Preanalysis Plan for the Chelsea Eats Study

Jeffrey Liebman, Kathryn Carlson, Eliza Novick, and Pamela Portocarrero

August 13, 2021

Introduction

This document describes our analysis plan for the Chelsea Eats study. The purpose of this document is to pre-specify our primary outcomes, data cleaning methodology, weighting scheme, and statistical specifications.

1. The Intervention

Chelsea, Massachusetts, a city of 40,000 people just north of Boston, is among the places in the country hardest hit by Covid-19, both from a health and an economic perspective. Its heavily Latino population is concentrated in sectors of the economy that were shut down when the pandemic hit, and Chelsea residents are also disproportionately likely to be front-line service workers exposed to infection risk. In April 2020, local community organizations and the City of Chelsea responded to the economic crisis facing jobless Chelsea residents by mounting an unprecedented food distribution effort.

In September 2020, after five months of running its food distribution sites, the City decided to redirect its efforts toward distributing financial support so that residents could purchase their own food through a program called Chelsea Eats. By combining city general revenue funds, state aid, and philanthropic contributions, the City assembled enough resources to distribute Chelsea Eats cash cards to approximately 2,000 households and to replenish the cards on a monthly basis for a total of six months. The card amounts vary with household size. Most households are receiving $400 per month, but one- and two-person households receive $200 and $300, respectively. The cards can be spent anywhere Visa is accepted. In total, 3,615 households applied for the cards, and 2,074 were chosen to receive the cash assistance cards via a lottery. Gift cards were credited with the first payment on November 18th, 2020 and with the second payment on December 18th, 2020. The program is currently expected to continue with monthly credits through August 2021.

2. Study Details

The lottery creates an opportunity to evaluate the impact of the Chelsea Eats program by comparing the outcomes of lottery winners to those of applicants who were unsuccessful in the lottery. We enrolled 1951 applicants in our study prior to randomization, 1121 who ultimately won the lottery and 830 who did not, and administered two baseline surveys, one in September 2020 prior to randomization and a second in November 2020 that was after randomization but before the Chelsea Eats cards were distributed. Because we collected baseline data, we will be able to describe the circumstances Chelsea Eats applicants were facing at the beginning of the program as well as how the circumstances of both lottery winners and unsuccessful applicants subsequently evolve.

The surveys are administered through Qualtrics, with approximately two-thirds of the sample taking the surveys on-line themselves and about one-third via a telephone interview. 74 percent of the baseline surveys were conducted in Spanish. In appreciation of their time, respondents receive a $20 gift card after every completed survey.

The research study is focusing on five primary domains where it is unambiguous what result would represent an improvement for participating households

- Financial distress
- Food insecurity
- Food satisfaction
- Health, including mental health
- K-12 school attendance

These are the five outcomes that will be evaluated in assessing the overall success of the intervention.

There are two other key outcomes where it is less clear what result would be a positive outcome.

- Employment. While there are benefits to the economy of increased output, if the extra income permits people to delay going back to unsafe jobs during the Covid epidemic or to stay home and supervise their children in online learning, then a decline in employment could represent a beneficial outcome.

- Diet quality. If the extra income permits people to switch from eating rice and beans to eating beef and pork is that a good outcome or a bad outcome?

We will not include these latter two outcomes in the omnibus test of the overall success of the intervention.

In addition to the two baseline surveys, we administered pulse surveys in December 2020 and February 2021 that measured a small number of key outcomes and enabled us to stay in touch with households, a final survey in April and May 2021, and a 24-hour food recall collected in April and June 2021. We also obtained administrative attendance data from the Chelsea Public Schools. Finally, we are trying to obtain spending data from the gift cards we used to compensate respondents, but as of this writing it is unclear if we will be able to obtain those data.

3. Primary Outcomes

Because the study is measuring a wide range of outcomes, we do not have statistical power to test them all independently. Instead, we are specifying a subset of outcomes as the primary outcomes and collapsing related outcomes into indices as in Kling, Liebman, and Katz (2007) and Kling and Liebman (2004).[1] Results of all of the other outcomes will be considered descriptive and exploratory.

The five primary outcomes of the study are:

1. Financial distress. We will form an index that equally weights measures of three concepts: whether the household's financial situation has improved over the past six months; whether the household can currently afford its essential expenses, and whether the family would be able to pay for a $300 emergency expense. The first and third concepts will be measured directly from individual survey questions. The second concept will be measured by creating a subindex combining several different survey questions. In particular, half of the weight for this subindex will be from the summary question of whether the household can currently meet its essential expenses and the other half will be constructed by taking a weighted average of the answers to whether the

[1] Kling, J.R., J. B. Liebman, and L.F. Katz (2007). Experimental Analysis of Neighborhood Effects. Econometrica, 75:1, 83-119. Kling, J.R. and J.B Liebman (2004). "Experimental Analysis of Neighborhood Effects for Youth." Working Paper 483. Industrial Relations Section, Princeton University.

family is behind on utilities, behind on rent, and delaying health care expenditures with the weights on these three elements coming from the relative expenditure shares for low-income families for utilities, housing, and out of pocket medical expenditures as measured in the Consumer Expenditure Survey.

2. Food insecurity. Food insecurity will be measured using the USDA/Current Population Survey Food Security Scale, scored using the procedures described in USDA's Guide to Measuring Food Security (2000 edition).

3. Food satisfaction. Food satisfaction will be measured by constructing an equally weighted index of three survey questions. The first component will be the fraction answering "enough of the kinds of foods we want" to the standard one question food insecurity question. The second component will be the fraction of respondents saying they are "very satisfied" with their family's food situation these days. The third component will be the fraction saying their family's food situation is "better" than six months ago.

4. Health. Health will be as measured by an equally weighted index of two concepts. The general health component will measure the fraction of respondents saying their health is "fair" or "poor." The mental health component will measure the fraction of respondents who score 3 or above on either the GAD-2 or the PHQ-2.

5. School attendance. Child school attendance will be measured as the number of days of school missed during the 2020-2021 school year (as of our May measurement date).

In addition to presenting separate tests of statistical significance for each of these outcomes, we will create an overall index that equally weights the five primary outcomes (each converted to standard deviation units) and perform a test for the overall impact of the intervention. We will also present results for the individual components of the indices. These individual results will show both the per comparison p-values and p-values adjusted using the modified Westfall-Young procedure described in Kling and Liebman (2004).

For the other two key outcomes domains – employment and diet quality – our primary measures will be the percent of households in which either the respondent or the respondent's partner is employed and the respondent's score on the Healthy Eating Index, respectively. Because of the

complexity of analyzing the food recall data, we (along with a different team of coauthors) expect to present those results in a separate paper that is likely to be completed after the paper describing the rest of the survey results.

4. Accounting for Survey and Item Non-response

Because we were only able to complete final surveys with 95 percent of our sample, there is a possibility that the attrition process could bias our results. For our primary results, we will estimate an equation that predicts attrition using baseline characteristics and then use this equation to reweight the results. We also plan to produce sensitivity analyses in which we try to model the relationship between survey response date and outcome measures (e.g., employed respondents may be harder to reach because they are busy at work during the day, so later respondents may have higher employment rates) and use the estimated relationship to impute responses for nonrespondents.

For missing values of regression covariates, we will impute the mean value. For item non-response for primary outcomes that are measured by a series of questions (e.g., the mental health questions and the food insecurity questions), if a respondent answers most of the questions in an index, we will impute any missing questions based on the responses to the other elements in the index.

5. Weighting

The lottery was a weighted lottery. Household received additional "lottery tickets" if no one in the household was currently working, if the household was not receiving unemployment insurance, if the household was not receiving food assistance like SNAP, if there was a disabled household member, if there was a household member over 65 years of age, if there was a household member who was a veteran, if there was a household member who was under 6 years of age, and if there was a member between 6 and 17. The total number of lottery tickets ranged from 1 to 8. In order to produce balanced samples between the treatment group and control group we will be weighting the sample by the inverse probability of winning the lottery. The weights will be created so that the treatment and control groups are each representative of the full applicant population.

6. Duplicate Households and Waitlist Households

We have recently learned about two administrative details that will complicate our analysis. First, although it was the city's intention to restrict households to a single application, some households managed to enter the lottery twice. In the cases in which a household won the lottery twice, the household received only a single Chelsea Eats card. We are currently trying to identify all of the duplicate applicants using only data collected pre random assignment. We will be combining the records for the duplicate applicants into a single record, summing the lottery tickets across the duplicate records to determine the household's overall probability of winning the lottery. In cases in which the duplicate records resulted in households taking our surveys more than once, we will first examine to see if one survey is significantly more complete than the other and include only the more complete survey in our analysis. In cases in which multiple surveys are largely complete, our current plan is to include both of the surveys in our analysis but give them each half of the weight that they would otherwise receive (and cluster the standard errors at the household level). If that proves infeasible, we will include only the first survey from a household as measured by the survey complete date and time.

The second complication is that some households never picked up their Chelsea Eats cards. The city then gave these unused cards to households from a randomly generated waitlist that was created at the time of random assignment. The waitlist households received their cards approximately one month later than the primary lottery winners, but the cards were credited with both the payments for the first and second months – so the total payments received by waitlist households are the same as initial lottery winners. We are currently in the process of obtaining the list of waitlist households and will consider them to be lottery winners (members of the treatment group). We will adjust the probabilities of winning the lottery to include both the chance of winning the initial lottery and the chance of winning via the wait list. This means our treatment group will be larger and our control group smaller than the headline numbers we described on page one and two.

7. Statistical Specifications

We will estimate the experimental impacts by regressing outcomes on an indicator variable for being a lottery winner. As mentioned above, these regressions will be weighted to account for the varying probabilities of winning the lottery. Binary outcomes will be estimated with linear probability models. We will present results both with and without covariates. Covariates improve estimation precision and account for chance differences between the treatment and control group in the distribution of pre-random-assignment characteristics. We will use robust standard errors (clustered within households for the households with duplicate observations). The regression specification is the standard one

$$Y = Z\pi + X\beta + \epsilon$$

where Z is an indicator for being a lottery winner and X is a matrix of baseline covariates. The baseline covariates will be:

Age (indicator variables for five-year age groups)

Sex (indicator variable for female)

Race and ethnicity (indicator variables for Latino, White, and Black)

Number of adults in the household

Number of children younger than 5 in the households

Number of children between 5 and 17 in the household

Financial situation compared to last year (indictor variables for better and worse; omitted indicator is "the same")

Financial situation compared to last month (indictor variables for better and worse; omitted indicator is "the same")

Expectations of future financial situation (indictor variables for better and worse; omitted indicator is "the same")

Indicator for whether the household reported experiencing financial difficult in the past year

Indicator for financial difficultly from respondent having lost a job

Indicator for financial difficulty from spouse having lost a job

Indicator for financial difficultly from respondent having reduced hours or lower pay

Indicator for financial difficulty from spouse having reduced hours or lower pay

Indicator for financial difficulty from health issues

Indicator for having expenses in August that household could not pay

Indicator for having unpaid rent

Indicator for having unpaid utilities

Indicator for having utilities, cable, or phone disconnected because of unpaid bills

Food insecurity (three indicator representing three of the four responses to the standard one-question food insecurity question)

Food insecurity index (a count of the number of the five additional household food insecurity questions with affirmative responses)

Child food insecurity (two indicators representing two of the three responses to the standard one-question child food insecurity question)

Monthly income before Covid (will adjust for household size using the recommended adjustment in the National Academy of Sciences *Measuring Poverty* book, will include both linearly and squared. Will likely winsorize at the fifth percentile and 95th percentile).

Current income as percentage of pre-covid income (will include both linear and squared terms, will likely winsorize to reduce impact of measurement error).

General health (indicator variables for four of the five response categories)

Mental health (K6 score)

Covid (an indicator variable for whether the respondent or another member of the respondent's household has had Covid).

Employment (indicator variables for unemployed looking for work and unemployed not looking for work – omitted category is "working for pay")

Indicator for disability being the reason for not working

Indicator for either "family responsibilities" or "taking care of kids" as reason for not working.

8. Main Table for Assessing Intervention Impact

Table 1: Summary Indices

	Without covariates				With covariates	
	Treatment	**Control**	**Difference**	**P-value**	**Difference**	**P-value**
Financial distress						
Food insecurity						
Food satisfaction						
Health						
School attendance						
Overall						