Journal of Development Economics 158 (2022) 102896

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Regular article

Evaluating the distributive effects of a micro-credit intervention[☆]

Pushkar Maitra[a], Sandip Mitra[b], Dilip Mookherjee[c], Sujata Visaria[d,*]

[a] Department of Economics, Monash University, Clayton Campus, VIC 3800, Australia
[b] Sampling and Official Statistics Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata 700108, India
[c] Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, USA
[d] Department of Economics, Lee Shau Kee Business Building, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong

ARTICLE INFO

JEL classification:
D82
O16
C93
H21

Keywords:
Distributive impacts
Program evaluation
Agricultural finance

ABSTRACT

Most analyses of randomized controlled trials of development interventions estimate an average treatment effect on the outcome of interest. However, the aggregate impact on welfare also depends on distributional effects. We propose a simple method to evaluate efficiency–equity trade-offs in the utilitarian tradition of Atkinson (1970). This involves an estimation of the average treatment effect on a monotone concave function of the outcome variable, whose curvature captures the degree of inequality aversion in the welfare function. We argue this is preferable to the current practice of examining distributional impacts through sub-group analysis or quantile treatment effects. We illustrate the approach using data from a credit delivery experiment we implemented in West Bengal, India.

1. Introduction

Randomized controlled trials allow researchers to cleanly identify the average treatment effect (ATE) or the expected change in the beneficiary's outcome caused by a development intervention. However a more complete evaluation of the program's effects on welfare would not only account for average effects, but also for how equitably the benefits are distributed.

When the current literature examines heterogeneity in treatment effects, it usually follows one of two methods. Sub-group analysis involves classifying beneficiaries by certain fixed characteristics (e.g., gender) or baseline levels of well-being (e.g., baseline income or wealth), and estimating treatment effects for each sub-group separately. This provides a partial ordering of the effect of the intervention. For instance, a particular intervention may improve the outcome for a particular sub-group but worsen it for another. However, the method provides no guidance on the trade-offs between these different effects.

Quantile treatment effects are inherently different: rather than telling us how the intervention affects different sub-groups of beneficiaries, classified by fixed characteristics, they estimate the effect of the intervention on the different quantiles of the outcome distribution, such as, for instance, the 25th, 50th or 75th percentile (Bedoya et al., 2017). While informative, the method also does not explicitly build in any assumptions about the weights to assign to the average versus the distributional impact of the intervention. As a result it does not allow us to evaluate the aggregate welfare impact of intervention, or compare different interventions that may have different efficiency and equity impacts.

In this paper we propose an alternative approach that allows such welfare evaluations. It is rooted in the utilitarian tradition of public economics going back to Atkinson (1970). In this approach, the welfare impact of an intervention is given by the average treatment effect on welfare, where welfare is represented by a given increasing, concave

☆ Funding was provided by the Australian Agency for International Development (CF09/650), the International Growth Centre, United Kingdom (1-VRA-VINC-VXXXX-89120), United States Agency for International Development (AID-OAA-F-13-00007) and the Hong Kong Research Grants Council (GRF16503014). We are grateful to Shree Sanchari for collaborating on the project. Jingyan Gao, Rina Lookman Jio, Arpita Khanna, Clarence Lee, Daijing Lv, Foez Mojumder, Moumita Poddar and Nina Yeung provided excellent research assistance at different stages of the project. Elizabeth Kwok provided exceptional administrative support. We thank three anonymous referees, the editor Andrew Foster, Gaurav Datt, Lakshmi Iyer, Christina Jenq, Xun Lu, Farshid Vahid, Diego Vera-Cossio and seminar participants at the Institute for Emerging Market Studies at HKUST, UNSW, the Workshop on The Role of the Private Sector in Development at the University of Sydney, the IEMS-CAG Workshop on Financial Inclusion in Asia and the Italian Summer School in Development Economics held in Prato, Italy for helpful feedback and comments. Internal review board clearance was received from Monash University, Boston University and The Hong Kong University of Science and Technology. The authors are responsible for all errors.

* Corresponding author.
E-mail addresses: pushkar.maitra@monash.edu.au (P. Maitra), sandipisi@gmail.com (S. Mitra), dilipm@bu.edu (D. Mookherjee), svisaria@ust.hk (S. Visaria).

https://doi.org/10.1016/j.jdeveco.2022.102896
Received 9 June 2021; Received in revised form 15 May 2022; Accepted 16 May 2022
Available online 31 May 2022

function of the outcome variable. The slope and concavity of the welfare function incorporate the efficiency and distributive objectives respectively. The estimation is a straightforward exercise that does not require any additional assumptions beyond those used to estimate the average treatment effect on the outcome variable itself.

In particular, social welfare is represented as the sum of the welfare of the individuals in a population. The welfare of an individual can be written as $U_i = U(y_i) \equiv \frac{y_i^{1-\theta}}{1-\theta}; \theta > 0$, where the individual's wellbeing y_i could be proxied by income or consumption.[1] The parameter θ represents the degree of *inequality aversion*. Thus the welfare function $U(\cdot)$ reflects ethical judgments by the external observer, aid donor or social planner, in the "extended sympathy" approach to social choice theory.[2] Roberts (1980) provides axioms of cardinality and comparability of wellbeing that characterize this class of welfare functions.

When $\theta = 0$, the measure reduces to the sum of incomes, thus ignoring income distribution entirely. As θ increases, the social welfare function places greater weight on the wellbeing of worse-off individuals, and thus becomes more responsive to the distribution. As $\theta \longrightarrow +\infty$, social welfare approaches the Rawlsian maximin criterion $\min_i\{y_i\}$, thereby placing all weight on the welfare of the worst-off individual. Hence by varying the value of θ, the external evaluator can assess how distributive considerations affect the assessment.

A well established literature in public economics has similarly used Atkinson welfare functions to evaluate the distributional impacts of taxes, government transfers and price changes, especially when they have general equilibrium effects (see, for example Braverman et al., 1987; Ahmad and Stern, 1987; Newbery and Stern, 1987; Hughes, 1987; Newbery, 1995; Coady and Harris, 2004; Bhattacharya and Komarova, 2021). Our method shows how this approach can be applied to evaluate efficiency and equity impacts even for small-scale randomized interventions. The approach delivers a single summary measure of the impact of the intervention on welfare, which can be useful when policy makers evaluate interventions, or compare different interventions and choose which one to implement.

Section 2 shows how this methodology can be applied in a general setting with a randomized policy intervention, with differing underlying assumptions about the specific context. We allow a first stage where (a subset of) individuals in treated villages are selected as beneficiaries. This may be the result of a screening procedure or explicit criteria. It may depend on household-specific observable and unobservable characteristics, and therefore is not necessarily random. After this, the intervention is offered to a random subset of the selected group. If the research design includes both a treatment arm of villages where the intervention is conducted and a control arm where it is not introduced, then the evaluation is straightforward. We also show how the methodology can be applied in a more parsimonious research design where there is no control arm. In that case, the intervention can be evaluated relative to a hypothetical counterfactual, provided that only a random subset of the selected individuals receive the intervention, and there are no spillovers to untreated individuals.

We illustrate this methodology to evaluate the distributive impacts of three experimental micro-credit interventions that were implemented in West Bengal, India, during 2010–13. In two of these interventions, a commission agent was asked to select eligible borrowers for individual liability loans. In the Trader-Agent Intermediated Lending (TRAIL) arm, this agent was selected from among private traders operating in the village. In the Gram Panchayat Agent Intermediated Lending (GRAIL) arm, the agent was appointed by the local government. In a third intervention (Group Based Lending or GBL), borrower groups could self-form and all members could apply for joint liability loans.[3]

Our estimates show that the TRAIL intervention increased aggregate welfare for the entire range of θ values that we consider. In other words, any increase in inequality appears to be small enough that even at high levels of inequality-aversion, it does not outweigh the positive efficiency effects. In contrast, the welfare impacts of the GRAIL and GBL schemes are non-significant at all values of θ. Both the difference in the welfare effects of the TRAIL and GRAIL schemes, and of the TRAIL and GBL schemes are statistically significant.

In order to examine the underlying changes in outcomes that underlie this aggregate welfare effect, we decompose the average treatment effects on welfare into the impact on households in different land classes. This illustrates how our approach leads to different inferences than the quantile treatment effects (QTE) approach. Our decomposition exercise shows that the TRAIL intervention had the largest positive impact on the welfare of households in the intermediate range of landholding. On the other hand, the QTE approach shows that the TRAIL intervention increased the top deciles of the income distribution, thereby increasing inequality. We also know that the intervention had a positive average treatment effect on incomes (documented previously in Maitra et al., 2017, 2021). The QTE approach cannot assess how these different effects on efficiency and equity should be traded off to estimate an overall welfare effect. However as discussed above, our approach is able to evaluate this trade-off at different degrees of inequality aversion.

2. Methodology

The 72 villages in our study were randomly assigned to the three intervention arms. In each intervention arm, the credit intervention scheme was introduced in each village. As a result, the design did not include any control villages. Instead, within each village, potential beneficiaries were selected via a non-random procedure, and then a subgroup of these selected households were randomly chosen to receive a loan offer. Thus each village consists of three groups of households: (1) not selected to be potential beneficiaries; (2) selected but not randomly assigned to treatment; and (3) selected and also randomly assigned to treatment. Conditional on selection, the welfare effect of the intervention is the expected difference between the Atkinson transform of household farm income of selected households randomly assigned to treatment, and that of selected households not assigned to treatment.[4]

Since our goal is to estimate the aggregate or unconditional welfare effect of the scheme, we then scale the estimated conditional treatment effect by the proportion of households who participated in the loan scheme. This "treatment proportion" is the product of the objective average probability that a random household in the village was selected to be a potential beneficiary, written as $\Pr(s)$, and the fraction of the selected who were actually offered the treatment (p). Below we describe more formally how our approach yields an unbiased estimate of the welfare impact of the intervention.

For expositional purposes, we start with a context where the design allows for both treatment and control villages, and then explain how the standard methodology can be extended to the design we will be working with in subsequent sections. Consider a set of villages, which are randomized into a treatment group denoted T and a control (or

[1] For semantic convenience, in what follows we shall refer to the measure of wellbeing y_i as "income". Our specific empirical application also uses income, but the method can be applied to consumption as well.

[2] This is in contrast to the notion of a utility function, which determines the household's actual behavior, and represents the household's own subjective sense of wellbeing, which may incorporate considerations of status or relative income.

[3] In previous work (Maitra et al., 2017, 2021), we estimated the average treatment effects of these interventions on farm incomes. In this paper we propose and implement a method to evaluate the distributive impacts.

[4] This is the result of the additive decomposability of the Atkinson welfare measure.

counterfactual) group C. Each village consists of two types of individuals, $\sigma = s, n$, where s is the type that is selected for the intervention, and n is the type that is not selected. Selection can take place on criteria that are observable to the researcher, such as landholding or occupation, or on others that are unobservable, such as when individuals opt in, or when an intermediary selects beneficiaries. Since the intervention is randomly assigned across villages, $\Pr(s)$ describes the expected fraction of s type individuals in both T and C villages.

Next, let $e \in \{0,1\}$ denote whether a specific individual receives the treatment, and $p \equiv P(e = 1|s,T)$ denote the fraction of s types in a T village that receive the treatment. By construction, the scheme is available to none of the individuals in the C villages. It is also unavailable to type-n households in T villages. Hence we have

$$P(e = 1|n,T) = P(e = 1|\sigma,C) = 0 < p \equiv P(e = 1|s,T); \text{ for } \sigma = s,n \qquad (1)$$

Let the endline outcome (income or consumption) for an individual be represented by random variables $y(\sigma,e,T)$ and $y(\sigma,e,C)$ in T and C villages respectively. Then social welfare in a T village can be written as

$$W(T) = \underbrace{p\Pr(s)E[U(y(s,1,T))]}_{\text{selected and treated}} + \underbrace{(1-p)\Pr(s)E[U(y(s,0,T))]}_{\text{selected but untreated}} + \underbrace{[1-\Pr(s)]E[U(y(n,0,T))]}_{\text{not selected}} \qquad (2)$$

while social welfare in control (or counterfactual) villages can be written as

$$W(C) = \Pr(s)E[U(y(s,0,C))] + [1-\Pr(s)]E[U(y(n,0,C))]$$
$$\Longrightarrow W(C) = E[U(y(\sigma,0,C))] \qquad (3)$$

If the research design includes control villages, and data are collected from a random sample of households, then $W(C)$ can be directly estimated.[5]

In treatment villages, we assume income is measured for random samples within each of the three relevant groups: "selected and treated" $(s,1)$, "selected but untreated" $(s,0)$ and "not selected" $(n,0)$.[6] This allow us to estimate $W(T)$, and in turn to directly estimate the welfare impact of the intervention $W(T) - W(C)$.

The welfare impact can also be estimated if the research design is more parsimonious, in that there are no control villages. This requires that two conditions hold. First, we need $p < 1$, i.e., some selected subjects do not receive the intervention. This implies there is a non-null group of selected but untreated, $(s,0)$. The second condition is that there are no spillovers from treated to untreated subjects, or that the treatment does not affect untreated subjects of either type:

$$y(n,0,T) = y(n,0,C) \qquad (4)$$

$$y(s,0,T) = y(s,0,C) \qquad (5)$$

Observe that if C denotes the counterfactual that would have occurred if the T villages had not received the intervention, then Eqs. (4) and (5) imply that the welfare impact of the intervention equals

$$\begin{aligned} W(T) - W(C) &= p\Pr(s)\{E[U(y(s,1,T))] - E[U(y(s,0,C))]\} \\ &+ (1-p)\Pr(s)\{E[U(y(s,0,T))] - E[U(y(s,0,C))]\} \\ &= p\Pr(s)\{E[U(y(s,1,T))] - E[U(y(s,0,C))]\} \\ &= p\Pr(s)\{E[U(y(s,1,T))] - E[U(y(s,0,T))]\} \end{aligned} \qquad (6)$$

The first equality relies on the assumption of the absence of spillovers among the non-selected, while the second and third equalities rely on the assumption that there are no spillovers among the selected but untreated. Intuitively, when $p \in (0,1)$, we are able to derive unbiased estimates of the individual welfare of both the treated and untreated eligible households in the treated villages. The "no spillover" assumption implies that the welfare of the selected but untreated equals the welfare that the selected would have had, if the intervention had not been conducted. Hence the difference between the welfare levels of the treated and selected but untreated within treated villages $E[U(y(s,1,T))] - E[U(y(s,0,T))]$ is an unbiased estimate of the impact of the intervention on the treated $E[U(y(s,1,T))] - E[U(y(s,0,C))]$. By scaling this by the proportion of individuals treated $p\Pr(s)$, we obtain an unbiased estimate of the welfare impact relative to no intervention.

We apply this methodology to evaluate the welfare effect of our credit interventions, relative to a no-intervention counterfactual. Since only about 2.5 percent of the relevant population received the program, it is unlikely that there were spillover effects on the population that did not receive the program credit. Since the assignment of treatments across villages was randomized, this provides an unbiased estimate of the welfare impact.

3. The interventions

A non-profit microfinance institution conducted the agricultural credit interventions in 72 villages in the districts of Hugli and West Medinipur in the state of West Bengal, India.[7] The schemes were primarily designed to facilitate the cultivation of potatoes, the most profitable cash crop in this region. The loan size, duration, interest rate and dynamic repayment incentives were identical across the three interventions. The loans had a 4-month duration and were offered at an annual interest rate of 18 percent. This was considerably lower than the 25% per annum average interest rate that prevailed in the informal loan market. The first loans were offered in October 2010. Repayment was due in a single lumpsum at the end of 4 months, at which point the next cycle of loans began. Borrowers who repaid successfully were eligible for a larger loan in the subsequent cycle; those who did not were not allowed to borrow again. The program ran for three years in all, and data were collected through detailed farm surveys every four months throughout this period.

The study attempted to compare different approaches to identifying beneficiaries for subsidized agricultural credit. The agent-intermediated lending (AIL) approach taps into the knowledge and information about local residents that exists within a community but might be unobservable to researchers. Borrower selection was delegated to a local intermediary, who was appointed as the MFI's agent. In the 24 randomly selected villages assigned to the TRAIL intervention, the agent was a trader who bought agricultural output from village farmers and/or sold them farm inputs. In the 24 villages assigned to the GRAIL intervention, the agent was a political appointee. Each agent was asked to recommend 30 smallholder borrowers to potentially receive individual liability loans. In the 24 villages assigned to the Group Based Lending (or GBL) intervention, any smallholder village resident could form a group with four other members. After they had attended group meetings and made savings deposits for 6 months, all the members of two randomly selected groups were offered the program loans. Group members were jointly liable for each others' loans: if any member defaulted, all other group members were cut off from all future loans.[8] In all three interventions, households were only eligible for loans if they owned no more than 1.5 acres of agricultural land.

[5] In fact, the interventions we consider were implemented through a parsimonious design randomized controlled trial, involving three different treatment arms and no control arm, and with $p < 1$ for each of the three interventions. However we provide a general method that could be used for cases where a randomly selected subset of villages are pure control, that is, do not receive the intervention.

[6] This applies even when all selected subjects are treated (or $p = 1$), in which case there are no "selected but untreated" individuals.

[7] Here we provide a brief summary of the experimental details; these are discussed in greater detail in Maitra et al. (2017, 2021).

[8] The agents in the TRAIL and GRAIL schemes were incentivized through commissions equal to 75 percent of the interest paid by borrowers they had

Table 1
Household characteristics by village treatment arm.

	All (1)	TRAIL (2)	GRAIL (3)	GBL (4)
Head's education: primary or more	0.420	0.407	0.420	0.433
	(0.01)	(0.02)	(0.02)	(0.02)
Head's occupation: cultivation	0.431	0.441	0.415	0.438
	(0.01)	(0.02)	(0.02)	(0.02)
Head's occupation: Labor	0.335	0.340	0.343	0.323
	(0.01)	(0.02)	(0.02)	(0.02)
Non-Hindu	0.172	0.210	0.151	0.155
	(0.01)	(0.02)	(0.01)	(0.01)
Low caste	0.388	0.383	0.355	0.423
	(0.01)	(0.02)	(0.02)	(0.02)
Area of house and homestead (acres)	0.053	0.052	0.052	0.054
	(0.00)	(0.00)	(0.00)	(0.00)
Brick-and-mortar house	0.294	0.288	0.334	0.259
	(0.01)	(0.01)	(0.02)	(0.01)
Electrified house	0.751	0.740	0.752	0.760
	(0.01)	(0.01)	(0.01)	(0.01)
Separate toilet in house	0.575	0.564	0.608	0.552
	(0.01)	(0.02)	(0.02)	(0.02)
Owns radio/ TV/ VCR/ DVD	0.464	0.450	0.486	0.456
	(0.01)	(0.02)	(0.02)	(0.02)
Owns telephone (mobile or landline)	0.590	0.573	0.590	0.607
	(0.01)	(0.02)	(0.02)	(0.02)
Has savings bank account	0.456	0.447	0.475	0.446
	(0.01)	(0.02)	(0.02)	(0.02)
Farm income	7696.08	6138.99	8170.97	7333.40
	(357.8)	(312.4)	(370.9)	(201.3)
Observations	3120	1032	1050	1038
Test of joint significance for assignment to treatment				
Using a multinomial logit	χ^2 statistic		10.33	
	p-value		*0.993*	

Notes: Other than farm income, data are from the first wave of household surveys conducted in the 72 sample villages in 2010. Only eligible households are included in the sample. Since we drew a purposive sample of Treatment, Control 1 and Control 2 households, we do not expect our sample means to be representative of the village populations. To correct for the non-representativeness of our sample, we assign each household a weight that is in inverse proportion to the probability that they would be selected into the sample. Thus, Treatment and Control 1 households each receive a weight of $\frac{30}{N}$ and Control 2 households receive a weight of $\frac{N-30}{N}$, where N denotes the total number of households in the village, as reported in the 2011 Census. The test of joint significance does not include farm income, since it is endogenous to the intervention. Standard errors are in parentheses.

The experiment was designed to separately identify how selected borrowers differed from those not selected (selection effects), and the effect of the intervention conditional on selection (conditional treatment effects). Specifically, in the TRAIL and GRAIL arms, loans were offered to only 10 households randomly chosen from the 30 whom the agent had recommended in the village. By comparing these treated households with those that were also recommended but were not chosen to receive the loans, we can identify the effect of the TRAIL/GRAIL loans conditional on selection. Similarly, in the GBL arm only 2 of the joint liability groups that had formed in the village were randomly selected, and each member was offered the loans. Comparing them with households that also formed groups but were not randomly selected to receive loans, identifies the effect of the GBL loans, conditional on self-selection. Importantly, there were no pure control villages in the research design, and therefore we follow the methodology for the parsimonious design discussed in Section 2.

4. Data and descriptive statistics

Between 2010 and 2013, we conducted eight waves of surveys with a sample of 50 households in each of the 72 villages. In each village, the sample includes all 10 households that were randomly assigned to receive the subsidized loans (Treatment households) and 10 randomly selected households from among those that were also selected, but not assigned to receive the loans (Control 1 households). In addition, it includes 30 Control 2 households, which are randomly chosen from the set of households that were not selected (they were not recommended by the TRAIL/GRAIL agents or did not form GBL groups).[9]

In Table 1 we present summary statistics of household characteristics, farm and non-farm income, as reported in the first round of household surveys, for households that owned no more than 1.5 acres of land. As columns 2–4 show, the characteristics of households in the three treatment arms are very similar. The pair-wise differences are almost always statistically non-significant (results available upon request). Using a multinomial logit regression, we cannot reject the null hypothesis that on average, observable household characteristics do not explain assignment to treatment arm (p-value = 0.993).[10]

recommended. In the GBL scheme, the MFI that organized the group meetings received as commission 75 percent of the interest paid by GBL borrowers.

[9] We do not include Control 2 households in our sample when we estimate the conditional treatment effects in the next section.

[10] Since we drew a purposive sample of Treatment, Control 1 and Control 2 households, we do not expect our sample means to be representative of the village populations. To estimate representative means, we re-weight the sample to inflate each household in inverse proportion to the probability that they would be selected into the sample. Thus, Treatment and Control 1 households each receive a weight of $\frac{30}{N}$ and Control 2 households receive a weight of $\frac{N-30}{N}$, where N denotes the total number of households in the village, as reported in the 2011 Census. In this way we can scale up the sample proportions in each land category to arrive at the population proportions.

Table 2
Estimated Welfare Impacts of the Loan Interventions.

	$\theta = 0$ (1)	$\theta = 0.5$ (2)	$\theta = 1$ (3)	$\theta = 1.5$ (4)	$\theta = 2$ (5)
Panel A: Conditional treatment effects on household welfare					
TRAIL	2177.856	12.866	0.079	5.08E−04	3.43E−06
	(1388.13, 3008.86)	(8.027, 17.920)	(0.048, 0.112)	(2.92E−04, 7.41E−04)	(1.75E−06, 5.16E−06)
GRAIL	391.313	3.602	0.030	2.51E−04	2.07E−06
	(−569.11, 1339.80)	(−2.30, 9.44)	(−0.010, 0.069)	(−3.78E−05, 5.35E−04)	(−2.13E−07, 4.35E−06)
GBL	59.175	0.944	0.008	5.22E−05	2.30E−07
	(−1249.20, 1717.19)	(−6.69, 10.29)	(−0.040, 0.063)	(−2.73E−04, 4.03E−04)	(−2.11E−06, 2.66E−06)
Panel B: Change in aggregate welfare (using treatment proportions)					
TRAIL	49.845	0.294	0.002	1.16E−05	7.86E−08
	(19.057, 71.584)	(0.112, 0.426)	(0.001, 0.003)	(3.83E−06, 1.74E−05)	(2.14E−08, 1.21E−07)
GRAIL	9.162	0.084	0.001	5.87E−06	4.84E−08
	(−16.370, 31.429)	(−0.072, 0.221)	(−0.0003, 0.002)	(−2.08E−06, 1.27E−05)	(−1.50E−08, 1.03E−07)
GBL	1.006	0.016	0.000	8.87E−07	3.90E−09
	(−20.683, 32.898)	(−0.108, 0.188)	(−0.001, 0.001)	(−4.17E−06, 7.33E−06)	(−3.34E−08, 5.12E−08)
Panel C: Comparison of welfare effects					
P(GRAIL = TRAIL)	0.000	0.000	0.000	0.000	0.000
P(GBL = TRAIL)	0.000	0.000	0.000	0.000	0.000
P(GBL = GRAIL)	0.000	0.000	0.000	0.000	0.000

Notes: Welfare impacts are estimated following the procedure outlined in Section 2, where θ indicates the value of the inequality-aversion parameter. In Panels A and B, the terms in parentheses denote the bootstrapped 90% confidence interval (with 2000 replications). To compute the changes in aggregate welfare in Panel B we use the empirical treatment proportions: these are 2.3%, 2.3% and 1.7% for the TRAIL, GRAIL and GBL schemes respectively. Panel C presents the p-values from Mann–Whitney rank-sum tests using 2000 bootstrap replications, comparing the aggregate welfare effects of the schemes.

Table A.1 in the Appendix shows summary statistics about village demographics and infrastructure, computed using data from the 2011 Census of India and from a 2007 pre-intervention village census conducted for a different project (see Mitra et al., 2018).[11] We can reject the null hypothesis that these village-level characteristics can jointly explain assignment to treatment arm, indicating that the villages were balanced on these observables.

Recall that selected households were randomly assigned to Treatment and Control 1 groups. Maitra et al. (2017) (Table 1) and Maitra et al. (2021) (Table 1) show that they were balanced on observable household characteristics.

5. Empirical estimates

Following the methodology laid out in Section 2, below we describe our empirical estimates for the impact of the three interventions on aggregate welfare. We start by estimating the conditional treatment effects. These are then multiplied by the treatment proportions to arrive at unconditional estimates of the aggregate welfare effects. We also discuss the results of statistical tests comparing the welfare effects of the three interventions.

5.1. Computing the CTEs

To estimate the conditional treatment effects, we restrict the sample to Treatment and Control 1 households and run the following regression:

$$\begin{aligned} U(y_{ivt};\theta) = \beta_0 + \beta_1 \text{TRAIL}_v + \beta_2 \text{GRAIL}_v + \beta_3 \text{Treatment}_{iv} \\ + \beta_4 (\text{TRAIL}_v \times \text{Treatment}_{iv}) \\ + \beta_5 (\text{GRAIL}_v \times \text{Treatment}_{iv}) + \xi \mathbf{X}_{ivt} + \epsilon_{ivt} \end{aligned} \tag{7}$$

Here $\mathbf{X}_{ivt}$ is a set of household characteristic variables, including landholding, household caste and religion, the age, education and occupation of the oldest male in the household, year dummies and a dummy for the village information treatment.[12]

The dependent variable in the regression is household welfare $U(y_{ivt}) = \frac{y_{ivt}^{1-\theta}}{1-\theta}$, corresponding to a specific non-negative value of θ. The value $\theta = 1$ corresponds to the log utility function.[13] Here y_{ivt} is aggregate farm income for household i in village v in year t, which is the sum of value-added from the four major crop categories: potatoes, paddy, sesame and vegetables.[14] The explanatory variable TRAIL_v indicates whether village v was assigned to the TRAIL intervention and GRAIL_v indicates whether it was assigned to the GRAIL intervention. The default category includes villages assigned to the GBL intervention. Treatment_{iv} indicates if the household was assigned to receive the loan, thus these are intention to treat estimates. This allows us to estimate the conditional treatment effects on household welfare in the TRAIL villages as $\beta_3 + \beta_4$, in the GRAIL villages as $\beta_3 + \beta_5$, and in the GBL villages as β_3.

In each column of Table 2 Panel A, we present the results from running this regression for different values of the inequality-aversion parameter θ. When θ takes value 0, the welfare impact represents the change in average farm income. In line with results in Maitra et al. (2017, 2021), we find in column 1 that the TRAIL scheme increased the average farm income of recommended households by Rs.2178. The 90% bootstrapped confidence interval does not include zero.[15] The point estimates for both the GRAIL (Rs.391) and the GBL (Rs.59)

[11] This survey was conducted in 72 villages. However Maoist violence in 2010 forced us to replace four of the 72 villages from our 2007 sample. Therefore Table A.1 uses a sample of only 68 villages.

[12] The information intervention was undertaken for a separate project examining the effect of providing information about potato prices to farmers and is similar to the public information treatment described in Mitra et al. (2018). Villages were assigned to the information treatment randomly and orthogonally to the credit intervention we examine here. The results are unchanged if we exclude this dummy variable from the regression specification.

[13] To deal with the problem of negative farm income, we winsorize the top and bottom 1% of the distribution of farm income and add a constant of Rs. 15,000.

[14] Our farm surveys were timed with the planting and harvest seasons, and asked detailed questions about input quantity, input price, and output quantity data for each crop the household planted. Each individual sales transaction was recorded, and we tracked sales out of storage, as well as delayed payments. We believe this approach minimized measurement error. However in the Appendix we discuss how the welfare estimates can be corrected for possible measurement error.

[15] Bootstrap samples were drawn using a stratified (by treatment arm) clustered (by village) random procedure, to ensure that each sample contained an equal number of randomly drawn TRAIL, GRAIL and GBL villages. Once a village was drawn into the sample, all original sample households from that

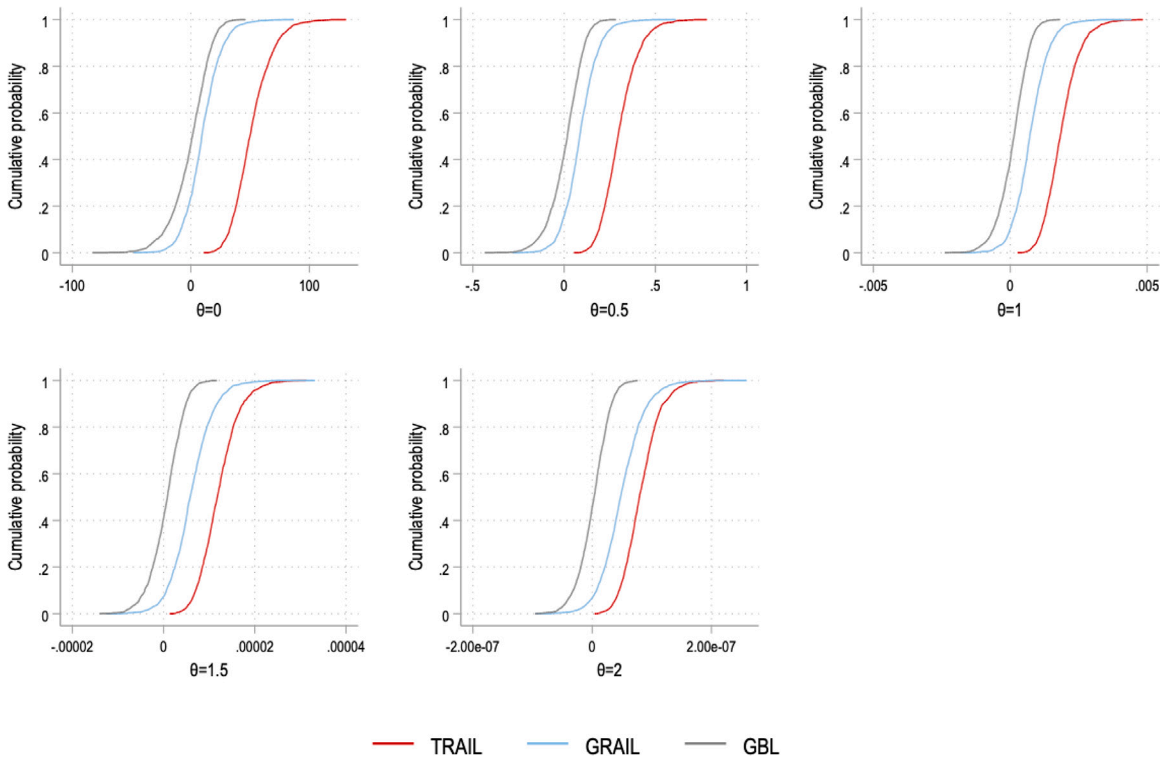

Fig. 1. Cumulative distribution functions of estimated changes in aggregate welfare for different inequality-aversion parameters.
Notes: Cumulative distribution functions are drawn from 2000 bootstrap estimates of aggregate welfare impacts of the TRAIL, GRAIL and GBL schemes.

schemes are smaller in magnitude and statistically indistinguishable from zero. In Table 2 columns 3–6, we estimate the welfare effect for four different θ values, increasing from 0.5 to 2 in 0.5 unit increments.[16] At higher values of θ, the farm incomes of low income households receive greater weight in the welfare calculation. Therefore, the same increase in farm income would have a larger impact on welfare if it accrued to a lower-income household than a higher-income household, and the welfare reduction caused by an increase in the inequality of farm income distribution is more likely to overpower the increased welfare caused by higher average incomes. However, we see that the TRAIL CTEs remain positive and statistically significant. This suggests either that very low-income households benefited from the TRAIL scheme relatively more than higher-income households, or that any increase in inequality is very small compared to the efficiency increase. The CTEs of the GRAIL and GBL schemes continue to have positive point estimates, but remain non-significant. Thus not only did these schemes have a small effect on the average household, there is also no evidence to suggest that they particularly benefited lower income households.

5.2. Computing the treatment proportions

Recall that households had to satisfy two conditions to receive loans offers: they had to be selected as potential beneficiaries (either through recommendation by the TRAIL/GRAIL agent, or through self-selection in GBL groups), and then had to be randomly assigned to the treatment group. The probability that a household would be selected is $\Pr(s)$. Next, the likelihood of being assigned to treatment is denoted by probability p. Thus the treatment proportion is given by $p\Pr(s)$. Our calculations show that across the 24 TRAIL villages, 6.9 percent of our sample households were recommended for loans. Since one-third of the recommended households were offered loans, the treatment proportion for the TRAIL intervention is ($6.9 \times 0.33 =$) 2.3 percent. Similarly, across the 24 GRAIL households, 7 percent of our sample households were recommended, and since one-third of them were offered loans, the treatment proportion is 2.3 percent. In GBL villages, 5.1 percent of our sample households formed groups, and the treatment proportion is 1.7 percent.

5.3. Change in aggregate welfare as a result of the interventions

In Panel B of Table 2, we compute the implied change in aggregate village welfare as the product of the conditional treatment effects as presented in Panel A, and the treatment proportions described above. This measures the change in aggregate welfare that would be expected if the intervention were introduced in a representative village. Once again, only the TRAIL scheme increased aggregate welfare at the θ values we consider. The point estimates for the GRAIL and GBL schemes are non-significant at all values of the inequality aversion parameter. We present confidence intervals from bootstrapped estimates of the aggregate welfare effects.[17] In Fig. A.1 we present the change in aggregate welfare as a proportion of base welfare (average welfare of Control 1 and Control 2 households, re-weighted) for $\theta \in \{0, 0.5, 1, 1.5, 2\}$. For every value of θ, the aggregate welfare impact of the TRAIL scheme exceeds that from the GRAIL scheme, which in turn is larger than the welfare impact of the GBL scheme.

village were included. In each bootstrap sample we estimate the conditional treatment effects of each scheme.

[16] The literature in public economics has generally considered values of θ between 0 and 2 (Braverman et al., 1987; Newbery, 1987; Bhattacharya and Komarova, 2021). In Table A.2 we present estimates with θ = 3, 4 and 5. As the value of θ increases, the comparisons become stable.

[17] Since the treatment proportions are sample estimates, the conversion from CTEs to aggregate welfare effects has introduced sampling error. Therefore, in each bootstrap sample we estimate the treatment proportion and multiply this with the estimated conditional treatment effect to obtain the bootstrapped treatment effect on average welfare.

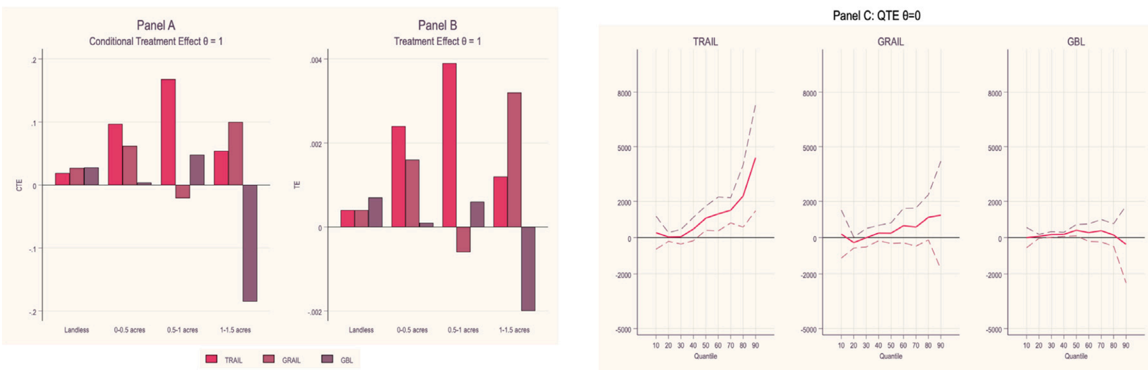

Fig. 2. Conditional and unconditional treatment effects by land class and QTEs.
Notes: Panel A shows the estimated CTEs across the four land classes, for $\theta = 1$. The unconditional treatment effects in Panel B are obtained by multiplying the CTE estimate for each land class by the treatment proportions, or the fraction of group g households in the villages that received the treatment, so that $\text{TE}_g = \alpha_g s_g \text{CTE}_g$. Panel C presents the Quantile Treatment effects for $\theta = 0$. The treatment proportions are 0.020, 0.025, 0.023 and 0.022 for the four land classes in the TRAIL intervention 0.013, 0.026, 0.029 and 0.032 for the GRAIL intervention, and 0.024, 0.018, 0.013 and 0.011 for the GBL intervention. See Fig. A.2.

5.4. Comparing the welfare impacts of the interventions

Welfare estimates at different inequality aversion levels are not comparable. Instead, our approach is more useful for comparisons across interventions, for a given level of inequality aversion (θ).

In Panel C of Table 2, we compare the average welfare impacts of the three interventions. We conduct pair-wise Mann–Whitney rank-sum tests on the 2000 bootstrap replications. At all θ values, in each pair-wise comparison we can reject the null hypothesis that the interventions generate identical welfare effects.

In Fig. 1 we present the cumulative distribution functions of these estimated changes in aggregate village welfare from our 2000 bootstrap replications. These corroborate our findings from the rank-sum tests. At every value of θ, the TRAIL scheme first order stochastically dominates both the GRAIL and the GBL schemes.

6. Welfare decomposition

A decomposition of the aggregate welfare effects by sub-group can help understand how the effects of the intervention differ among households with different characteristics. Since the Atkinson welfare function is additively separable, the average welfare impact of the intervention can be written as a weighted average of the conditional treatment effects on welfare of the different groups. Letting g denote the subgroup that the household lies in, we can write:

$$W(T) - W(C) = \sum_g \alpha_g s_g \text{CTE}_g \tag{8}$$

where α_g denotes the demographic weight of group g, and s_g denotes the fraction of group g that was treated.[18] The conditional treatment effect CTE_g equals $E[U(y)|s,1,T,g] - E[U(y)|s,0,T,g]$, where $E[U(y)|s,1,T,g]$ and $E[U(y_i)|s,0,T,g]$ denote average utility among Treatment and Control 1 subjects respectively, within group g in treatment villages. Thus, the overall welfare impact can be expressed as the (population share) weighted average of the product of $\alpha_g s_g$ and the CTEs across different groups.

6.1. Empirical decomposition results by landholding groups

In principle, this decomposition could be conducted along any dimension of inter-household heterogeneity. Below, we decompose the welfare effect across different classes of land ownership. We make this choice for multiple reasons. One, landholding is one of the most important and stable observable determinants of variation in farm income.[19] In an analysis of variance (ANOVA) between farm income and explanatory variables landholding, area of the homestead, presence of an in-house toilet, brick-and-mortar house construction, electricity connection, ownership of a radio/television, a savings account and a telephone (landline or mobile), household size, head's education and indicators for the head's occupation, landholding explains 42% of the total explained variation from the model. Two, as Table A.3 in the Appendix shows, landholding is also a good predictor of households' socio-economic status more broadly. In households that owned more land, heads were more likely to have completed primary school, and were more (less) likely to report their main occupation as cultivation (casual labor). Households with more land lived in larger houses that were more likely to be constructed with brick and mortar (*pucca*), have electrical connections, and an in-house toilet. They were also more likely to own televisions or other audio-visual electrical appliances, telephones, and bank savings accounts.[20]

To estimate the conditional treatment effects (CTE_g) on welfare for each land group g, we run the following regression:

$$\begin{aligned} U(y_{igvt}) = {} & \sum_{g=1}^{G} \gamma_g (Z_{igv}) + \sum_{g=1}^{G} \delta_g (Z_{igv} \times \text{TRAIL}_v) + \sum_{g=1}^{G} \zeta_g (Z_{igv} \times \text{GRAIL}_v) \\ & + \sum_{g=1}^{G} \eta_g (Z_{igv} \times \text{Treatment}_{igv}) \\ & + \sum_{g=1}^{G} \xi_g (Z_{igv} \times \text{TRAIL}_v \times \text{Treatment}_{igv}) \\ & + \sum_{g=1}^{G} \kappa_g (Z_{igv} \times \text{GRAIL}_v \times \text{Treatment}_{igv}) + \lambda \mathbf{X}_{ivt} + \epsilon_{igvt} \end{aligned} \tag{9}$$

where Z_{igv} is an indicator for whether household i in village v belongs to land category g. The sample is restricted to Treatment and Control 1 households. The TRAIL, GRAIL and GBL conditional treatment effects for a household in land group g are given by $\eta_g + \xi_g$, $\eta_g + \kappa_g$ and η_g respectively; $\mathbf{X}_{ivt}$ is as defined earlier.

At positive levels of the inequality aversion parameter θ, Eq. (9) delivers a sub-group analysis of the welfare effect, where higher θ

[18] We can think of $\alpha_g s_g$ as the treatment proportion for group g.

[19] In the three years that we followed our sample, only 6% of households bought or sold any land.

[20] Within each land category household characteristics were balanced across treatment arms. In each land class, we cannot reject the null hypothesis that these household characteristics do not predict assignment to treatment arm.

values represent welfare functions that place more weight on lower-income households within a land class.[21] Following Eq. (8), we multiply the estimates with s_g, or the fraction of group g households that received the treatment, to obtain the (unconditional) treatment effects for each land class.

This decomposition also illustrates how our welfare estimation approach differs from the QTE approach. In Fig. 2, Panels A and B show the land group-specific conditional and unconditional treatment effects on welfare when $\theta = 1$. In this specification, changes in individual welfare correspond to proportional income changes. We restrict the discussion to the TRAIL intervention, since the welfare effects of the GRAIL and GBL interventions are never statistically different from zero. The CTE estimates in Panel A show that the TRAIL scheme increased farm income for landless households by 1.9 percent, for the land class 0–0.5 acres by 9.7 percent, for those with 0.5–1 acre by 16.8 percent, and for the highest land class of 1–1.5 acres by 5.4 percent. Thus, the scheme benefited those in the intermediate land groups by more than those in either the lowest or highest land categories. Our data suggest that TRAIL agents were equally likely to recommend households in all four land class groups (see Fig. A.2).[22] As a result, the unconditional treatment effects in Panel B follow a very similar pattern to the conditional treatment effects.

On the other hand, the quantile treatment effects presented in Panel C show that the TRAIL scheme increased all income deciles above the 5th decile, and the effect was not different from zero at lower deciles. Thus they indicate that the scheme increased inequality, but do not provide a method to evaluate how this trades off against the increased average incomes. In particular, the QTE indicates that the total impact on farm income was highest for the largest landholding group, obscuring the fact that this group experienced a smaller proportional change in income compared to intermediate landholding groups. As the decomposition exercise indicates in Panels A and B, the larger proportional increase in income for the intermediate landholding group contributed to a reduction in inequality.

7. Conclusion

Our approach has several advantages. Its conceptual underpinnings provide a clear rationale for the welfare evaluations that it generates. It is simple to implement and to understand, and does not require strong assumptions for a clear-cut interpretation: the average treatment effect on welfare is consistently estimated under the same assumptions used to estimate the average treatment effect on the outcome variable itself. Importantly, in contrast to commonly used approaches that generate a vector of heterogeneous treatment effects for different beneficiary groups, or at quantiles of the outcome distribution, our method provides a single summary quantitative measure of the welfare impact.

We also highlight the fact that the Atkinson approach and the quantile treatment effects approach provide different insights. Quantile treatment effects estimate impacts of the intervention at different quantiles of the outcome distribution. The Atkinson approach explicitly builds in a consideration of the distributional effects, but does not estimate the effects at different points of the outcome distribution. The aggregate welfare impact accounts for both the efficiency and equity impacts of the intervention, for any chosen level of inequality aversion.

[21] Note that when $\theta = 0$, Eq. (9) delivers sub-group analysis of treatment effects on income, similar to the standard in the literature. As we see in Fig. A.3 in the Appendix, the qualitative patterns are similar for $\theta = 0.5$.

[22] While TRAIL agents were equally likely to recommend households in any of the four landholding classes; the selection pattern of GRAIL agents was regressive: they were more likely to recommend households with larger landholdings. The GBL scheme disproportionately attracted landless households.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data is available online.

Appendix

A.1. Additional tables and figures

A.2. Estimation of distributive impacts when income is measured with error

We discuss here how the calculation of the average treatment effects on welfare is affected by classical measurement error in income.

Let the model be written as

$$Y_i = \alpha + \beta \mathrm{T}_i + \epsilon_i \tag{A.1}$$

where Y_i is true income of farmer i in a given land class, T_i is an indicator for the farmer's treatment status and ϵ_i is a zero-mean i.i.d. residual.

Suppose instead the researcher observes

$$y_i = Y_i + m_i \tag{A.2}$$

where m_i is classical zero-mean measurement error.

Then we have

$$y_i = \alpha + \beta \mathrm{T}_i + e_i \tag{A.3}$$

where $e_i (\equiv \epsilon_i + m_i)$ has zero mean, and a variance greater than ϵ_i. In other words, the measurement error causes the variance of observed income to be greater than those of true income. Also since errors are normal, odd order moments are not affected.

The welfare of household i, given by $U(\cdot)$, is a continuous, increasing, concave function of Y_i. Then, if this household is in the treatment group, its expected welfare is given by

$$\mathrm{U}_T = EU(\alpha + \beta + \epsilon_i) \tag{A.4}$$

whereas if instead it is in the control group, its expected welfare is given by

$$\mathrm{U}_C = EU(\alpha + \epsilon_i) \tag{A.5}$$

giving us this expression for the average treatment effect

$$\mathrm{ATE} = \mathrm{U}_T - \mathrm{U}_C = EU(\alpha + \beta + \epsilon_i) - EU(\alpha + \epsilon_i) \tag{A.6}$$

whereas measured ATE equals:

$$\mathrm{ATE}^m = EU(\alpha + \beta + \epsilon_i + m_i) - EU(\alpha + \epsilon_i + m_i)$$

Using Taylor expansions up to the kth degree to approximate ATE and ATE^m, we obtain

$$\mathrm{ATE} = \sum_{k=1}^{K} (\mu_{1k} - \mu_{0k}) \sigma^k \tag{A.7}$$

$$\mathrm{ATE}^m = \sum_{k=1}^{K} \mu_{1k} - \mu_{0k} \sigma_m^k \tag{A.8}$$

where σ^k and σ_m^k respectively denote kth order moments of ϵ_i and e_i respectively, and

$$\mu_{1k} = \frac{1}{k!} u^k (\alpha + \beta) \tag{A.9}$$

$$\mu_{0k} = \frac{1}{k!} u^k (\alpha) \tag{A.10}$$

where u^k is the kth order derivative of the welfare function.

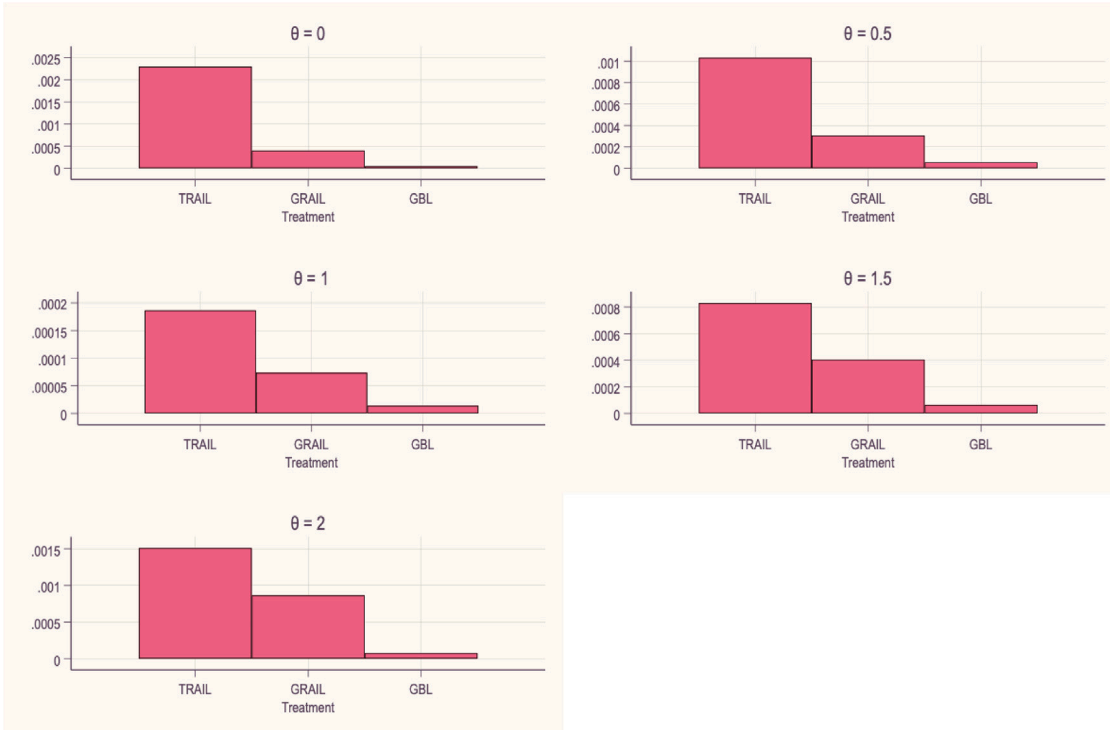

Fig. A.1. Change in aggregate welfare expressed as ratio of base welfare.
Notes: The height of each bar shows the change in aggregate welfare caused by the intervention, as a ratio of base welfare (average welfare of Control 1 and Control 2 households, re-weighted) in the villages where the intervention was implemented. The absolute values of change in aggregate welfare are presented in Table 2.

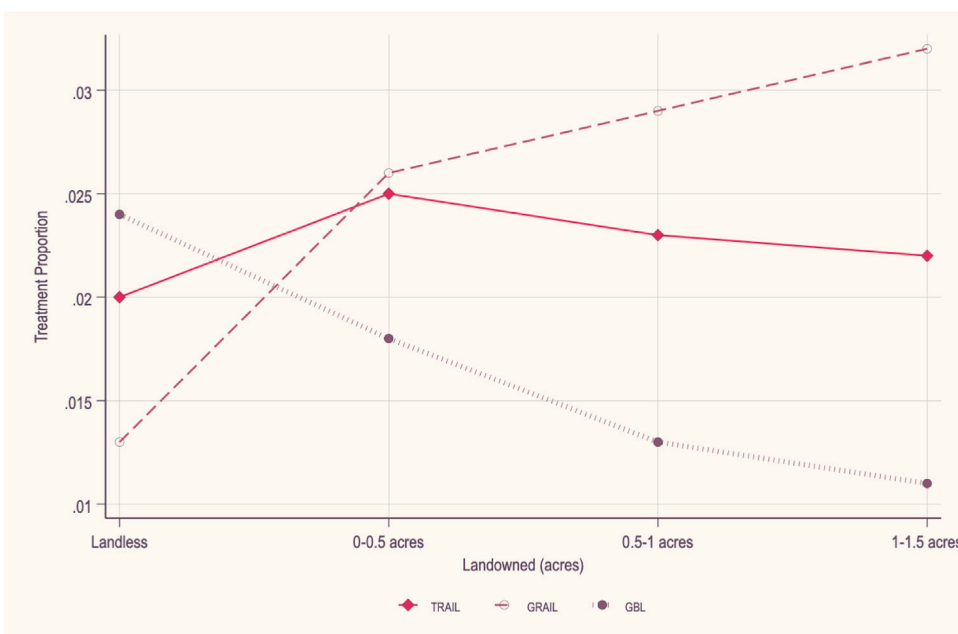

Fig. A.2. Treatment proportions, by intervention and land class.
Notes: The points represent the treatment proportions $p\Pr(s)$, given as the product of the probability a given land class is selected as a potential beneficiary of the intervention scheme, and the fraction of selected households that were randomly chosen to be offered the program loan.

If for example, the welfare function is quadratic in income, given by $U(y) = y^2$, then $U^1(y) = 2y$ and $U^2(y) = 2$, then the bias in estimating ATE is approximated by $\text{ATE}^m - \text{ATE} = \sum_{k=1}^{K}[\mu_{1k} - \mu_{0k}][\sigma_m^k - \sigma_m]$. Then, when $k = 1$, $\sigma^1 = E(e_i) = 0$. When $k = 2$, $\mu_{12} = \mu_{02} = 2$, while $\mu_{1k} = \mu_{0k} = 0 \forall k > 2$. The third term is zero, and measurement error does not add any bias.

If instead, as in our paper, the welfare function belongs to the family of Atkinson functions, then we have

$$U(y) = \frac{y^{1-\theta}}{1-\theta}; \theta \neq 1 \quad \text{(A.11)}$$

and $\mu_{12} = -\theta(\alpha + \beta)^{-\theta-1}$; $\mu_{02} = -\theta(\alpha)^{-\theta-1}$.

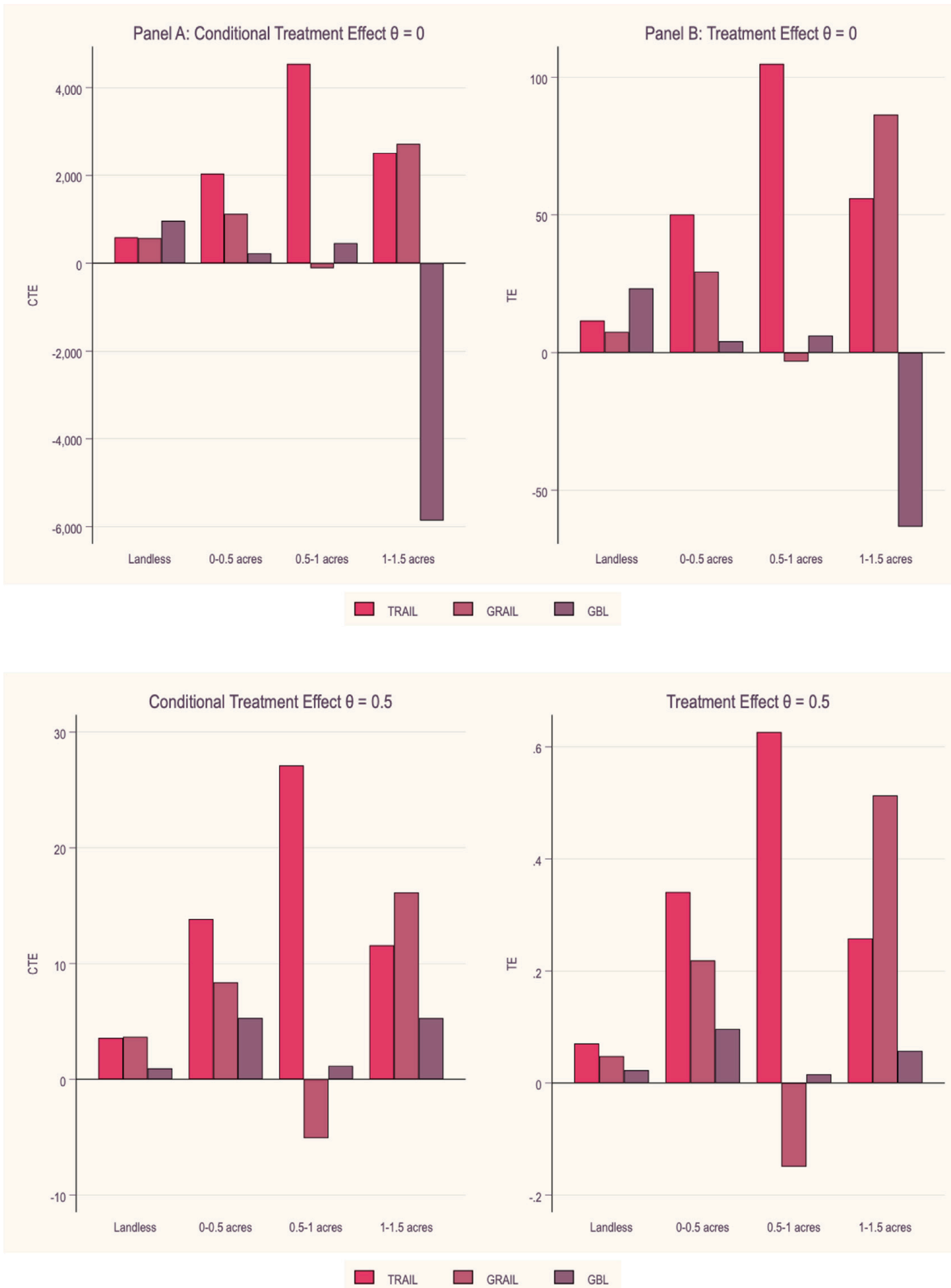

Fig. A.3. Decomposition of Welfare Effects by Land Class, $\theta = 0$ & 0.5.
Notes: Panel A shows the estimated CTEs across the four land classes following Eq. (9), for θ values 0 and 0.5. The unconditional treatment effects in Panel B are obtained by multiplying the CTE estimate for each land class by the treatment proportions, as per Eq. (8). The treatment proportions are 0.020, 0.025, 0.023 and 0.022 for the four land classes in the TRAIL intervention 0.013, 0.026, 0.029 and 0.032 for the GRAIL intervention, and 0.024, 0.018, 0.013 and 0.011 for the GBL intervention. See Fig. A.2.

Then, if measurement error inflates measured income variance by x%, the second moment is overestimated by $x\sigma^2$, and contributes to the bias

$$\frac{\sigma^2 x}{2}[\mu_{12} - \mu_{02}] = \frac{-\theta\sigma^2 x}{2}[(\alpha+\beta)^{-\theta-1} - \alpha^{-\theta-1}] \tag{A.12}$$

where σ^2, α is the mean of the control group, and β is the estimated average treatment effect on income. Ignoring moments of order higher than 2, we can use Eq. (A.12) to calculate the bias resulting from a given level (x) of income measurement error.

To check whether our welfare effect comparisons across the three treatments are robust to different levels of measurement error, we

Table A.1
Balance of Village-level Characteristics, by Village Treatment Arm.

	All	TRAIL	GRAIL	GBL	Differences: Two-way comparisons		
	(1)	(2)	(3)	(4)	(2)–(3)	(2)–(4)	(3)–(4)
Number of households	377.75	340.77	327.65	457.58	-13.14	−116.81	−129.95
	(42.57)	(56.24)	(69.74)	(113.93)	(89.59)	(106.88)	(113.93)
Proportions by landholding class							
Landless	0.18	0.18	0.18	0.17	0.00	0.02	0.02
	(0.01)	(0.03)	(0.03)	(0.02)	(0.04)	(0.03)	(0.03)
0–0.5 acres	0.34	0.32	0.36	0.34	−0.04	−0.02	0.02
	(0.02)	(0.03)	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
0.5–1 acre	0.22	0.23	0.20	0.22	−0.03	0.01	0.02
	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
1–1.5 acres	0.10	0.10	0.11	0.11	−0.01	−0.01	0.00
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)
> 1.5 acres	0.17	0.17	0.15	0.17	0.02	−0.00	0.02
	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)
Percent households electrified	0.615	0.603	0.652	0.591	−0.049	0.01	0.061
	(0.03)	(0.06)	(0.05)	(0.05)	(0.08)	(0.08)	(0.08)
Has primary school	0.779	0.773	0.773	0.792	0.00	−0.02	−0.02
	(0.05)	(0.09)	(0.09)	(0.08)	(0.129)	(0.12)	(0.12)
Has primary health center	0.221	0.273	0.182	0.208	0.09	0.06	−0.03
	(0.05)	(0.10)	(0.08)	(0.08)	(0.13)	(0.13)	(0.12)
Has bank branch	0.074	0.00	0.045	0.167	−0.05	−0.17	−0.12
	(0.03)	(0.00)	(0.05)	(0.08)	(0.05)	(0.08)	(0.09)
Has MFI	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Has tarred road	0.353	0.27	0.36	0.42	−0.09	−0.14	−0.05
	(0.06)	(0.10)	(0.11)	(0.10)	(0.14)	(0.14)	(0.15)
N	68	22	22	24			
F-test of joint significance					0.45	1.11	0.51
p-value					*0.81*	*0.37*	*0.77*

Notes: The number of households in the 72 sample villages is taken from the 2011 Census of India village directory. Proportions of households in each landholding class are calculated from the 2007 house-listing exercise we conducted in 68 of these 72 villages for a previous studies reported in Maitra et al. (2017, 2021). Other village-level characteristics are sample means from 68 village surveys conducted in 2007. Four villages from the Maitra et al. (2021) study were replaced in 2010 because of Maoist conflict, and we do not have pre-intervention village census or village survey data for the replacements. Number in parentheses are standard errors.

Table A.2
Estimated Welfare Impacts of the Loan Interventions for Higher Values of θ.

	$\theta = 3$	$\theta = 4$	$\theta = 5$
	(1)	(2)	(3)
Panel A: Conditional treatment effects on household welfare			
TRAIL	7.49E−10	1.21E−13	2.04E−17
	(2.816E−10, 1.211E−09)	(3.004E−14, 2.073E−13)	(3.670E−18, 3.632E−17)
GRAIL	6.09E−10	1.22E−13	2.38E−17
	(−1.900E−10, 1.320E−09)	(−8.067E−14, 2.863E−13)	(−2.192E−17, 5.851E−17)
GBL	−1.03E−10	−5.72E−14	−1.58E−17
	(−6.755E−10, 4.923E−10)	(−1.579E−13, 4.942E−14)	(−3.316E−17, 3.380E−18)
Panel B: Change in aggregate welfare (using treatment proportions)			
TRAIL	1.71E−11	2.78E−15	4.68E−19
	(3.545E−12, 2.776E−11)	(4.093E−16, 4.770E−15)	(7.330E−20, 8.269E−19)
GRAIL	1.43E−11	2.86E−15	5.58E−19
	(−5.860E−12, 3.142E−11)	(−2.196E−15, 6.768E−15)	(−5.729E−19, 1.388E−18)
GBL	−1.76E−12	−9.72E−16	−2.68E−19
	(−1.094E−11, 1.063E−11)	(−2.633E−15, 1.489E−15)	(−5.607E−19, 1.870E−19)
Panel C: Comparison of welfare effects			
P(GRAIL = TRAIL)	0.000	0.000	0.761
P(GBL = TRAIL)	0.000	0.000	0.000
P(GBL = GRAIL)	0.000	0.000	0.000

Notes: Welfare impacts are estimated following the procedure outlined in Section 2, where θ indicates the value of the inequality-aversion parameter. In Panels A and B, the terms in parentheses denote the bootstrapped 90% confidence interval (with 2000 replications). To compute the changes in aggregate welfare in Panel B we use the empirical treatment proportions: these are 2.3%, 2.3% and 1.7% for the TRAIL, GRAIL and GBL schemes respectively. Panel C presents the p-values from Mann–Whitney rank-sum tests using 2000 bootstrap replications, comparing the aggregate welfare effects of the schemes.

do the following. First, for each loan intervention scheme and each inequality aversion parameter value $\theta \in \{0.5, 1, 1.5, 2\}$, we compute the bias for different assumptions about the extent by which measurement error inflated measured income variance, ranging from 5% to 20%, in increments of 5 percentage points. Next, we correct our actual estimates for this bias. To check if the welfare effects differ between the different interventions, we generate a distribution of corrected welfare estimates from a bootstrapped sample of 2000 draws from our dataset. We then carry out ranksum test on the corrected welfare estimates, in the same manner as we do for the actual welfare estimates (as reported in Table 2 in the paper).

In Fig. A.4 in the Appendix we present the bias-corrected conditional treatment effects as a percentage of the base welfare (average welfare of Control 1 and Control 2 households) for different assumed

Table A.3
Household Characteristics by Landholding Class, Pooled TRAIL, GRAIL & GBL Samples.

	Landless (1)	0–0.5 acres (2)	0.5–1 acre (3)	1–1.5 acres (4)
Head's education: primary or more	0.234	0.356	0.564	0.650
	(0.02)	(0.01)	(0.02)	(0.03)
Head's occupation: cultivation	0.056	0.381	0.689	0.696
	(0.01)	(0.01)	(0.02)	(0.02)
Head's occupation: Labor	0.677	0.404	0.089	0.041
	(0.02)	(0.01)	(0.01)	(0.01)
Non-Hindu	0.188	0.146	0.195	0.181
	(0.02)	(0.01)	(0.04)	(0.02)
Disadvantaged caste	0.565	0.417	0.286	0.203
	(0.02)	(0.01)	(0.02)	(0.021)
Area of house and homestead (acres)	0.037	0.048	0.063	0.074
	(0.00)	(0.00)	(0.00)	(0.00)
Brick-and-mortar house	0.207	0.280	0.344	0.379
	(0.02)	(0.01)	(0.02)	(0.03)
Electrified house	0.666	0.729	0.811	0.841
	(0.02)	(0.01)	(0.01)	(0.02)
Separate toilet in house	0.434	0.541	0.664	0.741
	(0.02)	(0.01)	(0.02)	(0.02)
Owns radio/ TV/ VCR/ DVD	0.350	0.420	0.541	0.639
	(0.02)	(0.01)	(0.02)	(0.03)
Owns telephone (mobile or landline)	0.446	0.528	0.706	0.796
	(0.02)	(0.01)	(0.02)	(0.02)
Has savings bank account	0.268	0.410	0.576	0.680
	(0.02)	(0.01)	(0.02)	(0.02)
Farm income	869.59	4660.63	12044.86	17871.63
	(139.84)	(197.25)	(440.04)	(867.43)
Observations	666	1263	796	365
Test of joint significance for assignment to treatment				
χ^2 statistic	32.64	28.30	24.09	32.08
p-value	*0.112*	*0.248*	*0.456*	*0.125*

Notes: Data are from the first round of household surveys. Disadvantaged caste includes Scheduled Caste (SC) and Scheduled Tribe (ST) households. The test of joint significance does not include farm income, since it is endogenous to the intervention. Standard errors are in parentheses.

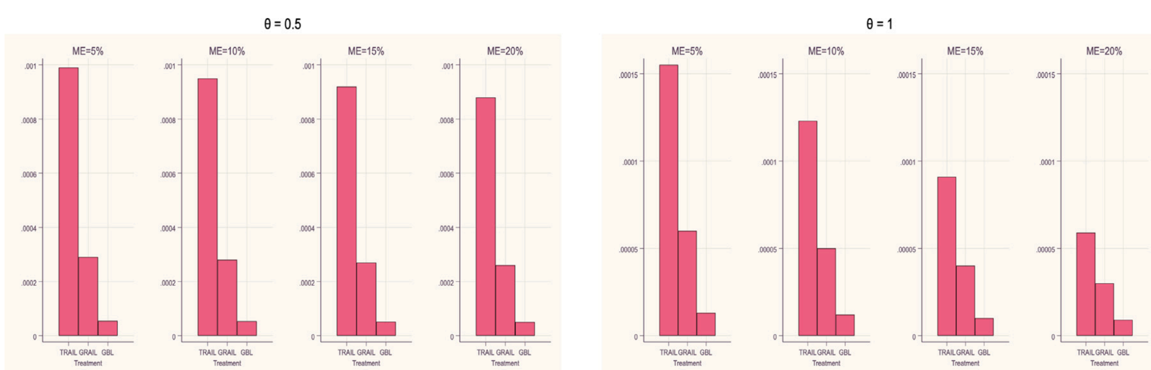

Fig. A.4. Conditional treatment effects corrected for measurement error bias.
Notes: Bias-corrected conditional treatment effects as a percentage of the base welfare (average welfare of Control 1 and Control 2 households, re-weighted) presented for different assumed extents of measurement error in farm income, by θ and intervention arm.

extents of measurement error in farm income, by θ and treatment. For all values of θ and all assumed levels of measurement error, the corrected CTE in TRAIL > GRAIL > GBL.

References

Ahmad, E., Stern, N., 1987. Alternative sources of government revenue: Illustrations from India 1979-80. In: Newbery, D., Stern, N. (Eds.), The Theory of Taxation for Developing Countries. World Bank: Oxford University Press.

Atkinson, A.B., 1970. On the measurement of inequality. J. Econom. Theory 2, 244–263.

Bedoya, G., Bittarello, L., Davis, J., Mittag, N., 2017. Distributional Impact Analysis: Toolkit and Illustrations of Impacts Beyond the Average Treatment Effect. Policy Research Working Paper 8139, World Bank.

Bhattacharya, D., Komarova, T., 2021. Incorporating Social Welfare in Program-Evaluation and Treatment Choice. Technical Report, Mimeo: Cambridge University and London School of Economics.

Braverman, A., Hammer, J., Ahn, C., 1987. Multimarket analysis of agricultural pricing policies in Korea. In: Newbery, D., Stern, N. (Eds.), The Theory of Taxation for Developing Countries. World Bank: Oxford University Press.

Coady, D., Harris, R., 2004. Evaluating transfer programs within a general equilibrium framework. Econom. J. 114, 778–799.

Hughes, G., 1987. The incidence of fuel taxes: A comparative study of three countries. In: Newbery, D., Stern, N. (Eds.), The Theory of Taxation for Developing Countries. Oxford University Press.

Maitra, P., Mitra, S., Mookherjee, D., Motta, A., Visaria, S., 2017. Financing smallholder agriculture: An experiment with agent-intermediated microloans in India. J. Dev. Econ. 127, 306–337.

Maitra, P., Mitra, S., Mookherjee, D., Visaria, S., 2021. Decentralized Targeting of Agricultural Credit Programs: Private versus Political Intermediaries. Working paper, http://people.bu.edu/dilipm/wkpap.

Mitra, S., Mookherjee, D., Torero, M., Visaria, S., 2018. Asymmetric information and middleman margins: An experiment with west bengal potato farmers. Rev. Econ. Stat. 100 (1).

Newbery, D., 1987. Identifying desirable directions of agricultural price reform in Korea. In: Newbery, D., Stern, N. (Eds.), The Theory of Taxation for Developing Countries. World Bank: Oxford University Press.
Newbery, D.M., 1995. The distributional impact of price changes in Hungary and the United Kingdom. Econom. J. 105 (431), 847–863.
Newbery, D., Stern, N. (Eds.), 1987. The Theory of Taxation for Developing Countries. The World Bank: Oxford University Press.
Roberts, K., 1980. Interpersonal comparability and social choice theory. Rev. Econom. Stud. 47 (2), 421–439.