Working Paper No. 2011/46

Poverty and Land Distribution

Evidence from a Natural Experiment

Malcolm Keswell[1] and Michael Carter[2]

August 2011

Abstract

While land reforms have long been motivated as a potential policy lever of rural growth and development, there is remarkably little evidence of the direct impacts of such reforms. In an effort to fill this lacunae, this paper examines South Africa's Land Redistribution for Agricultural Development (LRAD) program. We show that the implementation of this program operates as a natural experiment in which self-selected and administratively-filltered LRAD applicants receive land transfers at random points in time. This random exit from the application pipeline creates creates exogenous variation in treatment assignment as well as treatment duration. Exploiting both sources of exogenous variation, we estimate average and long-run treatment effects that imply a discounted gain in monthly per capita consumption of about fifty per cent after three years of exposure to the program.

Keywords: land reform, poverty, impact evaluation
JEL classification: O10, O12, O13

[1]Southern Africa Labour and Development Research Unit, School of Economics, University of Cape Town, email: malcolm.keswell@uct.ac.za; [2]Department of Agricultural and Resource Economics, University of California-Davis, email: mrcarter.ucdavis.edu.

This study has been prepared within the UNU-WIDER project on Land Inequality and Decentralized Governance in LDCs, directed by Pranab Bardhan and Dilip Mookherjee.

UNU-WIDER acknowledges the financial contributions to the research programme by the governments of Denmark (Royal Ministry of Foreign Affairs), Finland (Ministry for Foreign Affairs), Sweden (Swedish International Development Cooperation Agency—Sida) and the United Kingdom (Department for International Development).

ISSN 1798-7237 ISBN 978-92-9230-413-3

Acknowledgements

We especially thank Klaus Deininger as well as Pranab Bardhan, Hans Binswanger, Stephen Boucher, Ben Cousins, Alain de Janvry, Michael Kirk, Heinz Klug, Marcus Goldstein, Michael Lipton, Dilip Mookherjee, Jolyne Sanjak, and Francis Wilson for helpful comments on earlier versions of this work, as well as seminar participants at UCT, UJ, Madison-Wisconsin, Stellenbosch, Inter-American Development Bank, the BASIS CRSP conference on Escaping Poverty Traps, the World Bank, and the UNU-WIDER conference on Land Inequality and Decentralized Governance in LDCs. Tim Brophy, Susan Godlonton, Simon Halliday, Ronelle Ogle, Victor Orozco, and Heather Warren provided excellent research assistance.

UNU World Institute for Development Economics Research (UNU-WIDER)
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

A rich, though largely theoretical literature suggests that asset inequality can be economically detrimental if it leaves large numbers of low wealth agents unable to fully utilize their endowments and productive opportunities because they lack access to necessary contracts for financial or other resources.[1] Corespondingly, asset transfers to the poor are a potential win-win in this literature, reducing poverty and enhancing aggregate productivity.[2] By enabling lower wealth agents to realize their economic capabilities, asset transfer programs would thus seem to offer a distinct advantage as an anti-poverty program relative to the conditional cash transfer programs that have become a preferred tool in the anti-poverty toolkit.

Yet there has been relatively little empirical demonstration of the effectiveness of asset transfer programs, especially relative to the outpouring of empirical analyses of cash transfer programs. Taking advantage of a natural experiment that emerged from South Africa's market-assisted land reform program, this paper provides some of the first estimates of the poverty impacts of land redistribution that plausibly satisfy the strong exogeneity assumptions regarding treatment assignment. While our analysis primarily focuses on the intrinsic value of land to reform beneficiaries, our findings of large anti-poverty impacts and high rates of return to public funds invested in the program are consistent with the win-win logic that could make asset transfer programs smart public policy.

The relative scarcity of empirical evidence on asset transfers is understandable. While, the last century has seen any number of land redistribution schemes, they are hard to evaluate because the events giving rise to them are decidedly non-random, and options for credibly identifying their impacts are few. At the macro-level, much of the evidence on the instrumental value of egalitarian land ownership is derived from cross-country growth regressions, so the usual questions of data comparability and differences in historical legacies across societies confound causal interpretation.[3] At the micro-level, there is a hyper-abundant literature that explores the inverse farm size-farm productivity relationship.[4] While linked to the traditional "economic case for land reform," few, if any, of the studies in this literature explicitly concern redistributive reforms.[5] Most explore data generated by the historical farm size distribution,

[1]Foundations for this perspective are found in static models such as Dasgupta and Ray (1986), Eswaran and Kotwal (1986), Bardhan, Bowles and Gintis (2000), and Mookherjee and Ray (2002). Subsequent work has shown that overall economic performance can remain sensitive to asset inequality even in the context of dynamic models in which agents have time (and optimal savings plans) as another degree of freedom to work around missing contracts and financial markets. For example, see Ray and Streufert (1993), Carter and Zimmerman (2000) and Galor and Zeira (1993)

[2]As first noted by Dasgupta and Ray (1986), these transfers may not be strictly Pareto improving, despite their win-win characteristic.

[3]Deininger and Olinto (2000) is perhaps the most convincing analysis of this genre because of their use of fixed effects methods to control for intrinsic country heterogeneity.

[4]See for example Berry and Cline (1979), Carter (1984), Shaban (1987), Rosenzweig and Binswanger (1993), Binswanger, Deininger and Feder (1995), and Lipton, Eastwood and Newell (2009). See also Lipton, Ellis and Lipton (1996) and Van Zyl, Kirsten and Binswanger (1996) for the case of South Africa.

[5]Dorner (1970) and Kanel (1968) articulate the classic instrumental economic case for land redistribution, arguing that agency costs in agricultural labor markets create an inverse relationship between farm productivity

raising a plethora of identification concerns (i.e., are “naturally occurring” small farms intrinsically more productive because they have better soils, better farmers, etc.) More relevant to this paper, there are few studies that directly explore the causal impact of land redistribution on the living standards of reform beneficiaries.

The absence of such studies in part reflects the shift of land policy over the last 25 years away from redistribution and toward land titling and tenancy reform programs.[6] But it also reflects the complexity of evaluating interventions as politically potent and economically complicated as land redistribution. While randomized controlled trials (RCTs) have been used to study modest asset transfers, no country to date has implemented land redistribution by randomly distributing land to some, but not other, potential beneficiaries.[7]

This paper shows that the on-the-ground implementation of South Africa’s Land Redistribution for Agricultural Development (LRAD) programme mimics an RCT. LRAD makes available land purchase grants to landless farm workers and labor tenants. Because it does not mandate redistribution of land and operates on a willing buyer – willing seller basis, LRAD operates as a natural experiment in which self-selected and administratively-filtered LRAD applicants receive land transfers at random points in time unrelated to the characteristics of the applicant or of the land being purchased. This random exit from the LRAD application pipeline creates a natural control group of approved LRAD applicants who were, as of the survey date, denied asset transfers for exogenous reasons. It also creates random variation in the extent or duration of exposure to the asset transfer program amongst those who have received land transfers. We will here exploit both sources of exogenous variation in program exposure to identify both an average treatment effect and a long-run policy-relevant treatment effect. We find that land transfers associated with LRAD lead to very strong welfare gains for program participants. Accounting for heterogeneity in the length of exposure to the program, we find long-run treatment effects of 20-30%, which translate to an annual return that is sufficiently high to offset the upfront direct costs per household within three years.

The remainder of this paper is structured as follows. Section 2 outlines the core challenge of evaluating the impact of the LRAD program and shows how its implementation generates a natural experiment in which treatment assignment works as a virtual lottery. Section 3 then devises and implements a strategy for estimating the average treatment effect of LRAD land transfers under this virtual lottery. Section 4 then expands the analysis to look at treatment heterogeneity in the duration (or exposure) to the program, which we estimate through a duration of treatment response function. Section 5 concludes the paper with reflections on the comparative efficacy of land reform as an anti-poverty program.

---

and size such that aggregate output and economic performance increase with land redistribution.

[6] Both titling and tenancy reform policies have been extensively studied (see for example, the review in World Bank (2003)). The recent literature on tenacy reform has been dominated by studies of the Indian experience–see for example Banerjee, Gertler and Ghatak (2002) and Bardhan and Mookherjee (2008).

[7] See de Mel, McKenzie and Woodruff (2008) for an example of an RCT concerning physical capital.

2 Identifying the Impact of LRAD Land Redistribution on Poverty

The active dispossession of the land rights of South Africa's majority population was a cornerstone of colonial policy in South Africa and the apartheid era that followed. Consequently, it is not surprising that the first democratically elected government brought forward a land reform rhetoric and agenda in 1994. Ironically, the very thoroughness of dispossession had all but eliminated the conventional peasant smallholder class that was the economic focal point of land reform in other countries. In South Africa of the mid-1990s, land reform was consequently pursued as a restitution of legal rights, with lesser attention to securing the economic benefits typically associated with land reform efforts. The result, perhaps predictably, was an ineffective program that met with sluggish uptake. Reacting to this reality, the South African government overhauled its land reform program in 2001, creating the Land Redistribution for Agricultural Development (LRAD) program.

Styled on the market-assisted land reform model, LRAD was intended to provide land to individuals with an interest in farming. The program works on the basis of a grant that is awarded to beneficiaries on a sliding scale. The minimum grant of R20 000 requires a matching contribution (in cash or in-kind) of R5 000. The maximum grant of R100 000 requires a matching contribution amount of R400 000. In practice, grants are pooled into a fund that is administered on behalf of the beneficiaries by the state or a project formed by the prospective beneficiaries. These funds are then used to purchase land, which is then transferred to the beneficiaries.

As with other market-assisted programs, LRAD was intended to rely on beneficiary self-selection, assuring that resources would be channeled to those most interested in farming, and best able to benefit from the redistribution. While this feature is attractive from a program design perspective, this heavy reliance on self-selection makes more difficult the challenge of evaluating the impacts of LRAD on household economic well-being.

Using conventional notation, the impact evaluation challenge is to reliably estimate the counterfactual economic status of beneficiaries had they not been treated, $\mathrm{E}(y_0|\mathbf{x}, D = 1)$ so that the average treatment effect on the treated can be caluculated as:

$$\mathrm{ATT} = \mathrm{E}(y_1|, \mathbf{x}, D = 1) - \mathrm{E}(y_0|\mathbf{x}, D = 1)$$

Unfortunately, the very essence of the LRAD program and its reliance on self-selection, implies that the mean outcomes of the non-beneficiary population is unlikely to provide a plausible approximation to the counterfactual outcomes of a randomized out sample of would-be treated households.

Our solution to this evaluation problem rests upon a detailed understanding of the mechanics of the LRAD program. In brief, we will show that the implementation of LRAD creates a natural control group of individuals subject to the same double selection process as beneficiaries, but whose receipt of program benefits have been delayed for exogenous reasons. Before detailing this identification stragety, we need to first plumb the details of the LRAD program in practice.

2.1 The LRAD Application and Approval Process

Primarily targeted at small-scale black farmers and farmworkers, LRAD has 5 stages. Would be beneficiaries of LRAD grants must pass through the following five stages:

1. Project Registration: The first stage in the LRAD approval process is registration of applications to the program. Once an application is submitted, a state appointed official (hereafter referred to as the "planner") does a needs assessment by visiting the site on which the applicants live as well as the land they have applied to purchase (which need not coincide with the the current place of occupancy of the applicants). Once the application has been verified, the application is "registered" as as candidate land redistribution project.

2. Approval of Planning Grant: Stage 2 begins when the planner requests the district line authority of the land affairs department to release a nominal sum of money to begin developing a proposal on behalf of the applicants. The funds are meant to be used to commission various specialised activities that will culminate in a portfolio of sorts that will ultimately be used by the planner both to negotiate a purchase price for the land, as well as to serve as a basis for the deliberation that will occur over final approval of the application. Examples of such activities are valuations, soil assessments, land surveys, and business plans.

3. Preparation of Project Identification Report: Once these commissioned studies start to materialize, the planner begins to work with the applicants in an effort to iterate towards a workable proposal which will ultimately be submitted to the state as background motivation for the application. This proposal preparation stage is an important process that is handled through a series of workshops between relevant role-players and culminates in the preparation of a document called the project identification report (PIR), which summarizes the merits of the application. The existence of this document is an important milestone in the approval process as it signals that the applicants have a strong enough interest and background in farming to have warranted the release and expenditure of state resources to begin making the case for the grant.

4. Approval by District Screening Committee: In stage 4, the planner then submits this document to a district-level screening committee (DSC) of the land affairs department. The primary purpose of the DSC is to vet applications so as to improve their likelihood of approval when submitted for consideration to the provincial grants approval committee (PGAC), which is the the main body tasked with granting final approval of the application. This body has broad representation from all role players including officials from the agriculture department, surveyor general's office and local municipalities. The role of the DSC is to pre-screen applications before they are passed on for final approval by the PGAC. Once an application has been approved by the DSC, a formal request to designate the land for redistributive purposes is made. At this stage a quasi-legal document called the "designation memo" is prepared, which is what the provincial grants approval committee deliberates over when making their final decision. This document must ultimately be signed by the directors general and ministers of land affairs and agriculture. A key hurdle that applications reaching stage 4 must overcome is that there must be complementarity around basic service provision (roads, irrigation, electrification).

5. Final Transfer by Provincial Grants Approval Committee: Beyond its complexity, there is nothing especially remarkable about this multi-stage application and approval process. However, applicants often experience random delays in this final stage of approval. Interviews with program administrators in the field revealed that in practice stage 4-approved applicants often become mired in stage 5, facing considerable and highly variable delays in receiving program benefits, if they receive them at all. Two key reasons were invariably cited by program staff for why there might be a delay at stage 5. Firstly, there could be a competing claim for the land offered up by the seller. For example, a fairly frequent occurrence is where the land that is offered up for sale by the current landowner turns out to be the subject of a legal dispute between the seller and third-party descendants of previously dispossessed individuals under Apartheid.[8] Secondly, the seller might withdraw from the agreement at the last minute as a strategy to renegotiate pertinent terms of the agreement such as the selling price.[9] In probing specific examples where such events had taken place, it became clear that although not infrequent, these reasons for a delay in approval had little, if anything, to do with beneficiary characteristics. Reflecting the randomness of this process, we will refer to exit from this process, and receipt of a land transfer, as the Stage 5 Lottery.

Note that applicants caught in stage 5 of the pipeline have passed through the same two selection processes or filters passed by LRAD beneficiaries: they have self-selected into the program and their business plans have been similarly scrutinized, modified as needed and

[8]Forced removals under Apartheid had devastating consequences in both urban and rural areas. A major thrust of the Mandela-era land reforms focused on restoring the rights of the previously dispossessed to their descendants. This program, known as the "restitution" program, was terminated at the end of of the 1990's. However, many of the claims made under that program have proved to be quite complex to resolve. Cases are usually held by a special court set up to hear such disputes, which, in some instances can end up never being fully resolved. The overhang of this program, and the fact that the bureaucratic structures that handle these claims operate somewhat separately from programs like LRAD, often lead to situations where much progress can be made on an LRAD application, only to discover that the PGAC has to withhold final approval until the outcome of the dispute is resolved in a land claims court.

[9]A reason less frequently cited for delays in final approval of an application concerns a failure by local councils to meet their targeted roll-out of basic infrastructure in a district affected by an impending redistribution of land.

approved by the LRAD technical staff. The random exclusion of these individuals from benefits (or random delay in their receipt of benefits) provides the basis for the two identification strategies that we utilize here. One strategy will utilize the double filtered applicants still stuck in stage 5 (the Stage 5 Lottery losers) as our counterfactual of interest $\mathrm{E}(y_0|\mathbf{x}, D = 1)$, in order to estimate an average treatment effect of the LRAD program. The second strategy will utilize the random timing in the exit of applicants from stage 5 as the basis of a continuous treatment estimator that identifies the impact of length of exposure to the LRAD program on the outcomes of the treated.

2.2 Sample Construction and Characteristics

To implement the identification strategy outlined above, we began by constructing sampling frames comprising listings of all households that had either applied for or that had received land through the LRAD programme. The programme is implemented at the farm level, so the sampling follows a multi-level design, where the farm currently occupied (control group) or owned (treatment group) represents level 1 of the sampling design.[10] A random sample of farms with probability proportional to the number of beneficiary or applicant households on the farm was drawn from each sampling frame.

Level 2 of our sampling design concerns the process of screening and filtration that happens as an application nears final approval. As mentioned above, the existence of a "designation memo" measures whether an application successfully navigated the first four stages of the approval process. We therefore checked for the existence of this document for applicants in the LRAD pipeline that had been sampled for study. Where this type of documented evidence of stage 4 traversal could not be located, we interviewed LRAD program staff and delved further into archived PIRs of applicants so as to establish a timeline of milestones that would have been achieved as the application neared final approval. Ultimately, if we determined that a sampled application had not passed stage 4 of the approval process, a replacement applicant was found that did meet this criteria. The third and final stage of sampling happens at the farm level. For sampled treatment farms as well as control group farms that had survived the screening process, a random sample of households was then drawn and interviewed.

The final sample comprises 1650 households, of which 448 households are in the treatment group and 1202 households are the twice filtered, post-stage 4 control group households. Table 1 shows the means and standard deviations of the key variables used in our analysis.

Since we are interested in measuring the impact of LRAD on poverty, our welfare metric of choice is monthly per capita consumption.[11] We explicitly choose not to use a binary indicator

[10]Applicants are usually farm workers or labour-tenants who have applied to the LRAD programme to purchase the land which they farm, or are subsistence farmers who wish to purchase a subdivision of a neighbouring or nearby farm.

[11]We have also conducted our analysis using alternative deflators of household consumption than household size, such as the number of adult-equivalents in the household, estimated using Engel's method. These alternate forms of normalization do not change any of our substantive conclusions or interpretations, so we only report results for per-capita consumption.

of poverty status since this is a less efficient way of defining poverty outcomes when working with quasi-experimental data designs (Ravallion 2008).[12]

Table 1 shows that mean consumption in treatment households is 128.1 rands higher than the control group mean consumption, and this difference is significant at the 1% level. In anticipation of later continuous treatment analysis, Table 1 also disaggregates treated households into terciles based on the duration of time since they had received their land transfers. As can be seen, mean per-capita consumption consistently increases as the period of time since land transfers increases from less than two years, to 2-3 years to more than 3 years. If we rely solely on the efficacy of the natural experiment created by LRAD implementation, then these differences can be attributed to the effect of the program. However, as we argued earlier, it is possible that imperfections in the natural experiment may leave some differences between treatment and control groups. To test for the robustness of our measures, we construct and examine variables that could proxy for the processes that appeared to matter qualitatively in the screening processes discussed above.

A possible confound not addressed by the natural experiment that plays out through the stage 5 lottery is that applicants that applied later may be less likely to have received land transfers by the survey date, or might have received it for a lesser period of time. To explore this possibility, we formed the variable *application date*, which measures the number of days elapsed from the date of application to the LRAD program, to the date on which the household was first interviewed. As can be seen in Table 1, treated households on average applied 128 days (or about 4 months) earlier than did control households. Similarly, households with more than 3 years of exposure to land transfers applied on average 132 days earlier than households with less than 2 years of exposure to land transfers. While these 4 month differences are modest, there could be some concern that earlier applicants are somehow different than later applicants. As can be seen in Table 1, earlier applicants are in fact better educated. In addition, we might worry that these earlier applicants may be more entrepreneurial or exhibit other unobserved characteristic that predicts economic success. If this latter concern is in fact correct, then application date should in fact serve as a proxy for these unobserved characteristics. In both the binary and continuous treatment analysis that follows, we will employ application date as one of the variables on which we match across treatment status groups.[13]

In addition to these timing variables, Table 1 presents descriptive statistics on other characteristics of the control and treatment groups. Interestingly, gender, education, and a revealed preference to move in order to gain approval all appear to be statistically significant in sepa-

[12]We also do not normalise expenditure by a poverty line, because there is some controversy in the case of South Africa as to which is the most appropriate line to use (Woolard and Leibbrandt 2007).

[13]In addition to concerns about application date, we might also worry that a 'good' applicant will progress more rapidly through the LRAD approval process then a 'bad' applicant. Ideally, we would like to match on a measure of duration of time that the individual spent in the first four stages of application pipeline. Unfortunately, we do not have such a measure. Fortunately, given that we can control for application date (which is likely to also be correlated with this time in the pipeline) and given that all the households used for the study passed through the first four stages, this residual source of bias can at most be quite modest.

rating the treated from control group households, as is evidenced by the low p-values of a test of equality of means for these variables. To explore these issues further, the next section uses these variables to estimate the probability of being treated.

3 Binary Treatment Impacts of the LRAD Program

A possible concern with our identification strategy is that the initial application date might be factor that affects the timing of the exit from stage 5. That is, applicants that choose to apply earlier might be expected to exit the pipeline and receive a longer duration of LRAD treatment. We examine this concern empirically in this section. As we have already discussed in section 2.2, the implicit lottery that takes place in stage 5 of the approvals process does a reasonably good job of washing out the correlation between the date of application and observed treatment status. However, to better account for potential non-randomness of entry in the LRAD pipeline, we use propensity score matching.

3.1 Modeling the Probability of Treatment

A major drawback of models based on the selection-on-obsersables assumption is that there is no resolved method for choosing among possible sets of covariates. There is however some evidence from replication studies to suggest that a greater set of controls leads to a reduction in bias.[14] However, a mechanical approach whereby one tries to include the most exhaustive set of available covariates can often be a double-edged sword. Under conditional mean independence, matching on the propensity score must balance the data. While there are a range of tests to check for the balancing property, these tests are more informative about how to go about picking the correct functional form for a specific set of regressors, rather than how to choose between alternative sets of regressors (Smith and Todd 2005).[15] If one simply chooses the most exhaustive set of regressors that maximizes the prediction rate, say according to the hit-or-miss criterion, this might be at the cost of balancing the covariates.[16] Matching on the propensity score alone therefore does not necessarily solve the identification problem.[17]

However, by combining the insights gleaned from the administrative screening process outlined in section 2.1, with the apparent lottery that takes place in stage 5, we are able to model the selection probability such that the key assumptions required for matching are satisfied. Key

[14]By "replication study" we mean a study where an estimate of $\mathrm{E}(y_0|\mathbf{x}, D = 1)$ from a randomised-out control group of the $D = 1$ sub-population serves as a benchmark against which non-experimentally derived counterfactuals are compared. See for example Heckman, Ichimura, Smith and Todd (1998) and Smith and Todd (2005).

[15]For example, Dehejia and Wahba (2002) present an algorithm that is based on Theorem 2 of Rosenbaum and Rubin (1983). See Eichler and Lechner (2002) for an alternative approach.

[16]The hit or miss method to maximize the within-sample correct prediction rate counts as a correct prediction $p(\hat{\mathbf{x}}) > p(\tilde{\mathbf{x}})$. If one assumes a symmetric loss function, then it is natural to set $p(\tilde{\mathbf{x}}) = 0.5$. See Heckman et al. (1998) for an example.

[17]See Heckman, Ichimura, Smith and Todd (1996), Heckman, Ichimura and Todd (1997), Heckman et al. (1998), and Smith and Todd (2005) for more on these issues.

to our approach are our qualitative observations of the approvals process: specifically, that the final approval (and hence selection into the treatment group) is tantamount to a lottery for applicants that successfully traverse stages 1–4. If this is true, all that remains is to control for entry into the pipeline (i.e., a measure of application date), as well as variables that are likely to covary with application date. The most important variables (already outlined in table 1), again are informed from our qualitative study of the LRAD approvals process, and from detailed face-to-face interviews with staff of the LRAD program.

We model the probability of selection into the treatment group parametrically by running a logit regression. The dependent variable takes a value of 1 if a household is in the treatment group, and a value of 0 if it is in the control group and has passed stage 4 of the approval process. Table 2 shows two specifications of the logit regression. The first specification contains a parsimonious list of regressors, whereas the second contains additional controls and a more flexible specification.

The signs of the variables in the parsimonious regression can be read off the table directly, and are informative in terms of understanding the screening process of LRAD. Two immediately noticeable features, already highlighted by the descriptive statistics, is that the apparent targeting of women by the LRAD program does not seem to be borne out by the data, since female-headed households have a lower probability than male-headed households of gaining admission into the program. Another interesting finding is the role of farming experience. In the parsimonious specification, this variable appears insignificant, which would accord with our intuitions about the LRAD screening process: individuals with little or no farming experience would be screened out prior to stage 4 so it is not surprising that this variable turns out to be insignificant in the propensity score regression. Yet, it would be reasonable to conjecture that the effect of farming experience and other characteristics of the applicant like education affect the selection probability non-linearly.

The second specification of the propensity score regression lends credence to this suspicion. While the direction of the effects of education and farming experience are not as readily apparent because of the non-linear specification, it is clear that both of these variables combine non-linearly with the timing of entry into the applicant pool to significantly affect the treatment probability. This specification also dominates more parsimonious specifications (including several not reported here) in separating the treated from the control households, with a prediction rate of 76.36%.

A key result from modeling the treatment probability in this way is shown in figure 1, which depicts the supports of the estimated log-odds ratio. This figure allows a graphical assessment of the common-support assumption that is required for use of matching estimators. Figure 1 clearly shows considerable overlap in supports for parts of each of the two empirical densities with greatest concentrations of mass. Moreover, it is clear that even at very high probabilities (log-odds of over 1.5 say), there are more than a handful of observations in the control group. Thus, our binary treatment estimator that matches households non-parametrically over this space will generate treatment effects that are valid for virtually the full support of our matched sample.

3.2 Balancing Tests

The balance test we employ works as follows: the support of the propensity score is partitioned into disjoint sets such that: (i) the average propensity score within each of these sets is uncorrelated with treatment assignment and; (ii) the covariates are uncorrelated with treatment assignment within each set. This idea is operationalized by arbitrarily grouping the data by intervals of the propensity score, where initially the scores within each interval are quite similar. An equality of means test between treatment and control observations is performed for each of the regressors contained in $\mathbf{x}$. If there are no statistically significant differences between treatment and control for each of the covariates in the propensity score regression, then the regressors are balanced. If a particular regressor is unbalanced for a particular interval, then that stratum is split into further strata and the test is conducted again. This iterative process continues until all the regressors are balanced. Tables 3 – 4 shows a summary of the results from this test.[18]

There are 7 bins of the propensity score in the final iteration of this balancing algorithm, and as table 3 shows, in each case the p-value for an equality of means test of the propensity score between the treatment and control groups is small enough to fail to reject the null hypothesis of equality of means.

The final step in checking for covariate balance is shown in table 4, which presents the p-values of an equality of means test for each of the regressors used in the propensity score regressions reported in table 2. As is evident from the results, each of the covariates contained in $\mathbf{x}$ are well balanced across the 7 bins.

3.3 Empirical Estimates of Average Treatment Effects on Treated

We start with the standard GLM framework by modeling treatment assignment as a Bernoulli trial. The probability mass function therefore is

$$f(D_i|\mathbf{x}_i) = p_i^{D_i}(1-p_i)^{1-D_i}, \qquad D_i = 0, 1 \tag{1}$$

where $D_i = G(\mathbf{x}_i'\boldsymbol{\beta})$, and where $\mathbf{x}$ is a $k \times 1$ column vector of covariates. Thus, the odds of being treated are given by $f(1) = p^1(1-p)^0 = p$, whereas the odds of not being treated are given by $f(0) = p^0(1-p)^1 = 1-p$.

As is implied by our discussion in section 3.1, we specialize $G(\mathbf{x}_i'\boldsymbol{\beta})$ to the logistic cdf when modeling p_i. This representation of p_i is especially useful in our context because $e^{\mathbf{x}'\boldsymbol{\beta}}/(1+e^{\mathbf{x}'\boldsymbol{\beta}})$ is the canonical link function of the Bernoulli density, and this link function is sufficient to satisfy the key first-order conditions of the maximized value of the log of equation 1.

[18] See Dehejia and Wahba (2002) for more on this algorithm.

To define an average treatment effect on the treated, we start by denoting S_p as the region of common support of p_i between the $D = 1$ and $D = 0$ distributions. Let N_1 denote the set of households that have already received land through LRAD, and let N_0 denote the set of households still awaiting final approval (i.e., households that have passed stage 4 of the approval process). Now denote as n_1 the number of treated households falling into the common support region of the estimated propensity score density; i.e., the number of households falling into the set $N_1 \cap S_p$. Our matching estimator is then given by

$$\begin{aligned}
\delta &= (n_1)^{-1} \sum_{i \in N_1 \cap S_p} \Big(y_{1i} - \hat{\mathrm{E}}(y_{0i} | D_i = 1, p_i) \Big) \\
&= (n_1)^{-1} \sum_{i \in N_1 \cap S_p} \left(y_{1i} - \sum_{j \in N_0} \omega(i,j) y_{0j} \right) \qquad (2)
\end{aligned}$$

where $i \in N_1 \cap S_p$ denotes the ith treated household from the set of households with common support on p_i. The second term in this expression serves as a matched substitute for the outcomes of a randomized-out household of the treatment group, where the imputed counterfactual outcome $\sum_{j \in N_0} \omega(i,j) y_{0j}$ is a kernel-weighted average over the set of possible matches, with weight function:

$$\omega(i,j) = K \left(\frac{\mathbf{x}_j' \boldsymbol{\beta} - \mathbf{x}_i' \boldsymbol{\beta}}{h_n} \right) / \sum_{k \in N_0} K \left(\frac{\mathbf{x}_k' \boldsymbol{\beta} - \mathbf{x}_i' \boldsymbol{\beta}}{h_n} \right) \qquad (3)$$

where K is a kernel function, h_n is a bandwidth parameter and $\mathbf{x}_i' \boldsymbol{\beta}$ is the log of odds ratio.[19] A notable feature of this weight function is that it accounts for the choice-based nature of our survey design without the need for precise sampling weights, because matching takes place on the log of odds ratio. As Heckman and Todd (2009) have shown, since the odds ratio that is estimated using incorrect weights is a scalar multiple of the true odds ratio, and because the odds-ratio is a monotonic transformation of the propensity score, one can match on the log-odds ratio without reweighting equation 2.[20]

In our empirical estimates, we check that our average treatment effect δ is robust to bandwidth specification and kernel choice.[21] Table 5 shows our estimates of the average treatment effects that are generated when matching non-parametrically against the log-odds ratios that

[19] It is usual to match directly on p_i. However, here we match on the log-odds ratio, for reasons we discuss below. Recall that $p_i = G(\mathbf{x}_i' \boldsymbol{\beta}) = e^{\mathbf{x}' \boldsymbol{\beta}} / (1 + e^{\mathbf{x}' \boldsymbol{\beta}})$, and $1 - p_i = 1/(1 + e^{\mathbf{x}_i' \boldsymbol{\beta}})$. Thus $\ln(p/(1-p)) = \mathbf{x}_i' \boldsymbol{\beta}$.

[20] The true sampling weights are unknown in our survey design because both the control group and treatment group sampling frames contained many listings with missing information on the number of applicants at the farm level (which was the first stage of sampling). This missing information matters because the design was meant to be self-weighted at the farm level. To correct for this missing information cell-median imputation was used. These imputations potentially distort the sampling weights, thus exacerbating the choice-based nature of the quasi-experimental survey design. Matching on the odds ratio therefore has the strong benefit of holding this aspect of the design constant.

[21] We also assume this kernel function is symmetric around zero and integrates to one. In estimating this kernel, we experiment with a range of functional forms, but we benchmark against the Guassian kernel function, given by $K = (\sqrt{2\pi})^{-1/2} \exp(-s^2/2)$.

are derived from the estimates presented in table 2 over the region of common support shown in figure 1. Estimates are presented for each of five choices of functional form for K, with optimal bandwidths computed according to the plug-in approach of Silverman (1986). This means that the optimal bandwidth will vary by kernel choice. We benchmark these estimates against the Guassian kernel with a fixed global bandwidth of 0.05, which gives an upper bound average treatment effect of 28.36%. This estimate declines to 22.77% when we apply the optimal plug-in bandwidth of 0.26 for the Gaussian kernel. This estimate is remarkably similar to the average treatment effects found when applying the Epanechnikov and Quartic kernels.

The general conclusion we reach when comparing the results of table 5 is that our estimates of average treatment effects do vary by choice of bandwidth, but in every case, our estimates of the average impact of LRAD for LRAD beneficiaries always exceeds a 20% change in per capita monthly consumption.

These are very sizable effects. In the next section, we explore further whether these average impacts mask other features of selectivity that a binary treatment approach cannot account for.

4 Measuring Dynamic Impacts

Because LRAD land transfers emerged as part of a natural, as opposed to a controlled experiment, beneficiary households received land grants at different points in time. Within our sample, some treated households had received land five years before the survey date, while others had only recently received transfers. The impact on the level of economic well-being of the latter group might well be negligible if no income had yet been generated by the newly acquired land. Indeed, if these households were busy investing in the newly acquired land, the short run impact on household income and consumption could even be negative. In this case the significantly positive average treatment effect reported in the prior section would be a data-weighted average of the zero or negative impact of the program on recent transfer recipients combined with the positive impact on longer duration transfer recipients.

Full evaluation of the efficacy of the LRAD program requires information on the shape of the "duration response function" and the long-run treatment effect to be enjoyed by asset transfer beneficiaries.

In addition to this possible first year dip in living standards when individuals receive the asset transfer, there are two other reasons why the impact of an asset transfer may change over time. First, the beneficiary may experience a learning effect with technical and entrepreneurial efficiency improving over time. Second, and consistent with the theoretical literature on asset inequality discussed earlier, the asset transfer may create a crowding-in effect if the beneficiary further invests in the new enterprises made possible by the LRAD grant. Whether this additional investment occurs because of improved access to financial markets, or because of learning effects, the overall impact on inequality in the long run is likely to be significant, if

LRAD transfers suffice to lift households over the critical threshold asset level in a multiple equilibrium model above which it is dynamically rational to accumulate further assets and move ahead.[22] It is these second round effects that are likely to distinguish asset transfer programs from the many other policy instruments that might constitute a package of reforms aimed at rural development.

If these observations are correct, then the duration response function – meaning the relationship between program impact and duration of time since the asset was transferred – is unlikely to be a simple step function that can be approximated with a binary treatment estimate. The goal of this section is to estimate the impact dynamics and duration response function and recover both the long-run impact of land transfers and their time path. Both are of particular relevance from a policy perspective. Indeed, it is the prospect that asset transfer programs, such as land reform, facilitate and crowd-in additional asset building that makes them especially interesting as an anti-poverty program.

4.1 Modeling the Duration Response Function

A natural starting point for this analysis ordinarily would be to consider a random coefficients model (Heckman and Robb 1985). However, as we have argued above, the variables predicting treatment status enter non-linearly in the logit regression of treatment, which implies that the minimum mean square error approximation to the underlying CEF (i.e., the non-linear analogue to a linear regression control-function) does not have a straightforward interpretation, because the regression coefficients in such a model would actually represent a matrix-weighted average of the gradient of the CEF (Chamberlain and Leamer 1976). This problem can be overcome if the CEF of our treatment duration variable were restricted to be linear, as in Angrist (1998), but then we would still require a large number of observations for each value of treatment duration to justify this approach. An alternative approach that would not necessitate such an assumption is the extension of the propensity score approach developed by Hirano and Imbens (2004).

We begin by restricting attention to the sample of households in the treatment group, $i \in N_1$. We then postulate a duration-response function $y_i(d)$ for all $d \in \mathcal{D}$ given that $i \in N_1$; i.e., each household could have any potential outcome from the set $\mathcal{D}$ depending on its duration of treatment. When treatment status is binary, we have $\mathcal{D} = \{0, 1\}$, but here we let $\mathcal{D} = \{d_0, d_1\}$. In the empirical implementation, we measure duration as the number of days between land transfer and the date of the survey.

[22] Formal models of the emergence of poverty traps highlight the interplay of investments and occupational structure. A key feature of this literature is the idea that non-convexities in the production of human capital are induced by indivisibilities in its investment as well as imperfections in credit markets. In this class of models, the shape of the aggregate distribution of occupations (and therefore long-run inequality) is strongly dependent on the opportunity sets of the previous generation. Galor and Zeira (1993) is the classic reference in this literature. Galor, Moav and Vollrath (2009) and Carter and Barrett (2006) and Mookherjee and Ray (2002) are recent extensions to this framework that suggest a strong role for asset redistribution of the kind that the LRAD program is concerned with.

The evaluation problem of course results from the fact that each household realizes exactly one outcome, associated with its actual duration in the LRAD program $y_i = y_i(D_i)$, where $D_i \in [d_0, d_1]$. However, under the continuous treatment case, the problem is further complicated by the fact that there is more than one counterfactual duration, because we have $D_i \in \mathcal{D}$. We therefore define the impact of LRAD in this continuous case in terms of an average duration-response function, $\mu(d) = E[y_i(d)]$. Our goal then is to uncover non-constant treatment effects by taking the difference between this average and some benchmark level of treatment; namely

$$\theta(d) = \mu(d) - \mu(\tilde{d}) = E[y(d)] - E[y(\tilde{d})] \qquad \tilde{d}, d \in \mathcal{D} \tag{4}$$

where $\tilde{d}$ serves as the benchmark duration.[23] In our empirical estimates, we set $\tilde{d}$ to be the lowest length of exposure observed in the data.

As in the binary approach, statistically valid identification depends on an independence assumption regarding treatment assignment. Following Imbens (2000), we define weak unconfoundedness:[24]

$$y(d) \perp D | \mathbf{x} \qquad \forall d \in \mathcal{D}$$

To fix ideas, define $r(d, x)$ as the conditional density of treatment duration given the covariates

$$r(d, x) = f_{D|\mathbf{x}}(d, x) \tag{5}$$

and define a generalized propensity score (GPS) $R = r(D, X)$. Using this framework, Hirano and Imbens (2004) then show that assignment to treatment duration (or equivalently, the timing of treatment), is unconfounded when $f_D(d|r(d,x), Y(d)) = f_D(d|r(d,X))$.[25]

As discussed in Section 1, all LRAD beneficiaries passed through a rigorous 5 stage approval process. As detailed in that earlier section, time spent in the approval process appears to be orthogonal to beneficiary characteristics and expected gains from the program, $Y(d)$. Conditional on original program application date (which is observed), the LRAD natural experiment conforms to the unconfoundedness assumption.[26]

[23] To simplify the notation, we drop the i subscripting when making reference to realised outcomes or treatment levels.

[24] This is essentially a weaker version of the Rosenbaum and Rubin (1983) "strong ignorability" assumption, generalized to multi-valued treatments. Hirano and Imbens (2004) refer to this assumption as weak unconfoundedness as it does not require joint independence of all potential outcomes, $\{y(d)\}_{t \in [t_0, t_1]}$ but rather that conditional independence holds for each value of D. For alternative approaches to non-binary treatments, see Behrman, Cheng and Todd (2004) and Imai and van Dyk (2004)

[25] For the proof, as well as further details about assumptions regarding the probability space used in the proof, see their theorem 1. Two key assumptions are that D is continuously distributed with respect to Lebesgue measure on $\mathcal{D}$, and that $r(d, \mathbf{x})$ is measurable with respect to the σ-algebra generated by $\mathbf{x}$. Based on these assumptions, they then show that using standard results on iterated integrals, that the right hand side of this expression can be written as $f_D(d \mid r(d, \mathbf{x})) = \int f_D(d \mid x, r(d, \mathbf{x})) \mathrm{d}F_{\mathbf{x}}(x \mid r(d, \mathbf{x})) = r(d, \mathbf{x})$. Then by weak unconfoundedness, they show that this quantity must equal $f_D(d \mid r(d, \mathbf{x}) Y(d)) = \int f_D(d \mid x, r(d, \mathbf{x}), Y(d)) \mathrm{d}F_{\mathbf{x}}(x \mid Y(d), r(d, \mathbf{x}))$, which in turn is equal to the left hand side.

[26] There is a strong analogy between our analysis of LRAD and the Hirano and Imbens (2004) study of the impact of lottery winnings on labor supply. In both cases, the focus is only on 'players,' meaning those self-selected individuals who either applied for an LRAD grant or who purchased lottery tickets. Also in both

Under the assumption of unconfoundedness, Hirano and Imbens (2004) then show how the GPS can be used to identify $\mu(d)$. Two steps are involved in this proof. First, using Bayes rule and their theorem 1, they show that in estimating the conditional expectation of the outcomes, all relevant information about the conditional density of the treatment is controlled for by directly conditioning on the treatment level D and the generalized propensity score $\hat{R}_i$. Second, to estimate the duration-response function, $\beta(d, r(d, X))$, at a particular level of the treatment they average this conditional expectation over the GPS at that particular level of the treatment, $\mu(d) = \mathrm{E}[\beta(d, r(d, X))]$ and then by iterated expectations, $\mathrm{E}[\beta(d, r(d, \mathbf{x}))] = \mathrm{E}[\mathrm{E}[y(d) \mid r(d, \mathbf{x})]] = \mathrm{E}[y(d)]$ obtains.[27] Thus knowledge of $\beta(D, R)$ will identify the average duration-response function, under weak unconfoundedness conditional on the GPS.

To implement this estimator, we follow Hirano and Imbens (2004) and assume that the conditional density of the duration of treatment is normally distributed with mean $\mathbf{x}\boldsymbol{\gamma}^{\mathbf{c}}$ and variance σ^2. These parameters can be estimated by maximum-likelihood, and the estimated GPS recovered as:

$$\hat{R}_i = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}}\big(- \frac{1}{2\hat{\sigma}^2}(D_i - \mathbf{x}_i\hat{\boldsymbol{\gamma}^{\mathbf{c}}})\big) \tag{6}$$

To estimate the duration-response function, we model the conditional expectation of y_i, as a flexible function of D_i and R_i

$$\begin{aligned} \beta(D_i, R_i) = E[Y_i|D_i, \hat{R}_i] \quad = \quad & \alpha_0 + \alpha_1 D_i + \alpha_2 D_i^2 + \alpha_3 \hat{R}_i + \\ & \alpha_4 \hat{R}_i^2 + \alpha_5 D_i \hat{R}_i \end{aligned} \tag{7}$$

Equation 7 is then estimated by OLS.[28] Once we have the α_k, we can then recover the average duration response function $\mathrm{E}[y(d)]$. Recall that $\mathrm{E}[y(d)]$ is identified for particular levels of duration, so the average must be taken over all households at duration level d. This effectively equates to averaging over the GPS for each duration level d. By changing the treatment level at which the averaging takes place, we recover an estimate of the entire duration-response

cases, the unconfoundedness is met as the intensity of treatment (timing of land transfer or amount of lottery winnings) was generated by essentially a random process unrelated to expected impacts of the treatment.

[27] Importantly, note that under this approach, the averaging that is used to construct $\mu(d)$ takes places over the GPS score evaluated at the treatment level of interest, $r(d, \mathbf{x})$, and not over the GPS itself.

[28] It should be stressed that the regression function $\beta(d, r)$ does not have a causal interpretation. In particular, the derivative with respect to the treatment level d does not represent an average effect of changing the level of treatment for any particularly subpopulation. We also experimented with various specifications for this regression and conclude that not much additional explanatory power is added by including higher than second-order polynomials in D and $\hat{R}$.

function.[29] This gives a treatment effect estimator of the form

$$\hat{\mu}(d) = \widehat{E[Y(d)]} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\alpha}_0 + \hat{\alpha}_1 \cdot d + \hat{\alpha}_2 \cdot d^2 + \hat{\alpha}_3 \cdot \hat{r}(d, \mathbf{x}_i) + \hat{\alpha}_4 \cdot \hat{r}(d, \mathbf{x}_i)^2 + \hat{\alpha}_5 \cdot d \cdot \hat{r}(d, \mathbf{x}_i)) \tag{8}$$

Finally, to compute a non-constant effect of treatment on treated, we estimate

$$\hat{\theta}(d) = \hat{\mu}(d) - \hat{\mu}(\tilde{d}) \qquad \forall d \in \mathcal{D} \tag{9}$$

where $\tilde{d}$ has been fixed at the benchmark level discussed earlier.

4.2 Generalized Propensity Score Estimates and Balancing Tests

As noted above, our treatment variable is the duration of exposure to the LRAD program, measured as the number of days elapsed since the date of transfer of the land. Table 6 presents maximum likelihood estimates of the conditional distribution of this treatment variable. We use the same specification to that chosen for the binary case. These estimates assume that our treatment variable conditioned on the covariates is normally distributed. The assumption turns out to be satisfied if we normalize the treatment variable by the maximum number of days a household could be exposed to the program. Specifically, we find that the assumption of normality is satisfied at the 1% level using the Kolmogorov-Smirnov test of normality of the errors.

Our next step is to recover the GPS according to equation 6. To investigate whether the specification used in equation 6 is adequate, we test for balancing of the covariates once we condition on the estimated GPS. The test has the same intuition as that of the binary case, and is based on the algorithm followed by Hirano and Imbens (2004), and the implementation developed by Bia and Mattei (2008).

The basic structure of the test is as follows: we partition the support of $\mathcal{D}$ into three mutually exclusive intervals, denoted as $G_1, \ldots, G_3$. Within each treatment interval G_k, we compute the GPS $r(d_{G_k}, \mathbf{x}_i)$ at the mean of the interval $d_{G_k} \in G_k$. Then, for each of the three intervals we estimate the GPS at these treatment interval means d_{G_k} and then discretize the distribution of the GPS evaluated at this representative point. In our model, we chose 6 mutually exclusive blocks, denoted by $B_1^{(k)}, \ldots, B_6^{(k)}$. Within each interval $B_j^{(k)}$ for $j = 1, \ldots, 6$ and $k = 1, 2, 3$, we compute the difference in means for each covariate across different treatment intervals, but in the same GPS interval (i.e., j is held constant while k is varied). This results in 6 mean differences for each $d_{G_k} \in G_k$. This information is then collapsed into a single metric, by taking a weighted average of the differences at each representative point, where the

[29] We estimate standard errors and confidence intervals for each point along the duration-response function using bootstrap methods. However, in principal, analytical standard errors can also be computed given the parametric forms of the GPS and $\beta(D, R)$.

weights are equivalent to the number of observations within each block $B_j^{(k)}$. This procedure is repeated for each covariate, so that in total, we have 39 such weighted averages (i.e., 13 covariates across 3 treatment intervals). In a final step, these weighted averages are then used to construct test statistics. Table 8 reports Bayes factors that are computed off of these weighted averages. Overall, the model is well balanced as the lowest Bayes factor of 0.3364 falls within an acceptable range of the Jeffery's order of magnitude criterion. Decisive rejection of the null that the data are balanced requires a minimum Bayes factor of less than 0.01.

4.3 Empirical Estimates of Dynamic Treatment Effects

Figure 2 graphs the estimated GPS-adjusted average treatment effect as a function of the duration of treatment. While the model was estimated using the natural logarithm of per-capita monthly expenditures, the estimates were converted back to natural units (South African Rand per-person, per-month) in order to ease interpretation. The corresponding 95% confidence interval (calculated using bootstrap estimates of the standard error) is shown by dashed lines. As can be seen, a household with a trivial treatment duration is estimated to have a per-capita living standard of 400 Rand per-month. By way of comparison, the average per-capita living standard of approved applicant, non-beneficiary household is 466 Rand per month (Table 1) and the South African poverty line–shown on the figure–is approximately 430 Rand.

As is apparent from Figure 2, the treatment response function is far from the simple step function that would indicate binary treatment estimates capture all the relevant dimensions of program impact. Noticeable is the sharp drop in consumption over the first year of exposure, and then a period of quick recovery over the second year, reaching impact magnitudes similar to the binary non-parametric binary ATT estimates reported in Table 5. Beyond the second year, the estimated average treatment effect rises further to about 300 Rand, an increase of almost 75% over the low dosage level, or 45% higher than the mean of the non-treated control group.

As can be seen, increases of this magnitude take households that are in the vicinity of the poverty line and lift them well above it. In addition, with an average of 5 individuals per-family, it takes only a short time before the direct program benefits swamp the direct monetary value of the asset transfer.

In summary, these patterns accord with what one would expect of an asset transfer program like LRAD, where participants inevitably must confront the vagaries entailed in a farming enterprise. More generally, as King and Behrman (2009) have argued, a variety of factors might account for why there could be a lag, or more gradual realization of the benefits of anti-poverty programs. This is especially the case for a program like LRAD, which is aimed not only at affecting rural livelihoods directly, by changing the incentives governing work effort, but also indirectly through changing the underlying wealth distribution in rural areas.

5 Conclusion

The largely theoretical literature on asset inequality has long suggested that asset transfers – such as land redistribution – can be an effective anti-poverty tool. Not only should they be constrained by fewer questions about work disincentives than are cash transfers, they should be able to generate very high rates of return if they succeed in unlocking the productive potential of the poor by improving their market access and perhaps getting them over a critical minimum threshold.

Despite this promise, well-identified empirical evidence on efficacy of land redistribution has been scarce, in no small part because the most interesting reforms arise endogenously, either as a response to, or as a cause of, conflict. This fact places limits on the uses of experimental approaches to identify impacts. For this paper, we have been able to explore a relatively low conflict situation (South Africa) and exploit the fact that the implementation of its market-assisted land reform program generated a natural experiment, allowing identification of the impact of land transfers on the economic well-being of poor and near poor households.

Standard binary treatment effect estimates indicate that the land transfers boosted household living standards by 25%. More interestingly, our continuous treatment estimates, which exploits variations in the period of ownership of the redistributed land, show that living standards initially dip with the land transfers, but then after three years rise to levels that imply a 50% increase in living standards of the treated households who entered the program with poverty line standards of living. Both the temporal pattern of this impact, and its magnitude, are consistent with the theoretical literature on asset transfers and their potential to crowd-in investment, learning and income increases beyond what would be expected from the direct transfer alone.

Compared to cash transfers, where it is possible to "just give the poor, any poor, the money" (Hanlon, Barrientos and Hulme 2010), asset transfers are clearly more complicated and have a more limited scope as not everyone can be a successful small-scale farmer. It also remains to be determined if the increases in family well-being detected by the South African land redistribution program spill over into the kinds of investment in child human capital detected in conditional and unconditional cash transfer programs. Subject to these limitations, the impacts detected here would seem to motivate further experimentation with asset transfer programs.

Table 1: Descriptive Statistics

Variables	Treated	Controls	$0 < d < 2$	$2 < d < 3$	$d > 3$	All
Per-capita consumption	594.5	466.4	486.6	548.1	749.1	501.2
	(1098.1)	(692.0)	(906.2)	(713.1)	(1485.7)	(824.1)
Application date	1803.1	1675.7	1775.6	1717.6	1908.0	1710.3
	(761.1)	(1302.1)	(865.9)	(714.9)	(670.2)	(1181.0)
Transfer date	808.6	0	337.6	837.9	1278.3	219.9
	(456.6)	(0)	(305.0)	(55.93)	(239.1)	(431.4)
Family Labor	0.787	0.436	1.156	0.496	0.658	0.532
	(1.291)	(0.942)	(1.789)	(0.791)	(0.870)	(1.060)
Household head is male	0.754	0.667	0.725	0.719	0.816	0.690
	(0.431)	(0.472)	(0.448)	(0.451)	(0.389)	(0.462)
Education of household head (yrs)	6.447	5.843	5.088	7.193	7.217	6.007
	(4.880)	(4.496)	(4.058)	(5.084)	(5.198)	(4.610)
Mean farming experience (yrs)	1.594	1.464	1.361	1.292	2.108	1.500
	(3.706)	(3.784)	(3.754)	(2.213)	(4.581)	(3.763)
Household Size	6.060	6.138	6.675	5.852	5.599	6.117
	(3.532)	(3.732)	(4.013)	(3.192)	(3.192)	(3.678)
Relocated to Participate	0.208	0.0727	0.206	0.126	0.283	0.109
	(0.406)	(0.260)	(0.406)	(0.333)	(0.452)	(0.312)

Notes: Sample sizes of treatment group = 448; sample size of control group = 1202. Monthly per capita consumption expenditure is in 2005 Rands. The *Application Date* and *Transfer Date* are measured as the number of days elapsed between the relevant milestone and the date of commencement of fieldwork. The duration of treatment, d is measured in years. The final column shows means (standard deviations) for the combined sample of treated and control households. Tests of the restriction that the difference in means for each variable between the treatment and control groups equal zero: *pce* (p = 0.0053); *doseapp* (p = 0.0502); *onfarmemp* (p = 0.0000); *sexhead* (p = 0.0008); *hheadeduc* (p = 0.0147); *farmexper* (p = 0.5434); *residents* (p = 0.7425); *moved* (p = 0.0000).

Table 2: Propensity Score Regressions

Variable	(1)	(2)
Family Labour	.252	.342
	(.052)***	(.064)***
Application Date	.623	44.787
	(.082)***	(3.745)***
Household Head is Male	.385	.353
	(.131)***	(.146)**
Education of Household Head (yrs)	.052	-.867
	(.013)***	(.259)***
Mean Farming Experience (yrs)	.009	-1.819
	(.015)	(.587)***
Household Size		.001
		(.020)
Relocated to Participate		1.321
		(.205)***
Application Date Squared		-3.058
		(.253)***
Education of household head Squared		.006
		(.004)*
Mean farming experience Squared		-.002
		(.002)
Education × Experience		.146
		(.081)*
Application Date × Experience		.243
		(.078)***
Application Date × Education		.108
		(.034)***
Application Date × Experience × Education		-.019
		(.011)*
Const.	-6.272	-164.461
	(.638)***	(13.878)***

Notes: The regressions are based on the logit model. The dependent variable equals one if the household is in the LRAD treatment group and zero if it is in the LRAD control group.

Table 3: Propensity Score Balance

Strata	$\hat{p}(\mathbf{x})$ cutoff	N_0	N_1	p-value
1.00	0.02	227	11	0.2605
2.00	0.1	160	32	0.2504
3.00	0.2	318	139	0.9713
4.00	0.4	180	173	0.1200
5.00	0.6	24	23	0.0156
6.00	0.7	7	32	0.1618
7.00	0.8	8	38	0.2496

Notes: "Strata" refers to mutually exclusive intervals of the estimated propensity score distribution. These intervals are defined by the cut-off points given by $\hat{p}(\mathbf{x})$. The fifth column in the table reports the p-value of a two-sided test of the null hypothesis of a difference in the mean value of the estimate propensity score between treatment and control within each stratum.

Table 4: Covariate Balance

	Strata of $\hat{p}(\mathbf{x})$						
Variable	1	2	3	4	5	6	7
Family Labour	0.316	0.277	0.353	0.054	0.351	0.504	0.648
Application Date	0.101	0.600	0.012	0.630	0.121	0.776	0.679
Household Head is Male	0.060	0.078	0.450	0.532	0.045	0.026	0.518
Education of Household Head (yrs)	0.058	0.435	0.098	0.314	0.391	0.944	0.019
Mean Farming Experience (yrs)	0.126	0.734	0.381	0.478	0.583	0.427	0.660
Household Size	0.651	0.025	0.746	0.026	0.040	0.161	0.396
Relocated to Participate	0.417	0.426	0.561	0.917	0.684	0.082	0.020
Application Date Squared	0.094	0.569	0.011	0.656	0.120	0.806	0.690
Education of household head Squared	0.096	0.522	0.044	0.107	0.371	0.922	0.055
Mean farming experience Squared	0.405	0.505	0.403	0.197	0.847	0.565	0.620
Education	0.092	0.796	0.990	0.770	0.816	0.708	0.239
Application Date × Experience	0.165	0.700	0.417	0.473	0.642	0.427	0.658
Application Date × Education	0.165	0.550	0.189	0.361	0.410	0.962	0.016
Application Date × Experience × Education	0.107	0.861	0.954	0.759	0.867	0.725	0.254

Notes: The table shows that the covariates are balanced once we condition on the propensity score. The column headings refer to the 7 intervals of the propensity score distribution within which the estimated propensity score is balanced (see table 3. The entries in each table report the p-value for an equality of means test of each regressor by treatment status.

Table 5: Average Treatment Effect on Per Capita Consumption (Percentage Change)

Kernel	Kernel Function $K(s)$	Optimal Bandwidth	ATT	t-ratio Analytical	t-ratio Bstrap
Gaussian	$(2\pi)^{-1/2}\exp(-s^2/2)$	–	28.36	2.04	2.41
Gaussian	$(2\pi)^{-1/2}\exp(-s^2/2)$	0.26	22.77	1.72	1.84
Epanechnikov	$\frac{3}{4}(1-s^2)\times\mathbf{1}(\lvert s\rvert<1)$	0.59	22.34	1.69	1.88
Quartic	$\frac{15}{16}(1-s^2)^2\times\mathbf{1}(\lvert s\rvert<1)$	0.69	22.45	1.69	1.69
Rectangular	$\frac{1}{2}(\lvert s\rvert<\mathbf{1})$	0.23	25.16	1.85	1.96
Tricube	$\frac{70}{80}(1-s^3)^3\times\mathbf{1}(\lvert s\rvert<1)$	0.50	24.39	1.82	1.76

Notes: The first ATT estimate using the Gaussian Kernel uses a fixed global bandwidth of 0.05, whereas the remaining Kernel estimators use an optimal bandwidth calculated according to Silverman's (1986) plug-in formula. Bootstrapped standard errors are over 100 replications. Matching occurs over the common support of the log-odds ratio. In addition, a further trimming rule of 2% is used.

Table 6: Maximum Likelihood Estimates of the Parameters of the Generalized Propensity Score

Variable	**Coefficient**	**(Std. Err.)**
Household head is male	0.055*	(0.023)
Education of household head (yrs)	-0.014	(0.012)
Mean farming experience (yrs)	0.053**	(0.019)
Household Size	-0.006†	(0.003)
Relocated to Participate	0.032	(0.025)
Application Date	0.000**	(0.000)
Application Date × Farming Experience	0.000*	(0.000)
Education of the Household Head Squared	0.001*	(0.001)
Farming Experience Squared	0.000	(0.000)
Education × Experience	-0.006*	(0.003)
Application Date × Education	0.000	(0.000)
Application Date Squared	0.000**	(0.000)
Application Date × Experience × Education	0.000*	(0.000)
Intercept	-0.261*	(0.127)
MLE of σ	0.204**	(0.007)
N	438	
Log-likelihood	73.929	
$\chi^2_{(13)}$	92.533	

Significance levels : † : 10% ∗ : 5% ∗∗ : 1%

Table 7: OLS Estimates of the Parameters of the Conditional Expectation of Monthly Per Capita Consumption given D_i and $\hat{R}_i$

Variable	**Coefficient**	**(Std. Err.)**
Normalized Duration of Treatment	-1.334	(1.034)
Normalized Duration of Treatment Squared	1.476	(1.112)
Estimated GPS	-1.353**	(0.484)
Estimated GPS Squared	0.500**	(0.184)
Normalized Duration of Treatment × Estimated GPS	1.345**	(0.361)
Intercept	5.98**	(0.221)
N	438	
F	9.38	

Significance levels : † : 10% ∗ : 5% ∗∗ : 1%

Table 8: Bayes Factor Tests of Equality of the Conditional Means of the Covariates given the Generalized Propensity Score

	Normalized Treatment Intervals		
Variable	[0,0.37]	[0.37,0.47]	[0.47,1]
Household head is male	3.0899	3.9107	5.8918
Education of household head (yrs)	0.8949	2.2702	5.2837
Mean farming experience (yrs)	5.5048	4.1895	5.7830
Household Size	6.4096	5.3034	4.9582
Relocated to Participate	5.9554	0.4677	2.8339
Application Date	6.1249	3.8364	1.9113
Application Date $\times$ Farming Experience	5.9762	3.7087	5.7806
Education of the Household Head Squared	0.3364	1.4644	4.9125
Farming Experience Squared	6.2360	3.4174	5.9819
Education $\times$ Experience	2.5461	4.8163	4.7199
Application Date $\times$ Education	0.8618	4.0728	5.8074
Application Date Squared	6.1396	4.0811	2.7045
Application Date $\times$ Experience $\times$ Education	3.1035	3.7053	3.3773

Notes: Table entries are the Bayes factor test statistics of the hypothesis that the mean in one of the three treatment groups is not statistically different to the mean in the other two groups combined. The specific algorithm we use is based on Hirano and Imbens (2004). See Bia and Mattei(2008) for details on the implementation. The main idea of the algorithm is estimate the GPS at a representative point within each treatment interval (here chosen to be the mean) and then discretize the distribution of the GPS evaluated at this representative point. In our model, we chose 6 mutually exclusive blocks. At the mean of each treatment interval, a Bayes factor in excess of 1 therefore counts as decisive evidence that a covariate is adequately balanced between interval k, block j and interval not k, block j. Overall, the model is well balanced as the lowest Bayes factor of 0.3364 falls within an acceptable range of the Jeffery's order of magnitude criterion. Decisive rejection of the null that the data are balanced requires a minimum Bayes factor of less than 0.01.

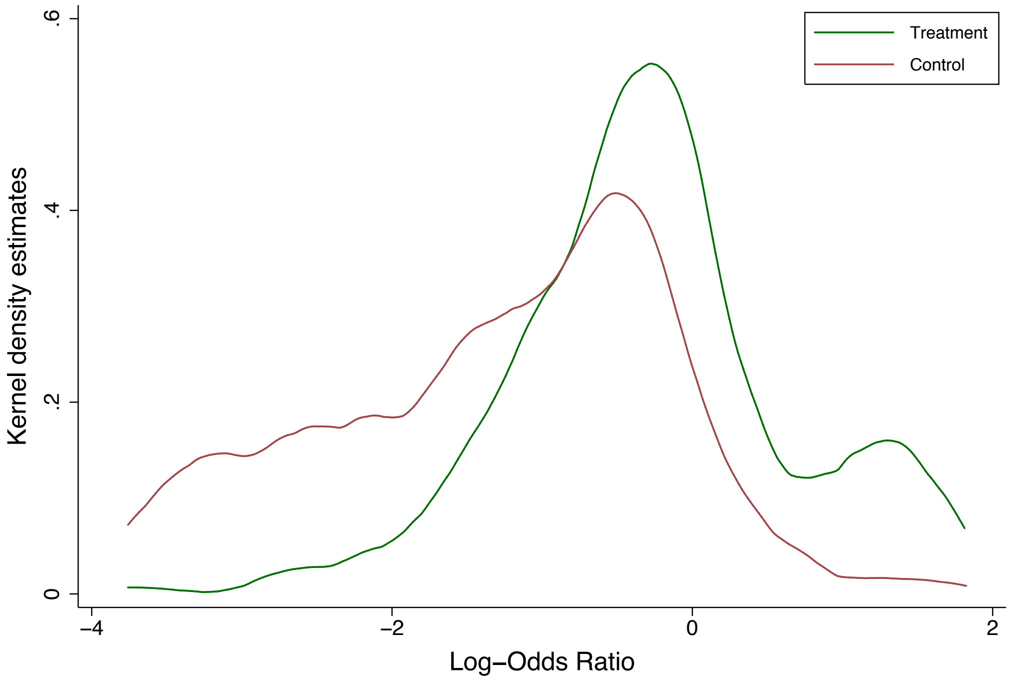

Figure 1: Distribution of Estimated Log Odds Ratios

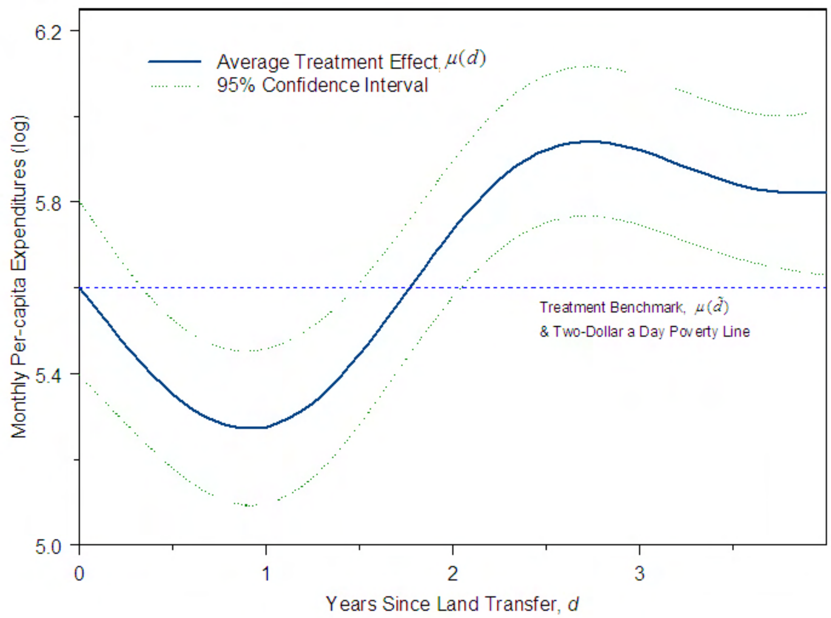

Figure 2: Duration Response Function

References

Angrist, Joshua, "Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants," *Econometrica*, 1998, *66* (2), 249–288.

Banerjee, A., P. Gertler, and M. Ghatak, "Empowerment and Efficiency: Tenancy Reform in West Bengal," *Journal of Political Economy*, 2002, *110* (2), 239–280.

Bardhan, P., S. Bowles, and H. Gintis, "Wealth Inequality, Wealth Constraints and Economic Performance," in A.B. Atkinson and F. Bourguignon, eds., *Handbook of Income Distribution*, Elsevier-Science, North-Holland, 2000.

Bardhan, Pranab and Dilip Mookherjee, "Productivity Effects of Land Reform: A Study of Disaggregated Farm Data in West Bengal, India," 2008.

Behrman, Jere R., Yingmei Cheng, and Petra E. Todd, "Evaluating Preschool Programs when Length of Exposure to the Program Varies: A Nonparametric Approach," *Review of Economics and Statistics*, 2004, *86* (1), 108–132.

Berry, R. A. and W. R. Cline, *Agrarian Structure and Productivity in Developing Countries*, Baltimore: Johns Hopkins University Press, 1979.

Binswanger, H. P., K. Deininger, and G. Feder, "Power, Distortions, Revolt and Reform in Agricultural and Land Relations," in "Handbook of Development Economics," Vol. 3B, North-Holland: Amsterdam: Elsevier-Science, 1995.

Carter, Michael, "Identification of the Inverse Relationship between Farm Size and Productivity," *Oxford Economic Papers*, 1984, *March.*

Carter, Michael R and Christopher Barrett, "The Economics of Poverty Traps and Persistent Poverty: An Asset-based Approach," *Journal of Development Studies*, 2006.

Carter, Michael R. and Frederick J. Zimmerman, "The dynamic cost and persistence of asset inequality in an agrarian economy," *Journal of Development Economics*, December 2000, *63* (2), 265–302.

Chamberlain, Gary and Edward E. Leamer, "Matrix Weighted Averages and Posterior Bounds," *Journal of the Royal Statistical Society*, 1976, *Series B* (38), 73–84.

Dasgupta, Partha and Debraj Ray, "Inequality as a Determinant of Malnutrition and Unemployment: Theory," *Economic Journal*, December 1986, *96*, 1011–1034.

de Mel, Suresh, David McKenzie, and Christopher Woodruff, "Returns to Capital in Microenterprises: Evidence from a Field Experiment*," *Quarterly Journal of Economics*, 2008, *123* (4), 1329–1372.

Dehejia, R.H. and S. Wahba, "Propensity Score Matching Methods for Nonexperimental Causal Studies," *Review of Economics and Statistics*, 2002, *84*, 151–161.

Deininger, Klaus and Pedro Olinto, "Asset distribution, inequality, and growth," Policy Research Working Paper Series 2375, The World Bank June 2000.

Dorner, Peter, *Land Reform*, New York: Penguin Press, 1970.

Eichler, M. and M Lechner, "An evaluation of public employment programmes in the East German state of Sachsen-Anhalt," *Labour Economics*, 2002, *9*, 143–186.

Eswaran, Mukesh and Ashok Kotwal, "Access to Capital and Agrarian Production Organisation," *Economic Journal*, June 1986, *96* (382), 482–98.

Galor, Oded and Joseph Zeira, "Income Distribution and Macroeconomics," *Review of Economic Studies*, January 1993, *60* (1), 35–52.

____ , Omer Moav, and Dietrich Vollrath, "Inequality in Land Ownership, the Emergence of Human Capital Promoting Institutions, and the Great Divergence," *Review of Economic Studies*, 2009, *76* (1), 143 – 179.

Hanlon, Joseph, Armando Barrientos, and David Hulme, *Just give money to the poor: the development revolution from the global south*, Sterling, VA.: Kumarian Press, 2010.

Heckman, J., H. Ichimura, and P. Todd, "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program," *Review of Economic Studies*, 1997, *64*, 605–654.

____ , ____ , J. Smith, and P. Todd, "Sources of Selection Bias in Evaluating Social Programs: an interpretation of conventional measures and evidence on the effectiveness of matching as a program evaluation method," *Proceedings of the National Academy of Sciences*, 1996, *93* (23), 13416–13420.

____ , ____ , ____ , and ____ , "Characterizing Selection Bias Using Experimental Data," *Econometrica*, 1998, *66*, 1017–1098.

Heckman, James J. and Petra E. Todd, "A note on adapting propensity score matching and selection models to choice based samples," *Econometrics Journal*, 01 2009, *12* (s1), S230–S234.

____ and Richard Jr. Robb, "Alternative methods for evaluating the impact of interventions : An overview," *Journal of Econometrics*, 1985, *30* (1-2), 239–267.

Hirano, Keisuke and Guido Imbens, "The Propensity Score with Continuos Treatments," in Andrew Gelman and Xiao-Li Meng, eds., *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*, West Sussex, UK.: Wiley, 2004, pp. 73–84.

Imai, Kosuke and David A. van Dyk, "Causal Inference With General Treatment Regimes: Generalizing the Propensity Score," *Journal of the American Statistical Association*, January 2004, *99*, 854–866.

Imbens, Guido W., "The Role of the Propensity Score in Estimating Dose-Response Functions," *Biometrica*, 2000, *87* (3), 706–710.

Kanel, Don, "The Economic Case for Land Reform," *Land Economics*, 1968.

King, Elizabeth M. and Jere R. Behrman, "Timing and Duration of Exposure in Evaluations of Social Programs," *World Bank Research Observer*, 2009, *24* (1), 55–82.

Lipton, Michael, Frank Ellis, and Merle Lipton, "Introduction," in Michael Lipton, Mike de Klerk, and Merle Lipton, eds., *Land, Labour and Livelihhods in Rural South Africa*, University of Natal, Durban: Indicator Press, 1996, pp. v – xvii.

____, Robert Eastwood, and Andrew Newell, "Small Farms," in "Handbook of Agricultural Economics," Vol. 4, North-Holland: Amsterdam: Elsevier-Science, 2009.

Mookherjee, Dilip and Debraj Ray, "Contractual Structure and Wealth Accumulation," *American Economic Review*, September 2002, *92* (4), 818–849.

Ravallion, M., "Evaluating Anti-Poverty Programs," in R. E. Evenson and T. Paul Schultz, eds., *Handbook of Development Economics: Volume 4*, Amsterdam, North-Holland: Elsevier, 2008, pp. 3787–3846.

Ray, Debraj and Peter A Streufert, "Dynamic Equilibria with Unemployment Due to Undernourishment," *Economic Theory*, January 1993, *3* (1), 61–85.

Rosenbaum, Paul R. and Donald Rubin, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 1983, *70* (1), 41–55.

Rosenzweig, Mark and Hans Binswanger, "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments," *Economic Journal*, 1993, *103* (416), 56–78.

Shaban, R., "Testing Between Competing Models of Sharecropping," *Journal of Political Economy*, 1987, *95*, 893–920.

Silverman, B., *Density Estimation for Statistics and Data Analysis*, London: Chapman and Hall, 1986.

Smith, Jeffrey and Petra Todd, "Does Matching Overcome LaLonde's Critique of NX Estimators," *Journal of Econometrics*, 2005, *125* (12), 305–353.

Woolard, Ingrid and Murray Leibbrandt, "The Measurement of Poverty in South Africa: some technical issues," 2007.

World Bank, *Land Policies for Growth and Poverty Reduction*, Vol. World Bank Policy Research Report, New York: Oxford University Press, 2003.

Zyl, J. Van, J. Kirsten, and H.P. Binswanger, *Agricultural Land Reform in South Africa*, Cape Town: Oxford University Press, 1996.