

# Lotteries of Temporary Housing Assistance during the COVID-19 Epidemic – King County

## Pre Analysis Plan

February 2021

### 1. Introduction

This evaluation will examine lotteries of rental assistance in King County, WA (Seattle) for low-income households affected by the COVID-induced economic crisis. King County will offer to pay 3 months of back rent to landlords of lottery winners. We will use lottery assignment as an instrument for payment of funds and study how such assistance affects housing stability. For approximately 6,000 participants in these lotteries, we will collect administrative records and initiate surveys to measure address changes, evictions, homelessness, payment of rent, financial status, and health. We are coordinating with researchers studying similar lotteries in other cities.

King County's Department of Community and Human Services (DCHS) recently created a package of financial assistance programs to efficiently distribute COVID-19 relief funds contributed by the CARES Act, the federal Eviction Rental Assistance Program, COVID-19 Community Development Block Grants, and other sources. These programs include a program to pay back rent of tenants of large landlords and mobile home parks as well as a more traditional program assisting tenants who apply for assistance. This evaluation pertains to the latter program, and people eligible for other programs will not be eligible for the lotteries we will study.

### 2. Key Data Sources

#### A. King County Homeless Management Information Systems

The Homeless Management Information System collects client-level data from all publicly contracted homeless service providers in King County. From this information system, we can observe date-specific service records to track outcomes for both treatment and control group participants. Among other variables, HMIS data will allow us to observe vulnerability assessment scores and entry into emergency shelters or transitional housing. HMIS will act as the primary database for the purposes of merging County records. HMIS ID, name, SSN, date of birth, and other characteristics will be used to link records together.

#### B. Infutor (Consumer Reference Database)

We will also measure housing stability using comprehensive consumer reference data on household addresses in the U.S. These data allow us to determine how frequently study participants move or have an official address—information that is difficult to obtain through surveys or administrative data records. Infutor is a consumer reference data company that aggregates records from various sources (e.g. utility bills) to construct address histories for most people in the United States. Similar address history data has been used, for example, to longitudinally follow residents affected by rent control in San Francisco (Diamond et al. 2018). In the context of housing stability, this outcome is particularly useful. Compared to shelter entry, the existence of a formal address in such data is a more comprehensive measure of homelessness as it can provide information on unsheltered homelessness and those depending on others for housing.

#### C. Web Survey

We will attempt to contact the same sample to survey lottery applicants about their housing, financial and health situations in greater detail using a web survey instrument adapted from one used in Chicago by Bertrand et al. We plan to conduct this survey in February 2021. It will ask retrospectively about their circumstances in the lead-up to applying for assistance (as well as pre-COVID situations), and ask about current (post-treatment) outcomes. The survey will measure outcomes that are difficult to capture from purely administrative sources. For housing stability, this includes measures such as: doubling-up or “couch-surfing”, staying in places not meant for human habitation (such as sleeping in a vehicle or outdoors), reasons for moves, reasons for not-renewing a lease, and details on arrangements the tenant has made with the landlord regarding repayment of rental arrears. For financial stability, the survey captures details on missed, late, or incomplete payment for key necessities such as utilities and rent; missed bill payments for common bills (car payments, health insurance etc); a measure of food insecurity; and borrowing behavior (e.g. use of payday loans). Finally, from the survey, we can examine effects on broad measures of health such as mental health troubles and illness due to COVID-19.

#### **D. Experian**

We also intend to match the sample Experian credit records. These data provide information on credit scores, debt, credit inquiries, and delinquency. While we will attempt to match the entire sample, we may be limited by privacy considerations to focus on the sub-sample who provide permission in the survey.

### **3. Experimental Design**

The tenant rental assistance intervention we will study is aimed at providing King County residents negatively impacted by COVID-19 with emergency assistance funding to assist with disruptions in their ability to pay rent. Residents complete an interest form administered by King County and, if eligible for the program, receive a program application. Because demand for the program exceeds available funding, funds are offered by lottery. The program offers those tenants who are selected three months of back rent paid directly to the landlord. The interest form became available August 20, 2020, and lotteries began on September 14. A short research consent statement and some identifiers were added to the interest form on September 16; we focus on people applying after that time.

Our evaluation design centers on the tenant assistance lotteries. Each week, DCHS conducts a lottery among applicants. They screen out ineligible people, multiple applications from the same house/apartment, and people eligible for landlord-based programs. Among these eligible applicants, they randomly select the number of applicants for whom they have staff capacity to offer assistance that week. Both winners and non-winners receive an e-mail informing them of the week’s result. Staff at delegate agencies then follow-up with winners to offer the assistance and work with landlords to pay it. People assigned to the control group will have access to usual care that does not include significant access to similar financial assistance. This process was repeated weekly until the end of 2020.

The series of lotteries creates a natural randomized controlled trial. The lottery creates a comparison group of eligible applicants who do not win. Since applicants are eligible for all lottery weeks after their application date, the probability of winning is conditional the week the applicant applied, but this date can be observed. Take-up will be incomplete and endogenous due to incorrect contact information and rejection of program terms by some landlords, so we will use the offer of assistance as an instrument for receiving assistance.

### **4. Ethics**

This research protocol was reviewed and approved by the Institutional Review Board of the University of Notre Dame. The lotteries of emergency financial assistance were created outside of the research process, so research ethics pertain only to data collection via administrative data and the survey. Privacy, data storage, and security are maintained in keeping with that review, as well as data sharing agreements with the data owners.

## 5. Hypotheses - Main Analysis

We state four primary hypotheses:

1. Receipt of emergency financial assistance increase the likelihood of paying rent in full.
2. Receipt of emergency financial assistance increases housing stability.
3. Receipt of emergency financial assistance increases the likelihood of paying non-housing bills and subsequent financial stability.
4. Receipt of emergency financial assistance improves health status, particularly mental health.

We will analyze outcomes by each of these four domains. Within each domain, we indicate a primary measure, which will be the focus of our analysis. We also list alternative measures of interest, which we may investigate to further understand the main effects on the primary outcome measure.

### A. Payment of Rent

Primary measure: payment of rent

- a) Percent of rent paid
  - (1) For a given month, divide rent paid by total rent
  - (2) Constructed using survey
- b) Hypothesis: expect treatment group to pay more rent

Secondary measures

- Survey: presence of agreement with landlord, characteristics of landlord agreement, reason for no agreement with landlord, expectations of future payment

### B. Housing Stability

Primary measure: either changed addresses or used homelessness services

- a) Indicator for if the participant either changed addresses or used homeless services
  - 1) Indicator for whether a participant has changed addresses since random assignment
    - (i) Dummy for whether any address ends or starts after random assignment.
    - (ii) Constructed using consumer reference database Infutor.
  - 2) Indicator for use of homelessness services
    - (i) Dummy for whether or not participant received any of the following services: Emergency Shelter, Street Outreach, Permanent Supportive Housing, Rapid Re-Housing, Transitional Housing, Diversion, or Coordinated Entry
    - (ii) Constructed using HMIS data
  - 3) Outcome is the union of these two measures
- c) Hypothesis: expect treatment group to be less likely to move or use homeless services

Alternative measures

- HMIS: use of homeless services, use of emergency shelter
- Infutor: address moves, former address ending, new address beginning, neighborhood characteristics of most recent address
- Survey: unstable housing, doubling up, homelessness, expectations of housing stability, address moves, reasons for moving, evictions, lease non-renewal, knowledge of eviction moratorium

### C. Financial Stability

Primary measure: % of non-rent bills paid

- a) Estimate % of bills paid from responses about particular bills in the past month
  - (1) For each particular bill, count fully paid bill as 1, partial payment as 0.5 and no payment as 0.
  - (2) Sum over internet, phone, gas, electricity, car payment, car insurance, health insurance, student loans
  - (3) Divide by the number of bills, excluding those for which they did not have a bill
  - (4) Constructed using survey
- b) Hypothesis: expect treatment group to be more likely to pay bills

Alternative measures

- Survey: payment of bills by category, food security, borrowing/debt
- Experian: credit score, delinquency, collections, debt, credit inquiries

### D. Health

Primary measure: mental health

- a) Number of days anxious and depressed
  - (1) For all four questions in the PHQ-2 and CAG-2 scales, translate the responses into a number of days (not at all = 0; couple days = 2; more than half = 4; nearly every = 6). Then find the mean across the four questions.
  - (2) Constructed using survey
- b) Hypothesis: expect treatment group to be less likely to be depressed and anxious

Alternative measures:

- Survey: GAD-2 and PHQ-2 separately; COVID infections
- Administrative records: COVID positive test results

## 6. Sub-Group Analysis

We are interested in determining whether the intervention is relatively more impactful for certain populations relative to others.

### A. Split by predicted outcomes

- a) For each outcome, we construct predict the outcome in the control group using a regression on the full set of variables listed as control variables in section 7. We then predict the outcome for all observations using observed characteristics and the coefficients from that

regression. We finally split the sample into thirds based on that index and test for heterogeneous effects in the three groups. To avoid endogenous stratification we compute these statistics with a repeated split sample procedure as in Abadie, Chingos, and West (2018, RESTAT).

B. Causal forest

- a. For each outcome, we estimate a causal forest (Wager and Athey, 2018, JASA) using the full set of variables listed as control variables in part 8 as predictors.

## 7. Data Analysis

A. Estimates

We will estimate a two-equation instrumental variables specification.

$$T_{it} = \alpha_0 + \alpha_1 Z_{it} + \mu_t + \epsilon_{it} \quad (1)$$

$$Y_{it} = \beta_0 + \beta_1 T_{it} + \eta_t + u_{it} \quad (2)$$

We consider a person  $i$ . The data will be a cross-section, but we use the subscript  $t$  to denote the week in which the person applied to the program.  $Y_{it}$  measures an outcome, such as whether the applicant changes addresses after random assignment;  $T_{it}$  indicates whether the person received assistance from the program;  $Z_{it}$  indicates whether the person was a lottery winner;  $\mu_t$  and  $\eta_t$  are application week fixed effects; and  $\epsilon_{it}$  and  $u_{it}$  are error terms. Equation (1) is the first stage, in which lottery assignment predicts treatment status, conditional on week of application. Equation (2) is the second stage. The coefficient of interest is  $\beta_1$ .

B. Covariates

We plan to use double LASSO to select covariates. For each outcome, LASSO will select covariates from the following list as well as two-way interactions among all of these variables:

- i. Value of dependent variables at baseline, if available
- ii. Gender
- iii. Set of mutually exclusive race/ethnicity dummies
- iv. Age
- v. City dummies (for all cities with more than 1% of the sample)
- vi. Disability status
- vii. Set of preferred language dummies

C. Standard Errors

We will use heteroscedasticity-robust standard errors.

D. Multiple Hypothesis Testing

We have limited our primary outcomes to a causal chain with one outcome per domain, making multiple hypothesis testing less of a concern. Nonetheless, we will report classic p-values. This provides the reader with full information that they can use to make MHT corrections if they desire.

E. Variables with Limited Variation

Questions with a non-response rate of 90% or higher will be omitted from analysis.

F. Missing Data

If data are missing for some observations for a particular outcome, we will calculate treatment effects among the sample with observed values for that outcome. If data are missing for some observations for a control variable, we will fill in a zero value and also include a dummy for whether that variable is missing at baseline.