INTERNATIONAL FOOD
POLICY RESEARCH INSTITUTE
sustainable solutions for ending hunger and poverty

Supported by the CGIAR

IFPRI Discussion Paper 00733

December 2007

Impact of Soil Conservation on Crop Production in the Northern Ethiopian Highlands

Menale Kassie, Environmental Economics Policy Forum for Ethiopia and Ethiopian Development Research Institute

John Pender, International Food Policy Research Institute

Mahmud Yesuf, Environmental Economics Policy Forum for Ethiopia and Ethiopian Development Research Institute

Gunnar Kohlin, University of Göteborg

Randy Bluffstone, Portland State University

and

Elias Mulugeta, International Livestock Research Institute

Environment and Production Technology Division

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

The International Food Policy Research Institute (IFPRI) was established in 1975. IFPRI is one of 15 agricultural research centers that receive principal funding from governments, private foundations, and international and regional organizations, most of which are members of the Consultative Group on International Agricultural Research.

FINANCIAL CONTRIBUTORS AND PARTNERS

IFPRI's research, capacity strengthening, and communications work is made possible by its financial contributors and partners. IFPRI gratefully acknowledges generous unrestricted funding from Australia, Canada, China, Denmark, Finland, France, Germany, India, Ireland, Italy, Japan, the Netherlands, Norway, the Philippines, Sweden, Switzerland, the United Kingdom, the United States, and the World Bank.

INTERNATIONAL FOOD
POLICY RESEARCH INSTITUTE
sustainable solutions for ending hunger and poverty
Supported by the CGIAR

IFPRI Discussion Paper 00733

December 2007

Impact of Soil Conservation on Crop Production in the Northern Ethiopian Highlands

Menale Kassie, Environmental Economics Policy Forum for Ethiopia and Ethiopian Development Research Institute

John Pender, International Food Policy Research Institute

Mahmud Yesuf, Environmental Economics Policy Forum for Ethiopia and Ethiopian Development Research Institute

Gunnar Kohlin, University of Göteborg

Randy Bluffstone, Portland State University

and

Elias Mulugeta, International Livestock Research Institute

Environment and Production Technology Division

PUBLISHED BY

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

2033 K Street, NW
Washington, DC 20006-1002 USA
Tel.: +1-202-862-5600
Fax: +1-202-467-4439
Email: ifpri@cgiar.org

www.ifpri.org

Notices:

[1] Effective January 2007, the Discussion Paper series within each division and the Director General's Office of IFPRI were merged into one IFPRI-wide Discussion Paper series. The new series begins with number 00689, reflecting the prior publication of 688 discussion papers within the dispersed series. The earlier series are available on IFPRI's website at www.ifpri.org/pubs/otherpubs.htm#dp.

[2] IFPRI Discussion Papers contain preliminary material and research results. They have not been subject to formal external reviews managed by IFPRI's Publications Review Committee, but have been reviewed by at least one internal and/or external researcher. They are circulated in order to stimulate discussion and critical comment

Contents

List of Tables

List of Figures

ACKNOWLEDGMENTS

The authors gratefully acknowledge funding for this research from the International Food Policy Research Institute, the Swedish International Development Cooperation Agency (Sida)-supported Environment for Development Project of Göteborg University in Sweden, the Research Council of Norway, and the Norwegian Ministry of Foreign Affairs and the Swiss Agency for Development and Cooperation.

ABSTRACT

Land degradation, in the form of soil erosion and nutrient depletion, threatens food security and the sustainability of agricultural production in many developing countries. Governments and development agencies have invested substantial resources in promoting soil conservation practices, in an effort to improve environmental conditions and reduce poverty. However, very limited rigorous empirical work has examined the economics of adopting soil conservation technology. This paper investigates the impact of stone bunds[1] on crop production value per hectare in low and high rainfall areas of the Ethiopian highlands using cross-sectional data from more than 900 households having multiple plots per household. We use modified random effects models, stochastic dominance analysis (SDA) and matching methods to ensure robustness. The parametric regression and SDA estimates are based on matched observations obtained from nearest neighbor matching using propensity score estimates. This is important because conventional regression and SDA estimates are obtained without ensuring the existence of comparable conserved and non-conserved plots within the distribution of covariates. Here, we use matching methods, random effects and Mundlak's approach to control for selection and endogeneity biases that may arise due to correlation of unobserved heterogeneity and observed explanatory variables. The three methods used herein consistently show that plots with stone bunds are more productive than those without such technologies in semi-arid areas but not in higher rainfall areas, apparently because the moisture-conserving benefits of this technology are more beneficial in drier areas. This implies that the performance of stone bunds varies by agro-ecological type, suggesting a need for the design and implementation of appropriate site-specific technologies.

Keywords: Ethiopia; Soil conservation; crop production; agro-ecology; matching method; stochastic dominance, modified random effects model

---

[1] A bund is a barrier that prevents soil and water from escaping the plot.

1. INTRODUCTION

Land degradation, in the form of soil erosion and nutrient depletion, threatens food security and the sustainability of agricultural production in Sub-Saharan Africa (SSA). In response, governments and development agencies have invested substantial resources in promoting soil conservation practices as part of efforts to improve environmental conditions and ensure sustainable and increased agricultural production. However, there has been limited rigorous empirical work examining the economics of soil conservation. This paper attempts to partly close this gap by assessing returns to the use of stone bunds in high and low rainfall areas of the Ethiopian highlands.

Since the distribution and amount of rainfall obviously varies both in spatial and temporal terms across Africa, the distribution of rainfall should be considered when making a variety of decisions, including those related to soil conservation. Almost half of the area of Africa, which includes more than 14 percent of the low-income countries in the world, is arid or semi-arid, and over 90 percent of agricultural production is rain-fed (Fischer et al. 2004; WDI 2005). Since rainfall is often inadequate and there are extreme fluctuations in the availability of water, food production in these agro-ecological zones is a serious challenge (Fischer et al. 2004). Furthermore, climate change is causing rainfall variability in many African countries that are already at least partly semi-arid and arid. This is seriously affecting the sustainability and productivity of agriculture, and will continue to do so unless farm households adopt appropriate mitigation mechanisms to conserve the available rain, such as soil and moisture conservation (SMC) technologies (IPCC 2001).

Whether SMC technologies increase crop yields may depend on the agro-ecology and technology in question. Sutcliffe (1993) concluded that physical soil conservation activities are justifiable in moisture-stressed areas of the Ethiopian highlands, where moisture conservation plays an important role in increasing yield. This suggests that agro-ecological conditions may be a particularly important determinant of SMC profitability. However, despite the likely importance of rainfall patterns for determining whether SMC technologies improve the welfare of farming households, very little economic research has explicitly incorporated this feature. Indeed, similar soil and moisture conservation technologies, such as stone bunds, soil bunds and *fanya juu,*[2] are promoted in Ethiopia and many other countries without regard for the performance of these chronologies in different agro-ecologies.

This paper contributes to several aspects of the literature. First, earlier studies comparing plots with and without conservation failed to analyze plots that were similar in terms of observable characteristics. As a result, the analyses might have compared incomparable observations, possibly

---

[2] *Fanya juu* (a Swahili term meaning "to throw up") is a soil bund type wherein a ditch is dug along the contour and the soil is thrown up to form a ridge above; a natural bench terrace will subsequently form over the next few years.

leading to biased conclusions concerning the impacts of conservation (Heckman et al. 1998). Unlike previous econometric studies of the economic impacts of soil and moisture conservation measures, in which all observations were pooled and used (e.g. Shively 1998a and 1998b; Holden et al. 2001; Benin 2006; Kassie and Holden 2006; Pender and Gebremedhin 2006), our regression and stochastic dominance analysis estimates are based on matched observations. Second, our statistical tests reject the assumption of homogeneous impacts implicit in the econometric approaches used in other studies. We therefore use a switching regression approach that allows for differential impacts of covariates on conserved vs. non-conserved plots. Third, our data is cross-sectional with multiple plots per household, allowing us to control for the influence of unobservable household characteristics on outcomes. We control for plot quality characteristics using a detailed dataset that contains measures of many potentially important plot quality characteristics. Finally, we compare the performance of stone bunds on crop production value in high rainfall (Amhara region) and low rainfall (Tigray region) areas of the Ethiopian highlands, providing new insight into how the type of agro-ecology affects the performance of soil conservation measures.

The paper is organized as follows. Section 2 gives a brief review of previous empirical works related to our estimation methods. Following a discussion on econometric methodology in section 3, a description of the dataset is presented in section 4. The empirical results are presented in section 5. Finally, section 6 summarizes the main findings and concludes the paper.

2. LITERATURE REVIEW

Few empirical studies have used econometric and cross-sectional data to directly examine the impacts of soil conservation measures on mean yield in developing countries (e.g. Byiringrio and Reardon 1996; Shively, 1999; 1998a; 1998b; Holden, Shiferaw and Pender 2001; Bekele 2003; Kaliba and Rabele 2004; Benin 2006; Kassie and Holden 2006; Pender and Gebremedhin 2006). Byiringiro and Reardon (1996), using farm-level data in Rwanda, found that farms with greater investments in soil conservation had much greater land productivity than other farms. The type of conservation, however, was not specified. In the Philippines, Shively (1998a; 1998b) found that conservation via contour hedgerows had a positive and statistically significant impact on yield, as assessed using farm-level data. In Lesotho, Kaliba and Rabele (2004) found that short- and long-term soil conservation measures had significantly positive effects on wheat yield.

Several studies have estimated the impacts of soil and moisture conservation (SMC) measures in the Ethiopian highlands using econometric analysis of cross-sectional survey data. In one peasant association of northern Ethiopia (Andit Tid in the North Shewa zone of the Amhara region), Holden, Shiferaw and Pender (2001) found that SMC measures (soil bunds and *fanya juu* terraces) had statistically insignificant impacts on land productivity in all regressions (but a negative coefficient in almost all regressions). Benin (2006), based on a survey of 434 households representing the highlands of the Amhara region as a whole, found that stone terraces had significantly positive impacts (a 42 percent increase) on average crop yields in a reduced form regression for lower-rainfall parts of the Amhara region, but insignificant impacts in a structural regression. This suggested that the impacts of stone terraces were due to their ability to enable more productive input use. In contrast, Benin found that stone terraces had insignificant impacts on yields in the high rainfall parts of the region. Kassie and Holden (2006) found that physical conservation measures (*fanya juu*) resulted in lower yield in a high rainfall area of the Ethiopian highlands in western Amhara, compared to non-conserved plots. Finally, Pender and Gebremedhin (2006) conducted a survey of 500 households representing the semi-arid highlands of Tigray, and found higher crop yields from plots with stone terraces (an average yield increase of 23 percent), and estimated the average rate of return to stone terrace investment to be 46 percent.

Studies based on farm-level data from the Ethiopian highlands have also found differences in yields and economic returns between different agro-ecologies. Shiferaw and Holden (2001) and Gebremedhin et al. (1999) used a cost-benefit analysis to evaluate the economic benefits of structural and biological conservation in different parts of the northern Ethiopian highlands. Shiferaw and Holden (2001), using data from experimental trials conducted by the Soil Conservation Research program in two high-rainfall highland sites (Anjeni in western Amhara and Andit Tid in eastern Amhara) of Ethiopia concluded that structural technologies (graded bund and *fanya juu* terraces) have very low payoffs and do not seem to offer poor farmers sufficient economic incentives to pay for the necessary investments. The authors showed that investment in grass strips appeared promising (yielding a positive net present value), but only in Anjeni. In contrast, Gebremedhin et al. (1999) estimated that stone terraces yielded a 50 percent rate of return, based on experimental evidence collected in the semi-arid central Tigray region.

These experimental results, consistent with the results of econometric analysis of household survey data (see above), suggest that the economic returns to SMC investments are greater in lower rainfall areas, such as central Tigray, than in higher rainfall areas, such as western Amhara. However, the prior experimental studies are limited in their coverage of different situations in Ethiopia and suffer from a lack of important data, such as the absence of data on conventional inputs. Furthermore, it is not clear whether the observations with and without conservation technology were strictly comparable. The prior econometric studies also suffer from several econometric shortcomings, as discussed further below.

Some studies have used stochastic dominance analysis to assess the impacts of SMC measures on yield distribution. Using non-experimental farm-level data collected in the Philippines, Shively (1999) compared observed yields obtained from farmers' fields with and without contour hedgerows, and found that the use of hedgerow technology did not constitute an unambiguously dominant production strategy. Bekele (2003), using results from experimental trials of the Soil Conservation Research Project in a low rainfall area of eastern Ethiopia, found that physical conservation (level bunds) had an unambiguous dominance over the no-conservation condition. Kassie and Holden (2006) used cross-sectional farm-level data from a high rainfall area in northwestern Ethiopia to show that yield distributions without conservation unambiguously dominated yield distributions with conservation (graded *fanya juu*) for all yield levels. Again, these results from the Ethiopian highlands suggest that SMC measures perform better in lower rainfall environments.

These studies, however, suffered from a number of methodological problems that may have led to under- or over-estimation of the productivity impacts of the analyzed technologies. First, some of the comparisons were not based on comparable observations, which can yield biased estimates (Heckman et al. 1998). Second, all of the prior studies assumed a single equation model in which technology had only intercept effects, and the same set of variables was taken as equally affecting both technology adopters and non-adopters. These assumptions were not tested empirically. Third, except for Shively (1998b; 1999) and Kassie and Holden (2005), none of the studies accounted for the endogeneity of the technology and the self-selection problem. Fourth, none of the studies accounted for unobserved heterogeneity, which might have affected their findings. Kaliba and Rabele's (2004) study suffered from a small sample size (50 households) and did not control for plot characteristics variables. If plot quality is asymmetrically distributed across plots and households, and conservation correlates with plot quality, estimation of conservation impact on yield without controlling for these factors may lead to inconsistent results. Furthermore, some studies used only a partial cost-benefit analysis that did not capture the effects of important variables, such as conventional inputs and household characteristics and endowments.

3. METHODOLOGY: ESTIMATION CHALLENGES, TECHNIQUES, AND PROCEDURES

Estimation Challenges and Techniques

It is difficult to assess productivity gains from soil conservation based on non-experimental observations, because the counterfactual outcome of what production would have been without conservation on conserved plots is not observed. In experimental studies, this problem is addressed by randomly assigning plots to treatment and control status, which assures that the outcomes observed on the control plots without conservation are statistically representative of what would have occurred without conservation on the treatment plots. However, in real farming situations, farmers and plots are not randomly assigned to the two groups (adopters and non-adopters), but rather make their own adoption choices, or are systematically selected by development agencies based on their propensity to participate in technology adoption. In addition, farmers (or development agencies) are likely to select plots non-randomly based on their quality attributes, which are often unobservable by the researcher. Therefore, adopters and non-adopters may be systematically different, and conserved and non-conserved plots may also be systematically different, and these differences may manifest themselves in differences in farm performance that could be mistakenly attributed to differences in adoption behavior. Thus, it is difficult to perform ex-post assessment of gains from conservation using observational data, because of possible selection bias due to observed and unobserved plot and household characteristics. Failure to account for this potential selection bias could lead to inconsistent estimates of the impact of technology adoption.

Two-step Heckman (Heckman 1979) and matching (Heckman and Robb 1985) approaches are possible solutions to the selection problem. The Heckman two-step approach assumes that selection is affected by unobservable variables ("selection on unobservables"), whereas the Heckman and Robb approach assumes selection on observables. The Heckman two-step approach addresses selection on unobservables by imposing distributional and functional form assumptions (e.g., the outcome equation is usually linear) and extrapolating over regions having no common support, i.e. lacking similar conserved and non-conserved plots observations. In such cases, the findings of Heckman et al. (1998), Dehejia and Wahba (1999; 2002) and Smith and Todd (2005) suggest that selection bias may be reduced by avoiding functional form assumptions and imposing a common support condition.

Conventional regression and stochastic dominance analysis estimates are usually obtained without ensuring that the treated and non-treated observations are actually comparable in terms of their covariate distributions (i.e. they may lack common support), possibly resulting in substantial biases (Heckman et al. (1998); Dehejia and Wahba (1999; 2002); Smith and Todd (2005). To deal with this problem, we based our regression and stochastic dominance analyses on matched samples of conserved and non-conserved plots, using propensity score matching to select the matched samples (Rosenbaum and Rubin 1983). The basic idea of the propensity score matching approach is to match observations of conserved and non-conserved plots according to the predicted propensity of the plots to have conservation. The resulting comparisons occur between conserved and non-conserved plots having characteristics that are similar and relevant to the technology choice. This reduces the potential for bias arising from the comparison of non-comparable observations (in terms of observables), although selection bias may still be caused by differences in unobservables.

The other econometric issue is that even if no selection bias problem exists or we can account for the selection process, it may be inappropriate to use a pooled sample of adopters and non-adopters (i.e. a dummy regression model wherein a binary indicator is used to assess the effect of soil conservation on productivity). This is because pooled model estimation assumes that the set of covariates has the same impact on adopters and non-adopters (i.e. common slope coefficients for both regimes). This implies that soil conservation has only an intercept shift effect, which is always the same irrespective of the values taken by other covariates that determine yield. However, a Chow test of equality of coefficients for the adopters and non-adopters of stone terraces in our sample rejected the equality of non-intercept coefficients [$\chi^2(63) = 142.51$, $p= 0.000$ and $\chi^2(68) = 147.49$, $p = 0.000$ for the Amhara and Tigray regions, respectively].[3] This supports the idea that it may be helpful to use a regression approach that differentiates each coefficient for adopters and non-adopters.

We herein use parametric switching regression and non-parametric techniques to overcome the econometric problems and assess the robustness of our results. The non-parametric methods include stochastic dominance analysis and propensity score matching. The parametric regression equation to be estimated using multiple plots per household is:

$$1)\quad \begin{cases} y_{hp1} = x_{hp}\beta_1 + u_h + e_{hp1} \text{ if } C_{\mathrm{hp}} = 1 \\ y_{hp0} = x_{hp}\beta_0 + u_h + e_{hp0} \text{ if } C_{\mathrm{hp}} = 0 \end{cases}$$

where y_{hp} is the value of crop production per ha obtained by household h on plot p, depending on its conservation status (C_{hp}); u_h captures unobserved household characteristics that affect productivity, such as farm management ability, land fertility, etc.; e_{hp} is the random variable that summarizes the effects of plot-specific unobserved components on productivity, such as unobserved variation in plot quality and plot-specific production shocks (e.g. plot-level variation in rainfall, frost, floods, weeds, pests

---

[3] The result is the same excluding the conventional inputs (fertilizer, seed, labor and oxen use per ha). The Chow test statistics for this case are [$\chi^2(53) = 125.14$, $p = 0.0000$ and $\chi^2(60) = 111.11$, $p = 0.0001$] for the Amhara and Tigray region datasets, respectively. Although not reported, similar results were found without Mundlak's approach (more on this issue later).

and disease); x_{hp} includes both plot-specific and household-specific observed explanatory variables, and β is a vector of the parameters to be estimated.

Estimation Procedures

To obtain consistent estimates of the effects of conservation, we need to control for unobserved heterogeneity (u_h) that may be correlated with observed explanatory variables. One way to address this issue is to exploit the panel nature of our data (repeated cross-sectional plot observations per household), and use household-specific fixed effects. The main shortcoming of using fixed effects in this case is that many of the surveyed households have only a single plot, and therefore do not play a role in a fixed-effects analysis. Random effects and pooled ordinary least square (OLS) models are consistent only under the assumption that unobserved heterogeneity is uncorrelated with the explanatory variables. As an alternative, we herein use the modified random effects model framework proposed by Mundlak (1978), whereby the right hand-side of each equation includes the mean value of the plot-varying explanatory variables. Mundlak's approach relies on the assumption that unobserved effects are linearly correlated with explanatory variables as specified by:

$$2)\ u_h = \bar{x}\gamma + \eta_h,\ \eta_h \sim \text{iid}(0, \sigma_\eta^2)$$

where $\bar{x}$ is the mean of the plot-varying explanatory variables within each household (cluster mean), γ is the corresponding vector coefficient and η is a random error unrelated to $\bar{x}'s$. In our case, it is most important to include average plot characteristics, such as average plot fertility, soil depth, slope, and conventional input use, which we believe have large impacts on production and technology adoption decisions. The vector γ will be equal to zero if the observed explanatory variables are uncorrelated with the random effects.

The selection process in the parametric switching regression model can be addressed using the selectivity terms (inverse Mills' ratio) derived from the criterion equation (probit model), which addresses the problem of selection on unobservables. However, the criterion models turned out to be insignificant [i.e. the result of the overall model significance test (Wald chi^2) is insignificant] for both with- and without Mundlak's approach as well as for both regions. This is perhaps not surprising since we use matched samples (based on observable variables) obtained from a nearest neighbor matching method using estimated propensity scores. The inverse Mills' ratio derived from such insignificant models, assuming functional form identification (nonlinearity of the first-step probit estimators), is marginally significant (at 10 percent) for only two of the 16 models estimated. This suggests that by addressing selection on observables using propensity score matching, we have also reduced problems with selection on unobservables.[4] Therefore, we do not include regression results with the inverse Mills' ratio (endogenous switching regression), but rather provide the predicted mean output value differences in order to save space. The regression results are available from the authors.

The selection process and endogeneity bias can also be addressed using the panel nature of our data and Mundlak's approach, if the selection and endogeneity bias are due to plot-invariant unobserved factors, such as household heterogeneity (Wooldridge 2002). If the unobserved plot component (e_{hp}) is correlated with the decision to adopt stone bunds and other observed regressors, the parameter estimates from equation (1) will be inconsistent. If we fail to control for these factors, we will not obtain the true effect of conservation.

Controlling for plot heterogeneity is a bit more difficult than addressing household heterogeneity. Fortunately, our dataset offers a richer characterization of plot quality than that found in most of the other studies. It is likely that observed plot quality would be positively correlated with unobserved plot quality. In terms of plot characteristics, the dataset includes plot slope, position on slope, plot size, soil fertility, soil depth, soil color, soil texture, presence of gullies, plot distance from the homestead, rainfall, altitude, and input use. Including these variables in our model allows us to address selection due to idiosyncratic errors (e.g. plot heterogeneity) using observed plot quality characteristics and inputs (Fafchamps 1993; Levinsohn and Petrin 2003; Assunção and Braido 2004). The use of input use to control for plot heterogeneity is based on the notion that farmers may respond to positive and negative shocks by increasing or decreasing their input use. With regard to using the matching method, matching on every covariate is difficult to implement when the set of covariates is large. To overcome the curse of dimensionality, Rosenbaum and Rubin (1983) showed that if matching on vector x_{hp} is valid, so is matching on the propensity score. This allows us to match on a single index rather than on the multidimensional x_{hp} vector.

Our main goal in the matching method is to identify the average treatment effect on the treated (ATT) plots, and obtain matched treated and non-treated observations. This is achieved using a two-step procedure. In the first step, we use a probit model to estimate the propensity score, which is defined as the conditional probability that plot p receives conservation treatment given the covariates. In the second stage, we use nearest neighbor (NN) matching based on propensity score estimates as an input to obtain the ATT. The NN matching method, unlike other weighted matching methods such as kernel matching,

---

[4] Weak identification of the selectivity correction terms (resulting from using mostly the same variables in both the selection and productivity equations) could also be responsible for this statistical insignificance. Thus, it is more precise to claim that we do not have evidence that suggests selection on unobservables is affecting our regression results, rather than saying that the problem is clearly solved. Without experimental evidence, it is difficult to know whether such problems are fully solved.

allows us to identify the specific matching observations that entered the ATT calculation, and which will be used for parametric regressions and stochastic dominance analysis.

Matching methods assume that the selection process is based only on observable characteristics (i.e. conditional independence). To adjust for unobservables, we include the means of the plot-varying covariates following Mundlak's approach and Wooldridge's (1995) panel data sample selection estimation approach. Controlling for the above econometric problems and incorporating equation (2) into (1), the expected yield difference between adoption and non-adoption of stone bunds is estimated as:

$$3)\; E(y_{hp1} \big| x_{hp}, u_h, C_{hp} = 1) - E(y_{hp0} \big| x_{hp}, u_h, C_{hp} = 1) = x_{hp}(\beta_1 - \beta_0) + \bar{x}(\gamma_1 - \gamma_0).$$

The second term on the left-hand side of (3) is the expected value of y if the plot had not received soil conservation (counterfactual outcome), which will be approximated by non-conserved plot observations after we account for the selection process. This is our parameter of interest in the parametric regression analysis. Equation (3) is also estimated without including the second term of the right-hand side equation (without Mundlak's approach), for comparison purposes and to generate a greater degree of confidence in the robustness of the econometric results.

Finally, we check for multicollinearity (MC) and non-linearity problems for all regression models. MC is inevitable for regression analyses using the mean of plot-varying explanatory variables, although it is not a problem if the goal is simply to predict a dependent variable from a set of explanatory variables, such as the case here, where our objective in the parametric regression is to assess the predicted mean yield impacts of stone bunds. MC is not a problem for the regression model without Mundlak's approach, as the maximum variance inflation factor (VIF) is less than 10. Graphical (augmented component residual plot) and statistical (ovtest) tests indicate that non-linearity is not a problem in our regression analysis. The STATA software package was used for the regression analysis, the results were corrected for clustering, (possible non-independence of the errors in observations of multiple plots from the same household) and bootstrapped standard errors were used in the propensity score matching and endogenous switching regression models to account for the additional error caused by the two-stage nature of the estimation.

4. DATA SOURCES AND TYPES

The data used in this study come from a farm survey conducted in 1999 and 2000 in the Tigray and Amhara regions of Ethiopia. The analyzed plots were all located in the highlands above 1500 meters above sea level. The Amhara region dataset includes 435 farm households, 98 villages, 49 *kebeles*[5] and about 1365 plots, after removal of missing observations for some variables. The Tigray dataset includes 500 farm households, 100 villages, 50 *kebeles* and 965 plots after deletion of missing observations.[6] Using the nearest neighbor matching method based on propensity score estimates and Mundlak's approach, we obtain a sample of 382 (232 conserved and 150 non-conserved) and 573 (390 conserved and 183 non-conserved) plots in the Amhara and Tigray regions, respectively. Without using Mundlak's approach, we obtain a sample of 391 (232 conserved and 159 non-conserved) and 590 (390 conserved and 190 non-conserved) plots in the Amhara and Tigray regions, respectively.

Tables 1 and 2 present descriptive statistics by region for the two sub-samples before and after matching, along with those for the conserved and non-conserved plots after matching.

Table 1. Descriptive statistics of variables for the Tigray region, with Mundlak's approach*

Variables	Mean 1	Mean 2	Mean 3	Mean 4
Value of crop production, ETB/ha	1809.982	1614.470	1670.133	1495.844
	(2445.045)	(1715.694)	(1718.317)	(1708.725)
Plot size, ha	0.271	0.284	0.305	0.239
	(0.226)	(0.216)	(0.224)	(0.191)
Other plot size (total farm size - plot size)	0.813	0.793	0.795	0.789
	(0.774)	(0.870)	(0.993)	(0.519)
Middle slope position	0.221	0.281	0.310	0.219
Bottom slope position	0.245	0.276	0.272	0.284
Not on slope	0.422	0.295	0.262	0.366
Deep soil plots	0.374	0.356	0.359	0.350
Medium soil plots	0.417	0.485	0.495	0.464
Brown soil plots	0.147	0.176	0.192	0.142
Gray soil plots	0.231	0.264	0.254	0.284
Red soil plots	0.389	0.384	0.382	0.388
Gently sloped plots	0.305	0.393	0.415	0.344
Steeply sloped plots	0.091	0.136	0.146	0.115
Loam soil plots	0.357	0.414	0.421	0.399
Clay soil plots	0.305	0.304	0.300	0.311
Sandy soil plots	0.108	0.115	0.118	0.109
Moderately eroded plots	0.280	0.328	0.356	0.268
Severely eroded plots	0.065	0.094	0.097	0.087
Fenced plots	0.048	0.056	0.056	0.055
Gully plots	0.035	0.042	0.041	0.044

---

[5] A *kebele* is a higher administrative unit than a village, usually constituted of three or four villages, and is often translated as a "peasant association."

[6] For more details on the study areas, sampling techniques and criteria used to select the sample areas, please see Pender and Gebremedhin (2006) and Benin (2006).

Table 1. Continued

Variables	Mean 1	Mean 2	Mean 3	Mean 4
Plot distance to residence, walking hrs	0.315	0.303	0.281	0.350
	(0.365)	(0.388)	(0.375)	(0.411)
Household altitude, masl	2176.428	2163.726	2170.431	2149.437
	(339.661)	(320.252)	(317.159)	(327.166)
Fertilizer use, kg/ha	41.682	39.878	43.408	32.353
	(97.997)	(102.641)	(114.779)	(69.775)
Seed use, kg/ha	157.134	129.370	129.279	129.564
	(244.867)	(145.267)	(141.872)	(152.650)
Labor use, days/ha	75.845	64.861	63.964	66.771
	(112.606)	(54.914)	(55.658)	(53.392)
Oxen use, days/ha	29.652	26.873	27.082	26.428
	(35.820)	(14.989)	(16.237)	(11.929)
Rented plots	0.133	0.096	0.097	0.093
Reduced tillage plots	0.124	0.120	0.131	0.098
Irrigated plots	0.036	0.009	0.008	0.011
Residence distance to market, walking hrs	2.872	3.137	3.132	3.150
	(2.320)	(2.529)	(2.405)	(2.781)
Male household head	0.902	0.911	0.913	0.907
Household head age	48.398	49.349	49.628	48.754
	(12.672)	(12.296)	(12.173)	(12.566)
Family size, number	5.997	6.084	6.138	5.967
	(2.065)	(1.997)	(1.987)	(2.019)
Education between grade one & two	0.079	0.087	0.092	0.077
Education above grade three	0.056	0.051	0.046	0.060
Oxen holding, number	1.418	1.358	1.326	1.426
	(0.912)	(0.867)	(0.869)	(0.860)
Other cattle, number	3.605	3.234	3.185	3.339
	(3.660)	(3.352)	(3.438)	(3.168)
Small ruminant, number	5.876	5.874	5.708	6.230
	(9.014)	(8.604)	(8.523)	(8.788)
Pack animals, number	0.997	0.902	0.926	0.852
	(1.493)	(1.378)	(1.513)	(1.035)
Mean annual rainfall, mm	649.783	652.657	653.736	650.357
	(100.883)	(95.923)	(93.003)	(102.093)
Population density, /km^2	141.640	148.858	151.729	142.741
	(69.559)	(72.901)	(75.156)	(67.636)
Number of observations	935	573	390	183

Mean1 = Refers to mean and standard deviations (sd) of variables from total sample before matching
Mean2= Refers to mean and standard deviations of variables from total matched sample
Mean3 = Refers to mean and standard deviations of variables from matched sample with conservation
Mean4 = Refers to mean and standard deviations of variables from matched sample without conservation
* Standard deviations are not reported for dummy variables. We did not find statistically significant differences in input use between conserved and non-conserved plots.

Table 2. Descriptive statistics of variables for Amhara region, with Mundlak's approach*

Variables	Mean 1	Mean 2	Mean 3	Mean 4
Value of crop production, ETB/ha	2232.405	1683.896	1582.921	1840.071
	(4007.41)	(1368.983)	(1117.251)	(1678.262)
Male household head	0.954	0.982	0.983	0.980
Family size, number	6.817	6.605	6.591	6.627
	(2.567)	(2.142)	(2.156)	(2.125)

Table 2. Continued

Variables	Mean 1	Mean 2	Mean 3	Mean 4
Household head age	44.789	46.644	47.151	45.860
	(12.408)	(12.592)	(12.183)	(13.204)
Livestock holding, TLU	2.931	2.526	2.332	2.825
	(2.416)	(1.938)	(1.774)	(2.139)
Education level	2.592	2.534	2.603	2.427
	(3.375)	(3.246)	(3.186)	(3.345)
Residence distance to market, walking hrs	2.322	2.956	2.892	3.055
	(3.030)	(3.214)	(2.693)	(3.893)
Plot slope, degree	5.552	8.042	8.034	8.053
	(6.003)	(7.122)	(5.609)	(8.996)
Black soil plots	0.308	0.380	0.358	0.413
Brown soil plots	0.277	0.335	0.362	0.293
Gray soil plots	0.070	0.089	0.086	0.093
Deep soil plots	0.240	0.147	0.138	0.160
Medium soil plots	0.536	0.586	0.591	0.580
Moderately eroded plots	0.304	0.526	0.556	0.480
Severely eroded plots	0.098	0.113	0.108	0.120
Clay soil plots	0.122	0.073	0.078	0.067
Loam soil plots	0.432	0.361	0.362	0.360
Sandy soil plots	0.117	0.154	0.147	0.167
Highly fertile plots	0.108	0.045	0.039	0.053
Medium fertile plots	0.694	0.702	0.724	0.667
Plot distance to residence, walking hrs	0.283	0.274	0.254	0.305
	(0.807)	(0.340)	(0.276)	(0.421)
Plot distance to main road, walking hrs	1.301	1.442	1.437	1.449
	(1.028)	(1.122)	(1.135)	(1.106)
Gully plots	0.045	0.063	0.052	0.080
Plot altitude, masl	2344.017	2395.906	2377.013	2425.127
	(468.414)	(485.541)	(454.888)	(529.724)
Population density, persons/km^2	143.897	141.832	141.328	142.613
	(84.268)	(89.612)	(87.558)	(92.992)
Mean annual rainfall, mm	1980.048	1899.536	1851.093	1974.461
	(592.056)	(645.389)	(601.234)	(703.814)
Fertilizer use, kg/ha	92.196	47.124	39.732	58.557
	(209.972)	(125.669)	(115.928)	(139.034)
Seed use, kg/ha	152.800	147.893	186.232	188.596
	(730.379)	(1034.306)	(1323.488)	(118.622)
Labor use, days/ha	123.106	96.233	94.613	98.740
	(227.831)	(101.152)	(98.510)	(105.393)
Oxen use, day/ha	56.555	48.103	45.878	51.545
	(64.738)	(46.480)	(41.515)	(53.222)
Rented plots	0.108	0.050	0.052	0.047
Plot size, ha	0.385	0.418	0.431	0.398
	(0.350)	(0.274)	(0.276)	(0.271)
Other plot area (total farm size- plot size)	1.396	0.968	0.962	0.978
	(1.127)	(0.625)	(0.617)	(0.639)
Number of observations	1320	382	232	150

* See note for Table 1.

We do not find statistically significant differences in input use between conserved and non-conserved plots, with one exception: fertilizer use is significantly higher (to a marginal degree at 10%) on non-conserved plots.

In the interest of space, we do not show or discuss the descriptive statistics for the data without Mundlak's approach, but these results are available upon request. About 37 percent of the sample plots in Tigray and 17 percent of the sample plots in the Amhara region had stone bunds. Soil bunds are also used on some of the plots, but there are not enough observations to run parametric regressions on soil bund-containing plots.

There are two dominant sources of investment supporting the construction of stone bunds: private investments and labor mass mobilization campaigns. About 64 percent of the investments in stone bunds in Amhara between 1991 and 2000 and 37 percent of the investments in stone bunds in the Tigray region during 1997 to 1999 were private investments, often promoted by government extension workers[7] and peasant association officials. In addition, officials also mobilized community labor for construction of 29 and 55 percent of the stone bunds found in the Amhara and Tigray regions, respectively ("mass mobilization investment"). In Tigray, a combination of private and mass mobilization investment was used for 4 percent of the investments in stone bunds during 1997 to 1999. The rest were built through other investment sources, such as food-for-work. We observe no statistically significant mean yield difference between plots having private vs. mass mobilization investments; the mean differences between the values of outputs from plots with private and mass mobilization investments were Ethiopian Birr (ETB) 36 (se = 167) per ha and ETB 73 (se = 312) per ha in the Amhara and Tigray regions, respectively. However, the estimated differences in mean yield are positive in favor of private investments.

The mean plot altitude, which is associated closely with temperature and microclimate, is 2176 and 2344 meters above sea level for the Tigray and Amhara regions, respectively. The average annual rainfall in Amhara is about 1980 mm per year and that for Tigray is 650 mm.[8] The rainfall in our Amhara study sites therefore averages approximately three times that of Tigray, and the rainfall differences across the two regions are very large. The mean population density, however, is similar across the two regional sub-samples at 142 to 144 persons per square kilometer. Fertilizer use averages about 40 kg per ha in Tigray and 47 kg/ha in Amhara for matched sample households. Although multivariate analysis is important to compare input use on conserved and non-conserved plots, a two-sample t test failed to find a

---

[7] In Amhara, the survey asked about investments made on the plot between 1991 and 2000, including the source (e.g. private investment, labor mass mobilization campaigns, food-for-work, or other), amount and cost of investments. In Tigray, the survey asked for comparable information during the years 1997, 1998 and 1999. These figures reflect the flows of recent investments and do not fully account for the entire stock of SMC investments on the surveyed plots, since many plots had SMC investments by 1991 in Amhara and by 1997 in Tigray. Hence, it is not possible to estimate the economic impacts of SMC investments by the source of investment, since we do not have this information for all plots. We also do not have information on the age of the stocks of investments, and in many cases, the age of investments may vary even on the same plot, since additional investments on particular plots may occur from year to year.

[8] The mean rainfall data are based on long-term rainfall averages, spatially interpolated using a climate model (Corbett and White 2001). The minimum and maximum rainfall averaged over the Amhara region for the last fifty years (1953-2003) was 1303 and 2457 mm, respectively. Even the minimum average rainfall in Amhara is higher than the maximum annual rainfall (994 mm) of the drier region, Tigray.

statistically significant difference (at the 5 percent level) in input use between conserved and non-conserved plots.[9]

In addition to these variables, plot characteristics, household endowments and indicators of access to infrastructure are included in the empirical model based on the guidance of economic theory and previous empirical research. In the presence of missing and/or imperfect markets, a household's initial resource endowments and characteristics may play a role in investment and production decisions (Holden et al. 2001; Pender and Kerr 1998) and are therefore included.

In the empirical model, variables such as inputs (e.g. fertilizer use and improved seed use per ha) are potentially endogenous variables.[10] We do not believe this is a problem, however, because explanatory variables covering input use are also included. In addition, our modified random effects estimation approach helps control for unobserved effects that may correlate with input use decisions.[11] Nevertheless, to assure robustness we estimate the parametric switching regression models both with (structural model) and without (reduced model) these variables.

---

[9] Fertilizer use in Amhara was higher on non-conserved plots, with marginal statistical significance. We would like to thank an anonymous reviewer for raising this important issue.

[10] Conventional inputs are not included in the propensity score matching procedure, since the matching procedure requires inclusion of variables that simultaneously affect the adoption decision and the outcome variable (agricultural productivity) (Heckman et al. 1998). We do not expect input use per ha to influence long-term investments; however, soil conservation may influence input use.

[11] Traditionally, farm households retain their own seeds from the previous harvest for use in the following season. Seed use is therefore a pre-determined variable. Improved seeds were used only on 3 and 1 percent of the sample plots in the Tigray and Amhara regions, respectively. We assume labor and oxen use are fixed in the short-term, since households usually depend on family resources because of limited labor and oxen markets.

5. RESULTS AND DISCUSSION

Matching Method: Propensity Scores and Nearest Neighbor Matching Estimates

The objectives of the propensity score matching used herein are to estimate the ATT, obtain matched treated and non-treated observations, and use them as inputs for switching regression and SDA. The propensity score matching method was estimated with and without Mundlak's approach for comparison purposes, although the statistical evidence found in the correlation between the observed explanatory variables and unobserved effects (Tables 3 and 4) suggests that ignoring this might lead to biased estimates. The results of the probit models used to estimate the propensity scores for conservation investments in Tigray and Amhara regions are reported in Tables 3 and 4. Table 5 provides the NN matching method estimates of the ATT for crop yields in the two regions (with and without using Mundlak's approach).

Table 3. Propensity score estimates of stone bund adoption in the Tigray region

Explanatory variables	With Mundlak's approach	Without Mundlak's approach
Deep soil plots	-0.046	0.173
	(0.164)	(0.130)
Medium Soil plots	0.027	0.309**
	(0.171)	(0.133)
Gently sloped plots	0.265	0.350***
	(0.180)	(0.129)
Steeply sloped plots	0.356	0.438**
	(0.278)	(0.195)
Brown soil plots	-0.074	0.221
	(0.267)	(0.196)
Gray soil plots	-0.228	0.024
	(0.287)	(0.192)
Red soil plots	-0.405	-0.238
	(0.257)	(0.182)
Loam soil plots	0.117	0.257
	(0.248)	(0.182)
Clay soil plots	0.344	0.233
	(0.260)	(0.184)
Sandy soil plots	0.511	0.434**
	(0.319)	(0.222)
Moderately eroded plots	0.146	0.152
	(0.156)	(0.110)
Severely eroded plots	0.217	0.234
	(0.285)	(0.202)
Plot distance from residence	-0.747***	-0.434***
	(0.203)	(0.137)
Rented plots	-0.282	-0.385***
	(0.187)	(0.143)
Reduced tillage plots	0.201	0.101
	(0.225)	(0.147)
Gully plots	0.025	0.017
	(0.354)	(0.259)

Table 3. Continued

Explanatory variables	With Mundlak’s approach	Without Mundlak’s approach
Ln(plot size)	0.593***	0.329***
	(0.126)	(0.062)
Ln(other plot size)	0.306	-0.058
	(0.274)	(0.060)
Irrigated	-0.913**	-0.872**
	(0.435)	(0.365)
Fenced plots	-0.054	0.130
	(0.274)	(0.211)
Middle slope position	0.219	0.006
	(0.247)	(0.173)
Bottom slope position	0.193	-0.090
	(0.245)	(0.172)
Not on slope	-0.303	-0.454***
	(0.274)	(0.175)
Population density	0.003***	0.003***
	(0.001)	(0.001)
Ln(rainfall)	0.800**	0.757**
	(0.397)	(0.352)
Ln(altitude)	0.149	0.460
	(0.415)	(0.368)
Distance to market	0.023	0.045*
	(0.026)	(0.023)
Male household head	0.080	0.155
	(0.180)	(0.164)
Ln(household head age)	0.400**	0.285
	(0.175)	(0.004)
Ln(family size)	0.229*	0.042*
	(0.131)	(0.024)
Education grades one & two	0.414**	0.305*
	(0.183)	(0.174)
Education above grade three	0.061	-0.058
	(0.229)	(0.219)
Joint chi^2 test for significance of mean of plot-varying explanatory variables (vector γ)	44.45***	
Constant	-9.596**	-9.622**
	(4.480)	(3.963)
Wald chi^2	280.893***	227.750***
Pseudo R^2	0.2211	0.1793
Number of observations	935	935

* $p<0.10$, ** $p<0.05$, *** $p<0.01$
Standard errors are in parenthesis

Table 4. Propensity score estimates of stone bund adoption in the Amhara region

Explanatory variables	With Mundlak's approach	Without Mundlak's approach
Plot slope in degrees	0.029**	0.029***
	(0.013)	(0.008)
Black soil plots	0.118	0.293**
	(0.185)	(0.131)
Brown soil plots	-0.087	0.252*
	(0.195)	(0.130)
Gray soil plots	-0.333	0.171
	(0.260)	(0.188)
Deep soil plots	0.002	-0.074
	(0.235)	(0.188)
Medium soil plots	0.155	0.071
	(0.170)	(0.125)
Moderately eroded plots	0.545***	0.646***
	(0.147)	(0.106)
Severely eroded plots	0.127	0.218
	(0.242)	(0.172)
Clay soil plots	-0.601**	-0.226
	(0.247)	(0.179)
Loam soil plots	-0.120	-0.133
	(0.163)	(0.109)
Sandy soil plots	-0.274	-0.138
	(0.208)	(0.150)
Highly fertile plots	-0.123	-0.386
	(0.308)	(0.238)
Medium fertile plots	0.355**	0.058
	(0.179)	(0.128)
Plot distance to residence	-0.210	-0.235*
	(0.175)	(0.140)
Gully plots	0.089	-0.193
	(0.305)	(0.222)
Rented plots	0.044	-0.183
	(0.239)	(0.191)
Reduced tillage plots	-0.293	0.348***
	(0.315)	(0.119)
Plot altitude	0.000	0.388
	(0.000)	(0.330)
Irrigated plots	-0.344	-0.651**
	(0.388)	(0.296)
Ln(plot size)	0.698***	0.298***
	(0.221)	(0.072)
other plot area	0.746	-0.361***
	(0.492)	(0.072)
Residence distance to market	0.093	0.068
	(0.236)	(0.048)
Male household head	0.683**	0.743***
	(0.291)	(0.278)
Family size	-0.009	-0.020
	(0.024)	(0.022)
Household head age	0.012***	0.013***
	(0.005)	(0.004)
Education level	0.024	0.015
	(0.016)	(0.015)
Population density	-0.002**	-0.162*

Table 4. Continued

Explanatory variables	With Mundlak's approach	Without Mundlak's approach
	(0.001)	(0.091)
Ln(rainfall)	-1.344***	-1.074***
	(0.246)	(0.195)
Joint chi^2 Significance of mean of plot-varying explanatory variables (vector γ)	55.10***	
Constant	8.238***	3.787*
	(1.800)	(2.126)
Wald chi^2	341.005***	278.746***
Pseudo R^2	0.2778	0.2271
Number of observations	1320	1320

* p<0.10, ** p<0.05, *** p<0.01

Standard errors adjusted for clustering effects are in parenthesis (adoption model)

Table 5. Nearest neighbor matching estimates of the effects of stone bunds on value of crop production (the dependent variable is the value of crop production per ha, ETB/ha)[*]

With Mundlak's approach	No. of conserved plots	No. of non-conserved plots	ATT	se	T
Tigray region	390	183	412.034	140.867	2.925
Amhara region	232	150	-178.897	177.460	-1.008
Without Mundlak's approach					
Tigray region	390	190	298.822	141.694	2.109
	232	159	-145.584	149.784	-0.972
Amhara region					

Bootstrapped standard errors are used to account for the estimated propensity score used in the second stage (nearest neighbor matching estimator)

The outcome variable is the value of crop production per ha (hereafter referred to as yield). Our estimates show the existence of a positive additional significant yield premium of Ethiopian Birr (ETB)[12] 412 (US$ 59) and ETB 299 (US$ 47) per ha with and without Mundlak's approach, respectively, for conserved plots compared to non-conserved plots in a low rainfall area (Tigray region) of the Ethiopian highlands.[13] These estimated impacts are fairly large relative to the average value of crop production in the Tigray highlands, which averaged ETB 1816 per ha in the survey sample. Given all other variables constant, if all comparable non-conserved plots in the highlands of Tigray had been covered with stone bunds, the estimated total benefit would have been about ETB 52 million (US$ 7 million) and ETB 38 million (US$ 6 million) with and without Mundlak's approach, respectively.[14] In contrast, no significant

---

[12] The official exchange rate averaged about 7 ETB per U.S. dollar in 1998.

[13] This result is consistent with results obtained using alternative matching methods such as the kernel [ATT = 333.805 (se =111.649)***] and stratification [ATT = 378.734 (se=115.431)***] matching methods. Although we did not show these results in order to conserve space, they are consistent with those obtained without Mundlak's approach.

[14] This benefit is calculated based on the following assumptions. There are about one million hectares of cropland cultivated by smallholder farmers in the Tigray region (Tsegay 1996). The proportion of terraced and un-terraced cropland in our survey is about 37 and 63 percent, respectively. We assumed all untreated plots in the region need conservation and will be treated only by stone bunds. We also assumed that the proportion of matched untreated plots that are comparable to treated plots (about 20 percent for both with and without Mundlak's approach) would also hold if we analyze the population (all plots) of treated and

differences in mean crop production value were found between conserved and non-conserved plots in the higher rainfall area (Amhara region), although the estimated yield difference is negative (ETB -179 and ETB -145 per ha with and without Mundlak's approach, respectively).[15]

Notably, the results obtained without Mundlak's approach underestimate the impact of stone bunds, implying that impact assessment without controlling for unobservable effects can lead to biased estimates.

Stochastic Dominance Analysis Estimates

Stochastic dominance analysis (SDA) is used to compare and rank the distributions of alternative risky outcomes according to their level and dispersion (riskiness) of returns (Mas-Colell et al. 1995). The comparison and ranking is based on cumulative density functions. Unlike matching and linear regression models, SDA does not focus only on the mean yield, but rather examines the entire density of yields. Similar to the propensity score matching method, SDA makes no assumption about the relationships between the regressors and outcome variables and does not require distributional assumptions.

Here, the SDA estimates are based on matched observations, in order to control for the impacts of other factors (apart from stone bunds) on production. The SDA therefore determines the difference in yield distribution between the two states (conservation and no conservation) due only to the effects of this technology. Figures 1-4 show cumulative density functions for yields obtained from conserved and non-conserved plots. As illustrated in the figures, the yield cumulative distribution with conservation is entirely to the right of the no conservation yield distribution for the Tigray region, indicating that yield with conservation unambiguously holds first-order stochastic dominance over that without conservation. The non-parametric Kolmogorov-Smirnov statistics test for first-order stochastic dominance (or the test for the vertical distance between the two cumulative density functions (CDFs)) also confirmed this result [$D = 0.1270$ ($p = 0.028$)** and $D = 0.2471$ ($p = 0.011$**) with and without Mundlak's approach, respectively].[16] These results imply that the chance of getting higher yields is greater for plots with conservation compared to plots without conservation, given a matched sample of conserved and non-conserved plots. However, we do not see this dominance for the Amhara region dataset. It appears that the yield distribution of non-conserved plots first-order stochastically dominates the yield distribution of conserved plots in this dataset. However, the results from a Kolmogorov-Smirnov test for first-order

---

untreated plots in the region. Based on this, the total benefit of investing in untreated plots = total cropland area*0.63*0.2* the estimated per ha benefit of treatment (ETB 412 or 299). However, further work will be required to robustly confirm this.

[15] This result also is consistent with those from the kernel [ATT = -115.248 (se =130.245)] and stratification [ATT = -106.628 (se= 133.132)] matching methods.

[16] The null hypothesis of this statistic is that the empirical CDFs of conserved and non-conserved plots have the same distribution function, while the alternative is that the CDF of conserved plots first-order stochastically dominates the CDF of non-conserved plots.

stochastic dominance analysis fails to indicate a statistically significant difference between the two distributions [$D = 0.0915$ ($p = 0.431$) and $D = 0.1571$ ($p = 0.353$) with and without Mundlak's approach, respectively]. These results agree with those obtained from the propensity score nearest neighbor matching approach.

Figure 1. The impact of stone bunds on value of crop production in Tigray region with Mundlak's approach: First order stochastic dominance analysis

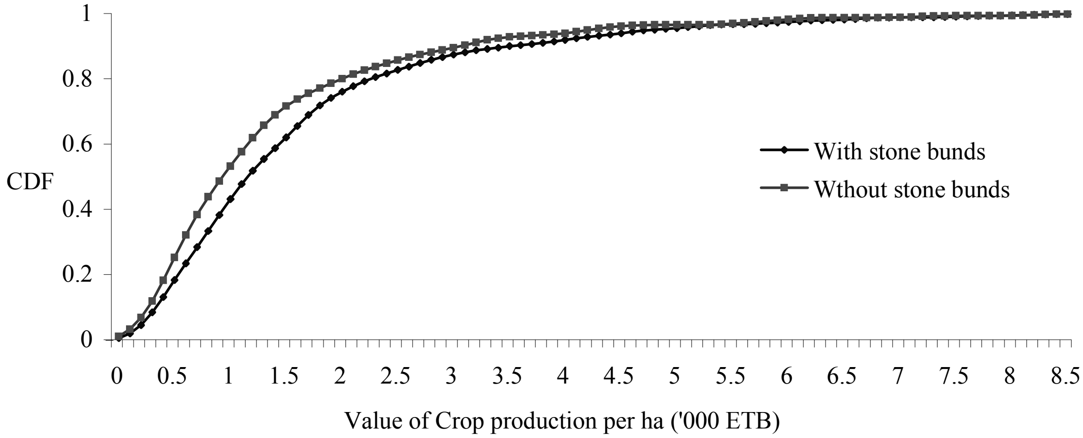


Figure 2. The impact of stone bunds on value of crop production in Amhara region with Mundlak's approach: First order stochastic dominance analysis

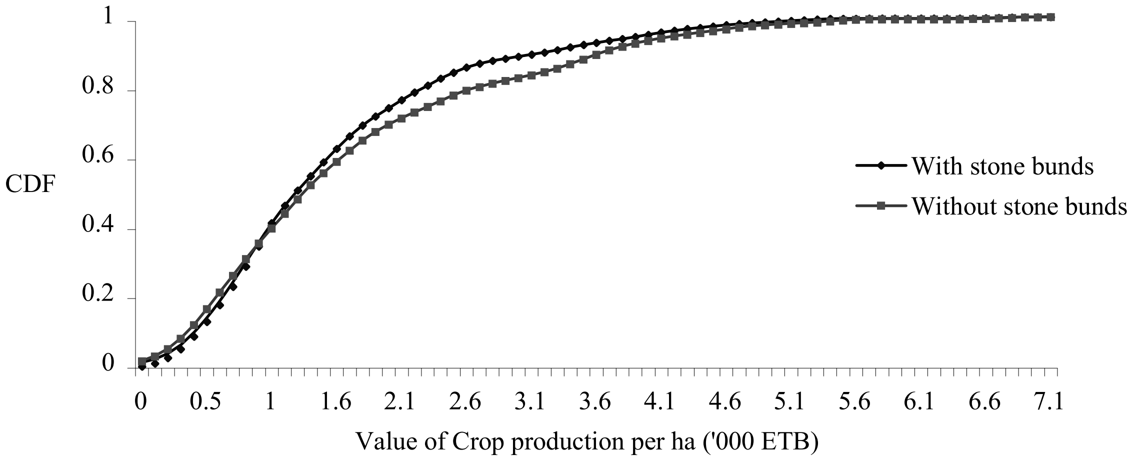


Figure 3. The impact of stone bunds on value of crop production in Tigray region with Mundlak's approach: First order stochastic dominance analysis

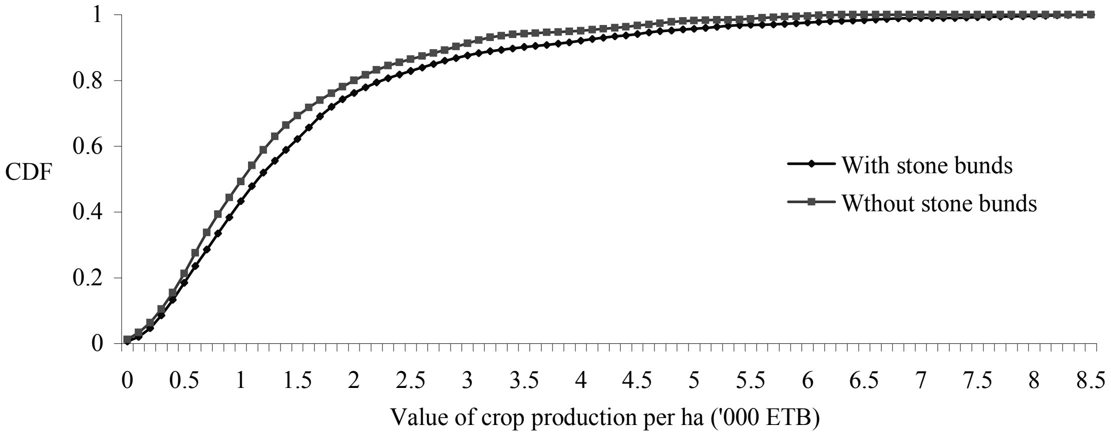
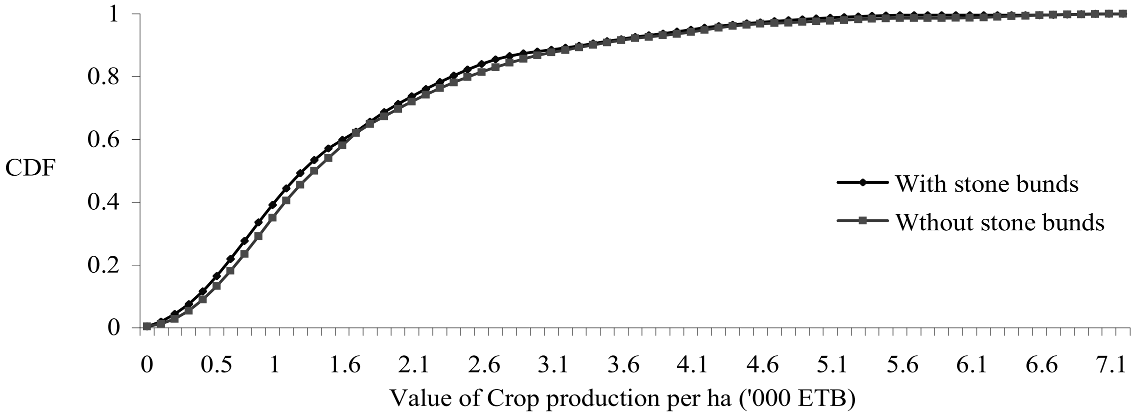


Figure 4. The impact of stone bunds on value of crop production in Amhara region without Mundlak's approach: First order stochastic dominance analysis



Parametric Regression Estimates

Modified random effects models[17] are based on matched observations that are similar in the distribution of propensity scores and covariates. The structural and reduced forms of the models are estimated with and without Mundlak's approach, even though our statistical evidence indicates that the vector γ is statistically different from zero, implying that there is a correlation between observed regressors and unobserved random effects. The dependent variable is the log of value of crop production per hectare.

Our parameter of interest is the mean yield gap between conserved and non-conserved plots. In the interest of brevity, we do not discuss the details of the estimated coefficients of the explanatory variables of the exogenous switching regressions herein, but these results are available in Tables 6-9.

---

[17] We use the modified random effects models for each specification, except for the regression model without conservation with potential endogenous conventional inputs (third column of Table 7) of the Amhara region dataset with Mundlak's approach, where we used pooled ordinary least square estimation because we had insufficient observations to run the random effects model (RE) on matched samples.

We also estimate the endogenous switching regression models; however, because of the large amount of output, which in turn is related to the large number of utilized regressors, the predicted yield is reported but the overall regression results are not, in order to save space.[18]

Table 6. Exogenous switching regression results of the determinants of crop production value in the Tigray region with Mundlak's approach (the dependent variable is the natural logarithm of crop production value per ha, ETB/ha)

Explanatory variables	Model 1	Model 2	Model 3	Model 4
Deep soil plots	-0.131	-0.498**	-0.139	-0.462**
	(0.107)	(0.215)	(0.115)	(0.196)
Medium soil plots	-0.362***	-0.606***	-0.386***	-0.709***
	(0.130)	(0.202)	(0.134)	(0.183)
Gently sloped plots	-0.282**	-0.246	-0.237*	-0.307*
	(0.116)	(0.152)	(0.123)	(0.177)
Steeply sloped plots	-0.259*	-0.203	-0.270*	-0.335
	(0.139)	(0.303)	(0.140)	(0.302)
Brown soil plots	-0.166	-0.394*	-0.173	-0.371
	(0.176)	(0.213)	(0.171)	(0.291)
Gray soil plots	-0.070	-0.896***	-0.088	-0.930***
	(0.213)	(0.231)	(0.207)	(0.235)
Red soil plots	-0.071	-0.695**	-0.069	-0.699**
	(0.177)	(0.296)	(0.175)	(0.322)
Loam soil plots	-0.025	0.713***	-0.000	0.810***
	(0.133)	(0.225)	(0.129)	(0.226)
Clay soil plots	0.120	0.689***	0.149	0.697**
	(0.180)	(0.265)	(0.182)	(0.275)
Sandy soil plots	-0.110	0.666**	-0.083	0.676**
	(0.189)	(0.304)	(0.184)	(0.295)
Moderately eroded plots	0.098	-0.255	0.082	-0.129
	(0.084)	(0.185)	(0.095)	(0.205)
Severely eroded plots	0.003	0.226	0.025	0.346
	(0.172)	(0.295)	(0.172)	(0.334)
Plot distance to residence	-0.196	-0.314*	-0.333**	-0.310*
	(0.139)	(0.163)	(0.156)	(0.183)
Reduced tillage plots	0.471***	-0.118	0.336**	-0.169
	(0.153)	(0.231)	(0.154)	(0.267)
Gully plots	0.050	0.005	-0.030	0.086
	(0.354)	(0.232)	(0.332)	(0.272)
Ln(plot size)	-0.247***	-0.351***	-0.379***	-0.594***
	(0.081)	(0.125)	(0.088)	(0.149)
Ln(other plot)	-0.265*	-0.392	-0.368**	-0.714**
	(0.159)	(0.271)	(0.176)	(0.315)
Fenced plots	0.014	0.564**	0.009	0.447
	(0.179)	(0.277)	(0.184)	(0.317)
Middle slope position	0.123	-0.773**	0.133	-1.047***
	(0.120)	(0.325)	(0.120)	(0.335)
Bottom slope position	0.144	-0.774**	0.166	-1.047***
	(0.124)	(0.358)	(0.126)	(0.351)

---

[18] Although not reported, the criterion models were not statistically significant for both with and without Mundlak's approach. The inverse Mills ratio was marginally significant (at 10%) for only two models of the 16 models estimated. We have assumed functional form (non-linearity of Mills ratio) identification since we do not have a valid instrument for inclusion in the first stage of the Heckman model.

Table 6. Continued

Explanatory variables	Model 1	Model 2	Model 3	Model 4
Not on slope	-0.089	-0.706**	-0.096	-1.006***
	(0.145)	(0.325)	(0.158)	(0.356)
Rented plots	-0.146	0.024	-0.219*	0.049
	(0.119)	(0.187)	(0.126)	(0.235)
Residence distance to market	-0.035	-0.076*	-0.042*	-0.072**
	(0.028)	(0.039)	(0.025)	(0.035)
Male household head	0.664***	0.847***	0.478**	0.984***
	(0.214)	(0.291)	(0.226)	(0.243)
Ln(household head age)	0.014	-0.316	-0.005	-0.290
	(0.179)	(0.345)	(0.203)	(0.310)
Ln(family size)	-0.057	-0.371**	-0.005	-0.431**
	(0.137)	(0.188)	(0.136)	(0.178)
Education between grades one & two	-0.049	0.158	-0.010	0.140
	(0.183)	(0.375)	(0.169)	(0.344)
Education above grade three	-0.057	0.361	0.040	0.465
	(0.239)	(0.364)	(0.228)	(0.355)
Population density	-0.001	0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Ln(rainfall)	0.442	-1.007*	-0.044	-0.922*
	(0.493)	(0.573)	(0.412)	(0.492)
Ln(household altitude)	-0.745*	-0.628	-0.103	-0.166
	(0.450)	(0.570)	(0.450)	(0.565)
Oxen holding	-0.062	-0.064	-0.053	-0.088
	(0.064)	(0.108)	(0.061)	(0.103)
Other cattle	0.032*	0.068**	0.042**	0.076***
	(0.017)	(0.028)	(0.018)	(0.026)
Small ruminant	0.011*	0.001	0.004	-0.004
	(0.006)	(0.009)	(0.006)	(0.009)
Pack animals	-0.065	-0.102	-0.058	-0.085
	(0.043)	(0.095)	(0.045)	(0.085)
Ln(fertilizer use)	0.013	0.005		
	(0.020)	(0.044)		
Ln(seed use)	0.254***	0.171**		
	(0.070)	(0.081)		
Ln(labor use)	0.102	0.058		
	(0.072)	(0.110)		
Ln(oxen use)	0.005	0.297*		
	(0.102)	(0.171)		
Joint chi^2 test for significance of mean of plot- varying explanatory variables (vector γ)	40.99**	41.47**	29.08*	43.51***
constant	8.651**	17.928***	7.582*	14.557**
	(4.285)	(6.691)	(4.442)	(6.377)
Model Wald chi^2 test	470.621***	448.395***	240.997***	309.971***
Overall R^2	0.3870	0.5892	0.3027	0.5678
Number of conservations	390	183	390	183

Note: *p<0.10, ** p<0.05, *** p<0.01
Standard errors adjusted for clustering are in parenthesis

Model1, regression estimates with conservation including potential endogenous conventional inputs
Model2, regression estimates without conservation including potential endogenous conventional inputs
Model3, regression estimates with conservation without potential endogenous conventional inputs
Model4, regression estimates without conservation without potential endogenous conventional inputs

Table 7. Exogenous switching regression results of the determinants of crop production value in the Tigray region without Mundlak's approach (the dependent variable is the natural logarithm of crop production value per ha, ETB/ha)

Explanatory variables	Model 1	Model 2	Model 3	Model 4
Deep soil plots	-0.096	-0.130	-0.087	-0.112
	(0.100)	(0.194)	(0.109)	(0.195)
Medium soil plots	-0.323***	-0.248	-0.309**	-0.246
	(0.121)	(0.192)	(0.124)	(0.193)
Gently sloped plots	-0.313***	0.077	-0.288***	0.133
	(0.097)	(0.177)	(0.106)	(0.184)
Steeply sloped plots	-0.270**	0.178	-0.274**	0.005
	(0.131)	(0.299)	(0.132)	(0.287)
Brown soil plots	0.051	0.129	0.038	0.128
	(0.146)	(0.283)	(0.144)	(0.334)
Gray soil plots	-0.063	0.225	-0.027	0.205
	(0.170)	(0.287)	(0.165)	(0.296)
Red soil plots	0.018	0.146	0.054	0.092
	(0.146)	(0.273)	(0.143)	(0.274)
Loam soil plots	-0.012	-0.042	-0.016	0.119
	(0.127)	(0.306)	(0.128)	(0.305)
Clay soil plots	0.138	-0.196	0.089	-0.009
	(0.151)	(0.284)	(0.153)	(0.286)
Sandy soil plots	-0.039	-0.185	-0.053	0.011
	(0.160)	(0.317)	(0.158)	(0.324)
Moderately eroded plots	0.030	-0.210	0.022	-0.167
	(0.081)	(0.172)	(0.084)	(0.170)
Severely eroded plots	0.025	-0.210	0.011	-0.206
	(0.150)	(0.248)	(0.155)	(0.256)
Plot distance from residence	-0.183	0.195	-0.324**	0.183
	(0.128)	(0.171)	(0.142)	(0.193)
Rented plots	-0.135	-0.099	-0.239**	-0.128
	(0.112)	(0.152)	(0.118)	(0.200)
Reduced tillage plots	0.457***	0.120	0.337**	-0.014
	(0.141)	(0.205)	(0.132)	(0.238)
Gully plots	0.097	-0.368	0.098	-0.407
	(0.300)	(0.402)	(0.271)	(0.347)
Ln(plot size)	-0.238***	0.006	-0.320***	-0.224***
	(0.054)	(0.082)	(0.056)	(0.086)
Ln(other plot size)	-0.167**	0.037	-0.208***	-0.030
	(0.065)	(0.089)	(0.074)	(0.100)
Fenced plots	0.014	0.740***	0.030	0.858***
	(0.139)	(0.229)	(0.142)	(0.289)
Middle slope position	0.092	0.013	0.107	-0.201
	(0.096)	(0.244)	(0.098)	(0.282)
Bottom slope position	-0.017	0.067	0.014	-0.152
	(0.109)	(0.247)	(0.113)	(0.295)
Not on slope	-0.181	-0.044	-0.126	0.040
	(0.129)	(0.239)	(0.133)	(0.256)
Distance to market	-0.039	-0.065*	-0.052**	-0.067*
	(0.024)	(0.033)	(0.024)	(0.034)
Male household head	0.544**	0.554***	0.416*	0.792***
	(0.214)	(0.207)	(0.226)	(0.184)
Ln(household head age)	-0.009	-0.290	0.016	-0.326
	(0.185)	(0.221)	(0.205)	(0.216)

Table 7. Continued

Explanatory variables	Model 1	Model 2	Model 3	Model 4
Ln(family size)	-0.100	-0.047	-0.008	-0.079
	(0.133)	(0.163)	(0.134)	(0.179)
Education between grades one & two	-0.026	0.054	-0.029	0.233
	(0.151)	(0.240)	(0.144)	(0.229)
Education above grade three	0.047	-0.088	0.102	-0.130
	(0.199)	(0.333)	(0.223)	(0.292)
Population density	-0.000	-0.001	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Ln(rainfall)	0.288	-0.860*	-0.375	-1.648***
	(0.373)	(0.443)	(0.372)	(0.436)
Ln(altitude)	-0.854**	-0.895**	-0.627	-0.434
	(0.388)	(0.445)	(0.415)	(0.480)
Oxen holding	-0.068	-0.114	-0.047	-0.112
	(0.065)	(0.078)	(0.063)	(0.082)
Other cattle	0.033**	0.048***	0.045***	0.053***
	(0.016)	(0.017)	(0.016)	(0.019)
Small ruminant	0.012**	0.003	0.006	-0.003
	(0.005)	(0.005)	(0.005)	(0.005)
Pack animals	-0.039	0.024	-0.028	-0.007
	(0.041)	(0.036)	(0.039)	(0.047)
Ln(fertilizer use)	0.028	0.059**		
	(0.017)	(0.025)		
Ln(seed use)	0.255***	0.222***		
	(0.054)	(0.070)		
Ln(labor use)	0.034	0.182*		
	(0.068)	(0.098)		
Ln(oxen use)	0.090	0.317**		
	(0.090)	(0.147)		
Constant	9.921**	17.619***	13.654***	21.339***
	(3.947)	(5.023)	(4.118)	(5.576)
Model Wald chi^2 test	262.851***	219.587***	161.142***	153.716***
Overall R^2	0.3009	0.5195	0.2254	0.3911
Number of observations	390	190	390	190

Note: $^*p<0.10$, ** $p<0.05$, *** $p<0.01$
Standard errors adjusted for clustering are in parenthesis

Model1, regression estimates with conservation including potential endogenous conventional inputs
Model2, regression estimates without conservation including potential endogenous conventional inputs
Model3, regression estimates with conservation without potential endogenous conventional inputs
Model4, regression estimates without conservation without potential endogenous conventional inputs

Table 8. Exogenous switching regression results of the determinants of crop production value in the Amhara region with Mundlak's approach (the dependent variable is the natural logarithm of crop production value per ha, ETB/ha)

Explanatory variables	Model 1	Model 2	Model 3	Model 4
Plot slope in degrees	-0.003	0.007	-0.012	0.017
	(0.011)	(0.015)	(0.013)	(0.014)
Black soil plots	-0.014	-0.172	0.102	-0.296
	(0.134)	(0.199)	(0.159)	(0.262)
Brown soil plots	0.071	-0.119	0.128	-0.327
	(0.161)	(0.262)	(0.189)	(0.300)
Gray soil plots	0.044	-0.229	0.196	-0.954**
	(0.217)	(0.431)	(0.220)	(0.409)

Table 8. Continued

Explanatory variables	Model 1	Model 2	Model 3	Model 4
Deep soil plots	0.106	0.149	0.141	0.438
	(0.139)	(0.345)	(0.151)	(0.354)
Medium soil plots	-0.100	-0.044	-0.146	0.159
	(0.091)	(0.263)	(0.116)	(0.281)
Moderately eroded plots	-0.008	-0.005	0.096	0.136
	(0.135)	(0.226)	(0.149)	(0.238)
Severely eroded plots	0.023	-0.477	0.030	0.295
	(0.199)	(0.376)	(0.273)	(0.438)
Clay soil plots	-0.119	-0.129	-0.153	-0.188
	(0.174)	(0.391)	(0.209)	(0.366)
Loam soil plots	-0.139	-0.286	-0.106	-0.255
	(0.121)	(0.189)	(0.146)	(0.195)
Sandy soil plots	-0.136	0.279	-0.092	0.219
	(0.159)	(0.276)	(0.177)	(0.284)
Highly fertile plots	0.624**	-0.624*	0.572	-0.866**
	(0.313)	(0.338)	(0.351)	(0.408)
Medium fertile plots	0.164	-0.229	0.127	-0.170
	(0.109)	(0.224)	(0.144)	(0.251)
Plot distance to main road	-0.253**	0.194	-0.187	-0.020
	(0.123)	(0.233)	(0.142)	(0.235)
Plot distance to residence	0.295*	0.001	0.230	0.018
	(0.157)	(0.269)	(0.177)	(0.340)
Gully plots	-0.116	-0.123	0.037	-0.325
	(0.211)	(0.386)	(0.245)	(0.370)
Rented plots	0.204	-0.013	0.231	-0.235
	(0.164)	(0.323)	(0.170)	(0.416)
Reduced tillage plots	-0.217	-0.207	-0.185	-0.299
	(0.278)	(0.670)	(0.391)	(0.685)
Plot altitude	0.000	0.001	0.000	0.001
	(0.000)	(0.001)	(0.000)	(0.001)
Ln(plot size)	0.002	0.331	0.226	0.819***
	(0.227)	(0.364)	(0.270)	(0.268)
Other plot area	-0.283	-0.605	-0.329	-1.043
	(0.534)	(0.755)	(0.620)	(0.644)
Residence distance to market	-0.076	0.162	-0.253	0.326
	(0.152)	(0.419)	(0.171)	(0.486)
Male household head	-0.075	0.792	-0.179	1.329*
	(0.239)	(0.626)	(0.212)	(0.753)
Family size	0.017	-0.013	0.042	0.009
	(0.024)	(0.031)	(0.027)	(0.036)
Household head age	0.002	0.003	0.001	0.001
	(0.005)	(0.006)	(0.005)	(0.007)
Livestock holding	0.032	0.014	0.058**	0.103**
	(0.026)	(0.045)	(0.028)	(0.044)
Education level	0.011	-0.027	0.018	-0.036
	(0.013)	(0.018)	(0.014)	(0.023)
Population density	-0.001	0.001	-0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
Ln(rainfall)	0.135	0.223	0.415	0.611
	(0.242)	(0.357)	(0.261)	(0.376)
Ln(fertilizer use)	-0.008	0.043		
	(0.022)	(0.056)		

Table 8. Continued

Explanatory variables	Model 1	Model 2	Model 3	Model 4
Ln(seed use)	0.189***	0.129		
	(0.048)	(0.100)		
Ln(labor use)	0.197*	0.416***		
	(0.102)	(0.148)		
Ln(oxen use)	0.245**	0.043		
	(0.113)	(0.160)		
Joint chi^2/F test for Significance of mean of plot-varying explanatory variables (vector γ)	43.44**	1.77**	40.43***	31.60*
constant	3.706**	2.119	3.480*	0.736
	(1.825)	(2.530)	(1.889)	(2.667)
Model Wald chi^2 (F) test	581.768***	6.11***	253.184***	249.412***
Over all R^2	0.5779	0.6508	0.4100	0.4576
Number of observations	232	150	232	150

Note: *p<0.10, ** p<0.05, *** p<0.01
Standard errors adjusted for clustering are in parenthesis

Model1, regression estimates with conservation including potential endogenous conventional inputs
Model2, regression estimates without conservation including potential endogenous conventional inputs
Model3, regression estimates with conservation without potential endogenous conventional inputs
Model4, regression estimates without conservation without potential endogenous conventional inputs

Table 9. Exogenous switching regression results of the determinants of crop production value in the Amhara region without Mundlak's approach (the dependent variable is the natural logarithm of crop production value per ha, ETB/ha)

Explanatory variables	Model 1	Model 2	Model 3	Model 4
Plot slope in degrees	0.009	0.010	0.004	0.012
	(0.007)	(0.007)	(0.008)	(0.009)
Black soil plots	0.057	-0.108	0.076	-0.111
	(0.096)	(0.235)	(0.121)	(0.240)
Brown soil plots	0.111	-0.040	0.066	-0.193
	(0.098)	(0.201)	(0.121)	(0.212)
Gray soil plots	0.252*	0.460*	0.332**	0.024
	(0.130)	(0.254)	(0.152)	(0.340)
Deep soil plots	0.135	0.188	0.127	0.406
	(0.114)	(0.281)	(0.134)	(0.272)
Medium soil plots	-0.061	-0.301	-0.126	-0.111
	(0.078)	(0.274)	(0.095)	(0.226)
Moderately eroded plots	-0.072	-0.044	-0.012	-0.045
	(0.092)	(0.178)	(0.108)	(0.195)
Severely eroded plots	-0.175	-0.064	-0.038	-0.197
	(0.135)	(0.253)	(0.208)	(0.230)
Clay soil plots	-0.125	-0.919**	-0.226	-0.660
	(0.153)	(0.400)	(0.182)	(0.430)
Loam soil plots	-0.108	-0.279**	-0.098	-0.323**
	(0.096)	(0.134)	(0.114)	(0.157)
Sandy soil plots	-0.200**	0.080	-0.158	0.129
	(0.097)	(0.213)	(0.103)	(0.240)
Highly fertile plots	0.603***	0.013	0.699***	0.152
	(0.201)	(0.269)	(0.259)	(0.328)
Medium fertile plots	0.171*	0.041	0.235**	0.052
	(0.096)	(0.184)	(0.110)	(0.226)

Table 9. Continued

Explanatory variables	Model 1	Model 2	Model 3	Model 4
Plot distance to main road	-0.002	0.022	-0.007	0.025
	(0.016)	(0.018)	(0.018)	(0.017)
Plot distance to residence	0.194	-0.184	0.063	-0.309
	(0.120)	(0.243)	(0.150)	(0.296)
Gully plots	-0.010	-0.479*	0.083	-0.215
	(0.142)	(0.249)	(0.168)	(0.264)
Rented plots	0.253**	-0.266	0.382***	-0.460*
	(0.118)	(0.334)	(0.135)	(0.239)
Reduced tillage plots	-0.122	0.092	-0.081	0.216
	(0.090)	(0.228)	(0.105)	(0.172)
Ln(plot altitude)	0.026	0.980	-0.194	0.276
	(0.218)	(0.625)	(0.284)	(0.637)
Ln(plot size)	0.018	0.053	0.299***	0.349***
	(0.080)	(0.129)	(0.079)	(0.117)
Other plot area	-0.170***	-0.136	-0.369***	-0.158
	(0.063)	(0.094)	(0.083)	(0.115)
Residence distance to market	0.007	-0.073	-0.022	-0.054
	(0.034)	(0.082)	(0.042)	(0.092)
Male household head	-0.098	0.377	-0.165	1.226*
	(0.276)	(0.498)	(0.246)	(0.671)
Family size	-0.006	0.052	0.028	0.051
	(0.022)	(0.035)	(0.028)	(0.037)
Household head age	0.003	-0.001	0.001	-0.007
	(0.005)	(0.007)	(0.005)	(0.007)
Livestock holding	0.054**	0.041	0.080***	0.109***
	(0.024)	(0.034)	(0.029)	(0.040)
Education level	0.014	0.011	0.016	0.006
	(0.012)	(0.023)	(0.012)	(0.025)
Population density	-0.082	0.115	-0.195**	-0.033
	(0.094)	(0.120)	(0.096)	(0.152)
Ln(rainfall)	0.203	-0.633*	0.432**	-0.125
	(0.154)	(0.345)	(0.187)	(0.347)
Ln(fertilizer use)	0.014	0.088**		
	(0.017)	(0.036)		
Ln(seed use)	0.187***	0.156**		
	(0.030)	(0.067)		
Ln(labor use)	0.207***	0.244**		
	(0.072)	(0.109)		
Ln(oxen use)	0.177**	0.135		
	(0.085)	(0.128)		
Constant	3.382*	1.159	5.963***	4.574
	(1.834)	(3.812)	(1.814)	(3.477)
Model Wald chi^2 test	312.106***	273.532***	100.905***	69.609***
Overall R^2	0.5025	0.4503	0.2978	0.2532
Number of observations	232	159	232	159

Note: *p<0.10, ** p<0.05, *** p<0.01
Standard errors adjusted for clustering are in parenthesis

Model1, regression estimates with conservation including potential endogenous conventional inputs
Model2, regression estimates without conservation including potential endogenous conventional inputs
Model3, regression estimates with conservation without potential endogenous conventional inputs
Model4, regression estimates without conservation without potential endogenous conventional inputs

The predicted yields from regression equation (1) (augmented by equation (2)) are used to examine the mean yield gap between conserved and non-conserved plots. The results from the exogenous switching regressions indicate that the mean value of crop production difference between conserved and non-conserved plots is negative, but not statistically significant for the Amhara region (Table 10). However, it is positive and significant for the Tigray region, which is in line with the results from our NN propensity score matching and stochastic dominance analysis. The mean yield gaps are significantly lower and less statistically significant for the dataset without Mundlak's approach. When the inverse Mills' ratio is included (endogenous switching regression model) in the models, the results are consistent and robust (Table 11). No significant differences in estimated yield impacts are observed between the endogenous and exogenous switching regression models.

Table 10. Exogenous switching regression estimates of the effects of conservation on predicted mean value of crop production with Mundlak's approach

Model types	**Predicted mean yield with stone bunds**	**Predicted mean yield without stone bunds**	**Predicted mean yield difference [se]**
A	**B**	**C**	**D =B-C**
1. Exogenous switching regression estimates			
1.1. With Mundlak's approach			
With conventional inputs			
Tigray Region	7.050	6.901	0.149 (0.055)***
Amhara region	7.131	7.180	-0.049 (0.064)
Without conventional inputs			
Tigray Region	7.056	6.890	0.166 (0.050)***
Amhara region	7.130	7.181	-0.050 (0.053)
1.2. Without Mundlak's approach			
With conventional inputs			
Tigray Region	7.049	6.954	0.095 (0.048)**
Amhara region	7.109	7.129	-0.020 (0.058)
Without conventional inputs			
Tigray Region	7.049	6.953	0.097 (0.042)**
Amhara region	7.113	7.127	-0.014 (0.045)

Note: Standard errors in parentheses

Table 11. Endogenous switching regression estimates of the effects of conservation on predicted mean value of crop production with Mundlak's approach

Model types	**Predicted mean yield with stone bunds**	**Predicted mean yield without stone bunds**	**Predicted mean yield difference [se]**
A	**B**	**C**	**D =B-C**
1. Endogenous switching regression estimates			
1.1. With Mundlak's approach			
With conventional inputs			
Tigray Region	7.049	6.901	0.148 (0.055)***
Amhara region	7.131	7.180	-0.049 (0.064)
Without conventional inputs			
Tigray Region	7.055	6.891	0.165 (0.050)***
Amhara region	7.131	7.181	-0.050 (0.054)
1.2. Without Mundlak's approach			
With conventional inputs			
Tigray Region	7.049	6.954	0.095 (0.048)**
Amhara region	7.110	7.129	-0.019 (0.057)
Without conventional inputs			
Tigray Region	7.049	6.952	0.097 (0.042)**
Amhara region	7.113	7.113	-0.015 (0.047)

Note: Standard errors in parentheses

Based on the results obtained using three different methods, we therefore conclude that soil and moisture conservation is more productive in low rainfall areas compared to high rainfall areas of Ethiopia. We believe this is due to greater benefits of moisture conservation in low rainfall areas, whereas moisture conservation in high rainfall areas may contribute to problems such as waterlogging, increased weed growth and enhanced pest infestation. This finding is consistent with those from other studies in the Ethiopian highlands (Herweg 1993; Benin 2006; Kassie and Holden 2006; Pender and Gebremedhin 2006). However, even though the use of physical structures for moisture conservation in high rainfall areas may not increase short-term productivity, this does not mean that no conservation techniques are warranted. In fact, appropriate conservation measures (e.g. drainage ditches) could help protect soils during extreme rainfall events even in higher rainfall areas.

6. SUMMARY AND CONCLUSION

The primary objective of this paper is to investigate the impact of stone bund adoption on crop yields using multiple plot observations per household in low and high rainfall areas of the Ethiopian highlands. It is important to consider rainfall in such studies, because rainfall varies substantially across and within countries, including those in SSA such as Ethiopia. It is therefore important to consider the distribution of rainfall when making decisions about SMC technologies. Indeed, agro-ecological conditions may be a particularly important determinant of whether SMC adoption increases household welfare. Oddly, however, SMC technologies have been actively promoted in Ethiopia and many other countries without accounting for agro-ecological conditions.

We used propensity score matching, stochastic dominance analysis (SDA), and parametric regression (modified linear random effects and pooled OLS models) to ensure robustness of our findings. The parametric regression and SDA estimates are based on matched samples obtained from a nearest neighbor matching method using propensity score estimates. This is important because conventional regression and stochastic dominance analysis estimates are obtained without ensuring that comparable conserved and non-conserved plots actually exist within the distribution of covariates.

The estimates from the three methods consistently indicate that stone bunds have a positive and statistically significant impact on productivity in low rainfall areas. For instance, the results from propensity score estimates show the existence of a positive additional significant crop production value premium of ETB 412 (US$ 59) and ETB 299 (US$ 47) per ha with and without Mundlak's approach, respectively, for conserved plots compared to non-conserved plots in a low rainfall area (the Tigray region) of the Ethiopian highlands. However, this impact is not observed in a high rainfall area (Amhara). These findings suggest that the productivity impact of stone bunds is agro-ecology-specific. This highlights the importance of developing and disseminating agro-ecology-specific soil conservation technologies to increase agricultural productivity, rather than making blanket recommendations that promote similar conservation measures to all farmers. For instance, in high rainfall areas, moisture conservation using physical structures may not be important, but placing appropriate drainage measures could help protect soil during extreme rainfall events.

REFERENCES

Assunção, J.J. and B.H.L. Braido. 2004. Testing among competing explanations for the inverse productivity puzzle. Accessed on November 2006, available at: http://www.econ.puc-rio.br/PDF/seminario/2004/inverse.pdf.

Bekele, W. 2005. Stochastic dominance analysis of soil and water conservation in subsistence crop production in the Eastern Ethiopian highlands: The case of Hunde-Lafto area. *Environ. Res. Econ*. 32(4): 533-550.

Benin, S. 2006. Policies and programs affecting land management practices, input use and productivity in the highlands of Amhara region, Ethiopia. In *Strategies for sustainable land management in the East African Highlands*, ed. Pender, J., F. Place, and S, Ehui. Washington, D.C.: International Food Policy Research Institute.

Byiringiro, F. and T. Reardon. 1996. Farm productivity in Rwanda: effects of farm size, erosion and soil conservation investments. *Agric. Econ.* 15: 127-136.

Corbett, J.D. and J.W. White. 2001. Using sub-annual climate models for meso-resolution spatial analyses (for *Agronomy Journal*).

Dehejia, R., and S. Wahba. 1999. Causal effects in non-experimental studies: Re-evaluating the evaluation of training programs. J. *Am. Stat. Association* 94(448): 1053-1062.

Dehejia, H.R., and Wahba, S. 2002. Propensity score matching methods for non-experimental causal studies. *The Rev.Econ.Stat*. 84(1): 151-161.

Fafchamps, M. 1993. Sequential labour decisions under uncertainty: An estimable household model of West Africa farmers. *Econometrica* 61: 1173-1197.

Fischer T., N. Turner, J. Angus, L. McIntyre, M. Robertson, A. Borrell, and D. Lloyd. 2004. New directions for a diverse planet: Proceedings for the 4th International Crop Science Congress, Brisbane, Australia, September 26–October 1, 2004

Gebremedhin, B, S.M. Swinton, and Y. Tilahun. 1999. Effects of stone terraces on crop yields and farm profitability: results of on-farm research in Tigray, Northern Ethiopia. *J. Soil and Water Conservation* 54 (3): 568-573.

Heckman, J. 1979. Sample selection bias as a specification error. *Econometrica* 47: 153-61.

Heckman, J., Robb, R. 1985. Alternative methods for evaluating the impact of interventions. In *Longitudinal analysis of labor market data,* ed. Heckman, J. and B. Slinger. New York: Cambridge University Press.

Heckman, J., H. Ichimura, J. Smith, and P. Todd. 1998. Characterizing selection bias using experimental data. *Econometrica* 66 (5), 1017-1098.

Herweg, K. 1993. Problems of acceptance and adoption of soil conservation in Ethiopia. *Tropics Applied Resource Management* 3: 391-411.

Holden, S.T., B. Shiferaw, and J. Pender. 2001. Market imperfections and profitability of land use in the Ethiopian Highlands: A comparison of selection models with heteroskedasticity. *J. Agric. Econ.* 52(2): 53-70.

IPPC (Intergovernmental Panel on Climate Change). 2001. *Climate change 2001*. Synthesis report. Cambridge University Press. Cambridge.

Kaliba, A.R.M. and T. Rabele. 2004. Impact of adopting soil conservation practices on wheat yield in Lesotho. In *Managing nutrient cycles to sustain soil fertility in Sub-Saharan Africa.* Tropical Soil Biology and Fertility Institute: CIAT.

Kassie, M. and T.S. Holden. 2006. Parametric and non-parametric estimation of soil conservation adoption impact on yield. Contributed paper prepared for presentation at the international Association of Agricultural Economists Conference, Gold Coast, Australia, August 12-18, 2006.

Levinsohn, J. and A. Petrin. 2003. Estimating production functions using inputs to control for unobservable. *Rev. Econ. Studies* 70: 317–341.

Mas-Colell, A., M.D. Whinston, and J.R. Green. 1995. *Microeconomic theory*. New York, Oxford University Press.

Mundlak, Y. 1978. On the pooling of time series and cross section data. *Econometrica* 64(1): 69-85.

Pender, J. and J. M. Kerr. 1998. Determinants of farmer's indigenous soil and water conservation investments in semi-arid India. *Agric. Econ.* 19: 113-125.

Pender, J. and B. Gebremedhin. 2006. Land management, crop production and household income in the highlands of Tigray, Northern Ethiopia: An econometric analysis. In *Strategies for sustainable land management in the East African Highlands*, ed. Pender, J., F. Place, S. Ehui. Washington, D.C.: International Food Policy Research Institute.

Rosenbaum, P.R. and D.B. Rubin. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1): 41-55.

Shiferaw, B. and S.T. Holden. 2001. Farm-level benefits to investments for mitigating land degradation: empirical evidence from Ethiopia. *Environ. Dev. Econ.* 6: 335-358.

Shively, G.E. 1999. Risks and returns from soil conservation: evidence from low-income farms in the Philippines. *Environmental Monitoring and Assessment* 62: 55-69.

Shively, G.E. 1998a. Modelling impacts of soil conservation on productivity and yield variability: Evidence from a heteroskedastic switching regression. Selected paper at annual meeting of the American Agricultural Economics Association, August 2-5, 1998, Salt Lake City, Utah.

Shively, G.E. 1998b. Impact of contour hedgerow on upland maize yields in the Philippines. *Agroforestry systems* 39: 59-71.

Smith, J. and Todd, P. 2005. Does Matching Overcome LaLonde's Critique of Non-experimental Estimators? *J. Econometrics* 125(1-2): 305-353.

Sutcliffe J.P. 1993. Economic assessment of land degradation in the Ethiopian Highlands. A Case Study. National Conservation Strategy Secretariat, Ministry of Planning and Economic Development. Addis Ababa, Ethiopia: Transitional Government of Ethiopia.

Wooldridge, J.M. 1995. Selection correction for panel data models under conditional independence assumptions. *J. Econometrics* 68: 115-132.

Wooldridge, J.M. 2002. *Econometric analysis of cross section and panel data*. Cambridge, Mass.: The MIT Press.

WID. 2005. *World development indicators database.* Washington, D.C.: World Bank.

RECENT IFPRI DISCUSSION PAPERS

For earlier discussion papers, please go to www.ifpri.org/pubs/pubs.htm#dp.
All discussion papers can be downloaded for free.

732. *Innovation systems governance in* Bolivia: Lessons for agricultural innovation policies. Frank Hartwich, Anastasia Alexaki, and René Baptista, 2007.
731. *The human right to food as a U.S. nutrition concern, 1976-2006.* Ellen Messer and Marc Cohen, 2007.
730. *Measuring and accounting for community capabilities in Kordofan, Sudan.* Khalid El Harizi and Heather Klemick, 2007.
729. *How to make agricultural extension demand-driven?: The case of India's agricultural extension policy.* Regina Birner and Jock Anderson, 2007.
728. *Urbanization, educational expansion, and expenditures inequality in Indonesia in 1996, 1999, and 2002*. Takahiro Akita and Sachiko Miyata, 2007.
727. *Diversification in Indian agriculture towards high-value crops: The role of smallholders*. P. S. Birthal, P. K. Joshi, Devesh Roy, and Amit Thorat
726. *Farmer preferences for milpa diversity and genetically modified maize in Mexico: A latent class approach*. Ekin Birol, Eric Rayn Villalba, Melinda Smale, 2007.
725. Generating *plausible crop distribution and performance maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach*. Liangzhi You, Stanley Wood, and Ulrike Wood-Sichra, 2007.
724. *Assessing the impact of the national agricultural advisory services (NAADS) in the Uganda rural livelihoods.* Samuel Benin, Ephraim Nkonya, Geresom Okecho, John Pender, Silim Nahdy, Samuel Mugarura, Edward Kato, and Godfrey Kayobyo, 2007.
723. *Rural Investments to Accelerate Growth and Poverty Reduction in Kenya*. James Thurlow, Jane Kiringai, and Madhur Gautam, 2007.
722. *Smallholders' Commercialization through Cooperatives: A Diagnostic for Ethiopia.* Tanguy Bernard, Eleni Gabre-Madhin, Alemaheyu Seyoum Taffesse, 2007.
721. *Understanding Policy Volatility in Sudan.* Khalid El Harizi, El Sayed Zaki, Betina Prato, and Ghada Shields, 2007.
720. *The impact of the Central America free trade agreement on the Central American textile maquila industry.* Hans G.P. Jansen, Sam Morley, Gloria Kessler, Valeria Piñeiro, and Marco Sánchez, 2007.
719. *The Food retail revolution in poor countries: Is it coming or is it over?: Evidence from Madagascar*. Bart Minten, 2007.
718. *The economic impact and the distribution of benefits and risk from the adoption of insect resistant (bt) cotton in West Africa*. Jose Falck-Zepeda, Daniela Horna, and Melinda Smale, 2007.
717. *Quality control in non-staple food markets: Evidence from India*. Marcel Fafchamps, Ruth Vargas Hill, and Bart Minten, 2007.
716. *Investment, subsidies, and pro-poor growth in rural India*. Shenggen Fan, Ashok Gulati, and Sukhadeo Thorat, 2007.
715. *Risk aversion in low income countries: Experimental evidence from Ethiopia.* Mahmud Yesuf and Randy Bluffstone, 2007.
714. *Micro-level analysis of farmers' adaptation to climate change in southern Africa.* Charles Nhemachena, and Rashid Hassan, 2007.
713. *Resource abundance and regional development in China.* Xiaobo Zhang, Li Xing, Shenggen Fan, Xiaopeng Luo, 2007.
712. *Is food insecurity more severe in south Asia or sub-saharan Africa?: A comparative analysis using household expenditure survey data.* Lisa C. Smith and Doris Wiesmann, 2007.
711. *Managing conflict over natural resources in greater Kordofan, Sudan: Some recurrent patterns and governance implications*. El Fatih Ali Siddig, Khalid El-Harizi, and Betinna Prato, 2007.
710. *Mortality, mobility, and schooling outcomes among orphans: Evidence from Malawi.* Mika Ueyama, 2007.

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

www.ifpri.org

IFPRI HEADQUARTERS
2033 K Street, NW
Washington, DC 20006-1002 USA
Tel.: +1-202-862-5600
Fax: +1-202-467-4439
Email: ifpri@cgiar.org

IFPRI ADDIS ABABA
P. O. Box 5689
Addis Ababa, Ethiopia
Tel.: +251 11 6463215
Fax: +251 11 6462927
Email: ifpri-addisababa@cgiar.org

IFPRI NEW DELHI
CG Block, NASC Complex, PUSA
New Delhi 110-012 India
Tel.: 91 11 2584-6565
Fax: 91 11 2584-8008 / 2584-6572
Email: ifpri-newdelhi@cgiar.org