Evaluating the distributive impacts of a micro-credit intervention

Evaluating the Distributive Effects of a Micro-Credit Intervention*

Pushkar Maitra† Sandip Mitra‡ Dilip Mookherjee§ Sujata Visaria¶

June 2021

Abstract

Most analyses of randomized controlled trials of development interventions estimate an average treatment effect. However, the aggregate impact on welfare also depends on distributional effects. We propose a simple approach to evaluate efficiency-equity trade-offs, that follow in the utilitarian tradition of Atkinson (1970). The method does not impose additional assumptions or data requirements beyond those needed to estimate the average treatment effect. We illustrate the approach using data from a credit delivery experiment we implemented in West Bengal, India.

Key words: Distributive Impacts, Program Evaluation, Agricultural Finance

JEL Codes: D82, O16, C93, H21

*Funding was provided by the Australian Agency for International Development (CF09/650), the International Growth Centre (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, 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 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.

†Pushkar Maitra, Department of Economics, Monash University, Clayton Campus, VIC 3800, Australia. pushkar.maitra@monash.edu.au

‡Sandip Mitra, Sampling and Official Statistics Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata 700108, India. Sandip@isical.ac.in

§Dilip Mookherjee, Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215, USA. dilipm@bu.edu

¶Sujata Visaria, Department of Economics, Lee Shau Kee Business Building, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong. svisaria@ust.hk

1 Introduction

It is now common to conduct randomized controlled trials to evaluate the impact of development policy interventions. The focus of most such evaluations is the *average treatment effect* (ATE), or the expected change in the beneficiary's outcome due to the intervention. However, the average treatment effect masks the fact that impacts may differ for program participants, with resulting implications for inequality. 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. In the first, beneficiaries are classified into sub-groups according to certain fixed characteristics (e.g., gender) or baseline levels of well-being (e.g., baseline income or wealth), and treatment effects are estimated for each sub-group separately. However this does not provide a quantitative summary measure of the impact on distributive equity. The second method is to estimate the treatment effects on different quantiles of the outcome distribution. For example, one may estimate the effect of the intervention on the median of the outcome variable, or its 25^{th} or 75^{th} percentile. However, quantile treatment effects do not allow straightforward inferences about the distributional impacts of the intervention (Bedoya et al., 2017). Crucially, the treatment effect on a particular quantile of the outcome distribution is only equivalent to the treatment effect on that quantile of the baseline outcome distribution if beneficiaries maintain the same rank in the outcome distribution in both treatment and counterfactual conditions (rank ivariance). This is a non-trivial requirement, and there is little to suggest that it is generally satisfied in the data.

In this paper we propose an alternative approach rooted in the utilitarian tradition of public economics going back to Atkinson (1970). A well established, although sparse, 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 Newbery and Stern, 1987, Hughes, 1987, Newbery, 1995, Coady and Harris, 2004). Our method shows how this approach can be applied to evaluate efficiency and equity impacts even for small-scale randomized interventions.

In the Atkinson approach, social welfare is represented as the sum of the welfare of the individuals in a population, as evaluated by an impartial observer, aid donor or social planner. This incorporates both efficiency and distributive implications. The welfare of an individual is an increasing, concave, iso-elastic function of the individual's wellbeing $U_i = U(y_i) \equiv \frac{y_i^{1-\theta}}{1-\theta}$, where $\theta > 0, \neq 1$ and wellbeing y_i is proxied by income or consumption. The welfare function $U(\cdot)$ reflects ethical judgments of the external observer, in the "extended sympathy" approach to social choice theory.[1]

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

Roberts (1980) provides axioms of cardinality and comparability of wellbeing that characterize this class of welfare functions.

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 methodology can be applied to consumption as well. The parameter θ represents the degree of *inequality aversion* incorporated into the welfare function. When $\theta = 0$, the measure reduces to the sum of incomes, thus ignoring income distribution entirely. When $\theta = 1$, welfare $U(y) = \log y$ and marginal welfare weights are inversely proportional to income; so income changes of poorer households receive greater weight. 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 θ approaches $+\infty$, the resulting expression for 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. This requires that we estimate the average impact on a given monotone function of well-being, rather than well-being itself. This is a straightforward exercise that does not require any additional assumptions beyond those used to estimate the standard average treatment effect. Note, in particular, that we do not need to assume rank-invariance. Moreover, the sensitivity of the assessed impacts to the value of θ, tells us whether distributive considerations greatly change the welfare assessment.

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 as well as 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 apply 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 he was appointed by the local government. In a third intervention (Group

Based Lending or GBL), borrower groups could self-form and apply for joint liability loans.[2]

We find that the TRAIL intervention had statistically significant positive welfare impacts across 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 at any of the θ values. By examining how the impacts differ across four different landholding classes, and decomposing the average treatment effects by land category, we are better able to understand the underlying mechanisms.Specifically, the TRAIL scheme has a larger impact on welfare irrespective of the degree of inequality aversion, because it increased the welfare of landless households, the poorest group by more.

Section 2 below explains our empirical methodology. The rest of the paper is devoted to the particular application. Section 3 describes the interventions in more detail. Further details about the data and descriptive statistics are presented in Section 4, while Section 5 presents the welfare ATE estimates for the three schemes and for different values of θ. In Section 6 we show the welfare decomposition of ATEs by land class.

2 Methodology

Consider a population of villages, which are randomized into a treatment group denoted T and a control (or 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. As we mentioned above, selection could take place on criteria that are observable to the researcher (e.g. landholding, occupation) or on others that are unobservable (such as when individuals self-select to opt in, or when an intermediary selects individuals to offer the program to). Since the intervention is randomly assigned across villages, $\Pr(s)$ describes the expected fraction of s types in both T and C villages.

Next, consider whether an individual actually receives the intervention. Let $e \in \{0, 1\}$ denote whether a specific individual receives the intervention, and $p \equiv P(e = 1|s, T)$ denote the fraction of s types in a T village that receive the intervention. By construction, the intervention is available to none of the individuals in the C villages. It is also unavailable to type–n households

[2]In previous work (Maitra et al., 2017, 2021), we estimated the average treatment effects of these interventions on farm incomes. In this paper we apply the methodology described above to study distributive impacts.

in T villages. Hence we have

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

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 the social welfare in T villages can be written as

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

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

$$\begin{aligned} W(C) &= \Pr(s)E[U(y(s,0,C))] + [1-\Pr(s)]E[U(y(n,0,C))] \\ \implies W(C) &= E[U(y(\sigma,0,C))] \end{aligned} \tag{3}$$

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

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" $(n,0)$.[3] 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 is no control arm. 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:

$$y(n,0,T) = y(n,0,C), y(s,0,T) = y(s,0,C) \tag{4}$$

Observe that if C denotes the counterfactual that would have occurred if the T villages had not received the intervention, then equation (4) implies that the welfare impact of the intervention

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

equals

$$
\begin{aligned}
W(T) - W(C) &= p\Pr(s)\{E[U(y(s,1,T))] - E[U(y(s,0,C))]\} \\
&\quad + (1-p)\Pr(s)\{E[U(y(s,0,T))] - E[U(y(s,0,C))]\} \\
&= p\Pr(s)\{E[U(y(s,1,T))] - E[U(y(s,0,C))]\} \\
&= p\Pr(s)\{E[U(y(s,1,T))] - E[U(y(s,0,T))]\} \qquad (5)
\end{aligned}
$$

The first equality relies on the assumption of the absence of spillovers among the non-selected, while the second and third equalities relies 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 income of both the treated and untreated eligibles in the treated villages. The "no spillover" assumption implies that the income of the selected but untreated equals the income that the selected would have had, if the intervention had not been conducted. Hence the difference between incomes 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. These interventions 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. Only about 2.5 percent of the relevant population received the program, making it unlikely that there were spillover effects on the population that did not receive the program credit. We apply the methodology described above to assess the welfare impact of each intervention relative to a no-intervention counterfactual. 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 three agricultural credit interventions in the districts of Hugli and West Medinipur in the state of West Bengal, India.[4] The schemes were primarily designed to facilitate the cultivation of potatoes, the most profitable cash crop in this region. Loan size, duration, interest rate and dynamic repayment incentives were identical across the three interventions. 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

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

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.[5] 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 experiment was designed 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 is delegated to a local intermediary appointed as the MFI's agent. He/she is incentivized through commissions that depend on the interest payments made by the borrowers they recommend.

In the 24 randomly selected villages that were assigned to the TRAIL intervention, a local trader was appointed as the agent. The field research team randomly selected the agent from a list of local traders who had at least 50 clients, or had been operating in the village for longer than 3 years. Another 24 villages were assigned to the GRAIL intervention, where the agent was a political appointee. The field research team randomly selected one individual from a list recommended by the local government (gram panchayat or GP) of persons who had lived in the village for at least 3 years, were personally familiar with farmers in the village and had a good local reputation.[6] Each agent was asked to recommend 30 borrowers who owned no more than 1.5 acres of land. Ten of these 30 were randomly selected to receive the program loans. At the end of each loan cycle, the agents received a commission equal to 75 percent of the interest paid by borrowers they had recommended.

A third group of 24 villages was assigned to the Group Based Lending (or GBL) intervention. Village residents who owned no more than 1.5 acres of land could form 5-member groups, attend regular group meetings and make savings deposits. At the end of 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. The MFI that organized the group meetings received commissions equaling 75 percent of the interest paid by GBL borrowers.

To enable us to understand the underlying mechanisms, the experiment was designed to separately identify how selected borrowers differed from those not selected (selection effects), and to estimate 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. We refer to these as the TRAIL

[5]The program ran for three years in all.

[6]In all TRAIL villages, the first randomly chosen trader approached accepted the contract. In the GRAIL villages, one individual refused to participate for religious reasons; he was replaced by a second randomly chosen individual from the list.

Treatment and GRAIL Treatment households respectively. In the GBL arm, only 2 of the joint liability groups that had formed in the village were randomly selected, and each member offered the loans. The households of these group members are the GBL Treatment households. Of the 20 recommended TRAIL and GRAIL households that were not randomly assigned to receive the loans, 10 were randomly drawn into the survey sample; we refer to them as Control 1 households. In GBL villages, two of the groups that did not receive the loans were randomly chosen and all of their member households surveyed, these are the GBL Control 1 households.[7] 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

Table 1 shows village characteristics, computed using the 2011 Census of India and data from a 2007 pre-intervention village census conducted for a different project (see Mitra et al., 2018).[8] As the table shows, 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.

Over 2010 to 2013, we conducted eight waves of surveys with a sample of 50 households in each of the 72 villages where the loan schemes ran. In each village, the sample includes the 10 Treatment and 10 Control 1 households, as defined above.

In Table 2 we present summary statistics from the first wave of household survey data. These statistics pertain to the characteristics of 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=0.998).[9]

[7] In addition, 30 non-selected households were included in the sample. We refer to these as Control 2 households. In TRAIL and GRAIL villages, Control 2 households were randomly selected from those that were not recommended by the agent. In GBL villages they were randomly selected from households that did not participate in joint-liability groups. We do not include Control 2 households in our sample when we estimate the conditional treatment effects in the next section.

[8] 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 1 uses a sample of only 68 villages.

[9] 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 ensure that we 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. Thus we can scale up the sample proportions in each land category to arrive at the population proportions.

Table 2 also shows how the summary statistics vary by land ownership class. Looking across the columns of the table, it is clear that landholding is a good proxy for a household's socio-economic status. 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 labour).[10] Households with more land lived in larger houses that were more likely to be 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 and telephones, and to have bank savings accounts. We find very few land transactions across the three survey years, so that households' land class remains largely fixed over this short time horizon. Thus, it is informative to conduct sub-group analysis across different land categories.

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.

5 Empirical Estimates

5.1 Computing the CTEs

We start by estimating the conditional treatment effects. To do this, we restrict the sample to Treatment and Control 1 households in the TRAIL, GRAIL and GBL villages, and run the following regression:

$$\begin{aligned} U(y_{ivt}) = \beta_0 + \beta_1 \text{TRAIL}_v + \beta_2 \text{GRAIL}_v + \beta_3 \text{Treatment}_{iv} + \beta_4 (\text{TRAIL}_v \times \text{Treatment}_{iv}) \\ + \beta_5 (\text{GRAIL}_v \times \text{Treatment}_{iv}) + \xi \mathbf{X}_{ivt} + \epsilon_{ivt} \end{aligned} \tag{6}$$

Here $\mathbf{X}_{ivt}$ is a set of variables measuring household characteristics consisting of 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.[11]

The dependent variable in the regression is $U(y_{ivt}) = \frac{y_{ivt}^{1-\theta}}{1-\theta}$, corresponding to a specific value of θ. Here y_{ivt} is aggregate farm income for household i in village v in year t, calculated as the sum of value-added earned from the four major crop categories: potatoes, paddy, sesame and

[10] Note, however, that there is an active land rental market, so that even landless households do engage in agriculture.

[11] The information intervention was undertaken for a separate project aimed at 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 that is the focus of this paper. The results are unchanged if we do not include this dummy variable in the regression specification.

vegetables. The explanatory variable TRAIL_v indicates whether village v was assigned to the TRAIL intervention, GRAIL_v indicates whether it was assigned to the GRAIL intervention, and Treatment_{iv} indicates if the household was assigned to receive the loan. 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.[12]

In Panel A of Table 3, each column presents results from running this regression using a different inequality-aversion parameter θ ranging from 0 to 5. When θ takes value 0, the welfare impact represents the change in average farm income. In line with our previous result, we find in column 1 that the TRAIL scheme increased the average farm income of recommended households by ₹2546. This is significant at the 10% level. The point estimates for both the GRAIL (₹125) and the GBL (₹34) schemes are smaller in magnitude and not statistically different from zero.

At higher values of θ, the farm incomes of low income households receive greater weight in the welfare calculation. When $\theta = 1$ the welfare function is logarithmic, so that the conditional welfare impacts correspond to proportional changes in farm income. 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. As we see in Column 2, the TRAIL scheme generated a significant increase in welfare for selected households, even according to this metric. This pattern is repeated as we increase θ to higher values. This suggests that the TRAIL scheme benefited poorer households in particular.

The results in columns 2–6 also suggest that neither the GRAIL nor the GBL schemes changed welfare significantly: even when the welfare function is highly sensitive to inequality, the point estimates continue to be positive but are never statistically significant. Thus not only is the average effect of these schemes small, there is no evidence to suggest that lower income households benefited either.

5.2 Computing the Treatment Proportions

Recall that loans were offered to households that satisfied two conditions. First, the agent recommended households (in the TRAIL and GRAIL interventions), or the households formed a group with other village households (in the GBL intervention). The probability that a household would satisfy this condition is given by $\Pr(s)$. Next, from this set of households, the research team

[12] In order to estimate the conditional treatment effects (for values of $\theta = 1$ and above), we add 10,00,000 to the farm imputed profits. This helps to ensure that we can take the log (for $\theta = 1$), and for the other values of θ we are not working with very small numbers. The regressions give us point estimates and standard errors for this transformed variable. We transform them back to arrive at estimated treatment effects on welfare. The re-transformation introduces a stochastic element, making it difficult to analytically calculate the standard errors for the point estimates on welfare. Therefore, we present bootstrapped standard errors.

randomly picked the household to whom the loan was offered; for any household in this set, the likelihood of being picked this way is denoted probability p. Thus the treatment proportion is given by $p\Pr(s)$. Our calculations show that across the 24 TRAIL households, 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 2.3 percent. Similarly, across the 24 GRAIL households, 7 percent of our sample households were recommended, and 2.3 percent were offered loans. In GBL villages, 5.1 percent of our sample households formed groups, and 1.7 percent were offered loans.

5.3 Change in Aggregate Welfare as a Result of the Interventions

Panel B of Table 3 shows the implied change in aggregate welfare, calculated 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. It is clear once again that the TRAIL scheme increased aggregate welfare at all inequality-aversion levels. However, neither the GRAIL nor the GBL schemes had a significant effect on welfare.

5.4 Comparing the Welfare Impacts of the Interventions

In Panel C of Table 3, we compare the welfare impacts of the three interventions. We conduct pair-wise Mann-Whitney rank-sum tests on 2000 bootstrap replications of the aggregate welfare effects. 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 village were included. In each bootstrap sample we estimate the treatment proportion for each scheme, and the conditional treatment effects of each scheme. The product is the treatment effect on average welfare.

At $\theta = 0$, the null hypothesis that the GRAIL and TRAIL schemes generate identical aggregate welfare effects is rejected with a p-value < 0.00. Similarly, we can reject the null that the GRAIL scheme had a larger welfare effect at the 5% level when $\theta = 1$, and at the 10% level when $\theta = 2$ and $\theta = 3$. We can also reject the null that the GBL scheme generated a larger aggregate welfare effect than the TRAIL scheme for $\theta = 0, 1, 2$, and the p-values range from 0.14 to 0.20 for $\theta = 3, 4, 5$. We are never able to reject the null hypothesis that the GBL scheme generated a larger aggregate welfare effect than the GRAIL.

The cumulative distribution functions of these estimated changes in welfare for TRAIL and GBL are presented in Figure 1. These corroborate our findings from the rank-sum tests. For every value of θ, the aggregate welfare effects of the TRAIL scheme stochastically dominate those of the GRAIL and GBL schemes. However at low values of θ, there is no clear pattern of stochastic dominance between the GRAIL and GBL schemes. As the value of θ increases, the GBL scheme stochastically dominates the GRAIL scheme, suggesting that the GBL scheme was relatively better for low-income households than the GRAIL scheme.

6 Welfare Decomposition

We obtain further insight by decomposing the aggregate welfare effects by sub-groups. The decomposition is straightforward because the Atkinson welfare function is additively separable. Letting g denote the group, the average village welfare impact of the intervention can be written as a weighted average of the conditional treatment effects on welfare of the different groups:

$$W(T) - W(C) = \sum_g \alpha_g s_g \mathrm{CTE}_g \tag{7}$$

where α_g denotes the demographic weight of group g, and s_g denotes the fraction of group g that was treated. The conditional treatment effect CTE_g equals $\{E[U(y)|s,1,T,g]-E[U(y_i)|s,0,T,g]\}$. $E[U(y)|s,1,T,g]$ 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. The decomposition represented in equation (7) shows that the overall welfare impact can be expressed as the (population share) weighted average of the product of $\alpha_g \cdot s_g$ and the CTEs across different groups.

6.1 Empirical Decomposition Results by Landholding Groups

We decompose the aggregate welfare effects of each intervention in our experiment, using four different landholding groups. To estimate the conditional treatment effects (CTE_g) on welfare for each land group g, we run the following regression:

$$\begin{aligned} U(y_{igvt}) = & \sum_{g=1}^{G} \gamma_g(Z_{igv}) + \sum_{g=1}^{G} \delta_g(Z_{igv} \times \mathrm{TRAIL}_v) + \sum_{g=1}^{G} \zeta_g(Z_{igv} \times \mathrm{GRAIL}_v) + \sum_{g=1}^{G} \eta_g(Z_{igv} \times \mathrm{Treatment}_{igv}) \\ & + \sum_{g=1}^{G} \theta_g(Z_{igv} \times \mathrm{TRAIL}_v \times \mathrm{Treatment}_{igv}) + \sum_{g=1}^{G} \kappa_g(Z_{igv} \times \mathrm{GRAIL}_v \times \mathrm{Treatment}_{igv}) \\ & + \lambda \mathbf{X}_{ivt} + \epsilon_{igvt} \end{aligned} \tag{8}$$

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 $\delta_g + \theta_g$, $\zeta_g + \kappa_g$ and η_g respectively; $\mathbf{X}_{ivt}$ is as defined earlier.

The top panel of Figure 2 shows the values of s_g, or the fraction of group g households that received the treatment: these range from 1.1 to 3.2%, although they decline with landholding in the GBL treatment arm, are flat in the TRAIL am and increase in the GRAIL arm.

The middle panel of Figure 2 shows the estimated CTEs across the four groups for two values of $\theta = 0, 1$. For $\theta = 0$ these represent the conditional treatment impacts on the absolute level of income of each group, while for $\theta = 1$ they represent the corresponding proportional impacts. While the absolute income impacts of the TRAIL scheme were largest for the intermediate land groups, the corresponding *proportional* impact was the largest for the landless group. Similarly, the absolute difference in CTEs between the TRAIL scheme and the other interventions were larger for intermediate land groups, but the proportional difference was largest for the landless.

The bottom panel shows the corresponding treatment impacts across the different groups (i.e, the products $\alpha_g s_g \text{CTE}_g$). The patterns resemble the corresponding patterns in the CTEs. Hence the larger proportional impacts on the incomes of the landless seem to drive the superior aggregate welfare impact of the TRAIL scheme when $\theta = 1$. This effect is accentuated further as the welfare function takes on a larger inequality aversion parameter.

Note that in all three schemes, the evidence suggests that the landless benefited disproportionately more than the landed. Moreover, even among landless households, mean baseline income differed significantly across the three schemes. Landless households in the TRAIL villages who were selected but were not assigned the loan (Control 1 households) had an average farm income of only ₹362, whereas the corresponding landless households in GRAIL and GBL villages earned more than four times as much.[13] This helps explain why a relatively small ₹519 increase in farm income translated into a 142 percent increase for landless households in the TRAIL scheme, while a similar ₹679 increase translated into a much smaller 49 percent increase for landless households in the GBL scheme. In addition, even though there was no explicit gatekeeper in GBL preventing particular landless households from participating, it appears that landless households that earned very low incomes were either unable to or unwilling to form joint-liability groups.

[13] Thus even within the landless group, the TRAIL agent appears to have targeted poorer households than the GRAIL agent did.

7 Conclusion

Our approach has several advantages. One, its conceptual underpinnings provide a clear rationale for the welfare evaluations that it generates. Two, unlike other approaches it does not require strong assumptions for a clear-cut interpretation: the welfare effects are identified under the same assumptions by which a randomized controlled trial delivers consistent estimates of the average treatment effect. Three, the approach provides a single summary quantitative measure of the welfare impact for any given level of inequality aversion. Four, it is simple to implement and does not impose any additional data requirements beyond those used to estimate the standard average treatment effects. It also allows for decomposition analysis across population sub-groups, allowing a closer examination of the patterns that drive the aggregate effects.

References

Atkinson, A. B. (1970). On the Measurement of Inequality. *Journal of Economic Theory*, 2:244 – 263.

Bedoya, G., Bittarello, L., Davis, J., and Mittag, N. (2017). Distributional Impact Analysis: Toolkit and Illustrations of Impacts Beyond the Average Treatment Effect. Policy research working paper 8139, World Bank.

Coady, D. and Harris, R. (2004). Evaluating Transfer Programs Within a General Equilibrium Framework. *Economic Journal*, 114:778 – 799.

Hughes, G. (1987). The Incidence of Fuel Taxes: A Comparative Study of Three Countries. In Newbery, D. and Stern, N., editors, *The Theory of Taxation for Developing Countries*. Oxford University Press.

Maitra, P., Mitra, S., Mookherjee, D., Motta, A., and Visaria, S. (2017). Financing Smallholder Agriculture: An Experiment with Agent-Intermediated Microloans in India. *Journal of Development Economics*, 127:306 – 337.

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

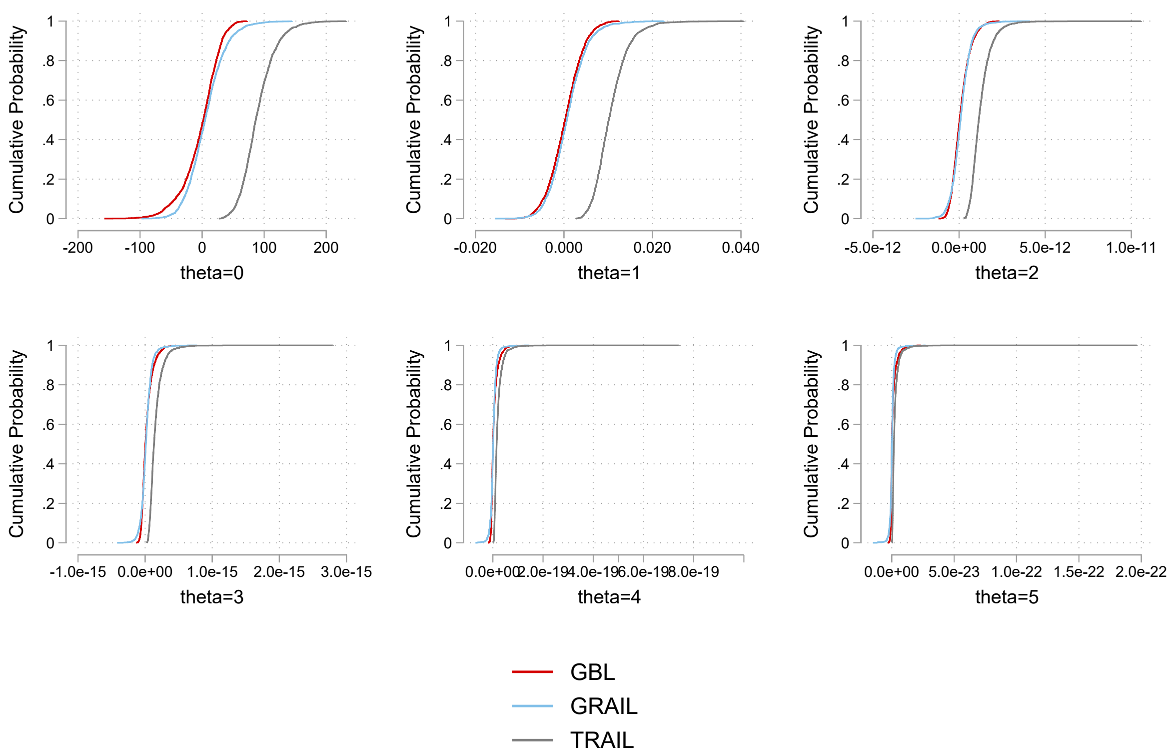
Mitra, S., Mookherjee, D., Torero, M., and Visaria, S. (2018). Asymmetric Information and Middleman Margins: An Experiment with West Bengal Potato Farmers. *Review of Economics and Statistics*, 100(1).

Newbery, D. and Stern, N., editors (1987). *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. *Economic Journal*, 105(431):847 – 863.

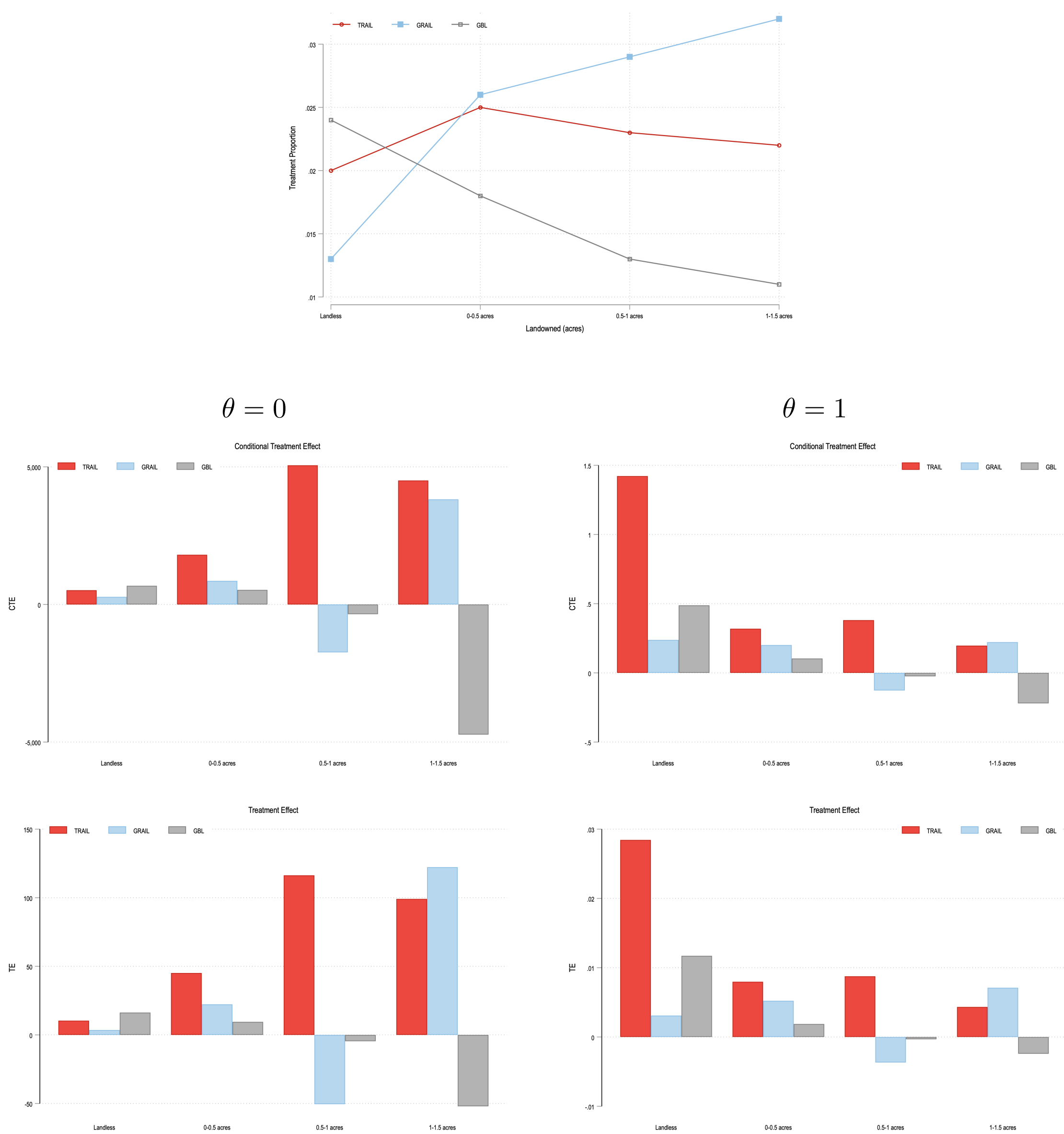
Roberts, K. (1980). Interpersonal Comparability and Social Choice Theory. *Review of Economic Studies*, 47(2):421–439.

Figure 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.

Figure 2: Proportions Treated and Treatment Effects, by Intervention and Land Class



Notes: The top panel shows the values of s_g, or the fraction of group g households that received the treatment. The middle panel shows the estimated CTEs across the four groups for two values of $\theta = 0, 1$. The bottom panel shows the corresponding treatment impacts across the different groups (i.e, the products $\alpha_g s_g \text{CTE}_g$).

Table 1: **Balance of Village-level Characteristics, by Village Treatment Arm**

	All (1)	TRAIL (2)	GRAIL (3)	GBL (4)	Differences: Two-way comparisons (2)-(3)	(2)-(4)	(3)-(4)
Number of households	365.32	327.63	310.71	457.58	16.92	-129.96	-146.88
	(40.66)	(52.28)	(64.87)	(88.35)	(83.32)	(102.66)	(109.61)
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 centre	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 *pucca* 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)
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.

Table 2: Household Characteristics, by Village Treatment Arm and Land Category

	All (1)	TRAIL (2)	GRAIL (3)	GBL (4)	Landless (5)	0–0.5 acres (6)	0.5–1 acre (7)	1–1.5 acres (8)
Head's education: primary or more	0.420	0.407	0.420	0.433	0.234	0.356	0.564	0.650
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.03)
Head's occupation: cultivation	0.431	0.441	0.415	0.437	0.056	0.381	0.689	0.696
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.02)	(0.02)
Head's occupation: Labor	0.335	0.340	0.343	0.323	0.677	0.404	0.089	0.041
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)
Non-Hindu	0.172	0.21	0.151	0.155	0.188	0.146	0.195	0.181
	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.04)	(0.02)
Low caste	0.387	0.383	0.355	0.423	0.565	0.417	0.286	0.199
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	0.023
Area of house and homestead (acres)	0.053	0.052	0.052	0.054	0.037	0.048	0.063	0.074
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Pucca House	0.294	0.288	0.334	0.259	0.207	0.280	0.344	0.379
	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.02)	(0.03)
Electrified house	0.751	0.740	0.752	0.760	0.666	0.729	0.811	0.841
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)
Separate toilet in house	0.575	0.564	0.608	0.552	0.434	0.541	0.664	0.741
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Owns radio/ TV/ VCR/ DVD	0.464	0.450	0.486	0.456	0.350	0.420	0.541	0.639
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.03)
Owns telephone (mobile or landline)	0.590	0.573	0.590	0.607	0.446	0.528	0.706	0.796
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Has savings bank account	0.456	0.447	0.475	0.446	0.268	0.410	0.576	0.680
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Test of joint significance for assignment to treatment								
Chi-squared statistic			10.32		32.64	28.30	24.09	32.27
p-value			*0.993*		*0.112*	*0.248*	*0.456*	*0.121*

Notes: Household characteristics 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.

Table 3: Estimated Welfare Impacts of the Loan Interventions

Value of θ	$\theta = 0$	$\theta = 1$	$\theta = 2$	$\theta = 3$	$\theta = 4$	$\theta = 5$
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Conditional treatment effects on household welfare						
TRAIL	2456.36	0.283	3.205E-11	3.699E-15	4.271E-19	4.934E-23
Confidence interval	(1493.55, 3399.04)	(0.134, 0.401)	(4.52E-12, 4.75E-11)	(-1.02E-15, 5.81E-15)	(-3.78E-19, 7.01E-19)	(-8.55E-13, 8.39E-23)
GRAIL	125.83	0.017	2.205E-12	2.942E-16	3.921E-20	5.221E-24
Confidence interval	(-1099.77, 1298.32)	(-0.146, 0.180)	(-1.97E-11, 2.44E-11)	(-2.81E-15, 3.45E-15)	(-4.17E-19, 5.01E-19)	(-6.65E-23, 7.23E-23)
GBL	33.5	0.0048	7.146E-13	1.012E-16	1.428E-20	2.010E-24
Confidence interval	(-1488.72, 1988.19)	(-0.251, 0.206)	(-4.35E-11, 2.14E-11)	(-7.67E-15, 2.44E-15)	(-1.34E-18, 2.89E-19)	(-2.31E-22, 3.61E-23)
Panel B: Change in aggregate welfare (using treatment proportion)						
TRAIL	56.22	0.006	7.336E-13	8.467E-17	9.776E-21	1.129E-24
	(23.71, 79.64)	*(0.002, 0.009)*	*(-3.638E-14, 1.120E-12)*	*(-4.522E-17, 1.354E-16)*	*(-1.192E-20, 1.624E-20)*	*(-2.328E-24, 1.946E-24)*
GRAIL	2.946	0.000	5.163E-14	6.889E-18	9.182E-22	1.222E-25
	(-30.89, 30.29)	*(-0.004, 0.004)*	*(-5.47E-13, 5.81E-13)*	*(-7.98E-17, 8.29E-17)*	*(-1.18E-20, 1.18E-20)*	*(-1.81E-24, 1.74E-24)*
GBL	0.569	8.00E-05	1.214E-14	1.719E-18	2.427E-22	3.417E-26
	(-24.69, 38.29)	*(-0.004, 0.004)*	*(-6.81E-13, 4.10E-13)*	*(-1.18E-16, 4.58E-17)*	*(-2.10E-20, 5.46E-21)*	*(-3.68E-24, 6.82E-25)*
Panel C: Comparison of welfare effects						
Pr(GRAIL = TRAIL) : $p - value$	0.000	0.000	0.000	0.000	0.000	0.000
Pr(GRAIL > TRAIL) : $p - value$	*0.021*	*0.034*	*0.057*	*0.088*	*0.122*	*0.155*
Pr(GBL = TRAIL) : $p - value$	0.000	0.000	0.000	0.000	0.000	0.000
Pr(GBL > TRAIL) : $p - value$	*0.003*	*0.020*	*0.062*	*0.114*	*0.162*	*0.203*
Pr(GBL = GRAIL) : $p - value$	0.000	0.000	0.181	0.569	0.034	0.001
Pr(GBL > GRAIL) : $p - value$	*0.447*	*0.467*	*0.488*	*0.505*	*0.519*	*0.531*

Notes: Welfare impacts are estimated following the procedure outlined in Section 2, where θ indicates the value of the inequality-aversion parameter in the individual utility function that is the building block of the social welfare function. The terms in parentheses denote the bootstrapped 90% confidence interval (with 2000 replications). Panel C presents the results of Mann-Whitney rank-sum tests using 2000 bootstrap replications, comparing the aggregate welfare effects of the schemes. The p-value is for a test that the two unconditional treatment effects come from the same population. $Pr(X > Y)$ indicates an estimate of the probability that a random bootstrapped estimate of the unconditional treatment effect of scheme X is larger than a randomly drawn bootstrapped estimate of the unconditional treatment effect of scheme Y, where $X; Y \in$ {TRAIL, GRAIL, GBL}

Data Availability Statement

The data underlying this article will be made available in online supplementary material.