

Evaluating the basic income using an experiment

Gregory Mason

Gregory.Mason@umanitoba.ca

October 31, 2018

Abstract

The basic income experiments of the seventies and the recent termination of the Ontario Basic Income Pilot (OBIP) offer important lessons for the design and conduct of large-scale social experiments. Much is at stake, since the projected gross cost for implementing a basic income in Canada is \$79 billion over 2018-23. A high-quality policy experiment promises to test the validity of the expected consequences and map unintended consequences. Yet attempts to test policy using experiments *ex ante* have failed. Three reasons exist for this. First, basic income pilots require a level of logistical support and patience that few governments seem willing to support. Second, proponents have imagined many outcomes and it has become a general antidote for many of modern society's ills. The theories of change that link the intervention to outcomes have become complex and resistant to simple specification. Finally, most pilots continue to rely on survey data that introduces a range of sampling and response variability that undermines the ability to measure impacts. The paper concludes with an experimental design intended to test the main behavioural hypotheses believed to be associated with a basic income.

Draft: For comment only.

1. Introduction

The universal basic income (UBI) has endured as central policy for combating poverty and inequality.¹ For many it has assumed iconic status as a core policy prescription with far reaching outcomes and capacity to transform society. It also has been subject to the most ambitious large-scale social experiments attempting to evaluate economic policy *ex-ante* using.

This paper reviews the concept of a UBI, explains the underlying theory of change and causal structure believed to characterize the policy, assesses attempts to evaluate it, and present an experimental design that overcomes shortcomings that have marked in prior evaluations. The paper starts with a review of the UBI concept and its associated theories of change that link the income supplementation with the intended/expected outcomes. Following this, I present a brief review of findings from prior income maintenance studies comprising work incentive effects, the focus for research most studies and other “non-economic” research.

The next part of the paper, reviews the research designs common to negative income tax experiments, using Mincome and the Ontario Basic Income Pilot. This section explores the conceptual models such as sampling, experimental scope, and data development, identifying key weaknesses in the methodologies.

Of importance to the critical assessment of UBI experiments is the “political economy” for a negative income tax. Most studies focus on measuring the impact on those who receive the treatment. A central idea of this paper, is that the UBI has political economic effects on those who are not the direct targets for such income supplementation. This implies that the scope of the experiments need to be wider than is commonly imagined.

The paper concludes with a proposed design for a basic income experiment, that draws from the social experimental literature, while attempting to avoid the traps of the income maintenance experiments of the seventies and the most recent Ontario study.

2. What is the universal basic income?

2.1. Basic form of a UBI

¹ The terms universal basic income (UBI), basic income (BI), negative income tax (NIT), and guaranteed annual income (GAI) are interchangeable.

According to the Basic Income Canada Network an ideal basic income has the following features:²

- Enables individuals to have both:
 - autonomous income to meet their needs; and
 - access to public services that benefit all of us;
- Replaces income provided through social assistance systems and other supports such as GST rebate;
- May not replace other income support such as Employment Insurance, old age security and public pension;
- Is inflation adjusted and declines as other income increases;
- Can be adjusted to meet specific needs (lone parenthood);
- Makes no one worse off by the transition from the existing system;
- Does not negate the need for labour adjustment programs or education;
- Does not eliminate the income tax system.

Two forms of a basic income comprise a demogrant and a negative income tax (NIT). A demogrant consists of regular payments based on some form of demographic eligibility (age, residency or citizenship), but not income, wealth or work participation. A typical proposed format for a demogrant involves tax free payments received by eligible households, with any additional income received (regardless of source) above this guaranteed level taxed, either proportionately or progressively.

A negative income tax, sets a threshold and then awards income supplements to eligible households to ensure that income always reaches this level. Those without income receive an amount to bring them to this threshold, those with income is 50% of threshold receive a portion of that level, and those whose income is \$1 less than the threshold, receive some fraction

A negative income tax assures all eligible households receive minimum income, with payments declining as other income increases as show in *Figure 1*.

² https://www.basicincomecanada.org/about_basic_income

A regime with a 0 NIT, would have incomes before and after the NIT being identical. Defining a threshold value G or support level, shows the NIT payment received by someone with no other income. The offset tax rate t , shows the rate at which the NIT payment declines as other income ceases. Eventually the NIT payment falls to 0 at the breakeven level B . In this formulation $B = G/t$, so specifying two of the three parameters of a NITY identifies the third. Note that this representation abstracts from the tax system that may apply to earnings.

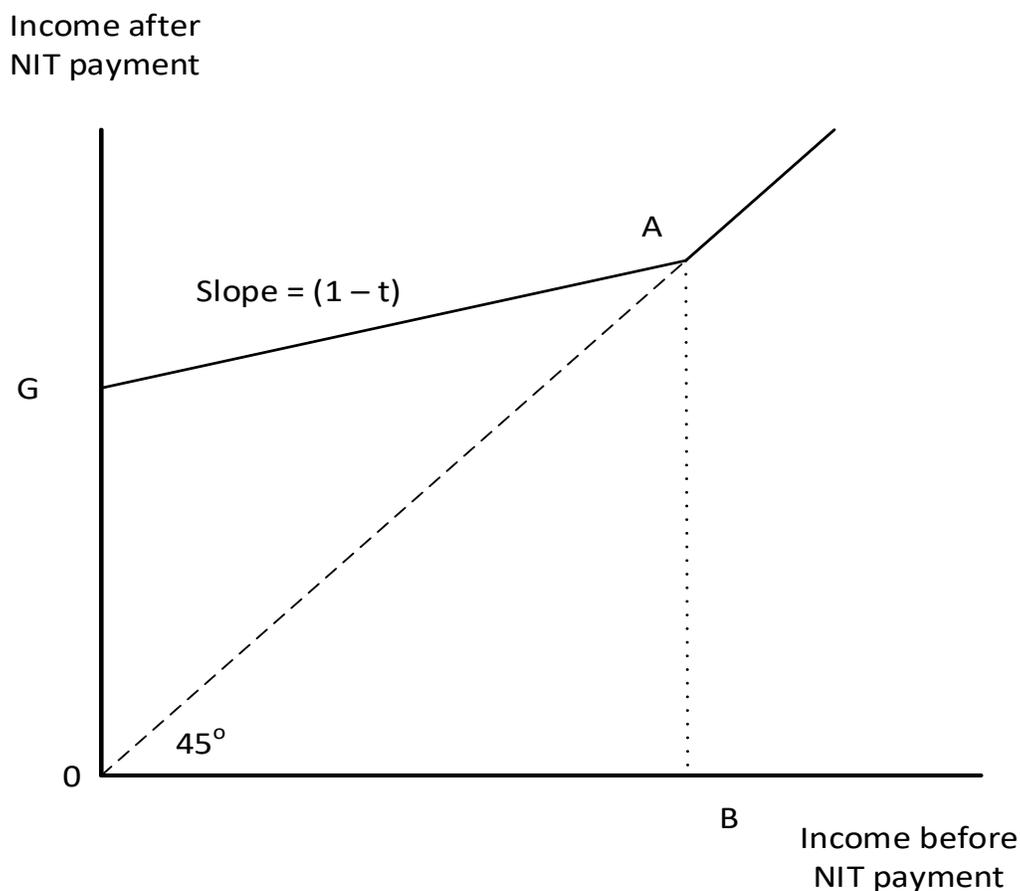


Figure 1 Negative Income Tax

Canada already has forms of a demogrant and NIT. Old Age Security is payment for which everyone over 65 is entitled and the Guaranteed Income Supplement is a payment inversely linked to income for seniors as well. The Canada Child benefit is also an income tested supplement for parents based on income and numbers of children under 18.

2.2. The results of the NIT experiments

The income maintenance experiments of the seventies³ focused on estimating the short-run labour market response of a negative income tax. Aside from the ethical issue of whether a NIT would encourage idleness, a large withdrawal from work could increase the costs of the program. Ancillary goals include testing the effect of a NIT on health outcomes, educational attainment, etc. As I discuss in the next section, the structure of the large scale NIT experiments thus far, have been aligned to test labour market impacts, which can circumscribe the capacity to properly evaluate ancillary impacts.

Both in the seventies and today, policy activists promote the NIT as the most effective antidote to poverty and inequality. These earlier studies revealed two important findings that promote the value of this approach.

- First, some members of the work force would reduce their participation, either in part or entirely. Typically, these are members of a household that had the lower incomes, and elected to engage in other activities, most commonly parenting. In the context of the seventies these were most often women (Prescott and Swidinsky ((1985)).
However, the measured *short-run* effects were quite small. The emphasis on short-run is important, since none of the experiments persisted sufficiently long to assess changes over a longer period than 5 years and most finished after 3. Moreover, these results apply to the labour markets of forty years ago, which may have been more institutionally “rigid” than the current so called “gig” economy.
- Second, these modest impacts encouraged many to see a NIT as financially feasible, since the labour adjustments would not lead to mass work withdrawal that requires large increases to public budgets.

³ New Jersey Graduated Income Work Incentive (1968-72), the Rural Income Maintenance Experiment (RIME) (1970-72), the Seattle-Denver Income Maintenance Experiment (SIME/DIME) (1970-76), the Gary Indiana Experiment (1971-74), and the Manitoba Basic Annual Income Experiment (Mincome) (1974-78).

As Wilderquist (2002) writes

Because the work disincentive effects of the NIT were greater than negligible but not so large as to make the programme unaffordable, the meaning of the figures depends on how large is large and how small is small. The work disincentive effect seems to have been just enough that supporters can claim it to be small and opponents can claim it to be large.

This remains the current state of the debate. The fact that research has neither measured long-term impacts or the role of labour market flexibility in work effort adjustment remain important issues and a major motivation for new experiments.(Hum and Simpson (2001). As Burtless and Hausman (1977) note, depending on who exactly adjusts their work effort can have a major impact on the final budget. Someone earning a minimum wage in part-time casual work, but whose income is quite close to the guarantee threshold reduces effort by a few hours a month, has a lower impact on public budgets than a similarly situated individual who decides to stop work entirely. The income maintenance experiments cannot reveal the nature of the adjustments in such detail, which is why [projections of the cost of a basic income](#), such as issued by the Parliamentary Budget Office lack the required context.

The results from the NIT experiments, especially the news that a discernible work disincentive did exist, halted any political movement on this policy in the US (Moffit (2003). Even a strong advocate such as Daniel Monyihan withdrew his support in the face of the negative news that recipients tended to reduce their work effort. Neither the researchers, nor the press picked up on the potential benefits of such a work withdrawal, such as the value of increased attention to parenting or that with a reduction in the supply of labour, the wages of those who remained employed might increase. Despite, very significant research budgets, these massive social experiments failed to clarify core issues around a basic income, and appear to have offered evidence to both opponents and proponents. Of note is that these studies failed to shed light on the capacity of the basic income to reduce inequality or poverty, which was the original rationale for the UBI.

In the last decade, the apparent increase in inequality has fueled a renewed attention on the basic income and a tool to promote economic fairness. This increase in support has occurred despite evidence that suggests both the poverty rate and measures of inequality in Canada show signs of reversing.

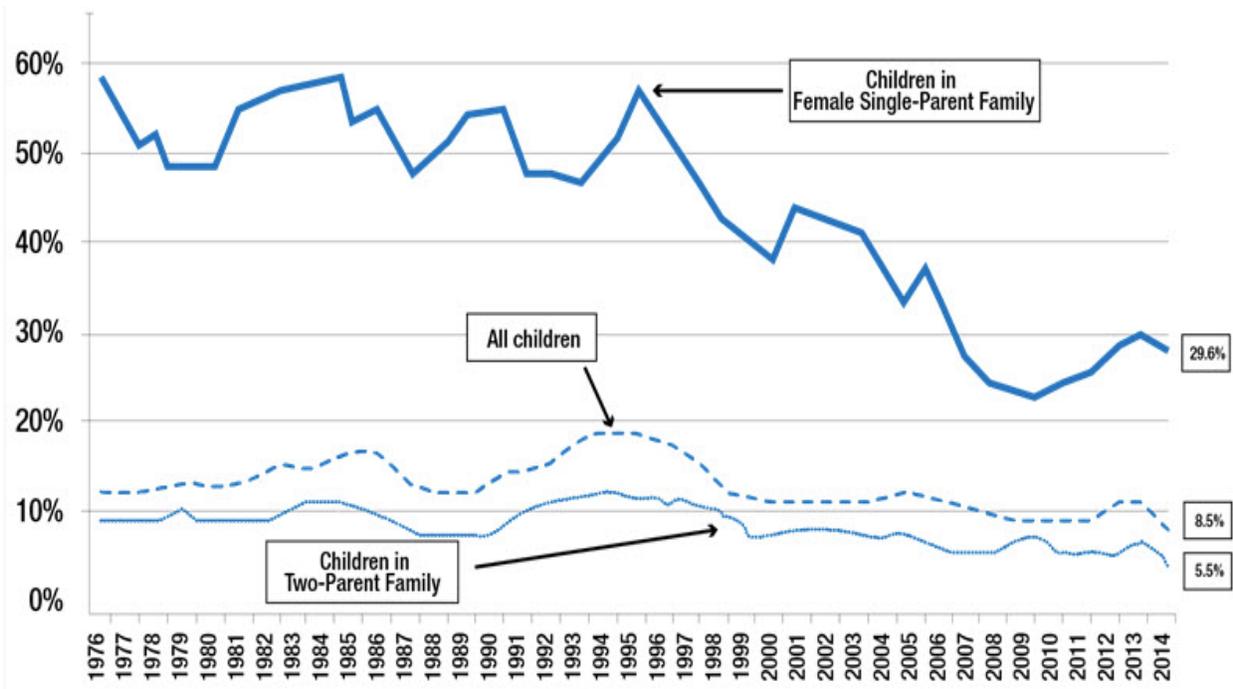


Figure 2 Low Income Measure in Canada

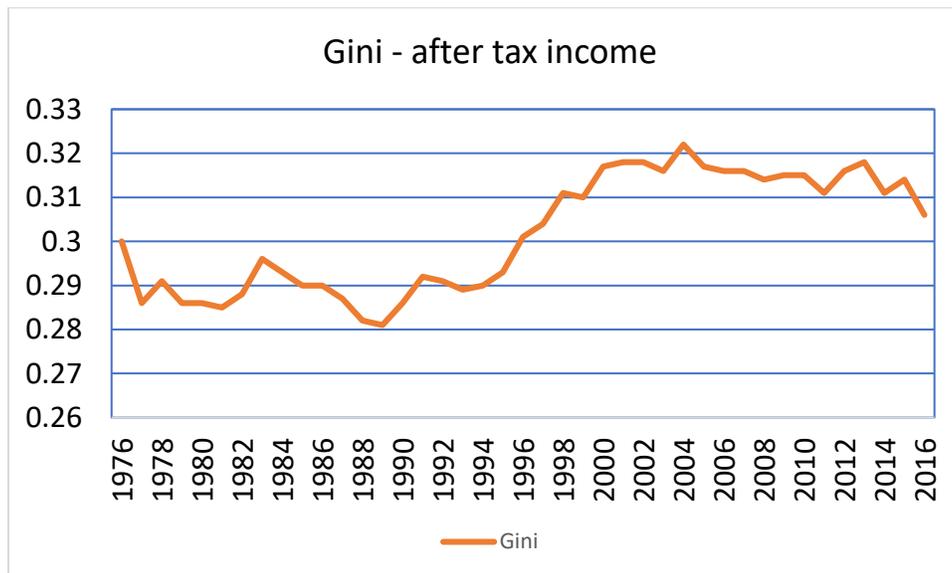


Figure 3 After Tax Gini - Canada

Setting aside the important question whether these measures are valid, proponents of a basic income now advance a more complex and diverse narrative to support this policy.

2.3. Current theories of change about the UBI

In the program evaluation literature, a theory of change refers to the behavioural and social processes that both lead to the development of an intervention as well a causal logic for exploring the range of enduring outcomes. Such theories rest on a theoretical framework and typically employ a logic model closely resembling a path diagram interrelating causes and effects.

A key feature of theories of change is their implied temporal element, unlike the comparative statics that characterizes most of economic policy. Policies are neither instantaneous nor final and outcomes occur over time. The value of logic models is their ability to make causal connection explicit.

Theories of change also feature multidimensional and contextual causal relations, which become simultaneously a major element in creating a realistic causal structure, but also set the potential for crippling complexity that subverts tractable analysis. Balancing realism with tractability is an important goal for social experiments of “big” policy such as testing a basic income before the fact.

The recent literature on the UBI has mushroomed, especially with respect to speculating on the benefits of this form of support. Now, much more than just reducing poverty, many see the UBI as having the potential to influence a wider range of social ills as well as creating a foundation for a better society.

A common argument in favour of a BI is that increased income and increased income security will yield important health benefits. The work of Forget ((2011) is noteworthy for having attempted to test this using a quasi-experimental approach. Flowing from this idea is the notion that a UBI will affect all aspects of health outcomes specifically reduced use of mental health services, increased visits to primary care, and increased active management of chronic conditions. Another potential avenue for increased income to support health is the increase in food security. Willows et al (2009) do not connect food security to a basic income, but the connection seems plausible. However, Thoits and Hannan (1979) use the Seattle-Denver experimental results to show that some increased distress occurred among recipients, a finding that is consistent with the notion that income changes, up or down, represent life events that tend to confront recipients opportunities to change aspects of their lives.

Associated with the notion that a basic income addresses inequality is the notion that technical change will displace and increasing share of lower skilled labour. Many have observed that the last decade has seen a “skills” divide. Technological advance displaces those with lower skills from the workplace, and a basic income both “recycles” income from rich to poor as well as offers the essential economic support. Whether this is actually occurring is a matter of debate (Autor(2014)), but one prediction of proponents is that the UBI will allow individuals with low skills to regroup and invest in education. Finally, others extend this to the idea that increased economic security will allow recipients to engage in more creative pursuits with the rejuvenation of arts and culture.

A common thread in the support for a basic income is the notion of universal justice. Pasma and Mulvale (2009) see the basic income as fundamental to “economic democracy that provides economic support and to all citizens.” Further they argue a range of social benefits to those covered by social assistance programs, increases social cohesion and protection for vulnerable persons. Zelleke (2005) evaluates a UBI using the principles of distributive justice and concludes the unconditional nature of a basic income aligns well with the concept of a property owning democracy articulated by Meade (1964). Huws (2017) argues from a feminist perspective that a universal income would compensate household members whose work is not recognized by the market economy, thereby becomes a policy supporting women’s liberation.

Finally, many authors mention the idea that a basic income reduces the stigma associated with social assistance. To be sure with the era of direct deposit, the spectre of forming a queue in public view to receive a cheque no longer exists. Stigma can remain perceived by recipients of social assistance, if they view that somehow the payments are not normal or “deserved.” A very common reason for refusing to participate in Mincome was the perception that these payments were a form of handout (Kurz 1978). However, recent qualitative analysis of Mincome data by Calnitsky (2016) appears to suggest that those receiving payments under a NIT may well view such support less moralistically and less tainted by social stigma.

One emerging issue may be termed the “political economy” of a basic income. Identified by those working from a left perspective, this view attempts to identify the impact of a basic income on institutions such as unions and worker cohesion. As D’Ippoliti (2018) argues a basic income has potential negative impacts on these more abstract concepts. If we assume that a basic income is enough for any competent person to manage their affairs, then the market economy will take precedence over public provision of services. Subsidized housing is an example, where recipients of

a BI would be expected to fend for themselves which in turn may create pressure on scarce affordable housing. Second, a potential exists for devaluing the public sector as payments become solely conditioned by income. For some, such as Milton Friedman (), the removal of welfare state bureaucracy represents a benefit. For others, however, a purely financial support system cannot compensate for the non-financial services that often attend the provision of social services. Assuming a purely cash transfer represents the sole support needed by low-income families may over simplify the needs of families. Third a UBI may reduce union power even further. Fourth, it is possible that a UBI would remove low skilled persons even more from the labour market and create a permanent underclass who could never advance their economic situation. Fifth a UBI may attract increasing international migration, further alienating low income working populations.

To extend the idea of political economic effects, the focus of basic income research is overwhelmingly on households that receive support. Indeed, much of the analysis resembles a neo-classical “two-factor” model ... the low-income recipients and the high-income payers. In other cases, the dichotomy comprises the recipients and “government”, aka tax-payers. In fact, it important to insert at least one additional income tranche, namely those whose income is just above the breakeven level. For example, in the recently cancelled Ontario Basic Income Pilot, the threshold for a single person was just above \$16,000 annually. Any earnings above that resulted in an offset of 50% reduction in the basic income payment. Therefore, someone earning \$14,000 (either in wages or other income), would receive an OBIP payment of almost \$10,000, for a total income of \$24,000.

Now here is the issue. Statistics Canada reports that the median income for Ontario in 2016 was \$32,000. A check of popular job search sites such as Workopolis reveals that incomes for occupations such as data entry clerk typically start at around \$30,000. With the advent of the so-called gig economy increasing numbers of younger Canadians can expect to cobble together several part-time jobs to make a living. Even though this strategy might result in incomes that would push someone beyond the reach of a basic income, it is conceivable that the availability of a basic income could induce those in the income tranche above the reach of the program to adjust their work effort. The elasticities measured in the seventies may have been unduly low because the focus was on those who received the benefit and because labour markets may have been more institutionally rigid.

This is similar, but not identical to what Kurz (1978) terms the “truncation” problem common to all social experiments that evaluate means tested programs. Most policy analysts failed to consider

impacts of a means tested program on the income tranche just above the breakeven level. Further the data may not include sufficient numbers of those whose income lies between the payment (G) and the breakeven level (B). These households may self-select themselves out of the experiment under the belief they are not eligible. Hausman and Wise (1977) present techniques for estimating parameters in the face of such distortion.

Consideration of wealth remains an important omission from the entire inequality and basic income literature. The focus has been entirely on income. Yet examples of wealthy households with low income are increasingly common...these are seniors. Mincome did condition payments on wealth, but OBIP did not, with the result that it is entirely possible that unemployed singles could remain in their parents, rent free. Social assistance programs typically limit the wealth recipients may retain, meaning that for most must deplete their assets, with the exception of primary residence. Most modern proposals for a basic income do not consider how wealth would condition benefits.

Finally, administrative issues have received limited attention from the NIT experiments. For most commentators, the issue of sending money to low income households appears simple, but complexities abound. A basic problem is “cadasterability” were a “cadaster” is a comprehensive list of land holdings (Wispelaere and Stirton (2007)). In the context of a basic income, it refers to the list of all eligible households. In principle, the income tax files could form a cadaster, but for two limitations. Obviously, income tax records track wealth poorly. More important is that many low-income individuals do not file tax returns. Social assistance payments are untaxed, and recipients would typically not need to file a return unless they received earnings where the employer deducted tax at source and filed statements with the income tax authority. Low-income households often work in the so-called underground economy and have incomes too low to trigger attention. They do not leave an administrative footprint.

Limited cadasterability has forced all NIT experiments to rely on a series of screening surveys to qualify participants. This raised the costs of the experiments and if maintained for the actual program, would blunt the potential of a basic income to trim bureaucracy. Reliance on surveys to monitor program outcomes remains a core defect of all NIT experiments.

This brief overview of the current thinking about the impacts of a UBI, needs joining with the traditional theory of change for a NIT to create a research program for an updated policy evaluation. Figure 2 presents a high-level logic flow for the basic income.

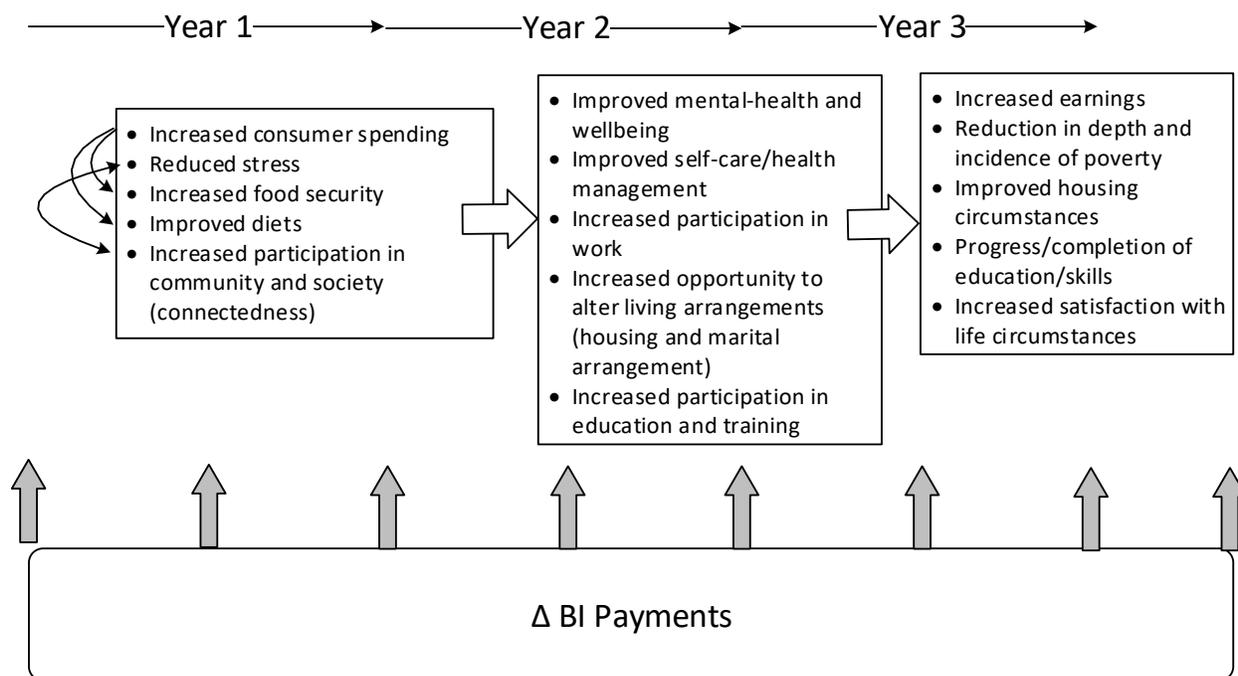


Figure 4 High Level logic model for Ontario Basic Income Pilot

(Explain diagram more)

2.4. The main themes and questions for a BI experiment

In most BI models, while households receive the benefits, it is important to track some responses at the individual adult member. This section poses some central issues for evaluating a basic income, grouped in three main themes with specific sub-questions. The final questions that will be posed will reflect pragmatic concerns about data availability, the sustainability of data collection, and ethical/privacy constraints governing the use of administrative data and limits to primary data collection.

- *Theme 1: Changes in economic well-being and work effort*

Questions about how a basic income reduces poverty remain fundamental. Key sub-questions include

- a. Do recipient households experience changed income at the individual and the household level?

- b. Have reductions in earnings occurred for individual household members?
- c. Has work effort changed for individual and household members (measured in hours of work)? Have these changes remained stable?
- d. Have members of recipient invested in vocationally related education or training?
- e. What non-work activities do members of recipient household engage in and how does this vary by major demographic?
- f. Do recipients change housing consumption broadly defined? By broadly defined, is meant changes in housing condition, crowding, and location.
- g. How to non-recipients in the immediately high tranche that the breakeven adjust their work effort? Is there any evidence that some reduce work effort to qualify for a BI?

- *Theme 2: Health and social outcomes*

The potential for a basic income to have positive impacts on health remain a central idea. Testing this idea within the context of a universal health care system such as exists in Canada, presents important measurement issues. Canadians typically do not face barriers to accessing primary care, except possibly in remote locations. More likely is that the “working poor” do not pay for access to supplementary health services. Those on social assistance usually obtain supplementary health services as part of the non-financial benefits package associated with this program.

Therefore, the causal link between an increased and more stable income as health outcomes is likely indirect, working through food security and changed stress associated with poverty. It is unclear whether increased resources will lead to reduced use of mental health services, reduced hospitalizations, and changed morbidity/mortality.

Based on the discussion in the previous section, another claim is that increased income supports increased engagement with society. Children especially may benefit when a household gains access to the internet and can afford a computer. Again, the causal relationships may be complex and indirectly lead to other outcomes, such as improved academic performance.

Key questions for the basic income – health link may include:

- a. Do household members change (reduce/increase) use of mental health services?
- b. Does the hospitalization rate change for members of recipient households?

- c. How does the consumption of pharmaceutical change in terms of quality and type? What does any changed pattern indicate about changes in health status?
- d. How do primary care change for household members (visits, immunizations, screening, etc...)
- e. How does the BI affect family stability (divorce/marriage, adult children at home)?

It is important to stress that it is difficult to predict direction of some measures of health status or whether they imply an improvement in health status. Increased use of primary care is probably an indicator of future improvements, but more visits for mental health counselling may indicate the household members are taking a more proactive approach to care, or may be experiencing increased stress and reflect worsening mental health. A clear issue for any evaluation of the basic income is to undertake follow-up analysis to understand the changes in measured outcomes.

- *Theme 3: Administration of a BI*

The BI requires a means test on income. This requires identifying eligible responses, either by going to administrative data and/or enrolling using a constructed sample frame. The next section discussed the theoretical and practical issues of creating assembling the data for analyzing BI outcomes.

Some important administrative issues that need to be evaluated include the following:

- a. Will social assistance recipients (SAR) accept a transfer to a basic income? Will they remain within the BI experiment? What factors explain why SAR return to welfare?
- b. Will the “working poor” accept an invitation to participate in a BI experiment?
- c. What is the projected cost of a BI based on take-up and work effort adjustment?

These themes and sub-questions represent the core of avenues of inquiry for a BI pilot. Many have identified other social and cultural impacts, but no practical way exists to measure outcomes such as increase in artistic outputs. Changes in social cohesion and increase in community spirit are similarly vague notions that are both hard to measure and would likely fail to motivate acceptance of a BI.

Answering these questions requires varying forms of data. The NIT experiments thus far, including the most recent OBIP, relied exclusively on survey data. This created serious biases due to non-

response, item unreliability (respondent/interviewer variation in question interpretation), self-selection into the study, and self-selection out of the study (attrition). These data distortions undermine consistent estimation of impacts.

Given the evolution of large administrative datasets, an important goal for any future will be to align the themes/questions to the appropriate data sources. **Table 1** presents a comparison of administrative data (tax files and health records), sample surveys, and qualitative interviews in supporting different question constructs.

	Description	Advantages	Disadvantages
Administrative data	Information routinely collected due to respondent participation in service consumption activity (taxation records, health services,	<ul style="list-style-type: none"> • Fixed variable definition among respondents • Temporal alignment usually assured • High coverage rates across subject population • Low collection costs • Option to use entire population (if experimental costs permit) • Temporal alignment high 	<ul style="list-style-type: none"> • Variable definition may not align with theoretical constructs • Privacy barriers may impede use • Information sources may not share a common key (which creates barriers to merging datasets merged.)
Sample survey (fixed response categories)	Set questions with limited response options posed to a large sample (n>100)	<ul style="list-style-type: none"> • Samples usually randomly selected supporting statistical inference to population. • Question construct can closely align with theoretical concept. • Low cost per unit for collection and analysis. 	<ul style="list-style-type: none"> • Interviewer/respondent misunderstanding of question intent. • Response categories and questions construct fail to capture theoretical concept • Non-response and selection biases. • The field of knowledge defined by researcher. • Temporal alignment careful logistical management.
Qualitative interview (open responses)	Open questions (respondent is free to respond in own words) posed to a small sample (n<50)	<ul style="list-style-type: none"> • Samples selected to include only respondents with relevant information • Respondents can expand the field of knowledge 	<ul style="list-style-type: none"> • High cost per unit for collection and analysis • Analyst interpretation reduces reliability and validity • Temporal alignment low

Two elements of Table 1 may require explanation.

- A large sample survey requires fixed questions with stable response categories. The researcher defines the scope of responses and respondents must align their answers fit one (and usually only one) response category. The respondent cannot modify the answer except by adding a verbal comment or inserting a written comment if the researcher has allowed space in the margin or as an “Other(specify)” option. The cold reality is that most of the time researchers ignore such parenthetical comments.

A qualitative interview, either singly or severally (focus group) poses general questions where respondents offer comments that researchers transcribe and then code into categories. Here the nature of responses shapes the categories, and in this sense the respondent shapes the information or field of knowledge, mediated of course by the researcher.

- The concept of temporal alignment receives little attention, largely because most longitudinal research uses administrative files. The NIT experiments have used survey data, where re-interviewing respondents at fixed intervals becomes almost impossible. Where external influences, such as a sharp change in unemployment or prices do not occur, temporal variability in interviewing may not pose much problem. Figure 5 shows this issue, while.

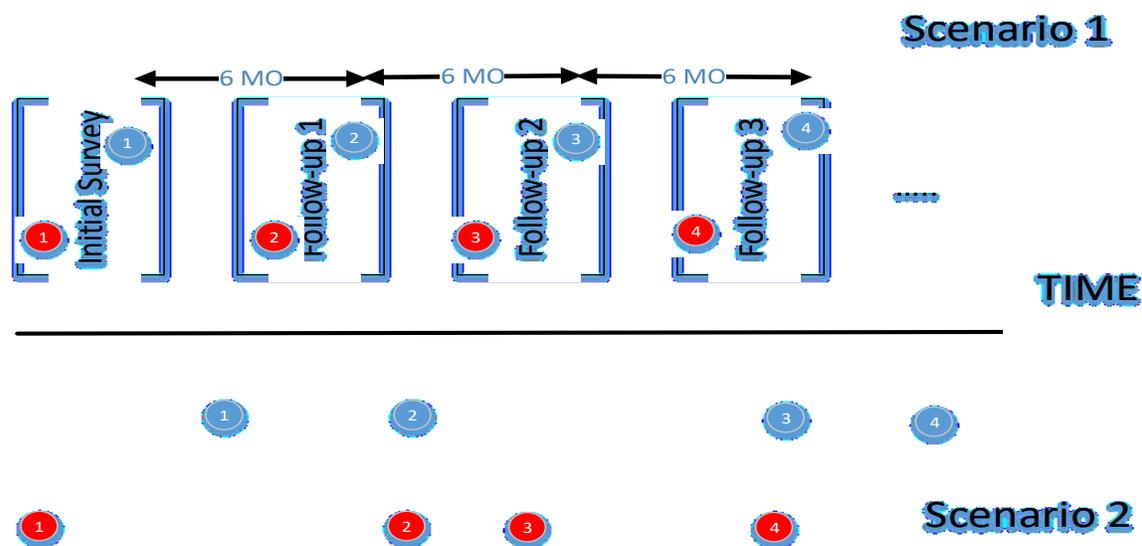


Figure 5 Temporal Alignment for panel data

Figure 6 illustrates the typical “spread” of interview timings of the Mincome periodic surveys. Note how different survey waves overlap. In the case of Mincome, where both adult household heads (if present) and children participated in an interview, revisits to secure all the interviews increased the costs considerably. Note the interviews conducted by Mincome, featured in-person, in-home interviews that typically required three hours and collected information on several hundred items. The modal interview numbers noted in Figure 3 illustrate the intensity of field effort required to collect primary data for a complicated panel study. In a modern context, where telephone or on-line techniques may be the method of data collection it may be possible to “tighten” the data collection, but this still requires substantial logistical control and incentives for respondents to complete surveys at the prescribed times.

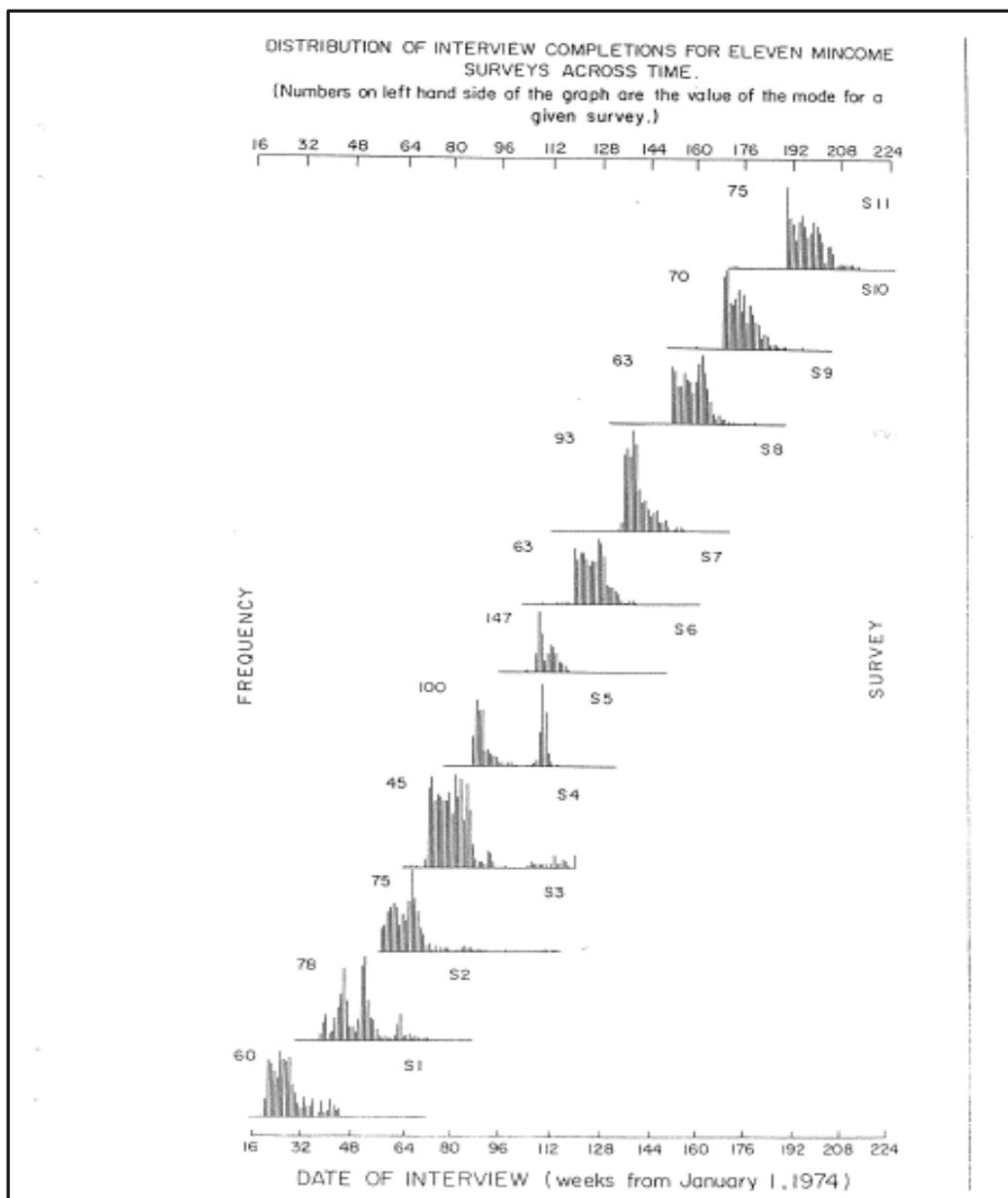


Figure 6 Distribution of Interview Completions for Mincome Survey

Table 2 aligns the evaluation themes/questions with the data source.

Table 2: Evaluation theme/question aligned to data source				
Validity/reliability/cost rating ●●●high ●● moderate ● low - not useful	Administrative data		Sample Survey	Interview/Focus Group
	Tax files	Health Records		
<i>Theme 1: Changes in economic well-being and work effort</i>				
a. Do recipient households experience changed income at the individual and the household level?	●●●	-	●●	-
b. Have reductions in earnings occurred for individual household members?	●●●	-	●●	-
c. Has work effort changed for individual and household members (measured in hours of work)? Have these changes remained stable?	●●●	-	●	-
d. Have members of recipient invested in vocationally related education or training?	●●	-	●●	-
e. What non-work activities do members of recipient household engage in and how does this vary by major demographic?	●●	●●	●●●	●●
f. Do recipients change housing consumption broadly defined? By broadly defined, is meant changes in housing condition, crowding, and location.	●●	●●	●●	-

Table 2: Evaluation theme/question aligned to data source				
Validity/reliability/cost rating ●●●high ●● moderate ● low - not useful	Administrative data		Sample Survey	Interview/Focus Group
	Tax files	Health Records		
g. How to non-recipients in the immediately high tranche that the breakeven adjust their work effort? Is there any evidence that some reduce work effort to qualify for a BI?	●●●	-	●	●●
<i>Theme 2: Health and social outcomes</i>				
h. Do household members change (reduce/increase) use of mental health services?	-	●●●	●●	
i. Does the hospitalization rate change for members of recipient households?	-	●●●	●	●
j. How does the consumption of pharmaceutical change in terms of quality and type? What does any changed pattern indicate about changes in health status?	-	●●●	-	-
k. How do primary care change for household members (visits, immunizations, screening,...)?	-	●●●	●●	-

Table 2: Evaluation theme/question aligned to data source				
Validity/reliability/cost rating ●●●high ●● moderate ● low - not useful	Administrative data		Sample Survey	Interview/Focus Group
	Tax files	Health Records		
1. How does the BI affect family stability (divorce/marriage, adult children at home)?	●●●	●●●	●●	-
<i>Theme 3: Administration of a BI</i>				
a. Will social assistance recipients (SAR) accept a transfer to a basic income? Will they remain within the BI experiment? What factors explain why SAR return to welfare?	●●●	-	●●	●●●
b. Will the “working poor” accept an invitation to participate in a BI experiment? Why do they decline to participate?	●●●	-	●●	●●●
c. What is the projected cost of a BI based on take-up and work effort adjustment?	●●●	-	●●	-

This discussion reinforces the idea that administrative data, specifically tax files and health records should form the data foundation for any future BI pilot/experiment. In Canada, with its universal health records, provinces have access to health records on the population. Data on health provider billings, pharmaceutical usage, primary care screening, and other health services are becoming

increasingly complete and timely, especially as electronic medical records are finally permeating the system. Supplementary and complementary health services would require sample surveys.

The most serious defect of administrative data is that research must adapt concepts to the variables as defined, usually for administrative purposes and not with research needs in mind. Survey questionnaires can include latent and abstract measures such as food security as well as attitudinal and psychological states.

Income tax records can record earnings, other income sources, medical deductions, work deductions and education/training expenditures that also generate deductions. The tax records become the primary data source for predicting uptake in the BI and cost projections.

In this view, sample surveys that have formed the data “backbone” for prior NIT experiment become supplementary. Qualitative research, remains important to answer the “why” questions. The next section presents a methodological design for a NIT experiment.

3. Design for a Basic Income Pilot

3.1. Review of NIT experiments – Mincome and OBIP

The main features of Mincome and OBIP, especially their successes and missteps supports the formulation of a basic income experiment.

3.2. Mincome design features

Mincome focused on testing a single null hypothesis “*a NIT will not reduce labour market attachment*” All household members over 15 provided details on their labour force participation using extensive, in person, in home sample surveys. Other data collected served as mediators for testing this core hypothesis. Over time, researchers attempted to include other research objectives (such as impact on wealth, family relationships...) but these were always seen as secondary to the main purpose of testing labour force attachment.

Mincome collected data from three sites

- Winnipeg was the main site and featured a full randomized design featuring 8 “treatments” and a control group
- Rural dispersed site, with a random selection and control, but only one treatment.
- Dauphin site (saturation) with a volunteer sample, single place and no control. As a saturation site, any qualified Dauphin resident could enrol.

The experimental design (for Winnipeg) appears below (Table 3). A treatment comprised a level of annual support and a tax on total income (earnings, other income and Mincome supplement). It is important to understand that most Mincome recipients had some form of other earnings, but once enrolled in the experiment respondents had to withdraw from social assistance. At the time of enrollment, no social assistance recipient of Mincome became worse off financially

Guarantee at enrolment, \$	Tax Rate on Total Income		
	35%	50%	75%
3,800	Plan 1 (n = 55)	Plan 3 (n = 61)	Plan 6 (n = 49)
4,600	Plan 2 (n = 67)	Plan 4 (n = 70)	Plan 7 (n = 29)
5,400	Plan X	Plan 5 (n = 56)	Plan 8 (n = 45)
		Plan 9 (Control: n = 94)	

Mincome had a complex formula for determining payment at enrollment, based on family size and net worth. The value appearing in this table refers to a specific family size and wealth, with no other income. The receipt of earnings and other income (such as pensions or disability) mediated the actual payments receive which would vary to reflect a family’s changing circumstances.

Many see Mincome as a standard RCT, but allocation to the treatment and control groups did not feature the usual randomization process. Rather an “assignment” process placed participants in a treatment or control group based on their expected “informativeness”. As explained by Hum et al. (1974) the allocation results from a solution to a linear programming algorithm designed to maximize informativeness subject to a cost constraint. Most of the later income maintenance

experiments used this approach, and while it may have managed costs, has been subject to criticism (Lyll (1975), Bernstein (1975)). The assignment model results in a non-orthogonal design that impedes straightforward multivariate analysis. Most important is that the information function which the assignment model seeks to maximize, must include all relevant factors that determine work behaviours, but how can one know these before the fact?

Development of the sample, proceeded using a series of preliminary surveys, before the formal enrollment and assignment of participants to the treatment and control group. After enrollment, a series of periodic surveys tracked respondents over 224 weeks

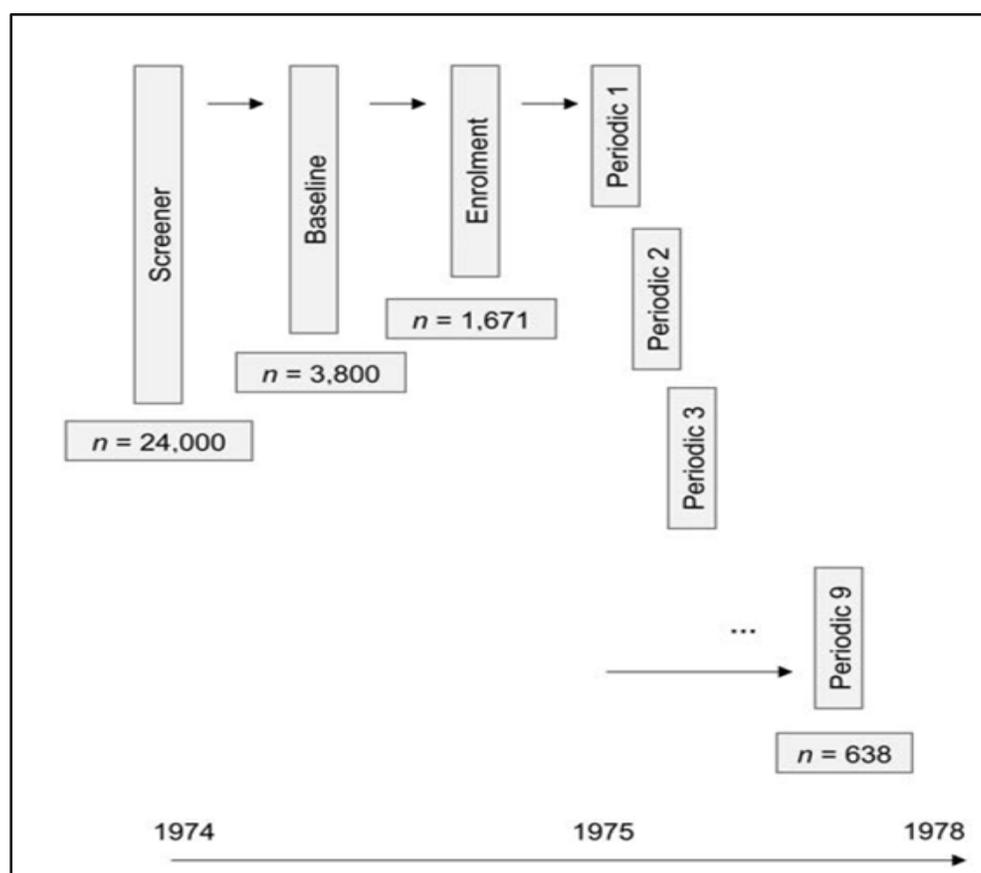


Figure 7 Survey structure for Mincome

Mincome samples shrink by about 30% during the panel period. Participants left the experiment for many reasons, but often because their earnings rendered them ineligible for benefits. While they could receive modest compensation for completing surveys, some elected not to once their payments from Mincome dwindled. Kurz (1978) expressed concern about the level of attrition and

initial refusals, especially for the saturation site, Dauphin, but finally did conclude that the Mincome data were usable.

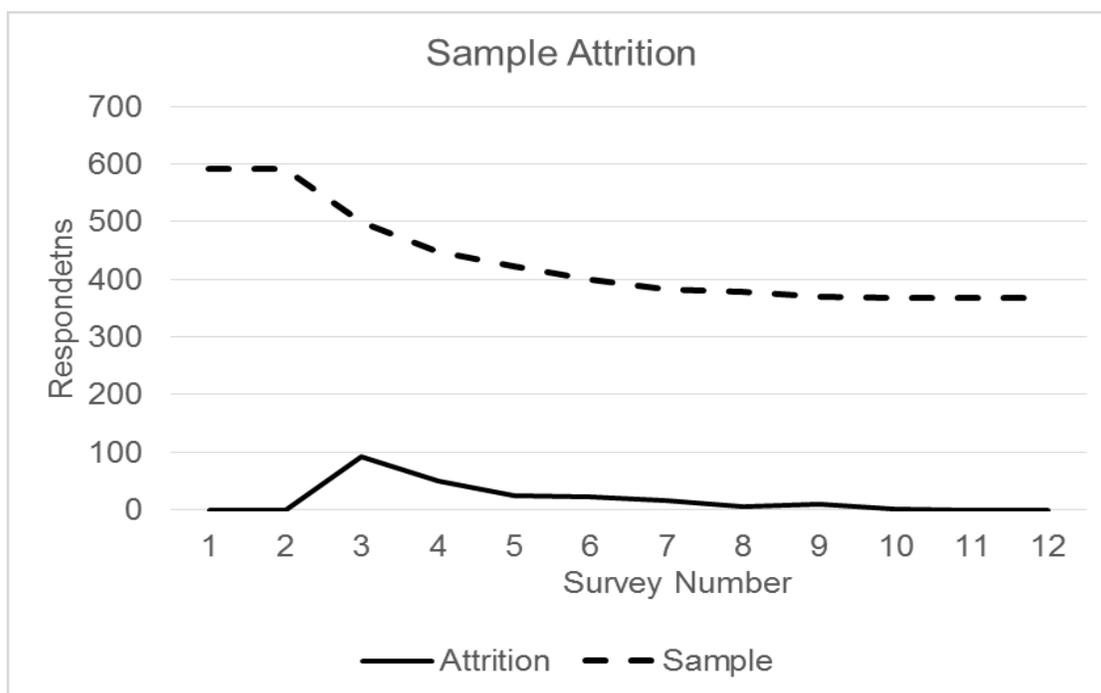


Figure 8: Typical Attrition in Mincome

A serious issue arose for Mincome at periodic 2. It appears that in the original sample development (Screener survey) a frame error resulted in under representation of a specific low-income cohort (Sabourin 1974). Further more, many respondents received unexpectedly high transfers (not social assistance) that rendered them eligible to anything but minimal payments. This blunted the experiment and raised serious concerns about the validity of the experiment. Accordingly starting with Periodic 3 a supplementary sample was initiated which explains why the surveys persisted beyond the originally planned nine periodic surveys. The relationship between the main and supplementary samples remains unclear.

The assignment allocation and attrition create non-random disturbances. In varying degrees, similar issues occurred for the other income maintenance experiments. Researchers have typically ignored

these issues and proceeded as if the resulting data were orthogonal. Since the estimation of corrective weights seems difficult, this approach is understandable.

A unique and important feature of Mincome was the administrative process. The sponsors (Canada and Manitoba) created a separate non-governmental non-profit entity to administer all aspects of the program including surveys, enrollment and payments. Mincome Inc. also became the face of government and prepared tax returns on behalf of the participants.

3.3. Ontario basic income pilot design features

OBIP emerged some 40 years after Mincome. It reflects a complete current view of basic income, where estimating the labour market response to a NIT was but one of several objectives. OBIP was intended to be an RCT, with two main sites (Hamilton and Thunder Bay) in Phase 1, and then with a saturation site (Lindsay) added in the second year (Phase 2).

An early problem occurred since those on social assistance (SA) have non-financial benefits. Health benefits (HB) were the most important. A key tenet of a basic income is that former SA recipients transitioning to the BI need to pay for a range of goods and services formerly covered. Everyone moving from SA to OBIP retained their supplementary health benefits which created a separation within the recipient group.

To illustrate a negative income tax, the Ontario Basic Income Pilot devised the scheme shown in **Table 4** and . A couple, assessed as disabled, is guaranteed slightly more than \$30,000, and will receive a declining payment until their earnings reach slightly more than \$60,000. OBIP is a tax free, but any other income was to be taxed at normal rates.

Table 4: Ontario Basic Income Pilot - NIT Thresholds		
	Single	Couple
Maximum Basic Income Amount	\$16,989	\$24,027

Maximum Basic Income Amount plus Disability Supplement for one person with a disability	\$22,989	\$30,027
---	----------	----------

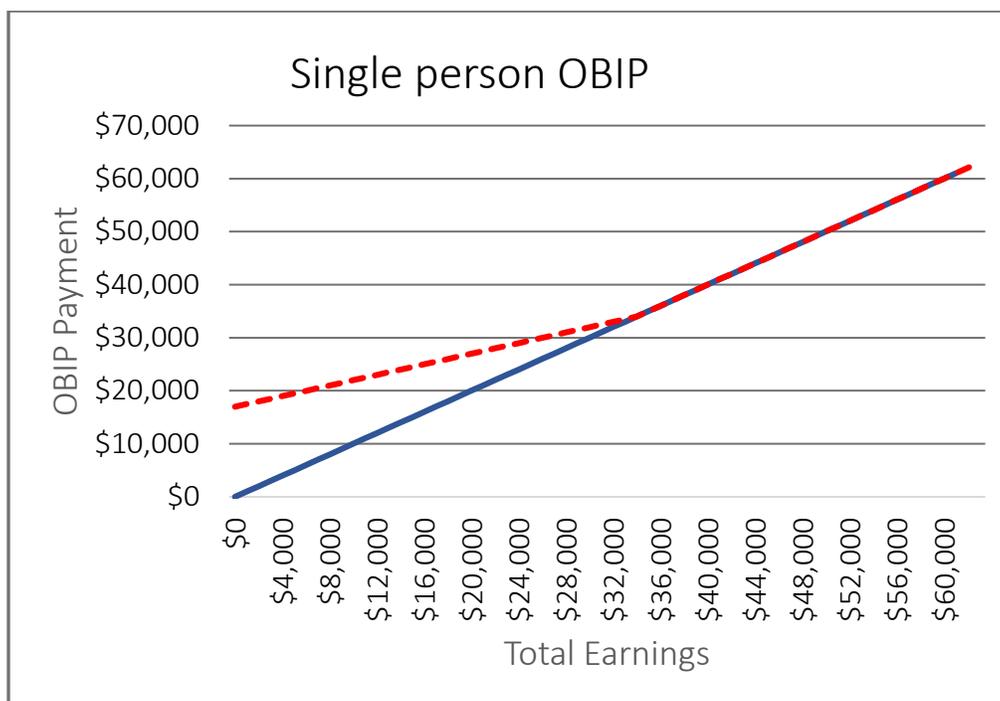


Figure 9 OBIP Payments

OBIP used a much simpler algorithm than Mincome for determining payments, with no consideration of net worth and no provision for children. Applicants with children could apply to the Canada Child Benefit Program, which offered a significant top up as shown in Table 5.

Table 5 Examples of support levels under OBIP

- A single individual, without a disability, earning \$28,000, will receive a BI payment of about \$2,989 to make total income to \$30,989.
- A couple, with one person disabled, and with part-time employment between the two of them generating \$16,009, will receive \$22,027 in BI with will leave them with a total income of \$38,027.
- A single parent with two children under 6 and no earned income, will receive a BI of \$16,989 plus the CCB of \$12,800 to reach a total income of \$29,789 tax free.

- In the case of the single individual earning \$28,000, the combined federal and provincial tax (on earnings) will be about \$3,500 reducing their after-tax income (with the BI on top) to about \$27,500

The original design plan for OBIP involved a simple RCT, with the 2 x2 payment scheme of **Table 4: Ontario Basic Income Pilot – NIT Thresholds** Table 5.

Table 6: Design matrix for OBIP			
Intervention		Comparison	
Single (WO HB)	Couple (WO HB)	Single (WO HB)	Couple (WO HB)
Single (W HB)	Couple (W HB)	Single (W HB)	Couple (W HB)
Single D	Couple D	Single D	Couple D
WO HB – no health benefits, W HB – health benefits on top of OCIP, D – disabled			

The OBIP had two phases

- Phase 1 (May 2017 – Dec 2017) Baseline study:
 - Test participation and sample maintenance (Hamilton and Thunder Bay)
 - Assess enrolment procedures
 - Validate questionnaire and survey logistics
 - Guide the development of Phase 2.
- Phase 2 (January 2018 - ?):
 - Inclusion of the saturation site – Lindsay
 - Generate periodic samples to support hypothesis testing
 - Outcome monitoring (semi-annually with cyclical themes)
 - Link to health records to assess impacts on mental and physical health
 - Projection of a final cost for province-wide rollout.

The Ontario government terminated the pilot in the summer of 2018, ostensibly because of cost, but also likely because of its opposition to the basic philosophy of a basic income. Many rued this

termination but the pilot encountered some significant problems that may have limited its capacity to test key hypotheses. These limitations included:

- Distorted sample frame development that failed to enumerate the eligible population

One might imagine, that is easy to select a sample of low-income households. After all, there are income tax and social assistance records. Three things undermine the use of these records in social experiments.

First, Canada Revenue Agency (CRA) has always been very careful about sharing tax information. For example, only in 2016 could Statistics Canada join income tax records to census information. Provinces that wish to use federal income tax data for program administration need to negotiate a data sharing agreement with CRA. This was not done prior to OBIP.

Second, while Ontario has access to records since it levies a personal income tax, conducting such an experiment was not within the mandate of the Ministry of Finance. A negative income tax is a tax program and needs to be delivered by a tax authority, namely the Ministry of Finance. When the Ministry of Finance could not accept responsibility for OBIP, the Ministry of Community and Social Services became the home for the pilot. In principle, this should have opened access to social assistance records but privacy barriers forestalled access to these data as well.

Third and most importantly, many individuals who would qualify for a basic income “fly under the administrative radar”. Those on social assistance and many low-income households do not file tax returns. So, sampling from income tax and social assistance records misses many potentially eligible participants.

- Flawed randomization process due to a convoluted and complex enrolment process.

OBIP to use an open enrollment process in the two main sites – Thunder Bay and Hamilton. The result was that enrollment relied on a letter sent to a general sample within the two test sites. This failed dramatically. First, after the privacy lawyers finished with the introductory materials, the invitation became long, legal, and impenetrable. It took more than a month to trim the introductory materials further delaying enrolment. Second, it always surprises planners, how many low-income

individuals fail to enrol in programs that would increase their financial well-being. Everyone involved was astonished at the low take-up of the pilot. Third, it required extensive support and re-contacting to secure tax and banking (for direct deposit) information from applicants to finalize their eligibility

The Pilot started mailing invitations in June of 2017 and by September, after mailing 37,000 invitations it had managed to enroll barely 150 participants, well short of the original target of 2000. This prompted a revised enrollment process that involved direct solicitation through community organizations, which after great effort did manage to raise enrolment. However, we are now a long way from a random allocation of participants into treatment and control groups.

- Collapse of the control group

OBIP sent conformations to participants that suggested the recipient would be receiving cheques. In fact, the province then informed a subset that they would not receive benefits but would receive \$50 for completing periodic interviews. Little wonder than many declined this offer, with the result that the control group never formed.

- Costly data collection processes

Even though most respondent had access to a smartphone, most respondents preferred to complete a printed questionnaire they needed to mail. Anyone familiar with survey methodology understands that such data collection requires persistent and costly follow-up to the point of harassment. A printed questionnaire with conventional mail back requires the most follow-up, as many as 10 attempts before the respondent is abandoned. Non-response is usually high in these situations.

- Multiple hypotheses require highly reliable and valid data and a strictly managed RCT.

The last question is probably the most important. The large-scale NIT experiments used sample designs that focused on core hypotheses about the labour market response of recipients. The researchers optimized the designs to detect the relatively small effects expected. Other hypotheses have been tested using these designs, but ideally each causal theme, such as the impact of a NIT on health outcomes requires a sample (treatment and controls) developed specifically for that set of

hypotheses. OBIP was attempting the test a range of hypotheses, where the effects were all likely to be relatively small. This requires high quality data, and with an RCT that is strictly maintained over time. OBIP was unlikely to achieve this and researchers would have had to resort to complex statistical procedures to extract causal relationships. This would have likely created more debate rather than answering key questions. Going back to how the results of modest labour market impacts derailed the major income maintenance experiments of the seventies, it is crucial that any future experiments of major social policy produce unambiguous results. Contested outcomes, where researchers engage in dueling op eds is not the way to move the basic income forward.

It is important to underscore that these criticisms of OBIP rest on Phase 1 results. It appears that Phase 2 was more successful in recruitment, but without any reporting of interim results from this phase, one can only speculate on the quality of the data and the integrity of the sample design. Specifically, would there have been an orthogonal control group? Would respondents have participated in follow-up surveys? Would respondents have agreed to link their survey responses to health records? Would respondents have allowed their tax information to the survey and health data? Would the cells of the design matrix (**Table 6**) retain sample over several survey cycles?

3.4. Methodological proposal for a basic income pilot

It is now possible to sketch the main parameters for a new basic income pilot, anchored around some core ideas

1. Data development should rely on administrative information

The most important principle is that administrative data must form the core of hypothesis testing. Specifically, tax and health records will support key hypotheses testing. In Canada, this means that the federal and provincial governments must co-sponsor a pilot, unlike OBIP where the provincial government remained the sole sponsor.

Such cooperation is essential to ensuring that tax records and social assistance case files can form the basis for random selection. Data sharing agreements and privacy issues require careful scrutiny to smooth legal impediments to using confidential information to anchor a pilot.

2. *Designating geographic sites should be avoided*

Mincome and OBIP designated specific sites. For both OBIP and Mincome, part of the rationale was to minimize data collection costs. Other reasons for selecting regional cities for OBIP was to avoid the overheated Toronto housing market, although it is unclear that those who would participate in a basic income are affected by runaway price inflation in Rosedale or Richmond. A specific geographic site also controls for local economic changes. More likely is that politicians like the idea of defining geography where funding will flow to create ribbon cutting opportunities. Political considerations were at play when the provincial selected Dauphin as the so-called saturation site.

The notion of a saturation site to test basic income impacts is especially suspect. Much of the rationale centers on rather obscure notions of community effects and the potential for such “isolated” communities to serve as a utopian petri dish. Finding such communities is quite difficult, since in a digital age, the concept of isolation, where migration is low, this tends to occur only in remote settlements with constrained economic opportunity. The generalizability of any results emerging from such settings would be very low.

Prior to OBIP designating Lindsay as a saturation site, Mincome was unique in the NIT experiments in creating Dauphin as a test site. It is remarkable that the Mincome is often referred to as the “Dauphin experiment”, when it was in many ways a complete failure. First the take-up rate was very low. At most perhaps as many as 50% of the eligible population was enrolled to receive benefits. Second, many recipients received other forms of income as the experiment progressed and the share of Mincome payments in the Dauphin “GDP” dwindled quite naturally. Third, and relate to the second point, the mid-seventies marked the first of a series of block buster sales of Canadian grain to Russia. As a regional agricultural centre, Dauphin residents and business could have benefitted differentially compared to the rest of Manitoba. In that context, it is entirely possible that observed change in employment, social, and health outcomes may have had little to do with Mincome payments and more related to standard economic developments.

If administrative data become the core of any future pilot, it makes sense to “marble” participants throughout the population.

3. *Conducting a “quiet” experiment*

The last two points lead to the creation of a “quiet” experiment. Media just loved to locate participants in these social experiments and feature case studies. This undermines measurement as existing participants become self-conscious and researchers, especially interviewers, communicate intent unwittingly. Large social experiments, like all experiments must be quiet and anonymous, if we are to extract valid parameters.

4. Multiple sub-experiments

In line with the use of administrative data, which lowers costs, rather than a single experiment, why not conduct several experiments. One can use tax records to assess labour market interactions and health records to assess health outcomes. This will mean trimming the range of hypotheses being tested, but it is far superior to conduct successful analysis on a few well-defined questions than to try to show-horn in every possible interaction within a single sample.

5. Surveys have a place

Some important questions will not be supported by administrative data alone. Placing the bulk of the data burden on administrative will allow surveys to become targeted and much shorter. This increases the likelihood of successful on-line survey management, thereby disciplining costs. The Mincome survey of 1 – 2 hours and hundreds of items and even the OBIP survey that typically required 30 minutes on the phone are just too long to maintain reliable and valid information over a series of panels.

6. Administrative structure

Mincome created a separate corporate entity to manage all aspects of the experiment. This became the face of government for recipients and aside from disbursing payments and managing the surveys, it prepared tax returns. This served to increase the accuracy of information collected in the periodic surveys. It also allowed government to fund the entity using annual allocations rather than incrementally extending budgets.

The danger is that the separate entity can become disconnected from the political realities of a major policy study. A critical error of Mincome Inc. was to focus on administrative and scientific issues, and not maintain and information flow of interesting results. Funders became impatient that the increasing expenditure had not resulted in any release of results.

4. Summary and conclusions

The basic income endures as major policy proposal, yet empirical social science has been unable to shed light on the potential outcome. The NIT experiments have been marked by a series of missteps. However, with the increased use of “big data”, it is possible to create a valid and lower cost evaluation process that could produce important insights to support a revolutionary anti-poverty policy.

Bibliography (*incomplete*)

- Autor, D. (2014). *Polanyi's Paradox and the Shape of Employment Growth* (No. w20485). Cambridge, MA: National Bureau of Economic Research. <https://doi.org/10.3386/w20485>
- Bernstein, R. N. (1975). Validity Issues in Evaluative Research: An Overview. *Sociological Methods & Research*, 4(1), 3–12. <https://doi.org/10.1177/004912417500400101>
- Burtless, G., & Hausman, J. A. (1977). *The effect of taxation on labour supply: evaluating the Gary negative income tax experiment* (Working Paper - Department of Economics No. 211) (p. 48). Massachusetts Institute of Technology.
- Calnitsky, D. (2016). More normal than welfare_The Mincome Experiment., Stigma, and Community Experience. *Canadian Research in Sociology*, 53(1).
- D'Ippoliti, C. (2018, October 1). The Political-Economy Fallouts Of Universal Basic Income Schemes • Social Europe. Retrieved October 3, 2018, from <https://www.socialeurope.eu/the-political-economy-fallouts-of-universal-basic-income-schemes>
- Forget, E. L. (2011). The town with no poverty: the health effects of a Canadian guaranteed annual income field experiment. *Canadian Public Policy*, 37(3), 283–305.
- Hausman, J. A., & Wise, D. A. (1977). Social experimentation, truncated distributions, and efficient estimation. *Econometrica*, 45(4).
- Hum, D., & Simpson, W. (2001). A guaranteed annual income: From Mincome to the millennium. *POLICY OPTIONS-MONTREAL*-, 22(1), 78–82.
- Huws, U. (2017). Universal Basic Income and Women's Liberation | Compass. Retrieved March 21, 2017, from <http://www.compassonline.org.uk/universal-basic-income-and-womens-liberation/>

- Lyall, K. C. (1975). Some Observations On Design Issues in Large-Scale Social Experiments. *Sociological Methods & Research*, 4(1), 54–76. <https://doi.org/10.1177/004912417500400104>
- Moffitt, R. A. (2003). The Negative Income Tax and the Evolution of U.S. Welfare Policy. *Journal of Economic Perspectives*, 17(3), 119–140.
- Pasma, C., & Mulvale, J. (2009). Income Security for All Canadians: Understanding Guaranteed Income. *Ottawa: Basic Income Earth Network Canada*. Retrieved from http://www.assembly.gov.nt.ca/sites/default/files/td_63-175.pdf
- Prescott, D., Swidinsky, R., & Wilton, D. A. (1985). Labour supply estimates for low-income female households using Mincome data. *Canadian Journal of Economics*, 86, 134–141.
- Thoits, P., & Hannan, M. (1979). Income and Psychological Distress: The Impact of an Income-Maintenance Experiment. *Journal of Health and Social Behavior*, 20(2), 120. <https://doi.org/10.2307/2136434>
- Widerquist, K. (2002). A Failure to Communicate: The Labour Market Findings of the Negative Income Tax Experiments and Their Effects on Policy and Public Opinion. *Basic Income European Network*, 52.
- Willows, N. D., Veugelers, P., Raine, K., & Kuhle, S. (2009). Prevalence and sociodemographic risk factors related to household food security in Aboriginal peoples in Canada. *Public Health Nutrition*, 12(08), 1150. <https://doi.org/10.1017/S1368980008004345>
- Wispelaere, J. D., & Stirton, L. (2007). The Public Administration Case against Participation Income. *Social Service Review*, 81(3), 523–549. <https://doi.org/10.1086/520939>
- Zelleke, A. (2005). Distributive justice and the argument for an unconditional basic income. *The Journal of Socio-Economics*, 34(1), 3–15. <https://doi.org/10.1016/j.socec.2004.09.051>