Design for "The causal effects of sustained unconditional cash transfers: Experimental evidence from two U.S. states" – Outcomes on Childbearing and Children

Sarah Miller[*] Elizabeth Rhodes[†] Alex Bartik[‡] David E. Broockman[§]
Eva Vivalt[¶]

PRELIMINARY

Abstract

The regular provision of unconditional cash transfers to individuals is a tactic to fight poverty that has attracted significant interest from researchers and policymakers. Despite this interest, many fundamental questions about the effects of receiving a sustained unconditional cash transfers remain. Open Research Lab, a nonprofit research organization, aims to help address this absence of data by conducting the U.S.'s first large-scale randomized trial of a guaranteed income. This document describes the design and analysis plan for the study. In the experiment, 1,000 participants will receive $1,000 per month for 3 years. A control group of 2,000 individuals who receive $50 per month will serve as the comparison group. The study offers an opportunity to inform both the debate over unconditional cash assistance and other questions about the effects of income that typically elude causal identification. This document focuses on the design of the study and outcomes we will measure related to children.

[*]Assistant Professor, University of Michigan Ross School of Business mille@umich.edu
[†]Research Director, Open Research Lab. elizabeth@openresearchlab.org
[‡]Assistant Professor, University of Illinois
[§]Associate Professor, University of California, Berkeley
[¶]Assistant Professor, University of Toronto

Contents

1 Introduction **2**

2 Existing Research **4**
2.1 Early experiments on unconditional cash transfers 5
2.2 Evidence from the Earned Income Tax Credit (EITC) 6
2.3 Natural Experiments 7
2.4 Unconditional Cash Transfers in Developing Countries 9
2.5 Recent Experiments 10

3 Sample Definition and Sampling Procedures **11**
3.1 Population 11
3.2 Sampling Frames 13

4 Recruitment and Randomization Procedures **16**
4.1 Recruitment to Eligibility Survey 16
4.2 Randomization 1: To In-Person Enrollment or Passive Monitoring 18
4.3 In-Person Enrollment 19
4.4 "Long Baseline" 21
4.5 Randomization 2: Treatment and Control Groups 22
4.6 Intervention 24
4.7 Outcome Measurement 25
4.8 Mobile Phone Application 27

5 Estimation **27**
5.1 Waitlist 28
5.2 Regression Adjustment to Increase Precision 28
5.3 Adjusting for Multiple Comparisons 28
5.4 Combining Data from Multiple Sources 29
5.5 Attrition 30
5.6 Heterogeneous Treatment Effects 30
5.7 Characterizing "Treatment" of Control Group Participants 31
5.8 Elicitation of Forecasts 31
5.9 Other Notes 32

6 Income and Child Development and Well-Being **32**

7 Outcomes **32**
7.1 Family 1: Birth Outcomes and Fertility 32
7.2 Family 2: Home and Neighborhood Environment 34
7.3 Family 3: Child health care use and access 35
7.4 Family 4: Child health 36
7.5 Family 5: Stress and social development 37

7.6 Family 6: Educational outcomes 37
7.7 Family 7: Parental Behaviors and Investments 38
7.8 Family 8: Amount and Quality of Non-parental Care 40
7.9 Heterogeneity Analysis 40

8 Conclusion **40**
8.1 Known Limitations 40

1 Introduction

Since the late 1960s, income inequality in the United States has risen dramatically and the share of income going to the bottom half of the income distribution has fallen by over a third (Piketty, Saez and Zucman 2019). Intergenerational mobility has fallen, wage growth has stagnated for all but the most skilled, and the official poverty rate remains essentially unchanged despite decades of robust economic growth (Chetty and Hendren 2018*a*;*b*; Congressional Research Service 2019; U.S. Department of Health and Human Services 2016). Individuals and communities are struggling as opportunities are increasingly concentrated in urban areas and among the highly skilled. These trends have increased political and social divisions (e.g., Dorn et al. 2016), and the ability of existing social programs to stem them is limited.

Research shows that the current social safety net leaves many Americans cycling in and out of poverty and/or categorically ineligible for aid (Shaefer and Edin 2013; Danziger 2010; Ben-Shalom, Moffitt and Scholz 2012). The patchwork of programs is complex, costly to administer, and difficult to navigate. Take-up rates are often low, particularly among those most in need (Bhargava and Manoli 2015; Finkelstein and Notowidigdo 2019). Due to the high marginal tax rates and eligibility "cliffs" introduced at moderate income levels, families who do find work often face a difficult trade-off between earnings and the benefits they rely on for survival.

In response to these challenges, policymakers at state and local levels around the country have become increasingly interested in exploring unconditional cash transfers as a solution. Research points to negative economic, social, and psychological feedback loops that keep individuals with-

out a steady income "trapped" in poverty. Sustained unconditional cash transfers seek to break these feedback loops. Interest in unconditional cash assistance has recently skyrocketed, but the debate often relies on conjecture, stereotypes, and studies that are out-of-date, have important methodological shortcomings, or were conducted in very different contexts. This lack of data and experience impedes rigorous policy analyses and data-driven political debate.

To help guide academic, policy, and political debates, we plan to conduct an experiment that will provide new evidence about the effects of sustained unconditional cash transfers in the U.S. We are collaborating with two non-profit organizations that will implement a cash assistance program. Our partners will recruit approximately 3,000 individuals across two U.S. states and randomly assign 1,000 in total to receive $1,000 per month for 3 years. We will conduct extensive quantitative measurement of outcomes related to individuals' economic, social, and physiological self-sufficiency and well-being, as well as gather data on how individuals use their time and money and how their receipt of monthly cash transfers impacts their children and those in their households. We are partnering with state and local government agencies and private entities to measure many outcomes with administrative data. A single study cannot answer all questions about the effects of a guaranteed income, but we view this experiment as the strong foundation for a broader research agenda moving forward.

The experiment also offers the opportunity to speak to policy debates about unconditional cash assistance programs. Most directly, the study will provide evidence that will inform debates about the design of public benefits, including whether benefits should be provided as cash or in-kind, whether they should be provided monthly or annually, and whether transfer programs should be extended to groups that they do not traditionally target (such as young adults without children). More broadly, the study will allow us to better understand the relationship between income, work, and well-being generally, and it can provide new evidence on the mechanisms underlying rich-poor gaps in policy-relevant outcomes such as education, health, and time use. For example, unearned income may relax liquidity constraints and facilitate investments in health, human capital,

or geographic mobility that may provide long-run returns to households. Unearned income may also change individual bargaining power with employers, landlords, family members, romantic partners, and others. Additionally, unearned income may reduce the cognitive burdens that may be created by scarce resources (Mani et al. (2013)), causing individuals to make different decisions. We discuss a broad array of additional channels through which unearned income may influence outcomes in subsequent sections.

2 Existing Research

Much of the existing literature on unconditional cash transfers in developed countries focuses on estimating effects on labor supply. Traditional economic theory predicts that unconditional cash transfers should cause individuals to work less (e.g., Becker 1965), while also consuming more of most goods. By providing nonwage income, cash transfers make household incomes less dependent on labor market earnings; this "income effect" allows households to consume more leisure. Based on this insight, much of the literature on unconditional cash transfers and welfare programs more broadly has focused on quantifying and understanding the determinants of income effects (Chan and Moffitt 2018).

Less work has been done measuring how unconditional cash transfers influence household consumption, which is the other impact of unconditional cash transfers predicted by traditional economic theory. Moreover, richer models suggest that unconditional transfers could have more nuanced effects than those predicted by traditional theory due to liquidity constraints, behavioral mechanisms, social interactions and spillovers, and other factors. More recent research has started to provide evidence on these broader effects of unconditional cash transfers.

In this section, we summarize this literature. Later, we go into more detail and characterize the contribution of this project relative to the existing literature for particular topics and outcomes.

2.1 Early experiments on unconditional cash transfers

To examine the effects of a negative income tax (NIT) on the labor supply of recipients, the U.S. government conducted four randomized experiments between 1968 and 1980, while the Canadian government sponsored one. A number of studies have aggregated the findings on reduced labor supply among participants across the four U.S. experiments, and these estimates range between a 5% and 7.9% reduction in the number of hours worked annually per individual for men; a 17% to 21.1% reduction for married women with children; and a 7% to 13.2% reduction for single women with children (Burtless 1986; Keeley 1981; Robins 1985).

The goal of the experiments was to examine the effect of a guaranteed income on labor supply, but supplemental analyses revealed positive effects on birth weight, homeownership, health, children's academic achievement, the number of adults pursuing continuing education, and other indicators of well-being (see, e.g., Hanushek et al. 1986; Widerquist et al. 2005; Murnane, Maynard and Ohls 1981; Weiss, Hall and Dong 1980; Rea 1977; Kehrer and Wolin 1979; Keeley 1980*b*;*a*; Baumol 1974; Maynard 1977; Elesh and Lefcowitz 1977; Maynard and Murnane 1979; Kaluzny 1979; O'Connor and Madden 1979). Similarly, a reexamination of Canada's guaranteed annual income experiment in the 1970s using health administration data shows a significant decrease in hospitalizations—particularly due to accident, injury, and mental health concerns—and an overall reduction in health service utilization among guaranteed income recipients relative to controls (Forget 2011; 2013). These overall improvements in health may lead to significant savings in health system expenditures.

Despite their path-breaking design, these experiments were plagued by nonrandom selection, errors in randomization protocols, differential attrition, nonparticipation, and systematic income misreporting, calling their results into question (Hausman and Wise 1979; Greenberg and Halsey 1983). Even without these empirical issues, the experiments were begun a half-century ago in a different economic and political context, so the results may not generalize to the present day. Moreover, the 1970s studies also did not track a number of outcomes that more recent research

suggests may play key mediating roles in the effects of unconditional cash transfers. The proposed study will employ research tools unavailable during the NIT experiments to generate a more holistic picture of the effects of the supplemental income on individuals. Tracking expenditures and financial data and leveraging a mobile application and web-based surveys to gather data on time use enable us to investigate how the cash transfers are spent and whether individuals are able to make investments that promote long-term economic self-sufficiency and build savings to help weather shocks and reduce vulnerability.

2.2 Evidence from the Earned Income Tax Credit (EITC)

The expansion of the Earned Income Tax Credit (EITC) in the early 1990s provided another opportunity to examine the effects of exogenous increases in income. Because it is linked to the amount earned, the EITC also affects beneficiaries' incentives to be employed and the number of hours worked, creating a substitution effect in addition to the income effect discussed above. Empirical research has suggested that the EITC increased labor force participation but had negligible impacts on hours worked (Eissa and Liebman 1996; Meyer and Rosenbaum 2001; Nichols and Rothstein 2016). Eissa and Hoynes (2004) show that while there is a positive increase in the labor supply of married men, the increase is more than offset by the reduction in labor force participation by married women, leading to an overall decrease in the total labor supply of married couples. There is ongoing debate about these estimates, however, as more recent analyses suggest that the observed effects on the extensive margin may be confounded by the simultaneous effects of welfare reform and a strong economy (Kleven 2018; 2020).

Additional research has investigated the effects of the EITC beyond measures of labor supply. By transferring money to lower-income households, the EITC substantially reduces the fraction of households in poverty. These gains are concentrated among families near the poverty level, however, and the EITC has little impact on those who are very poor (Meyer 2010). One analysis of maternal health before and after the expansion documented improvements in self-reported health

and mental health as well as reductions in the counts of risky biomarkers for cardiovascular diseases, metabolic disorders, and inflammation (Evans and Garthwaite 2014). Another EITC study found reductions in low infant birth weight that may be at least partially attributable to notable decreases in smoking during pregnancy and increases in prenatal care. More generally, the authors highlight that there are positive externalities to safety net programs that may lead policymakers to underestimate the benefits (Hoynes, Miller and Simon 2015). Other welfare reforms, such as Connecticut's Jobs First program, bundled multiple reforms together, making it difficult to determine the effects of individual components (Kline and Tartari 2016).

2.3 Natural Experiments

Unlike unconditional cash transfers, programs like the EITC affect beneficiaries' incentives to be employed and the number of hours worked because the amount of the benefit is linked to the amount of earned income. To address this limitation, several studies have examined the labor supply of lottery winners. Lottery studies generally find that the income effects of these transfers are modest. Using earnings data from the tax records of consenting Massachusetts lottery players, Imbens, Rubin and Sacerdote (2001) estimate that individuals with winnings up to $100,000 reduce their earnings from labor by about 11 percent of the exogenous increase in income provided by their prize. The effect is larger for individuals between 55 and 65, and the marginal propensity to earn actually increases for those with the lowest pre-lottery earnings, although the effect is not statistically significant.

In a study of Swedish lottery winners, Cesarini et al. (2016) also find negative effects on labor supply, though much smaller in magnitude than earlier studies. The authors report that pretax earnings decrease by approximately 1.1 percent of the payout amount per year, mainly due to a reduction in wages from working fewer hours. It is also important to note that, for lottery winners with a large lump sum or large monthly payments, negative effects on labor supply could also be attributed to higher marginal tax rates on wages. Furthermore, the lottery studies generally

either had small samples (Imbens, Rubin and Sacerdote 2001) or took place in policy contexts very different from the U.S. (Cesarini et al. 2016).

Other recent quasi-experimental evidence of responses to exogenous increases in income comes from examinations of the Alaska Permanent Fund and casino disbursements to Native American families in the U.S. The Alaska Permanent Fund provides an annual unconditional cash transfer to every resident of the state. In 2019, this transfer amounted to $1,606. Feinberg and Kuehn's (2018) analysis using data from the American Community Survey shows a negative effect of dividend receipt on hours worked. In contrast, Jones and Marinescu (2018) employ synthetic controls using data from the Current Population Survey and find no effect on the extensive margin and a small positive effect on the intensive margin. Available data was insufficient to determine if the latter is a result of people shifting from full to part time work or more people entering the labor force part time. A study of the effects of casino disbursements to Native American families found that a $4,000 annual increase in income per adult had no effect on parental labor force participation (Akee et al. 2010).

In addition to the effects on labor supply, some of the recent quasi-experimental papers have examined broader outcomes. Research on casino disbursements to Native American families finds that an average increase in annual household income of $1,750 is associated with statistically significant reductions in obesity, hypertension, and diabetes (Wolfe et al. 2012). Casino windfall cash disbursements have also been linked to higher achievement and educational attainment, reduced incidence of risk behaviors in adolescence, improvements in children's mental health, and better parent-child relationships (Akee et al. 2010; 2018; Costello et al. 2003). The Swedish lottery study found that winners consumed fewer mental health medications after winning, particularly those targeting anxiety (Cesarini et al. 2016). Though they did not report statistically significant changes in health service utilization and other indicators of health, the generalizability of the results to the U.S. context is questionable given the presence of universal health coverage and a generous social safety net.

2.4 Unconditional Cash Transfers in Developing Countries

There is also an important literature on cash transfers in a developing country context. Most of this work focuses on conditional cash transfers and children's outcomes (reviewed, for example, in Fiszbein et al. 2009). However, some studies leverage unconditional cash transfers and consider employment outcomes. Banerjee et al. (2017) review seven government-run cash transfer programs, plus Haushofer and Shapiro's evaluation of a Give Directly program in Kenya (2016), and find no systematic effect on labor supply on either the intensive or extensive margin.

One of the largest and most widely available of these recent cash transfer programs was the 2011 policy enacted in Iran that distributes the equivalent of 28% of the median per capita household income to over 70 million individuals. Despite the size of these transfers, no impacts were found on labor force participation (Salehi-Isfahani and Mostafavi-Dehzooei 2018). Individuals under thirty worked slightly less, though the effect was not statistically significant, and there were very small positive effects on labor supply for some groups (e.g., women and men in industrial and service sectors). These results may not generalize to the U.S., given the significant contextual differences.[1]

Other studies have focused on the impacts of cash transfers targeted at business owners or workers in particular industries (de Mel, McKenzie and Woodruff 2008; Blattman, Fiala and Martinez 2014; Fafchamps and Quinn 2017; McKenzie 2015). Schady and Rosero's (2007) analysis of data from an Ecuadorian unconditional cash transfer program reveals no impact on the labor supply of recipients. In a study of three-generation households in South Africa, Bertrand, Mullainathan, and Miller 2003 find a sharp decline in both the extensive and intensive margin in working-age individuals' labor supply when an older individual in the household receives a pension.

[1]There is also a large literature on conditional cash transfers in developing countries we do not review here.

2.5 Recent Experiments

More recently, there have been a growing number of conditional and unconditional cash transfer pilots in high-income countries. In the U.S., there have been two recent experiments with conditional cash transfers (CCTs) in New York City and Memphis, Tennessee, but results were mixed. The transfers reduced poverty and led to modest improvements in other areas that varied across sites, but researchers did not observe expected gains in academic achievement, employment, and health (Miller et al. 2016; Riccio and Miller 2016). However, a disproportionate amount of the cash rewards went to more advantaged families; in households that earned more rewards, parents had higher education levels and were more likely to be employed and married. There are a number of possible explanations for the lack of impact, including challenges with implementation, the complexity of the incentives, the process of documenting participation, and the small amount of money relative to the cost of living.

Finland recently piloted a basic income scheme targeted to those experiencing long-term unemployment. Two thousand unemployed individuals were randomly selected to receive 560 euros per month unconditionally for two years in lieu of traditional unemployment benefits. Final results are due in 2020, but no significant impacts were found on labor market participation in preliminary analyses (Kangas et al. 2019). It is important to note, however, that the control group was asymmetrically affected by changes to the unemployment system implemented in the middle of the experiment that require unemployment benefit recipients to prove they are looking for a job in order to continue receiving financial assistance. Though survey response rates were low, survey data indicated that basic income recipients experienced less stress, fewer symptoms of depression, and better cognitive functioning than the control group. Positive effects were also found on financial well-being, trust, and confidence in their future possibilities (Kangas et al. 2020).

3 Sample Definition and Sampling Procedures

3.1 Population

3.1.1 Eligibility Criteria

We define the population of interest as all individuals with Social Security Numbers between the ages of 21 and 40, inclusive, whose self-reported total household income in the calendar year prior to enrollment did not exceed 300% of the federal poverty level (FPL). In addition, we will exclude individuals that receive Supplemental Security Income (SSI) or Social Security Disability Income (SSDI), live in public housing or have a Section 8 voucher (also called Housing Choice Voucher) or other housing subsidy, and live in households in which another member receives SSI. Receiving an income supplement could jeopardize individuals' eligibility for housing assistance and SSI, and getting back on these benefits is very difficult and may take years. Losing this assistance could cause permanent harm, so these individuals will be excluded from the study.

3.1.2 Geography

The study will be conducted in regions in two states. Within each state, we chose a mixture of urban counties with large city centers, urban counties with medium-sized city centers, suburban counties, and rural counties.[2] We selected 1-5 counties of each type in each state that are demographically representative of counties of that type in the region. Nationally, roughly 19% of households that meet the eligibility criteria for the cash assistance program live in rural areas, 35% live in suburban

[2]Counties are divided into rural, suburban, small urban, medium urban, and large urban based on the share of households living in rural census tracts, the population density, whether the county is the largest in its metropolitan or micropolitan area, and population. Rural counties are those that have at least 50% of the population living in rural census tracts or population densities of less than 100 per square mile. Suburban counties are those that are not rural counties, but are not the largest city in their metropolitan or micropolitan area and have populations of less than two million. Small urban counties are those non-rural counties that are the largest in their micropolitan area but have urban cores of smaller than 40,000 people. Medium urban counties are those that are the largest in their metropolitan area, but have population densities of less than 10Expenses00 per square mile and populations of less than one million. Large urban counties are those that are the largest in their metropolitan area and have populations of at least one million or densities of greater than 1000 per square mile.

areas, less than 1% live in small urban areas, 17% live in medium-sized urban areas, and 28% live in large urban areas. Small urban counties make up a small share of the overall eligible population (less than 1%), so we excluded them from the sample. We aimed to recruit a sample that roughly matched these population shares, but we oversampled large urban areas to reduce recruitment and survey costs. This approach resulted in a sample of program participants composed of 13% individuals living in rural counties, 18% living in suburban counties, 16% living in medium urban counties, and 53% living in large urban counties.

3.1.3 Demographic Characteristics

In additional to the geographically stratified sampling described above, we used stratified random sampling to ensure that low-income individuals are over-represented in the sample of program participants and the share of males and females is approximately proportionate to their shares of the eligible population (which is roughly 62% female). Table 1 reports basic summary statistics of both eligible mailer respondents and enrolled program participants and compares both groups to the population mean characteristics computed using the American Community Survey for eligible households living in study counties. We report estimates of the eligible population both unweighted and reweighted to reflect the FPL group and county type stratification variables that were used.

On most dimensions, the characteristics of the sample closely match the eligible population in study counties. Our sample is slightly poorer, less likely to be Hispanic, and more likely to be female than eligible households as a whole. The biggest differences between our sample and the full eligible population are that our sample is more likely to report having a college degree and to be a renter than the eligible population.

Table 1: Study Sample Characteristics Compared to Eligible Population

	Eligible Populaton Comparison(ACS)				Study Sample		
	Full US Population			Study Counties	Eligible Mailer Respondents		Enrolled Active Survey Group
	Unweighted	Reweighted to Match Enrolled Sample FPL Distribution	Reweighted to Match Enrolled Sample FPL and County	Reweighted to Match Enrolled Sample FPL and County Type	Unweighted	Reweighted to Match Enrolled Sample FPL and County Type Distribution	Unweighted
	(1)	(2)	(3)	(3)	(4)	(5)	(6)
Panel A. Key active group stratification variables							
Income < 100% of FPL	0.25	0.37	0.37	0.37	0.38	0.37	0.37
Income 100-200% of FPL	0.36	0.39	0.39	0.40	0.31	0.40	0.40
Income 200% + of FPL	0.38	0.23	0.23	0.23	0.30	0.23	0.23
Rural County	0.26	0.22	0.13	0.13	0.15	0.13	0.13
Suburban County	0.32	0.34	0.18	0.18	0.18	0.18	0.18
Medium-Sized Urban County	0.16	0.18	0.16	0.16	0.18	0.16	0.16
Large Urban County	0.24	0.26	0.53	0.53	0.49	0.53	0.53
Panel B. Demographic Characteristics							
Any Children	0.59	0.60	0.59	0.62	0.62	0.59	0.58
HH Size	3.4	3.3	3.2	3.3	3.1	3.2	3.0
Age < 30	0.52	0.53	0.54	0.54	0.41	0.54	0.54
White (non-hispanic)	0.59	0.53	0.46	0.40	0.45	0.46	0.46
Black (non-hispanic)	0.17	0.22	0.25	0.30	0.30	0.27	0.30
Hispanic	0.17	0.18	0.21	0.25	0.19	0.23	0.22
Female	0.57	0.59	0.59	0.61	0.69	0.69	0.67
HH Income	36,204	29,822	29,549	30,158	28,715	28,297	28,800
College Degree or more	0.17	0.15	0.16	0.15	0.28	0.29	0.26
Renter	0.56	0.66	0.69	0.67	0.79	0.84	0.85
	919,395	904,792	904,792	35,086	14,708	14,708	3,000

Notes: This table compares the study sample to estimates of the characteristics of the study in the US as a whole. Eligible individuals are thoe ages 21-40 with household incomes of less than 300% of the federal poverty line. Columns (1) - (4) report estimates of the characteristics of eligible households using the American Community Survey (ACS) 2013-2017 pooled sample. Column (1) presents the unweighted means for eligible individuals, Column (2) reweights this sample to match the enrolled sample distribution of income groups as a share of the FPL (which was a stratification target when assigning individuals to the active survey group), Column (3) reweights the ACS sample to match both the income group distribution and the county-type distribution in the enrolled active survey group sample, and Column (4) presents estimates of characteristics of eligible individuals in sttudy counties, reweighted to match the enrolled sample FPL group and county type distribution. Columns (5)-(7) report characteristics of the study sample. Columns (5) and (6) report characteristics of eligible respondents to the mailer and online advertisement reruitment methods. Column (5) is unweighted, while Column (6) is reweighted o match the enrolled sample FPL and county type distribution. Column (7) reports the unweighted mean of the ultimate enrolled actve survey group (i.e. the 3000 individuals assigned to the active group who answered the baseline survey).

3.2 Sampling Frames

3.2.1 Address-based Sampling

The majority of the sample—approximately 87%—was recruited through mailers. We selected addresses in eligible Census tracts from Target Smart (`targetsmart.com`). This vendor appends commercial data on name, income, race, and other available information to addresses from a variety of state and commercial sources. We understand that the accuracy of these commercial data varies widely, but using the data for targeting significantly improved the efficiency and cost of recruitment in pilots of the mailing strategy. About 69% of mailers were targeted to individuals

who appear income and age eligible on the basis of these commercial data. We refer to these as the "targeted mailers".

To ensure that we did not systematically exclude from the sample individuals who are income and age eligible but did not appear as eligible in the commercial data (for example, because they moved or lost a job recently, they have missing or incomplete information in the commercial data, or they do not appear in any of the commercial data), the remaining 31% of the mailers were sent to addresses that were chosen randomly without regard to information from the Target Smart data. We refer to these as the "untargeted mailers." Where data on names was available, we randomly selected one name per household to whom to address the letter.[3] We appended "or Current Resident" to the end of each name.

We sent mailers to Census tracts roughly in proportion to their share of the eligible population within the county type in the region. For example, if a Census tract contains 2% of the eligible households in rural counties in a state, that county was sent roughly the number of mailers required to ensure that the tract represents 2% of the ultimate sample. The number of mailers this procedure required for each tract depended on the share of households in the tract that are eligible for the program, the targeting effectiveness of the commercial data, and the share of respondents we aimed to recruit using targeted versus untargeted mailers. Ultimately, we sent mailers to 1,138,130 unique addresses, making up about 23% of households in the average Census tract in the study.[4]

To identify the optimal mailing strategy and generate variation in selection into the study, we randomized both the number of letters sent to each address (ranging from one to four) and the gift card incentive offered for completing the online screening questionnaire, which ranged from \$0 to \$20. Roughly 2% of mailed households received one letter, 55% received two letters, 26% received three letters, and 17% received four letters. In terms of gift cards amounts, 37% of households received no gift card, 21% received \$5, 17% received \$10, 2% received \$15, and 23%

[3]For the "targeted" mailers and 50% of the "untargeted" mailers, we randomly selected one name per household among those names that appear age eligible in the commercial mailer data.

[4]The exact share varies with response and eligibility rates across different geography types.

received $20.

3.2.2 Alternative Recruitment Methods

In an effort to include in the sample participants selected differently from those who chose to respond to mailers, we employed two alternative methods to recruit the remaining 13% of the sample. First, the partner organizations purchased ads on the Facebook and Instagram platforms that were shown to all age eligible individuals located in program counties. Participants recruited through this method make up about 1 percent of study participants.

Second, the partners placed ads on the Fresh EBT platform. FreshEBT is a free mobile application developed by Propel (`www.joinpropel.com`) that allows Supplemental Nutrition Assistance Program (SNAP, also known as food stamps) recipients to check their balance and manage their benefits. FreshEBT has over 4 million users nationwide, including more than 180,000 active users in the program counties. The partner organization recruited app users in eligible zip codes by placing ads for the study within the app. Participants recruited through this method comprise roughly 12% of study participants.

3.2.3 Mitigating Spillovers Between Participants

We took three primary measures to reduce potential spillovers between study participants (either through direct interactions or through changing housing or labor market conditions). First, we sent mailers in 6 waves, composed of 0.4%, 9.5%, 19%, 25%, 20%, and 26% of the total mailers, spread out over 8 months. We stratified the number of mailers sent across each wave within a Census tract. This meant that, at most, 6% of households in the average tract received a mailer during any given mailer wave.[5]

Second, we capped the number of households we randomized into the program participation

[5]There are a few rural counties where we needed to send mailers to essentially all households within the county during the course of recruitment.

group at 2 for each Census block and 20 for each Census tract. This reduces the probability that participants in the program interact socially.

Third, prior to randomization into treatment and control, we conducted a survey of study participants to ask if they knew anyone else in the study and, if so, who that person was. Individuals who knew another person in the program were randomized in clusters with the other person(s) they knew in the study to avoid spillovers between people with different treatment status. For more details, see Section 4.5 below.

4 Recruitment and Randomization Procedures

4.1 Recruitment to Eligibility Survey

4.1.1 Mailers

The non-profit organizations implementing the cash assistance program first sent the mailers described above, informing individuals they may be eligible to participate in a new program in which participants receive "$50 or more" per month for three years. The mailers directed recipients to a website where they could register their interest in the program and complete a short eligibility screening survey. This screening survey collected demographic data that was used to verify eligibility for the program (e.g., household size and income to determine if respondents' incomes were below the cap, age, participation in public assistance programs). Respondents were also presented with an e-consent form to give the research team permission to access their administrative data. In order to facilitate linkages to administrative data, individuals who consented to share admin data had the option of providing their social security numbers during this process. Consent to share admin data was not a requirement for program participation, and it did not affect the probability of being selected for the program or randomized into the treatment group.

The partner organizations provided a phone number on the letter that people could call with

questions or to receive assistance accessing and completing the survey. Ultimately, 38,823 individuals responded to the mailers and completed the eligibility survey, of whom 12,745 were program eligible (33%).

4.1.2 Facebook and Instagram

As described above, each implementing partner organization purchased ads that appeared on Instagram and in the Facebook news feeds of users in all eligible counties who are predicted to be age-eligible for the program. The ads ran for 1-3 weeks and had varied levels of concentration, as measured by ad spending, by zip code group in each state; more money was spent on ads in zip code groups with the highest poverty rates.

The ad included a thumbnail picture of a calculator and a notepad with a list of monthly bills and text announcing a new program in which "Participants will receive $50 or more per month." Clicking a button that said "Learn more" directed respondents to a website hosted by each partner organization that included a brief description of the program, contact information for questions, and a link to complete the same online eligibility survey that mailer recipients completed.

4.1.3 FreshEBT

Also as described above, each implementing partner organization posted ads on the FreshEBT app to users in eligible counties. These notices ran for 1-2 weeks and advertised a "new financial assistance program" in which "selected participants receive $50 or more per month." When a user clicked the "Learn More" button, they were directed to a short form that collected their email address, phone number, age, and zip code. Age-eligible respondents who confirmed that they live in an eligible zip code were sent an email that provided instructions to complete the same online eligibility survey administered to individuals recruited through other methods.

4.2 Randomization 1: To In-Person Enrollment or Passive Monitoring

We then randomized individuals to be targeted for in-person enrollment or to remain in an "administrative data only" control group. Though individuals in the latter group will not participate in any research activities, their de-identified administrative data can be used for comparison on outcomes measured using these data.

Once we had a pool of eligible individuals, we blocked participants by demographics (age, gender, and race) and pre-treatment values of high-priority outcomes collected in the eligibility survey. We randomly assigned participants to the **"administrative data control"** or the **"program participation"** sample. To ensure that we met our demographic quotas[6] in the program participation group, we sent a larger number of mailers than required to reach our sample size and then randomly selected the program participation group to satisfy the demographic quotas. This means that participants had different probabilities of assignment to the "administrative data control." We include all eligible screener respondents who are not randomized into the program participation group in the administrative data control group, but we will reweight the administrative data control group to have the same demographic averages as the program participation group.

In total, 9,504 individuals were placed in the "administrative data control" group, of whom 55% consented to share their non-health related administrative data, yielding an admin control group of 5,266.[7] [8]

We plan to compare outcomes measured using administrative data for the administrative data control group to the control group enrolled in the main study (as described in Randomization 2

[6]There are three demographic quotas that we targeted for the sample. Specifically, we designed the randomization to ensure that i) the share of women in the sample resembles the share of women in the eligible population in study counties; ii) the sample is least 20% non-Hispanic White, 20% Black, and 20% Hispanic; and iii) the household income of at least 30% of the sample is 0-100% of the federal poverty level (FPL), the household income of at least 30% is 101-200% of FPL, and the household income of no more than 25% of the sample is 201-300% of FPL.

[7]Individuals in the admin control group are disproportionately in the middle and high income groups (with household incomes of 101%-200% and 201%-300% of the FPL) given the need to assign households with incomes of 0-100% of the FPL to the program participation group with higher probability in order to achieve our sample income group target goals.

[8]A smaller proportion, 51%, agreed to also share health related administrative data.

below). This comparison will reveal whether participation in the study and receipt of the $50 per month transfer had any effects on outcomes.[9]

4.3 In-Person Enrollment

The partner organizations then attempted to enroll individuals who had been randomized into the group targeted for in-person enrollment into the cash assistance program. As part of this enrollment, we administered the baseline survey to program participants who consented to take part in the research. We contracted with the University of Michigan Survey Research Center (SRC), a survey research firm with extensive experience fielding national studies, to manage recruitment and conduct in-person enrollment and baseline surveys. SRC employees aimed to ultimately complete 3,000 enrollments from the larger pool of possible participants. During the first 3 weeks of an attempted enrollment, interviewers made a total of 12 phone calls to primary and secondary phone numbers and sent follow up emails and text messages. The non-profit partner reached out to the individual at least once during week 4 if no contact had been made, and a different interviewer attempted 3 additional phone calls in week 5. If there had been no response after 6 weeks, we put contact on hold for two months before making another call and sending another text. If there was still no response, interviewers continued to call and text at least once per month until 3,000 participants had been enrolled.[10]

The in-person enrollment proceeded as follows:

- SRC staff first explained the purpose of the cash assistance program and the program pro-

[9]When conducting any such estimation, our estimand will be the average treatment on treated effect (ATT), weighting to the sample actually targeted for enrollment in the program. We had originally planned to conduct pooled analyses that estimated treatment effects by pooling our main analysis with an analysis that compared this "administrative data control" group to the treatment group that received the cash assistance. However, due to many participants having either very low or very high probabilities of assignment to the administrative data control group and the lower than anticipated take-up rate of the study among those assigned to the group targeted for in-person enrollment (due in part to COVID-19, which required enrollment to be done over the phone rather than in person), we do not plan to pursue this estimator for our final analysis. Our power calculations indicated that it would only increase our statistical precision by approximately 2%.

[10]Depending on response rates after the two-month break, interviewers in some cases attempted to reach individuals by visiting their home up to three times. In-person outreach stopped in March 2020 due to the COVID-19 pandemic.

cedures. Everyone was informed that they will receive "$50 or more" each month for three years and that the specific amount will be randomly assigned, but the fact that some participants will receive $1000 each month was not disclosed. This reduces the likelihood that the control group will know they are in the control group, as that knowledge may change their behavior in ways that would bias the results (including differential take up or attrition and a negative reaction to learning one is receiving less than others). Additionally, we did not want the prospect of a large cash transfer to coerce anyone into participating in the study.

- Individuals who agreed to participate in the program were enrolled in accordance with the procedures established by the non-profit organizations implementing the program.
- SRC staff then explained the purpose of the research and the study procedures.
- The explanation included the incentive structure for participation in research activities: $50 each for completing in-person baseline, midline, and endline surveys, $15 for each mobile baseline survey, $10 for each short monthly survey, and $10 per month for completing short activities on a mobile app. These incentives are taxable (unlike the cash assistance gifts), so we will send participants a 1099 if the participation incentive payments exceed $600 per calendar year, although we intend to keep incentives under the threshold.

During study enrollment, the enumerators:

- Obtained informed consent and contact information for friends and family that can help us locate the participant if we cannot reach them.
- Collected names and demographic information for other members of the household and a description of their relationship to the participant, to help document spillover effects.
- Helped the participant install the custom mobile app and showed participants how to use it, if the participant had a smartphone and consented to using a mobile app.

- Administered the first and most comprehensive baseline survey, including collecting biomarkers (height, weight, and blood pressure).

- Helped the participant set up direct deposit for the research incentive payments. If the participant already had a bank account, the interviewer logged in to a custom-built payments processing system and allowed the participants to verify their bank account information. If participants did not have a bank account, they were given the option of opening an account at Chime Bank, an online bank with no monthly fees, no minimum balance, and no overdraft fees. If they chose this option, they received a Visa debit card in the mail within 7 business days.

4.3.1 Changes to Enrollment in Response to COVID-19

Enrollment began in October 2019, and 1,317 individuals were enrolled and completed the in-person baseline survey by March 14, 2020. On March 15, 2020, the University of Michigan imposed restrictions prohibiting all in-person research activities in response to the COVID-19 pandemic. All outreach was suspended and no enrollments were conducted for approximately six weeks. During that time, we worked with SRC to make the necessary adjustments so that interviewers could enroll participants and administer the baseline survey over the phone. With the exception of biomarkers and the cognitive tasks, all other data could be collected over the phone. Enrollments resumed in late April and all remaining participants were enrolled remotely by October 6, 2020. Ultimately, 44% (1317) of enrolled individuals were enrolled via an in-person baseline survey and 56% (1683) were enrolled via phone.

4.4 "Long Baseline"

Enrollments took place over a 12 month period (the "long baseline"). During this time, random assignment to treatment had not yet taken place; all participants who had been enrolled were

receiving the control group cash assistance gift of $50 per month. In the month after a participant was enrolled, we administered three additional waves of web-based baseline surveys, notifying participants by text and email. These "mobile baselines" allowed us to collect data on outcomes that were not included in the in-person baseline. We also began distributing short web-based surveys each month that took approximately 10 minutes to complete. The purposes of these surveys are 1) to gather additional pre-treatment data to increase the precision of the estimates, and 2) to identify individuals likely to attrit from the study under the $50 condition.

The desire to identify participants likely to attrit is primarily driven by concerns over differential attrition. As previously noted, the 1970s NIT experiments were plagued by differential attrition. Differential attrition also seems likely *ex ante*; even though participants will continue receiving their $50 (in the control group) or $1,000 (in the treatment group) monthly payments regardless of whether they participate in all of the surveys, individuals receiving $1,000 per month may nevertheless be significantly more responsive than those receiving only $50. In case this differential attrition occurs, we hope we can identify a large subsample *ex post* that did not exhibit differential attrition, as defined by their *ex ante* responsiveness. For example, we might conclude: "We see differential attrition on average, but among those who answered at least 2 of the 3 pre-randomization baseline surveys, we do not." We will not, however, exclude any participants from randomization or change the probability of assignment to the treatment group based on whether they continue responding to surveys during the "long baseline."

4.5 Randomization 2: Treatment and Control Groups

After all 3,000 individuals had been enrolled, we randomly assigned them to the **"treatment"** ($1,000 per month) and **"program control"** (remain at $50 per month) groups.

We used blocked and clustered random assignment as follows:

1. *Clustering.* We first formed clusters of individuals based on information that a small num-

ber of study participants knew each other. We placed individuals who reported knowing each other into the same cluster, such that they would always receive the same treatment assignment.

2. *Selecting the Waitlist.* We next selected a stratified random sample of 300 individuals in each state to be placed in a waitlist group. Only individuals not in a cluster with other individuals were eligible for this waitlist group. Within this waitlist group in each state, we formed 10 blocks of 30 observations, blocking on a number of pre-treatment characteristics. We then placed the observations on the waitlist in order such that each 10 observations contained one randomly sampled observation from each of the 10 blocks.

3. *Blocking.* We next "collapsed" the data to the cluster level to conduct a cluster-level random assignment. (The vast majority of individuals are in a cluster of size one with no other observations, but around a dozen clusters were of size two or three.) We then formed blocks of clusters as follows. We first formed strata based on race/ethnicity, income group, and state; any clusters with more than one individual within them were placed in their own strata. Within these strata, we formed blocks of three based on several dozen pre-treatment covariates using the `blockTools` package in R. When the number of clusters in a strata did not evenly divide into three, there were either one or two leftover clusters in a strata after the first round of blocking. We then conducted a second round of blocking for these leftover clusters, again forming blocks based on a set of pre-treatment covariates using `blockTools`.

4. *Random Assignment: blocks.* Within each block of three, we selected one of three observations to be in the treatment group and placed the remaining two in the program control group. Given that the number of clusters did not evenly divide into three, within the final block we sampled from the vector $\{0, 0, 1\}$ without replacement to assign treatment within the final block.

5. *Random Assignment: waitlist.* After the first random assignment, we computed the number

of *individuals* (not clusters) in each state that had been placed in the treatment group. Because the clusters are not of equal size, the number of individuals placed in the treatment group during the first random assignment step varies by randomization. We then calculated how many remaining individuals N from the waitlist would need to be placed into the treatment group in order for 1/3 of each state to be in the treatment group. For example, our target was to place 501 participants in one state (1/3 of the 1503 enrolled) into the treatment group; if 401 participants had been randomly assigned to the treatment group in the first randomization, we would place 100 of the state's 300 observations on the waitlist into the treatment group.

Recall that the waitlist had already been placed in a random order within each state. To select the individuals on the waitlist that would be initially placed in the treatment group, we simply selected the top N individuals on the waitlist.

6. *Re-randomization.* After conducting a randomization, we conducted a series of balance checks across several dozen pre-treatment covariates. Each pre-treatment covariate was associated with a different p-value floor, with covariates we deemed to be more important assigned a higher floor. We rejected any randomization where the p-value on a t-test was below the p-value floor for any of the individual variables. We also conducted an F-test for the joint significance of all of the same set of pre-treatment variables by outcome area and rejected a randomization if the p-value on any of these F-tests was over 0.25.

Through simulation, we verified that this procedure resulted in all observations having an exactly 1/3 probability of being in the treatment group.

4.6 Intervention

After random assignment, participants in the treatment and control groups will be notified about the amount of the cash transfer they will receive each month and the schedule for disbursements.

The intervention in this study is an exogenous increase in income in the form of unconditional cash transfers. The transfers ($50 monthly for the program control group and $1,000 monthly for the program treatment group) will be delivered by the implementing non-profit organizations via direct deposit to the participants' bank accounts.[11] All participants will be notified monthly when the payment is deposited into their account.

Receipt of the treatment transfers and the nominal transfer for the control group is not conditional on participation in any of the research activities and individuals can use the money however they choose. Note that the transfers are provided as a gift from a non-profit organization and will not be subject to income tax.

4.6.1 Waitlist

Participants may not wish to receive the $1,000 per month transfer (e.g., because they do not feel comfortable taking money they did not "earn," or because it affects their eligibility for other benefits). During the first three months of the program, if any individuals assigned to the treatment group refuse the $1,000 per month transfer, we will go to the next person on the randomized waitlist in their state and offer that person the transfer instead.

4.7 Outcome Measurement

4.7.1 Monthly Surveys

We plan to use Qualtrics to conduct monthly web-based surveys. Participants will be notified by a text message and an e-mail containing a personalized link to the survey, and we will ask them to complete the questionnaire at their convenience within 2 weeks. We will send reminders to nonresponders, and $10 will be deposited to participants' bank accounts immediately upon

[11] The implementing partner organizations work with participants who do not have a bank account and who decline to or are unable to open a Chime account to ensure that they are able to receive direct deposits via a reloadable debit card or payment transfer app.

completion. We plan to keep the surveys very short to reduce fatigue.

Maintaining regular contact allows us to identify changes in employment, housing, education, and other variables for which a change will trigger an additional module asking about the reasons for the change and collecting new data on relevant measures (e.g., housing quality following a move, job satisfaction and earnings for new job, etc.). We will spread the modules to be administered less frequently across months to keep the length fairly consistent. Questions pertaining to variables with higher likelihood for measurement error or misreporting due to difficulty remembering will be asked more frequently.

If we see large differential attrition from these surveys, we may abandon them and focus on collecting data during the midline and endline surveys. However, we do see the monthly surveys as an important way to maintain contact with respondents, and response rates were very high (over 90%) throughout the pilots.

4.7.2 Midline Survey

The survey firm will administer an in-person midline survey 15-18 months after the treatment group begins receiving $1000 per month.

4.7.3 Endline Survey

The survey firm will administer an in-person endline survey towards the end of year 3, several months before the cash transfers will end. Respondents in the treatment group may behave differently during the last few months of the program in anticipation of the payments ending, so we will conduct this survey a bit early, starting at 2.5 years into the program and ending at least 3 months before the transfers cease. We hope to conduct long-run follow ups in the future after the program has ended to observe whether effects persist.

4.7.4 Administrative Data

We will gather a variety of administrative data which is described in more detail below.

4.8 Mobile Phone Application

Participants have the option to download a mobile phone application created for the study. We will use this mobile app for both passive and active data collection for consenting participants. We will administer 2-4 short activities each month through the app; participants who choose not to or are unable to download the app will be able to complete these activities via a web interface. From the subset of participants who consent to share anonymized location data, we will passively collect GPS location and accelerometer data from the participants' phones that we can connect to other data sources to potentially improve the precision of our estimates.

5 Estimation

To estimate treatment effects, we will compare outcomes for individuals who were assigned to the treatment group to individuals who were assigned to the "program control" group. In an alternative specification, we will take advantage of the fact that we collect repeated measures over time to analyze treatment effect dynamics. When necessary to combine multiple outcomes (e.g. to form indices) we will estimate models using seemingly unrelated regression. To be more specific, we will run one seemingly unrelated regression for each outcome in an index, and then test the joint hypothesis that the weighted sum of the treatment effects are zero while using robust standard errors. The weights will be determined as described below in the "combining data from multiple sources" section.

5.1 Waitlist

Within the waitlist group, we will follow the approach of (De Chaisemartin and Behaghel 2020). We will separately estimate the TOT of the $1,000 per month among the observations not in the waitlist. Finally, for our estimates, we will compute a precision-weighted average pooling the estimates for the waitlist group and for the observations not in the waitlist.

5.2 Regression Adjustment to Increase Precision

In general, we will compute regression-adjusted treatment effects using the procedures outlined in Bloniarz et al. (2016), using the LASSO to select baseline covariates to use for regression adjustment, then including the selected covariates in an OLS regression with the treatment indicator present. These OLS regressions with clustered standard errors will represent our main estimates and standard errors. For robustness, given the re-randomization process, we will also compute a set of standard errors by permutation, using 100,000 permutations that also passed our randomization criteria.

In some instances, we will be unable to merge our survey data with the administrative data outcomes for the TOT component of our estimator. In these cases, we will always include all of the pre-treatment values of the administrative data outcomes on the right hand side of our regressions unless otherwise specified.

We will present unweighted estimates for our primary results.

5.3 Adjusting for Multiple Comparisons

We will organize our outcomes at four levels:

1. **Topic**. E.g., political outcomes, health, time use, labor supply, geographic mobility, financial health, child outcomes, material hardship, cognitive, intrahousehold, psychosocial outcomes. One can think of each topic as representing one academic paper.

2. **Family**. This is the level at which we will conduct the multiple comparison adjustment. Therefore, each paper will make false discovery rate (FDR) adjustments within each family of outcomes in the paper. E.g., intergroup attitudes, political attitudes, political participation.

3. **Outcome**. Each family will have multiple outcomes. E.g., attitudes on social issues, attitudes on economic issues.

4. **Outcome Measures**. An outcome may be composed of multiple measurements. E.g., an economics attitudes index might be composed of ten different survey items about different economic issues.

We will categorize all outcomes into outcomes, families, and topics *ex ante*.

We plan to compute "sharpened" FDR-adjusted p-values that control the rate of false positives within the family of tests to be no more than the nominal level. We will use the procedure as outlined by Anderson (2008). We will report per comparison p-values in addition.

We will treat ordinal outcomes as continuous by default.

We will place secondary outcomes in separate families from primary outcomes and clearly label them as secondary.

5.4 Combining Data from Multiple Sources

Unless specified otherwise, in cases where we collect both midline and endline survey outcomes, we will combine the midline and endline outcomes to increase precision (McKenzie 2012); the main outcomes of interest will be a weighted average of the midline and endline outcomes, with 30% of the weight on the midline outcomes and 70% of the weight on the endline outcomes. We will also report the midline and endline results separately. We will also report a version that incorporates the data collected via the monthly or online surveys, collapsing the monthly surveys by pre/post midline and constructing a monthly survey outcome putting an analogous 30/70 weight on pre versus post-midline quarterly surveys. This set of monthly survey outcomes will then be

combined with the midline/endline aggregated outcomes with 30% weight being applied to the monthly surveys and 70% applied to the midline/endline surveys. Note that we will estimate all effects on individual $\times$ time period data (i.e., data will not be collapsed to the individual level). For outcomes collected at frequencies other than midline and endline (e.g., monthly), results will be reported by year unless otherwise noted in the PAP. For selected outcomes collected on a frequent basis, we will look at time trends; these cases will be specified in the discussion of the outcome measures.

5.5 Attrition

We will test for differential attrition from the surveys and, should this prove to be an issue, we will present a set of results correcting for differential attrition. We will check for balance in attrition rates using the same set of covariates that we used to test for balance at randomization. We are fortunate that we will have a variety of administrative data outcomes which are significantly less likely to be subject to substantial attrition.

We will conduct two-stage sampling for midline and endline data collection to minimize attrition-related bias by concentrating resources and efforts on a randomly chosen subset of the cases that are the most difficult to reach (and adding weights accordingly). We will also keep track of the number of contacts required to reach each participant for each survey. We will consider using the randomly assigned intensive follow-up and number of contacts required to reach each participant to construct attrition adjusted treatment effect estimates.

5.6 Heterogeneous Treatment Effects

Given the sample size and the many hypothesis tests we already plan to conduct, we are concerned about statistical power. Therefore we will pre-register that all heterogeneous treatment effect estimates will be considered exploratory unless explicitly pre-specified otherwise. PAPs for some

outcome areas may specify hypothesis tests for heterogeneous treatment effects and note them as exploratory or non-exploratory.

5.7 Characterizing "Treatment" of Control Group Participants

Not all eligible respondents who complete the online eligibility screener will be randomly selected to participate in the program and study. As a result, we have access to an additional "control" group of individuals who consented to passively provide administrative data but will not be contacted by the research team. Using this "administrative control" group can help us shed light as to whether the program has any effects on the "program control" group, either as a result of the $50 monthly payments, the survey incentives, or the act of completing surveys themselves. We will use this group to characterize any such effects on outcomes measured using administrative data that might be present in the program control group.

5.8 Elicitation of Forecasts

We will be eliciting forecasts for several key outcomes on the Social Science Prediction Platform. We expect to receive forecasts from other researchers, those working in policy or non-profit organizations, and the general public. These forecasts can help in gauging the novelty of our results. There are not currently standard ways of presenting comparisons of *ex ante* forecasts with research results, but we anticipate including some comparisons, if only in an appendix. In comparing our research results to the *ex ante* forecasts, we will focus on comparing our results to the predictions of researchers in economics unless otherwise specified. The outcomes that we will forecast are indicated with an asterisk in the section on outcomes below.

5.9 Other Notes

The survey questions and analyses described here are contingent on securing sufficient funding to gather the requisite data.

6 Income and Child Development and Well-Being

A large and growing literature has shown a strong causal relationship between parental resources and child outcomes. The effect of parental income emerges at birth and appears persistent throughout childhood, with implications for adult well-being, economic security, and health. Future work will incorporate the impact of the transfers on children's involvement with the criminal justice system and longer-run outcomes. These outcomes will be described in a separate pre-analysis plan. Outcomes related to children will be conducted just using children in the household at the time of the transfer, to avoid issues of selection if the treatment itself affects childbearing.

7 Outcomes

7.1 Family 1: Birth Outcomes and Fertility

A strong correlation between health at birth and family socioeconomic status has been documented across several contexts (e.g., Hoynes, Miller and Simon 2015; Strully, Rehkopf and Xuan 2010; Currie 2011; Case, Lubotsky and Paxson 2002; Buckles 2018; Blumenshine et al. 2010; Aizer and Currie 2014), but relatively few studies provide insight into whether this relationship is causal and potentially mediated by policy interventions to increase income. This is a critical gap in the literature as we know from an extensive research in economics, public health, and epidemiology that health at birth has outsized effects on adult health and achievement (Currie and Almond 2011). One exception is a quasi-experimental evaluation of the Earned Income Tax Credit (EITC), which

found that birth outcomes, such as birthweight and the incidence of low birthweight, improved after eligibility for the EITC expanded (e.g., Hoynes, Miller and Simon 2015).

Given the strong relationship between health at birth and lifelong health and achievement, the health of any children born to program participants is a natural outcome of interest for this study. However, we anticipate that sample sizes may be too small to be well-powered for this analysis. We therefore will only proceed if there are at least **100** births to women who are either in the treated group or partners of individuals who are treated over the 3-year period of the study.[12] Given that the current fertility rate is 62.5 births per 1000 women of childbearing age, a back of the envelope calculation suggests 180 births will occur over our sample period.

In addition to birth outcomes, the receipt of basic income could directly affect fertility decisions. Fertility decisions have been shown to be sensitive to economic conditions such as house prices and macroeconomic factors (e.g., Dettling and Kearney (2014), Kasey Buckles and Lugauer (2017)). If basic income results in different women selecting into fertility, analysis of birth outcomes alone may generate misleading results. We therefore plan to evaluate the effect of basic income both on fertility decisions directly and the selection into fertility by mothers with different characteristics.

Our outcomes will be grouped in the following way. Some fertility and health at birth measures will be derived from linked birth certificate records, whereas the measure of composite health of mothers will be measured from survey instruments described in the Health PAP:

- Overall fertility: number of births.*
- Health at birth: birth weight, length of gestation, indicator that child is pre-term, low birthweight, or very low birthweight, presence of medical conditions at birth, and APGAR score from birth certificate.
- Pre-birth health interventions: first prenatal visit, number of prenatal visits during pregnancy

[12]Regardless of births during the study period, we plan to conduct longer-term analysis of birth outcomes for participants over a longer follow-up period. These will be detailed in a separate pre-analysis plan.

- Childbearing intention: Are you IF MALE insert ”and a partner” intending to become pregnant in the next year?” if yes: ”On a scale of 1-10, how much do you want to become pregnant with a partner in the next year? if no: ”On a scale of 1-10, how much do you want to avoid becoming pregnant with a partner in the next year? We will use these questions to form a scale of the desire or intention to have children.

- Use of contraceptives (among those who indicate a score of 5 or higher on desire to avoid pregnancy): an indicator that they use contraception, and the method's efficacy, measured as one minus the CDC's published failure rate with typical use for the method used most often.

Fertility and health at birth will be measured using individually-linked birth certificate data. We will also explore how the composition of mothers changed based on pre-randomization characteristics such as BMI, income, and educational attainment.

7.2 Family 2: Home and Neighborhood Environment

The second set of topics considered will involve directly measuring improvements in material well-being for the children of those in the treatment group. This family will contain the following outcomes:

- USDA child food security survey: this battery of food security questions was developed as a screening tool. We will use a modified (shortened) version of this tool, and score it following guidelines from USDA.[13]

- Home Environment CHAOS Scale (Matheny et al. (1995)). Given constraints in survey space, we will use a shortened version of this scale.

[13] See https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/survey-tools.aspx#youth for more details.

- Neighborhood environment: child opportunity index (http://diversitydatakids.org/child-opportunity-index), area deprivation index (https://www.neighborhoodatlas.medicine.wisc.edu/), and exposure to air pollution.

7.3 Family 3: Child health care use and access

Parental resources may affect a child's ability to use the health care system. To evaluate the impact of basic income on children's use of health care, we will survey parents with the following questions:

- Medical Care Access: During the past 12 months, was there any time when a child in your household needed medical care, but did not get it because you couldn't afford it or take time off from work?,
- Use of care: "Is child current on all required vaccinations?", "About how long has it been since you last saw or talked to a doctor or other health care professional about child's health? Include doctors seen while a patient in a hospital.", "About how long has it been since child last saw a dentist? Include all types of dentists, such as orthodontists, oral surgeons, and all other dental specialists, as well as dental hygienists."

In addition to these survey questions, we will also use linked hospital/emergency department administrative data to evaluate each child's use of emergency department and hospital services. These will only be available for children with social security numbers; children without social security numbers at baseline will be excluded from the analysis. We will examine the following outcomes based on administrative data; these measures will be reported separately from the survey outcomes. We will only conduct this analysis if these data are available and we receive permission to pursue the linkages.:

1. Hospital care use: a) # of hospitalizations in the last 12 months and the # of "preventable"

hospitalizations, based on AHRQ classification algorithm b) total number of hospital days in the last 12 months, c) total hospital charges in the last 12 months.

2. ED care: a) # of ED visits in the last 12 months, b) total ED charges.

3. # of ED visits for non-urgent, urgent but primary-care treatable, and urgent but preventable conditions, as assigned using the NYU ED Algorithm.

If we observe fewer than 10 hospitalizations for children in the treated group, we will consider all hospital usage outcomes exploratory. If we observe fewer than 10 ED visits for children in the treated group, we will consider all ED use outcomes exploratory.

7.4 Family 4: Child health

It is widely-known that family income is positively associated with childhood health (see, e.g., Case, Lubotsky and Paxson (2002)), although the direction of the causal effect remains unclear. Recent work Cesarini et al. (2016) finds significant reductions in obesity rates at age 18 among children whose parents were assigned additional wealth via a lottery relative to those with smaller winnings but no effects on other measures. We will estimate the impact of basic income on child health using the following survey-elicited measures.

- Self-reported health: parent's report of the child's health on a 5-point scale.

- Whether the child has limitations preventing them from participating in usual activities, and (if such a condition exists), the degree to which it limits their activities.

- Child's BMI as reported by parent, and an indicator BMI is in the obese range (exceeding the 95th percentile for age/sex according to the CDC growth charts).

To improve interpretability we will also report results for an indicator that the child's health is very good or excellent.* This binary variable will not itself be inputted into the index.

7.5 Family 5: Stress and social development

We will also investigate the relationship between the income payments and a child's stress and social development.

- For parents with children from ages 5 to 17, we will ask parents to complete the PROMIS proxy psychological stress experience questionnaire (Bevans et al. 2013) to measure the child's stress level.
- We will assess social and emotional development for children ages 2 to 17 using the Strengths and Difficulties questionnaire to measure the child's emotional state, conduct problems, hyperactivity/inattention, peer relationship problems, and prosocial behavior (Goodman 2001).

For this family, we will examine heterogeneity in the effect by the age of the child at the time the basic income payment starts. We will consider children under age 5, age 5-10 and over age 10 separately.

7.6 Family 6: Educational outcomes

We will leverage available administrative data to evaluate outcomes related to school performance. We are currently working with local partnerships to develop data use agreements surrounding these outcomes, and outcomes will be based on data availability. We anticipate that these outcomes will include: standardized test scores for English/reading and math for grades 3-8 (STAAR in Texas, and the Illinois Assessment of Readiness in Illinois); end-of-course test results in grades 9-12 in Texas and PSAT/SAT scores in grades 9-11 in Illinois; disciplinary actions such as suspensions and expulsions; grade retention; absences. The standardized tests in Texas and Illinois are different from one another, but we will use a mixture model to determine how to best pool standardized effects. This is a key benefit of these types of models; they can estimate how much to pool different variables. As we do not anticipate many children in any given grade, we will consider results

within a grade to be exploratory, but also present the results of an analysis that partially pools across grades. In situations where we have both administrative and survey based measures of the outcome, we will rely primarily on the administrative measure as long as it is available for participants at both sites.

In addition to the administrative records, we will elicit information on whether the school asked the parent to come in and talk about problems with schoolwork or behavior. We will also ask via survey the following questions:

1. Enrollment and age for grade: Enrollment status of children 5-17, has child 5-17 ever repeated a grade in school.

2. Attendance: Number of days child was absent from school in the most recent complete school year.

3. Disciplinary or behavioral issues: has child been suspended or expelled from school, even for 1 day, in the past year, Has school asked someone to come in and talk about problem child 5-17 was having with schoolwork or behavior in past 2 years, Has child 5-17 gone to a special class or school or gotten special help in school for behavioral or emotional problems in past 2 years, Has child 5-17 gone to a special class or school or gotten special help in school for learning problems in past 2 years.

4. Performance: Parent's report of child's grades in school.

5. School quality: we will us numeric rating for schools attended by children based on the GreatSchools.org website, which we will validate with state statistics if available.

7.7 Family 7: Parental Behaviors and Investments

We will measure the quality of child-parent interactions using several measures.

- Parental satisfaction: On a 5 point scale, do parents agree or disagree with the following statements: "When it comes to raising my child/children, I feel alone most of the time," "I get as much satisfaction from parenting as other parents do," "I spend a great deal of time with my children."

- Quantity of parental interaction: The number of times in the last week the parent ate dinner with their child and put their child to bed. For children under 5, we will ask parents how often they or someone in their household takes the child on an outing (such as to the park or to visit a friend), how often they or someone in their household plays with the child, and how often they or someone in their household reads to the child, how much time is typically spent helping with homework. From the time diaries, we will also include number of hours spent on childcare on average. We will also ask whether or not a parent has gone to a general meeting at the child's school (like a back-to-school night), gone to a school event (such as a play or sporting event), attended a parent-teacher conference, volunteered at the child's school, worked with a youth group, sports team, or club outside of school in the past year.

- Parental anxiety and stress: parental stress using the Perceived Stress Scale (Cohen, Kamarck, and Mermelstein 1983), parental anxiety as measured with our anxiety screen questions, and parental psychological distress as measured on the Kessler 6. See Health PAP for more details.

- Expenditures on children overall. We will also report separately expenditures in the following categories (considered exploratory and not separate outcomes within the family): food/clothing, entertainment (toys, games, etc), education and enrichment (books, tuition, tutors, etc), activities (sports, lessons). The Income, Expenditures and Financial Health PAP contains this outcome as part of an analysis of other expenditures.

7.8 Family 8: Amount and Quality of Non-parental Care

We will also investigate the quality of non-parent care received by the child. Specifically, we will ask about:

- The parent's perceived quality of the non-parental care: level of satisfaction with child care, an indicator that the child was switched to more reliable childcare, an indicator the child was switched to higher quality child care.
- Stability of the care: number of non-parental caregivers within past month, number of child care arrangement changes in past year.
- Number of hours of non-parent childcare received by children under age 5.

7.9 Heterogeneity Analysis

In analysis of fertility, we will examine heterogeneity based on whether the respondent reported wishing to have more children in the future at baseline. For birth outcomes, we will examine heterogeneity based on income at baseline (less than 100% FPL vs 100% FPL and above). In analysis of children's outcomes, we will examine heterogeneity based on income at baseline (less than 100% FPL vs. 100%FPL and above), child age groups, and child's gender.

8 Conclusion

8.1 Known Limitations

Our study has several limitations. First, the limited nature of the RCT does not permit us to simulate the macroeconomic conditions of the government introducing an unconditional cash transfer program to all residents of the United States who meet broad eligibility criteria. If recipients are spending the money helping friends and family who would receive their own cash transfer under

the policy, the treatment is diluted and the likelihood of the hypothesized effects is undermined. Similarly, the dispersed sample precludes our ability to capture the multipliers and general equilibrium effects identified in the theoretical literature and observed in studies in developing countries. The dispersed study also precludes studying the effect of sustained unconditional cash transfers on cultural attitudes towards work and other social spillovers. Despite these limitations, we selected a geographically dispersed population for several reasons. Most importantly, the intervention is very expensive and our sample size is constrained by the budget. A geographically saturated study would likely cost billions of dollars, and we would not have enough statistical power to detect effects with a geographically saturated study with our budget.

A second limitation is the time-bound nature of our treatment. The 3-year timespan of the intervention is obviously not the same as a perceived long-term guarantee, and individuals may behave differently knowing that the transfers are time-limited (Hoynes and Rothstein (2019)). Nevertheless, a study at the scale proposed in this analysis plan will allow us to provide timely evidence to inform ongoing policy debates and future research on this topic.

References

Aizer, A and J Currie. 2014. "The intergenerational transmission of inequality: Maternal disadvantage and health at birth." *Science* 344(6186):856–861.

Akee, Randall K Q, William E Copeland, Gordon Keeler, Adrian Angold and E Jane Costello. 2010. "Parents' Incomes and Children's Outcomes: A Quasi-Experiment Using Transfer Payments from Casino Profits." *American Economic Journal Applied Economics* 2(1):86–115.

Akee, Randall, William Copeland, E Jane Costello and Emilia Simeonova. 2018. "How Does Household Income Affect Child Personality Traits and Behaviors?" *American Economic Review* 108(3):775–827.

Anderson, Michael L. 2008. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American statistical Association* 103(484):1481–1495.

Banerjee, A, R Hanna, G Kreindler and B A Olken. 2017. "Debunking the Stereotype of the Lazy Welfare Recipient: Evidence from Cash Transfer Programs." *The World Bank Research Observer* 32(2):155–184.

Baumol, William J. 1974. *The Journal of Human Resources* 9(2):253–264, title = An Overview of the Results on Consumption, Health, and Social Behavior.

Becker, G S. 1965. "A Theory of the Allocation of Time." *The Economic Journal* 75(299):493–517.

Ben-Shalom, Y, R Moffitt and J K Scholz. 2012. An assessment of the effectiveness of anti-poverty programs in the United States. In *The Oxford Handbook of the Economics of Poverty*, ed. P Jefferson. New York: Oxford University Press pp. 709–749.

Bertrand, M, S Mullainathan and D Miller. 2003. "Public Policy and Extended Families: Evidence from Pensions in South Africa." *World Bank Economic Review* 17(1):27–50.

Bhargava, Saurabh and Dayanand Manoli. 2015. "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment." *American Economic Review* 105(11):3489–3529.

Blattman, C, N Fiala and S Martinez. 2014. "Generating Skilled Self-employment in Developing Countries: Experimental Evidence from Uganda." *The Quarterly Journal of Economics* 129(2):697–752.

Bloniarz, Adam, Hanzhong Liu, Cun-Hui Zhang, Jasjeet S Sekhon and Bin Yu. 2016. "Lasso ad-

justments of treatment effect estimates in randomized experiments." *Proceedings of the National Academy of Sciences* 113(27):7383–7390.

Blumenshine, P, S Egerter, C J Barclay, C Cubbin and P A Braveman. 2010. "Socioeconomic Disparities in Adverse Birth Outcomes." *American Journal of Preventative Medicine* 39(3):263–272.

Buckles, Kasey S. 2018. Maternal Socioeconomic Status and the Well-Being of the Next Generation(s. In *The Oxford Handbook of Women and the Economy*, ed. Susan L Averett, Laura M Argys and Saul D Hoffman. Oxford: Oxford University Press.

Burtless, G. 1986. The work response to a guaranteed income: A survey of experimental evidence. In *Lessons from the income maintenance experiments*, ed. A H Xavier. Boston, MA: Federal Reserve Bank of Boston and The Brookings Institution pp. 22–54.

Case, Anne, Darren Lubotsky and Christina Paxson. 2002. "Economic Status and Health in Childhood: The Origins of the Gradient." *American Economic Review* 92(5):1308–1334.

Cesarini, David, Erik Lindqvist, Robert Östling and Björn Wallace. 2016. "Wealth, health, and child development: Evidence from administrative data on Swedish lottery players." *The Quarterly Journal of Economics* 131(2):687–738.

Chan, Marc K and Robert Moffitt. 2018. "Welfare Reform and the Labor Market." *Annual Review of Economics* 10(1):347–381.

Chetty, Raj and Nathaniel Hendren. 2018*a*. "The Impact of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *Quarterly Journal of Economics* 113(3).

Chetty, Raj and Nathaniel Hendren. 2018*b*. "The Impact of Neighborhoods on Intergenerational Mobility II: County-Level Estimates." *Quarterly Journal of Economics* 113(3).

Congressional Research Service. 2019. "Real Wage Trends, 1979 to 2018." CRS Report.

Costello, E Jane, Scott N Compton, Gordon Keeler and Adrian Angol. 2003. "Relationships Between Poverty and Psychopathology: A Natural Experiment." *Journal of the American Medical Association* 290(15):2023–2029.

Currie, Janet. 2011. "Inequality at Birth: Some Causes and Consequences." *American Economic Review* 101(3):1–22.

Currie, Janet and Douglas Almond. 2011. Human capital development before age five. In *Handbook of Labor Economics*, ed. David Card and Orley Ashenfelter. Vol. 4B Elsevier B.V. chapter 15, pp. 1315–1486.

Danziger, Sandra K. 2010. "The Decline of Cash Welfare and Implications for Social Policy and Poverty." *Annual Review of Sociology* 36(1):523–545.

De Chaisemartin, Clement and Luc Behaghel. 2020. "Estimating the effect of treatments allocated by randomized waiting lists." *Econometrica* 88(4):1453–1477.

de Mel, S, D McKenzie and C Woodruff. 2008. "Returns to Capital in Microenterprises: Evidence from a Field Experiment." *The Quarterly Journal of Economics* 123(4):1329–1372.

Dettling, Lisa and Melissa Kearney. 2014. "House Prices and Birth Rates: The Impact of the Real Estate Market on the Decision to Have a Baby." *Journal of Public Economics* 110:82–100.

Dorn, David, Gordon Hanson, Kaveh Majlesi et al. 2016. Importing political polarization? the electoral consequences of rising trade exposure. Technical report National Bureau of Economic Research.

Eissa, N and J Liebman. 1996. "Labor Supply Response to the Earned Income Tax Credit." *Quarterly Journal of Economics* 111:605–637.

Eissa, Nada and Hilary Hoynes. 2004. "Taxes and the labor market participation of married couples: the earned income tax credit." *Journal of Public Economics* 88:1931–1958.

Elesh, D and M J Lefcowitz. 1977. The effects of health on the supply of and returns to labor. In *The New Jersey Income-Maintenance Experiment*, ed. H W Watts and A Rees. New York, NY: Academic Press pp. 289–320.

Evans, William N and Craig L Garthwaite. 2014. "Giving mom a break: The impact of higher EITC payments on maternal health." *American Economic Journal: Economic Policy* 6(2):258–290.

Fafchamps, M and S Quinn. 2017. "Aspire." *The Journal of Development Studies* 53(10):1615–1633.

Feinberg, Robert M and Daniel Kuehn. 2018. "Guaranteed Nonlabor Income and Labor Supply: The Effect of the Alaska Permanent Fund Dividend." *The B.E. Journal of Economic Analysis & Policy* 18(3):350–13.

Finkelstein, Amy and Matthew Notowidigdo. 2019. "Take-up and Targeting: Experimental Evidence from SNAP." *Quarterly Journal of Economics* 134(3).

Fiszbein, A, N Schady, F H G Ferreira, M Grosh, N Keleher, P Olinto and E Skoufias. 2009. "Conditional Cash Transfers : Reducing Present and Future Poverty." World Bank Policy Research Report.

Forget, Evelyn 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.

Forget, Evelyn L. 2013. "New questions, new data, old interventions: The health effects of a guaranteed annual income." *Preventive Medicine* 57(6):925–928.

Greenberg, David and Harlan Halsey. 1983. "Systematic misreporting and effects of income maintenance experiments on work effort: evidence from the Seattle-Denver experiment." *Journal of Labor Economics* 1(4):380–407.

Hanushek, Eric A et al. 1986. Non-labor-supply responses to the income maintenance experiments. In *Lessons from the Income Maintenance Experiments: Proceedings of a Conference Held at Melvin Village, New Hampshire*. pp. 106–21.

Haushofer, Johannes and Jeremy Shapiro. 2016. "The Short-term Impact of Unconditional Cash Transfers to the Poor: Experimental Evidence from Kenya*." *The Quarterly Journal of Economics* 131(4):1973–2042.

Hausman, Jerry A and David A Wise. 1979. "Attrition bias in experimental and panel data: the Gary income maintenance experiment." *Econometrica: Journal of the Econometric Society* pp. 455–473.

Hoynes, Hilary, Doug Miller and David Simon. 2015. "Income, the earned income tax credit, and infant health." *American Economic Journal: Economic Policy* 7(1):172–211.

Hoynes, Hilary and Jesse Rothstein. 2019. "Universal Basic Income in the United States and Advanced Countries." *Annual Review of Economics* 11:929–958.

Imbens, G W, D B Rubin and B I Sacerdote. 2001. "Estimating the effect of unearned income on labor earnings, savings, and consumption: Evidence from a survey of lottery players." *American Economic Review* 91(4):778–794.

Jones, Damon and Ioana Marinescu. 2018. "The Labor Market Impacts of Universal and Permanent Cash Transfers: Evidence from the Alaska Permanent Fund." NBER Working Paper 24312.

Kaluzny, Richard L. 1979. "Changes in the Consumption of Housing Services: The Gary Experiment." *The Journal of Human Resources* 14(4):496–506.

Kangas, Olli, Signe Jauhiainen, Miska Simanainen and Minna Yilkanno. 2019. "The basic income experiment 2017-2018 in Finland. Preliminary results." Working Paper.

Kangas, Olli, Signe Jauhiainen, Miska Simanainen and Minna Yilkanno. 2020. Suomen perustulokokeilun arviointi. Technical report Ministry of Social Affairs and Health Helsinki: .

Kasey Buckles, Dan Hungerman and Steven Lugauer. 2017. "Is fertility a leading indicator?" NBER Working Paper.

Keeley, M C. 1981. *Labor supply and public policy: A critical review*. New York NY: Academic Press.

Keeley, Michael C. 1980*a*. "The Effect of a Negative Income Tax on Migration." *The Journal of Human Resources* 15(4):695–706.

Keeley, Michael C. 1980*b*. "The Effects of Negative Income Tax Programs on Fertility." *The Journal of Human Resources* 15(4):675–694.

Kehrer, B H and C M Wolin. 1979. "Impact of income maintenance on low birth weight: Evidence from the Gary Experiment." *The Journal of Human Resources* 14(4):434–462.

Kleven, Henrik. 2018. "Taxation and Labor Force Participation: The EITC Reconsidered." NBER conference presentation.

Kleven, Henrik. 2020. "The EITC and the Extensive Margin: A Reappraisal." NBER Working Paper No. 26405.

Kline, Patrick and Melissa Tartari. 2016. "Bounding the Labor Supply Responses to a Randomized Welfare Experiment: A Revealed Preference Approach." *American Economic Review* 106(4):972–1014.

Mani, Anandi, Sendhil Mullainathan, Eldar Shafir and Jiaying Zhao. 2013. "Poverty Impedes Cognitive Function." *Science* (341):976–980.

Matheny, Adam, Theodore Wachs, Jennifer Ludwig and Kay Pihllips. 1995. "Bringing order out of chaos: psychometric characteristics of confusion, hubbub, and order." *Journal of Applied Development Psychology* 16(3).

Maynard, R. 1977. "The effects of the rural income maintenance experiment on the school performance of children." *American Economic Review* 67(1):370–375.

Maynard, Rebecca A. and Richard J. Murnane. 1979. "The Effects of a Negative Income Tax on School Performance: Results of an Experiment." *The Journal of Human Resources* 14(4):463–476.

McKenzie, D. 2015. "Identifying and spurring high-growth entrepreneurship: Experimental evidence from a business plan competition." World Bank Policy Research Working Paper No. 7391.

McKenzie, David. 2012. "Beyond baseline and follow-up: The case for more T in experiments." *Journal of Development Economics* 99(2):210–221.

Meyer, Bruce. 2010. "The Effects of the Earned Income Tax Credit and Recent Reforms." *Tax Policy and the Economy* 24(1).

Meyer, Bruce and Dan Rosenbaum. 2001. "Welfare, the Earned Income Tax Credit, and the Labor Supply of Single Mothers." *Quarterly Journal of Economics* 116(3).

Miller, Cynthia, Rhiannon Miller, Nandita Verma, Nadine Dechausay, Edith Yang, Timothy Rudd, Jonathan Rodriguez and Sylvie Honig. 2016. "Effects of a Modified Conditional Cash Transfer Program in Two American Cities.".

MSchady, N and J Rosero. 2007. "Are Cash Transfers Made to Women Spent Like Other Sources of Income?" World Bank Policy Research Working Paper No. 4282.

Murnane, R, R Maynard and J Ohls. 1981. "Home resources and children's achievement." *Review of Economics and Statistics* 63(3):369–377.

Nichols, A. and J. Rothstein. 2016. The Earned Income Tax Credit. In *Economics of Means-Tested Transfer Programs in the United States*, ed. R.A. Moffit. Chicago: University of Chicago Press pp. 137–218.

O'Connor, J. Frank and J. Patrick Madden. 1979. "The Negative Income Tax and the Quality of Dietary Intake." *The Journal of Human Resources* 14(4):507–517.

Piketty, Thomas, Emmanuel Saez and Gabriel Zucman. 2019. "Distributional National Accounts: Methods and Estimates for the United States." Forthcoming, Quarterly Joural of Economics.

Rea, S A. 1977. "Investment in human capital under a negative income tax." *Canadian Journal of Economics* 10(4):607–620.

Riccio, James A and Cynthia Miller. 2016. "New York City's first conditional cash transfer program: What worked, what didn't.".

Robins, Philip. 1985. "A Comparison of the Labor Supply Findings From the Four Negative Income Tax Experiments." *Journal of Human Resources* 20(4):567–582.

Salehi-Isfahani, Djavad and Mohammad H Mostafavi-Dehzooei. 2018. "Cash transfers and labor supply: Evidence from a large-scale program in Iran." *Journal of Development Economics* 135:349–367.

Shaefer, H Luke and Kathryn Edin. 2013. "Rising Extreme Poverty in the United States and the Response of Federal Means-Tested Transfer Programs." *Social Service Review* 87(2):250–268.

Strully, Kate W, David H Rehkopf and Ziming Xuan. 2010. "Effects of Prenatal Poverty on Infant Health." *American Sociological Review* 75(4):534–562.

U.S. Department of Health and Human Services. 2016. "Poverty in the United States: 50-Year Trends and Safety Net Impacts." US Department of Health and Human Services, ASPE Report.

Weiss, Y, A Hall and F Dong. 1980. "The Effect of Price and Income on Investment in Schooling." *The Journal of Human Resources* 15(4):611–640.

Widerquist, Karl, Robert Levine, Alice O'Conner, Harold Watts, Robinson Hollister and Walter Williams. 2005. "A retrospective on the negative income tax experiments: Looking back at the most innovative field studies in social policy." *The ethics and economics of the basic income guarantee* pp. 95–106.

Wolfe, Barbara, Jessica Jakubowski, Robert Haveman and Marissa Courey. 2012. "The income and health effects of tribal casino gaming on American Indians." *Demography* 49(2):499–524.