The Experiment Design and Sample Selection

Tuna Dökmeci

November 19, 2019

Towards the end of 60s and in 70s, there was a heightened interest in assessing the impact of a guaranteed annual income (GAI) or a negative income tax (NIT) in the United States and in Canada, motivated by high levels of poverty. In 1964, President Lyndon Johnson called for a war against poverty in the United States. The same year, the Congress established the Office of Economic Opportunity to design programs to combat poverty and inequality of opportunities. In 1968, the Economic Council of Canada had declared that poverty among Canadians was widespread beyond belief (Hum and Simpson 1991). Much of the debate concentrated on the effect of such programs on the labor market and on the work incentives of such an income transfer.

These policy discussions and concerns over work behavior motivated experimental research in both countries. In the United States, the first such experiment was conducted in New Jersey in 1967 (Kershaw and Skidmore, 1973). Similar programs followed in Seattle, Washington; Denver, Colorado and Gary, Indiana. In Canada, a negative income tax program was launched in 1974 in Manitoba to be continued until 1979 with the main purpose of estimating the labor supply effects of this program. These experiments were all conducted with the explicit aim of informing policy discussions on transfer schemes (Hum and Simpson 1991). The experimental data from Manitoba and similar programs in the United States are unique in the sense that they are collected during the first and last such randomized controlled trials conducted in an industrialized country.

The provincial government of Manitoba in Canada submitted a proposal on the project mainly with the intention to understand the effect of a GAI on labor supply as well as the feasibility of this transfer scheme (Farthing 1992). The Manitoba Basic Annual Income Experiment (Mincome) was conducted in this context between 1974 and 1979 under the joint sponsorship of the Manitoba government and the federal government of Canada.

The design of the treatment groups was based on a combination of three different levels of guaranteed income payments and three tax levels. The annual support level represented what the household would receive if they had no other source of income or wealth. This amount was

calculated for a four-person household and was then adapted for other households. The assigned tax rate determined the rate at which the income support declined with this an increase in income or net worth. For instance, if a household was in a treatment group with 50% tax rate, every dollar the household made would reduce the payment by 50 cents. The final payment of the household i in treatment group k was given by $P_{ik} = G_k - t_k * Y_i - r * W_i$ where G_k is the support level, t_k the group-specific offset tax rate, Y_i monthly earnings of the household, W_i household wealth, and r the tax rate on wealth, which was fixed for all groups (Hum, Laub and Powell, 2016: 1). The overview of different treatment groups, or plans can be seen in the following table:

Table 1: Treatment Plans

GAI Level / Tax Level	$3800	$4800	$5480
%35	Plan 1	Plan 2	-
%50	Plan 3	Plan 4	Plan 5
%75	Plan 6	Plan 7	Plan 8

As mentioned above, the main purpose of the Mincome experiment was to provide evidence for the implementation and design of a negative income tax policy. The treatment with the highest income level and lowest tax rate would be more generous than any feasible negative income tax scheme that the Canadian government would implement. This treatment plan was thus considered to be irrelevant for the main purpose of the experiment and was eliminated from the start. The least generous plan, Plan 6, was merged with Plan 7 after a while due to the high rate of dropping out (User Manual).

A few words should be said on the sample selection. Ideally, both the selection of the sample and the allocation to treatment and control groups should be completely random, with each group composed of more or less the same number of units, in this case households.

The sample selection for Mincome consisted of several steps. The experiment was launched on three sites: the urban dispersed site in the provincial capital city Winnipeg, a rural dispersed sample composed of households in 18 rural communities in Manitoba, and a saturation site in Dauphin. The dispersed samples composed of randomly selected families in Winnipeg and rural areas who were then screened for eligibility. Eligibility was based on the income and assets of the household, and only households under a certain threshold could participate in the program. In Dauphin on the other hand, the whole population was eligible to receive the payments. While enrolling in the program was possible for all in Dauphin, the payments were down to zero if the household had more earnings than a certain amount that month. One of the reasons why the researchers wanted to have two dispersed and one saturation site was to estimate the community

effects on labor supply when the whole population can receive a guaranteed income.As mentioned above, all Dauphin residents were automatically eligible. In the other two sites, in Winnipeg and in rural areas, a large sample of 24,000 households was randomly chosen from the population. As the policy targeted low-income families, these households were then interviewed and screened for eligibility for the first time. Out of 6,372 families found eligible, baseline survey could be completed for 3,819 in order to do a second screening. Out of those found eligible after the baseline survey, 1,862 households enrolled in the program by responding to the second survey. While the treated households in the rural areas and in Dauphin were all assigned to Plan 3, there were initially eight treatment and one control group for the urban site in Winnipeg.

In Winnipeg, the allocation to different treatment groups was realized using the Conlisk-Watts model, following the example of similar NIT experiments in the United States in Seattle and Denver (Keeley and Robbins, 1980). The reason was that, in the context of a policy experiment with limited budget, several trade-offs were considered. First of all, the cost of assigning a household to a specific treatment group depended on the family size and income, meaning that some units were more expensive than the others. Minimizing this cost per unit leads to a larger sample size by giving the possibility to make payments to more households. The second trade-off stemmed from the heterogeneity of treatment plans in terms of their policy relevance. The policy makers had an interest in having more reliable samples for the more "relevant" treatment plans. Some treatment plans were more relevant than the others in the sense that they were more likely to be put in place as a policy country-wide and it was thus more important to have precise estimates of the effect of these programs on work behavior.

The trade-off between randomization and cost and policy considerations was tackled with the Conlisk-Watts assignment model, which optimizes the experimental design given certain constraints. The sample was stratified using two variables: family structure and income. The family structure was stratified along four states: double-headed families with multiple earners, double-headed families with a single earner, single-headed families and single-headed individuals (Hum et al., 2016). The assignment is of particular importance for this study looking at the effect of guaranteed income on fertility decisions. It implies that the assignment to treatment cells is not orthogonal to household characteristics that also affect the outcome variable. It is therefore important to employ a multivariate analysis where the stratifying variables, family structure and income group of the household, are included in the regression (Keeley and Robins, 1980).

A further concern for the sample selection is non-participation and attrition. If we can assume that refusing to participate and dropping out of the program is random, this does not pose any problem other than for the sample size. If however, households with certain characteristics are

3

more prone to refuse to participate in the interviews or to drop out of the program, we would have a bias in the sample. Especially if the reasons of self-selection also positively affect the outcome variable, this would lead to the overestimation of the effect of the treatment.

The participation in an experiment such as Mincome is associated with certain costs, not least in terms of time. The Mincome experiment required the participants to participate in periodic interviews every three months and to file monthly reports of income. The interviews were composed of a very extensive set of questions that went beyond labor supply and included socio-psychological variables, leisure time use, satisfaction with marital life and so on. In our case, a family planning on having a child might find it more important to receive the benefits and thus be less likely to drop out of the program. This would create the problem of self-selection into the sample and overestimate the effect of the program on fertility. It is plausible to think that families for whom the benefits are less important will have less incentives to go through the interview and filing processes.

In the table below, we look at the differences in observed characteristics between the non-participants and the participants to see if any observed characteristic could shape the decision of the household to enroll and continue to participate in the program or drop out. The p-value in the last column tells the probability of rejecting the null hypothesis; that the mean of the given variable is different among two groups. A small p-value indicates that the two groups differ significantly for that variable.

Table 2: Summary descriptives table for non-participants and participants

	Non-Participant N=1421	Participant N=752	p-Value
Double-Headed Household:			0.359
0	622 (43.8%)	313 (41.6%)	
1	799 (56.2%)	439 (58.4%)	
Single-Headed Household:			0.865
0	1164 (81.9%)	613 (81.5%)	
1	257 (18.1%)	139 (18.5%)	
Single Individual	0.26 (0.44)	0.23 (0.42)	0.186
Age of Female Householder	34.7 (12.5)	34.0 (12.2)	0.202
Age of Male Householder	36.7 (13.0)	36.1 (11.9)	0.436
Number of Children in the Household:			0.001
0	937 (65.9%)	434 (57.7%)	

continued on next page

Table 2 – *continued from previous page*

	Non-Participant N=1421	Participant N=752	P-Value
1	320 (22.5%)	196 (26.1%)	
2	134 (9.43%)	96 (12.8%)	
3	25 (1.76%)	23 (3.06%)	
4	3 (0.21%)	3 (0.40%)	
5	2 (0.14%)	0 (0.00%)	
Homeowner Dummy:			0.470
0	887 (62.4%)	482 (64.1%)	
1	534 (37.6%)	270 (35.9%)	
Assigned Treatment Group [1]:			<0.001
Not enrolled from the start	516 (36.3%)	4 (0.53%)	
1	32 (2.25%)	54 (7.18%)	
2	20 (1.41%)	66 (8.78%)	
3	370 (26.0%)	270 (35.9%)	
4	35 (2.46%)	70 (9.31%)	
5	24 (1.69%)	55 (7.31%)	
6	36 (2.53%)	49 (6.52%)	
7	29 (2.04%)	27 (3.59%)	
8	26 (1.83%)	44 (5.85%)	
Control	333 (23.4%)	113 (15.0%)	
Liquid Assets	1895 (5539)	1790 (4964)	0.665
Debt	1336 (2348)	1281 (2045)	0.574
Labor Participation (Male Householder)	0.71 (0.45)	0.83 (0.37)	<0.001
Reason why male householder did not look for work:			.
Employed	780 (79.7%)	426 (84.5%)	
Laid off	4 (0.41%)	0 (0.00%)	
Labor dispute	0 (0.00%)	1 (0.20%)	
Unpaid vacation	5 (0.51%)	4 (0.79%)	
No jobs available	7 (0.72%)	1 (0.20%)	
Bad weather	2 (0.20%)	0 (0.00%)	
Wanted to take care of family	2 (0.20%)	3 (0.60%)	
Child care too expensive	0 (0.00%)	1 (0.20%)	

Table 2 – *continued from previous page*

	Non-Participant N=1421	Participant N=752	P-Value
In-job training	3 (0.31%)	1 (0.20%)	
In school	19 (1.94%)	12 (2.38%)	
Available wages too low	3 (0.31%)	2 (0.40%)	
Did not want to work	91 (9.30%)	34 (6.75%)	
Ill or disabled	50 (5.11%)	14 (2.78%)	
Self-employed	3 (0.31%)	2 (0.40%)	
Retired	2 (0.20%)	1 (0.20%)	
Has a job but has not started working	8 (0.82%)	2 (0.40%)	
Number of Hours Worked Past Week (Male Householder)	334 (178)	365 (147)	<0.001
Job Satisfaction Index (Male Householder):			<0.001
0	268 (27.5%)	80 (16.0%)	
Very satisfied	220 (22.5%)	137 (27.3%)	
Somewhat satisfied	346 (35.5%)	197 (39.3%)	
Neither nor	73 (7.48%)	30 (5.99%)	
Somewhat unsatisfied	47 (4.82%)	37 (7.39%)	
Very dissatisfied	22 (2.25%)	20 (3.99%)	
Expected or actual childcare cost (Male Householder)	0.59 (5.08)	0.31 (3.10)	0.187
Number of Jobs Worked in 1974 (Male Householder)	0.91 (0.74)	1.06 (0.70)	<0.001
Total Earnings in 1974 (Male Householder)	4725 (3637)	4977 (3300)	0.191
Total Earnings in 1973 (Male Householder)	3928 (3356)	4511 (2986)	0.001
Number of Weeks Employed in 1974 (Male Householder)	28.6 (24.1)	33.0 (22.8)	0.001
Number of Weeks Employed in 1973 (Male Householder)	30.4 (21.9)	34.9 (19.4)	<0.001
Years of Schooling Completed (Male Householder)	8.28 (4.79)	9.44 (4.28)	<0.001
Completed high school (Male Householder):			<0.001
0	153 (15.6%)	35 (6.94%)	
1	82 (8.38%)	51 (10.1%)	
2	744 (76.0%)	418 (82.9%)	
Labor Participation (Female Householder):			0.036
.	3 (0.24%)	2 (0.29%)	
0	688 (55.5%)	420 (61.1%)	
1	549 (44.3%)	265 (38.6%)	

continued on next page

Table 2 – *continued from previous page*

	Non-Participant N=1421	Participant N=752	P-Value
Job Satisfaction Index (Female Householder):			.
.	3 (0.24%)	2 (0.29%)	
0	608 (49.2%)	386 (56.4%)	
Very satisfied	203 (16.4%)	106 (15.5%)	
Somewhat satisfied	285 (23.1%)	127 (18.6%)	
Neither nor	66 (5.34%)	18 (2.63%)	
Somewhat dissatisfied	49 (3.97%)	31 (4.53%)	
Very dissatisfied	21 (1.70%)	14 (2.05%)	
Completed high school (Female Householder):			0.080
.	3 (0.24%)	2 (0.29%)	
0	76 (6.13%)	51 (7.42%)	
1	68 (5.48%)	55 (8.01%)	
2	1093 (88.1%)	579 (84.3%)	
Years of Schooling Completed (Female Householder)	9.17 (3.50)	9.27 (3.65)	0.561
Expected or actual childcare cost (Female Householder)	9.06 (16.7)	12.9 (19.4)	<0.001
Reason why female householder did not look for work:			.
.	3 (0.24%)	2 (0.29%)	
Employed	551 (44.4%)	266 (38.7%)	
Laid off	14 (1.13%)	6 (0.87%)	
In school	21 (1.69%)	21 (3.06%)	
Institutionalized	0 (0.00%)	1 (0.15%)	
Available wages too low	17 (1.37%)	8 (1.16%)	
Did not want to work	105 (8.47%)	60 (8.73%)	
Ill or disabled	5 (0.40%)	3 (0.44%)	
Self-employed	9 (0.73%)	4 (0.58%)	
Retired	22 (1.77%)	10 (1.46%)	
Has job but has not started working	8 (0.65%)	6 (0.87%)	
Too difficult to get to town	0 (0.00%)	1 (0.15%)	
Wanted to help with family farm	1 (0.08%)	0 (0.00%)	
Labor dispute	1 (0.08%)	1 (0.15%)	

Table 2 – *continued from previous page*

	Non-Participant N=1421	Participant N=752	P-Value
Unpaid vacation	8 (0.65%)	3 (0.44%)	
No jobs available	14 (1.13%)	8 (1.16%)	
Wanted to take care of family	430 (34.7%)	268 (39.0%)	
Child care too expensive	15 (1.21%)	11 (1.60%)	
Pregnancy	14 (1.13%)	8 (1.16%)	
In-job training	2 (0.16%)	0 (0.00%)	
Number of Hours Worked Past Week (Female Householder)	330 (142)	319 (145)	0.100
Total Earnings in 1973 (Female Householder)	1203 (1674)	1018 (1524)	0.015
Total Earnings in 1974 (Female Householder)	1564 (2102)	1257 (2006)	0.002
Number of Weeks Worked in 1974 (Female Householder)	21.7 (23.8)	18.0 (23.0)	0.001
Number of Weeks Worked in 1973 (Female Householder)	16.9 (20.2)	14.5 (19.1)	0.010

The descriptive statistics show that the participants and the non-participants do not differ significantly in terms of being single individuals, single-headed households or double-headed households. Furthermore, differences in the age of householders, whether the household owns a house, their liquid asset holdings or liabilities are not statistically significant.

As can be expected, the assignment to plans seems to be very significant in determining whether or not the household continues with the experiment. Drop out rate is much higher in the control group and much lower in more generous treatment groups such as plans 1 and 2.

Furthermore, the male householders in the participating group are more educated, have worked more hours in the past year, are more likely to be employed and have higher earnings as compared to the non-participating male householders. As for female householders, the most significant difference between the two groups seems to be the expected or actual cost of childcare with the participating female householders reporting higher costs. As opposed to the male householders, the female householders in the participating group typically work less hours and weeks in a year and have less reported earnings than those that did not enroll or dropped out. This might suggest that in the participating group, the fact that the female householder is at home and can file the

---

[1]Households that are marked with "not enrolled from the start" are those that completed the baseline surveys and then did not participate at all in the experiment. Households that were assigned to a treatment group are those that initially enrolled in the program, but then dropped out within the first two years of the experiment.

monthly report and respond to interviews made it easier for these families to participate in the experiment.

An important difference between the families that participated in the program and those that dropped out for our question is the number of children. The participating families have on average more children. 65.9% of households in non-participants have no children, which is 57.7% for the participating group. Together with the difference in childcare cost reported by female householders in participating and non-participating groups, this difference suggests that if the families self-selected themselves to treatment, it is most likely due to the motives of supporting their already born children rather than planning to have children.

References

[1] Gerald Brian Farthing. *Social experiments and social policy formulation: A study of the Manitoba basic annual income experiment.* PhD thesis, London School of Economics and Political Science (United Kingdom), 1992.

[2] Derek Hum and Wayne Simpson. *Income Maintenance, Work Effort, and the Canadian Mincome Experiment.* Economic Council of Canada, 1991.

[3] Michael; Hum, Derek; Laub and Brian Powell. The objectives and design of the manitoba basic annual income experiment (mincome technical report 1), 2016.

[4] Michael; Metcalfe Charles; Hum, Derek; Laub and Don Sabourin. The sample design and assignment model (mincome technical report 2), 2016.

[5] Michael C Keeley and Philip K Robins. Experimental design, the conlisk-watts assignment model, and the proper estimation of behavioral response. *Journal of Human Resources*, pages 480–498, 1980.

[6] David Kershaw and Felicity Skidmore. The New Jersey Graduated Work Incentive Experiment. Technical report, 1973.