Pre-anlysis plan:
Experimental evaluation of a Basic Income Pilot in Germany

Sandra Bohmann Susann Fiedler Maximilian Kasy Jürgen Schupp
Frederik Schwerter

May 27, 2021

1 Introduction

Basic income A policy proposal that has received much attention in recent years is the introduction of a (universal) basic income. Relative to existing forms of social protection and redistribution, the distinguishing feature of a basic income is its lack of conditionality: Income receipt is not contingent on requirements such as a prior work history, demonstrated willingness to work, sickness, or old age. This lack of conditionality has implications for various dimensions of welfare and behaviour, mediated by channels including the bargaining power of (potential) recipients, increased income security, and increased coverage of populations who are otherwise excluded.

In this study, we will evaluate a basic income pilot program in Germany using a blocked randomized controlled trial design. Our evaluation will focus on dimensions of behavior and welfare which might be distinctively impacted by a basic income, relative to other forms of social protection. We will consider, in particular, labor market outcomes, expenditures, time use, and subjective wellbeing.

Program description We study the effect of a basic income on recipients in a randomized controlled trial (RCT) that will last for 3 years. The RCT has 1,484 participants, of which 107 are assigned to the treatment group, while 1,377 are assigned to the control group. Members of the treatment group receive unconditional cash transfers of 1,200.00 EUR, paid monthly, over the course of three years. Members of the control group do not receive unconditional cash transfers, but instead receive incentives to participate in research-related activities. Participants in the treatment group are required to participate in at least seven online surveys, one before the start of treatment and, thereafter, every 6 months until the end of the trial. Participants in the control group receive (at least) 10 EUR for each completed survey, plus (at least) 30 EUR if all surveys were completed.

Implementation partner The study is financed by the German NGO "Mein Grundeinkommen e.V." (MG henceforth). MG is financed through (mostly small) private donations. MG has previously used their donations to finance smaller basic income packages of 1,000 EUR per month for a single year, paid to randomly assigned people from a pool of people who signed up for them. Between 2014 and 2021, MG has financed 818 of these smaller basic income packages.

2 Sampling and treatment assignment

In the following we describe in detail the multi-step sampling and treatment assignment procedure used to construct our study sample. The steps in this procedure are (i) a public call and voluntary registration of potential participants, (ii) selection of a subsample based on demographic and economic eligibility criteria, (iii) stratified sampling of eligible registrants to construct a representative baseline sample, members of which were then invited to fill out a longer baseline survey, (iv) blocking of participants in the baseline sample who have a completed survey, based on a rich set of baseline covariates, and random assignment to treatment within blocks, and (v) selection of a representative subsample of blocks based on the budget constraints of the study.

2.1 Sampling

Signup call and registrations In August 18, 2020, MG and the German Institute for Economic Research (DIW Berlin) publicly announced the launch of the RCT during Spring/Sommer 2021 and made a public call to register to participate in the RCT. The announcement included a description of the main features of the study: Selected participants of the study would be randomly assigned to a treatment group or a control group; treatment and control groups would participate in biannual online surveys; members of the treatment group would receive monthly payments of 1,200.00 EUR for three years; members of the control group would receive monetary incentives to complete the surveys; additional research activities may be offered. During signup, we collected the following demographic and socioeconomic information: Age, gender, education, monthly net income, number of people living in their household, number of kids, zip code, and their general attitude towards universal basic income. Between August 18 and December 10 in 2020, 2,048,370 potential participants registered in response to this public signup call.

Eligibility criteria We then invited a subsample of registered individuals (called "baseline sample") to complete the baseline survey. Selection into the baseline sample is based on the following eligibility criteria with respect to participants' demographic and socioeconomic characteristics. These eligibility criteria were largely determined by our implementation partner, MG.

1. Participants have to be between 21 and 40 years old.
2. Households of size greater than one, and individuals with dependent children, are excluded from participation.

   Participants of our study whose household size changes, or who have a child, will, however, not lose their participation status.
3. Participants are required to be German residents and to have a monthly net income between 1,100.00 and 2,600.00 EUR.
4. Individuals who, at the time of the baseline survey, were receiving social benefits for long term unemployment are excluded from participation.[1]

   Participants of our study who transition to unemployment and receipt of social transfers will, however, not loose their participation status.

Baseline sample Among the potential participants who satisfied these criteria, our implementation partner next sampled 20,000 individuals who were invited to participate in a baseline survey. Sampling of these individuals was based on the following criteria. First, the sample was supposed to contain an equal number of proponents and opponents of a universal basic income. Second, potential participants in both of these groups were sampled using a weighted sampling procedure to generate a sample that is close to being representative for the (eligible) German population, and similar across both groups, in terms of age, gender, income, education, employment status, and state ("Bundesland").[2]

Baseline survey Before the invitations to the baseline survey were sent out, one person requested to be excluded from the RCT. The baseline survey resulted in 14,420 completed surveys. Of the remaining invitations,

- 51 invitations were sent to recipients with multiple registrations These participants were in turn excluded since potential participants were allowed to register only once.
- 3,359 invitations were sent to recipients who subsequently never started the baseline survey.

[1] Given current benefit eligibility rules, such social benefits would have been cut by up to the full amount of the cash transfer by MG, if these individuals were to participate in our study. The net transfer to such individuals would thus have been significantly below the expenditure for MG.

[2] The exact sampling procedure is unknown to us. This does not affect, however, the internal validity or correctness of inference for the study design described below.

- 328 invitations were sent to recipients who then started but did not complete the baseline survey.
- 1,841 recipients completed the survey, but did not sign the required data sharing consent forms.

Amongst the 14,420 individuals who completed the baseline survey and gave consent, 8,971 participants are considered in the randomized block assignment discussed next. The remaining 5,449 individuals are dropped because their eligibility status with respect to their characteristics listed above in criteria 1-4 changed and/or they had missing responses in baseline variables that were used in the randomized block assignment.[3]

2.2 Blocking and treatment assignment

Blocking We use the answers to the baseline survey to sort participants into homogenous blocks. The 28 variables used are listed in column 1 of Table 1. Appendix A provides further details on the definition of these variables.

Pairwise distances between observations are calculated using the Mahalanobis distance.[4] We construct blocks containing 32 observations each. The blocks are chosen to minimize the total sum of distances between pairs of observations within blocks. We do so using the R package *blockTools* (Moore and Schnakenberg, 2016). We then discard all blocks with a maximum within-block distance greater than 14 (to avoid poorly matched observations), as well as one block with less than 32 observations.

Random assignment within blocks Within each block, treatment is assigned uniformly at random. We assign 2 out of the 32 observations in a block to the treatment group, 26 observations to the control group, and the remaining 4 observations to a "reserve," which is to be sampled in case of attrition of observations from the treatment or control group.

These numbers are chosen based on the following considerations: We want two treated units per block, in order to be able to calculate standard errors for the sample average treatment effect; cf. Athey and Imbens (2017) and our discussion of inference below. We don't want more treated units per block, to keep blocks as homogenous as possible. The budget constraints of our implementation partner are furthermore such that we can survey 13 control units for every treated individual.

Lastly, because we have 107 treated individuals in total (an odd number), one additional individual from one block is chosen at random to participate in the treatment.

Weighted sampling of blocks This procedure results in 273 blocks, while our project budget allows for 53 blocks. These blocks are furthermore not fully representative for the baseline sample, because not all individuals who were invited to participate in the baseline survey passed eligibility and had non-missing responses in the questions we used for blocking (see above) and because of our discarding of poorly matched blocks.

In order to obtain a representative sample of blocks, we create block level sampling weights. These weights are chosen so as to match the distribution of gender, education groups, and income groups of eligible participants in the screening survey. We then draw a sample of 53 blocks from the 273 available blocks using these sampling weights, to obtain a representative subsample.

This results in 107 individuals assigned to treatment, 1377 assigned to the control group, and 212 individuals assigned to the "reserve," distributed evenly across 53 blocks.

[3] Additionally, our implementation partner selected a group of 15 individuals who will be treated (that is, who will receive the basic income). These additional individuals indicated in the baseline survey that they were willing to participate in qualitative surveys (which are not conducted by the authors of this preregistration and are not part of this preregistration) and in interviews with journalists to publicly share their own experiences with the basic income *during* the RCT. Since any public appearance of these participants *may* bias their responses in our online surveys, we exclude these "media participants" from our study.

[4] The Mahalanobis distance of two covariate vectors x_1 and x_2 that are realizations of a random vector X is given by $d(x_1, x_2) = \sqrt{(x_1 - x_2) \cdot Var(X)^{-1} \cdot (x_1 - x_2)}$.

The resulting treatment assignment Table 1 summarizes the resulting study sample. The second and third columns show covariate averages for the 28 covariates used for blocking, for the treated and control group. This table drops observations in the reserve. The remaining columns show standard errors, confidence intervals, and p-values as discussed below.[5] As can be seen from this table, we have achieved an extraordinary degree of balance between the treated and control group.

At this point, the selected participants in the treatment group and control group were informed about their treatment status. 7 individuals in the control group wanted to be excluded from the study sample, 1 individual in the treatment group resigned his/her spot in the treatment group because of a job opportunity outside of Germany, and 1 individual in the treatment group could not be reached. For each of these missing individuals, we sampled one individual from the replacement sample within the same block, to receive the corresponding treatment status.

Hair analysis After participants were informed about their treatment status, participants in the treatment and control groups were asked whether they would like to participate in another research activity, hair analyses. Participation in the hair analyses requires them to send three samples of their hair (before the basic income treatment starts, 1.5 years after treatment begins, and when the treatment is completed in 3 years) to a research lab that measures only the level of cortisol (a proxy for stress) in their hair. For each correctly submitted hair sample, participants receive a compensation of 50 EUR.

Among the participants from the treatment group, 71 participants indicated to be willing to participate in the hair analyses and were selected to participate in the hair analyses. For each block with treated participants included in the hair analyses, twice as many control participants as treated participants willing to participate in the hair analyses were randomly selected. This resulted in 142 control participants that were selected to be included in the hair analyses. Since financial resources for an additional number of 24 hair analyses were available, 24 additional control participants willing to be included in the hair analyses were randomly selected amongst those blocks with hair-analyses participants.

[5]Inference should not be taken literally here, and is only including for illustration. In particular, because of our blocked assignment procedure, which aims for balance, p-values are expected to be systematically larger than suggested by the uniform distribution under the "null" of no effect.

Table 1: Balance of baseline covariates in the study sample

Outcome	Treated	Control	Difference	SE	t-statistic	p-value
Age 29-32	0.355	0.331	0.024	0.048	0.498	0.619
Age 33-40	0.336	0.373	-0.036	0.048	-0.757	0.449
Female	0.477	0.412	0.065	0.050	1.290	0.197
German citizen	0.981	0.981	0.000	0.014	0.014	0.989
UBI proponent	0.505	0.547	-0.042	0.050	-0.837	0.403
Tenure	0.766	0.766	0.000	0.043	0.005	0.996
Education: Hauptschule	0.037	0.038	0.000	0.019	-0.020	0.984
Education: Realschule	0.215	0.214	0.001	0.041	0.035	0.972
Education: Fachabitur	0.243	0.241	0.002	0.043	0.044	0.965
Education: Abitur	0.037	0.054	-0.016	0.019	-0.843	0.399
Net monthly income	1944.888	1925.767	19.121	40.181	0.476	0.634
Monthly saving	271.607	296.407	-24.800	24.742	-1.002	0.316
Wealth	25327.103	25392.157	-65.054	4450.093	-0.015	0.988
Debt	10170.374	9077.122	1093.252	2655.173	0.412	0.681
High financial security	0.327	0.312	0.016	0.047	0.329	0.742
Working for money	0.935	0.944	-0.010	0.025	-0.383	0.702
In training or education	0.178	0.151	0.027	0.038	0.691	0.489
In vocational training	0.411	0.432	-0.021	0.050	-0.421	0.674
Searching work	0.037	0.038	0.000	0.019	-0.020	0.984
Sick days	7.776	10.850	-3.075	1.152	-2.669	0.008
Weekly hours worked	37.826	37.346	0.480	1.458	0.329	0.742
Political preferences (PC1)	0.015	0.142	-0.127	0.142	-0.893	0.372
Political preferences (PC2)	0.164	0.053	0.112	0.125	0.893	0.372
Subjective wellbeing (PC1)	-0.360	-0.129	-0.231	0.183	-1.263	0.207
Body mass index	24.656	25.452	-0.797	0.490	-1.627	0.104
Transfers to others	363.551	330.733	32.819	103.753	0.316	0.752
Donations in 2020	100.664	96.562	4.101	21.002	0.195	0.845
Binary gender	1.000	1.000	0.000	0.000	–	–

Notes: This table shows averages of baseline covariates for the treated and control group in our study sample, as well as their difference. The table additionally shows "naive" standard errors (ignoring blocked assignment), as well as the corresponding t-statistic and p-value. As this table shows, we were able to achieve a very high degree of balance for almost all variables.

3 Inference

We denote individual treatment assignment by D and outcomes by Y. Throughout, we consider our primary object of interest to be the sample average treatment effect

$$\Delta = \sum_i (Y_i^1 - Y_i^0), \tag{1}$$

for various individual-level outcomes Y_i for individuals i and corresponding potential outcomes Y_i^0, Y_i^1.

Our primary estimator will be based on block-level differences in mean outcomes, averaged across blocks b:

$$\begin{aligned} \bar{Y}_b^1 &= \frac{1}{n_b^1} \sum_{i:\ b_i=b} D_i Y_i & \bar{Y}_b^0 &= \frac{1}{n_b^0} \sum_{i:\ b_i=b} (1-D_i) Y_i \\ \hat{\Delta}_b &= \bar{Y}_b^1 - \bar{Y}_b^0 & \hat{\Delta} &= \frac{1}{N} \sum \hat{\Delta}_b, \end{aligned} \tag{2}$$

where n_b^1 and n_b^0 are the number of treated and untreated individuals in block b, and N is the number of blocks. For robustness checks, we will consider alternative estimators using regressions controlling for baseline covariates.

Inference will be based on two alternative methods, both of which yield valid inference for the sample average treatment effect: Standard errors and confidence intervals based on a normal approximation, and randomization inference.

Standard errors To calculate a standard error for $\hat{\Delta}$ as an estimator of Δ, we calculate block-level standard-errors (allowing for arbitrary heteroskedasticity), and aggregate:

$$\begin{aligned} \hat{\sigma}_b^{21} &= \frac{1}{n_b^1 - 1} \sum_{i:\ b_i=b} D_i \cdot (Y_i - \bar{Y}_b^d)^2 & \hat{\sigma}_b^{20} &= \frac{1}{n_b^0 - 1} \sum_{i:\ b_i=b} (1-D_i) \cdot (Y_i - \bar{Y}_b^d)^2 \\ \hat{\sigma}_b^2 &= \frac{1}{n_b^1} \hat{\sigma}_b^{21} + \frac{1}{n_b^0} \hat{\sigma}_b^{20} & \hat{\sigma}^2 &= \frac{1}{N} \sum_b \hat{\sigma}_b^2. \end{aligned} \tag{3}$$

95% confidence intervals for Δ are then calculated as

$$CI = [\hat{\Delta} - 1.96 \cdot \hat{\sigma}^2, \hat{\Delta} + 1.96 \cdot \hat{\sigma}^2]. \tag{4}$$

Neyman p-values are similarly based on these standard errors and the normal approximation for the distribution of $\hat{\Delta}$.

Randomization inference Our second approach toward inference is based on permutations of treatments, that is, based on randomization inference. This approach allows us to test the null hypothesis that the intervention had no effect of any kind, that is, $Y_i^1 = Y_i^0$ for all individuals i and potential outcomes Y_i^1, Y_i^0.

We re-assign treatment at random *within* each of the blocks b. For this counterfactual treatment assignment, we re-calculate any given test-statistic. Repeating this process many times, we can calculate the share of re-assignments for which the test-statistic is bigger than the realized value of the test-statistic. This share is the p-value for the null hypothesis of no effects.

Compound hypotheses In order to deal with the issue of multiple testing in a principled manner, we will use the Benjamini–Hochberg procedure, which allows us to control the false discovery rate, that is, the share of rejected hypotheses which in fact hold true.

This procedure works as follows. Sort the p-values for each of the m hypotheses tested by size, resulting in ordered values $P_{(j)}$. For a critical value α, find the largest value k such that

$$P_{(k)} \leq \frac{k}{m} \alpha.$$

Reject the null hypothesis for all $i = 1, \ldots, k$.

Groups of outcome variables We will apply this procedure separately for each group of outcomes. The exact list of outcomes remains to be determined, but likely include will include the following:

- Expenditures and household finance
- Administrative data on labor supply
- Survey responses related to labor supply
- Psychological concepts related to autonomy
- Subjective well-being
- Physical well-being
- Political preferences
- Economic preferences
- Social attitudes

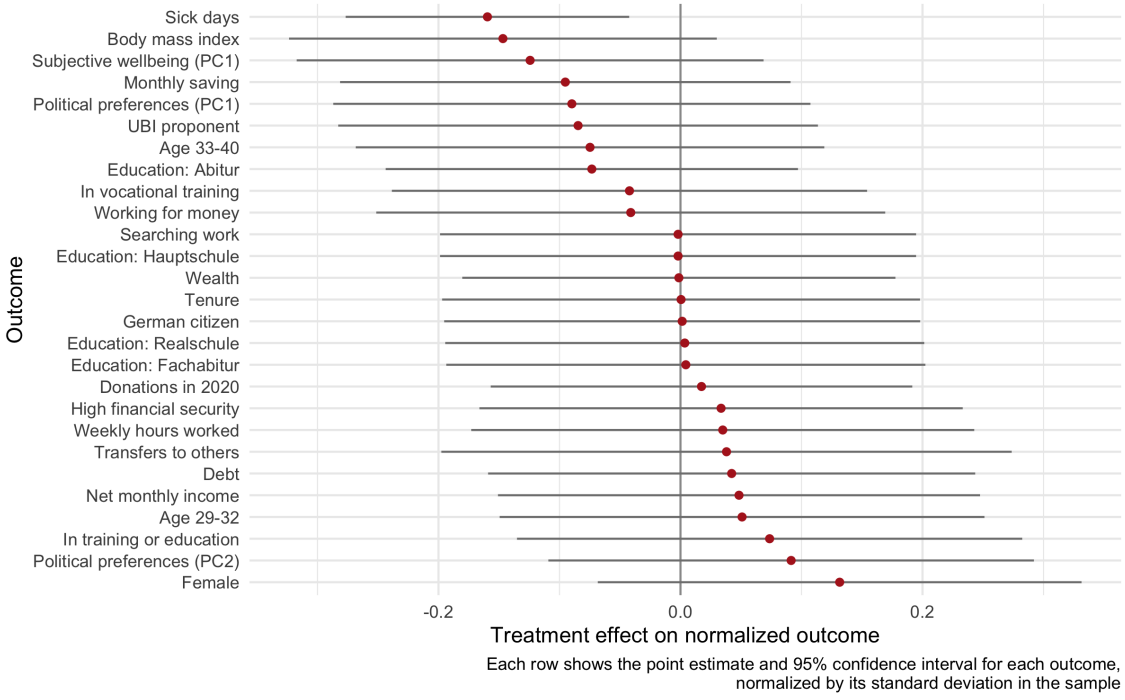
A list of outcomes, and of survey instruments to measure them, will be uploaded separately before the start of data collection.

A Description of baseline variables used for treatment assignment

- Age 29-32: a dummy variable =1 if individuals' age is between 29 and 32 years, =0 if individuals' age is below 29 or above 32 years.
- Age 33-40: a dummy variable =1 if individuals' age is between 33 and 40 years, =0 if individuals' age is below 32 years.
- Female: a dummy variable =1 if individuals' gender is female, =0 if individuals' gender is not female.
- German citizen: a dummy variable = 1 if individual is a german citizen, =0 if not =1
- UBI proponent: a dummy variable = 1 if individuals' general attitude towards universal basic income is positive, =0 if it is negative.
- Tenure: a dummy variable =1 if the individual has (at least one) tenured job, =0 if the individual has no tenured job.
- Education: Hauptschule: a dummy variable = 1 if highest education level qualifies for vocational training, =0 if not =1.
- Education: Realschule: a dummy variable = 1 if highest education level qualifies for high school, =0 if not =1.
- Education: Fachabitur: a dummy variable = 1 if highest education level qualifies for vocational academy, =0 if not =1.
- Education: Abitur: a dummy variable = 1 if highest education level qualifies for university, =0 if not =1.
  (Note that the omitted education category is college or more.)
- Net monthly income: net monthly income available to the individual.
- Monthly saving: amount of money saved per month.
- Wealth: individuals' level of wealth.
- Debt: individuals' level of debt.
- High financial security: a dummy variable = 1 if individual states that she could finance herself (with help of others but absent social security benefits) for one year without receiving any income, =0 if not =1.
- Working for money: a dummy variable = 1 if individual works and receives a financial compensation in return, =0 if not =1.
- In training or education: a dummy variable = 1 if individual is in vocational training or receives higher education (undergraduate, graduate, or doctoral level), =0 if not =1.
- In vocational training: a dummy variable = 1 if individual is in vocational training, =0 if not =1.
- Searching work: a dummy variable = 1 if looking for a job, =0 if not looking for a job.
- Sick days: number of workdays missed because of health.
- Weekly hours worked: number of hours worked per week
- Political preferences (PC1): first component of a principle component analysis that is based on an individual's response to how likely (in percent) it is that they vote for either party currently in the German parliament.

- Political preferences (PC2): second component of a principle component analysis that is based on an individual's response to how likely (in percent) it is that they vote for either party currently in the German parliament.
- Subjective wellbeing (PC1): first component of a principle component analysis that is based on an individual's responses to questions related to several dimensions of their subjective well-being (life satisfaction, emotional wellbeing, depression, eudaimonie, and subjective health).
- Body mass index.
- Transfers to others: how much money did the individual give to family members or friends (or others) in 2020.
- Donation in 2020: how much money was donated in 2020.
- Binary gender: a dummy variable =1 if binary gender, =0 if not =1

Figure 1: Naive confidence intervals for treatment effect on baseline covariates



References

Athey, S. and Imbens, G. W. (2017). The econometrics of randomized experiments. In *Handbook of Economic Field Experiments*, volume 1, pages 73–140. Elsevier.

Moore, R. T. and Schnakenberg, K. (2016). blocktools: Blocking, assignment, and diagnosing interference in randomized experiments. R package version 0.6-3.