AGRICULTURAL ECONOMICS

Agricultural Economics 46 (2015) 757–768

# Does land registration and certification boost farm productivity? Evidence from Ethiopia

Mequanint B. Melesse*, Erwin Bulte

*Development Economics Group, Wageningen University, P.O. Box 8130, 6700 EW Wageningen, The Netherlands*

Received 20 September 2013; received in revised form 6 January 2015; accepted 7 January 2015

**Abstract**

According to economic theory, tenure security is an important determinant of agricultural investment and productivity. Land titling has been at the center stage of development efforts of many African countries to boost tenure security. We investigate the productivity impacts of the Ethiopian land registration and certification program, employing propensity score matching method in an effort to create a credible counterfactual. Consistent with theory, we find land registration and certification has robust positive effects on farm productivity. More tentatively, we identify the assurance effect as one probable channel for impact. Households with land certificates are more likely to adopt soil-fertility management strategies on their plots than households without certificates.

*JEL classifications*: Q12, Q15, R52

*Keywords:* Land certification; Land productivity; Propensity score matching; Ethiopia

## 1. Introduction

Land remains an important resource for a large portion of the African population (Cotula et al., 2004; Toulmin, 2009), and land reforms have been at the heart of development efforts of several African countries. Particularly, since millions live and work on land they do not legally own, the quest for secure land rights features high on the development agenda of governments and the international development community.

Theoretically, the literature identifies several channels through which tenure security can contribute to economic growth (Besley, 1995; Brasselle et al., 2002). First, secure tenure rights can boost long term land-related investments by providing assurance that returns on these investments will not be appropriated. Second, secure and transferable land rights allow factor mobility, and facilitate efficiency gains by reallocating land to more efficient users via land markets. Third, secure land rights may increase access to formal credit by using land as collateral. Brasselle et al. (2002) coin these three effects of secured land rights as, respectively, the "assurance," "realizability," and "collateralizability" effect.[1]

Based on these arguments, many countries have launched land registration and certification programs. However, for the African context it is surprisingly hard to come by conclusive empirical evidence supporting the view that land registration and certification are instrumental to development. While evidence from Latin America and Asia tends to corroborate this proposition (e.g., Deininger and Chamorro, 2004; Galiani and Schargrodsky, 2010; Jacoby et al., 2002), evidence is rather mixed for Africa. While some report that stronger land rights are associated with an increased likelihood of investment (Abdulai et al., 2011; Besley, 1995; Deininger et al., 201; Holden et al., 2009), others do not (Braselle et al., 2002; Jacoby and Minten, 2007; Place and Migot-Adholla, 1998). Empirical work in this domain faces serious challenges in light of (self) selection issues, and the notion that investment behavior

*Corresponding author: Tel: +31-317-317484-360. *E-mail address:* mequanint.melesse@wur.nl (M.B. Melesse).

**Data Appendix Available Online**

A data appendix to replicate main results is available in the online version of this article.

[1] Secured land rights may also reduce the level and likelihood of land conflicts (Fenske, 2011). In the same vein, secure tenure is more likely to reduce the need to expend private resources to defend land rights and preclude the need for presence of landholders on a continuous basis to safeguard their land claims, thus increasing their participation in off-farm labor markets (Deininger and Feder, 2009).



DOI: 10.1111/agec.12191

of households may affect tenure security (Sjaastad and Bromley, 1997).[2]

Amid such mixed experiences, new waves of land certification programs have been fueled by several factors, including new legislation, low-cost titling methods, and increasing demand for land (Cotula et al., 2004). The Ethiopian land registration program has been viewed as a pilot for such certification programs. With the goal of issuing every rightful holder a certificate of usufructs, the program is the largest delivery of nonfreehold rights in a short period (Deininger et al., 2011). A few careful empirical papers have documented its effects. Deininger et al. (2008) describe the implementation process and provide an evaluation of its early impacts based on subjective assessments by households. Holden et al. (2011) find positive effects of land certification on allocative efficiency of land rental markets in Tigray region. Likewise, Deininger et al. (2011) report that land certification improved tenure security, and increased investment and supply of land to rental markets in Amhara region.

Little work has been done on the program's impact on agriculture productivity. One key exception is Holden et al. (2009), who study productivity effects using panel data in Tigray. Following Holden et al. (2009), our aim is to study the effect of the Ethiopian land registration program on productivity. Our study complements Holden et al. (2009) because we focus on another region with a different certification program—allowing us to assess robustness of earlier findings to another context and program design. For example, while certificates are issued only in the name of the head in Tigray, the Amhara region requires joint certification in the names of both spouses.[3] Moreover, the process was more intensive and costly than in other regions due to a survey and land administration team involvement (Deininger et al., 2008). In addition, while the threat of redistribution never appears far off the agenda in Ethiopia, it was only in Amhara that land was redistributed as recently as 1997.

Unfortunately, we do not have access to panel data. Relying on cross-sectional data, we use propensity score matching for causal inference based on comparing certified plots in certified Kebeles with plots in Kebeles that were left uncertified for arguably exogenous reasons (see below). We employ several approaches to test major identifying assumptions behind our method to diagnose its credibility and performance. However, as with other (nonexperimental) cross-section studies, we realize that reservations remain about our ability to attribute causal effects to specific treatments (certification). While we seek to minimize concerns about unobserved heterogeneity, we cannot eliminate all potential concerns about endogeneity. Hence, we also provide qualifications to our findings in the concluding section.

Our main result is consistent with theory and with Holden et al. (2009). Land registration and certification has large and robust positive effects on farm productivity. A simple comparison of costs and benefits reveals that benefits substantially exceed the cost of implementation. Based on reasoning that discounts alternative mechanisms we conjecture that our results are presumably driven by the so-called assurance effect, but also argue that further research into the mechanism linking titling to productivity is needed.

The remaining part of the article is organized as follows. Section 2 provides a brief account of the Ethiopian tenure system and describes the land registration and certification program in some detail. Section 3 describes the data and the empirical strategy, and Section 4 contains the results. Section 5 presents robustness checks and sensitivity analysis, while Section 6 summarizes and puts the results into perspective.

## 2. The Ethiopian tenure system and land certification program

In Ethiopia, the land tenure system has been at the forefront of policy debates for generations, owing to the prominence of land as a source of livelihood and as a source of political power (Crewett and Korf, 2008). In the decades prior to and during the imperial era, land was concentrated in the hands of absentee landlords, and arbitrary evictions posed a serious threat to tenant farmers (Deininger and Jin, 2006). After overthrowing the imperial regime of Haile Selassie through a military coup, the socialist Derg regime implemented radical reforms that altered the agrarian structure and access to land, transferring land ownership to the state (Holden et al., 2011).

After the fall of the Derg regime in 1991, the current government embarked on liberalizing the economy. However, the set of reforms largely "overlooked" the land problem, and the legacy of the Derg continued to define key elements of the current land policy (Crewett and Korf, 2008). Land rights continue to be vested in the state. On the other hand, the current government has made several changes. First, it devolved responsibility for land issues to regions. Second, the frequency of land redistribution has been stalled, but land redistribution is not completely off the agenda. Third, land rentals are allowed officially, though some regions still impose restrictions on the terms of rental contracts.

Overall, the state continues to be a source of tenure insecurity. The government remains critical of privatizing land holdings, retaining a discourse of social equity and protection of land concentration in the hands of the few. However, some have argued that the government uses land rather as a "carrot and stick" to fulfill political ends (Crewett and Korf, 2008). Partly in an effort to respond to these critics, and to bolster farmers' tenure security, the government commissioned a large-scale land registration program in 1998.

The Ethiopian registration program has delivered the largest nonfreehold land rights in a short period in Sub-Saharan Africa

[2] Empirical work is also complicated because of restrictions on certain forms of land use (e.g., tree planting), or because of public investment programs that might crowd in or crowd out private investments in land quality.

[3] Joint certification is not a peculiar feature of the Amhara region—the Oromia and the Southern Nations and Nationalities regions also issued joint certificates.

(Deininger et al., 2011). The Amhara region, where our survey was conducted, started a pilot in 2002, and scaled up the program in 2003. Since 1998, more than 20 million parcels have been registered across the country (Deininger et al., 2011). Certificates are issued with names of both spouses, and contain details related to size, location, and quality of plots. The certification process starts with election of a land use and administration committee (LAC).[4] The LAC assumes responsibility for the main registration tasks, including plot identification, demarcation, measurement, and field-based conflict resolution. The registration process is implemented in a decentralized, high-speed and participatory approach (Toulmin, 2009). It has been viewed as a right response to local needs, as evidenced by the recipients' readiness to pay for replacing lost certificates, high demand for a spatial reference, and positive assessment of likely impacts (Deininger et al., 2008).

Another feature of the program is its cost effectiveness. With estimated average cost of about US$1 per plot or US$3.2 per hectare[5] (Deininger et al., 2008),[6] the program has set a new standard. This cost is an order of magnitude below what has been reported elsewhere in the literature for similar first-time land registrations, with average costs between US$20 and US$60 per parcel, at times even exceeding US$100 (Deininger et al., 2011; Jacoby and Minten, 2007). The Ethiopian land registration costs are low and affordable, owing mainly to the employment of local tools and techniques with local participation (Holden et al., 2011).

## 3. Data and empirical strategy

### *3.1. Data description*

Data used for this study were collected in Amhara region, Ethiopia. The survey was conducted in 2012, covering districts of Womberma, Bure, and Jabitehinan. Five certified and seven uncertified Kebeles were randomly sampled from these three districts.[7] Districts and Kebeles are gradually enrolled in the certification program. We focus on uncertified Kebeles from a large set of uncertified Kebeles, making sure to select Kebeles that were uncertified for plausibly exogenous reasons (i.e., unrelated to farm management and productivity). Such reasons include human resource and certificate constraints at the level of the district, and inter-Kebele conflicts. Such conflicts concerned migration issues or disputed grazing land claims—relevant domains from the perspective of rural households, but with no bearing on farming land. Accordingly, we believe there is no *a priori* reason to assume that certified and uncertified households are systematically different.

Within each Kebele, we randomly sampled households proportional to the population, resulting in a sample of 325 farm households, of which 141 were certified. In total, 1,678 plots were included in the survey. Identification of average treatment effects (ATE) rests upon a comparison of households from certified Kebeles, and households in uncertified Kebeles. Within certified Kebeles, there are also uncertified households, but it is easy to see how including such households could introduce (self) selection bias since such households are likely to be systematically "different."[8] However, we use data from these households to estimate intention to treat (ITE) effects, as a robustness check for our results. For our core specification, we also exclude rented-in and rented-out plots, allowing us to identify the assurance effect more clearly (as opposed to the realizability effect).[9] This yielded 1,509 plots for the analysis, of which 588 were owner-operated certified and 921 were uncertified.

To further ensure comparability of certified and uncertified households, we selected Kebeles from a restricted region where economic opportunities and livelihood options are sufficiently similar. We picked the districts based on the notion of "development domains," defined as geographic locations sharing broadly similar rural development constraints and opportunities based on agricultural potential, access to markets and population density (Chamberlin et al., 2006). They are characterized as high population density and moisture-reliant highlands with productivity growth and market improvement as priority livelihood strategies. Households also predominantly produce cereals and operate in similar markets.

Table 1 summarizes the data used in our analysis. Our explanatory variables include household and plot level variables. Household variables are age and education of heads, household size, land and livestock of households, and market access. The interpretation of most variables is straightforward. Average

[4] The process of the program varies to some extent from region to region. Here the descriptions are mainly based on the experience of the Amhara region, where our survey was conducted.

[5] US$1 was approximately equal to 9.6 Ethiopian Birr in 2007/2008.

[6] The costs were computed using data from Amhara region where registration was much more intensive and costly than in other regions. The costs include public costs of the program and private costs for books and certificates. Though LAC members work voluntarily and for free, the project cost estimation accounts for their opportunity costs as well (see Deininger et al., 2008: 1803-05 for a detailed explanation of the cost estimation).

[7] Kebeles are the lowest administrative unit in Ethiopia. We selected five Kebeles from Womberma, four Kebeles from Bure, and three Kebeles from Jabitehinan.

[8] Households residing in certified areas might not be certified for such reasons as absence during registration, failure to collect certificates, household establishment after the registration, and changes in household heads.

[9] This also accentuates incentive problems related to such plots, and rules out the possibility of double counting of plots. Holden et al. (2011) report that land certification has enhanced participation of households in land rental markets. However, Deininger et al. (2013) find that such rentals lead to lower levels of efficiency because of incentive problems and inefficient contractual arrangements. Tenants would lose incentive to undertake long-term investments on a land they rented in for a limited period, for which we find tentative supportive evidence in our data. A simple two-sample $t$-test for the productivity of rented and owned plots suggests that the productivity of owner-operated plots is significantly ($3.21 > 2.99$; $P$-value $= 0.005$) larger than that for rented plots.

Table 1
Variable definition and descriptive statistics[a]

| Variable | Description | All data Mean | Min | Max | Certified Mean | Uncertified Mean | t-test |
|---|---|---|---|---|---|---|---|
| HH-age | Age of household head in years | 44.99 (10.49) | 24 | 71 | 45.19 (9.68) | 44.86 (10.98) | 0.552 |
| HH-education | Education of household head in years | 1.85 (2.56) | 0 | 10 | 2.29 (2.67) | 1.57 (2.44) | 0.000 |
| HH-size | Household size | 6.89 (2.13) | 2 | 12 | 6.99 (2.04) | 6.82 (2.19) | 0.066 |
| HH-land size | Total household land size in timad | 10.69 (4.97) | 2 | 33 | 10.00 (4.33) | 11.13 (5.29) | 0.000 |
| Total livestock | Total livestock in tropical livestock units (tlu)[11] | 7.22 (3.88) | 0.08 | 21.01 | 7.24 (4.05) | 7.21 (3.77) | 0.883 |
| Distance to market | Distance to a daily market center in hours on foot | 2.08 (0.41) | 1.15 | 4.22 | 2.03 (0.38) | 2.12 (0.42) | 0.000 |
| Plot size | Size of plot in timad | 1.87 (1.21) | 0.25 | 16 | 1.90 (1.13) | 1.85 (1.26) | 0.421 |
| Distance to plot | Distance to plot from homestead in minutes | 31.72 (35.73) | 1 | 320 | 26.15 (27.90) | 35.27 (39.54) | 0.000 |
| Very fertile plot | Dummy for plot is of very good quality | 0.69 (0.46) | 0 | 1 | 0.70 (0.46) | 0.68 (0.47) | 0.416 |
| Fertile plot | Dummy for plot is of medium quality | 0.24 (0.43) | 0 | 1 | 0.26 (0.44) | 0.23 (0.42) | 0.140 |
| Poor fertile plot | Dummy for plot is of poor quality | 0.07 (0.25) | 0 | 1 | 0.04 (0.19) | 0.09 (0.28) | 0.128 |
| Flat plot | Dummy for plot is of flat slope | 0.896 (0.31) | 0 | 1 | 0.91 (0.29) | 0.89 (0.32) | 0.215 |
| Gently sloping plot[12] | Dummy for plot is of gently sloping | 0.104 (0.305) | 0 | 1 | 0.09 (0.29) | 0.11 (0.32) | 0.893 |
| Plot no conflict | Dummy for plot is not affected by conflicts | 0.75 (0.43) | 0 | 1 | 0.76 (0.42) | 0.74 (0.44) | 0.291 |
| Plot conflict | Dummy for plot is affected by conflicts | 0.25 (0.43) | 0 | 1 | 0.24 (0.43) | 0.26 (0.44) | 0.855 |
| Value added in log | Value added per plot in log | 3.20 (0.43) | 1 | 4.59 | 3.41 (0.35) | 3.07 (0.42) | 0.000 |

[a]Standard deviations in parentheses.

household land size is about 11 timads,[10] which is larger than the national average of 6 timads (Future Agricultures, 2006). Plot level variables include biophysical features, such as fertility, slope and plot size, and conflict status of plots. Two-sample *t*-tests for equality of means of the covariates of certified and uncertified plots indicate that the two groups are not balanced for all relevant covariates (last column of Table 1), hence we will control for these differences in our analysis below.

Our outcome variable for land productivity is value added per timad, which we transform by taking its logarithm. When computing the value added for plots, we have included agricultural working capital and opportunity costs of family labor at local wage rates. Figure A in the online appendix provides the value added distribution for certified and uncertified plots. A two-sample *t*-test for equality of means of the two plot groups shows that the mean for certified plots is significantly greater than that for uncertified plots ($P$-value $= 0.000$). Moreover, a two-sample Kolmogorov–Smirnov test for equality of distributions reveals that the distributions are significantly different ($P$-value $= 0.000$).

However, such differences between certified and uncertified plots should not necessarily be attributed to the certification program. Causal inference requires controlling for a number of potential sources of biases. Certified and uncertified plots may differ in a number of cultivator and plot characteristics, like cultivator's resources and plot fertility. To navigate these concerns, we use propensity score matching.

[10] One timad is the land area ploughed by a pair of oxen in a day, and approximately equals 0.25 hectare.

[11] Tropical livestock unit is a common unit used to quantify a wide range of various livestock species and sizes as a single figure indicating the total amount of livestock owned by a household. Conversion factors vary across geographical regions. We employed a TLU for SSA with conversion factors: matured cow (1.0), ox (1.42), small cattle (0.73), goats/sheep (0.2), horses/mules/donkeys (0.8), and poultry (0.04).

[12] In the survey instrument, we captured the plot slope variable in three constructs of flat, gently, and steeply sloping. However, we included steeply sloping plots in the gently sloping plots' category, as there were only 19 observations.

### 3.2. Propensity score matching

We employ propensity score matching (PSM) to mitigate the problem of selection bias. Let $D_i$ be an indicator of whether plot $i$ is certified or uncertified. The potential productivity outcome of certification, represented by $\Gamma_i$, for each plot $i$ is defined as $\Gamma_i(D_i)$. The average treatment effect (ATT) on the treated is computed as:

$$\Delta_{ATT} = E(\Delta|D=1) = E[\Gamma(1)|D=1] - E[\Gamma(0)|D=1] \quad (1)$$

where $\Delta_{ATT}$ is the average treatment effect on the treated; $E[\Gamma(1)|D=1]$ is expected productivity for certified plots; and $E[\Gamma(0)|D=1]$ is expected productivity of certified plots if they had not been certified. The latter is a counterfactual, and not observed.

Computing ATT requires obtaining a credible counterfactual. PSM helps to construct the counterfactual from uncertified plots. To control for selection bias, the comparison group must be statistically equivalent to the treated group, and one needs to match on all observable covariates. To circumvent the curse of dimensionality, Rosenbaum and Rubin (1983) propose instead to match on the propensity score, $p(X)$, which is the probability of receiving treatment conditional on all covariates, $X$. PSM requires imposing conditional independence (CIA) and common support assumptions for identification (Heckman et al., 1997). If these two assumptions are met, the PSM estimator for $\Delta_{ATT}$ is given as:

$$\Delta_{ATT}^{PSM} = E_{p(X)|D=1}\{E[\Gamma(1)|D=1, p(X)] \\ -E[\Gamma(0)|D=0, p(X)]\}. \quad (2)$$

We estimate propensity scores using a logit model. We are guided by theory, previous empirical studies, and the institutional setting to include only explanatory variables that influence participation and the outcome (Heckman et al., 1997). We use the kernel matching algorithm for our main discussion, which uses weighted averages of all subjects in the comparison group to construct the counterfactual (Caliendo and Kopeinig, 2008). However, we probe the model's robustness using alternative matching methods.

Finally, PSM estimators do not account for selection on unobservables. However, we believe such selection bias has little impact on our results. First, while land certification is a voluntary decision of households, the intervention is driven by the government, and is exogenous to cultivators and communities' unobservables. The compliance rate is also very high. Moreover, certified and uncertified plots are from restricted geographical areas, sharing similar ecology, infrastructure, market development, and returns to land-specific investments (Chamberlin et al., 2006). In our base model, we exclude uncertified plots in certified areas to mitigate unobservable bias *a priori*. Finally, we carry out a robustness analysis to unobserved heterogeneity using the standard bounding approach test due to Rosenbaum (2002).

## 4. Results and discussion

### 4.1. Estimating propensity scores

Table 2 reports regression estimates of marginal effects of household and plot level observables used to estimate propensity scores. Both household and plot level observables determine the likelihood of a plot to be certified. Predicted propensity scores are then used to match certified plots to comparable uncertified plots. Some of the variables included in Table 2 are potentially endogenous to the certification decision. For example, plot fertility may capture (inherent) dimensions of soil quality but may also capture management decisions of the household. Since we collected our data sometime after certification, some of these management decisions may have occurred in response to the certification. To attenuate such endogeneity concerns in column (1), we also estimate a more parsimonious model that only contains variables that are arguably exogenous to certification status. This parsimonious model, presented in column (2), is used for a robustness analysis below.

### 4.2. Effect of land certification on plot productivity

Once propensity scores are estimated, the next step is to match certified and uncertified plots based on "closeness" of propensity scores. We report and discuss the results using the kernel matching algorithm,[13] which lowers variance as it uses more information for matching (Heckman et al., 1998). Matching has been done only on plots that are on the common support. The results of the match are summarized in Table 3. Table 3 presents estimates of the average effects of certification on certified plots. The productivity of matched certified plots is 35.4 percentage points higher than matched uncertified plots. This difference is statistically significant ($P = 0.000$). The number of certified plots used for the estimation of ATT after matching is 569, which is a vital diagnostic indicator of success in matching power. The information loss is low since only 19 certified plots are off the support region.

Given this statistically significant and large impact of certification on productivity, what is its economic relevance? The increase in productivity due to certification of plots is equivalent to an annual gain of 1,281 Birr (US$75.40)[14] in value added per plot.[15] This increment in value added per plot exceeds the program's one-off social costs (about US$1 per plot) by a wide margin.[16] Net benefits are likely to remain positive even if the land administration is comprehensively restructured to

[13] The results remain robust for other well-known matching algorithms (see Section 5 for robustness checks).

[14] US$1 was approximately equal to 17 Ethiopian Birr in 2011/2012.

[15] The increase in value added per plot in Birr is computed based on the average value added per timad of unmatched certified plots, which is equal to 3,619.081 Birr. The same value for the matched certified plots is slightly greater than this value.

[16] US$1 was approximately equal to 9.6 Ethiopian Birr in 2007/2008 when the program costs had been valued.

Table 2
Estimates of the logit model (standard errors are in parentheses)

| Variable | 1 | 2 |
|---|---|---|
| HH-age head | 0.222 (0.046)*** | 0.240 (0.046)*** |
| HH-age head squared | −0.002 (0.0005)*** | −0.002 (0.0005)*** |
| HH-education head | 0.167 (0.024)*** | 0.164 (0.024)*** |
| HH-land size | −0.150 (0.048)*** | −0.093 (0.044)** |
| HH-land size squared | 0.007 (0.002) | 0.001 (0.003) |
| Total livestock | 0.073 (0.019)*** | |
| Distance to market | −0.820 (0.163)*** | −0.722 (0.159)*** |
| Plot size | 0.461 (0.149)*** | 0.201 (0.052)*** |
| Distance to plot | −0.009 (0.002)*** | −0.009 (0.002)*** |
| Very fertile plot | 0.925 (0.280)*** | |
| Fertile plot | 0.970 (0.286)*** | |
| Flat plot | −0.443 (0.357) | 0.153 (0.185) |
| Plot no conflict | 0.084 (0.133) | |
| Interaction plot size and flat plot | −0.285 (0.156)* | |
| _cons | −4.453 (1.122) | −4.222 (1.089) |
| Number of observations | 1,509 | 1,509 |
| Prob>chi$^2$ | 0.000 | 0.000 |
| Count $R^{2}$[a] | 68.00% | 66.00% |

*Significant at 10%; ** significant at 5%; *** significant at 1%.
[a]Count $R^2$ is computed as the ration of the number of correct predictions to total sample size. It indicates what proportion of the dependent variable is correctly predicted by the model.
Standard errors in parentheses.

Table 3
Average Treatment Effects: Kernel matching[a]

| | | Certified (treated) | Uncertified (control) | Difference |
|---|---|---|---|---|
| | Unmatched | 3.406 | 3.066 | 0.340 |
| ATT | Matched | 3.410 | 3.056 | 0.354 (0.023)*** |
| Number of plots on the common support | On support | 569 | 921 | |
| | Off Support | 19 | 0 | |

*Significant at 10%; **significant at 5%; ***significant at 1%.
[a]Matching is performed by using the psmatch2 program in STATA software. We use default kernel matching (kernel type = epanechnikov kernel; bandwidth = 0.06). Std. error in parentheses. The standard error for the ATT is bootstrapped standard error of 100 replications.

include a cadastral index map, common property resources, land used for residence and rural service facilities, and a mechanism to keep records up-to-date. Of course, one may question the necessity of such a "quality upgrade" in light of the impressive performance of the simple, low-cost technology.

What channel explains this productivity gain? As discussed, land is a state property in Ethiopia, and it can neither be sold nor mortgaged. This has important implications for the relative significance of potential channels. For example, land is unlikely to be used as collateral for accessing formal credit services and collateralizability effects are not anticipated from the program. Further, we have excluded all plots that were rented-in (or out) from the analysis, which most likely minimizes the observable role of realizability effects in driving our results. However, it is important to note that excluding rented-in (or out) plots from our analysis does not completely attaint the realizability channel. Farmers may undertake productivity-enhancing investments on a plot they own today, hoping to rent-out (or bequeath to their heirs) it to a more efficient farmer in the future. If so, productivity gains from such investments will appear in the data even if the time of the transaction has not come yet, making potential realizability effects unavoidable. Minding these caveats, we believe the assurance effect is the most likely candidate to explain the impact of certification.

Can we substantiate the importance of the assurance effect? We asked farmers about their subjective assessment of impacts of the program. About 96% of the certified households believed that certification enhanced their tenure security. Similarly, 91% of them stated that owning a certificate increased incentives to plant trees (see also Holden et al. 2009, 2011; Deininger et al., 2011). In addition, 83% of the households indicated that certification reduced plot border conflicts. Seventy-four percent also thought that certification increased the probability of receiving compensation in the event of land takings for public purpose. These statements were validated by actual land management practices "on the ground." Table 4a shows that certified households are more likely to adopt land management strategies than uncertified households. These differences still exist if we move to the plot level (Table 4b). Differences in land management practices between certified and uncertified plots are significant

Table 4
Land management practices of certified and uncertified households

4a. Household level analysis

| | Certified HHs | | Uncertified HHs | | |
|---|---|---|---|---|---|
| Land management practices | Mean | Std. Error | Mean | Std. Error | $t$-test |
| Dummy for household plants trees on its farm | 0.81 | 0.032 | 0.58 | 0.037 | *** |
| Dummy for household uses organic manure on its farm | 0.86 | 0.029 | 0.60 | 0.037 | *** |
| Dummy for household constructs stone terraces on its farm | 0.76 | 0.035 | 0.67 | 0.036 | * |

4b. Plot level analysis

| Variable | Sample | Mean Certified (treated) | Mean Uncertified (control) | % reduction (bias) | $t$-test p> \|t\| |
|---|---|---|---|---|---|
| Dummy for household plants trees on its farm | Unmatched | 0.845 | 0.477 | | 0.000 |
| | Matched | 0.852 | 0.491 | 2.0 | 0.000 |
| Dummy for household uses organic manure on its farm | Unmatched | 0.838 | 0.470 | | 0.000 |
| | Matched | 0.844 | 0.495 | 5.5 | 0.000 |
| Dummy for household constructs stone terraces on its farm | Unmatched | 0.810 | 0.503 | | 0.000 |
| | Matched | 0.807 | 0.583 | 27.0 | 0.000 |

** significant at 5%; *** significant at 1%.

after matching, providing some evidence that land management decisions are consequences of certification.

### 4.3. Assessing matching quality

We need to verify the performance of PSM in eliminating differences in observables between certified and uncertified plots. The credibility of PSM hinges on meeting the CIA and the common support condition. Regarding CIA, we employ a two-sample $t$-test to see whether PSM balances the distributions of relevant variables, and a chi-square test for joint significance of all variables in the logit model. Test results are provided in Table 5. The results in Table 5 indicate that, while not all variables are balanced before matching, all explanatory variables are balanced after matching. Matching has reduced differences between certified and uncertified plots sizably as indicated by the percentage of reductions in bias.[17] An exception is the *HH–age head squared* variable. However, the difference between the two groups for this variable was already small before matching.

The chi-square test indicates that all variables in the logit model are not jointly significant after matching (prob $> \chi^2 = 0.998$). Whereas, the same test is rejected before matching (prob $> \chi^2 = 0.000$). This is confirmed by the pseudo-$R^2$ of the model. Since there are no systematic differences in covariates between certified and uncertified plots after matching, the pseudo-$R^2$ is fairly low (0.002) compared with its value before matching (0.089).

To inspect whether the common support condition is met in estimating the counterfactual, we check the presence of enough overlap between certified and uncertified plots. The common support for the whole sample ranges between 0.012 and 0.924. The predicted propensity scores for certified and uncertified groups range from 0.029 to 0.924 and 0.012 to 0.741, respectively. Consequently, the effective common support ranges from 0.029 to 0.741, which suggests adequate overlap. In addition, we report the histogram of propensity scores for certified and uncertified groups after matching in Fig. 1 (we also present the density distribution of the propensity scores for the two groups before and after matching in the online appendix as Figure B). Both graphs reveal a clear overlap after matching.

Overall, matching quality tests indicate that matching has generated counterfactual samples of uncertified plots that are statistically similar to certified plots.

## 5. Sensitivity analysis and extension

### 5.1. Robustness of the ATT estimate to different matching algorithms

Some scholars argue that the choice of matching algorithms may influence estimated results. Baser (2006), for example, proposes that sensitivity analysis of matching algorithms is important since none is *a priori* superior to others. Table 6 reports estimates of ATT based on different algorithms. Tests for matching quality are satisfied for all algorithms. Across estimators, all explanatory variables are balanced after matching, and chi-square tests show that variables are not jointly significant after matching. Pseudo-$R^2$ values are also sizably lower after matching than before matching. Generally, the ATT remains robust not only to different matching algorithms, but also to the meanest specifications of each algorithm.

[17] The bias is calculated as the difference of the mean values of the certified group and the unmatched/matched uncertified group divided by the square root of the average sample variance in the certified group and the unmatched uncertified group.

Table 5
*t*-tests for equality of means for each variable before and after the match

| Variable | Sample | Mean Certified (treated) | Mean Uncertified (control) | % reduction (bias) | *t*-test p> \|t\| |
|---|---|---|---|---|---|
| HH-age | Unmatched | 45.194 | 44.864 | | 0.552 |
| | Matched | 45.195 | 44.907 | 12.5 | 0.616 |
| HH-age squared | Unmatched | 2,136.1 | 2,133.2 | | 0.956 |
| | Matched | 2,137.8 | 2,109.2 | −907.7 | 0.594 |
| HH-education | Unmatched | 2.2857 | 1.5722 | | 0.000 |
| | Matched | 2.1336 | 2.0799 | 92.5 | 0.730 |
| HH-land size | Unmatched | 10.001 | 11.127 | | 0.000 |
| | Matched | 10.056 | 9.9418 | 89.9 | 0.663 |
| HH-land size squared | Unmatched | 118.74 | 151.79 | | 0.000 |
| | Matched | 119.98 | 118.86 | 96.6 | 0.870 |
| Total livestock | Unmatched | 7.2379 | 7.2077 | | 0.883 |
| | Matched | 7.1808 | 7.1698 | 63.7 | 0.962 |
| Distance to market | Unmatched | 2.0285 | 2.1173 | | 0.000 |
| | Matched | 2.0333 | 2.0343 | 98.9 | 0.962 |
| Plot size | Unmatched | 1.9001 | 1.8485 | | 0.421 |
| | Matched | 1.8818 | 1.8716 | 80.2 | 0.887 |
| Distance to plot | Unmatched | 26.151 | 35.268 | | 0.000 |
| | Matched | 26.153 | 27.955 | 80.2 | 0.288 |
| Very fertile plot | Unmatched | 0.70068 | 0.68078 | | 0.416 |
| | Matched | 0.69772 | 0.71313 | 22.5 | 0.569 |
| Fertile plot | Unmatched | 0.26361 | 0.23018 | | 0.140 |
| | Matched | 0.26538 | 0.25059 | 55.8 | 0.569 |
| Flat plot | Unmatched | 0.90816 | 0.88817 | | 0.215 |
| | Matched | 0.91213 | 0.9104 | 91.4 | 0.918 |
| Plot no conflict | Unmatched | 0.76361 | 0.73941 | | 0.291 |
| | Matched | 0.76098 | 0.77662 | 35.4 | 0.532 |
| Interaction plot size and flat plot | Unmatched | 1.7079 | 1.6522 | | 0.411 |
| | Matched | 1.7096 | 1.6912 | 67.0 | 0.806 |

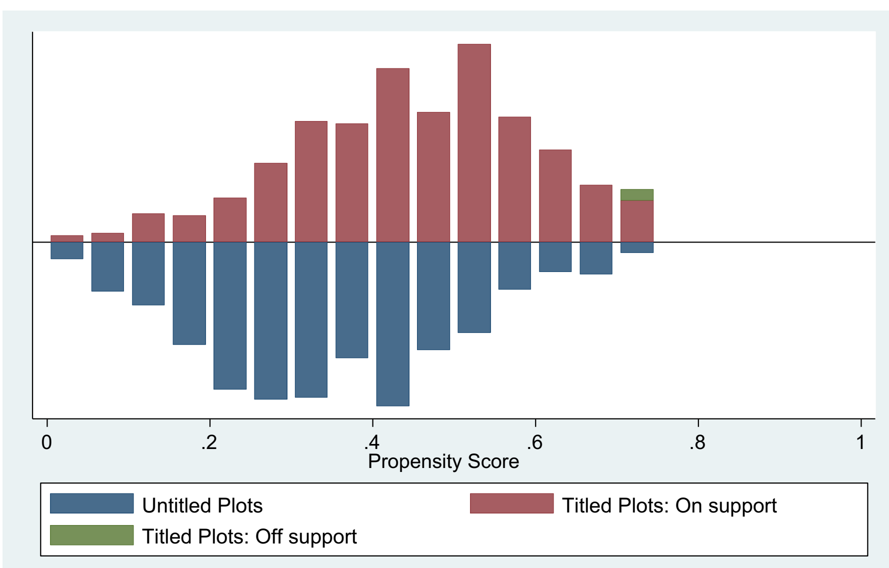


Fig. 1. Histogram distribution of propensity scores for certified and uncertified plots after matching.

## 5.2. *Exogenous estimation of the propensity scores*

Proper estimation of propensity scores requires employment of baseline data—data that are not available in our case. Some variables included in the logit model may be endogenous as we use cross-sectional data collected some time after the introduction of the certification program. Therefore, we estimate propensity scores including only exogenous

Table 6
Robustness of the ATT estimates for different algorithm estimators

| Matching algorithms | Observations on the common support | | ATT[a] | $t$-test for distributions of each variables |
|---|---|---|---|---|
| | Certified | Uncertified | | |
| Kernel matching | | | | |
| Bandwidth = 0.001 | 485 | 921 | 0.344 (0.029)*** | Satisfied |
| Badwidth = 0.01 | 569 | 921 | 0.351 (0.024)*** | Satisfied |
| Bandwidth = 0.05 | 569 | 921 | 0.354 (0.023)*** | Satisfied |
| Nearest neighbor[18] | | | | |
| NN = 1 | 569 | 921 | 0.347 (0.023)*** | Satisfied |
| (no replacement) | | | | |
| NN = 1 | 569 | 921 | 0.352 (0.034)*** | Satisfied |
| (with replacement) | | | | |
| NN = 2 | 569 | 921 | 0.340 (0.029)*** | Satisfied |
| NN = 3 | 569 | 921 | 0.358 (0.027)*** | Satisfied |
| Radius matching | | | | |
| Caliper = 0.001 | 365 | 921 | 0.345 (0.029)*** | Satisfied |
| (no replacement) | | | | |
| Caliper = 0.001 | 485 | 921 | 0.347 (0.034)*** | Satisfied |
| (with replacement) | | | | |
| Caliper = 0.01 | 473 | 921 | 0.334 (0.025)*** | Satisfied |
| (no replacement) | | | | |
| Caliper = 0.01 | 569 | 921 | 0.352 (0.034)*** | Satisfied |
| (with replacement) | | | | |
| Mahalanobis matching[19] | 588 | 921 | 0.370 (0.036)*** | Satisfied |

*Significant at 10%; **significant at 5%; ***significant at 1%.
[a]Standard errors are in parentheses.

Table 7
Average treatment effects: Kernel matching

| | | | Certified (treated) | Uncertified (control) | Difference |
|---|---|---|---|---|---|
| a) Exogenous propensity score estimation | | Unmatched | 3.406 | 3.065 | 0.341 |
| | | Matched(ATT) | 3.410 | 3.056 | 0.354 (0.023)*** |
| b) Intention to treat effect | | Unmatched | 3.399 | 3.066 | 0.333 |
| | | Matched (ITT) | 3.400 | 3.052 | 0.349 (0.023)*** |
| c) Heterogeneous treatment effect | Small group | Unmatched | 3.402 | 3.055 | 0.347 |
| | | Matched (ATT) | 3.405 | 3.067 | 0.337 (0.0367)*** |
| | Large group | Unmatched | 3.412 | 3.075 | 0.337 |
| | | Matched (ATT) | 3.410 | 3.062 | 0.348 (0.0398)*** |

We use the Epanechinikov kernel functional form with bandwidth = 0.01.
*Significant at 10%; **significant at 5%; ***significant at 1%.
Std. error for ATT in parentheses, which is bootstrapped standard error of 100 replications.

covariates (column (2) of Table 2). The new matching results are contained in Table 7(a). The ATT is 0.354, which is similar to the original estimate (Table 3). The matching quality indicators are reported in Tables A1 through A3 in the online appendix. Matching was done on the common support, and all tests reveal that quality of the match is good.

### *5.3. The intention to treat effect*

We have also estimated ITT effect. ITT gauges effects of an intervention by comparing groups based on whether or not groups received treatment, regardless of treatment status of units (Khandker et al., 2010). It provides a more conservative estimate of treatment effect as it includes noncompliers (i.e., households in certified kebeles who chose not to certify their land). To estimate the ITT effect, we include uncertified households in "certified locations" as certified ones. This procedure has resulted in 642 plots in treated villages, with 54 extra plots.

Table B1 in the online appendix provides estimates of marginal effects of covariates used to estimate propensity scores. Matching results are reported in Table 7(b). The ITT effect is 0.349, so plots in certified locations have on average 34.9 percentage points higher productivity than uncertified plots in comparison locations. This effect is comparable with the ATT effect (ATT = 0.351) for the same specification of the kernel matching algorithm (Table 6). This is not unexpected as compliance is high (about 92% of the households in the certified Kebeles are actually certified in our sample). Results for matching quality indicators are reported in Tables B2 through B4 in the online appendix and, generally, matching quality tests suggest the matching algorithm has performed well in terms of eliminating systematic differences in the distribution of covariates between the two groups.

### 5.4. Heterogeneous treatment effects

The effect of a treatment may vary across households, depending on certain characteristics. The Ethiopian land tenure is characterized by a land policy where land redistributions result in a dynamic process of leveling down households' land size. Households with more than average land size, who are more likely to be among losers in future land redistributions, tend to be more tenure insecure and may gain more from certification. We then hypothesize the ATT effect is large for households with large land size. Accordingly, we divide our data set into two groups at the median household land size to examine for heterogeneous treatment effects.[20] This procedure results in 754 farm plots of "small land size" group and 755 farm plots of "large land size" group.

Table C1 in the online appendix contains estimates of the marginal effects used to estimate propensity scores, and matching results are reported in Table 7(c) (results for matching quality indicators are given in Tables C2 through C5 in the online appendix). The ATT effects are 0.337 and 0.348 for small and large land groups, respectively. Both effects are quantitatively similar, so tenure security appears equally important for both groups. While periodic land redistributions are more likely a concern for households with large land size, other features of the land policy may create uncertainties for both groups alike. Major sources of uncertainty in the Ethiopian tenure system are the absence of a clear justice system for resolving land conflicts, frequent revisions of the law, authority given to different government agencies in land affairs, and rapid urban expansion and land grants to investors (Deininger et al., 2011). Also, competing claims from community members, such as land border conflicts and encroachment by neighbors, might be important sources of tenure insecurity for households with small land size and weak power.

### 5.5. Sensitivity of ATT estimate to unobserved heterogeneity

An important identification issue in PSM is that selection into treatment is based on only observed variables. Matching estimators are not robust to hidden bias due to unobserved variables (Caliendo and Kopeinig, 2008). One threat to identification is that unobserved heterogeneity may explain which kebeles receive the certification treatment. For example, one reason why some kebeles receive treatment earlier than others has to do with conflicts about demarcation of grazing lands (typically between kebeles). One might be concerned that this indicates a lack of well-recognized authority, be it an official or informal order, which could guarantee tenure security (affecting investments in agricultural productivity). While such an interpretation would not diminish our argument about the importance of tenure security (for agricultural productivity), it would compromise our ability to attribute the differences between certified and uncertified households to the certification process.

However, several features of the Ethiopian certification process attenuate this concern. First, we argue the role of heterogeneity is modest. Customary land laws do not prevail in Ethiopia. Competing claims based on conflicting rules, a relevant source of local heterogeneity elsewhere, are not a major source of tenure insecurity in Ethiopia. The Amhara region has a well-established conflict resolution system known as the *Shimagelle* system (Baker, 2013). Both certified and uncertified communities have a long tradition of solving land-related and other types of conflicts. While border conflicts between Kebeles regarding the demarcation of grazing lands may invite the postponement of certification, such conflicts do not signal weak local governance. Unobserved heterogeneity is also reduced because, as mentioned, certified and uncertified plots are from districts, sharing similar ecology, market development, and returns to land investments (Chamberlin et al., 2006).

Second, and minimizing the impact of remaining heterogeneity, the responsibility to the certification process has been bestowed on Woredas (districts), not on Kebeles or villages. Certification is not initiated at the local level (and compliance rates exceeding 90% imply that self-selection into certification at the household level is not likely to be relevant).[21] Certificate and human resource constraints of these districts—another leading explanation for uneven progress in certification across kebeles—are typically determined by factors beyond developments at specific Kebeles. These constraints should therefore not be viewed as indicative of unobserved heterogeneity at local levels.

Third, we also seek to formally check the robustness of our results to departures from the strong CIA assumption, and employ the bounding approach due to Rosenbaum (2002). This allows us to determine how strong an unobservable variable must be to influence the selection process so as to undermine our results. Results of Rosenbaum bound tests are reported in Table 8. We find that estimates of ATT are robust to unobservable covariates, even to the extent that would triple (gamma = 3)

[18] Matching without replacement is implemented only with 1-to-1 propensity score matching.

[19] This algorithm requires specifying a number of options at the same time. We just use the default option, which is matching on a quadratic metric with the generated weighting matrix.

[20] The median household land size is 10.25 timads while the mean household land size is 10.69 timads.

[21] Moreover, the source of tenure insecurity in Ethiopia has been the State itself through its land redistribution efforts. The decision to redistribute land is not made locally, but at the level of the regional parliament ("house of regional representatives"). Therefore, when it is decided to redistribute land, the decision will be effected throughout the region. Land proclamations do not leave discretion for making distribution decisions to zonal administrations, let alone to lower level units. This also considerably reduces the possibility of heterogeneity between villages and Kebeles on the source of the tenure insecurity.

Table 8
Robustness of the ATT estimates to unobserved heterogeneity (rbound tests)

| Gamma | sig+ | sig- | t-hat+ | t-hat- | CI+ | CI- |
|---|---|---|---|---|---|---|
| 1 | 0 | 0 | 0.3422 | 0.3422 | 0.3159 | 0.3696 |
| 1.1 | 0 | 0 | 0.3293 | 0.3559 | 0.3024 | 0.3837 |
| 1.2 | 0 | 0 | 0.3171 | 0.3683 | 0.2905 | 0.3968 |
| 1.3 | 0 | 0 | 0.3059 | 0.3803 | 0.2795 | 0.4086 |
| 1.4 | 0 | 0 | 0.2956 | 0.3911 | 0.2693 | 0.4204 |
| 1.5 | 0 | 0 | 0.2865 | 0.4011 | 0.2596 | 0.4312 |
| 1.6 | 0 | 0 | 0.2777 | 0.4105 | 0.2505 | 0.4417 |
| 1.7 | 0 | 0 | 0.2696 | 0.4201 | 0.2420 | 0.4514 |
| 1.8 | 0 | 0 | 0.2619 | 0.4286 | 0.2340 | 0.4603 |
| 1.9 | 0 | 0 | 0.2542 | 0.4372 | 0.2270 | 0.4690 |
| 2 | 0 | 0 | 0.2474 | 0.4453 | 0.2203 | 0.4775 |
| 2.1 | 0 | 0 | 0.2408 | 0.4527 | 0.2137 | 0.4858 |
| 2.2 | 0 | 0 | 0.2345 | 0.4597 | 0.2073 | 0.4937 |
| 2.3 | 0 | 0 | 0.2289 | 0.4667 | 0.2014 | 0.5014 |
| 2.4 | 0 | 0 | 0.2236 | 0.4733 | 0.1954 | 0.5084 |
| 2.5 | 0 | 0 | 0.2183 | 0.4799 | 0.1893 | 0.5150 |
| 2.6 | 0 | 0 | 0.2133 | 0.4862 | 0.1840 | 0.5218 |
| 2.7 | 0 | 0 | 0.2083 | 0.4923 | 0.1786 | 0.5282 |
| 2.8 | 0 | 0 | 0.2037 | 0.4981 | 0.1736 | 0.5344 |
| 2.9 | 0 | 0 | 0.1994 | 0.5036 | 0.1688 | 0.5407 |
| 3 | 0 | 0 | 0.1947 | 0.5091 | 0.1640 | 0.5470 |
| 3.1 | 0 | 0 | 0.1900 | 0.5141 | 0.1596 | 0.5529 |

Gamma: log odds of differential assignment due to unobserved factors.
Sig+ : upper bound significance level (assumption: overestimation of treatment effect).
Sig- : lower bound significance level (assumption: underestimation of treatment effect).
t-hat+ : upper bound Hodges-Lehmann point estimate.
t-hat-: lower bound Hodges-Lehman point estimate.
CI+ : upper bound confidence interval (a = 0.95 and assumption: overestimation of treatment effect).
CI- : lower bound confidence interval (a = 0.95 and assumption: underestimation of treatment effect).

log odds of differential assignment to treatment (which seems unlikely in light of observable heterogeneity in our sample).

## 6. Conclusions and discussion

Theory predicts that land certification is a catalyst for economic growth by providing incentives to boost agricultural investment and productivity. Yet, empirical evidence for Africa is mixed. A review of the empirical evidence shows that land reforms have not lived up to theoretical expectations, particularly in terms of productivity gains. Interestingly, a body of work is emerging that challenges this dismal outcome. Several studies of the Ethiopian land certification program produce favorable assessments of impact (Deininger et al., 2008, 2011; Holden et al., 2009, 2011).

Consistent with the gist of these studies, we find that land registration and certification has produced significant positive effects on farm productivity. We use propensity score matching as the basis of our empirical approach, and verify whether the so-called conditional independence assumption and common support condition are satisfied by our data. We perform various checks and tests supporting the validity of PSM for our analysis. Our results also survive several sensitivity analyses. Nevertheless, we acknowledge that attribution remains difficult for analyses based on nonexperimental cross-section data. While we demonstrate that unobserved heterogeneity is unlikely to drive our results, we cannot rule out that tenure security aspects beyond formal titles may explain part of the differences we find between certified and uncertified households. More speculatively, we conjecture that the assurance effect is the most likely channel via which certification affects agricultural productivity. Regardless of the mechanism, we find that the magnitude of benefits exceeds the program's implementation costs by a wide margin. The program managed to reduce tenure insecurity due to frequent land redistributions, and was tailored to fit local conditions with highly decentralized, transparent and cost-effective procedures.

Why are outcomes different for Ethiopia from other African countries? We believe the reason can be found in the local institutions history of Ethiopia. The impact of land registration and certification critically relies on the level of security that holders enjoy prior to the program. Assessing *de facto* and subjective tenure security prior to reform is difficult, as land tenure in Africa is complex, encompassing widely varying customary, religious, and statutory arrangements (Cotula et al., 2004). In many African countries, these complex customary tenure systems provide sufficient levels of *de facto* tenure security to induce farmers to invest in productivity-enhancing measures (Brasselle et al., 2002; Jacoby and Minten, 2007; Fenske, 2011; Toulmin, 2009). The variation in pre-formal arrangements is often only crudely mirrored in national land laws. Simplistic and blanket calls for formalization of land rights fail to account for the complexity and history of existing institutional arrangements, and may therefore not be appropriate. Such reforms can make matters worse if they create a parallel system that even increases insecurity (Bromley, 2009; Deininger and Feder, 2009). Land reforms are not a simple technical matter, but represent a complex and deep social intervention in a multidimensional construct (Galiani and Schargrodsky, 2010; Lemel, 1988).

The Ethiopian case is different. Customary land laws do not exist in Ethiopia, owing to the discontinuous evolution of its tenure system (Deininger et al., 2008). Before the certification program, the main source of tenure insecurity for farmers has been the state itself through its policy of land redistribution. Deininger and Jin (2006) characterize Ethiopia as more tenure insecure than other African countries, which explains the farmers' demand for land registration. Streamlining state-led redistribution initiatives had, therefore, tremendous effect on tenure security. However, the results of this study, and other studies of the Ethiopian certification program, should not be simply transferred to other African contexts—the external validity may be limited. In general, land certification reforms should not dismiss the role of competing interests and indigenous tenure forms, as these are fundamental to provide basic levels of tenure security where they exist. Neglecting such evidence through one-size-fits-all prescriptions may reduce rather than enhance *de facto* tenure security (Place, 2009).

## Acknowledgments

We would like to thank Adane Dabissa for excellent research support, and N.W.O. for financial support (N.W.O. grant # 453-10-001).

## References

Abdulai, A., Owusu, V., Goetz, R., 2011. Land tenure differences and investment in land improvement measures: Theoretical and empirical analyses. J. Dev. Econ. 96, 66–78.

Baker, B., 2013. Where formal and informal justice meet: Ethiopia's justice pluralism. Afr. J. Int. Comp. Law 21, 202–218.

Baser, O., 2006. Too much ado about propensity score models? Comparing methods of propensity score matching. Value Health 9, 377–385.

Besley, T., 1995. Property rights and investment incentives: Theory and evidence from Ghana. J. Polit. Econ. 103, 903–937.

Brasselle, A.-S., Gaspart, F., Platteau, J.-P., 2002. Land tenure security and investment incentives: Puzzling evidence from Burkina Faso. J. Dev. Econ. 67, 373–418.

Bromley, D.W., 2009. Formalizing property relations in the developing world: The wrong prescription for the wrong malady. Land Use Policy 26, 20–27.

Caliendo, M., Kopeining, S., 2008. Some practical guidance for the implementation of propensity score matching. J. Econ. Surv. 22, 31–72.

Chamberlin, J., Pender, J., Yu, B., 2006. Development domains for Ethiopia: Capturing the geographical context of smallholder development options. DSGD Discussion Paper 43, International Food Policy Research Institute (IFPRI), Washington, DC.

Cotula, L., Toulmin, C., Hesse, C., 2004. Land Tenure and Administration in Africa: Lessons of Experience and Emerging Issues. International Institute for Environment and Development (IIED), London.

Crewett, W., Korf, B., 2008. Ethiopia: Reforming land tenure. Rev. Afr. Polit. Econ. 35, 203–220.

Deininger, K., Chamorro, J. S., 2004. Investment and equity effects of land regularization: The case of Nicaragua. Agric. Econ. 30, 101–116.

Deininger, K., Castagnini, R., 2006. Incidence and impact of land conflict in Uganda. J. Econ. Behav. Org. 60, 321–345.

Deininger, K., Jin, S., 2006. Tenure security and land-related investment: Evidence from Ethiopia. Eur. Econ. Rev. 50, 1245–1277.

Deininger, K., Feder, G., 2009. Land registration, governance and development: Evidence and implications for policy. World Bank Res. Obs. 24, 233–266.

Deininger, K., Ali, D.A., Holden, S.T., Zevenbergen, J., 2008. Rural land certification in Ethiopia: Process, initial impact and implications for other African countries. World Dev. 36, 1786–1812.

Deininger, K., Ali, D.A., Alemu, T., 2011. Impacts of land certification on tenure security, investment, and land market participation: Evidence from Ethiopia. Land Econ. 87, 312–334.

Deininger, K., Ali, D.A., Alemu, T., 2013. Productivity effects of land rental market operation in Ethiopia: Evidence from a matched tenant–landlord sample. App. Econ. 45, 3531–3551.

Fenske, J., 2011. Land tenure and investment incentives: Evidence from West Africa. J. Dev. Econ. 95, 137–156.

Future Agricultures, 2006. Land, land policy and smallholder agriculture in Ethiopia. Policy Brief 001, University of Sussex, UK.

Galiani, S., Schargrodsky, E., 2010. Property rights for the poor: Effects of land certification. J. Pub. Econ. 94, 700–729.

Heckman, J., Ichimura, H., Todd, P., 1997. Matching as an econometric evaluation estimator: Evidence from evaluating a job training program. Rev. Econ. Stud. 64, 605–654.

Heckman, J., Ichimura, H., Todd, P., 1998. Matching as an econometric evaluation estimator. Rev. Econ. Stud. 65, 261–294.

Holden, S.T., Deininger, K., Ghebru, H., 2009. Impacts of low-cost land certification on investment and productivity. Am. J. Agric. Econ. 91, 359–373.

Holden, S.T., Deininger, K., Ghebru, H., (2011). Tenure insecurity, gender, low-cost land certification and land rental market participation in Ethiopia. J. Dev. Stud. 47, 31–47.

Jacoby, H.G., Minten, B., 2007. Is land certification in Sub-Saharan Africa cost-effective? Evidence from Madagascar. World Bank Econ. Rev. 21, 461–485.

Jacoby, H.G., Li, G., Rozelle, S., 2002. Hazards of expropriation: Tenure insecurity and investment in rural China. Am. Econ. Rev. 92, 1420–1447.

Khandker, S.R., Koolwal, G.B., Samad, H.A., 2010. Handbook on Impact Evaluation: Quantitative Methods and Practices. World Bank, Washington, D. C.

Lemel, H., 1988. Land certification: Conceptual, empirical and policy. Land Use Policy 5, 273–290.

Place, F., Migot-Adholla, S.E., 1998. The economic effects of land registration on smallholder farms in Kenya: Evidence from Nyeri and Kakamega districts. Land Econ. 74, 360–373.

Rosenbaum, P.R., 2002. Observational Studies. Springer, New York.

Rosenbaum, P.R., Rubin, D.B., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70, 41–50.

Sjaastad, E., Bromley, D.W., 1997. Indigenous land rights in Sub-Saharan Africa: Appropriation, security and investment demand. World Dev. 25, 549–562.

Toulmin, C., 2009. Securing land and property rights in Sub-Saharan Africa: The role of local institutions. Land Use Policy 26, 10–19.