Poverty and Land Redistribution: Evidence from a Natural Experiment

Malcolm Keswell and Michael Carter*

September 17, 2011

Abstract

The theoretical literature on asset inequality suggests that asset transfers – such as land redistribution – can be a particularly effective anti-poverty tool because they enhance the ability of beneficiaries to utilize their full set of endowments and opportunities and grow their way out of poverty. Despite this promise, well-identified empirical evidence on the efficacy of land redistribution has been scarce, in no small measure due to the fact that the more interesting redistributive reforms arise either as a response to, or as a cause of, conflict. In an effort to fill this gap, 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-filtered 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 50% after three years of exposure to the program.

JEL Keywords: Land Reform, Poverty, Impact Evaluation

JEL Codes: O10, O12, O13

---

*Keswell: Southern Africa Labour and Development Research Unit & School of Economics, University of Cape Town, Rondebosch 7700, Cape Town (*email*: malcolm.keswell@uct.ac.za); Carter: Department of Agricultural & Resource Economics, University of California-Davis, Davis, CA 95616, USA (*email*: mr-carter.ucdavis.edu). 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 Universities of Cape Town, Johannesberg, Wisconsin, and Stellenbosch, as well as seminar participants at the 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.

1 Introduction

A rich, 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] An increase in the wealth holdings of poor agents therefore has the potential to better align incentives thereby enabling lower wealth agents to unlock unrealized market potential, thus mitigating the type of informational asymmetries that give rise to inefficient outcomes in the first place. It is this direct connection to incentives (and the concomitant changes in contractual structure predicted by large scale wealth transfers) that make asset transfer programs an attractive policy option relative to other types of interventions (conditional cash transfer programs, for example) in the fight against rural poverty.[2]

Despite the attractiveness of this theoretical argument however, there has been relatively little empirical demonstration of the effectiveness of asset transfer programs, especially relative to the outpouring of empirical evidence on 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 reduction 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 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

[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.

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, raising a plethora of identification concerns (e.g., are "naturally occurring" small farms intrinsically more productive because they have better soils, better farmers, etc.). In contrast, 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 political potency and economic complexity of evaluating land reforms that entail an outright transfer of ownership from the wealthy to the poor. While randomized controlled trials (RCTs) have been used to study modest in-kind asset transfers,[7] no country to date has implemented land redistribution by randomly distributing land to some, but not other, potential beneficiaries.

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. The program does not mandate redistribution of land but rather operates on a willing buyer – willing seller basis. This "privacy-preserving" characteristic of the program makes it at once both interesting, but also more challenging to evaluate.[8] However, as we will show, the mechanisms which determine assignment to the program ultimately plays out a natural experiment: individuals self-select into the applicant pool, but are then subjected to an exhaustive screening process. This screening process leads to a homogenization of sorts among the applicant pool, where applicants with similarly high chances of succeeding as farmers are kept in the applicant pool, whilst applicants with little chance of success are dropped from the pool. This administratively filtered subset of the applicant pool end up receiving land transfers at random points in time. This combined process of beneficiary self-selection, administrative screening, and random timing of final assignment to the treatment group is the crux of our identification strategy: because treatment status is orthogonal to the characteristics of the applicant or of the land being purchased, we are able to construct a natural control group out of this subset of LRAD applicants, who, as

[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 and size such that aggregate output and economic performance improve 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) and Fafchamps, McKenzie, Quinn and Woodruff (2011) for examples of RCTs that transfer productive assets (e.g., sewing machines) worth a few hundred dollars to randomly selected beneficiaries.

[8] The LRAD program is privacy preserving because the reallocation works primarily through the price mechanism. However, as will become apparent in section 2, reliance on the price mechanism is helped considerably by state involvement in the process, such that the process is best viewed more as a case of the "helping hand" of the state, as opposed to the invisible hand of the market. We view the program as especially interesting because it involves a substantial transfer of *ownership* of land.

of the survey date, had been selected to receive land through the program, but who had not yet received the actual transfer for exogenous reasons. In addition, the random exit from the LRAD application pipeline 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 dynamic average treatment effect for the treated. We find that LRAD land transfers lead to at least a 20% increase in per-capita monthly consumption for program participants relative to the untreated control group. Exploiting variation in length of program exposure using a continuous treatment estimator, we find long-run treatment effects that lift the typical household from a \$2 per-day poverty line standard of living to almost \$3 per-day. Moreover, the total returns per-family are sufficiently large that they offset the direct costs of the asset transfer grant within three to four 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 variation in the duration (or exposure) to the program, estimating long-run impacts with 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.

2 The LRAD Program

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. Against this backdrop, 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 peasant smallholder class that has been the 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.

2.1 Key Features

Styled on the willing buyer-willing seller market-assisted land reform model, LRAD was intended to provide land to black South Africans (especially women) with an interest in farming. Applicants to the program have to live on or near the land they wish to acquire through the program. Individuals who hold public office, civil servants, or relatives of such individuals are ineligible.[9]

The program works on the basis of a grant that is awarded to beneficiaries on a sliding scale. The minimum grant of 20,000 South African rands requires a matching contribution (in cash or in-kind) of 5,000 rand. The maximum grant of 100,000 rand requires a matching contribution amount of 400,000 rand. In practice, grants are pooled into a fund that is administered on behalf of a group of beneficiaries. These funds are then used to purchase land, which becomes the property of 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. The basic idea is similar in spirit to Lavy (2002) who derives plausible counterfactuals to a non-randomly assigned treatment by backing out a natural experiment that results from the *observable* assignment mechanism.[10]

However, as will become apparent momentarily, our study differs methodologically in several fundamental ways from the approach taken in Lavy (2002). The first major difference, is that we study the dynamics of the duration of exposure to our program. Section 4 deals with this aspect of our study. The second major difference has to do with the fact that treatment

[9] As will become clear from the discussion in section 2.2, these criteria are both verifiable and enforced.

[10] The study by Lavy (2002) concerns the evaluation of a teacher incentives program introduced in schools in Israel. The schools which received the intervention had to be the only school of its kind in a given community (where school type could differ along some 20 plus dimensions). Lavy (2002) argues that the rules determining program assignment led to a natural experiment where schools quite close to the cut-off (i.e., schools falling into communities where there were two of a given type) were not selected for treatment. Identification is achieved by restricting the sample of untreated (comparison) schools to just these cases.

assignment in our study, relies both on beneficiary self-selection (i.e., entry in the applicant pool), and administrative screening. Our identification strategy will rely on both types of information to construct a comparison group who have the same characteristics as the treated, but fall just short of assignment to the treatment group. In this sense, our identification strategy is quite similar to that of Lavy (2002), but unlike in that study, here the assignment mechanism is not directly observed, but rather has to be deduced by tracing out the genesis from baseline entry into the applicant pool, till what is observed at end-line in our sampling frame.

We now turn to how we made this deduction. As we will elaborate on below, an understanding of this genesis is essential to our identification strategy as it reveals that LRAD operates as a natural experiment: from the universe of all applicants to the program, only those that survive several levels of administrative screening have any chance of selection for treatment. This subset of applicants forms the basis of our sampling approach. Details of the actual screening process discussed below reveals that a natural experiment plays out in how final assignment takes place from among this subset.

2.2 The Application and Approval Process

Primarily targeted at small-scale black farmers and farmworkers, would be beneficiaries of LRAD grants must pass through the following five stages that entail self-selection, administrative selection and a measure of luck:

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 (the "planner") does an assessment of the site on which the applicants live as well as the land they have applied to purchase (which need not coincide with the current place of occupancy of the applicants). The purpose of this initial assessment is both to verify the information of the applicant against the eligibility criteria and to form an initial impression of the resources available to the applicants as well as the constraints under which they operate. If the application appears to be a serious prospect from this initial impression, then it is "registered" as as candidate land redistribution project.

2. Approval of Planning Grant: Stage 2 begins when the planner requests the district land affairs department to release a nominal sum of money to begin developing a proposal on behalf of the applicant. The funds pay for various specialised studies that the planner uses to substantiate the beneficiaries' business plan and to negotiate a purchase price for the land. Examples of such studies are property valuations, soil assessments, and land surveys.

3. Preparation of Project Identification Report: Once these commissioned studies materialize, the planner works with the applicants to create the final business plan and proposal that is ultimately submitted to the state as background motivation for the LRAD transfer.

This proposal preparation stage is an important process that is handled through a series of workshops between relevant stake holders 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 the PIR 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 stake holders 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 the "designation memo" is prepared and used by the PGAC to make their final approval decision. A key hurdle that applications reaching stage 4 must overcome is that there must be complementarity between the proposed plan and the availability of basic services (roads, irrigation, electrification). If approved by the PGAC, the designation memo must ultimately be counter-signed by the directors general and ministers of land affairs and agriculture. Our study will focus only on the subset of LRAD applicants whose projects received this stage 4 approval.

5. Final Transfer by Provincial Grants Approval Committee: Once they receive stage 4 approval, LRAD applicants must wait before actually receiving the transfer of land. 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 their land, if they receive it at all.

Two key reasons were invariably cited by program staff for stage 5 delays. First, there could be a competing claim for the land offered up by the seller. For example, a fairly frequent occurrence is that land 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.[11] Second, an initially willing seller might withdraw from the

[11]The five-volume report of the *Surplus People's Project*, subsequently summarized in Platzky and Walker (1985), documented accounts of the devastating social consequences of the policy of "forced removals" under Apartheid. It is estimated that in the period 1960-1983, about 3.5 million people experienced evictions under this policy. In their book summarising the findings of the *Second Carnegie inquiry into Poverty and Development in Southern Africa*,Wilson and Ramphele (1989) provide a short summary of the main thrusts of this policy. While a major component of the policy concerned the enforcement of the "Group-Areas" Act, which sought to racially homogenize urban neighborhoods, an equally devastating component was the assault on blacks living in rural areas. The so-called "Bantustan Consolodation" evictions, together with the pre-Apartheid

agreement at the last minute as a strategy to renegotiate the selling price.[12] In probing specific examples where such events had taken place, it became clear that although not infrequent, these reasons for a delay in stage 5 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 that characterizes the key assignment mechanism which determines treatment status: they have self-selected into the program and their business plans have been similarly scrutinized, modified as needed and approved by the LRAD technical staff. The random exclusion of some 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) to estimate the 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.

2.3 Sample Construction and Characteristics

As described above, LRAD is implemented at the farm or project level, which typically brings together several beneficiary families. Our sampling followed a multi-level design in which multi-family projects were first randomly selected from the universe of all projects (approved or not) in the study regions, with probability of selection proportional to the number of beneficiary or applicant households who were part of the project. This random sample of LRAD projects was then subjected to a level 2 screening or filtering. As explained above, the existence of a signed "designation memo" indicates that a project and its applicants successfully navigated the first four stages of the approval process. We therefore checked for the existence of this document for sampled projects, screening out those for which the stage 4 designation memo did not exist.[13] Projects screened out at this level were replaced by newly randomly selected

"Betterment" programs led to massive movement of blacks living on or nearby white owned farmland, in order to consolidate white landholdings in rural areas, under the pretext of improving the living conditions of the rural black population. A major focus (perhaps *the* major focus) of the Mandela-era land reforms concerned the restoration of rights to this group of people and their descendants. This program, known as the "restitution" program, was terminated in 1999. 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.

[12] 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.

[13] While seemingly straightforward, it often took interviews with LRAD program staff and review of project PIRs to determine where or not the designation memo existed and had been approved.

projects and the process was repeated until the desired sample of stage 4-approved projects was obtained.

From this set of filtered projects, 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 primary outcome variable is monthly per capita consumption expenditures.[14] We explicitly choose not to use a binary indicator of poverty status since this is a less efficient way of defining poverty outcomes when working with quasi-experimental data designs (Ravallion 2008).[15]

Table 1 shows that mean per-capita consumption in treatment households is 128 rands (or 28%) higher than the control group mean consumption. This difference is significant at the 1% level. In anticipation of later continuous treatment analysis, Table 1 also disaggregates treated households into treatment 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. Beneficiary households that had enjoyed their land transfer for less than 2 years have mean per-capita consumption levels 4% higher than that of the untreated control groups. Households observed $2-3$ years after the transfer have 18% higher consumption than control households, while households with more than 3 years experience with the transferred lands have expenditures 61% higher than control households. As will be discussed more in Section 4, this dynamic pattern of impact is logical in the context of an asset transfer program. It also suggests that we should expect the long-run impact of LRAD to be larger than the impact suggested by a standard binary treatment estimator; i.e., the binary treatment effect is a simple data weighted average of the different dynamic treatment effects.

If we rely solely on the efficacy of the natural experiment created by LRAD implementation, then these single difference consumption changes in Table 1 could be interpreted as casual program impacts. However, it is possible that imperfections in the natural experiment may leave some differences between treatment and control groups. One possible such confound is that untreated control households (or households with low treatment durations) applied later to the program. If this lateness in applying is correlated with other productivity-relevant characteristics, then the single difference impact estimators would lose their validity.

To explore this possibility, Table 1 displays information on application date, defined as the number of days elapsed from the date of application to the LRAD program, to the date of the

[14]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.

[15]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).

household survey. As can be seen in Table 1, treated households on average applied 128 days (or about 4 months or 7%) earlier than did control households. Similarly, households with more than 3 years of exposure to land transfers applied 4 to 6 months earlier than beneficiary households in the low and medium treatment terciles. While these 4 month differences are modest (and would seem to verify the interpretation of treated households as Stage 5 Lottery winners), there could be some concern that earlier applicants are somehow different than later applicants. As can be seen in Table 1, treated applicants, and those beneficiaries with a longer treatment duration, are in fact better educated. In addition, we might worry that these earlier applicants may be more entrepreneurial or exhibit other unobserved characteristics that predict economic success. If this latter concern is in fact correct, then application date should in fact serve as a proxy for these unobserved characteristics.

A second possible confound to our identification strategy concerns non-random factors that might explain the time an applicant spends in the pipeline (that is the time between application date and grant of the land transfer). It is, however, unclear which way this confound might work. On the one hand, we might expect more able applicants to progress more rapidly, implying that the single difference estimates in Table 1 are upwardly biased. Unfortunately, we lack systematic information on the time it took an applicant to progress from stage 1 to stage 4. But 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.

On the other hand, it is also possible that time in the pipeline is longest for projects and applicants that are likely to generate the largest impacts. Under the willing seller/willing buyer model of LRAD, it is possible that potential sellers of higher quality, higher impact parcels are likely to negotiate harder and drag out time in the pipeline for applicants proposing to purchase their land with LRAD grants. While we lack data to corroborate whether or not such delays actually took place, if they did then our control and low treatment duration groups would contain disproportionate numbers of higher return projects. In this case, the estimators reported in this paper would underestimate average treatment effects.

In the remainder of this paper we test for the robustness of the single difference estimators in Table 1 by employing binary and continuous treatment methods that allow us to match on productivity-relevant variables that could differ between treatment groups. In addition to the timing variables already discussed, 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 separating 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 a 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 2.3, 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. In section 4 we apply a generalized version of this estimator to shed light on the dynamic pattern of impact.

3.1 Modeling the Probability of Treatment

A major drawback of models based on the selection-on-observables 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.[16] 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).[17] 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.[18] Matching on the propensity score alone therefore does not necessarily solve the identification problem.[19]

However, by combining the insights gleaned from the administrative screening process outlined in section 2.2, 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 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

[16] 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).

[17] 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.

[18] 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.

[19] 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.

for entry into the pipeline (i.e., a measure of application date), as well as variables that are likely to covary with application date.[20] 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

[20] A possible concern with this approach is that observed variation in the date of application might reflect the quality of services rendered by LRAD staff members. While we do not observe such a measure, we can proxy for it using geo-spatial data on the proximity of our surveyed households to the nearest district office of the department of land affairs, under the assumption that households located closer to such offices are bound to get better services, because of the likelihood of greater contact, initiated on both sides. Including this type of distance measure has no discernible effect on the effect of *Applicant Date* on the response probability. Indeed the distance measure is not significant in the regression, and this accords with the impressions gleaned from our qualitative work on the administrative side of the program: although the district offices are often located in small towns and not deep rural areas, the staffers of the LRAD program operate more like"traveling salesmen" where their vehicles function as mobile offices. While we cannot definitively rule out the possibility that *Applicant Date* is exogenous because of the "quality of services" argument, these combined insights give us a measure of confidence that such a confound is unlikely.

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. 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.[21]

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 are well balanced across the 7 bins.

3.3 Empirical Estimates of Average Treatment Effects on Treated

We start with the standard GLM framework used to model 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 $f(0) = p^0(1-p)^1 = 1-p$.

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

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 not only because $e^{\mathbf{x}'\boldsymbol{\beta}}/(1+e^{\mathbf{x}'\boldsymbol{\beta}})$ is the canonical link function of the Bernoulli density, but also because it possesses certain useful properties relevant to choice-based samples such as ours (see footnote 23).

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

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

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:

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

where K is a kernel function, h_n is a bandwidth parameter and $\mathbf{x}_i'\boldsymbol{\beta}$ is the log of odds ratio.[22] 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.[23]

[22] 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 $\ln(p/(1-p)) = \mathbf{x}_i'\boldsymbol{\beta}$.

[23] The true sampling weights are unknown in our survey design for two reasons. Firstly, both the control group and treatment group sampling frames contained a small but non-trivial number of 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 potentially canceling out the correction. Secondly, a small fraction (about 10%) of the sampled treatment households appear to have suffered attrition from the programme. These households were replaced

In our empirical estimates, we check that our average treatment effect δ is robust to bandwidth specification and kernel choice.[24] Table 5 shows our estimates of the average treatment effects that are generated when matching non-parametrically against the log-odds ratios that 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 expenditures. Similar to the unconditional single difference estimates calculated based on Table 1, these estimated impacts are sizable. By way of comparison, Behrman, Sengupta and Todd (2005) estimate that 7 years of monthly cash transfers in Mexico's Progresa program[25] would increase *total future* income by only 8%. Of course in the short-run (while receiving transfers), cash transfer beneficiaries experience substantial total income increases, as much as 20% in the case of Progresa program, an amount still well below the estimated binary treatment effect of the LRAD asset transfer program. We turn now to explore the dynamic pattern of LRAD impacts in order to derive an estimate of the long-run treatment effect.

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 simultaneously co-investing in the newly acquired land,

by resampling without replacement, so although these replacement households will have been drawn from the same strata as the replaced households, their probability of selection will have been different to the replaced households. Moreover, some of these replacement households could have been drawn from that subset where information on the number of applicants at the farm level was missing, thus potentially exacerbating the distortions created by the imputations mentioned above. Matching on the (log) odds ratio, as suggested by Heckman and Todd (2009), therefore has the strong benefit of holding this aspect of the design constant when we compute average treatment effects.

[24] 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.

[25] Excluding administrative costs, the direct cost of these cash transfers would be approximately $US4600, an amount similar to the direct costs of an LRAD grant.

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, as suggested by the Table 1 descriptive statistics.

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, especially if LRAD transfers suffice to lift households over the sort of critical threshold asset level that figures prominently in the poverty traps literature.[26] It is these second round multiplier effects that are likely to distinguish asset transfer programs from other anti-poverty policy instruments.

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 conditional expectation function (CEF) 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

[26]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.

(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 does 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.

The evaluation problem of course results from the fact that each household realizes exactly one outcome, that 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 possible counterfactual duration. 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:

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

where $\tilde{d}$ serves as the benchmark duration.[27] In our empirical estimates, we set $\tilde{d}$ to be the lowest length of exposure observed in the data.

As in the binary approach, valid identification depends on an independence assumption regarding treatment assignment. Following Imbens (2000), we define weak unconfoundedness:[28]

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

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 $f_D(d|r(d, x), Y(d)) = f_D(d|r(d, X))$.[29]

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

[28]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)

[29]For the proof, as well as further details about assumptions regarding the probability space used in

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.[30]

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.[31] 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

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

Equation 7 is then estimated by OLS.[32] Once we have the estimated α parameters, we can

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.

[30] 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 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.

[31] 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.

[32] 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

then recover the average duration response function $E[y(d)]$. Recall that $E[y(d)]$ 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 function.[33] This gives a treatment effect estimator of the form

$$\begin{aligned} \hat{\mu}(d) = \widehat{E[Y(d)]} &= \frac{1}{n_1}\sum_{i=1}^{n_1}(\hat{\alpha}_0 + \hat{\alpha}_1 \cdot d + \hat{\alpha}_2 \cdot d^2 + \hat{\alpha}_3 \cdot \hat{r}(d, \mathbf{x}_i) \\ &\quad + \hat{\alpha}_4 \cdot \hat{r}(d, \mathbf{x}_i)^2 + \hat{\alpha}_5 \cdot d \cdot \hat{r}(d, \mathbf{x}_i)) \end{aligned} \tag{8}$$

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 used to predict treatment 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 for the normalized treatment 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

regression and conclude that not much additional explanatory power is added by including higher than second-order polynomials in D and $\hat{R}$.

[33] 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 $\beta(D, R)$.

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 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 Long-run Treatment Effects

The solid line in Figure 2 graphs the estimated GPS-adjusted average treatment effect as a function of the duration of treatment. As discussed above, the model was estimated using the natural logarithm of per-capita monthly expenditures as the dependent variable.[34] The corresponding 95% confidence interval (calculated using bootstrap estimates of the standard error) is shown by dotted lines. The dashed horizontal line is the benchmark, null treatment estimate, $\hat{\mu}(\tilde{d})$. By coincidence, this line also marks the level of the two-dollar a day international poverty line converted to South African Rand using the PPP exchange rate for 2006. This two-dollar a day poverty line (which is equivalent to about 250 Rand per-person, per-month at the PPP exchange rate) is substantially lower than the national poverty line of 430 Rand per-person, per-month which has been suggested by the South African Treasury.[35]

As is apparent from Figure 2, the continuous treatment impact estimator is quite different 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. Beyond the second year, the estimated average treatment effect rises further to levels that are approximately 50% higher than the international poverty line and the null treatment benchmark. 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, the estimated increase in per-capita monthly expenditures implies a total annual household ex-

[34]This transformation imposes the assumption that any treatment duration has the same percentage impact on per-capita expenditures for all households. We also estimated the model in levels, implying that each treatment duration has the same absolute effect on per-capita expenditure. These results were qualitatively the same as the log estimates shown here, except that absolute average effects were somewhat larger than those implied by the log changes.

[35]See the comprehensive benchmarking exercises reported in Woolard and Leibbrandt (2007) for further background information on the measurement and reporting of poverty in South Africa. The R430 per-capita line recommended by Statistics South Africa is one of the lines interrogated by Woolard and Leibbrandt (2007).

penditure increase of some 7200 rand per-year. With a typical grant of 20,000 rand, it takes only a few years for these total returns to swamp the direct monetary value of the grant. Consistent with the theory of costly asset inequality, measured impacts of this magnitude imply substantial co-investment and improved returns to other beneficiary resources.[36]

Comparison with estimated income increases predicted to result from cash transfer programs is again instructive. As discussed in Section 3.3, Behrman, Sengupta and Todd calculate that cash transfers raise earnings by 8% in the long-run, implying an approximately 2% increase in per-capita consumption. Similarly, the Aguero, Carter and Woolard (2011) study of South Africa's Child Support Grant calculates that the early childhood nutritional impacts of that cash grant would increase adult earnings by some 3% in the long-run. Where it is possible to make these comparisons, the long-run impacts of these cash transfer programs, though not insignificant, are one to two orders of magnitude smaller than the estimated long-run effects of LRAD transfers. It is of course possible that monthly cash grants could also crowd-in additional investment and intermediate income growth. The only evidence we know on this point is the work by Gertler, Martinez and Rubio-Codina (2006) who find that Progresa cash transfers had long-run investment impacts on the order of about 18%.

In summary, the estimated duration response function 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. Asset transfer programs are confronted by fewer questions about work disincentives than are cash transfer programs. These programs also have the potential 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 the sort of critical minimum threshold hypothesized by the theoretical poverty traps literature.

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

[36]Consistent with our finding, Fafchamps et al. (2011) also find large marginal returns to incremental asset transfers in Ghana. By contrast the second arm of their study design dealing with *cash* as opposed to asset transfers reveals much smaller average treatment effects. See also de Mel et al. (2008) for a similar experiment that was carried out in Madagascar.

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 (land redistribution in post-apartheid 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 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	Treatment	Control	Duration of Treatment 0.01 − 2 Years	2 − 3 Years	3 − 5 Years	Combined Treatment and Control
Per-capita consumption (2005 Rands)	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)
n	448	1202	176	139	163	1770

Notes: *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 differences in mean application date between the 0.1 − 2 group and the 2 − 3 and 3 − 5 groups respectively, are not statistically significant. The difference in mean application date between the 2 − 3 and 3 − 5 groups is statistically significant. The stars of column 1 indicate whether the differences in means for the relevant variables between the treatment and control groups are statistically significant: ** = significant at 1% level; * = significant at 5% level; † = significant at 10% level

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 $\times$ Estimated GPS	1.345**	(0.361)
Intercept	5.98**	(0.221)
N	438	
F	9.38	

Significance levels : $\dagger$: 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 × 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 × Experience	2.5461	4.8163	4.7199
Application Date × Education	0.8618	4.0728	5.8074
Application Date Squared	6.1396	4.0811	2.7045
Application Date × Experience × 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.

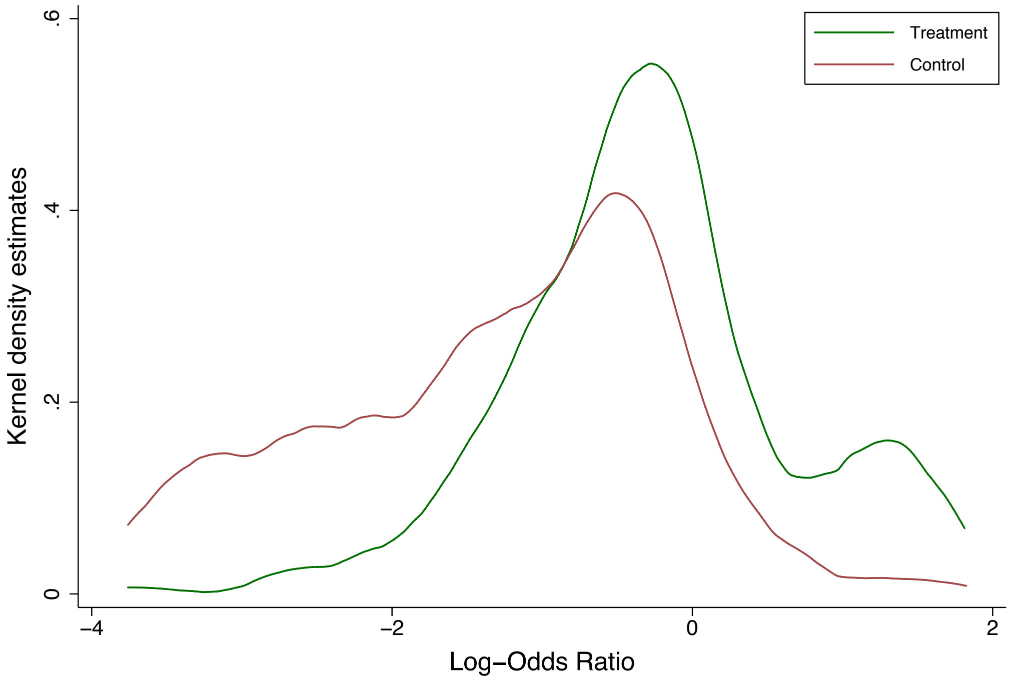

Figure 1: Distribution of Estimated Log Odds Ratios

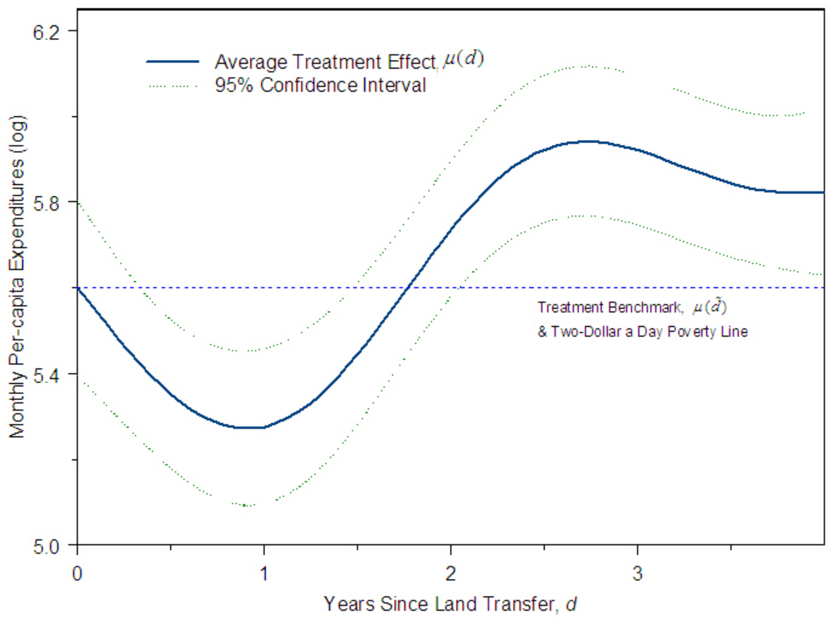

Figure 2: Duration Response Function

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. University of California, Berkeley.

Behrman, Jere R, Piyali Sengupta, and Petra Todd, "Progressing through PROGRESA: An Impact Assessment of a School Subsidy Experiment in Rural Mexico," *Economic Development and Cultural Change*, October 2005, *54* (1), 237–75.

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.

Fafchamps, Marcel, David J. McKenzie, Simon Quinn, and Christopher Woodruff, "When is capital enough to get female microenterprises growing? Evidence from a randomized experiment in Ghana," CEPR Discussion Papers 8466, C.E.P.R. Discussion Papers July 2011.

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.

Gertler, Paul, Sebastian Martinez, and Marta Rubio-Codina, "Investing cash transfers to raise long term living standards," Policy Research Working Paper Series 3994, The World Bank 2006.

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.

Lavy, Victor, "Evaluating the Effect of Teachers' Group Performance Incentives on Pupil Achievement," *Journal of Political Economy*, December 2002, *110* (6), 1286–1317.

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.

Platzky, L. and C. Walker, *The Surplus People: Forced Removals in South Africa*, Braamfontein, 1985.

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, Radwan Ali, "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.

Wilson, F. and M Ramphele, *Uprooting Poverty: The South African Challenge*, Cape Town: David Philip, 1989.

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

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.