) non-treated units where the weight given to non-treated unit j is in proportion to the closeness of the observables of i and j ('smoothed weighted matching estimators') ***kernel-based matching:***

(3.9) $$C^0(p_i) = \{D = 0\} \qquad w_{ij} \propto K\left(\frac{p_i - p_j}{h}\right)$$ (for Gaussian kernel) (Sianesi, 2001).[21]

[18] The Stata command psmatch2 (Leuven and Sianesi 2003) will perform PSM.
[19] Abadie and Imbens (2009) derive the large sample distribution of PSM estimators and propose an adjustment to the large sample variance of propensity score matching estimators that corrects for first step estimation of the propensity score (Grilli and Rampichini, 2011).
[20] Matching *with replacement* keeps bias low at the cost of larger variance. Matching *without replacement* keeps variance low at the cost of potential bias.
[21] Non-negative; symmetric and unimodal.

Nearest Neighbour match treated and control units taking each treated unit and searching for the control unit with the closest propensity score; i.e., the Nearest Neighbour.[22] Once each treated unit is matched with a control unit, the difference between the outcome of the treated units and the outcome of the matched control units is computed. The ATT of interest is then obtained by averaging these differences.

Given a treated unit *i*, let $l_{m(i)}$ denote the index of the non-treated unit that is the m-th closest to unit i in terms of the distance measure based on the norm ||.||.

(3.10) $$\sum_{j:D_j \neq D_i} \{||p_j - p_i|| \leq ||p_l - p_i||\} = m$$

Let $C(i)_M$ denote the set of indices for the first M matches for unit i:

$C(i)_M = \{l_1(i), \ldots, l_M(i)\}$

$$\hat{Y}(0) = \frac{1}{M} \sum_{j \in C(i)_M} Y_j^{obs}$$

The formula for of the NN matching estimator is:

(3.11) $$ATT^{NN} = \frac{1}{N^T} \sum_{i:D_i} Y_i^{obs} - \frac{1}{N^T} \sum_{j \in C(i)_M} w_j Y_j^{obs}$$

N^T is the number of observations in the treated group
N_i^C is the number of controls matched with treated observation i.
w_{ij} is equal to $\frac{1}{N_i^C}$ if j is a control units of i, and zero otherwise
$w_j = \sum_i w_{ij}$ (Grilli and Rampichini, 2011).

Identification Strategy

In general, if we compare the outcomes by treatment status, we obtain a biased estimate of the ATT. The difference between treated and non-treated outcomes (even) in absence of treatment is leading to the so-called *selection bias*.[23] The ATT – $E[Y_1 - Y_0 \mid D_i = 1]$ – is identified only if:

(3.12) $$E(Y_0 \mid D = 1) - (E(Y_0 \mid D = 0) = 0,$$

i.e. if the outcomes of individuals from the treatment and comparison groups would not differ in the absence of treatment. In experiments where assignment to treatment is random this is ensured and the treatment effect is identified. In observational studies, we must rely on some *identifying assumptions* to solve the selection problem (Grilli and Rampichini, 2011).

The underlying identifying assumption is **unconfoundedness** (selection on observables or conditional independence) (see Eq.(3.4) above). If the decision to take the treatment is purely

[22] Although it is not necessary, the method is usually applied with replacement, in the sense that a control unit can be a best match for more than one treated unit (Grilli and Rampichini, 2011).
[23] Sources of *Selection Bias*: (1) non-overlapping supports of X in the treated and comparison group (i.e., the presence of units in one group that cannot find suitable comparison in the other); (2) unbalance in observed confounders between the groups of treated and control units (selection on observables); (3) unbalance in unobserved confounders between the groups of treated and control units (selection on unobservables) (Grilli and Rampichini, 2011).

random for individuals with similar values of the pre-treatment variables, then we could use the average outcome of some similar individuals who were not exposed to the treatment. For each i, matching estimators impute the missing outcome by finding other individuals in the data whose covariates are similar but who were exposed to the other treatment. In this way, differences in outcomes of this well selected and thus adequate control group and of participants can be attributed to the treatment (Grilli and Rampichini, 2011).

Thus, to ensure that *the matching estimators identify* and consistently estimate the treatment effects of interest, we assume **unconfoundedness**: assignment to treatment is independent of the outcomes, conditional on the covariates:

(3.13) $(Y_0; Y_1) \perp\!\!\!\perp D \mid X$

overlap or common support condition:[24] the probability of assignment is bounded away from zero and one:

(3.14) $0 < \Pr(D = 1 \mid X) < 1.$[25]

Given these two key assumptions of unconfoundedness and overlap one can identify the average treatment effects (ATE) (ibid.).

With the observational post-harvest survey (PHS) data set, we try to structure it so that we can conceptualize the data as having arisen from an underlying regular assignment mechanism.[26] We will use the random sample statistics from the target areas, which Zambia's Central Statistical Office (CSO) collected in *the six year* period from **1996/1997 to 2001/2002**. This pseudo-panel dataset ideally should have presented us with an opportunity to use *panel data analysis* to test which factors that determine the variation of the productivity of cash crops in general, and cotton in particular. A panel data set would thus have allowed us to account for the *idiosyncratic* household level fixed effect with it's two components, namely: The farm effect, η_{ht}, and the cotton-specific effect, φ_{ht}.[27] However, the PHS dataset is unfortunately only *a repeated cross section of farmers*.

Another method to overcome the problem of the lack of panel data is by creating **a pseudo-panel**. In this method *groups of "like" households* are created and changes in their income over time are analysed.[28] The advantages of this method is that it allows us to make statements about changes that occur to different types of *similar households over time* but it involves loss of information on

[24] We can consider only the observations whose *propensity score belongs to the intersection of the supports* of the propensity score of treated and controls (Grilli and Rampichini, 2011).

[25] The assignment mechanism can be interpreted as if, within subpopulations of units with the same value for the covariate, completely randomized experiment was carried out. We can analyze data from subsamples with the same value of the covariates, as if they came from a completely randomized experiment (ibid.).

[26] **Regular designs** are like completely randomized experiments except that the probabilities of treatment assignment are allowed to depend on covariates, and so can vary from unit to unit.

[27] The unobservables are **indexed by ht** because, given the cross-sectional nature of the data, the unit of observation is a household-time period (ht) combination. However, if the data were a panel, the unobservables would be **indexed by h** only (Brambilla and Porto, 2006).

[28] The method is adopted by **cohort studies**, particularly in labour economics, where individuals are *grouped by age* (possibly gender and other attributes) and the cohort is compared with other cohorts over time (cf. Kingombe, 2012a).

the variation within "like" groups (McCulloch et al., 2001).[29] Thus, it is possible to *"create" a pseudo-panel* at a geographical scale by aggregation from our repeated independent cross-sectional PHSs with different households. This has proven to be quite useful for estimating structural relationships (Glewwe and Jacoby, 2000) to capture the short-to medium- run effects (see Kingombe, 2012a).[30]

There are some difficult methodological issues in assessing impacts rigorously. The number of *sources of bias* is more intractable with respect to rural roads. Policy and road placement is not random. Government does not randomly assign roads, because there are reasons for where they place roads. It is highly likely that the factors that attract better roads in certain areas also affect the agricultural productivity outcomes. Unless the comparison areas – the counterfactual – have the same factors as mentioned above, it will leave biased estimates. *Selection bias* occurs if for some reason roads are poor in participating area and being compared with places that don't have these factors.

Typically, ***the double-difference (DD)*** approach is undertaken to get rid of *endogeneity* (see Kingombe and di Falco, 2012). But this is not enough in a context where a lot of the initial conditions may affect the trajectory of the local communities. There are *time-varying initial conditions* that will not be purged with a DD approach. Thus, failure to adequately control for initial conditions that lead to the road placement can lead to very large biases in estimates of impacts.

Comparing changes in outcomes with changes in roads (difference-in-difference) does not eliminate the problem if roads are placed based on initial conditions that influence subsequent growth.[31] *Endogeneity* also arises if changes in placement are a function of time-varying factors, e.g.: when road expansions accord with changing economic conditions themselves correlated with changes in outcomes (van de Walle, 2009).[32]

4. Data

The Agricultural statistical system in Zambia has been producing both structural[33] and performance data.[34] In 1985/86 the two types of surveys were renamed **the Crop Forecasting Survey (CFS)** and **Post- Harvest Survey (PHS)**, respectively.[35] These surveys are conducted in an

[29] Other possible partitions include: the strata used by the sampling frame, i.e. low, middle and high cost housing areas in urban areas, and small, medium and large farmers and non-agricultural households in rural areas; employment sector (in urban areas) and main agricultural output (rural areas); age, gender.

[30] Banister & Berechman consider *10 years* as the time it takes for land use and travel markets to converge to a state of equilibrium following an external change. Thus, medium to long terms effects are to be *over 10 years*. Bourguignon, Ferreira, and Lustig (2001) in their review of income distribution dynamics, recommend at least a ten-year interval.

[31] Should be applied only if time-invariant unobservables are a problem. However, the problem with the DID approach is that it assumes away the following biases: (1) the decision to implement or participate in the intervention (D) is based in part on what people expect to gain from it (β)); (2) The impact of the intervention (β) is correlated with unobservables that determine the outcome (ε) (Arcand, 2012).

[32] That is common unobservables determine both treatment status and outcomes (Arcand, 2012).

[33] **Structural data or basic agricultural statistics** relate to characteristics of agricultural holdings that vary slowly over time (are normally collected in a Census of Agriculture, which is carried out at intervals of 10 years).

[34] **Performance data or current agricultural statistics** relate to: prices, quantities of inputs and outputs; enterprise costs and returns; and net farm incomes are collected mainly from current (annual) agricultural surveys. CSO and MAFF have been collecting current agricultural statistics since 1964.

[35] Up to 1978/79 agricultural season, the survey was called the Agricultural and Pastoral Production Survey, later renamed in 1982/83 as the Early Warning and Agricultural Survey to encompass **the Crop Forecasting** and **Post-Harvest** stages of the agricultural season during which period the two different types of surveys were conducted.

integrated manner and as the core of the National Household Survey Capability Programme (NHSCP), which has been implemented since 1983. However, The Agriculture and Environment Department of Zambia's CSO only have agricultural production data at the district level going back until 1995. We will be using the already existing PHSs of Zambia's Eastern Province exclusively.

A stratified multi-stage sample design was used for the Zambia PHS. The sampling frame was based on the data and cartography from the 1990 Census of Population, Housing and Agriculture.

The primary sampling units (PSUs) were defined as the CSAs delineated for the census. The CSAs were stratified by district within province and ordered geographically within district. A master sample of CSAs was selected systematically with probability proportional to size (PPS) within each district at the first sampling stage; the measure of size for each PSU was based on the number of households listed in the 1990 Census.[36]

The secondary sampling unit (SSU) is the SEA, that is, the sampling areas defined as the segment covered by one enumerator during the census. One SEA was selected within each sample CSA with PPS for the survey. A new listing of households was conducted within each sample SEA, and the farm size was obtained for each farm household. The listed households within each sample SEA were then divided into two groups based on farm size: **Category A** for households with less than 5 hectares (HAs.) and **Category B** for households with 5 or more HAs (**table 4.1**).

Table 4.1: Frequency of Holdings in Eastern Province, 1996-2002

	PHS 1996/97		PHS 1997/98		PHS 1998/99		PHS 1999/2000		PHS 2000/01		PHS 2001/2002	
	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent	Frequency	Percent
A-Small scale holding	956	78	1052	88	1111	88.5	1233	85.8	1060	84.9	1128	87.3
B-Medium scale holding	256	22	144	12	144	11.5	204	14.2	189	15.1	164	12.7
Total	1225	100	1196	100	1255	100	1427	100	1249	100	1292	100

Source: Author's calculation.

It was found that most sample SEAs had less than 10 households in Category B, in which case all of these households were included in the sample with certainty at the final stage of selection. In order to ensure *a sample of 20 households within each sample SEA*, the remaining households were selected from Category A (Megill 2000).

Specifically, **the objectives of the PHS** include provision of actual figures pertaining to: Area planted to individual crops (land usage - allocation); Realised Production quantities (output in physical units); Sales of produce and income realized; Numbers of livestock and poultry; Purchase and use of agricultural inputs; Capital formation and other operational expenses; Demographic characteristics of heads of rural households (household characteristics); Farming practices and soil conservation methods used; Access to agricultural loans; and, access to market prices information

[36] The project/catchment could be a local government area or community serviced by the road, or might consist of a number of communities in its vicinity. The set of all such areas defines the sampling frame from which one selects a random sample of primary sampling units (PSU) and within these, a random sample of beneficiaries. The beneficiaries of the project can be defined as the entire project area or the communities, firms, households or individuals located within the area (van de Walle, 2009).

and agricultural extension services in general. The reference period for this information is the **agricultural season** starting 1st October ending 30th September.

However, the PHS estimates for some crops which are rare or limited to particular geographic areas have relatively high sampling errors.[37] In order to **evaluate the effectiveness of the PHS sample design** in meeting these survey objectives, it is first necessary to measure the level of precision for the survey estimates based on this design. (Megill 2000) illustrates that the main limitation of the sample design was that it didn't not provide reliable results for minor crops such as rice, sorghum, *cotton*, and tobacco. Moreover, over the period during which the PHSs have been conducted, the survey questionnaire has undergone several major revisions and differences in questions asked.

The **PHS 2001/2002** also covered the whole country representing a sample proportion of about 5%. The survey was conducted in the same CSA and SEAs selected over the previous 4-5 years. The survey relied on the previous listing of household populations in **1999/2000 PHS** but with a new sample drawn from this listing.

In each district, the allocated sample size was shared **proportionately among the crop strata**, i.e., the more SEAs a crop stratum had the larger its share of the sample. This was done whilst ensuring that a minimum of two SEAs was selected from each stratum to facilitate computation of sampling error of the estimates.

Since the selection of participants in the PHS 2001/02 survey was not done with a simple random sample, a weight variable is used for our analysis. We use the overall household weight.[38] The district level weight is simply the probability that the number of households in a SEA will be selected as a primary unit from within a CSA within a particular District. After obtaining a complete list of the households in the SEA categorized as small or medium scale and the number of households to be sampled in each SEA, **the SEA level weight** is estimated. So with the District Level and SEA level weights, these two are multiplied and the product is *the boosting factor*.

Table 4.2: Post Harvest Survey (sample sizes) by District in Eastern Province, 1997-2002

District	1996/1997	1997/1998	1998/1999	1999/2000	2000/2001	2001/2002
Chadiza (301)	96	88	89	100	88	100
Chipata (303)	303	295	304	338	307	330
Katete (304)	198	198	199	220	184	212
Lundazi (305)	224	225	229	260	233	261
Petauke (308)	267	262	271	320	262	305
Total Catchment Districts	***1088***	***1068***	***1092***	***1238***	***1074***	***1208***
Chama (302)	37	36	76	80	70	77
Mambwe (306)	52	55	34	59	51	59
Nyimba (307)	48	37	53	60	54	59
Total Control Districts	***137***	***128***	***163***	***199***	***175***	***195***
Total	**1225**	**1196**	**1255**	**1437**	**1249**	**1403**

Source: Authors' calculations based on CSO's Post Harvest Surveys 1997-2002.

The number of sample household selected was on average 1,274 households, which were interviewed in the Eastern Province, during the period December and January using personal

[37] The definition of **in-scope farm households** for the survey should also be examined. Therefore a report by Megill recommends certain modifications to the sample design for **improving the sampling efficiency for future surveys**.

[38] The **Weights (Boosting Factors)** are the inverse of the probability that a given household has of being included in the sample. These factors are developed at the SEA level for each category of farmer.

interviews with qualified respondents in sample households in sample areas (see **table 4.2**). All PHSs were independent farm surveys and thus interviewed different households in each year. Consequently it is **not possible to construct a panel** of households using PHSs surveys in order to examine the correlates and causes of changes in the agricultural productivity of individual households over time (McCulloch, Baulch et al. 2001; UNECA 2005).

5. Estimation Results and Discussion

This section we are interested in estimating possible effect of the rehabilitation of the feeder road network in Eastern Province in the period from 1996 to 2001 (i.e. the EPFRP) on the **productivity of cotton production** in Zambia's Eastern Province from 1996/1997 to 2001/2002 by using the PHS dataset.[39]

5.1. Descriptive Statistics

We are interested in evaluating the effect of a binary rural road (AfT) intervention (i.e. access to local transport infrastructure or not) on a continuous outcome 'cotton yields per hectare' (i.e. farm productivity).

- The treatment variable is 'the Rural transport infrastructure (EPFRP)', which is discrete and of on/off variety.
- The outcome variable is 'the logarithm of cotton output (in Kg) per hectare' (or alternatively 'the Volume of cotton production per hectare produced (MT/HA)') a continuous variable with a mean ranging from 6.54 in 1996/1997 to 6.83 in 1997/1998 and a standard deviation from 0.71 in 2001/2002 to 1.40 in 1999/2000.
- The observable pre-treatment covariates (household determinants; household demographics; input use; assets; agricultural extension services; geographical variables) that we use to identify similar individuals are given in the **table 5.1** below.
- *The choice of covariates* from table 5.1 to insert in the propensity score model (PSM) is based on theory and previously empirical findings. However, a variable should only be excluded from analysis if there is consensus that the variable is either unrelated to the outcome or not a proper covariate.[40]

[39] The PHS dataset is available in STATA format upon request.

[40] Only variables that influence simultaneously the treatment status and the outcome variable should be included as covariates in the propensity model (see e.g., Sianesi, 2004; Smith and Todd, 2005). The set of X must credibly satisfy the unconfoudedness condition that the outcome variable to be independent of treatment conditional on the propensity score. In other words, only variables that are unaffected by treatment should be included in the model. To ensure this variables should either be fixed over time or measured before participation (Grilli and Ramphicini, 2011).

Table 5.1 Descriptive Statistics, 1996/1997 – 2001/2002

		1996/1997		1997/1998		1998/1999		1999/2000		2000/2001		2001/2002	
		Full Sample		Full Sample		Full Sample		Full Sample		Full Sample		Full Sample	
Variable	Variable	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Dependent variable	Volume of cotton production per hectare produced (MT)	1,33	2,31	1,48	2,09	1,62	3,06	1,64	3,02	0,97	0,68	0,97	0,68
	Log of cotton output (in kg) per hectare	6,54	1,10	6,83	0,96	6,75	1,17	6,55	1,40	6,65	0,71	6,64	0,71
Household determinants	Age of the household head	46,7	15,0	44,4	15,2	45,5	15,3	43,0	14,3	45,7	14,7	45,3	14,7
	Age Square of the household head	2404,0	1506,1	2205,4	1537,9	2307,8	1535,0	2056,0	1371,5	2309,7	1465,6	2270,4	1459,2
Household demographics	Size of the household	5,8	3,2	5,7	3,0	5,94	3,20	6,17	3,43	5,97	2,95	6,34	2,93
	Log of Size of the household	1,61	0,59	1,59	0,56	1,63	0,59	1,67	0,56	1,66	0,54	1,73	0,50
	Household category (stratum)	1,22	0,41	1,12	0,33	1,11	0,32	1,14	0,35	1,14	0,35	1,13	0,33
	Number of males in household	2,79	1,82	2,74	1,85	2,94	1,99	3,08	2,24	2,98	1,93	3,18	1,86
	Number of females in household	3,03	1,97	2,91	1,79	2,99	1,85	3,09	1,92	2,98	1,73	3,16	1,81
	Sex of head of household	1,23	0,42	1,23	0,42	1,24	0,43	1,24	0,43	1,25	0,43	1,25	0,44
Input use	Basal Quantity used (kg)	29,93	123,90	30,88	121,42	39,63	145,91	47,77	129,59	32,81	149,51	34,79	149,91
	Topdressing Quantity used (kg)	27,18	104,57	30,50	122,82	38,71	127,18	45,80	118,37	31,98	145,77	33,69	147,14
	Basal Fertilizers Used per cultiv. Area (kg per ha)	11,53	36,76	13,05	38,21	16,56	42,32	22,00	53,14	17,17	50,91	16,09	41,98
	Top Dressing Fertilizers Used per cultiv. Area (kg per ha)	10,43	28,63	13,32	41,86	16,74	38,74	21,01	47,71	16,10	40,45	15,56	37,37
	Value of Basal quantity used - (ZMK)	31920,3	92680,6	22202,3	238505,6	24409,9	87229,3	34564,7	95575,4	n.a.	n.a.	n.a.	n.a.
	Value of Topdressing quantity used - (ZMK)	27770,4	80701,8	23052,7	241685,4	25208,1	89689,2	33167,1	86535,9	n.a.	n.a.	n.a.	n.a.
	Expenditure on Basal fertilizers per cultivated area (ZMK/Ha)	12152,0	26389,7	7133,8	26384,2	10491,0	28902,8	15823,0	38979,2	n.a.	n.a.	n.a.	n.a.
	Expenditure on Topdressing fertilizers per cultivated area (ZMK/Ha)	10284,9	19167,9	8317,9	42566,3	10934,1	26732,4	15274,7	35724,6	n.a.	n.a.	n.a.	n.a.
Assets	Number of ploughs	0,374	0,865	0,29	0,77	0,30	0,77	0,27	0,65	n.a.	n.a.	n.a.	n.a.
	Number of draught animals	0,649	1,741	0,54	1,45	0,57	1,55	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
	Number of ploughs per household member	0,062	0,159	0,05	0,13	0,05	0,13	0,04	0,11	n.a.	n.a.	n.a.	n.a.
	Number of draught animals per household members	0,099	0,260	0,09	0,25	0,09	0,27	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
	Size of the land allocated to cotton	0,13	0,21	0,12	0,21	0,11	0,20	0,07	0,16	0,10	0,19	0,10	0,18
	Total area under crops (ha)	1,97	1,77	1,86	1,74	1,87	1,96	2,10	2,06	1,73	1,65	1,83	1,74
	Cultivated land per household member (ha)	0,38	0,33	0,37	0,33	0,35	0,32	0,39	0,45	0,34	0,37	0,36	0,31
	Livestock raising	0,58	0,49	0,48	0,50	0,48	0,50	0,50	0,50	0,55	0,50	0,47	0,50
	Usage of animal draught power for land preparation	0,27	0,45	0,25	0,43	0,24	0,43	0,28	0,45	0,35	0,48	0,35	0,48
	Received agricultural loan	0,323	0,468	0,265	0,441	0,32	0,47	0,16	0,37	n.a.	n.a.	n.a.	n.a.
EPFRP	***Rural transport infrastructure dummy (EPFRP)***	**n.a.**	**n.a.**	**n.a.**	**n.a.**	**0,84**	**0,37**	**0,83**	**0,37**	**0,83**	**0,37**	**0,83**	**0,37**
Aggregate agricultural - Year effects -	Length of Roads Network per total area of District (km / km2)	7,47	4,32	7,47	4,32	7,47	4,32	7,47	4,32	7,47	4,32	7,47	4,32
	Cotton-specific effect (OLS fitted values)	0,148	0,049	0,146	0,048	0,118	0,055	0,121	0,057	0,122	0,042	0,113	0,046
Agricultural extension services	Information on marketing for agricultural products	0,46	0,50	0,39	0,49	0,33	0,47	0,30	0,46	n.a.	n.a.	n.a.	n.a.
	Use any of the advice received on Crop husbandry	0,28	0,45	0,20	0,40	0,20	0,40	0,01	0,10	n.a.	n.a.	n.a.	n.a.
	Use any of the advice received on Crop diversification	0,23	0,42	0,12	0,32	0,16	0,37	0,14	0,35	n.a.	n.a.	n.a.	n.a.
	Information on agricultural input supply	0,41	0,49	0,35	0,48	0,32	0,47	0,23	0,42	n.a.	n.a.	n.a.	n.a.
Geographic Variables	Proportion of sample in Catchment Areas	0,85	0,36	0,85	0,36	0,84	0,37	0,84	0,37	0,83	0,37	0,83	0,37
	Proportion of sample in Control Areas	0,15	0,36	0,15	0,36	0,16	0,37	0,16	0,37	0,17	0,37	0,17	0,37
	Distance to the nearest all-weather road	1,374	0,603	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
	Distance to the nearest input market	1,855	0,784	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
	Rainfall	831,5	122,9	716,0	81,4	788,2	148,1	667,1	93,8	980,1	203,6	723,7	89,4
Cotton Observations		421		378		388		279		467		492	
Total number of Observations		1219		1197		1255		1427		1249		1403	

Source: Authors' estimations based on PHS.

Distribution of the Treatment and Comparison Samples

The sample characteristics of the comparison group and the treatment group highlight the role of randomization in the sense that the distribution of the covariates for the treatment and control groups are not significantly different. The age of the head of household in 1996/97 was only 2 years higher in the catchment districts, whereas in 2001/2002 it was almost similar. The size of the household was likewise equivalent in both 1996/97 and 2001/2002, although a bit higher in the catchment areas in entire period, exclusive in 1998/1999. The same could be said about the number of males in the household with the number in the catchment areas again being slightly higher (see **tables A7.1-2**). This implies that treatment with any of these covariates would allow us to find good comparisons in the control group (i.e. the overlap or the common support condition), or in other words, as mentioned above we can analyze data from subsamples with the same value of the covariates, as if they came from a completely randomized experiment (see section 3).[41]

[41] If the difference between the average values of the covariates in the two groups is large, the results are sensitive to the linearity assumption. More generally, because we do not know the exact nature of dependence of the assignment on the covariates, this results in increased sensitivity to model and a priori assumptions (Grilli and Ramphicini, 2011).

A more synoptic way to view these differences is to use the estimated **propensity score** as a summary statistic.

5.2. Evaluation of the EPFRP's impact on Cotton Productivity

The standard problem in treatment evaluation involves the inference of a causal connection between the treatment and the outcome. In our single-treatment case in each cross-section we observe (y_i, x_i, D_i; i = 1, …, N) the vector of observations on the scalar-valued outcome variable **y**, a vector of observable variables **x**, a binary indicator of a treatment variable **D**, and let **N** denote the number of randomly selected individuals who are eligible for treatment. Let $\mathbf{N_T}$ denote the number of randomly selected individuals who are treated and let $\mathbf{N_{NT}} = N - N_T$ denote the number of non-treated individuals who serve as a potential control group.

We would like to obtain a measure of the impact of the EPFRP intervention in **D** on **y**, holding **x** constant. The situation is akin to one of missing data, and it can be tackled by methods of causal inference carried out in terms of (policy-relevant) **counterfactuals**. We ask how the outcome of an average untreated individual household would change if such a person were to receive the treatment. That is, the magnitude $\mathbf{\Delta y/\Delta D}$ is of interest. Fundamentally our interest lies in the outcomes that result from or are caused by the EPFRP interventions. Here the causation is in the sense of *ceteris paribus* (Cameron & Triverdi, 2005).[42]

Using observational PHS data for Zambia, we first find that panel data doesn't exist (See **figures A5a-c**). Instead we find *repeated annual (i.e. equal spaced) sequence of independent*[43] *cross-sectional PHSs* based on a 'relative' large random sample of the population (see **table A5e vs. table A5h.2**). However, there is no random assignment mechanism for treatment. For this *cross-section survey*, it is impossible to track the same household over time as required in a genuine panel, because the sample design does not attempt to retain the same units in the sample. Instead, Deaton(1985, 1997) suggests tracking cohorts and estimating economic relationships based on **cohort means** rather than individual observations. Deaton(1985) argued that these pseudo-panels do not suffer the attrition problem that plagues genuine panels, and may be available over longer time periods compared to genuine panels (Baltagi, 2001).

$$E(Y^{obs} \mid D = 1) - E(Y^{obs} \mid D = 0) = E(Y_1 \mid D = 1) - E(Y_0 \mid D = 0) = E(Y_1 \mid D = 1) - E(Y_0 \mid D = 0) + [E(Y_0 \mid D = 1 - (E(Y_0 \mid D = 0)] = ATE + bias \tag{5.1}$$

The **average** **selection bias** is the *difference* between programme participants (i.e. the treated) and nonparticipants in the base state (Y_0) (i.e. non-treated outcomes in the absence of treatment) (Eq.5.1).[44] This effect cannot be attributed to the programme. Thus, *selection bias* arises when the treatment variable (D) is correlated with the error (ε) in the outcome equation.[45] This correlation could be included by incorrectly omitted observable variables that partly determine D and y. Then the omitted variable component of the regression error will be correlated with D – the case of

[42] The problem with Least Squares and matching approaches are that they simply assume away all three sources of bias (Arcand, 2012).
[43] Independence means that each subject appears in only one survey (Cameron & Triverdi, 2005:770f).
[44] ATT is identified only if $[E(Y_0 \mid D = 1 - (E(Y_0 \mid D = 0)] = 0$, i.e. if the outcomes of individuals from the treatment and comparison groups would not differ in the absence of treatment (Grilli and Rampichini, 2011).
[45] "Garden variety" endogeneity in which, for example, common unobservables determine both treatment status and outcomes (Arcand, 2012).

selection on observables (i.e. unbalance in observed confounders between the groups of treated and control units). Another source of *selection bias* comprises unobserved factors that partly determine both D and y. This is the case of *selection on unobservables* (i.e. unbalance in unobserved confounders between the groups of treated and control units) (op.cit., p.868; Grilli and Rampichini, 2011).

In our observational PHS data **the problem of selection of observables** is solved using regression and matching methods, which rely on the underlying identifying assumption *'unconfoundedness'* (selection on observables or conditional independence) (see Eq.(3.11)). The subsequent sections use these methods in order to avoid model dependence.

Matching and Propensity Score Estimators Approach

If the difference between the average values of the covariates in the two groups is large, the results are sensitive to the (simple or multiple) linear regression model linearity assumption. More generally, because we do not know the exact nature of dependence of the assignment on the covariates, this results in increased sensitivity to model and a priori assumptions. The choice of covariates to be included in the model strongly affects results (cf. specification of propensity score) (Grilli and Rampichini, 2011). In order to avoid model dependence in this section we apply matching techniques.

The question of how many *comparison units* to match with *each treatment unit* is closely related. One method of *selecting a set of comparison units* is ***the nearest-neighbor method***, which selects the *m comparison units* whose propensity scores are closest to the treated unit in question. Another method is ***caliper matching***, which uses all of the comparison units within a predefined propensity score ***radius (or "caliper")***. A benefit of caliper matching is that it uses only as many comparison units as are available within the calipers, allowing for the use of extra (fewer) units when good matches are (not) available (Dehejia and Wahba, 2002).

We consider a range of these simple estimators the results of which are shown in **table 5.2a.** For *matching without replacement*, we consider low-to-high, high-to-low, and random matching. In these methods, the treated units are ranked (from lowest to highest or highest to lowest propensity score, or randomly). The highest-ranked unit is matched first, and the matched comparison unit is removed from further matching. *For matching with replacement*, we consider ***single-nearest neighbor matching*** and ***caliper matching*** for a range of calipers.

We implement a ***full Mahalanobis matching*** and a variety of propensity score matching methods to adjust for pre-treatment observable differences between a group of treated and a group of untreated. Treatment status is identified by EPFRP==1 for the treated and EPFRP==0 for the untreated observations.

The propensity score - the conditional treatment probability - is estimated by the program on the independent variables. It is noted that the sort order of our data could affect the results when using *nearest-neighbor matching* on a propensity score estimated with categorical (non-continuous) variables. Or more in general when there are untreated with identical propensity scores. There are many options for *fine tuning the matching estimators* (Abadie et al., 2001). In **table 5.2a** we present the results of the following Matching methods: One-to-one (nearest neighbour or within caliper; with or without replacement), k-nearest neighbors, radius, kernel, local linear regression, 'spline-smoothing' and Mahalanobis matching using *'logyield'* as variable (alternatively using *'productivity'* as variable see **table A15.1).**

Table 5.2a: Matching and Propensity Score Estimators

Propensity score matching methods (i)	Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
	logyield	**Unmatched**	6,650	6,726	-0,075	0,043	-1,740
		ATT	6,701	6,924	-0,223	0,072	-3,080
		ATU	6,916	6,673	-0,243		
1. One-to-One propensity score matching (ii)		**ATE**			-0,233		
	logyield	**Unmatched**	6,650	6,726	-0,075	0,043	-1,740
		ATT	6,724	6,926	-0,201	0,073	-2,750
		ATU	6,909	6,761	-0,148		
2. K-nearest neighbors matching (iii)		**ATE**			-0,178		
	logyield	**Unmatched**	6,650	6,726	-0,075	0,043	-1,740
		ATT	6,729	6,919	-0,190	0,074	-2,560
		ATU	6,872	6,735	-0,137		
3. Radius matching (iv)		**ATE**			-0,168		
	logyield	**Unmatched**	6,650	6,726	-0,075	0,043	-1,740
		ATT	6,760	6,762	-0,003	0,052	-0,050
		ATU	6,852	6,852	0,000	-0,012	
4. Kernel (v)		**ATE**			-0,001		
	logyield	**Unmatched**	6,6503	6,7255	-0,0752	0,0432	-1,7400
		ATT	6,7638	6,7786	-0,0148	0,1456	-0,1000
		ATU	6,8330	6,8423	0,0093		
5.Local linear regression (vi)		**ATE**			-0,0045		
	logyield	**Unmatched**	6,650	6,726	-0,075	0,0432	-1,740
		ATT	6,724	6,923	-0,198	,	,
		ATU	6,909	6,700	-0,209		
6.'Spline-smoothing' (vii)		**ATE**			-0,203		
	logyield	**Unmatched**	6,650	6,726	-0,075	0,043	-1,740
		ATT	6,716	6,721	-0,005	0,063	-0,080
		ATU	6,690	6,700	0,009		
7. Mahalanobis matching (viii)		**ATE**			0,001		

Notes: (i) A variety of propensity score matching methods to adjust for pre-treatment observable differences between a group of treated and a group of untreated. Treatment status is identified by depvar==1 for the treated and depvar==0 for the untreated observations. (ii). (iii). (iv). (v) The uniform kernel type. (vi) The uniform kernel type. (vii) nknots(3). (viii) The uniform kernel type.
Source: Authors' estimations using the **PSMATCH2** Stata module.

The first estimator that we consider in **row one** of **table 5.2a** is the ***'One-to-One propensity score matching'***. We find that the difference between the matched treated and the matched controls is minus 0.223 while the T-statistics for H0 is minus 3.080 for ATT. In the **second row** we present the '***Nearest-neighbour matching'*** *without replacement* for which the treated unit i is matched to that non-treated unit j such that:

$$| p_i - p_j | = \min_{k \in \{D=0\}} \{| p_i - p_k |\} \qquad (5.2)$$

We calculate and display in **table 5.2a** the effect by the difference between the matched treated and the matched controls, which is minus 0.201 and T-statistics for H0 minus 2.75 in the case of ATT.[46] We achieve the best result by using ***'Kernel-based matching'*** as shown in **row 5**,

[46] The Abadie and Imbens (2002) procedure on match on the contrary allows individuals to be used as a match more than once, which generally lowers the bias but increases the variance.

that is the idea to associate to the outcome y_i of treated unit i a matched outcome given by a kernel-weighted average of the outcome of all non-treated units, where the weight given to non-treated unit j is in proportion to the closeness between i and j:

$$\hat{y}_i = \frac{\sum_{j \in \{D=0\}} K\left(\frac{P_i - P_j}{h}\right) y_j}{\sum_{j \in \{D=0\}} K\left(\frac{P_i - P_j}{h}\right)} \tag{5.3}$$

By choosing the *uniform kernel type* and imposing *common support* on the treated,[47] we find that the ATT difference between the treated and the control is almost zero (-0.003).

The difference is almost the same (-0.005) when carrying out ***Hahalanobis metric matching***, by replacing $p_i - p_j$ above with $d(i, j) = (P_i - P_j)' S^{-1} (P_i - P_j)$, where

- P_i is the (2x1) vector of scores of unit i
- P_j is the (2x1) vector of scores of unit j
- S is the pooled within-sample (2x2) covariance matrix of P based on the sub-samples of the treated and complete non-treated pool (Sianesi, 2001).

The fact that there is substantial overlap in the distribution of the propensity score between the comparison and treatment groups, explains why most of the matching algorithms yields similar results in **table 5.2a**. Therefore finding a satisfactory match by *matching without replacement* is appropriate given our PHS datasets.

In the output in **table 5.2b** above we estimate respectively the ATE; ATT; and ATC for the sample. Since cotton productivity is recorded in natural logarithm, the output in **row 1** in **table 5.2b** relying on only *a single match* implies that for the individual households in our sample, the SATE of benefiting from the EPFRP is a higher absolute increase for SATT of 0,192 compared to 0,057 for SATC. For all the specifications at hand we conclude that the sample ATTs are significantly different from zero at the 1% level, whereas the ATCs are insignificant, by using 3 matches.[48]

[47] Treated units whose p is larger than the largest p in the non-treated pool are left unmatched.

[48] We chose *3 matches* because it seemed to offer the benefit of not relying on too little information without incorporating observations that are not sufficiently similar. Like all smoothing parameters, the final inference can depend on the choice of the number of matches (Abadie et al., 2001, 2004).

Table 5.2b: Matching estimators for average treatment effects

No.	Matching estimator:	Number of matches m(#)	Number of matches, robust std. err. (h)	logyield	Coef.	Std.Err.	z	P>z	[95% Conf.	Interval]
1	Average Treatment Effect	1		SATE	-0,134	0,050	-2,680	0,007	-0,232	-0,036
	Average Treatment Effect for the Treated	1		SATT	-0,192	0,056	-3,410	0,001	-0,302	-0,082
	Average Treatment Effect for the Controls	1		SATC	-0,057	0,058	-0,970	0,330	-0,172	0,058
2 (i)	Average Treatment Effect	3		SATE	-0,148	0,046	-3,210	0,001	-0,238	-0,058
	Average Treatment Effect for the Treated	3		SATT	-0,216	0,050	-4,330	0,000	-0,314	-0,118
	Average Treatment Effect for the Controls	3		SATC	-0,057	0,051	-1,120	0,262	-0,156	0,042
3 (ii)	Average Treatment Effect	3		SATE	-0,187	0,046	-4,020	0,000	-0,277	-0,096
	Average Treatment Effect for the Treated	3		SATT	-0,268	0,051	-5,280	0,000	-0,368	-0,169
	Average Treatment Effect for the Controls	3		SATC	-0,078	0,050	-1,550	0,122	-0,176	0,021
4 (iii)	Average Treatment Effect	3	4	SATE	-0,187	0,044	-4,270	0,000	-0,272	-0,101
	Average Treatment Effect for the Treated	3	4	SATT	-0,268	0,046	-5,770	0,000	-0,359	-0,177
	Average Treatment Effect for the Controls	3	4	SATC	-0,078	0,049	-1,570	0,116	-0,174	0,019

Notes: 4662 observations dropped due to treatment variable missing. Number of observations = 2163. *Matching variables:* Age Agesq Sex shareofmale loghhsize stratum basalprha Topdresprha livestock Areapc Clandfrac rain_EP. *Bias-adj variables:* Age Agesq Sex shareofmale loghhsize stratum basalprha Topdresprha livestock Areapc Clandfrac rain_EP. (i) Homoskedastic errors are estimated. (ii) The *nnmatch* estimate heteroskedasticity-consistent standard errors using # matches in the second matching stage (across observations of the same treatment level). (iii-iv) We estimate the ATE; ATT and ATC with bias-adjustment. The k*k diagonal matrix of the inverse sample standard errors of the k variables in varlist_nnmatch is used. (iii) Exclusively use the Bias Corrected Matching Estimator. (iv) Whereas the variance Estimation allows for Heteroskedasticity.

Since the standard error of the SATEs underestimates the standard error of the PATE, it is possible that the PATE might not be significantly different from zero at either the 5% nor the 1% level (Abadie et al., 2001, 2004). However, when considering launching another rural road rehabilitation and/or maintenance programme in Eastern Province in which we would obtain another sample from the same population, the absolute increase in PATT of -0,208 is higher compared to PATC of -0,061 and that PATT is significantly different from zero at the 1% level. Moreover, since our productivity data are in terms of logarithms, our results would indicate a statistically significant but also economically important impact of the EPFRP on the individual rural household in the pooled PHS samples covering the period from 1996/1997 to 2001/2002.

Finally, as discussed in Imbens (2003) and Heckman et al. (1998) the effects of the treatment on the sub-population of treated units (SATTs) are more important than the effect on the population as a whole (SATE) as shown by our results displayed in **table 5.2b**.

The Bias Corrected Matching Estimator

The simple matching estimator will be **biased in finite samples** when the matching is not exact. In finite samples there is a trade-off between the plausibility of the unconfoundedness assumption and the variance of the estimates.[49] When using all the available covariates, bias arises from selecting a wide bandwidth in response to the weakness of the common support. Whereas when using a lower number of covariates, common support is not a problem but the plausibility of the unconfoundedness assumption is (Grilli and Ramphicini, 2011).

[49] Matching just one nearest neighbor minimizes bias at the cost of larger variance. Matching using additional nearest neighbors increase the bias but decreases the variance (Grilli and Ramphicini, 2011).

Abadie and Imbens (2002) show that with **k** continuous covariates the estimator will have a term corresponding to the matching discrepancies (the difference in covariates between matched units and their matches) that will be of the order $\mathbf{O_p(N^{-1/k})}$. In practice one may therefore attempt to *remove some of this bias term* that remains after the matching. *The bias-corrected matching estimator* adjusts the difference within the matches for the differences in their covariate values. The adjustment is based on an estimate of the two regression functions $\mu_\omega(x) = E[Y(\omega)|X = x]$.

Following Rubin (1973) and Abadie and Imbens (2002) we approximate these regression functions by linear functions and estimate them using least squares on the matched observations (Abadie et al., 2001, 2004).

Using the Bias Corrected Matching Estimator for the ATE:

$$\hat{\tau}_M^{bcm} = \frac{1}{N}\sum_{i=1}^{N}(\tilde{Y}_{1i} - \tilde{Y}_{0i})$$

And the bias-adjusted matching estimators for ATT and ATC:

$$\hat{\tau}_M^{bcm,t} = \frac{1}{N_1}\sum_{i:D_i=1}(\tilde{Y}_{1i} - \tilde{Y}_{0i}), \quad \text{and} \qquad \hat{\tau}_M^{bcm,c} = \frac{1}{N_0}\sum_{i:D_i=0}(\tilde{Y}_{1i} - \tilde{Y}_{0i})$$

We estimate the SATE, SATT and SATC in **rows 3**. We find that this approach both increase the absolute size of the coefficients and decrease the standard errors, while not changing our previous conclusion that EPFRP treatment had an effect on its participants that still is significant at the at the 1% level.[50]

Variance Estimation Allowing for Heteroskedasticity

In **row 4** we show the results for the variance of the **SATE**:

$$\hat{V}^{sample} = \frac{1}{N^2}\sum_{i=1}^{N}\left(1 + K_{Mi}\right)^2 \hat{\sigma}_{D_i}^2(X_i)$$

Similarly the variance for the estimator for **SATT** is:

$$\hat{V}^{sample,t} = \frac{1}{N_1^2}\sum_{i=1}^{N}\left(D_i - (1 - D_i)K_{Mi}\right)^2 \sigma_{D_i}^2(X_i)$$

and for **SATC**,

$$\hat{V}^{sample,c} = \frac{1}{N_0^2}\sum_{i=1}^{N}\left(D_i K_{Mi} - (1 - D_i)\right)^2 \sigma_{D_i}^2(X_i)$$

We estimate these variances by estimating the conditional outcome variance $\hat{\sigma}_{\varpi}^2(x)$, which is assumed *not to be constant* (i.e. ***heteroskedastic***) for both treatment groups (ω) and all values of the covariates (x). This is implemented using a second matching procedure, now matching treated units to treated units and control units to control units (Abadie et al., 2001, 2004). In other words, the SATE; SATT; and SATC is re-estimated in **row 4**, but compared to row 1-3 we estimate the standard error allowing for heteroskedasticity, while specifying 3 data matches in estimating the conditional variance functions. Our results show that when the standard error is estimated under these weaker conditions the estimated SATE and SATT are still significant at the 1% level. The in **row 4** the EPFRP appears to have had exactly the same significant impact on the beneficiaries as in **row 3**, although standard errors are slightly smaller by taking account of heteroskedasticity.

[50] The bias-adjustment does not affect the form of the estimator for the variance, although it may affect the numerical value. For the variance it does matter whether one is interested in the sample of population average treatment effect (or the average effect for the treated or controls) (Abedie et al., 2001, 2004).

Table 5.2c: One-to-One Matching: Sample characteristics and estimated impacts

Control Sample	No. Of Observation	Mean Propensity Score (i)	Age	Agesq	Sex	loghhsize	stratum	livestock	Areapc	Clandfrac	ATT: Treatment Effect (Diff. In Means) (ii)	Log (pseudo) likelihood	Pseudo R2
Probit (iii)	5276	0.52088	-0.0213516*	0.0002	0.084073*	0.3286929***	-0.202497***	0.0846107**	0.43318***	-0.2176546**		-3600.0551	0.0143
			0.012	0.0001	0.045	0.041	0.059	0.037	0.066	0.088			
Logit (iv)	5276	0.47465	-0.0493649**	0.000428*	0.123	0.5453476***	-0.2137555*	0.086	0.921236***	-0.434159**		-3591.3135	0.0147
			0.022	0.0003	0.081	0.078	0.114	0.069	0.163	0.174			
Probit (v)	0.0696437		-0.030978**	0.000269*	0.077	0.3381171***	-0.1317649*	0.054	0.5625045***	-0.26809**		-3591.5151	0.0147
Without replacement:													
Random (vi)	2163		-0.012	0.00004	0.2256***	0.14728**	-0.021	0.2518***	0.1656*	-1.086***	-0.120	-1422.666	0.0365
			0.020	0.0002	0.081	0.073	0.080	0.059	0.098	0.135	0.045		
Low to high (vii)	2163		-0.012	0.0000	0.2256***	0.14728**	-0.021	0.2518***	0.1656*	-1.08658***	-0.120	-1422.666	0.0365
			0.020	0.0002	0.081	0.073	0.080	0.059	0.098	0.135	0.045		
High to low (viii)	2163		-0.012	0.0000	0.22566***	0.14728**	-0.021	0.2518***	0.1656*	-1.086***	-0.030	-1422.666	0.0365
			0.020	0.0002	0.081	0.073	0.080	0.059	0.098	0.135	0.047		
With replacement:													
Nearest neighbor (ix)	2163		-0.012	0.0000	0.2256***	0.14728**	-0.021	0.2518***	0.1656*	-1.08658***	-0.216	-1422.666	0.0365
			0.020	0.0002	0.081	0.073	0.080	0.059	0.098	0.135	0.063		
Caliper, 0.00001 (x)	2163		-0.012	0.0000	0.2256***	0.14728**	-0.021	0.2518***	0.1656*	-1.08658***	-0.270	-1422.666	0.0365
			0.020	0.0002	0.081	0.073	0.080	0.059	0.098	0.135	0.167		
Caliper, 0.00005 (xi)	2163		-0.012	0.0000	0.2256***	0.1473**	-0.021	0.2518***	0.165*	-1.086***	-0.205	-1422.666	0.0365
			0.020	0.0002	0.081	0.073	0.080	0.059	0.098	0.135	0.100		
Caliper, 0.0001 (xii)	2163		-0.012	0.0000	0.2256***	0.1473**	-0.021	0.2518***	0.165*	-1.086***	-0.174	-1422.666	0.0365
			0.020	0.0002	0.081	0.073	0.080	0.059	0.098	0.135	0.077		

Notes: Variables: shareofmale; basalprha; Topdresprha and rain_EP are not balanced and therefore left out of the specification. (i) The propensity score is estimated using a logit of treatment status on. (v) [pweight=wgt] if productivity>0, pscore(mypscore6) comsup level(0.01). (vi) outcome(logyield) noreplacement common. (vii) outcome(logyield) noreplacement common. (ix) outcome(logyield) common. (x) outcome(logyield) neighbor(2) caliper(0.00001) common. (xi) outcome(logyield) neighbor(2) caliper(0.00005) common. (xii) outcome(logyield) neighbor(2) caliper(0.0001) common.
Source: Author estimation based on psmatch2 (Leuven and Sianesi, 2003) available from ssc desc psmatch2.

The results of the propensity score methods showed in **table 5.2c** assume a common support, i.e. the range of propensities to be treated is the same for treated and control cases, even if the density functions have quite different shapes (**figures 5.1a-b**).

Figure 5.1a. Histogram of estimated propensity score, Treated

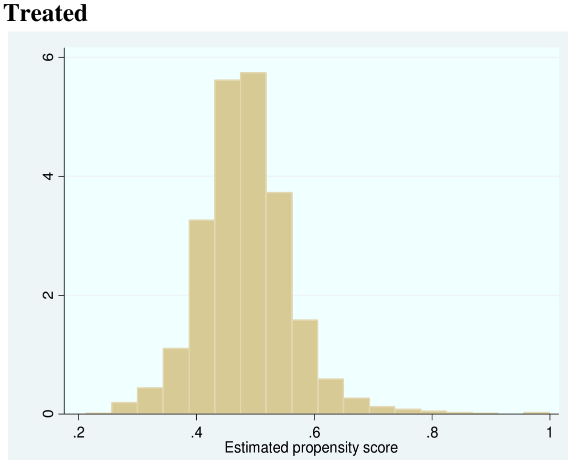
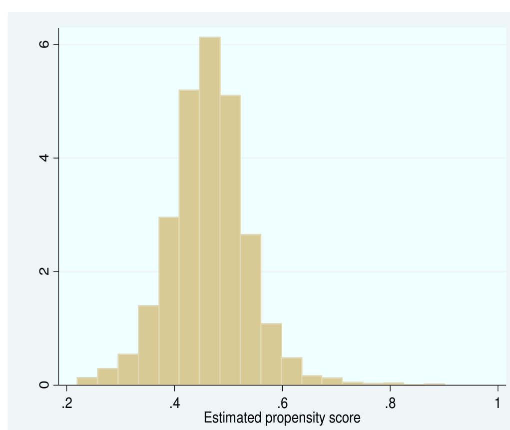


Figure 5.1b. Histogram of estimated propensity score, Controlled



Source: Authors' calculations.

5.3. Robustness Checks: Tests of the Matching Assumption and Sensitivity of Estimates

In this section we test the matching assumption and examine the sensitivity of our estimates to the specification of the propensity score.

Above (see section 5.2.1) we estimated the probability of getting the treatment as a function of observable pre-treatment covariates (using a Logit model and Probit model). We used the predicted values to generate propensity score p(x) for all treatment and control units (**table 5.2c** above). In this section we check **the balancing**, that is, we test that the means of each covariate do not differ between treated and control units, which is a precondition for trusting the ATT estimation.

In **table 5.4** we calculate several measures of the balancing of the independent variables before and after matching. For each regressor it calculates (a) *t-tests for equality of means* in the treated and non-treated groups, both before and after matching. For good balancing, these should be non-significant after matching. T-tests are based on a regression of the variable on a treatment indicator. Before matching this is an unweighted regression on the whole sample, after matching the regression is weighted using the matching weight variable and based on the on-support sample.

(b) *The standardised bias* before and after matching, together with the achieved percentage reduction in abs(bias). The standardised bias is the difference of the sample means in the treated and non-treated (full or matched) sub-samples as a percentage of the square root of the average of the sample variances in the treated and non-treated groups.[51] The standardized bias should be less than 5% after matching.

Table 5.4: Covariate imbalance testing

		Mean			%reduct	t-test	
Variable	Sample	Treated	Control	%bias	bias	t	p>t
Age	Unmatched	41.381	41.527	-1.3		-0.51	0.607
	Matched	40.607	39.982	5.4	-325.6	-4.29	0.000
Agesq	Unmatched	1845	1859.9	-1.5		-0.60	0.550
	Matched	1777	1726.5	5.0	-239.5	-5.17	0.000
Sex	Unmatched	1.2386	1.2319	1.6		0.63	0.527
	Matched	1.2169	1.1847	7.6	-385.9	-3.63	0.000
loghhsize	Unmatched	1.6963	1.6342	11.5		4.66	0.000
	Matched	1.7106	1.7673	-10.5	8.6	-4.12	0.000
stratum	Unmatched	1.1443	1.1397	1.3		0.52	0.602
	Matched	1.1572	1.2358	-22.5	-1637.3	-3.22	0.001
livestock	Unmatched	.51474	.4726	8.4		3.42	0.001
	Matched	.54003	.66925	-25.9	-206.6	-11.85	0.000
Areapc	Unmatched	.36411	.34372	5.9		2.34	0.019
	Matched	.38772	.48696	-28.8	-386.8	-2.11	0.035
Clandfrac	Unmatched	.09763	.12224	-12.7		-5.10	0.000
	Matched	.12518	.31539	-97.8	-672.8	-1.21	0.228

Note: **pstest** covariates, summary mweight(_weight) treated(EPFRP) support(comsup).
Source: Authors estimations.

We find that the means of all the covariates differ, but the *standardised bias (%bias)* is smaller before matching for all covariates than after matching, except for the logarithm of household size (loghhsize). In other words, the matching doesn't reduce the starting unbalancing. Moreover, the standardized bias is not less than 5% after matching for any of the covariates, which means that

[51] The stata formula is taken from Rosenbaum and Rubin, 1985.

none of the covariates are well balanced and the matching was not effective in building a good control group. Therefore that *the balancing property* is not satisfied and a less parsimonious specification of $h(X_i)$ is needed.[52]

We also find that the t-tests for equality of means in the treated and non-treated groups after matching are significant for all covariates except for the fraction of land devoted to cotton, which is insignificant (Clandfrac).

Finally, to assess *the bias of causal effect estimates* when the unconfoundedness assumption is assumed to fail in some specific and meaningful ways e.g. Ichino, Mealli and Nannicini(2006) proposed a strategy implemented in the **sensatt** command of STATA, which implements ***the sensitivity analysis for matching estimators***. The analysis builds on Rosenbaum and Rubin (1983) and Rosenbaum (1987), and *simulates a potential binary confounder* in order to assess the robustness of the estimated treatment effects with respect to specific deviations from the Conditional Independence Assumption (CIA).

As a first step, the **Average Treatment effect on the Treated (ATT)** is estimated by using one of the following propensity-score matching estimators: Nearest Neighbor (No.1 and No.3b-f); Radius (No.5a-f); Kernel (no.2) and Stratification (6a-b) (see **table 5.5**).[53]

As a second step, **a potential binary confounder (U)** is simulated in the data, on the basis of four parameters: pij (with i,j=0,1). Defining Y as the outcome (or as a binary transformation of the outcome in the case of continuous outcomes (i.e. cotton productivity)) and T as the binary treatment (i.e. EPFRP), each simulation parameter pij represents the probability that U=1 if T=i and Y=j (see **tables A14.3-6**).

Finally, U is considered as any other covariate and is included in the set of matching variables used to estimate *the propensity score and the ATT*. The imputation of U and the ATT estimation are replicated many times (we chose to perform 250 iterations), and a simulated ATT is retrieved as an average of the ATTs over the distribution of U. This estimate is robust to the specific failure of the CIA implied by the parameters pij. The comparison of the simulated ATT (**tables A14.3-6**) and the baseline ATT (**table 5.5; table A14.2)** tells us that the latter is robust with regards to the nearest neighbor matching method.

[52] The PS: $Pr(D_i=1|X_i) = F(h(X_i))$, where $h(X_i)$ is a function of covariates with linear (Age; Area per capita; log household size) and higher order terms (e.g. Age squared). F(.) is a cumulative distribution.
Since *balancing is not achieved/satisfactory*, the previous steps should be repeated by modifying the recursive matching algoritm and/or modifying the propensity score model.
[53] The options that are common to these commands specify how the baseline ATT is estimated (see Becker and Ichino, 2002) (Tommaso Nannicini, help sensatt).

Table 5.5: A simulation-based sensitivity analysis for matching estimators

	ATT Estimation	Baseline: With no simulated confounder														
		Analytical standard errors (i)					Bootstrapped standard errors					Options				
No.	Matching method	n.treat.	n. contr.	ATT	Std.Err.	t	n.treat.	n. contr.	ATT	Std.Err.	t	Rep	Logit	Comsup	Index	Radius
1.	Nearest Neighbor	4070	1818	-0.022	0.012	-1.947							No			
2.	Kernel	4070	2755	-0.023	0.008	-2.753	4070	2755	-0.023	0.008	-2.753		No			
3b	Nearest Neighbor	4070	1834	-0.032	0.011	-2.825	4070	1834	-0.032	0.011	-2.803	100	No			
3c	Nearest Neighbor	4070	1809	-0.027	0.012	-2.349	4070	1809	-0.027	0.011	-2.483	100	Yes			
3d	Nearest Neighbor	3870	1795	-0.031	0.011	-2.725	3870	1795	-0.031	0.011	-2.884	100	No	Yes		
3e	Nearest Neighbor	4070	1834	-0.032	0.011	-2.844							No		Yes	
3f	Nearest Neighbor	4070	1808	-0.027	0.012	-2.339							Yes	Yes	Yes	
5a	Radius	3869	2716	-0.024	0.009	-2.533										0.05
5b	Radius	3864	2716	-0.041	0.010	-4.279										0.01
5c	Radius	3870	2716	0.000	0.009	0.029										0.25
5d	Radius	3869	2716	-0.024	0.009	-2.562							Yes			
5e	Radius	3867	2716	-0.038	0.010	-3.918						100			Yes	
5f	Radius	3869	2716	-0.024	0.009	-2.533						100		Yes		
6a	Stratification	3967	2732	0.007	0.009	0.722	3967	2732	0.007	0.010	0.648	100				
6b	Stratification	3967	2732	0.007	0.009	0.722	3967	2732	0.007	0.010	0.650	100		Yes		

Source: Author's estimations.

6. Conclusions

Despite their popularity (the influence of rural roads is thought to be largely confined to well-defined zones (such as villages), which renders it suitable for impact evaluation (ADB, 2011)[54]) very few aid-financed rural road projects in developing countries have been the subject of rigorous impact evaluations (for recent attempts see Kingombe, 2012a; Kingombe and di Falco, 2012). Knowledge about their impacts and the heterogeneity in those impacts continues to be limited (van de Walle, 2009).

This paper has investigated the impacts of the improvements of rural feeder roads through the implementation of the EPFRP on farm cotton yields in rural Zambia's Eastern Province by using the non-parametric matching techniques. The outcome plays no role in the propensity score.[55]

We find the ATT estimation results are not the same when implementing various matching using 'the logarithm of (cotton) yield' compared to using 'cotton productivity' as variable.

In the latter case the following matching methods all have negative difference between treated and controls: 1-to-1 propensity score matching; k-nearest neighbors matching; radius matching; and 'spline-smoothing'. However, the Kernel matching has positive difference between treated and controls for the 'productivity' variable: Finally, some of the local linear regression and the Mahalanobis matching specifications yields positive difference between treated and controls for the 'logyield' variable, but not for the 'productivity' variable and not for all specifications either.

Through our robustness checks of the Matching Assumption and Sensitivity of Estimates we find that the matching doesn't reduce the starting unbalancing. The comparison of the simulated ATT and the baseline ATT tells us that the latter is robust. We conclude that the application of various non-parametric matching methods didn't enable us to identify a robust linkage, most likely due to the PHS data source and the evaluation design.

The main assumption underlying the matching approaches (estimating and assessing causal effects under unconfoundedness involving the propensity score) is the same as OLS. As with OLS, the matching is as good as its relevant pre-treatment characteristics / covariates (X) are. However, unlike the OLS approach, matching avoids potential misspecification of $E(Y_0 | X)$ and it allows for arbitrary heterogeneity in causal effects: $E(Y_1 – Y_0 | X)$ (Grilli and Rampichini, 2011).

Non-parametric (or semi-parametric if the pscores are estimated using a parametric model like Logit) matching techniques and OLS is only appropriate when unconfoundedness (selection on observables) is plausible. The PSM forces the researcher to design the evaluation framework and check the data before looking at the outcomes (this should avoid cheating from the evaluator according to Grilli and Rampichini, 2011). Moreover, PSM also makes the comparison of treated and control units more explicit than OLS.[56]

---

[54] It is argued that rural roads can raise living standards and improve the welfare of poor rural households by increasing access to goods and services, stimulating agricultural production and diversification, and creating off-farm employment (ADB, 2011).

[55] Similar to controlled experiments in which the design of the experiment has to be specified independently of the outcome (Grilli and Rampichini, 2011).

[56] If treatment effects are homogeneous (rarely) or you know the correct functional form (rarely), then regression-based estimators are more efficient (lower variance) (op.cit., p.74).

It is important to think about the distinction between Average Treatment Effect (ATE); Treatment on the Treated (TT); and Treatment on the Untreated (TUT), which is crucial in terms of policy relevance (Arcand, 2012). While, acknowledging that some consider the application of Matching Techniques as plain stupid, because they assume away all three sources of bias, which occur in any piece of empirical work (Arcand, 2012), we also admit that the relevance of matching methods depends on the data availability for the specific policy evaluation problem. Furthermore, we acknowledged that rural roads pose challenges for evaluation. The benefits of rural roads are indirect and conditional on interactions with the geographic, community and rural household characteristics of their location. Road locations are typically determined by those same characteristics confounding inferences based on comparisons of places with roads versus without them. Additionally, impacts may be distributional, felt across multiple outcomes and take a long time to emerge. Van de Walle(2009) argues that these features of rural roads have implications for evaluation design and data collection.

Through our relative small sample based on an existing cross-sectional PHS dataset an acceptable balance on important covariates is rarely achieved (see **table A10.1**).

Given the state of information, the supporting development assistance for rural infrastructure (i.e. Aid for Trade) must support rigorous evaluation of rural infrastructure projects.[57] To the extent possible, the gold standard of RCTs should be applied, it being understood that the standard itself may not be attainable in practice (Kanbur and Rauniyar, 2009).[58] Future rural roads impact evaluation requires panel (with pre-intervention) data for project and appropriate non-project areas; detailed information on outcome indicators, baseline attributes and controls for heterogeneity and exogenous time varying factors (Van de Walle, 2009).

According to van de Walle(2009) such a database allows for an evaluation design that combines a double difference with controls for initial conditions either through propensity score matching, regression controls or an IV. The DID method is commonly used for estimating rural road effects, which involves subtracting the difference in the outcome of non-project areas before and after the intervention from the difference of the project areas before and after the intervention. However, it is difficult and rare to find a qualified IV for rural road placement. If a large number of rural roads were under consideration and the rule was that those linking villages with 30% or higher poverty incidence were selected, regression discontinuity designs would offer an alternative method for impact evaluation (ADB, 2011).

[57] A large proportion of development assistance has been allocated to infrastructure investment, mainly transport, energy, and water supply. Accordingly, the number of impact evaluation studies in these subsectors has increased in recent years (ADB, 2011).
[58] RCTs are themselves controversial, and critics are agreed that too much is claimed for them, that the sort of power attributed to them in establishing causality cannot be fulfilled in practice (see Arcand, 2012).

References

Aakvik, A. 2001. Bounding a matching estimator: The case of a Norwegian training program. Oxford Bulletin of Economics and Statistics 63: 115-143.

Abadie, A., D. Drukker, J. L. Herr, and G. W. Imbens. 2004. Implementing matching estimators for average treatment effects in Stata. Stata Journal 4(3): 290-311.

Abadie, A., David Drukker, Jane Leber Herr, Guido W. Imbens. 2001. Implementing Matching Estimators for Average Treatment Effects in Stata. The Stata Journal (2001) 1, Number 1, pp. 1–18.

Asian Development Bank (ADB). 2011. A review of recent developments in impact evaluation. Mandaluyong City, Philippines: Asian Development Bank, 2011. http://203.131.219.245/oldbranches/n_puey/News/world%20bank/A%20Review%20of%20Recent%20Developments%20in%20Imapcct%20Evaluation.pdf

Arcand, Jean-Louis. 2012. Development and the effectiveness of social programmes. GTAP, Geneva, 27 June 2012.

Baffes, J., and Ajwad M.I. 1998. Detecting Price Links in the Cotton Market. World Bank Policy Research Working Paper 1944.

Balat, J.F., & Guido Porto, 2005. Globalization and Complementary Policies: Poverty Impacts in Rural Zambia. NBER Working Papers 11175, National Bureau of Economic Research, Inc.

Balat, Jorge F. & Porto, Guido G., 2005. The WTO Doha Round, cotton sector dynamics, and poverty trends in Zambia. Policy Research Working Paper Series 3697, The World Bank.

Baltagi, B.H. 2001. Econometric Analysis of Panel Data. 2nd ed. New York: Wiley.

Baltagi, B.H. 1998. Econometrics. Springer.

Baum, C.F.2006. An Introduction to Modern Econometrics Using Stata. Stata Press.

Block, S. and C. Peter Timmer. 1994. Agriculture and Economic Growth: Conceptual Issues and the Kenyan Experience." November 1994. 104 pp. http://www.cid.harvard.edu/hiid/498.pdf

Boughton, D., D. Tschirley, B. Zulu, A.O. Ofico, H. Marrule. 2003. Cotton Sector Policies and Performance in Sub-Saharan Africa: Lessons Behind the Numbers in Mozambique and Zambia. Paper prepared for presentation at the 25th International Conference of Agricultural Economists, August 16-22, 2003, Durban, South Africa.

Bellemare, M.F. 2009. The Productivity Impacts of de Jure and de Facto Land Rights.

Blundell, R. and MaCurdy, T.E., 2000. Labour Supply: A Review of Alternative Approaches, in Handbook of Labour Economics, O.C. Ashenfelter and D.E. Card (Eds.), Volume 3A, 1559-1695, Amsterdam, North-Holland.

Brambilla, Irene & Porto, Guido G., 2006. Farm productivity and market structure: evidence from cotton reforms in Zambia. Policy Research Working Paper Series 3904, The World Bank.

Brambilla, I., & Guido Porto, 2005. "Market Structure, Outgrower contracts and Farm Output. Evidence from Cotton Reforms in Zambia," NBER Working Papers 11804, National Bureau of Economic Research, Inc.

Brautigam, D. 2009. The Dragon's Gift. The Real Story of China in Africa. Oxford University Press.

Caliendo M. and Kopeinig S. 2008. Some practical guidance for the implementation of propensity score matching Journal of Economic Surveys , Vol. 22, No. 1, pp. 3172.

Caliendo, M., R. Hujer, and S. L. Thomsen. 2005. The employment effects of job creation schemes in Germany. IZA Discussion Paper No. 1512. Bonn, Germany.

Cameron, A.C. and Pravin K.Trivedi. 2009. Microeonometrics Using Stata. Stata Press.

Cameron, A.C. and Pravin K.Trivedi. 2005. Microeonometrics. Methods and Applications. Cambridge University Press.

Cochran, W. and Rubin, D.B. 1973. Controlling Bias in Observational Studies, Sankyha, 35, 417-446.

Conradie, B., Graham Cookson, and Colin Thirtle. 2006. Efficiency and farm size in Western Cape Grape Production: Pooling small datasets. South African Journal of Economics, 74:2, 334-43, June 2006.

Conway, G., Jeff Waage, and Sara Delaney. 2010. Science and Innovation for Development. UKCDS January 2010.

CSO. 2002. Republic of Zambia Post Harvest Survey 2001/2002 (for small and medium scale holdings). Interviewer's instruction manual. Central Statistical Office (agriculture and environment division) in conjunction with Ministry of Agriculture and cooperatives (database management and early warning unit). November 2002.

CSO. 2000a. Post Harvest Survey - Documentation of General Files.

CSO. Post Harvest Surveys (PHS) for Small and Medium-Sized Holdings (Synthetic Questionnaire).

CSO. 1997. Post-Harvest Survey 1996/97 Interviewer's Instruction Manual. November 1997.

Deaton, A. 1985. Panel Data from Time Series of Cross-Section. Journal of Econometrics, 30, 109-126.

Deaton, A. 1997. The Analysis of Household Surveys. A microeconometric Approach to Development Policy. John Hopkins University Press. Baltimore and London.

Dehejia, R.H and Wahba, S. 2002. Propensity score-matching methods for Non-experimental causal studies. The Review of Economics and Statistics, February 2002, 84(1): 151–161

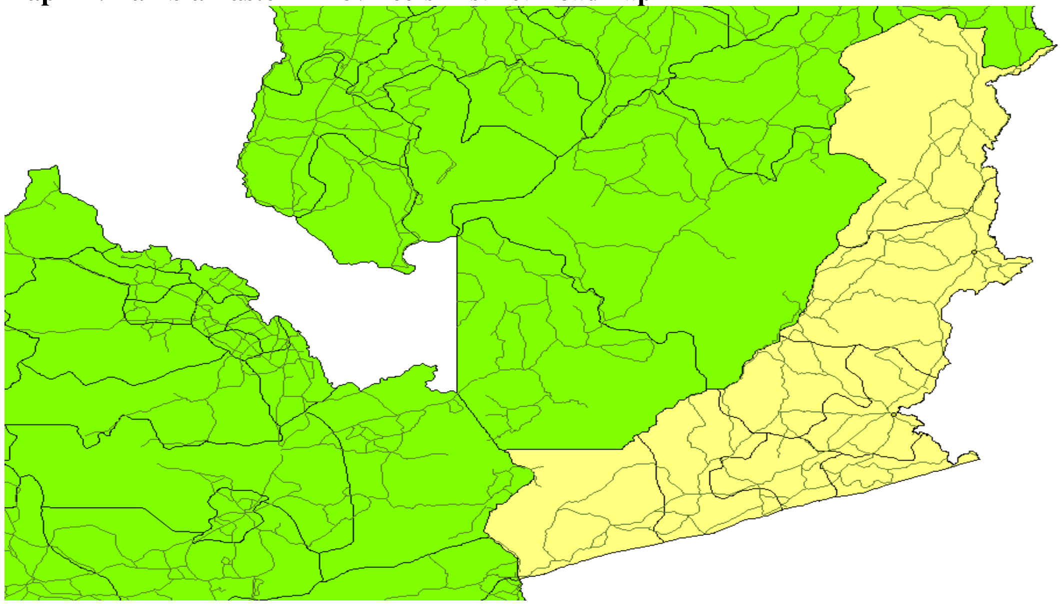
Dehejia, R.H and Wahba, S. 1999. Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programmes", Journal of the American Statistical Association 94, 1053-1062.
Donovan, C. 2007. Complex Survey Design & Other Challenges: Zambian Agricultural Surveys. Presentation to the Graduate-level Statistics Course STT 825: Sample Surveys. November 28, 2007. Wells Hall, Michigan State University.
Dorosh, P. A., Stifel, David and Minten, Bart, 2003. Transaction costs and agricultural productivity. MSSD discussion papers 56, International Food Policy Research Institute (IFPRI).
Dorward, A., Jonathan Kydd, Jamie Morrison and Ian Urey. 2004. A Policy Agenda for Pro-Poor Agricultural Growth. World Development Vol.32, No. 1, pp.73-89.
Fafchamps, M., Ruth Vargas Hill, Aliziki Kaudha and Robert Waggwa Nsibirwa. 2003. The transmission of international commodity prices to domestic producers. WPS/2003-14
FAO. 2009. Statistical On-line Crop Database. http://faostat.fao.org/site/567/DesktopDefault.aspx?PageID=567#ancor
Fibre2fashion News Desk – India. 2009. Mulungushi Textiles to be reopened after technical evaluation. 31st March.
FSRP Team. 2000. Key Challenges and Options Confronting Smallholder, Agribusiness and Government Leaders in Zambia's Cotton Sector. Number 1. 2000.
Govereh, J., Emma Malawo, Tadeyo Lungu, Thom Jayne, Kasweka Chinyama and Pius Chilonda. 2009. Trends and Spatial Distribution of Public Agricultural Spending in Zambia: Implication for Agricultural Productivity Growth-Draft for Review. Draft for Review. 2009.
Govereh, J., Michael Weber, Antony Chapoto and Steve Haggblade. 2008. Information and Analysis to Improve Agricultural Productivity & Reduce Rural Poverty in Zambia. Presentation at Seminar on Increasing Productivity in the Agricultural Sector to Contribute to the On-Going Macroeconomic Modeling Process Ministry of Finance and National Planning (MoFNP) - Tecla Lodge, Wed, 6th August, 2008.
Grilli, L., and Carla Rampichini. 2011. Propensity scores for the estimation of average treatment effects in observational studies. University di Firenze. Training Sessions on Causal Inference Bristol - June 28-29, 2011.
Guo S. and Fraser W.M. 2009. Propensity Score Analysis: Statistical Methods and Applications. Thousand Oaks, CA: Sage Publications.
Heckman, J.J., and Edward J. Vytlacil. 2007. Chapter 71 Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New Environments. Handbook of Econometrics. Volume 6, Part 2, 2007, Pages 4875-5143.
Heckman, J.J., Ichimura, H. and Todd, P.E. 1997. Matching As an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. Review of Economic Studies, 64, 605-654.
Heckman, J.J., Ichimura, H. and Todd, P.E. 1998. Matching as an Econometric Evaluation Estimator. Review of Economic Studies, 65, 261-294.
Heckman, J.J., Ichimura, H., Smith, J.A. and Todd, P. 1998. Characterising Selection Bias Using Experimental Data. Econometrica 66, 5.
Heckman, J.J., LaLonde, R.J., Smith, J.A. 1998. The Economics and Econometrics of Active Labour Market Programmes", in Ashenfelter, O. and Card, D. (eds.), The Handbook of Labour Economics Vol. 3A.
Hsiao, C. 1986. Analysis of Panel Data. Economics Society Monographs No.11. Cambridge University Press.
Ichino, A., F. Mealli, and T. Nannicini. 2006. From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity? IZA DP No. 2149.
Imbens, G. 2000. The Role of Propensity Score in Estimating Dose-Response Functions. Biometrika 87(3), 706-710.
Isiimwa, M. Crop Forecast Survey Methodology and Results. Ministry of Agriculture and Cooperatives.
Competition and Coordination in Cotton Market Systems in Southern and Eastern Africa Project. Imperial College, UK.
Kabwe, S. 2009. The Evolution of the Cotton and Textile Industry in Zambia. A presentation at ICAC Research ssociate Program, Washington DC, 10 April, 2009.
Kabwe, S. Effective Public-Private Coordination in Zambia's Cotton Sector: Deliberation on the Revised Cotton Act. FSRP.
Kabwe, S. and David Tschirley. 2007a. Increasing Demand for Quality in World Cotton Markets. How has Zambia Performed?
Kabwe, S., and David Tschirley. 2007b. An Effective Public-Private Coordination in Zambia's Cotton Sector: Deliberation on the Cotton Act. Presented at the Agricultural Consultative Forum. May, 2007.
Kabwe, S., and David Tschirley. 2007c. Input Credit Provision for Cotton Production: Learning from African Neighbors and Meeting Zambia's Challenges. Number 28. November 2007.
Kabwe, S., and David Tschirley. 2007d. Farm Yields and Returns to Farmers from Seed Cotton: Does Zambia Measure Up? Number 26. November 2007.
Kanbur, Ravi and Ganesh Rauniyar. 2009. Conceptualizing Inclusive Development: With Applications to Rural Infrastructure andDevelopment Assistance. ADB Occasional Paper No. 7.

Kasali, E. 2002. Report On The Consultancy To Measure Distances From The 1999-2000 Post Harvest Survey Sample Areas. To The Nearest Tarred Or Main Roads And District Towns.
Kingombe, C.K.M. 2012a. Regional Analysis of Eastern Province Feeder Road Project. District level estimation of the Poverty Alleviation Effects of Rural Roads Improvements in Zambia's Eastern Province. Graduate Institute of International and Development Studies Working Paper No: 10/2012.
Kingombe, 2012b. The Rural Transport Infrastructure and Marketing Linkage within the Context of the Sub-Regional Zambia-Malawi-Mozambique Growth Triangle. African Studies Association Biennial Conference 2012.
Kingombe, C.K.M., and Salvatore di Falco. 2012. The Impact of a Feeder Road Project on Cash Crop Production in Zambia's Eastern Province between 1997 and 2002. Graduate Institute of International and Development Studies Working Paper No: 04/2012.
Kydd, J., A., Dorward, Jamie Morrison and Georg Cadisch. 2004. Agricultural Development and Pro-poor Economic Growth in Sub-Saharan Africa: Potential and Policy. Oxford Development Studies, Vol. 32, No.1, March 2004.
Lalonde, R. J. 1986. Evaluating the Econometric Evaluations of Training Programs with Experimental Data. American Economic Review 76: 604–620.
Lechner, M. 2001. Identification and Estimation of Causal Effects of Multiple Treatments under the Conditional Independence Assumption, in: Lechner, M., Pfeiffer, F. (eds), Econometric Evaluation of Labour Market Policies, Heidelberg: Physica/Springer, p. 43-58.
Leuven, E., and B. Sianesi. 2003. "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.
Lokshin, M. Semi-parametric difference-based estimation of partial linear regression models.
Megill, D.J. 2005. Recommendations for Adjusting Weights for Zambia Post Harvest Survey Data Series and Improving Estimation Methodology for Future Surveys. Working Paper 13. March 2005.
Meyer, B.D. 1995. Natural and Quasi- Experiments in Economics. Journal of Business & Economic Statistics, American Statistical Association, vol. 13(2), pages 151-61, April.
Nichols, A. 2009. "Causal inference with observational data: Regression discontinuity and other methods in Stata," German Stata Users' Group Meetings 2009 02, Stata Users Group.
Austin Nichols, 2007. "Causal inference with observational data: Regression Discontinuity and related methods in Stata," North American Stata Users' Group Meetings 2007 2, Stata Users Group.
ODI. Zambia Agriculture Dataset: An Assessment of Trends in the Zambian Agriculture Sector.
Owens, T., John Hoddinott and Bill Kinsey. 2001. The impact of agricultural extension on farm production in resettlement areas of Zimbabwe.
People Daily.com. 2003a. New form of Chinese aid revives Zambia's largest textile company.
People Daily.com. 2003b.China, Zambia to strengthen economic ties.
Pletcher, J. 2000. The Politics of Liberalizing Zambia's Maize Markets. World Development Vol.28, No.1, pp.129-142.
Robinson, P. 1988. Root-n-consistent semi-parametric regression. Econometrica 56: 931-54.
Rosenbaum, P.R. 2002. Observational Studies. 2nd edition. New York: Springer.
Rosenbaum, P.R. and Rubin, D.B. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects", Biometrika 70, 1, 41-55.
Rosenbaum, P.R. and Rubin, D.B. 1985. Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score. The American Statistician 39(1), 33-38.
Sianesi, B. 2001. Implementing Propensity Score Matching Estimators with STATA. University College London and Institute for Fiscal Studies.
Smith, A, Jeffrey & E. Todd, Petra, 2005. "Does matching overcome LaLonde's critique of nonexperimental estimators?," Journal of Econometrics, Elsevier, vol. 125(1-2), pages 305-353.
Thurlow, J. & Wobst, P. 2005. Operationalizing Pro-Poor Growth. The Department for International Development.
Tschirley, D. Comparative Analysis of Cotton Sector Reforms in SSA. Overview of Approach and Broad Findings.
van Ark, B., and Marcel Timmer. 2006. Reader on productivity levels. Groningen Growth and Development Centre.
The Times of Zambia. 2008. Zambia: Cotton is Among the Foremost Export Crops.
Tschirley, D., and Stephen Kabwe. 2009. Cotton Sector Regulation in Zambia. A Preliminary Case Study. Presented at "Virtual Workshop on Cotton Sector Regulation in Africa" sponsored by the World Bank. 29 June 2009.
Tschirley, D., and Stephen Kabwe. Farm Yields and Returns to Farmers.
Tschirley, D. and Stephen Kabwe. 2007. Cotton in Zambia: 2007 Assessment of its Organization, Performance, Current Policy Initiatives, and Challenges for the Future. Working Paper No. 26.
Van de Walle, D. 2009. Impact evaluation of rural road projects. Journal of Development Effectiveness, Vol. 1, No.1, March, 15-36.
Weir, S., and John Knight. 2000a. Adoption and diffusion of agricultural innovations in Ethiopia: the role of Education. WPS/2000.5.

Weir, S., and John Knight. 2000b. Education externalities in rural Ethiopia: evidence from average and stochastic frontier production functions.
Wooldridge, J. 2002. Econometric Analysis of Cross Section and Panel Data. The MIT Press. Cambridge, MA.
Zambia Food Security Research Project. 2000. Improving smallholder and agri-business opportunities in Zambia's cotton sector: Key challenges and options. Lusaka, Zambia.
Zambia.Net. Business. http://www.zamnet.zm/newsys/news/viewnews.cgi?category=6&id=1068452021
Zambia.Net. Kabwe: From discovery to recovery. http://www.zamnet.zm/newsys/news/viewnews.cgi?category=4&id=1100549219
Zhang, Xiaobo & Fan, Shenggen, 2001. How productive is infrastructure?: new approach and evidence from rural India. EPTD discussion papers 84, International Food Policy Research Institute (IFPRI).
Zulu, B., and D. Tschirley. On Overview of the Cotton Sub-Sector in Zambia.
Zulu, B., and David Tschirley. 2004. An Assessment of Current Policy Initiatives in Zambia's Cotton Sector. Number 9.
Yatchew, A. 1998. Nonparametric regression techniques in economics. Journal of Economic Literature 36: 669-721.

Annexes

Map A1: Zambia Eastern Province's District Roadmap



Source: Authors' based on DIVA-GIS 5.2.

[redacted]

			Distance (km)		
	Longitude	Latitude	North-South	West-East	Area (km2)
Chadiza	32,3-33,2	13,9-14,3	28.33	75	2410.2
Chama	32,2-33,7	10,2-12,2	163.33	87.5	16209.6
Chipata	32,3-32,9	13-14	90	80	8166.3
Katete	31,5-32,3	13,7-14,4	51.67	55	3223.0
Lundazi	31,9-33,5	13,3-12,35	91.25	103.33	10694.6
Mambwe	31,5-32,2	13,5-12,9	90	60	6124.7
Nyimba	30-31	14-15	110	100	12476.2
Petauke	31-31,5	13,5-14,6	101.67	85	9801.4
Eastern	**32º 15' 0" E**	**13º 0' 0" S**	**726.25**	**645.833**	**69106.0**

Table A1b: Road and Population Density in the Districts of Eastern Province

Province	Capital	Area (km²)	Population	Density (people/km²)	Districts	Districts	Area (km²)	Population (2000)	Area Density (people/km²) (2000)	Total Length of Feeder Roads (*)	Road Density (km) Road length/ 100 km²	Road Density (km) Road length / 1000 population	Shape Area (GIS)
Central	Kabwe	94395	1,012,257	10.7	6	Chadiza	2410,18	83981	34,8	314,4	13,04	3,74	2170,16
Copperbelt	Ndola	31328	1,581,221	50.5	10	Chama	16209,64	7489	0,5	700,1	4,32	93,48	15304,33
Eastern	**Chipata**	**69106**	**1,306,173**	**18.9**	**8**	Chipata	8166,26	367539	45,0	402,35	4,93	1,09	4726,95
Luapula	Mansa	50567	775,353	15.3	7	Katete	3223,02	18925	5,9	506,8	15,72	26,78	3187,19
Lusaka	Lusaka	21898	1,391,329	63.5	4	Lundazi	10694,58	236833	22,1	727,6	6,80	3,07	11537,15
Northern	Kasama	147826	1,258,696	8.5	7	Mambwe	6124,69	70425	11,5	402,35	6,57	5,71	5112,32
North-Western	Solwezi	125827	583,35	4.6	12	Nyimba	12476,23	47376	3,8	404,05	3,24	8,53	6287,88
Southern	Livingstone	85283	1,212,124	14.2	11	Petauke	9801,40	235879	24,1	404,05	4,12	1,71	9198,50
Western	Mongu	126386	765,088	6.1	7	**Eastern**	**69106,00**	**1068447**	**15,5**	**3861,7**	**5,59**	**3,61**	**57524,48**
Zambia	**Lusaka**	**752616**	**9,885,591**	**13.1**	**72**								

Source: Author's calculations.

Map A2: Illustration of the Eastern Province Feeder Roads

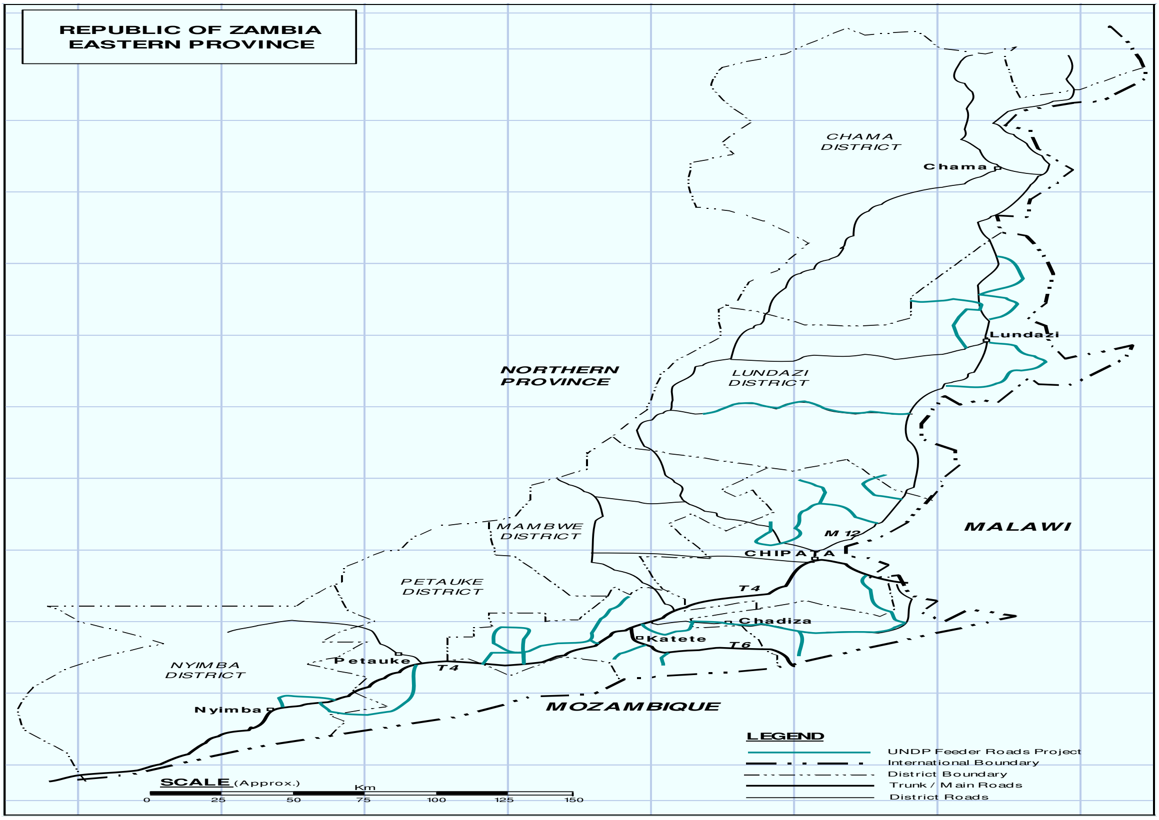


Table A2: Rainfall between 1994/95 and 2004/05 (in millimeter)

Year	1994/1995	1995/1996	1996/1997	1997/1998	1998/1999	1999/2000	2000/2001	2001/2002	2002/2003	2003/2004	2004/2005	Long-term Mean
Eastern	*528,55*	*805,71*	*813,72*	*719,02*	*759,38*	*678,55*	*914,78*	*700,40*	*835,42*	*781,11*	*788,61*	*756,84*
Chadiza (301) (i)	610,83	*868,50*	*945,33*	*708,50*	*901,42*	*584,83*	*1165,17*	*771,92*	*871,33*	*915,92*	*1007,92*	*850,15*
Chama (302) (iii)	494,08	*688,58*	*708,00*	*760,00*	*574,00*	*695,25*	*562,58*	*564,67*	*685,50*	*746,00*	*750,17*	*657,17*
Chipata (303) (i)	610,83	*868,50*	*945,33*	*708,50*	*901,42*	*584,83*	*1165,17*	*771,92*	*871,33*	*915,92*	*1007,92*	*850,15*
Katete (304)	214,17	*936,83*	*826,00*	*874,75*	*791,08*	*711,83*	*1163,50*	*846,17*	*877,47*	*878,01*	*895,40*	*819,56*
Lundazi (305)	693,58	*721,67*	*609,08*	*696,58*	*543,42*	*563,75*	*761,58*	*607,67*	*815,42*	*658,37*	*681,36*	*668,41*
Mambwe (306) (iii)	494,08	*688,58*	*708,00*	*760,00*	*574,00*	*695,25*	*562,58*	*564,67*	*685,50*	*746,00*	*750,17*	*657,17*
Nyimba (307) (ii)	555,42	*836,50*	*884,00*	*621,92*	*894,83*	*796,33*	*968,83*	*738,08*	*938,42*	*694,33*	*608,00*	*776,06*
Petauke (308) (ii)	555,42	*836,50*	*884,00*	*621,92*	*894,83*	*796,33*	*968,83*	*738,08*	*938,42*	*694,33*	*608,00*	*776,06*
Long-term Mean	756,84	*756,84*	*756,84*	*756,84*	*756,84*	*756,84*	*756,84*	*756,84*	*756,84*	*756,84*	*756,84*	

Notes: We assume the following coverage for the five weather stations: Chipata covers Chipata and Chadiza districts; Lundazi covers Lundazi district; Petauke covers Petauke and Katete districts; Msekere covers Katete district; and Mfuwe covers Chama and Mambwe districts.
Source: Author based on Zambia Meteorological Service data.

Table A3: Implementation of the EPFRP, 1996/1997 to 2001/2002

Agricultural Season	District Codes	1996/1997	1997/1998	1998/1999	1999/2000	2000/2001	2001/2002
Control Districts	302, 306, 307	Yes	Yes	Yes	Yes	Yes	Yes
Catchment Districts	301, 303, 304, 305, 308	No	No	Yes	Yes	Yes	Yes

Source: Authors.

The project area for a rehabilitated or new road can be defined as its "zone of influence", or *"catchment area,"* given by the area around the road where project impacts are expected (van de Walle, 2009).

Table A4: EPFRP treated roads in Catchment Districts

Road No.	Road Name	Rd Length (km)	Category	District
RD405	D128 - Zingalume - Mwangala	44	P	Chadiza
RD406	DR405 - D130	70	P	Chadiza
U3	T6 - Naviruli	7,6	P	Chadiza
Total EPFRP Primary Road Length		121,6	P	Chadiza
EPFRP Share of Primary of Feeder Roads in Chadiza		63,3%	Out of 192 km	
EPFRP Share of Total Length of Feeder Roads in Chadiza		***38,7%***	***Out of 314 km***	
RD118	M12-Tamanda Mission	6,7	P	Chipata
RD121	D104 - Chipalamba - D104	14,7	P	Chipata
RD400	D124 - Chiguya	12,2	P	Chipata
RD401	T4 - Madzimawe - D124	15,6	P	Chipata
RD595	T4 - Nzamane - Kazimuli	19,3	P	Chipata
RD596	RD595 - Sayiri - D128	25,5	P	Chipata
U33	Link RD 402 - Madzimoyo	8,3	P	Chipata
Total Primary Road Length		*102,3*	P	Chipata
EPFRP Share of Primary Length of Feeder Roads in Chipata		19,3%	Out of 529,4 km	
EPFRP Share of Total Length of Feeder Roads in Chipata		***12,7%***	***Out of 804,7 km***	
RD409	Chikonza - Walilanji - T6	14,9	S	Katete
RD585	T6 - Kalambana School	18,8	P	Katete
R292	Walilanji - Mbabala - T6	10	T	Katete
U23	Katete (T4) - Kazungulile - D598	33,2	P	Katete
U29	T6 - Mbinga - Katete	17,2	P	Katete
Total Primary Road Length		69,2	P	Katete
Total Secondary Road Length		14,9	S	Katete
Total Tertiary Road Length		10	T	Katete
Total EPFRP Road Length		94,1	Total	Katete
EPFRP Share of Primary Length of Feeder Roads in Katete		30,1%	Out of 229 km	
EPFRP Share of Total Length of Feeder Roads in Katete		***18,6%***	***Out of 506 km***	
RD107	D103 - Emusa - Chasefu - Chama Boundary	25,5	P	Lundazi
RD110	Lundazi (M12) - Mwase	27,6	P	Lundazi
RD110N	D109 - Kapachila - Mwase (RD110)	23,9	P	Lundazi
R243	Mphamba (D104) - Nyalubanga (D103)	35,4	P	Lundazi
R246	Phikhamalaza (R245) - R248	10	T	Lundazi
R250	RD110 N - Kanyunya School	8,2	T	Lundazi
R251	Mwase (RD 110) - Pono	17	S	Lundazi
R254	RD110 - Gwaba - Kamtande	17,2	P	Lundazi
R255	Mwase (RD110) - R254	9,5	T	Lundazi
U16	Gwaba (R254) - TBZ - Lumezi (M12)	25,1	P	Lundazi
U18	Kapachila - RD110	10,9	S	Lundazi
Total Primary Road Length		154,7	P	Lundazi
Total Secondary Road Length		27,9	S	Lundazi
Total Tertiary Road Length		27,7	T	Lundazi
Total Feeder Road Length		210,3	Total	Lundazi
EPFRP Share of Primary Length of Feeder Roads in Lundazi		31,1%	Out of 497,9 km	
EPFRP Share of Total Length of Feeder Roads in Lundazi		***28,9%***	***Out of 727,6 km***	
RD135	D139 - Sasare	21	P	Petauke
RD413	R12 Chataika - T4	21,2	P	Petauke
RD415	Minga (T4) - Nyalukomba (D414): 'D' State Road	19	P	Petauke
R13	T4 - Chikalawa School (R12)	25,1	P	Petauke
Total Primary Road Length		86,3	P	Petauke
EPFRP Share of Sub-Total Length of Feeder Roads in Petauke		16,5%	Out of 524 km	
EPFRP Share of Total Length of Feeder Roads in Petauke		***10,7%***	***Out of 808 km***	
Total Primary Road Length in Eastern Province		534,1	P	Eastern
EPFRP Share of Primary Length of Feeder Roads in Eastern		27,1%	Out of 1972 km	
EPFRP Share of Total Length of Feeder Roads in Eastern		***16,9%***	***Out of 3162 km***	

Source: Authors' calculations based upon EPFRP documents and Ministry of Local Government & Housing, 1998: Feeder Roads Support Programme. Volume 2: District Feeder Road Lists.
Notes: P = Primary Feeder Roads; S = Secondary Feeder Roads; and T = Tertiary Feeder Roads.

Table A5a: Number of Rural Households by Sex of Head by Province and District 2001/2002 Agriculture Season

Province	District	Gender of Household Head: Male	Female	Number of Rural Households
Eastern	Chadiza	10.385	3.856	14.241
	Chama	7.872	4.545	12.417
	Chipata	33.353	11.02	44.373
	Katete	21.347	11.058	32.405
	Lundazi	21.933	7.588	29.521
	Mambwe	8.143	1.102	9.245
	Nyimba	8.75	2.236	10.986
	Petauke	25.273	10.118	35.391
Province Total		**137.056**	**51.523**	**188.579**

Source: Author based on PHS 2001/02 dataset and CSO, 2002.

Table A5b: Number of Male Headed Households by Marital Status of Household Head during the 2001/2002 Agriculture Season by Province and District by Province and District

Province	District	Marital Status: Single	Monogamously Married	Polygamously Married	Divorced	Widowed	Seperated	Total
Eastern	Chadiza	174	7,897	616	-	132	-	10,385
	Chama	651	5,535	1,337	203	146	-	7,872
	Chipata	2,758	26,258	2,983	43	1 31	-	33,352
	Katete	2,625	15,812	1,916	289	705	-	21,347
	Lundazi	652	17,438	2,899	313	200	432	21,934
	Mambwe	357	7 03	387	182	187	-	8,143
	Nyimba	196	7,813	544	-	196	-	8,749
	Petauke	2,839	18,768	1,523	183	1,728	232	25,273
Province Total		**10,252**	**99,521**	**12,205**	**1,213**	**3,294**	**664**	**137,055**

Source: Author based on PHS 2001/02 dataset and CSO, 2002.

Table A5c: Number of Female Headed Households by Marital Status of Household Head during the 2001/2002 Agriculture Season by Province and District

Province	District	Marital Status: Single	Monogamously Married	Polygamously Married	Divorced	Widowed	Seperated	Total
Eastern	Chadiza	317	954	712	173	1,701	-	3,857
	Chama	151	1,576	630	333	1,559	297	4,546
	Chipata	1,466	1 97	789	1,395	4,863	536	11,019
	Katete	639	2,552	1,897	2,843	2,883	243	11,057
	Lundazi	92	1,757	2,636	134	2,969	-	7,588
	Mambwe	-	-	187	545	369	-	1,101
	Nyimba	212	196	271	1,088	468	-	2,235
	Petauke	435	2,325	152	970	5,725	512	10,119
Province Total		**3,312**	**9,360**	**7,274**	**7,481**	**20,537**	**1,588**	**51,522**

Source: Author based on PHS 2001/02 dataset and CSO, 2002.

Table A5d: Number of Male Heads of Households by Age Group by Province and District - 2001/2002 Agricultural season

Province	District	Age Group											Total Male Heads
		15 to 19	20 to 24	25 to 29	30 to 34	35 to 39	40 to 44	45 to 49	50 to 54	55 to 59	60 to 64	65 to 90	
Eastern	Chadiza	-	653	1.958	2.203	925	2.025	326	345	486	263	1.2	10.384
	Chama	-	-	1.633	1.151	594	982	483	1.002	-	602	1.425	7.872
	Chipata	-	647	4.631	5.005	6.46	3.681	2.715	1.83	1.447	1.881	5.058	33.355
	Katete	-	1.688	4.185	3.732	2.169	2.819	2.222	1.531	502	815	1.684	21.347
	Lundazi	-	660	2.313	4.565	3.693	2.43	1.793	1.452	805	735	3.486	21.932
	Mambwe	-	-	1.428	1.234	913	1.059	895	387	901	375	950	8.142
	Nyimba	-	212	1.752	2.098	739	544	784	547	424	333	1.316	8.749
	Petauke	-	476	1.791	4.073	2.733	5.506	2.734	2.044	1.969	1.139	2.807	25.272
Province Total		**-**	**4.336**	**19.691**	**24.061**	**18.226**	**19.046**	**11.952**	**9.138**	**6.534**	**6.143**	**17.926**	**137.053**

Source: Author based on PHS 2001/02 dataset and CSO, 2002.

Table A5e: Distribution of Households by Agricultural Activity by Province and District – 2001/2002 Agriculture Season

Province	District	Total Number of Rural Households	Crop Growing Households	Livestock Raising Households	Poultry Raising Households
Eastern	Chadiza	14.241	13.914	8.965	10.524
	Chama	12.417	12.417	3.716	6.934
	Chipata	44.373	43.689	18.034	28.528
	Katete	32.405	32.405	18.666	24.691
	Lundazi	29.521	29.338	8.116	18.944
	Mambwe	9.245	9.244	2.886	6.079
	Nyimba	10.986	10.502	7.255	8.974
	Petauke	35.391	35.051	22.758	28.408
Province Total		**188.579**	**186.560**	**90.396**	**133.082**

Source: Author based on PHS 2001/02 dataset and CSO, 2002.

Table A5f: Seed cotton: Number of Households Reporting, Area planted, Production and Sales by Province and District - 2001/2002 Agriculture Season

Province	District	HHs Reporting	Total Hectares planted	Quantity Harvested (metric tons)	Quantity sold (metric tons)	Yield (MT/Ha)
Eastern	Chadiza	3.71	3.423	2.419	2.419	0.71
	Chama	3.831	1.606	1.392	1.334	0.87
	Chipata	15.587	14.04	11.649	10.785	0.83
	Katete	16.299	12.01	11.885	11.831	0.99
	Lundazi	8.228	6.409	4.064	3.967	0.63
	Mambwe	5.92	4.483	4.974	4.974	1.11
	Nyimba	2.177	2.109	2.095	2.095	0.99
	Petauke	4.377	4.092	3.467	3.433	0.85
Province Total		**60.129**	**48.172**	**41.945**	**40.838**	**0.87**

Source: Author based on PHS 2001/02 dataset and CSO, 2002.

Table A5g: Seed cotton: Number of Households Applying Fertilizer and Lime, Quantity applied by province and district - 2001/2002 Agriculture Season

Province	District	BASAL FERTILIZER Number of Households	BASAL FERTILIZER Quantity (kg)	TOP FERTILIZER Number of Households	TOP FERTILIZER Quantity (kg)
Central	Chibombo	16	1.552	-	
	Kabwe Urban	13	651	-	
	Mumbwa	118	5.916	-	
Province Total		**147**	**8.119**	**-**	
Eastern	Chipata	108	2.704	108	2.704
	Katete	100	4.982	188	5.07
Province Total		**208**	**7.685**	**297**	**7.774**
Lusaka	Kafue	13	636	13	636
Province Total		**13**	**636**	**13**	**636**
Southern	Mazabuka	149	29.749	149	29.749
	Monze	11	1.676	11	1.676
Province Total		**160**	**31.425**	**160**	**31.425**
Zambia Total		**527**	**47.866**	**469**	**39.836**

Source: Author based on PHS 2001/02 dataset and CSO, 2002.

Table A5h.1: Percentage of Farmers Growing Cotton, Eastern Province, 1997 – 2002

District	1996/97	1997/98	1998/99	1999/2000	2000/2001	2001/2002
Chadiza (301)	44.79%	27.27%	11.24%	10.00%	26.14%	27.00%
Chipata (303)	40.92%	33.90%	34.54%	25.74%	36.16%	36.97%
Katete (304)	50.51%	53.03%	35.18%	39.55%	52.17%	52.36%
Lundazi (305)	24.11%	25.78%	43.23%	11.92%	27.04%	24.52%
Petauke (308)	24.34%	16.79%	20.30%	8.44%	21.37%	21.31%
Total Catchment Districts	35.48%	30.99%	31.04%	19.55%	32.50%	32.20%
Chama (302)	18.92%	33.33%	30.26%	11.25%	10.00%	16.88%
Mambwe (306)	59.62%	61.82%	50.00%	49.15%	54.90%	54.24%
Nyimba (307)	6.25%	8.11%	11.32%	6.67%	22.22%	23.73%
Total Control Districts	29.93%	38.28%	28.22%	21.11%	26.86%	30.26%
Total	34.86%	31.75%	30.68%	19.76%	31.71%	31.93%

Source: Author based on PHS dataset.

Table A5h.2: Number of Farmers Growing Cotton, 1997 – 2002

District	1996/97	1997/98	1998/99	1999/2000	2000/2001	2001/2002
Chadiza (301)	43	24	10	10	23	27
Chipata (303)	124	100	105	87	111	122
Katete (304)	100	105	70	87	96	111
Lundazi (305)	54	58	99	31	63	64
Petauke (308)	65	44	55	27	56	65
Total Catchment Districts	386	331	339	242	349	389
Chama (302)	7	12	23	9	7	13
Mambwe (306)	31	34	17	29	28	32
Nyimba (307)	3	3	6	4	12	14
Total Control Districts	41	49	46	42	47	59
Total Eastern Province	427	380	385	284	396	448
Observations	1225	1197	1255	1437	1249	1403

Source: Author based on PHS dataset.

Table A6a: Descriptive Statistics of Dependent Variable: Log Cotton Productivity

Treatment	Period	Variable	Obs	Mean	Std.Dev	Min	Max
Treatment	1996/97-2001/02	logyield	1238	6,650	0,995	2,286	10,697
Control	1996/97-2001/02	logyield	925	6,726	0,993	0,811	10,217
Treatment	1998/99-2001/02	logyield	1238	6,650	0,995	2,286	10,697
Control	1998/99-2001/02	logyield	196	6,803	0,845	2,773	8,490
Control	1996/97-1997/98	logyield	729	6,705	1,029	0,811	10,217

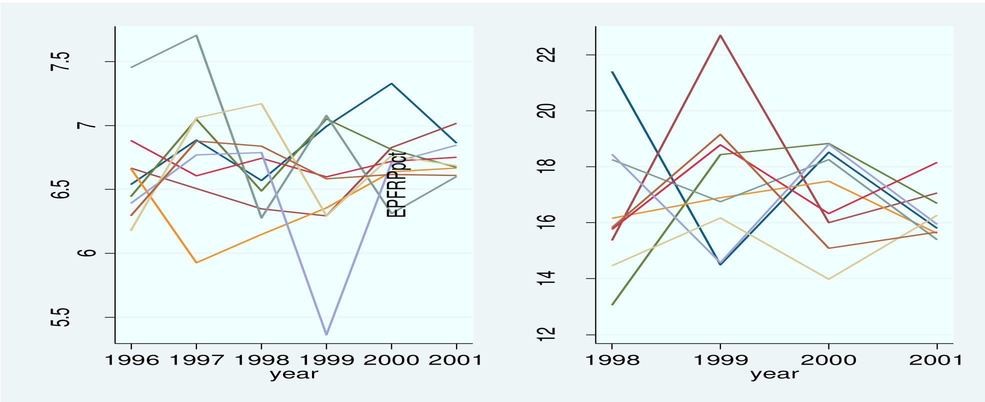
Source: Authors' calculations.

Table A6b: Descriptive Statistics of Dependent Variable: Cotton Productivity (kgs)

Treatment	Period	Variable	Obs	Mean	Std.Dev	Min	Max
Treatment	1996/97-2001/02	productivity	1238	1278,23	2252,89	9,84	44212,00
Control	1996/97-2001/02	productivity	925	1380,41	2051,91	2,25	27372,97
Treatment	1998/99-2001/02	productivity	1238	1278,23	2252,89	9,84	44212,00
Control	1998/99-2001/02	productivity	196	1174,38	792,85	16,00	4864,87
Control	1996/97-1997/98	productivity	729	1435,80	2271,79	2,25	27372,97

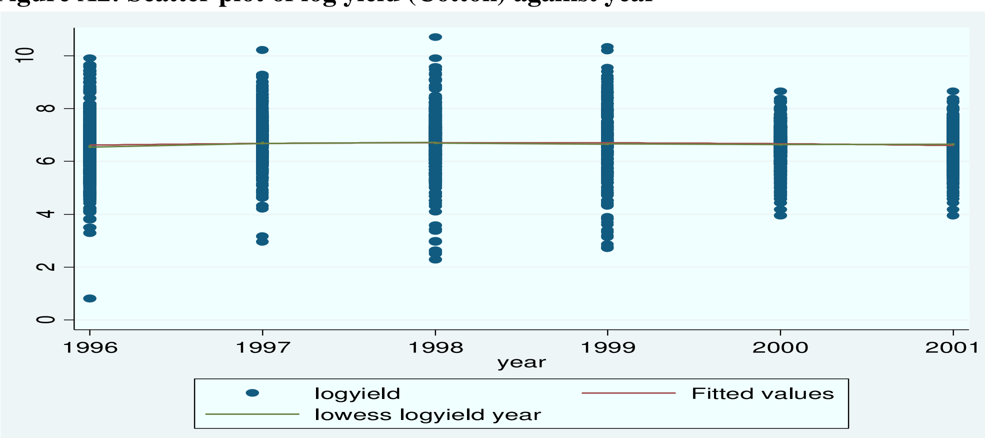
Source: Authors' calculations.

Figure A1: Time-series plot of log yield against year and EPFRPpct against year for ages 20-30



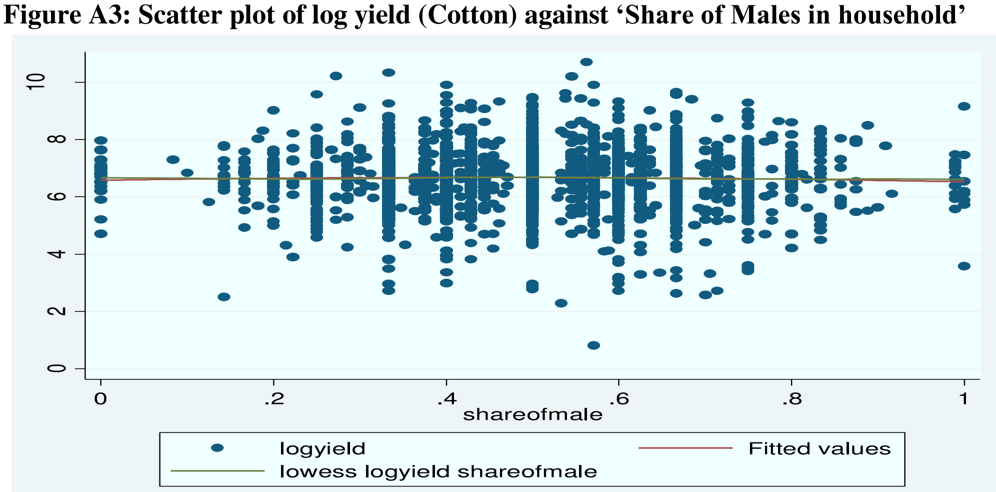
Source: Authors' estimations based on the Post Harvest Survey.

Figure A2: Scatter plot of log yield (Cotton) against year



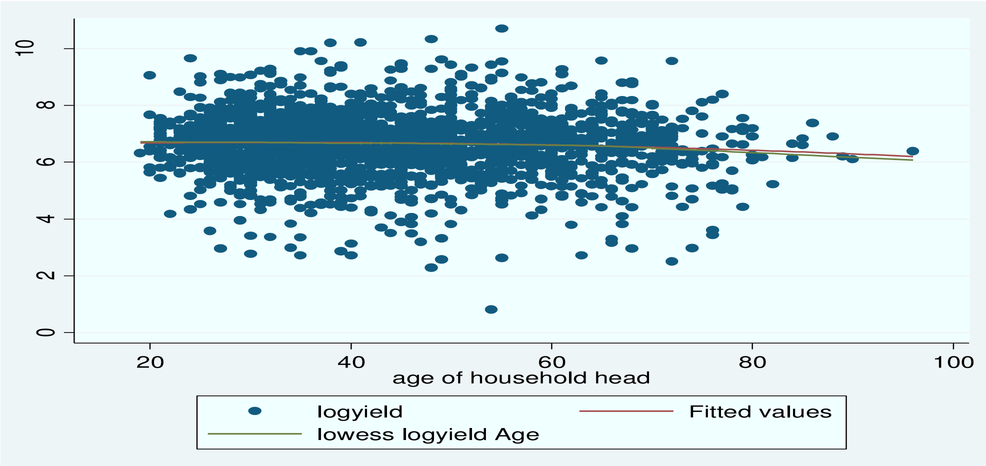
Source: Authors' estimations based on the Post Harvest Survey.

Figure A3: Scatter plot of log yield (Cotton) against 'Share of Males in household'



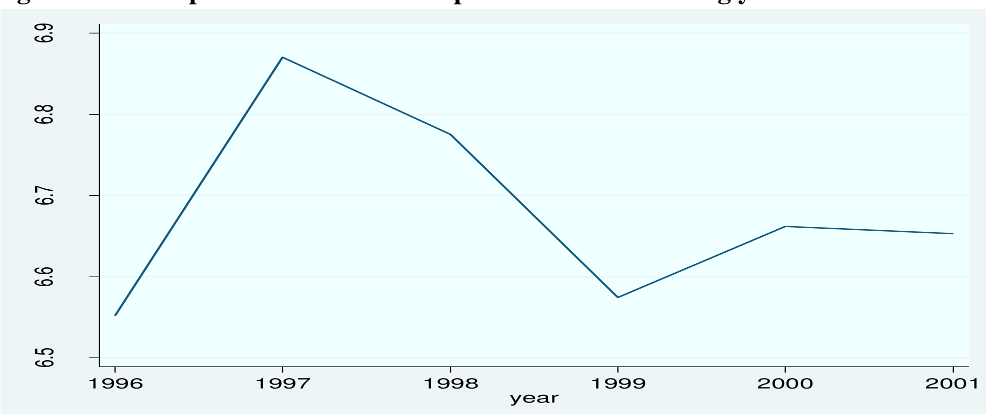
Source: Authors' estimations based on the Post Harvest Survey.

Figure A4: Scatter plot of log yield (Cotton) against 'Age of Household Head'



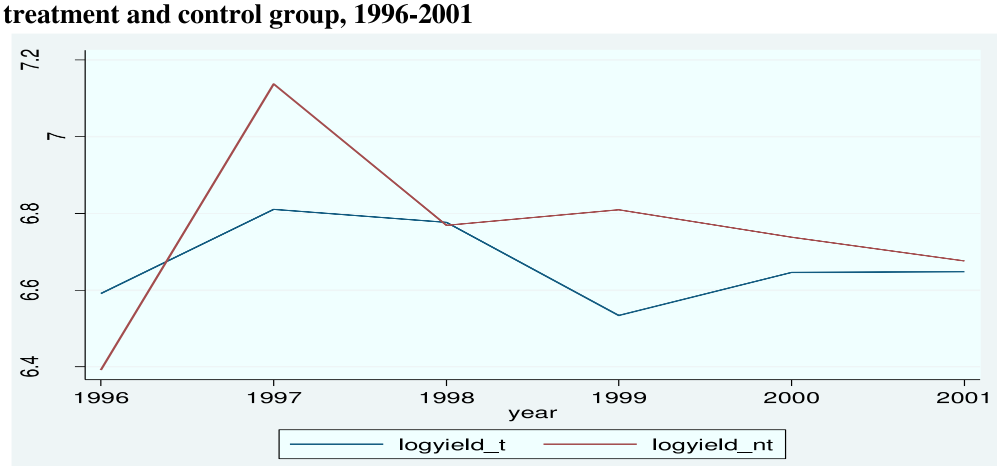
Source: Authors' estimations based on the Post Harvest Survey.

Figure A5a: Graphs of Panel data collapsed to time series: Log yield Cotton



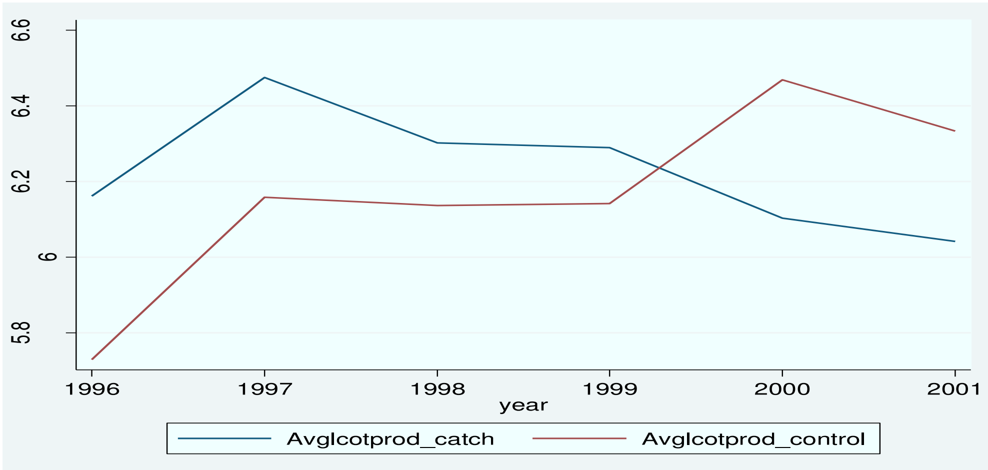
Source: Authors' estimations based on the Post Harvest Survey.

Figure A5b: Graphs of Panel data collapsed to time series: 'Trends' in the Log of Cotton Yield for treatment and control group, 1996-2001



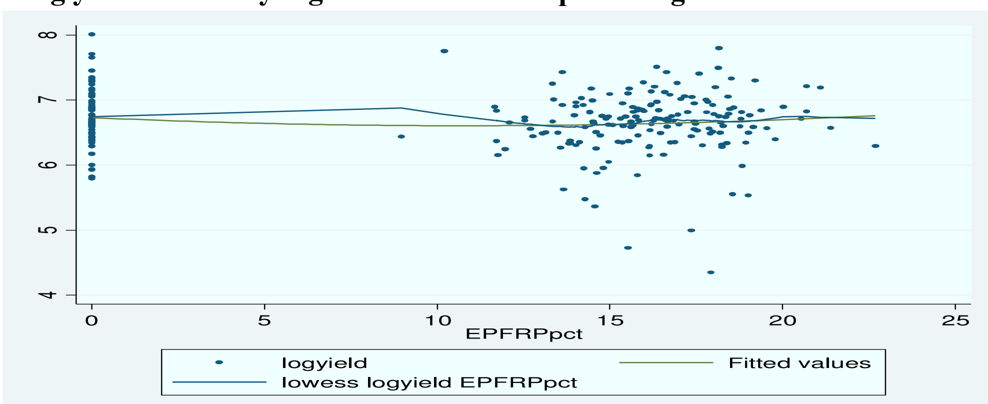
Source: Authors' estimations based on the Post Harvest Survey.

Figure A5c: Graphs of Panel data collapsed to time series: 'Trends' in the Log of Cotton Prod for treatment and control group, 1996-2001



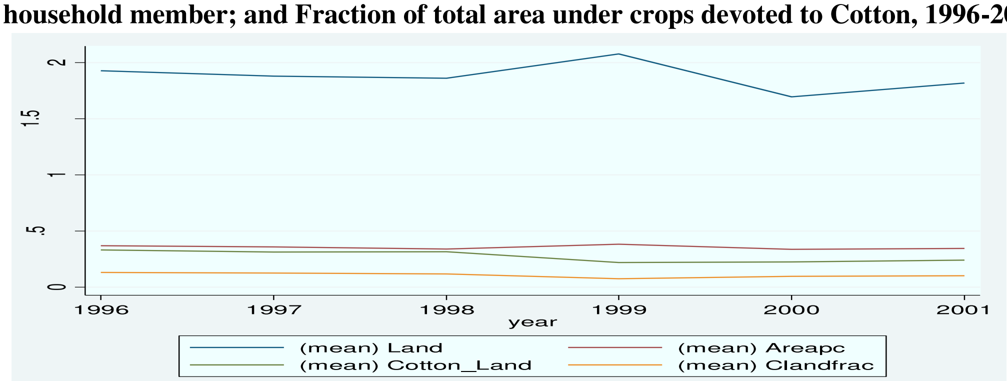
Source: Authors' estimations based on the Post Harvest Survey.

Figure A6: Fitted Quadratic Regression and Lowess Regression Curves added to Scatterplot of log yield on the key regressor – EPFRP's percentage share of total road network



Source: Authors' estimations based on the Post Harvest Survey.

Figure A7: Total area under crops (Ha); Size of Land allocated to Cotton; Cultivated land per household member; and Fraction of total area under crops devoted to Cotton, 1996-2001



Source: Authors' estimations based on the Post Harvest Survey.

Table A7.1: Descriptive Statistics: Treatment and Control districts, 1996/97 – 1998/1999

			1996/1997						1997/1998						1998/1999					
			Full Sample		Catchment Districts		Control Districts		Full Sample		Catchment Districts		Control Districts		Full Sample		Catchment Districts		Control Districts	
Type	Variable	Variable	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
CV	Dependent variable	Volume of cotton production per hectare produced (MT)	1.33	2.31	1.36	2.33	1.19	2.23	1.48	2.09	1.42	1.73	1.78	3.26	1.62	3.06	1.67	3.24	1.35	1.42
CV		Log of cotton output (in kg) per hectare	6.54	1.10	6.58	1.07	6.39	1.20	6.83	0.96	6.77	0.99	7.11	0.74	6.75	1.17	6.77	1.14	6.63	1.36
CV	Household determinants	Age of the household head	46.7	15.0	47.0	15.0	45.1	15.3	44.4	15.2	44.7	15.3	42.7	14.4	45.5	15.3	45.6	15.2	45.1	15.7
CV		Age Square of the household head	2404.0	1506.1	2428.6	1513.9	2265.3	1457.9	2205.4	1537.9	2236.0	1563.2	2031.1	1376.1	2307.8	1535.0	2313.1	1524.2	2280.0	1594.9
CV		Size of the household	5.8	3.2	5.9	3.3	5.6	2.9	5.7	3.0	5.7	3.0	5.6	2.9	5.94	3.20	5.91	3.20	6.06	3.19
CV		Log of Size of the household	1.61	0.59	1.61	0.60	1.58	0.56	1.59	0.56	1.59	0.57	1.58	0.56	1.63	0.59	1.62	0.60	1.67	0.53
DV		Household category (stratum)	1.22	0.41	1.23	0.42	1.16	0.37	1.12	0.33	1.13	0.33	1.09	0.29	1.11	0.32	1.12	0.33	1.07	0.26
CV		Number of males in household	2.79	1.82	2.83	1.86	2.58	1.55	2.74	1.85	2.75	1.87	2.72	1.74	2.94	1.99	2.94	2.01	2.97	1.87
CV	Household demographics	Number of females in household	3.03	1.97	3.03	1.95	3.03	2.08	2.91	1.79	2.92	1.78	2.87	1.83	2.99	1.85	2.98	1.82	3.09	1.98
DV		Sex of head of household	1.23	0.42	1.23	0.42	1.24	0.43	1.23	0.42	1.23	0.42	1.23	0.42	1.24	0.43	1.24	0.43	1.25	0.43
CV	Input use	Basal Quantity used (kg)	29.93	123.90	29.38	126.05	33.02	111.38	30.88	121.42	32.45	128.43	21.93	68.71	39.63	145.91	41.72	154.57	28.57	85.88
CV		Topdressing Quantity used (kg)	27.18	104.57	28.03	110.03	22.42	66.14	30.50	122.82	33.33	131.37	14.39	49.50	38.71	127.18	41.01	134.87	26.51	73.02
CV		Basal Fertilizers Used per cultiv. Area (kg per ha)	11.53	36.76	10.82	34.42	15.51	47.73	13.05	38.21	14.03	40.50	7.52	20.08	16.56	42.32	16.97	42.01	14.43	43.93
CV		Top Dressing Fertilizers Used per cultiv. Area (kg per ha)	10.43	28.63	10.37	28.27	10.72	30.63	13.32	41.86	14.68	44.78	5.60	15.83	16.74	38.74	17.67	39.56	11.85	33.79
CV		Value of Basal quantity used - (ZMK)	31920.3	92680.6	31978.9	94976.5	31653.9	81845.4	22202.3	238505.6	24114.1	257993.3	11329.6	42226.9	24409.9	87229.3	24982.4	89283.8	21371.9	75525.5
CV		Value of Topdressing quantity used - (ZMK)	27770.4	80701.8	29373.2	86486.7	20480.8	45590.6	23052.7	241685.4	25780.0	261674.1	7541.9	31132.9	25208.1	89689.2	26428.6	94747.7	18731.2	55349.4
CV		Expenditure on Basal fertilizers per cultivated area (ZMK/Ha)	12152.0	26389.7	11592.4	23957.6	14697.3	35444.5	7133.8	26384.2	7785.0	28226.1	3433.0	10465.8	10491.0	28902.8	10591.7	27646.4	9962.2	34842.7
CV		Expenditure on Topdressing fertilizers per cultivated area (ZMK/Ha)	10284.9	19167.9	10500.7	19129.1	9303.2	19405.7	8317.9	42566.3	9290.5	45947.1	2789.8	8789.9	10934.1	26732.4	11487.9	27268.4	8025.1	23568.2
DV		Any chemical fertilizers used in season	0.183	0.386	0.183	0.387	0.179	0.385												
CV	Assets	Number of ploughs	0.374	0.865	0.389	0.895	0.293	0.670	0.29	0.77	0.30	0.81	0.21	0.53	0.30	0.77	0.31	0.79	0.23	0.63
CV		Number of draught animals	0.649	1.741	0.674	1.809	0.505	1.289	0.54	1.45	0.55	1.49	0.47	1.19	0.57	1.55	0.60	1.61	0.44	1.16
CV		Number of ploughs per household member	0.062	0.159	0.065	0.168	0.042	0.100	0.05	0.13	0.05	0.14	0.03	0.09	0.05	0.13	0.05	0.13	0.04	0.12
CV		Number of draught animals per household members	0.099	0.260	0.103	0.269	0.073	0.203	0.09	0.25	0.09	0.25	0.08	0.21	0.09	0.27	0.10	0.27	0.07	0.25
CV		Size of the land allocated to cotton	0.13	0.21	0.12	0.20	0.17	0.23	0.12	0.21	0.117	0.203	0.143	0.218	0.11	0.20	0.11	0.20	0.11	0.20
CV		Total area under crops (ha)	1.97	1.77	2.03	1.82	1.65	1.40	1.86	1.74	1.901	1.790	1.652	1.421	1.87	1.96	1.92	2.01	1.57	1.62
CV		Cultivated land per household member (ha)	0.38	0.33	0.39	0.34	0.32	0.24	0.37	0.33	0.370	0.331	0.346	0.302	0.35	0.32	0.36	0.32	0.30	0.31
DV		Livestock raising	0.58	0.49	0.61	0.49	0.42	0.49	0.48	0.50	0.480	0.500	0.458	0.500	0.48	0.50	0.50	0.50	0.34	0.48
DV		Usage of animal draught power for land preparation	0.27	0.45	0.28	0.45	0.22	0.42	0.25	0.43	0.251	0.434	0.249	0.433	0.24	0.43	0.25	0.43	0.21	0.41
DV		Received agricultural loan	0.323	0.468	0.308	0.462	0.413	0.494	0.265	0.441	0.241	0.428	0.402	0.492	0.32	0.47	0.32	0.47	0.29	0.46
DV		Bank is too far away to apply for agricultural loan	0.053	0.225	0.056	0.230	0.033	0.178	0.081	0.273	0.075	0.263	0.117	0.323	0.05	0.22	0.05	0.22	0.06	0.23
	Districts effects (District Dummies)	Market access																		
		Local infrastructure																		
		Local knowledge																		
		Access to credit																		
FE	EPFRP	*Rural transport infrastructure dummy (EPFRP)*	n.a.	n.a.					n.a.	n.a.					0.84	0.37				
DV	Agricultural extension services	Information on marketing for agricultural products	0.46	0.50	0.45	0.50	0.53	0.50	0.39	0.49	0.37	0.48	0.47	0.50	0.33	0.47	0.33	0.47	0.31	0.46
DV		Use any of the advice received on Crop husbandry	0.28	0.45	0.29	0.45	0.23	0.42	0.20	0.40	0.20	0.40	0.18	0.38	0.20	0.40	0.20	0.40	0.17	0.37
DV		Use any of the advice received on Crop diversification	0.23	0.42	0.25	0.43	0.17	0.38	0.12	0.32	0.11	0.31	0.15	0.36	0.16	0.37	0.16	0.37	0.15	0.36
DV		Information on agricultural input supply	0.41	0.49	0.39	0.49	0.51	0.50	0.35	0.48	0.33	0.47	0.42	0.50	0.32	0.47	0.32	0.47	0.30	0.46
DV	Geographic Variables	Proportion of sample in Catchment Areas	0.85	0.36	1.00	0.00	0.00	0.00	0.85	0.36	1.00	0.00	0.00	0.00	0.84	0.37	1.00	0.00	0.00	0.00
DV		Proportion of sample in Control Areas	0.15	0.36	0.00	0.00	1.00	0.00	0.15	0.36	0.00	0.00	1.00	0.00	0.16	0.37	0.00	0.00	1.00	0.00
DV		Proportion of sample which shares a border with Mozambique	0.47	0.50	0.51	0.50	0.28	0.45	0.48	0.50	0.51	0.50	0.31	0.46	0.45	0.50	0.50	0.50	0.17	0.38
DV		Proportion of sample which shares a border with Malawi	0.50	0.50	0.49	0.50	0.52	0.50	0.49	0.50	0.49	0.50	0.49	0.50	0.49	0.50	0.50	0.50	0.45	0.50
CV		Distance to the nearest all-weather road	1.374	0.603	1.412	0.605	1.163	0.549	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
CV		Distance to the nearest input market	1.855	0.784	1.927	0.788	1.451	0.625	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
		Mode of transport	5.660	2.235	5.675	2.193	5.549	2.471	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
		Most important investment community want to invest in.	4.314	1.477	4.449	1.374	3.554	1.779	4.685	2.345	4.678	2.294	4.726	2.622	5.143	2.527	5.247	2.492	4.588	2.642
DV		Feeder Roads Community Investment priority	0.076	0.266	0.061	0.239	0.163	0.370	0.113	0.316	0.094	0.292	0.218	0.414	0.131	0.338	0.117	0.322	0.206	0.405
CV		*Rainfall*	831.5	122.9					716.0	81.4					788.2	148.1				
	Cotton Observations		421		341		80		378		310		68		388		332		51	
	Total number of Observations		1219		1035		184		1197		1018		179		1255		1056		199	

Source: Authors' estimations based on PHS.

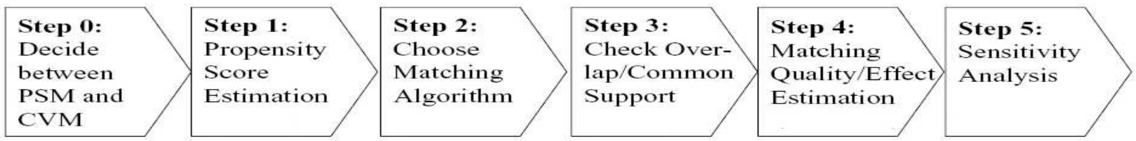
Table A7.2: Descriptive Statistics: Treatment and Control districts, 1999/2000 – 2001/2002

			1999/2000						2000/2001						2001/2002					
			Full Sample		Catchment Districts		Control Districts		Full Sample		Catchment Districts		Control Districts		Full Sample		Catchment Districts		Control Districts	
Type	Variable	Variable	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
CV	Dependent variable	Volume of cotton production per hectare produced (MT)	1.64	3.02	1.66	3.23	1.49	1.44	0.97	0.68	0.95	0.64	1.03	0.66	0.97	0.68	0.962	0.688	0.982	0.624
CV		Log of cotton output (in kg) per hectare	6.55	1.40	6.51	1.42	6.74	1.26	6.65	0.71	6.64	0.70	6.72	0.70	6.64	0.71	6.635	0.723	6.683	0.679
CV	Household determinants	Age of the household head	43.0	14.3	43.3	14.2	41.6	14.5	45.7	14.7	45.9	14.5	45.1	15.7	45.3	14.7	45.38	14.47	45.12	15.69
CV		Age Square of the household head	2056.0	1371.5	2078.6	1373.2	1940.5	1360.0	2309.7	1465.6	2315.4	1443.1	2281.6	1575.8	2270.4	1459.2	2268.21	1437.38	2280.91	1565.59
CV		Size of the household	6.17	3.43	6.29	3.54	5.56	2.75	5.97	2.95	6.01	2.98	5.76	2.77	6.34	2.93	6.35	2.94	6.28	2.92
CV		Log of Size of the household	1.67	0.56	1.69	0.57	1.59	0.54	1.66	0.54	1.66	0.55	1.63	0.53	1.73	0.50	1.74	0.50	1.72	0.51
DV		Household category (stratum)	1.14	0.35	1.16	0.37	1.05	0.21	1.14	0.35	1.17	0.38	1.05	0.21	1.13	0.33	1.14	0.35	1.04	0.20
CV		Number of males in household	3.08	2.24	3.15	2.34	2.71	1.64	2.98	1.93	3.00	1.96	2.90	1.77	3.18	1.86	3.18	1.88	3.17	1.78
CV	Household demographics	Number of females in household	3.09	1.92	3.14	1.93	2.86	1.86	2.98	1.73	3.01	1.73	2.86	1.74	3.16	1.81	3.17	1.79	3.11	1.89
DV		Sex of head of household	1.24	0.43	1.23	0.42	1.28	0.45	1.25	0.43	1.24	0.43	1.28	0.45	1.25	0.44	1.25	0.43	1.27	0.44
CV	Input use	Basal Quantity used (kg)	47.77	129.59	52.47	138.69	23.82	60.76	32.81	149.51	37.07	168.78	18.66	67.84	34.79	149.91	38.12	161.34	18.33	67.68
CV		Topdressing Quantity used (kg)	45.80	118.37	50.26	126.42	23.08	58.01	31.98	145.77	36.28	164.78	17.80	65.80	33.69	147.14	36.99	158.41	17.35	65.84
CV		Basal Fertilizers Used per cultiv. Area (kg per ha)	22.00	53.14	23.76	56.34	13.01	30.81	17.17	50.91	17.98	48.69	9.90	27.01	16.09	41.98	17.66	44.83	8.60	22.69
CV		Top Dressing Fertilizers Used per cultiv. Area (kg per ha)	21.01	47.71	22.72	50.60	12.32	27.32	16.10	40.45	17.42	40.70	10.90	33.87	15.56	37.37	16.88	38.67	9.24	29.64
CV		Value of Basal quantity used - (ZMK)	34564.7	95575.4	37868.0	102145.6	17723.5	46768.1	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
CV		Value of Topdressing quantity used - (ZMK)	33167.1	86535.9	36359.4	92354.9	16891.9	43323.7	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
CV		Expenditure on Basal fertilizers per cultivated area (ZMK/Ha)	15823.0	38979.2	17035.4	41278.3	9657.0	23206.0	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
CV		Expenditure on Topdressing fertilizers per cultivated area (ZMK/Ha)	15274.7	35724.6	16517.8	37887.3	8952.7	20524.3	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
DV		Any chemical fertilizers used in season																		
CV	Assets	Number of ploughs	0.27	0.65	0.30	0.69	0.12	0.37	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
CV		Number of draught animals	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
CV		Number of ploughs per household member	0.04	0.11	0.05	0.12	0.02	0.09	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
CV		Number of draught animals per household members	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
CV		Size of the land allocated to cotton	0.07	0.16	0.07	0.16	0.07	0.16	0.10	0.19	0.10	0.18	0.09	0.18	0.10	0.18	0.09	0.18	0.11	0.20
CV		Total area under crops (ha)	2.10	2.06	2.22	2.17	1.51	1.21	1.73	1.65	1.77	1.64	1.38	1.25	1.83	1.74	1.90	1.81	1.48	1.27
CV		Cultivated land per household member (ha)	0.39	0.45	0.40	0.47	0.33	0.33	0.34	0.37	0.35	0.38	0.30	0.36	0.36	0.31	0.37	0.33	0.28	0.20
DV		Livestock raising	0.50	0.50	0.52	0.50	0.39	0.49	0.55	0.50	0.58	0.49	0.44	0.50	0.47	0.50	0.49	0.50	0.37	0.48
DV		Usage of animal draught power for land preparation	0.28	0.45	0.30	0.46	0.21	0.41	0.35	0.48	0.38	0.48	0.21	0.41	0.35	0.48	0.38	0.49	0.21	0.41
DV		Received agricultural loan	0.16	0.37	0.16	0.37	0.15	0.36	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
DV		Bank is too far away to apply for agricultural loan	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
	Districts effects (District Dummies)	Market access																		
		Local infrastructure																		
		Local knowledge																		
		Access to credit																		
FE	EPFRP	*Rural transport infrastructure dummy (EPFRP)*	0.83	0.37					0.83	0.37					0.83	0.37				
DV	Agricultural extension services	Information on marketing for agricultural products	0.30	0.46	0.315	0.465	0.252	0.435	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
DV		Use any of the advice received on Crop husbandry	0.01	0.10	0.010	0.101	0.011	0.106	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
DV		Use any of the advice received on Crop diversification	0.14	0.35	0.137	0.344	0.158	0.366	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
DV		Information on agricultural input supply	0.23	0.42	0.234	0.423	0.205	0.405	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
DV	Geographic Variables	Proportion of sample in Catchment Areas	0.84	0.37	1.000	0.000	0.000	0.000	0.83	0.37	1.000	0.000	0.000	0.000	0.83	0.37	1.00	0.00	0.00	0.00
DV		Proportion of sample in Control Areas	0.16	0.37	0.000	0.000	1.000	0.000	0.17	0.37	0.000	0.000	1.000	0.000	0.17	0.37	0.00	0.00	1.00	0.00
DV		Proportion of sample which shares a border with Mozambique	0.45	0.50	0.498	0.500	0.231	0.422	0.46	0.50	0.506	0.500	0.251	0.435	0.46	0.50	0.51	0.50	0.25	0.43
DV		Proportion of sample which shares a border with Malawi	0.49	0.50	0.502	0.500	0.427	0.496	0.48	0.50	0.494	0.500	0.421	0.495	0.48	0.50	0.49	0.50	0.42	0.50
CV		Distance to the nearest all-weather road	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
CV		Distance to the nearest input market	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
		Mode of transport	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
		Most important investment community want to invest in.	5.026	2.359	4.982	2.333	5.248	2.481	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
DV		Feeder Roads Community Investment priority	0.142	0.349	0.137	0.344	0.162	0.370	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
CV		*Rainfall*	667.1	93.8					980.1	203.6					723.7	89.4				
	Cotton Observations		279		236		43		467		345		67		492		403		89	
	Total number of Observations		1427		1193		234		1249		1040		209		1403		1162		236	

Source: Authors' estimations based on PHS.

Implementation of PS matching (PSM)

Implementation of PSM requires the answer to a lot of questions. The following Figure (from Caliendo and Kopeinig, 2008) summarizes the necessary steps when implementing propensity score matching (Grilli and Rampichini, 2011:36):



CVM: Covariate Matching, PSM: Propensity Score Matching

Table A8.1: The treatment is EPFRP

EPFRP	Freq.	Percent	Cum.
0	2,755	40.37	40.37
1	4,070	59.63	100.00
Total	6.825	100.00	

Definitions	Unit
EPFRP dummy	DV
EPFRP percentage of length of feeder roads	%
Road Density	Road length/ 100 km2
	Road length / 1000 population (2000)

Step 1a: Algorithm to estimate the propensity score

Table A8.2

EPFRP	Age	Agesq	Sex	loghhsize	stratum	livestock	Areapc	Clandfrac	rain_EP	_cons
Coef.	-.0300551	.0002741	.0936159	.4819535	-.2823135	.0119289	.687302	-.9438211	.0017526	-1.008996
Std.Err.	.0176093	.0002009	.0641978	.0606429	.0866577	.0556247	.102049	.1348815	.0001645	.381528
z	-1.71	1.36	1.46	7.95	-3.26	0.21	6.74	-7.00	10.66	-2.64
P>z	0.088	0.172	0.145	0.000	0.001	0.830	0.000	0.000	0.000	0.008
[95% Conf.	-.0645688	-.0001196	-.0322095	.3630956	-.4521595	-.0970935	.4872896	-1.208184	.0014302	-1.756777
Interval]	.0044586	.0006679	.2194414	.6008115	-.1124675	.1209514	.8873143	-.6794583	.0020749	-.2612149
Number of obs	=	6586								
LR chi2(9)	=	235.37								
Prob > chi2	=	0.0000								
Pseudo R2	=	0.0264								
Log likelihood	=	-4345.76								

Note: the common support option has been selected

The region of common support is [.29481562, .99026378]

Table A8.3: Estimated propensity score

	Percentiles	Smallest		
1%	.3701691	.2948156		
5%	.4367709	.2958079		
10%	.4696934	.2960308	Obs	6585
25%	.5259147	.3009192	Sum of Wgt.	6585
50%	.5881604		Mean	.5876557
		Largest	Std. Dev.	.0919696
75%	.6484955	.9129527		
90%	.7053895	.9184439	Variance	.0084584
95%	.7422879	.9417921	Skewness	-.0004421
99%	.7905877	.9902638	Kurtosis	3.031821

Step Ia: Identification of the optimal number of blocks
The final number of blocks is 8. This number of blocks ensures that the mean propensity score is not different for treated and controls in each blocks.

Step IIa: Test of balancing property of the propensity score
The balancing property is not satisfied so we try a different specification of the propensity score.

Table A9.1: Specification of the propensity score

Block	Inferior of block of pscore	EPFRP 0	EPFRP 1	Total
2	.2	72	76	148
3	.4	1.685	1.811	3.496
4	.6	826	1.382	2.208
5	.7	78	187	265
6	.725	31	162	193
7	.75	20	207	227
8	.8	3	45	48
	Total	**2715**	**3870**	**6585**

Test in Block	Observations in Block: Total	Control	Treated
1	0	0	0
2	148	72	76
3	3496	1685	1811
4	2983	955	1938
4	2208	826	1382
5	685	129	556
5	485	109	349
5	265	78	187
6	193	31	162
7	227	20	207
8	48	3	45

We test that the mean propensity score is not different for treated and controls

Table A9.2: Test for block 2: Two-sample t test with equal variances

Group	Obs	Mean	Std.Err.	Std.Dev.	[95% Conf.	Interval]
0	72	.3666104	.0029989	.0254463	.3606309	.37259
1	76	.3637298	.0031321	.0273053	.3574902	.3699693
combined	148	.3651312	.0021674	.0263672	.3608479	.3694144
diff		.0028807	.0043446		-.0057058	.0114672
diff = mean(0) -	mean(1)				t =	0.6630
Ho: diff = 0			degrees	of	freedom =	146
Ha: diff < 0		Ha:	diff	!= 0	Ha: diff	> 0
Pr(T < t) = 0.7458	Pr(T >t)	= 0.5083			Pr(T > t) =	0.2542

The mean propensity score is not different for treated and controls in block 2.

Table A9.3: Test for block 3: Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	1685	.5269422	.0011822	.0485266	.5246235	.5292608
1	1811	.5304475	.0011512	.0489894	.5281897	.5327053
combined	3496	.528758	.0008252	.0487913	.5271401	.5303759
diff		-.0035054	.0016506		-.0067417	-.000269
diff = mean(0) -	mean(1)					t = -2.1236
Ho: diff = 0			degrees	of	freedom =	3494
Ha: diff < 0		Ha:	diff	!= 0		Ha: diff > 0
Pr(T < t) = 0.0169	Pr(T >t)	= 0.0338			Pr(T > t) = 0.9831	

The mean propensity score is not different for treated and controls in block 3.

The mean propensity score is different for treated and controls in block 4.

Table A9.4

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	955	.6536781	.0012743	.0393809	.6511773	.656179
1	1938	.6720885	.0011758	.051763	.6697825	.6743945
combined	2893	.6660111	.0009072	.0487973	.6642322	.66779
diff		-.0184103	.001899		-.0221338	-.0146869
diff = mean(0) -	mean(1)					t = -9.6949
Ho: diff = 0			degrees	of	freedom =	2891
Ha: diff < 0		Ha:	diff	!= 0		Ha: diff > 0
Pr(T < t) = 0.0000	Pr(T >t)	= 0.0000			Pr(T > t) = 1.0000	

Hence we **Split the block 4** and retest and check that blocks have shifted.

Table A9.5

Blocks of the pscore for treatment EPFRP	EPFRP 0	1	Total
2	72	76	148
3	1.685	1.811	3.496
4	955	1.938	2.893
6	3	45	48
Total	2.715	3.87	6.585

Table A9.6: Test for block 4: Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	826	.6426124	.0009914	.028492	.6406665	.6445583
1	1382	.6442874	.0007447	.0276831	.6428266	.6457482
combined	2208	.6436608	.0005957	.0279938	.6424925	.6448291
diff		-.001675	.0012309		-.0040889	.0007389
diff = mean(0) -	mean(1)				t = -1.3608	
Ho: diff = 0			degrees	of	freedom =	2206
Ha: diff < 0		Ha:	diff	!= 0		Ha: diff > 0
Pr(T < t) = 0.0869	Pr(T >t)		= 0.1737		Pr(T > t) = 0.9131	

The mean propensity score is not different for treated and controls in block 4.

Test for block 5: Two-sample t test with equal variances

Table A9.7: The mean propensity score is different for treated and controls in block 5

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	129	.7245333	.0019221	.0218308	.7207301	.7283365
1	556	.7411911	.0011506	.0271312	.738931	.7434512
combined	685	.7380541	.0010316	.0269989	.7360287	.7400795
diff		-.0166579	.0025624		-.0216889	-.0116269
diff = mean(0) -	mean(1)					t = -6.5010
Ho: diff = 0			degrees	of	freedom =	683
Ha: diff < 0		Ha:	diff	!= 0		Ha: diff > 0
Pr(T < t) = 0.0000	Pr(T >t)	= 0.0000			Pr(T > t) = 1.0000	

We **split the block 5** and retest, and then we check that blocks have shifted.

Table A9.8

Blocks of the pscore for treatment EPFRP	EPFRP 0	1	Total
2	72	76	148
3	1.685	1.811	3.496
4	826	1.382	2.208
5	129	556	685
7	3	45	48
Total	2.715	3.87	6.585

Test for block 5: Two-sample t test with equal variances

Table A9.9: The mean propensity score is different for treated and controls in block 5

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	109	.7171208	.0012953	.0135235	.7145532	.7196883
1	349	.7233854	.0007999	.0149441	.721812	.7249587
combined	458	.7218944	.0006937	.0148463	.7205312	.7232577
diff		-.0062646	.0016042		-.0094171	-.003112
diff = mean(0) -	mean(1)					t = -3.9051
Ho: diff = 0			degrees	of	freedom =	456
Ha: diff < 0		Ha:	diff	!= 0		Ha: diff > 0
Pr(T < t) = 0.0001	Pr(T >t)	= 0.0001			Pr(T > t) = 0.9999	

We **split the block 5** and retest, and then we check that blocks have shifted.

Table A9.10

Blocks of the pscore for treatment EPFRP	EPFRP 0	1	Total
2	72	76	148
3	1.685	1.811	3.496
4	826	1.382	2.208
5	109	349	458
7	20	207	227
8	3	45	48
Total	2.715	3.87	6.585

Table A9.11: Test for block 5: Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	78	.7095802	.0006479	.005722	.7082901	.7108703
1	187	.7112017	.0005279	.0072188	.7101603	.7122432
combined	265	.7107244	.0004203	.006842	.7098969	.711552
diff		-.0016215	.0009186		-.0034302	.0001871
diff = mean(0) -	mean(1)					t = -1.7653
Ho: diff = 0			degrees	of	freedom =	263
Ha: diff < 0		Ha:	diff	!= 0		Ha: diff > 0
Pr(T < t) = 0.0393	Pr(T >	t)	= 0.0787		Pr(T > t) = 0.9607	

The mean propensity score is not different for treated and controls in block 5.

Table A9.12: Test for block 6: Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	31	.7360939	.0013264	.0073852	.733385	.7388028
1	162	.7374492	.0005615	.0071473	.7363402	.7385581
combined	193	.7372315	.0005171	.0071838	.7362116	.7382514
diff		-.0013553	.0014086		-.0041337	.001423
diff = mean(0) -	mean(1)					t = -0.9622
Ho: diff = 0			degrees	of	freedom =	191
Ha: diff < 0		Ha:	diff	!= 0		Ha: diff > 0
Pr(T < t) = 0.1686	Pr(T >t)	= 0.3372			Pr(T > t) = 0.8314	

The mean propensity score is not different for treated and controls in block 6

Table A9.13: Test for block 7: Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	20	.7649312	.0026269	.011748	.759433	.7704294
1	207	.7712115	.0008879	.0127746	.769461	.7729621
combined	227	.7706582	.0008488	.012788	.7689857	.7723307
diff		-.0062803	.0029717		-.0121363	-.0004243
diff = mean(0) -	mean(1)					t = -2.1134
Ho: diff = 0			degrees	of	freedom =	225
Ha: diff < 0		Ha:	diff	!= 0		Ha: diff > 0
Pr(T < t) = 0.0178	Pr(T >t)	= 0.0357		Pr(T > t) = 0.9822		

The mean propensity score is not different for treated and controls in block 7.

Table A9.14: Test for block 8: Two-sample t test with equal variances

Group	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf.	Interval]
0	3	.8686915	.0238715	.0413467	.7659806	.9714024
1	45	.8390999	.0062846	.0421583	.8264341	.8517656
combined	48	.8409494	.006105	.0422969	.8286676	.8532311
diff		.0295917	.0251175		-.0209673	.0801506
diff = mean(0) -	mean(1)				t =	1.1781
Ho: diff = 0			degrees	of	freedom =	46
Ha: diff < 0		Ha:	diff	!= 0		Ha: diff>0
Pr(T < t) = 0.8776	Pr(T >t)	= 0.2448			Pr(T > t) =	0.1224

The mean propensity score is not different for treated and controls in block 8.

The final number of blocks is 8. This number of blocks ensures that the mean propensity score is not different for treated and controls in each blocks.

Step II: Test of balancing property of the propensity score

Table A10.1: Testing the Balancing Property for covariates

Covariate	Block	Balanced	Not Balanced	Comments
	1			does not have observations
Age	2	x		
Age square	2	x		
Sex	2	x		
loghhsize	2	x		
Stratum	2	x		
Clandfrac	2	x		
rain_EP	2	x		
loghhsize	3	x		
Stratum			x	
Areapc	3		x	
rain_EP	3		x	
Clandfrac	5	x		
rain_EP	5		x	
Clandfrac	6	x		
Stratum	7		x	
Areapc	7		x	
rain_EP	7		x	

The balancing property is not satisfied.
So we try a different specification of the propensity score.

Table A10.2

Inferior of block	EPFRP		
of pscore	0	1	Total
.2	72	76	148
.4	1.685	1.811	3.496
.6	826	1.382	2.208
.7	78	187	265
.725	31	162	193
.75	20	207	227
.8	3	45	48
Total	2.715	3.87	6.585

Note: the common support option has been selected.

Step 1b: Algorithm to estimate the propensity score

The treatment is still 'EPFRP', but we now remove the 'EP_Rain' variable and lower level to 0.0001.

Table A11.1

EPFRP	Age	Agesq	Sex	hhsize	shareofmale	stratum	livestock	Clandfrac	_cons
Coef.	-.0186295	.0001535	.1065173	.0392974	.3912748	-.0598325	.1873343	-.7536322	.3419837
Std.Err.	.0172813	.0001969	.0647828	.0093216	.1387936	.0775349	.0532151	.1313381	.3671838
z	-1.08	0.78	1.64	4.22	2.82	-0.77	3.52	-5.74	0.93
P>z	0.281	0.436	0.100	0.000	0.005	0.440	0.000	0.000	0.352
[95% Conf.	-.0525003	-.0002325	-.0204547	.0210274	.1192442	-.2117981	.0830345	-1.01105	-.3776834
Interval]	.0152412	.0005395	.2334893	.0575674	.6633053	.0921332	.291634	-.4962143	1.061651
Number of obs	=	6586							
LR chi2(8)	=	72.60							
Prob > chi2	=	0.0000							
Pseudo R2	=	0.0081							
Log likelihood	=	-4427.14							

Note: the common support option has been selected.

The region of common support is [.35378466, .85246556].

Table A11.2: Estimated propensity score

	Percentiles	Smallest		
1%	.4487029	.3537847		
5%	.4933914	.3656081		
10%	.5183903	.3768392	Obs	6586
25%	.5579566	.3831111	Sum of Wgt.	6586
50%	.5916011		Mean	.5876101
		Largest	Std. Dev.	.0516318
75%	.623515	.7717166		
90%	.6488504	.7785103	Variance	.0026658
95%	.6635021	.7902148	Skewness	-.4387265
99%	.6938927	.8524656	Kurtosis	3.715523

Step Ib: Identification of the optimal number of blocks

The final number of blocks is 9. This number of blocks ensures that the mean propensity score is not different for treated and controls in each blocks.

Step IIb: Test of balancing property of the propensity score

The balancing property is satisfied.

Table A11.3: The inferior bound, the number of treated and the number of controls for each block.

	Inferior of block	EPFRP		
Block	of pscore	0	1	Total
2	.2	3	5	8
3	.4	207	180	387
4	.5	452	503	955
5	.55	1003	1452	2455
6	.6	872	1299	2171
7	.65	166	397	563
8	.7	13	33	46
9	.8	0	1	1
	Total	2716	3870	6586

Step 1c: Algorithm to estimate the propensity score
The treatment is still 'EPFRP', but we now increase the level to 0.001.

Table A12.1

EPFRP	Age	Agesq	Sex	hhsize	shareofmale	stratum	livestock	Clandfrac	_cons
Coef.	-.0186295	.0001535	.1065173	.0392974	.3912748	-.0598325	.1873343	-.7536322	.3419837
Std.Err.	.0172813	.0001969	.0647828	.0093216	.1387936	.0775349	.0532151	.1313381	.3671838
z	-1.08	0.78	1.64	4.22	2.82	-0.77	3.52	-5.74	0.93
P>z	0.281	0.436	0.100	0.000	0.005	0.440	0.000	0.000	0.352
[95% Conf.	-.0525003	-.0002325	-.0204547	.0210274	.1192442	-.2117981	.0830345	-1.01105	-.3776834
Interval]	.0152412	.0005395	.2334893	.0575674	.6633053	.0921332	.291634	-.4962143	1.061651
Number of obs	=	6586							
LR chi2(8)	=	72.60							
Prob > chi2	=	0.0000							
Pseudo R2	=	0.0081							
Log likelihood	=	-4427.14							

Note: Exactly the same results as in table A11.2

Note: the common support option has been selected
The region of common support is [.35378466, .85246556]
Description of the estimated propensity score in region of common support

Table A12.2: Estimated propensity score

	Percentiles	Smallest		
1%	.4487029	.3537847		
5%	.4933914	.3656081		
10%	.5183903	.3768392	Obs	6586
25%	.5579566	.3831111	Sum of Wgt.	6586
50%	.5916011		Mean	.5876101
		Largest	Std. Dev.	.0516318
75%	.623515	.7717166		
90%	.6488504	.7785103	Variance	.0026658
95%	.6635021	.7902148	Skewness	-.4387265
99%	.6938927	.8524656	Kurtosis	3.715523

Step Ic: Identification of the optimal number of blocks
The final number of blocks is 9. This number of blocks ensures that the mean propensity score is not different for treated and controls in each blocks.

Step IIc: Test of balancing property of the propensity score
The balancing property is satisfied

Table A12.3: The inferior bound, the number of treated and the number of controls for each block

	Inferior of	EPFRP		
Block	block of	0	1	Total
2	.2	3	5	8
3	.4	207	180	387
4	.5	452	503	955
5	.55	1003	1452	2455
6	.6	872	1299	2171
7	.65	166	397	563
8	.7	13	33	46
9	.8	0	1	1
	Total	2716	3870	6586

End of the algorithm to estimate the pscore.

Table A13: Estimation of the propensity score

	EPFRP
	(1)
Age	-0.0186
	(0.0173)
Age squared	0.0002
	(0.0002)
Sex	-0.1065
	(0.0648)
hhsize	0.0393***
	(0.0093)
Share of male	0.3913***
	(0.1388)
Stratum	-0.0598
	(0.0775)
Livestock ownership	0.1873***
	(0.0532)
Cotton land fraction	-0.7536***
	(0.1313)
Constant	0.5550
	(0.3595)
Number of Obs.	6586
R^2	
F	
Log likelihood(i)	-4427.1405

Note: Standard errors in parentheses * p<0.10, ** p<0.05, *** p<0.01.

A simulation-based sensitivity analysis for matching estimators

Table A14.1: Estimation of Average Treatment Effects using Different Matching Methods

ATT estimation with	Number of (ii) Treated	Controls	ATT	Analytical Std.Err.	t	Bootstrapped Std.Err.	t
Nearest Neighbor Matching method	3870	530	-0.180	0.052	-3.447	0.080	-2.256
Kernel Matching method	3870	2716	-0.199	n.a.(i)	n.a.	0.050	-3.959
Stratification method	3869	2717	-0.298	0.026	-11.497	0.056	-5.360

Note: (ii) the numbers of treated and controls refer to actual nearest neighbour matches.
Bootstrap replications (REP) = 100.
Source: Authors estimations.

Table A14.2: The Baseline ATT estimation (with no simulated confounder)

Outcome	Type	ATT estimation with	Number of (i) Treated	Controls	ATT	Analytical Std.Err.	t
Logyield	Continuous	Nearest Neighbor Matching method	3870	516	-0.114	0.054	-2.127
LYIELDDV	Dummy	Nearest Neighbor Matching method	3870	516	-0.037	0.028	-1.355

Table A14.3: The Baseline ATT estimation (with simulated confounder)

Nearest Neighbor Matching method	The probability of having U=1 if T=1 and Y=1 (p11) is equal to:	0.98
	The probability of having U=1 if T=1 and Y=0 (p10) is equal to:	0.99
	The probability of having U=1 if T=0 and Y=1 (p01) is equal to:	0.99
	The probability of having U=1 if T=0 and Y=0 (p00) is equal to:	0.99
	The probability of having U=1 if T=1 (p1.) is equal to:	0.99
	The probability of having U=1 if T=0 (p0.) is equal to:	0.99
Kernel Matching method	The probability of having U=1 if T=1 and Y=1 (p11) is equal to:	0.98
	The probability of having U=1 if T=1 and Y=0 (p10) is equal to:	0.99
	The probability of having U=1 if T=0 and Y=1 (p01) is equal to:	0.99
	The probability of having U=1 if T=0 and Y=0 (p00) is equal to:	0.99
	The probability of having U=1 if T=1 (p1.) is equal to:	0.99
	The probability of having U=1 if T=0 (p0.) is equal to:	0.99
Nearest Neighbor Matching method	The probability of having U=1 if T=1 and Y=1 (p11) is equal to:	0.6
	The probability of having U=1 if T=1 and Y=0 (p10) is equal to:	0.5
	The probability of having U=1 if T=0 and Y=1 (p01) is equal to:	0.5
	The probability of having U=1 if T=0 and Y=0 (p00) is equal to:	0.2
	The probability of having U=1 if T=1 (p1.) is equal to:	0.55
	The probability of having U=1 if T=0 (p0.) is equal to:	0.36

Table A14.4: The simulated ATT estimation (with the confounder u):

	The probability of having U=1 if T=1 and Y=1 (p11) is equal to:	0.6
	The probability of having U=1 if T=1 and Y=0 (p10) is equal to:	0.5
	The probability of having U=1 if T=0 and Y=1 (p01) is equal to:	0.5
	The probability of having U=1 if T=0 and Y=0 (p00) is equal to:	0.2
	The probability of having U=1 if T=1 (p1.) is equal to:	0.55
	The probability of having U=1 if T=0 (p0.) is equal to:	0.36
50 centile	The probability of having U=1 if T=1 and Y=1 (p11) is equal to:	0.60
	The probability of having U=1 if T=1 and Y=0 (p10) is equal to:	0.50
	The probability of having U=1 if T=0 and Y=1 (p01) is equal to:	0.50
	The probability of having U=1 if T=0 and Y=0 (p00) is equal to:	0.20
	The probability of having U=1 if T=1 (p1.) is equal to:	0.58
	The probability of having U=1 if T=0 (p0.) is equal to:	0.45
25 centile	The probability of having U=1 if T=1 and Y=1 (p11) is equal to:	0.60
	The probability of having U=1 if T=1 and Y=0 (p10) is equal to:	0.50
	The probability of having U=1 if T=0 and Y=1 (p01) is equal to:	0.50
	The probability of having U=1 if T=0 and Y=0 (p00) is equal to:	0.20
	The probability of having U=1 if T=1 (p1.) is equal to:	0.59
	The probability of having U=1 if T=0 (p0.) is equal to:	0.47
	The probability of having U=1 if T=1 and Y=1 (p11) is equal to:	0.60
	The probability of having U=1 if T=1 and Y=0 (p10) is equal to:	0.50
	The probability of having U=1 if T=0 and Y=1 (p01) is equal to:	0.50
	The probability of having U=1 if T=0 and Y=0 (p00) is equal to:	0.20
	The probability of having U=1 if T=1 (p1.) is equal to:	0.58
	The probability of having U=1 if T=0 (p0.) is equal to:	0.43
	The probability of having U=1 if T=1 and Y=1 (p11) is equal to:	0.96
	The probability of having U=1 if T=1 and Y=0 (p10) is equal to:	0.99
	The probability of having U=1 if T=0 and Y=1 (p01) is equal to:	0.95
	The probability of having U=1 if T=0 and Y=0 (p00) is equal to:	0.99
	The probability of having U=1 if T=1 (p1.) is equal to:	0.97
	The probability of having U=1 if T=0 (p0.) is equal to:	0.96
	The probability of having U=1 if T=1 and Y=1 (p11) is equal to:	0.96
	The probability of having U=1 if T=1 and Y=0 (p10) is equal to:	1.00
	The probability of having U=1 if T=0 and Y=1 (p01) is equal to:	0.96
	The probability of having U=1 if T=0 and Y=0 (p00) is equal to:	0.99
	The probability of having U=1 if T=1 (p1.) is equal to:	0.97
	The probability of having U=1 if T=0 (p0.) is equal to:	0.96
	The probability of having U=1 if T=1 and Y=1 (p11) is equal to:	0.96
	The probability of having U=1 if T=1 and Y=0 (p10) is equal to:	0.99
	The probability of having U=1 if T=0 and Y=1 (p01) is equal to:	0.95
	The probability of having U=1 if T=0 and Y=0 (p00) is equal to:	0.99
	The probability of having U=1 if T=1 (p1.) is equal to:	0.97
	The probability of having U=1 if T=0 (p0.) is equal to:	0.96
	The probability of having U=1 if T=1 and Y=1 (p11) is equal to:	0.49
	The probability of having U=1 if T=1 and Y=0 (p10) is equal to:	0.40
	The probability of having U=1 if T=0 and Y=1 (p01) is equal to:	0.39
	The probability of having U=1 if T=0 and Y=0 (p00) is equal to:	0.56
	The probability of having U=1 if T=1 (p1.) is equal to:	0.48
	The probability of having U=1 if T=0 (p0.) is equal to:	0.40

Table A14.5: Att Estimation: General multiple-; Within; and Between Imputation Effect

Binary variable used to simulate the con-founder	Centile	ATT estimation with	ATT	General multiple-imputation Std.Err.	Effect Outcome	Effect Selection	ATT	Within-imputation Std.Err.	Effect Outcome	Effect Selection	ATT	Between-imputation Std.Err.	Effect Outcome	Effect Selection	The probability of having U=1 if T=1 and Y=1 (p11) is equal to:	The probability of having U=1 if T=1 and Y=0 (p10) is equal to:	The probability of having U=1 if T=0 and Y=1 (p01) is equal to:	The probability of having U=1 if T=0 and Y=0 (p00) is equal to:	The probability of having U=1 if T=1 (p1.) is equal to:	The probability of having U=1 if T=0 (p0.) is equal to:
Male	50	Nearest Neighbor Matching method	-0.071	0.035	3.536	0.606	-0.070	0.030	3.454	0.627	-0.074	0.018	3.051	0.578	0.98	0.99	0.99	0.99	0.99	0.99
Male	25	Nearest Neighbor Matching method	-0.070	0.034	2.728	0.664	-0.073	0.031	4.721	0.614	-0.068	0.018	2.876	0.619	0.98	0.99	0.99	0.99	0.99	0.99
Male	75	Nearest Neighbor Matching method	-0.070	0.034	4.861	0.677	-0.071	0.031	2.615	0.595	-0.073	0.017	5.128	0.553	0.98	0.99	0.99	0.99	0.99	0.99
	50	Nearest Neighbor Matching method	-0.121	0.038	4.249	2.067	-0.123	0.031	4.175	2.101	-0.124	0.021	4.204	2.080	0.6	0.5	0.5	0.2	0.55	0.36

Note: Both the outcome and the selection effect are odds ratios from logit estimations.

Table A14.6: ATT estimation with simulated confounder

	Outcome (Y)	Potential Binary variable used to simulate the con-founder (U)	Centile	ATT estimation with	General multiple-imputation ATT	General multiple-imputation Std.Err.	Effect Outcome	Effect Selection	Within-imputation ATT	Within-imputation Std.Err.	Effect Outcome	Effect Selection	Between-imputation ATT	Between-imputation Std.Err.	Effect Outcome	Effect Selection	The probability of having U=1 if T=1 and Y=1 (p11) is equal to:	The probability of having U=1 if T=1 and Y=0 (p10) is equal to:	The probability of having U=1 if T=0 and Y=1 (p01) is equal to:	The probability of having U=1 if T=0 and Y=0 (p00) is equal to:	The probability of having U=1 if T=1 (p1.) is equal to:	The probability of having U=1 if T=0 (p0.) is equal to:	Confounder U account for % of baseline estimate	Avg ATT
1	Binary	Male	50	Nearest Neighbor Matching method	-0.071	0.035	3.536	0.606	-0.07	0.03	3.454	0.627	-0.074	0.018	3.051	0.578	0.98	0.99	0.99	0.99	0.99	0.99	-94%	-0.072
2	Binary	Male	25	Nearest Neighbor Matching method	-0.07	0.034	2.728	0.664	-0.073	0.031	4.721	0.614	-0.068	0.018	2.876	0.619	0.98	0.99	0.99	0.99	0.99	0.99	-90%	-0.070
3	Binary	Male	75	Nearest Neighbor Matching method	-0.07	0.034	4.861	0.677	-0.071	0.031	2.615	0.595	-0.073	0.017	5.128	0.553	0.98	0.99	0.99	0.99	0.99	0.99	-93%	-0.071
4	Continuous	Male	50	Nearest Neighbor Matching method	-0.123	0.063	0.226	1.228	-0.121	0.054	0.204	1.217	-0.122	0.039	0.209	1.195	0.96	0.99	0.95	0.99	0.97	0.96	-7%	-0.122
5	Continuous	Male	25	Nearest Neighbor Matching method	-0.121	0.063	0.248	1.195	-0.121	0.055	0.227	1.144	-0.119	0.034	0.217	1.208	0.96	1.00	0.96	0.99	0.97	0.96	-6%	-0.120
6	Continuous	Male	75	Nearest Neighbor Matching method	-0.125	0.064	0.187	1.262	-0.12	0.054	0.15	1.228	-0.122	0.034	0.169	1.242	0.96	0.99	0.95	0.99	0.97	0.96	-7%	-0.122
7	Continuous		50	Nearest Neighbor Matching method	-0.198	0.071	4.101	1.657	-0.206	0.056	4.442	1.654	-0.207	0.037	4.49	1.659	0.60	0.50	0.50	0.20	0.58	0.45	-79%	-0.204
8	Continuous		25	Nearest Neighbor Matching method	-0.187	0.069	4.471	1.598	-0.184	0.055	4.395	1.609	-0.186	0.037	4.521	1.585	0.60	0.50	0.50	0.20	0.59	0.47	-63%	-0.186
9	Continuous		75	Nearest Neighbor Matching method	-0.201	0.069	4.397	1.778	-0.207	0.057	4.38	1.755	-0.195	0.042	4.372	1.761	0.60	0.50	0.50	0.20	0.58	0.43	-76%	-0.201
10	Continuous	Rain	50	Nearest Neighbor	-0.121	0.069	0.499	1.398	-0.124	0.055	0.464	1.399	-0.122	0.04	0.485	1.401	0.49	0.40	0.39	0.56	0.48	0.40	-7%	-0.122
11	Continuous	Rain	25	Nearest Neighbor	-0.127	0.074	0.487	1.414	-0.123	0.055	0.465	1.383	-0.12	0.042	0.481	1.404	0.49	0.40	0.39	0.56	0.48	0.40	-8%	-0.123
12	Continuous	Rain	75	Matching method	-0.128	0.071	0.725	1.404	-0.134	0.055	0.736	1.406	-0.134	0.046	0.742	1.385	0.49	0.40	0.39	0.56	0.48	0.40	-16%	-0.132
13	Continuous	Young	50	Nearest Neighbor	-0.128	0.071	0.962	1.069	-0.135	0.054	0.927	1.065	-0.129	0.04	0.93	1.068	0.23	0.26	0.22	0.25	0.24	0.23	-15%	-0.131
14	Continuous	Young	25	Matching method	-0.136	0.069	1.131	1.07	-0.128	0.055	1.127	1.064	-0.138	0.042	1.148	1.066	0.24	0.24	0.23	0.21	0.24	0.23	-18%	-0.134
15	Continuous	Young	75	Matching method	-0.131	0.068	0.883	1.074	-0.133	0.055	0.87	1.073	-0.134	0.038	0.832	1.071	0.23	0.26	0.22	0.25	0.24	0.23	-16%	-0.133

Table A15.1: Matching and Propensity Score Estimators

Propensity score matching methods (i)	Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
1. One-to-One propensity score matching (ii)	productivity	Unmatched	1278,227	1380,409	-102,182	94,277	-1,080
		ATT	1291,437	1660,469	-369,032	144,114	-2,560
		ATU	1640,836	1178,661	-462,176		
		ATE			-415,930		
2. k-nearest neighbors matching (iii)	productivity	Unmatched	1278,227	1380,409	-102,182	94,277	-1,080
		ATT	1383,466	1599,065	-215,600	163,080	-1,320
		ATU	1602,739	1389,113	-213,626		
		ATE			-214,717		
3. radius matching (iv)	productivity	Unmatched	1278,227	1380,409	-102,182	94,277	-1,080
		ATT	1420,120	1689,646	-269,526	172,248	-1,560
		ATU	1533,129	1327,329	-205,800		
		ATE			-242,665		
4. kernel (v)	productivity	Unmatched	1278,227	1380,409	-102,182	94,277	-1,080
		ATT	1192,912	1184,001	8,911	78,477	0,110
		ATU	1322,991	1317,266	-5,725		
		ATE			2,734		
5.local linear regression (vi)	productivity	Unmatched	1278,2270	1380,4087	-102,1817	94,2765	-1,0800
		ATT	1126,3095	1196,0222	-69,7127	195,5590	-0,3600
		ATU	1246,8842	1235,1213	-11,7629		
		ATE			-44,943443		
6.'spline-smoothing' (vii)	productivity	Unmatched	1278,227	1380,409	-102,182	94,277	-1,080
		ATT	1383,466	1629,728	-246,263	,	,
		ATU	1602,739	1303,153	-299,587		
		ATE			-270,129		
7. Mahalanobis matching (viii)	productivity	Unmatched	1278,227	1380,409	-102,182	94,277	-1,080
		ATT	1067,260	1080,069	-12,810	70,558	-0,180
		ATU	1075,650	1085,807	10,157		
		ATE			-2,529		

Source: Authors' estimations.

The following matching methods all have negative difference between treated and controls:

- 1-to-1 propensity score matching;
- k-nearest neighbors matching;
- radius matching;
- 'spline-smoothing'.

However, the following matching methods all have positive difference between treated and controls for the 'productivity' variable:

Kernel Matching:

Table A15.2: Kernel outcome(logyield productivity) kerneltype(normal) bwidth(0.2)

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
logyield	Unmatched	6.65031	6.72555	-0.07524	0.04321	-1.74000
	ATT	6.70968	6.72512	-0.01544	0.05876	-0.26000
	ATU	6.88367	6.88447	0.00081	0.00689	
	ATE			-0.00857		
productivity	Unmatched	1278.22703	1380.40869	-102.18167	94.27655	-1.08000
	ATT	1259.27287	1251.36738	7.90549	110.06478	0.07000
	ATU	1536.02831	1514.17336	-21.85496		
	ATE			-4.66532		

psmatch2:	psmatch2: Common		
Treatment	Common Support		
assignment	Off support	On support	Total
Untreated	397	528	925
Treated	516	722	1.238
Total	913	1.25	2.163

Note: mahalanobis(logyield productivity) common trim(0.05) odds index logit quietly ate.

Table A15.3: kernel outcome(logyield productivity) kerneltype(epan) bwidth(0.2)

Variable	Sample	Treated	Controls	Difference	S,E,	T-stat
logyield	Unmatched	6.65031	6.72555	-0.07524	0.04321	-1.74000
	ATT	6.75976	6.76215	-0.00240	0.05187	-0.05000
	ATU	6.85237	6.85289	0.00052	0.00033	
	ATE			-0.00117		
productivity	Unmatched	1278.22703	1380.40869	-102.18167	94.27655	-1.08000
	ATT	1192.91168	1184.75985	8.15183	79.00429	0.10000
	ATU	1322.99139	1318.65854	-4.33285		
	ATE			2.88287		

psmatch2:	psmatch2: Common		
Treatment	Common Support		
assignment	Off support	On support	Total
Untreated	427	498	925
Treated	556	682	1.238
Total	983	1.18	2.163

Note: mahalanobis(logyield productivity) common trim(0.05) odds index logit quietly ate.

Table A15.4: kernel outcome(logyield productivity) kerneltype(uniform) bwidth(0.2)

Variable	Sample	Treated	Controls	Difference	S,E,	T-stat
logyield	Unmatched	6.650	6.726	-0.075	0.043	-1.740
	ATT	6.760	6.762	-0.003	0.052	-0.050
	ATU	6.852	6.852	0.000	-0.001	
	ATE			-0.001		
productivity	Unmatched	1278.227	1380.409	-102.182	94.277	-1.080
	ATT	1192.912	1184.001	8.911	78.477	0.110
	ATU	1322.991	1317.266	-5.725		
	ATE			2.734		

psmatch2: Treatment assignment	psmatch2: Common Common Support		
	Off support	On support	Total
Untreated	427	498	925
Treated	556	682	1.238
Total	983	1.18	2.163

Note: mahalanobis(logyield productivity) common trim(0.05) odds index logit quietly ate.

Table A15.5: kernel outcome(logyield productivity) kerneltype(uniform) pscore(Clandfrac) bwidth(0.2)

Variable	Sample	Treated	Controls	Difference	S,E,	T-stat
logyield	Unmatched	6.650	6.726	-0.075	0.043	-1.740
	ATT	6.681	6.683	-0.002	0.046	-0.040
	ATU	6.688	6.687	-0.0005		
	ATE			-0.001		
productivity	Unmatched	1278.227	1380.409	-102.182	94.277	-1.080
	ATT	1180.566	1170.132	10.434	68.139	0.150
	ATU	1201.698	1185.119	-16.578		
	ATE			-2.126		

psmatch2: Treatment assignment	psmatch2: Common Common Support		
	Off support	On support	Total
Untreated	75	850	925
Treated	260	978	1.238
Total	335	1.828	2.163

Note: mahalanobis(logyield productivity) common trim(0.05) odds index logit ate.

Table A15.6: kernel outcome(logyield productivity) kerneltype(biweight) bwidth(0.2)

Variable	Sample	Treated	Controls	Difference	S,E,	T-stat
logyield	Unmatched	6.65031	6.72555	-0.07524	0.04321	-1.74000
	ATT	6.75976	6.76238	-0.00263	0.05222	-0.05000
	ATU	6.85237	6.85277	0.00040		
	ATE			-0.00135		
productivity	Unmatched	1278.22703	1380.40869	-102.18167	94.27655	-1.08000
	ATT	1192.91168	1185.55898	7.35270	79.54721	0.09000
	ATU	1322.99139	1318.81657	-4.17482		
	ATE			2.48769667		

psmatch2: Treatment assignment	psmatch2: Common Common Support		
	Off support	On support	Total
Untreated	427	498	925
Treated	556	682	1.238
Total	983	1.18	2.163

Note: mahalanobis(logyield productivity) common trim(0.05) odds index logit quietly ate.

Local linear regression:

Some of the local linear regression specifications also yields positive difference between treated and controls for the 'logyield' variable, but not for the 'productivity' variable and not for all specifications either.

Table A15.7: llr outcome(logyield productivity) kerneltype(epan) pscore(Clandfrac) bwidth(0.2)

Variable	Sample	Treated	Controls	Difference	S,E,	T-stat
logyield	Unmatched	6.650	6.726	-0.075	0.043	-1.740
	ATT	6.726	6.709	0.018	0.054	0.330
	ATU	6.670	6.659	-0.011		
	ATE			0.004		
productivity	Unmatched	1278.227	1380.409	-102.182	94.277	-1.080
	ATT	1138.840	1139.669	-0.829	65.987	-0.010
	ATU	1109.045	1048.998	-60.048		
	ATE			-28.955		

psmatch2: Treatment assignment	psmatch2: Common Common Support		
	Off support	On support	Total
Untreated	91	834	925
Treated	316	922	1.238
Total	407	1.756	2.163

Note: mahalanobis(logyield productivity) common trim(0.05) odds index logit quietly ate.

Table A15.8: llr outcome(logyield productivity) kerneltype(uniform) pscore(Clandfrac) bwidth(0.2)

Variable	Sample	Treated	Controls	Difference	S,E,	T-stat
logyield	Unmatched	6.650	6.726	-0.075	0.043	-1.740
	ATT	6.726	6.709	0.017	0.053	0.320
	ATU	6.670	6.657	-0.012		
	ATE			0.003		
productivity	Unmatched	1278.227	1380.409	-102.182	94.277	-1.080
	ATT	1138.840	1139.983	-1.143	65.263	-0.020
	ATU	1109.045	1045.906	-63.140		
	ATE			-30.588		

psmatch2: Treatment assignment	psmatch2: Common Common Support		
	Off support	On support	Total
Untreated	91	834	925
Treated	316	922	1.238
Total	407	1.756	2.163

Note: mahalanobis(logyield productivity) common trim(0.05) odds index logit quietly ate.

Mahalanobis matching:

Table A15.9: Mahalanobis(logyield productivity) add outcome(logyield productivity) pscore(Clandfrac)

Variable	Sample	Treated	Controls	Difference	S,E,	T-stat
logyield	Unmatched	6.650	6.726	-0.075	0.043	-1.740
	ATT	6.636	6.511	0.125	0.344	0.360
	ATU	6.725	6.707	-0.018		
	ATE			0.064		
productivity	Unmatched	1278.22703	1380.40869	-102.181666	94.276548	-1.08
	ATT	1176.32671	1184.76212	-8.43541104	628.78067	-0.01
	ATU	1333.6202	1159.12308	-174.497113		
	ATE			-79.4836194		

psmatch2: Treatment assignment	psmatch2: Common Support		
	Off support	On support	Total
Untreated	3	922	925
Treated	5	1.233	1.238
Total	8	2.155	2.163

Note: kernel(normal) llr bwidth(0.06) caliper(0.01) ate.

Table A15.10: mahalanobis(logyield productivity) add outcome(logyield productivity) pscore(Clandfrac)

Variable	Sample	Treated	Controls	Difference	S,E,	T-stat
logyield	Unmatched	6.65030632	6.72554821	-0.07524189	0.0432089	-1.74
	ATT	6.73601422	6.730604	0.005410218	0.166778	0.03
	ATU	6.71386027	6.74737519	0.033514915		
	ATE			0.017801096		
productivity	Unmatched	1278.22703	1380.40869	-102.181666	94.276548	-1.08
	ATT	1068.85187	1069.48209	-0.63022617	181.89105	0
	ATU	1086.7834	1086.79533	0.011923453		
	ATE			-0.34711344		

psmatch2: Treatment assignment	psmatch2: Common Support		
	Off support	On support	Total
Untreated	86	839	925
Treated	174	1.064	1.238
Total	260	1.903	2.163

Note: kernel(epan) llr bwidth(0.06) caliper(0.01) ate.

Table A15.11: mahalanobis(logyield productivity) add outcome(logyield productivity) pscore(Clandfrac)

Variable	Sample	Treated	Controls	Difference	S,E,	T-stat
logyield	Unmatched	6.65030632	6.72554821	-0.07524189	0.0432089	-1.74
	ATT	6.73601422	6.73065803	0.005356185	0.1643768	0.03
	ATU	6.71386027	6.74770943	0.033849155		
	ATE			0.017918246		
productivity	Unmatched	1278.22703	1380.40869	-102.181666	94.276548	-1.08
	ATT	1068.85187	1068.96396	-0.11209383	179.2766	0
	ATU	1086.7834	1086.95002	0.166615722		
	ATE			0.010784423		

psmatch2: Treatment assignment	psmatch2: Common Support		
	Off support	On support	Total
Untreated	86	839	925
Treated	174	1.064	1.238
Total	260	1.903	2.163

Note: kernel(uniform) llr bwidth(0.06) caliper(0.01) ate.

Table A15.12: mahalanobis(logyield productivity) add outcome(logyield productivity)

Variable	Sample	Treated	Controls	Difference	S,E,	T-stat
logyield	Unmatched	6.650	6.726	-0.075	0.043	-1.740
	ATT	6.716	6.721	-0.005	0.063	-0.080
	ATU	6.690	6.700	0.009		
	ATE			0.001		
productivity	Unmatched	1278.227	1380.409	-102.182	94.277	-1.080
	ATT	1067.260	1080.069	-12.810	70.558	-0.180
	ATU	1075.650	1085.807	10.157		
	ATE			-2.529		

psmatch2: Treatment assignment	psmatch2: Common Support		
	Off support	On support	Total
Untreated	79	846	925
Treated	194	1.044	1.238
Total	273	1.89	2.163

Note: kernel(uniform) llr bwidth(0.06) caliper(0.01) qui ate.

Table A16.1: Nearest-Neighbour Matching: Matching Estimator: Average Treatment Effect

Weighting Matrix	Number of Obs	Number of Matches (m)	logyield	Coef.	Std.Err.	z	P>z	[95% Conf.	Interval]	condition
Inverse Variance	2163	4	SATE	-.1624575	.0457288	-3.55	0.000	-.2520844	-.0728307	
Inverse Variance	1434	4	SATE	-.0448437	.102304	-0.44	0.661	-.2453559	.1556686	> 1997
Inverse Variance	1081	4	SATE	-.1358372	.1030121	-1.32	0.187	-.3377371	.0660627	> 1998
Inverse Variance	823	4	SATE	-.2515395	.0873619	-2.88	0.004	-.4227656	-.0803134	>1999

Note: Matching variables: Age Agesq Sex loghhsize stratum livestock Areapc Clandfrac rain_EP
Source: Author's estimations.

Table A16.2: Nearest-Neighbour Matching: Matching Estimator: Population Average Treatment Effect

Weighting Matrix	Number of Obs	Number of Matches (m)	logyield	Coef.	Std.Err.	z	P>z	[95% Con	Interval]	condition
Inverse variance	2163	4	PATE	-.1624575	.0458071	-3.55	0.000	-.2522379	-.0726772	
Inverse variance	1434	4	PATE	-.0448437	.102217	-0.44	0.661	-.2451854	.155498	>1997
Inverse variance	1081	4	PATE	-.1358372	.1028776	-1.32	0.187	-.3374737	.0657992	> 1998
Inverse variance	452	4	PATE	-.311264	.1197778	-2.60	0.009	-.5460241	-.0765039	>2000

Note: Matching variables: Age Agesq Sex loghhsize stratum livestock Areapc Clandfrac rain_EP
Source: Author's estimations.

Table A16.3: Nearest-Neighbour Matching: Matching Estimator: Average Treatment Effect for the Treated

Weighting Matrix	Number of Obs	Number of Matches (m)	logyield	Coef.	Std.Err.	z	P>z	[95% Conf.	Interval]	condition
Inverse Variane	2163	4	SATT	-.2237923	.0495116	-4.52	0.000	-.3208332	-.1267514	
Inverse variance	2163	1	SATT	-.2054818	.0569697	-3.61	0.000	-.3171404	-.0938232	
Inverse variance	1434	4	SATT	-0.02322	0.110252	-0.21	0.833	-0.2393077	0.192873	>1997
Inverse variance	1081	4	SATT	-0.11365	0.110238	-1.03	0.303	-0.3297126	0.102413	>1998
Inverse variance	823	4	SATT	-0.23514	0.094145	-2.5	0.013	-0.4196651	-0.05062	>1999
Inverse variance	452	4	SATT	-0.2993	0.129942	-2.3	0.021	-0.5539794	-0.04462	>2000

Note: Matching variables: Age Agesq Sex loghhsize stratum livestock Areapc Clandfrac rain_EP.
Source: Author's estimations.

Table A16.4: Nearest-Neighbour Matching: Matching Estimator: Average Treatment Effect for the Treated

Weighting Matrix	Number of Obs	Number of Matches (m)	logyield	Coef.	Std.Err.	z	P>z	[95% Conf	Interval]	condition
Inverse variance	2163	4	SATT	-.2654978	.0499541	-5.31	0.000	-.363406	-.1675897	
Inverse variance	1434	4	SATT	-0.02105	0.10987	-0.19	0.848	-0.2363872	0.194295	>1997
Inverse variance	1081	4	SATT	-0.06214	0.110297	-0.56	0.573	-0.2783159	0.154042	>1998
Inverse variance	823	4	SATT	-0.22318	0.096118	-2.32	0.02	-0.4115632	-0.03479	>1999
Inverse variance	452	4	SATT	-0.24791	0.134883	-1.84	0.066	-0.512279	0.016453	>2000

Notes: Matching variables: Age Agesq Sex loghhsize stratum livestock Areapc Clandfrac rain_EP.
Bias-adj variables: Age Agesq Sex loghhsize stratum livestock Areapc Clandfrac rain_EP.
Source: Author's estimations.

Table A17.1: Mantel-Haenszel bounds to check sensitivity of estimated ATT

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Logyield50	Unmatched	0.7806	0.8307	-0.0501	0.0109	-4.6000
	ATT	0.7845	0.8322	-0.0476	0.0115	-4.1400
	ATU	0.8307	0.7840	-0.0467		
	ATE			-0.0471		

psmatch2: Treatment assignment	psmatch2: Common support Off suppo	On suppor	Total
Untreated	0	2.528	2.528
Treated	418	2.33	2.748
Total	418	4.858	5.276

Noted: outcome(Logyield50) pscore(mypscore) neighbor(1) caliper(0.25) common noreplacement quietly ate
Logyield50 = 1 if logyield>=6.682483.

Table A17.2: Mantel-Haenszel (1959) bounds for variable Logyield50

Gamma	Q_mh+	Q_mh-	p_mh+	p_mh-
1	3.992	3.992	0.000	0.000
1.05	4.645	3.341	0.000	0.000
1.1	5.269	2.721	0.000	0.003
1.15	5.866	2.130	0.000	0.017
1.2	6.439	1.564	0.000	0.059
1.25	6.991	1.021	0.000	0.154
1.3	7.522	0.500	0.000	0.308
1.35	8.035	-0.001	0.000	0.500
1.4	8.530	0.409	0.000	0.341
1.45	9.010	0.875	0.000	0.191
1.5	9.474	1.325	0.000	0.093
1.55	9.925	1.761	0.000	0.039
1.6	10.363	2.183	0.000	0.015
1.65	10.789	2.592	0.000	0.005
1.7	11.204	2.990	0.000	0.001
1.75	11.608	3.376	0.000	0.000
1.8	12.002	3.752	0.000	0.000
1.85	12.386	4.118	0.000	0.000
1.9	12.762	4.475	0.000	0.000
1.95	13.129	4.823	0.000	0.000
2	13.487	5.162	0.000	0.000

Notes: performs sensitivity analysis at gamma = 1, 1.05, 1.10, ..., 2.
Gamma: odds of differential assignment due to unobserved factors; Q_mh+: Mantel-Haenszel statistic (assumption: overestimation of treatment effect); Q_mh-: Mantel-Haenszel statistic (assumption: underestimation of treatment effect); p_mh+: significance level (assumption: overestimation of treatment effect); p_mh-: significance level (assumption: underestimation of treatment effect).

Table A17.3: Mantel-Haenszel bounds to check sensitivity of estimated ATT

Variable	Sample	Treated	Controls	Difference	S.E.	T-stat
Logyield50	Unmatched	.780567686	.830696203	-.050128517	.010907704	-4.60
	ATT	.784549356	.832188841	-.047639485	.011512538	-4.14
	ATU	.830696203	.784018987	-.046677215		
	ATE			-.04713874		

psmatch2: Treatment assignment	psmatch2: Common support Off suppo	On suppor	Total
Untreated	0	2.528	2.528
Treated	418	2.33	2.748
Total	418	4.858	5.276

Notes: outcome(Logyield50) pscore(mypscore) neighbor(1) caliper(0.25) common noreplacement quietly ate

Table A17.4: Mantel-Haenszel (1959) bounds for variable Logyield50

Gamma	Q_mh+	Q_mh-	p_mh+	p_mh-
1	3.99211	3.99211	.000033	.000033
1.05	4.64509	3.34128	1.7e-06	.000417
1.1	5.26865	2.72131	6.9e-08	.003251
1.15	5.86592	2.12964	2.2e-09	.016601
1.2	6.43923	1.56369	6.0e-11	.058945
1.25	6.99061	1.02121	1.4e-12	.153577
1.3	7.52185	.500228	2.7e-14	.308457
1.35	8.03452	-.000992	4.4e-16	.500396
1.4	8.53002	.408573	0	.341427
1.45	9.0096	.874575	0	.190903
1.5	9.47437	1.32492	0	.092599
1.55	9.92533	1.76069	0	.039146
1.6	10.3634	2.18287	0	.014523
1.65	10.7893	2.59235	0	.004766
1.7	11.204	2.98993	0	.001395
1.75	11.6079	3.37634	0	.000367
1.8	12.0018	3.75225	0	.000088
1.85	12.3862	4.11826	0	.000019
1.9	12.7616	4.47492	0	3.8e-06
1.95	13.1285	4.82276	0	7.1e-07
2	13.4874	5.16223	0	1.2e-07

Note: gamma(1 (0.05) 2).

The following specifications don't change the result either:

- psmatch2 EPFRP, outcome(Logyield50) pscore(mypscore) neighbor(2) caliper(0.50) common quietly ate
- psmatch2 EPFRP, outcome(Logyield50) pscore(mypscore) neighbor(3) caliper(0.25) common quietly ate.