I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13559

Short Moves and Long Stays: Homeless Family Responses to Exogenous Shelter Assignments in New York City

Mike Cassidy

JULY 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13559

Short Moves and Long Stays: Homeless Family Responses to Exogenous Shelter Assignments in New York City

Mike Cassidy
Princeton University and IZA

JULY 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Short Moves and Long Stays: Homeless Family Responses to Exogenous Shelter Assignments in New York City*

Using an original administrative dataset in the context of a scarcity induced-natural experiment in New York City, I find that families placed in shelters in their neighborhoods of origin remain there considerably longer than those assigned to distant shelters. Locally-placed families also access more public benefits and are more apt to work. A fixed effects model assessing multi-spell families confirms these main results. Complementary instrumental variable and regression discontinuity designs exploiting policy shocks and rules, respectively, suggest difficult-to-place families – such as those that are large, disconnected from services, or from neighborhoods where homelessness is common – are especially sensitive to proximate placements. Better targeting through improved screening at intake can enhance program efficiency. The practice of assigning shelter based on chance vacancies ought to be replaced with a system of evidence-based placements tailored to families' resources and constraints.

JEL Classification: R28, I38, R20, H53, H75, D91, J22

Keywords: homelessness, neighborhoods, families, poverty alleviation, housing, public assistance, welfare policy, labor supply, program evaluation, causal inference

Corresponding author:
Mike Cassidy
Center for Health and Wellbeing
219 Julis Romo Rabinowitz Building
Princeton University
Princeton, NJ 08544-1013
USA

E-mail: miketcassidy@princeton.edu

* This research has been made possible through an abundance of kindness and support. Jennifer Hunt has been an extraordinary mentor and close reader. Roger Klein, Carolyn Moehling, Brendan O'Flaherty, and Rosanne Altshuler have offered many useful comments. The Center for Innovation through Data Intelligence (CIDI), a research unit within the Office of the Mayor of the City of New York, has graciously furnished data access and policy expertise. Particular thanks goes to Maryanne Schretzman, Eileen Johns, Andy Martens, and Jacob Berman. Generous funding has been provided by the National Academy of Education and the National Academy of Education/Spencer Dissertation Fellowship Program. Hilary Sigman, Douglas Blair, Jacob Bastian, Ira Gang, Anne Piehl, Eugene White, Norman Swanson, Bingxiao Wu, Robert Collinson, Emily Oster, Philip Oreopoulos, Ingrid Gould Ellen, Howard Bloom, Richard Murnane, and Jeremy Fiel, as well as participants in Rutgers University's Applied Microeconomics Seminar, have also volunteered helpful feedback.

1 Introduction

Housing is the most essential good people consume, besides, perhaps, food. Despite this, homeless families remain curiously ignored by economists. Housing instability is associated with worse physical and mental health, greater food insecurity, less labor market success, and more poverty (O'Flaherty, 2019; Ellen and O'Flaherty, 2010). Homeless children struggle in school (Buckner, 2008; Miller, 2011; Samuels, Shinn and Buckner, 2010). While causality primarily derives from deeper determinants (Cassidy, 2020), these compound challenges nevertheless mark homeless families as a population especially deserving of attention.

Nationwide, more than a third of America's homeless—some 180,413 individuals—are people in families (U.S. Department of Housing and Urban Development, 2018). Unlike the single adult street homeless who loom large in the public consciousness, homeless families—typically, young, African-American and Hispanic single moms with several kids and high school educations—reside out of view in government-provided homeless shelters often indistinguishable from the sorts of marginal housing stock from whence they came. Most of these families are neither addicted nor ill, but rather poor and unlucky[1].

Nowhere are the manifestations more obvious than in New York City, where the confluence of a legal right to shelter, high housing costs, and progressive governance (NYC Mayor's Office, 2017) led the shelter census to rise from 8,081 families in March 2009 to 13,164 in November 2018 (NYC Department of Homeless Services, 2019*b*)[2]. Sheltering these families costs taxpayers more than $1.2 billion annually (NYC Office of Management and Budget, 2019).

Reducing homelessness, a municipal priority for decades, has taken on increased urgency. The City maintains myriad programs intended to minimize shelter stays. Prevention services forestall entries. Rental subsidies speed exits. Traditional public assistance and work supports fill gaps. But accepting that some homelessness is unavoidable, a central element of the City's strategy is to make homeless spells less disruptive for families through neighborhood-based shelter placements. Since at least the late-1990's, the City has maintained a policy of placing families in shelters in the boroughs of their youngest children's schools. While the policy is predicated on minimizing educational hardship, community continuity—keeping families connected to friends, relatives, jobs, and places of worship—has increasingly been seen as a way of improving overall well-being and expediting returns to permanent housing (NYC Mayor's Office, 2017).

[1] O'Flaherty (2019); Evans, Philips and Ruffini (2019); O'Flaherty (2010); Culhane et al. (2007); Shinn et al. (1998); Curtis et al. (2013).

[2] This includes only families sheltered by the Department of Homeless Services (DHS). Since 2018, the family census has stabilized, standing at 12,195 as of September 2019.

In this study, I evaluate how families assigned to shelters in their neighborhoods of origin fare compared to those situated in less proximate shelters. I find that local placements result in considerably longer shelter stays. Proximity also promotes access to public benefits, as well as gains in employment and earnings. In other words, families do better when placed locally, but they remain homeless longer.

It is not immediately obvious that this would be the pattern of results. One could envision an alternative scenario where proximity-propagated labor market success is associated with shorter stays. Instead, the evidence suggests the comforts of being placed near one's networks (which encourage longer stays) outweigh any resource-augmentation they produce (which encourage shorter ones). Shelter satisfaction is more receptive to the effects of proximity than is labor, or at least more promptly so.

I explain my findings with a "search effort model of family homelessness," in which sheltered families choose how to allocate effort between housing search and other activities they value. Local placement is preferable, so families assigned there remain in shelter longer, diverting time that they would otherwise spend on housing search to other activities, like work and school. Locally-placed families may also require additional incentives—rental subsidies—to leave. Optimal search effort is increasing in family resources; the greater supports, or fewer constraints, a family is endowed, the less it gives up by searching.

My empirical results proceed from analysis of a novel administrative panel of all eligible families with children who entered the NYC Department of Homeless Services (DHS) family shelter system from 2010 to 2016. I construct it by linking Department records detailing family characteristics and shelter experiences with data on public benefit use and labor market experiences maintained by other agencies.

At the core of my research design is a natural experiment. Policy objectives notwithstanding, severe capacity limitations—the vacancy rate for traditional shelters was below 1 percent in 2016 (NYC Mayor's Office, 2017)—have meant that local placement is challenging to achieve. In 2010, 66 percent of families were placed in shelters in their boroughs of prior residence; by 2016, the local placement rate had dropped to 38 percent[3]. According to program administrators, conditional upon factors implicated as placement criteria—family size, health constraints, safety, and having a school-aged child—which families are placed locally is largely a matter of chance: what suitable units are available at the time of application[4].

[3]Calculations based on my sample and treatment definition. Officially, the City reports having placed 84 percent of families in the boroughs of their youngest children's schools in fiscal year 2010, declining to a range of 49–53 percent between FY15 and FY19 (NYC Mayor's Office of Operations, 2012; New York City Mayor's Office of Operations, 2019).

[4]During their stays, families may be offered transfers to more proximate shelters. Because these moves are at families' discretion, my treatment definition is based on initial assignment.

I demonstrate that this random assignment characterization is empirically apt. Assuming the same is true of unobservables, I can give causal interpretation to differences in average outcomes between locally- and distantly-placed families, after adjusting for placement factors. Nevertheless, I supplement OLS analysis with three complementary quasi-experimental identification strategies: instrumental variables, regression discontinuity, and family fixed effects. These can be viewed as guarding against endogeneity or as local average treatment effects reflecting heterogeneous responses.

The first strategy is an instrumental variable (IV) approach exploiting exogenous policy shocks. Although NYC has a legal right to shelter, families must prove their needs through a rigorous application process. City officials retain considerable discretion in making these determinations. The more lenient is eligibility policy, the faster the rate of shelter entry and the more competitive are local placements. Hence, my first instrument is the ineligibility rate: the higher is this rate, the better are the chances of in-borough placement for accepted families. While the applicant mix can influence the ineligibility rate, the most notable swings occur with changes of administration or other well-publicized policy initiatives. My second instrument, which I refer to as the "aversion ratio," extends the first by giving the rate of shelter stays averted—through ineligible applications and subsidized exits—per new entrant. During my study period, the City initiated and ended several rental assistance programs; as with eligibility, subsidy availability depends upon political priorities and budgetary constraints. I use these instruments separately, each characterizing an experiment influencing the treatment statuses of treatment-marginal families whose local placement responses may be different than average.

My second identification strategy takes advantage of exogeneity embedded in the neighborhood placement policy itself, isolating responses along a different margin. It is a regression discontinuity (RD) design based upon oldest children's ages. Neighborhood-based shelter placement is, first and foremost, an education policy, and so families with school-age children receive priority for in-borough placement. Because the timing of shelter entry is partly beyond families' control, those who enter shelter prior to their oldest children starting school (and are ineligible for the local placement boost) are counterfactuals for those who enter shelter after (and are eligible).

My third identification strategy is a family fixed effects approach. Repeat spells of homelessness are common. So long as outcome-relevant unobservables are spell-invariant, families who enter shelter multiple times with varying treatment assignments are counterfactuals for themselves.

Neighborhood placements have powerful impacts. Per OLS, families placed in-borough remain in shelter 12.7 percent longer, equivalent to about 50 days. Locally-placed families also access more public benefits and are better connected to the labor market. During the year following shelter entry, they are 1.4 percent (1.1 percentage points) more likely to receive Cash Assistance, 2.1 percent (1.0 pp) more likely to be employed, and have 9.9 percent higher earnings[5]. Elevated benefit use continues post-shelter. In-borough families are 4.6 percent (1.8 pp) more likely to exit shelter with a rental subsidy, and Cash Assistance receipt continues to be 2.3 percent (1.7 pp) higher during the ensuing year. However, labor market effects attenuate. Given capacity-based random assignment is the most broadly applicable experiment—all families are affected—these are my preferred estimates of *average* treatment effects (ATE's). Family fixed effects results—which are also informed by the natural experiment of shelter scarcity—reinforce these findings, with modestly larger effect estimates across outcomes.

On the other hand, my IV and RD results indicate that OLS may understate the potential of neighborhood-based placements. In the context of quasi-random assignment, I interpret IV and RD as dually-layered natural experiments identifying local average treatment effects (LATE's) among difficult-to-treat subgroups: "compliers" who are placed locally only when conditions are especially fortuitous[6]. I show that compliers exhibit characteristics one would expect of families facing augmented barriers to proximate placements: on average, complier families are large, young, disconnected from services, and from neighborhoods where homelessness is common. They are also more responsive to treatment. Ineligibility rate, aversion ratio, and school-starting compliers stay in shelter an order of magnitude longer than average homeless families when placed locally. They are as much as doubly likely to receive Cash Assistance compared to when they are placed out-of-borough. The evidence on labor market outcomes is more mixed. Policy compliers see large boosts to employment and earnings, while school-starters see diminished job prospects, especially post-shelter. The gap between ATE's and LATE's demonstrates the difference between average and marginal policy impacts. In other words, by carefully choosing policymaker-controlled instruments that affect treatment participation margins, I am able to identify policy relevant treatment effects in the spirit of Heckman and Vytlacil (1999, 2001, 2005, 2007).

Alternatively, under the assumption of constant treatment effects, another interpretation of IV estimates larger in absolute value than OLS is as evidence of endogeneity: OLS

[5]These outcomes may well be related. Longer stays allow more time for benefit and employment effects to percolate; at the same time, better connections to jobs and supports may encourage longer stays. In addition, Cash Assistance comes with work requirements and work supports.

[6]For further details on the LATE concept, as introduced by Imbens and Angrist (1994) and Angrist, Imbens and Rubin (1996), see Angrist and Pischke (2008). Heckman and Vytlacil (1999, 2005, 2007), Vytlacil (2002), Heckman, Urzua and Vytlacil (2006), and Heckman (2010) show LATE's can be constructed from a choice-theoretic primitive—the marginal treatment effect (MTE)—which has the additional appeal of unifying the treatment effects literature.

coefficients biased toward zero by unobservables correlated with treatment. In this telling of it, in-borough families are disproportionately those who would have left sooner on their own; they are also less likely to have their public benefit use or employment patterns impacted by local placement. One story consistent with these results is that unobservably well-resourced (or minimally-constrained) families are systematically more likely to secure favorable placements.

These findings complement those in Cassidy (2020), where, studying the same neighborhood placement policy, I find that local shelter assignment significantly improves homeless students' attendance, stability, and test scores. This pair of papers are the first (to my knowledge) to situate homeless families in an expressly microeconomic framework and assess, empirically, how they respond to the incentives of the shelter services system—as well as how their shelter usage patterns relate to labor supply, education, and participation in other government benefit programs.

Besides Cassidy (2020), the works most similar to my own are Curtis et al. (2013), who study health as an exogenous shock to family homelessness, Collinson and Reed (2018), who use a randomized judge design to study the effect of evictions on homelessness in NYC, Cobb-Clark et al. (2016), who use econometric methods to study homeless duration, and Cobb-Clark and Zhu (2017), who find that early-life homelessness is associated with worse education and employment outcomes in adulthood[7].

My work also contributes to two other literatures. The first is neighborhood effects[8]. The best studies have used natural experiments—typically the allocation of housing subsidies through lotteries—and have tended to find negligible effects on most economic outcomes[9]. However, some recent evidence suggests residential mobility improves families' contemporaneous physical and mental health and subjective well-being, as well as longer-term educational and labor market outcomes for children[10]. My work is the first to examine the effects of neighborhood specifically in the context of homeless families, a group less well-off than

[7]In contrast, most previous economic studies of homelessness have focused on one of five themes: macro issues (Cragg and O'Flaherty, 1999; O'Flaherty and Wu, 2006; Gould and Williams, 2010; Corinth, 2017); single adults (Allgood and Warren, 2003; Allgood, Moore and Warren, 1997); theory (Glomm and John, 2002; O'Flaherty, 1995; O'Flaherty, 2004, 2009); description (Shinn et al., 1998; Culhane et al., 2007; Ellen and O'Flaherty, 2010); prevention and prediction (Goodman, Messeri and O'Flaherty, 2014; Goodman, Messeri and O'Flaherty, 2016; Evans, Sullivan and Wallskog, 2016; O'Flaherty, Scutella and Tseng, 2018*a*; O'Flaherty, Scutella and Tseng, 2018*b*); or housing stability interventions (Wood, Turnham and Mills, 2008; Gubits et al., 2016). O'Flaherty (2019) and Evans, Philips and Ruffini (2019) provide two recent and comprehensive summaries of this literature from the perspective of economists.

[8]Topa, Zenou et al. (2015) summarize this literature.

[9]Oreopoulos (2003); Jacob (2004); Kling, Liebman and Katz (2007); Ludwig et al. (2008); Sanbonmatsu et al. (2011); Jacob and Ludwig (2012); Jacob, Kapustin and Ludwig (2015); Galiani, Murphy and Pantano (2015).

[10]Ludwig et al. (2012, 2013); Chetty, Hendren and Katz (2016); Andersson et al. (2016).

the low- and moderate-income families typically featured.

Second, an understanding of homeless family behavior can inform the design of poverty alleviation programs generally. Optimal programs must strike a balance between helping the truly needy and minimizing moral hazard[11]. My findings inform this trade-off. Abstractly, capacity-constrained shelter placements are exogenous variation in a public benefit program. Families that luck their way into more generous benefits have less incentive to give up those benefits and, simultaneously, wider latitude to pursue utility-augmenting possibilities.

But are neighborhood-based shelter placements a good idea? My findings indicate the answer is not unambiguous. When placed locally, homeless families will remain homeless longer (generally regarded as welfare-reducing) but they will be better connected to government services, jobs, and their children's schools (generally regarded as welfare-enhancing). In other words, the two current pillars of New York City's family homeless policy—stays that are short and comfortable—are not complementary. Nor are these stays cheap. At the City's average shelter cost of about \$200 per family per day (NYC Mayor's Office of Operations, 2018), neighborhood-based placements cost the City an additional \$10,000 per family. It is an open question whether 10 percent gains in school attendance and earnings are the best uses of the City's next \$10,000.

Recognizing these trade-offs is important. But complicated questions of budgetary optimization are not the first step; the more immediate point is that there remains ample room to enhance the efficiency of neighborhood placements. Outcomes among homeless families are highly variable. My IV and RD compliers—marginally-treated families—are highly policy elastic. This suggests potential gains from better targeting local placements to families most likely to benefit. Policy-relevant heterogeneity should be better screened at intake and explicitly factored into placement decisions using predictive models. Special attention should be afforded families who are difficult to place: my results suggest it is these families whose outcomes will be most sensitive to their assignments. Integrated support services should be correspondingly customized to families' comparative advantages and limitations, while respecting the influence placement proximity (and other characteristics) will have on families' incentives.

These insights derive from a natural experiment in shelter assignment. That experiment should be replaced with evidenced-based placements designed to allocate scarce resources in a welfare-maximizing manner.

[11]Nichols and Zeckhauser (1982); Besley and Coate (1992, 1995).

2 Policy Background

Neither homelessness nor poverty among families are foreign to municipalities anywhere in the United States, but in few places is the intersection starker than in New York City. Since 1994, New York's homeless census has nearly tripled, from 24,000 to 60,000 in 2019. Two-thirds are people in families (NYC Department of Homeless Services, 2019*b*). Overall, NYC accounts for about a quarter of sheltered homeless families in the U.S. (NYC Department of Homeless Services, 2019*d*; U.S. Department of Housing and Urban Development, 2018; Coalition for the Homeless, 2019)[12].

Family homelessness is particularly pronounced in New York City for two reasons. First, unique among U.S. cities, NYC has a legal right to shelter, the consequence of a series of consent decrees originating in the 1980s[13]. The City is legally obligated to provide emergency accommodations to any family able to demonstrate it has no suitable alternative. The second factor is NYC's relentless real estate market. In the decade ending in 2015, median rent in NYC grew three times the pace of median incomes (18.3 percent versus 6.6 percent). Vacancy rates are consistently below 4 percent (NYU Furman Center, 2016). According the City, demand for affordable apartments exceeds supply by a factor of two; approximately half of renters in the City are rent-burdened, defined as allocating more than 30 percent of household income to rent (NYC Mayor's Office, 2017).

Responsibility for managing shelters and supports for homeless families and individuals falls primarily to DHS, an agency under the purview of the City's much larger Department of Social Services (DSS)[14]. Families apply for shelter at a central intake center in the Bronx. The eligibility determination process requires families to prove they have no suitable housing alternative. State guidelines and court orders govern these determinations, but City policy-makers retain considerable discretion. Families deemed eligible are given shelter assignments by dedicated placement staff, who take into account such criteria as family size, health issues, safety, and proximity to children's schools[15].

[12]Los Angeles, which has a fifth the number of homeless families as NYC, has the second largest homeless family population among U.S. cities; 21 percent are unsheltered (U.S. Department of Housing and Urban Development, 2018).

[13]The state of Massachusetts also has such a right. See NYC Independent Budget Office (2014) and University of Michigan Law School (2017) for more detail.

[14]DHS was originally a part of DSS, but was spun off as an independent agency in 1993. In 2016, the two agencies were again consolidated under a single commissioner. Nevertheless, it remains conventional to refer to the departments as distinct. See NYC Department of Homeless Services (2019*a*) for more detail. DSS is also known as the Human Resources Administration (HRA). Accordingly, I use "DSS" and "HRA" interchangeably when referring to this agency.

[15]For more, see NYC Department of Homeless Services (2019*c*); NYC Independent Budget Office (2014). Additionally based on author's conversations with City officials.

The shelter system these families enter is vast and complex, consisting of traditional Tier II shelters[16], as well as “cluster” apartments scattered in otherwise private buildings and commercial hotels enlisted to expand capacity on-demand. In recent years, vacancy rates have hovered around one percent NYC Mayor’s Office (2017). Expanding capacity is complicated by the virulent community opposition that typically greets proposals for new shelters[17].

Shelter is also expensive. During fiscal year 2018, the average cost of sheltering one family for one night (inclusive of rent and services) was $192. Overall, DHS spent $1.2 billion on family homeless shelter—and this excludes administrative costs, prevention programs, and rental subsidies, as well as welfare benefits administered by other agencies (NYC Office of Management and Budget, 2019; NYC Mayor’s Office of Operations, 2018). While DHS does manage some shelters directly, most homeless services provision is carried out through contracts with community-based non-profit organizations who operate shelters[18].

Throughout this period, a pillar of the City’s homelessness strategy has been community continuity. To the extent capacity and other constraints allow, the City endeavors to place families in their neighborhoods of origin. Predicated on the goal of keeping children in their home schools, the policy reflects a more general premise—that families are better positioned to expeditiously return to permanent housing when they remain connected to their support networks, including relatives, friends, and places of work and worship (NYC Mayor’s Office, 2017). Since at least 1997, the share of families placed in shelters according to their youngest child’s school has been a DHS performance indicator. The official placement objective is the shelter nearest the youngest child’s school, but in practice DHS counts any placement within the school borough as successful (NYC Mayor’s Office of Operations, 2018). According to DHS officials, which families are given preferential local placement is essentially a function of what units are available at the time a family applies.

In recent years—after my study period—the emphasis on local placement has become even stronger, with the introduction of the School Proximity Project, through which DHS and DOE share data to identify homeless students and offer their families transfers to shelters closer to their schools.

[16] These are apartment buildings exclusively designated to serve homeless families.

[17] See, e.g, Stewart (2017).

[18] 82 percent of DHS’ budget consists of such contracts. This service arrangement is not unique to homeless services; most social service programs in the City are administered this way (NYC Mayor’s Office of Operations, 2017).

3 Theory

Homeless families' most pressing objective is to find permanent housing. Hence it is natural to adapt search theory to their context[19]. A *search effort model of family homelessness* parsimoniously characterizes my main results and offers generalizable insights.

Agents are homeless families, indexed by i and inhabiting a static, one period environment[20]. They start the period in shelter. Families value two goods, housing (H) and "consumption" (C), an aggregate good comprising everything besides housing, including leisure and work, that families value. Shelter (S) is a particular type of housing—namely, the least valuable kind: $S = \underline{H}$.

Families are endowed with a single resource: their own effort (e). Effort is normalized to a 0–1 scale, where 0 represents no effort expenditure and 1 represents maximal effort. A family's decision problem is to choose how to allocate effort[21] between housing search (e_S) and consumption ($e_C = 1 - e_S$). In choosing e_S, a family is choosing the probability it finds permanent housing.

Families' preferences are described by a continuously twice differentiable utility function $u(H, C)$, strictly increasing ($u_H, u_C > 0$) and strictly concave ($u_{HH}, u_{CC} < 0$) in both arguments (with subscripts denoting partial derivatives). In words, families value housing and consumption, there is diminishing marginal utility, and families are risk adverse. Also assume complementarity (or supermodularity), $u_{HC} > 0$. The pleasure of consumption increases with better housing, and housing is more satisfying when consumption is greater. Since shelter is the worst form of housing, it follows that $u_C(H, c) > u_C(S, c)$.

Neighborhoods affect families' valuation of homeless shelter as a housing good. The utility of families in shelter is $u(S(N), C)$, with N an indicator for local placement. I assume families prefer to be placed in their pre-shelter neighborhoods, so $u(S(N = 1), C) > u(S(N = 0), C)$.

Putting it all together, homeless families choose their housing search effort to maximize expected within-period utility:

$$\max_{0 \leq e_S \leq 1} (1 - e_S)u(S(N), C) + e_S u(H, C)$$

[19]Search theory, which typically considers job search, was pioneered by McCall (1970). Important contributions relevant for present purposes include Mortensen and Pissarides (1999); Pissarides (2000); Eckstein and Van den Berg (2007); Cahuc, Carcillo and Zylberberg (2014). Given that homeless families are in the receipt of government benefits (shelter) as they search for a good (housing), particularly useful are the insights of the unemployment duration and optimal unemployment insurance literatures (Chetty, 2008; Chetty and Finkelstein, 2013; Katz and Meyer, 1990; Lalive, Van Ours and Zweimüller, 2006; Spinnewijn, 2013).

[20]In what follows, I often omit the subscript i for simplicity.

[21]"Effort" does not imply that the object upon which it is expended is not enjoyable; excess can be thought of as being allocated to leisure. No effort is ever wasted.

$$\text{subject to}$$
$$C \leq w(1-e_S)$$

where w denotes the "wage" or, more generally, the return to effort not expended on housing search, inclusive of opportunity costs.

Assuming that the consumption constraint binds with equality at an interior solution, optimal housing search effort, e_S^* is implicitly defined by the first-order condition:

$$u(H,C) - u(S(N),C) - (1-e_S^*)wu_C(S(N),C) - e_S^* w u_C(H,C) = 0$$

Rearranging, I get the following expression, which makes makes the optimality condition intuitive to interpret.

$$\underbrace{u(H,C) - u(S(N),C)}_{\text{expected gain from search}} = \underbrace{w[(1-e_S^*)u_C(S(N),C) + e_S^* u_C(H,C)]}_{\text{expected loss from search}}$$

Families choose housing search effort so as to equate the (expected) benefit of search $(u(H,\cdot) - u(S(N),\cdot))$ with the expected utility cost of search, which is the product of the marginal opportunity cost of search (w) and the expected marginal utility of consumption, which depends on if the search is successful $((1-e_S^*)u_C(S(N),C) + e_S^* u_C(H,C))$.

Of primary interest is how this optimal effort changes based upon shelter neighborhood. Using the implicit function theorem, the comparative statics of neighborhood placement are straightforward to derive (with F denoting the implicit function defined by the FOC):

$$\frac{\partial e_S^*}{\partial N} = -\frac{\frac{\partial F}{\partial N}}{\frac{\partial F}{\partial e_S^*}} = \frac{\frac{\partial u(S)}{\partial S}\frac{\partial S}{\partial N} + w(1-e_S^*)\frac{\partial u_C(S)}{\partial S}\frac{\partial S}{\partial N}}{-w(\frac{\partial u(H)}{\partial C} - \frac{\partial u(S)}{\partial C})} = \frac{+}{-} < 0$$

where the consumption arguments in the utility function are suppressed for clarity and $\frac{\partial C}{\partial e_S^*} = -w$. Since the numerator is positive (being placed locally increases the marginal utility of being housed in shelter, and the marginal utility of consumption increases with being placed locally) and the denominator is negative (by exerting effort to search for housing, families give up consumption, which is valued more when in permanent housing), optimal search effort decreases when families are placed in their neighborhoods of origin[22].

Intuitively, families prefer permanent housing to shelter, but being placed in a local shelter narrows the gap. Thus, when placed locally, families have less incentive to search.

[22]Note that, in this setup, the level of intra-period consumption is the same whether or not families are successful at finding permanent housing.

Because e_S^* measures the probability of finding permanent housing,

$$E(Y) = \frac{1}{e_S^*}$$

gives the expected duration (length of stay) of the shelter spell. The model predicts families placed locally will remain in shelter longer because they allocate less effort to search.

On the other hand, since $e_C = 1 - e_S$, the effect of local shelter placement on "consumption" outcomes—of which labor market earnings and benefit receipt are of greatest interest—is positive.

$$\frac{\partial e_C^*}{\partial N} = -\frac{\partial e_S^*}{\partial N} > 0$$

That is, when families devote less effort to housing search, more effort is available to pursue earnings opportunities or apply for government benefits, like Cash Assistance.

I can also rearrange the FOC to get an expression for optimal search effort e_S^* in terms of the primitives of the model:

$$e_S^* = \frac{u(H) - u(S) - wu_C(S)}{w(u_C(H) - u_C(S))}$$

It is easy to show, given my assumptions, that this expression is strictly positive. Further, the following is a necessary and sufficient condition for an interior solution (i.e., optimal search effort less than unity):

$$wu_C(H) > u(H) - u(S)$$

In words, families will not spend all their effort on housing search when the utility of consumption they must give up to do so exceeds the utility of housing they gain[23].

A simple way to introduce heterogeneity is by allowing w, the opportunity cost of search, to depend on family characteristics $\mathbf{X}$. For simplicity, consider $\mathbf{X} = X$, a one-dimensional measure of resources (e.g., extended family support or savings); equivalently, it can be interpreted as an absence of constraints (e.g., having a small family). Assume that $\partial w / \partial X < 0$. The opportunity cost of search decreases with resources. The more supports or fewer constraints a family has, the less consumption it gives up by devoting effort to search. For any level of housing search effort, high resource families consume more.

Of primary interest is how optimal effort changes with resources. Differentiating the

[23]The term for consumption utility in shelter does not enter into the equation, as maximal search effort implies finding housing with certainty.

expression for e_S^* with respect to X,

$$\frac{\partial e_S^*}{\partial X} = \frac{(-w_X u_C(S))(w(u_C(H) - u_C(S))) - (u(H) - u(S) - wu_C(S))(w_X(u_C(H) - u_C(S))}{(w(u_C(H) - u_C(S))^2} = \frac{+}{+} > 0$$

where, as before, subscripts represent partial derivatives. The first term in the numerator is positive, as $w_X < 0$, as is the second term, given that the FOC implies $u(H) - u(S) > wu_C(S)$. The denominator is obviously positive, which means $\partial e_S^* / \partial X > 0$. Optimal search effort increases with resources; equivalently, it decreases with constraints.

4 Data and Sample

My data derives from administrative records linked across several City and State agencies. The main source is DHS' Client Assistance and Rehousing Enterprise System (CARES), the City's management information system of record for homeless families. My base data consists of all eligible family shelter entrants—adult(s) with one or more children under 21, or pregnant—who applied, were found eligible, and began their shelter stays in the period beginning January 1, 2010 and ending December 31, 2016. CARES provides detailed information characterizing family attributes and shelter stays. To this core DHS data, I append data on public benefit use and labor market experiences maintained by other agencies.

My unit of analysis is the *family-spell*. A homeless spell is defined as a shelter stay uninterrupted by a break of more than 30 days[24]; families returning after 30 days are considered to have begun a new spell. Many families experience multiple spells during the sample period. After removing from the raw data records with decisively missing data[25], my complete sample consists of 68,584 family-spells. This is a near-census of family homelessness. As shown in Table 1, my analytical sample shrinks for three reasons. First, 7,178 families originate from outside NYC. Another 286 spells lack data on borough of origin[26]. Finally, I limit my analytical sample to those families whose oldest child is under 18 years of age[27]. Henceforth I refer to these remaining 59,253 family-spells as my "Full Sample."

[24] This is the definition DHS conventionally uses in its own reporting.

[25] The unit of observation in the raw CARES data is the individual. Decisive fields include family identifier, entry dates, and the presence of children.

[26] My preferred measure of address of origin are geocoded addresses. 5,395 spells fail to geocode due to data entry errors. A redundant CARES "NYC Borough" field allows me to recover borough for 5,109 of these spells.

[27] Individuals 18 and over can be a head of household.

As robustness checks, I also consider three alternative samples: a "Non-DV" sample consisting of families eligible for shelter for reasons other than domestic violence (many DV families are deliberately placed out-of-borough for safety reasons), a "Pre-2015" sample consisting of all spells in the 2010–2014 period (to minimize censoring issues), and a "One School-Age Child" sample (to address potential multi-child confounding in my RD design).

Most variables are defined and measured at the time of shelter entry. For group characteristics shared by family members, like shelter assignment, I assign the shared value to the family. For individual characteristics that vary among members, such as age or sex, I assign the family the value of its (initial) head. For aggregate characteristics, like family size, I violate the "at-entry" rule and assign the family its maximum for the spell, to better reflect true composition.

In the remainder of this section, I discuss key variables conceptually and define their implementations in the data. Additional detail can be found in the Appendix.

4.1 Outcomes

The outcomes I assess are comprehensive, spanning shelter experiences, public benefit use, and employment. The most proximate and policy salient is length of stay (LOS) in shelter—a measure, in days, of the time between a family's entry into shelter and its exit, including gaps of up to 30 days[28]. As the most *immediate* shelter outcome, length of stay is the one most likely to be impacted by neighborhood placement; in turn, it impacts—and is impacted by—other outcomes, including families' experiences in the markets for labor and government benefits. In my analysis, I take the natural log of this duration.

Shelter exits must balance speed-of-transition with stability. A second outcome—return to shelter within a year of exit (after having been out of shelter for more than 30 days)—quantifies at this objective. My third outcome is an indicator for subsidy receipt; the presence of rental assistance is perhaps the most policy-relevant way to characterize shelter departures. I observe families' stays, exits, and returns through May 2019.

I also consider economic outcomes beyond housing: public benefit use and labor market experiences. The former, non-housing public benefits, derive from records maintained by the City's Department of Social Services (DSS), spanning 2001–2016. DSS, the City's designated Local Social Service Agency, oversees virtually all aspects of the social safety net, including the two most important income supports for homeless families: Cash Assistance[29] and Food

[28] In DHS parlance, this is known as "system" LOS, because it reflects a family's overall attachment to the homeless services system, regardless transient absences. It is not uncommon for families to leave shelter for a few days, then return. An alternative duration measure, "shelter" LOS, excludes the interludes from the count. The measures produce similar results.

[29] Cash Assistance consists of Temporary Assistance for Needy Families (TANF), which, in New York, is

Stamps[30]. I measure Cash Assistance and Food Stamps use with indicators for active cases at any time during a period of interest. I focus on two periods: the year post-shelter entry and the year post-shelter exit.

To assess labor market outcomes, I use quarterly earnings records from the New York State Department of Labor (DOL) spanning the first quarter of 2004 to the first quarter of 2017. Again focusing the years post-entry and post-exit, I construct indicators for positive earnings during any quarter as my measures of employment[31]. Correspondingly, my measure of earnings is log average quarterly earnings[32].

Public benefit and labor market outcomes require cross-agency data matches. Because individual identifiers vary by program (and are subject to administrative error), I use probabilistic matching techniques to link DHS and DSS data[33]. The DHS-DOL link is deterministic based on Social Security Number (SSN). I assume that non-linkages between DHS families and DSS/DOL records mean that families are truly not receiving benefits or not working.

4.2 Treatment

In my leading case, I define treatment as in-borough placement[34]. Origin address is defined as the family's "last known address" reported to DHS[35]. A small share of families (less than 4 percent) report other shelters as their prior addresses. In light of this, and given that unstably

referred to as Family Assistance, and its State counterpart for single adults and TANF time-limited families, Safety Net Assistance. Sometimes described as "public assistance" or "welfare," Cash Assistance provides unrestricted monetary transfers to poor individuals and families. Eligibility is limited to the very poorest and imposes work requirements. Benefits are similarly tight, topping out at $789 a month for a three-person family. 332,407 New York City residents were actively receiving Cash Assistance as of August 2019 (Cohen and Giannarelli, 2016; New York State Office of Temporary and Disability Assistance, 2016*b*, 2015*b*, 2017; NYC Human Resources Administration, 2019).

[30]Food Stamps, officially known as the Supplemental Nutrition Assistance Program (SNAP), provides low-income families with monthly dollars that must be spent on food. SNAP eligibility standards are less strict than Cash Assistance; correspondingly, its caseloads are much larger. In 2019, a family of three receives $509 monthly. 1.5 million NYC residents received SNAP as of August 2019 (New York State Office of Temporary and Disability Assistance, 2019; NYC Human Resources Administration, 2019).

[31]DOL data lacks information on work hours.

[32]Average quarterly earnings themselves are in real 2016 dollars, are inclusive of all quarters, whether working or not, and have one dollar added, so as to avoid omitting families with zero earnings when taking logs. For partially-censored spells, the earnings denominator is the minimum of four quarters or the number of quarters before censoring.

[33]There are several so-called "fuzzy matching" techniques standard in the computer science and statistics literatures. In this study, I primarily rely upon the user-written Stata command `reclink2`, which utilizes a bigram (two-character) string comparator (Wasi, Flaaen et al., 2015).

[34]NYC is comprised of five boroughs, which are analogous to counties: Manhattan, The Bronx, Brooklyn, Queens, and Staten Island.

[35]After cleaning, standardizing, and parsing addresses into distinct fields, I use the NYC Department of City Planning's Geosupport Desktop Edition application (GBAT), version 17.1, to classify origin and shelter addresses by borough, school district, and spatial X-Y coordinates.

housed family may move frequently, it is best to interpret origin addresses as places where families have preexisting community ties. Correspondingly, I define shelter neighborhoods in terms of *initial* shelter assignments. During their stays, families may be offered transfers to more proximate shelters; because within-spell moves are at families' discretion, I consider only initial assignment. Since some "control" families end up treated, this will have the effect, if any, of attenuating my results. In my Full sample, 51 percent of families are placed in their boroughs of origin.

For robustness, I also consider a continuous treatment definition: Euclidean (straight-line) distance, in miles, between origin and shelter addresses[36]. The average in-borough family is placed in a shelter 2.7 miles from its previous address, while the average out-of-borough one is placed 9.3 miles away. As a second check, I define neighborhoods in terms of the City's 32 geographical school districts, which are administrative boundaries for the public school system. 10 percent of families are placed in their neighborhood of origin by this standard.

4.3 Covariates

The extensive detail in my linked administrative data allows me to control for a rich set of observables. I group my covariates into three sets: placement characteristics, family characteristics, and shelter characteristics. Together, I refer to the complete collection of these variables as *Main covariates*.

Placement characteristics are factors upon which the natural experiment is conditioned. A cubic in year of shelter entry controls for time trends. Month fixed effects control for seasonal trends. Borough-of-origin dummies address systematic geographical disparities in treatment probabilities (i.e., boroughs are equal neither in shelter capacity nor shelter entrants). I also control for the four factors expressly considered as placement criteria. Family size is an integer count of unique individuals present at any time during a shelter stay. Number of children under 18 is analogously defined (both include non-relative case members). Health issue is a dummy equal to one if any family member has a medical, mental health, or substance abuse issue, and is based on screenings performed by DHS and providers at intake and during shelter stays[37]. Official eligibility reason is a set of six dummies: eviction, overcrowding, housing conditions, domestic violence, other, and unknown. DV status is particularly relevant to shelter placements, as safety concerns are paramount. I also include

[36]This measure is calculated from Cartesian geospatial coordinates.

[37]I interpret missing values of the health issues indicator as indicative of good health; families not receiving a screening are assumed not to have significant limitations. This assumption is strengthened by the fact that my data derives from authoritative administrative records.

an integer count of oldest child's (potential) grade—my RD running variable—both to ensure comparability between estimation methods and because this age factors into placement decisions.

Family characteristics describe families' compositions and circumstances, while proxying for unobservables. Female is a dummy that is equal to one for female head of family and zero otherwise. Age is a continuous measure, in years, of the duration between a head's date of birth and shelter entry date. Race consists of six mutually exclusive categories: White, Black, Hispanic, Asian, Other, and Unknown (if race is refused or missing). Partner present is a dummy equal to one if a head's significant other is present in shelter, whether or not such a partner is a married spouse. Pregnancy is a dummy equal to one if a family indicates a pregnant member at shelter entry. Education consists of four mutually exclusive categories: no degree (less than high school), high school graduate, some college or more, and unknown[38]. On Cash Assistance and On Food Stamps are dummies equal to one if a family has an active benefit case in the respective program at the time of shelter entry. Log average quarterly earnings in the year prior to shelter entry is analogous to the earnings outcomes defined above.

The final category of controls are *shelter characteristics*: variables related to a family's shelter assignment. These include four categories of facility type (Tier II shelter, cluster unit, commercial hotel, and other) and five dummies for shelter borough. In my "Shelter" specification, I also include dummies for the 271 individual "facilities" into which families in my sample are placed. These dummies proxy for unobservable shelter and provider characteristics[39].

4.4 Censoring

My analysis is complicated by the flow nature of my sample. I do not observe all families for the same length of time, and some outcomes for some families are right-censored. For outcomes derived from DHS records (length of stay, subsidized exits, and one-year returns), this issue is minimal, as my CARES data extends through May 2019. Only 2 percent of my sample have censored stays. Slightly more, 5 percent, are not observed for a full year following shelter exit (see Table A.1).

However, my DSS data only extends through 2016 and my DOL data through the first quarter of 2017. Thus, for these outcomes, I take care to define censoring-resilient measures,

[38] Education level derives from DSS records. While some families do not report education, non-matches between DHS and DSS account for most of the unknown cases.

[39] Given facility codes for cluster units encompass many distinct buildings, the latter interpretation of these fixed effects as indicative of provider influences is probably more accurate. Six facilities have singleton observations and are dropped from the Full sample in this specification.

focusing on one-year windows following shelter entry and exit, so as to put families on as equal footing as the data allows[40]. Because observations can still be censored within these year intervals, I also prioritize indicator or rate variables, which can at least be partially defined during partially-censored years. Nevertheless, I do not observe a full year of post-entry public benefit outcomes for 16 percent of family-spells. Post-exit, 34 percent of family-spells have incompletely observed benefit outcomes; 30 percent have censored labor market results.

The vast majority of this censoring occurs for family spells beginning in 2015 or 2016. Since the censoring mechanism is primarily an artifact of the data collection process, I make the standard assumption that it is as-good-as random and therefore will primarily attenuate my results toward zero. This assumption will hold so long as longer-staying early-year family shelter entrants are representative of longer-staying later-year ones. Nevertheless, for robustness, I replicate most of my main analyses for a sample of pre-2015 entrants[41].

5 Empirical Approach

5.1 OLS: A Shelter Scarcity Experiment

In my main analysis, I define treatment for family i during homeless spell p as an indicator in-borough placement, $N_{ip} = \mathbf{1}\{boro_{ip,origin} = boro_{ip,shelter}\}$. Correspondingly, Y_{Nip} is a potential outcome for family i. If as DHS suggests, shelter assignments are truly quasi-random once shelter entry contexts and placement criteria are taken into account, I can make the conditional independence assumption $\{Y_{ip0}, Y_{ip1}\} \perp N_{ip}|\mathbf{X}_{ip}$, where $\mathbf{X}_{ip}$ includes all covariates (including fixed effects and a constant) in a particular model. My general estimating equation is:

$$Y_{ip} = \mathbf{X}_{ip}\boldsymbol{\beta} + \tau^{OLS} N_{ip} + \varepsilon_{ip} \tag{1}$$

Under the CIA, unobservables, ε_{ip}, are unrelated to treatment ($E[\varepsilon_{ip}|\mathbf{X}_{ip}] = 0$), and so OLS consistently estimates the *average treatment effect* (ATE) of neighborhood placement, $ATE = E[Y_{1ip} - Y_{0ip}|\mathbf{X}_{ip}] = \tau^{OLS}$.

[40] When quarters are the unit of time, all such periods are defined as excluding the quarter of transition and inclusive of the following four quarters. When days are the time unit, periods begin on the day of transition and extend for the the next 365 days, inclusive. I also follow the same approach when controlling for pre-shelter earnings, considering the year prior to shelter entry.

[41] An earlier version of this paper, based on entirely on data observed through 2016, included an extensive discussion about the the econometrics of censoring and presented results for a variety of censoring methods, including survival analysis and selection models. The major prediction was that treatment effects would be attenuated in the presence of censoring, and indeed that is what I find. The earlier version of the paper is available upon request.

I focus on four covariate specifications, the components of which are described in Section 4. My *Base* specification is a simple bivariate mean comparison. My *Placement* specification controls for factors expressly implicated in families' placement assignments. My *Main* (preferred) specification augments the Placement specification with additional family and shelter characteristics. My *Shelter* specification includes facility fixed effects and narrows the unit of comparison to distantly- and locally-placed families in the same shelter. I cluster standard errors at the "family group" level[42].

5.2 Instrumental Variables: Exogenous Policy Shocks

In Section 6, I present evidence in favor of random assignment. But even when OLS consistently estimates ATE's, it is silent on response heterogeneity, τ_{ip}, which is particularly policy-relevant when resources are scarce. Instrumental variables identify natural experiments in their own right, estimating *local* average treatment effects (LATE's) among compliers whose treatment statuses are affected by the instrument[43]. By isolating impacts among families at various treatment margins (which, in general, differ by instrument), these localized experiments can reveal the distributional aspects of policy.

At the same time, the evidence for random assignment is favorable, but not dispositive; family unobservables, which even detailed administrative data cannot inform, may still bias results. Thus, IV can also play its more traditional role of guarding against endogeneity. The difference is one of interpretation.

My IV approach exploits exogenous variation in the City's homeless policy writ large. Neighborhood-based shelter placements are but one element of the City's complex and perpetually-evolving homelessness strategy. Front-door policies, like those influencing eligibility determinations, affect the pace of shelter entry, while back-door approaches, like rental subsidies, impact exit rates[44]. These flows influence the likelihood of local placement: the faster is the entry current or the slower is the exit stream, the worse is an eligible family's chance of a well-matched placement. Equally important, front-door and back-door policies are driven by political, budgetary, and operational considerations independent of families' potential outcomes and treatment statuses. In other words, these policy changes are exogenous shocks—a second layer of quasi-random variation—to doubly justify the natural

[42]Family groups, which I define with an algorithm linking all families with at least one overlapping member, address the evolution of family structures during my sample period as well as multi-spell families.

[43]For details on LATE's, introduced by Imbens and Angrist (1994) and Angrist, Imbens and Rubin (1996), see Angrist and Pischke (2008). Heckman and Vytlacil (1999, 2005, 2007), Vytlacil (2002), and Heckman, Urzua and Vytlacil (2006) show LATE's, as well as ATE's and other conventional treatment effect parameters, can be derived as weighted averages of underlying marginal treatment effects (MTE's).

[44]Also important are shelter conditions, but these are harder to measure.

experiment assumption.

I consider two such instruments. The first, borrowed from Cassidy (2020), focuses on the front door: the family shelter ineligibility rate. Although the City is legally required to house needy families, the rigor of the application process provides ample room for administrative discretion, typically with regard to the stringency with which disqualifying rules are enforced[45]. As can be seen in the top panel of Figure 1, the large changes in the ineligibility rate are associated with new commissioners, and the most striking shift came when Bill de Blasio replaced Mike Bloomberg as mayor in 2014. Other big swings coincide with well-publicized policy initiatives, such as the City-negotiated modifications to State eligibility rules that took place between September 2015 and November 2016[46]. The figure also makes plain the strong relationship between eligibility policy and in-borough placement[47].

Specifically, my first instrument is the 15-day moving average of the initial ineligibility rate for rolling 30-day application periods[48]. For family i entering shelter on day $D = d$, my instrument Z_{id}^{IE} is defined as average ineligibles divided average applications:

$$Z_{id}^{IE} = \frac{\frac{1}{15}\sum_{D=(d-14)}^{d}\sum_{i\in D}\mathbf{1}\{O_i = inel\}}{\frac{1}{15}\sum_{D=(d-14)}^{d}\sum_{i\in D}1}$$

with $\mathbf{1}\{\cdot\}$ the indicator function and O_i a random variable denoting family i's application outcome, which may be eligible, ineligible, diversion, or made own arrangement (voluntarily withdrawn or incomplete).

For the ineligibility rate instrument to be valid, it must satisfy four well-known conditions. First-stage relevance is empirically obvious. Monotonicity follows from the reasonable assumption that less competition means better chances of local placement for all families.

Independence requires that the ineligibility rate not influence the mix of shelter entrants; because ineligibility policy can select the eligible families who comprise my sample, this is a nontrivial concern. In Cassidy (2020), I present detailed evidence that this sort of sample

[45] Families are deemed ineligible for two broad reasons—failure to comply with application procedures or availability of other housing—both of which, in part, are subject to interpretation. For more detail, see the discussions in NYC Independent Budget Office (2014); Routhier (2017*a*); Harris (2016).

[46] O'Flaherty (2019) discusses these policy changes in detail. See also: Jorgensen (2017); New York State Office of Temporary and Disability Assistance (2016*a*); New York State Office of Temporary and Disability Assistance. (2015*a*); Fermino (2016*a*); Eide (2018); New York Daily News Editorial (2014); Fermino (2016*b*); Katz (2015); Routhier (2017*b*).

[47] Figure A.1 gives seasonally-detrended versions of these graphs, which makes the relationship even clearer.

[48] Families can apply for shelter multiple times; a month is the conventional agency standard for defining discrete spells of housing instability. New periods begin following gaps of more than 30 days from families' previous applications. Periods "roll" by resetting the 30-day clock with each application. "Initial" refers to the outcome of a family's first application within a period. The 15-day moving average includes each family's date of shelter entry and the 14 days prior, weighted in proportion to daily applications; it is simply a device to smooth out noise.

selection does not take place. Families entering during periods of high and low eligibility are remarkably similar. A major reason why is that families may apply for shelter as many times as desired. Even in strict policy environments, most are eventually determined eligible; tight policy operates primarily by slowing the pace of shelter entry rather than preventing entries completely.

Exclusion correspondingly demands that the effect of eligibility policy on outcomes operates entirely through its impact on local placement. One challenge is that eligibility policy may be correlated with other policy changes. I address this concern by including a cubic in years in all of my regressions, so as to capture general contextual trends without overfitting. What's more, eligibility policy is the most direct front-door intervention, so coincident policy changes that are part of the same broad homelessness strategy eligibility policy reflects can reasonably be seen as supplemental contributors. To err on the side of caution, I interpret my IV results as weakly satisfying the exclusion restriction: approximations of true LATE's that may be mildly influenced by the directs effects of related policies.

My second instrument, original to this paper, elaborates on the first by incorporating back-door policies—specifically, subsidized shelter exits. In an effort to shorten stays and strengthen housing stability, the City has implemented a variety of rental assistance programs over the years. Typically offering time-limited benefits and requiring family contributions, these programs, which are often conditioned on criteria such as employment and income, help families transition to permanent housing.

I refer to my second instrument as the "aversion ratio," Z_{id}^{AR}. It gives the shelter census averted by policy normalized by the number of entrants:

$$Z_{id}^{AR} = \frac{\overline{SE} + \overline{IN}}{\overline{EL}}$$

where SE is a count of subsidized exits, IN is a count of ineligible families, EL is a count of eligible families, and the bars denote 15-day moving averages, e.g., $\overline{SE} = \frac{1}{15}\sum_{D=(d-14)}^{d} SE_d$. As shown in the bottom panel of Figure 1, the aversion ratio has an even tighter correspondence with movements in the probability of in-borough placement than does ineligibility alone; accounting for both front- and back-door policies makes the instrument stronger. The arguments required to justify independence and exclusion are similar to before, with the obvious extension that the absence or presence of rental assistance programs doesn't alter potential outcomes except through their influence on treatment probabilities. As with front-door policies, the availability of rental assistance programs depends largely on political and budgetary factors orthogonal to family characteristics. For example, the primary rental assistance program during the Bloomberg years ended with great fanfare in 2011 due to funding dispute between the City and State (Secret, 2011; Edwards, 2012), while the the de Blasio administration was quick to roll out its successor, Living in Communities (LINC) upon taking office in 2014 (NYC Mayor's Office, 2017).

I use the ineligibility rate and aversion ratio instruments separately in standard two-stage least squares (2SLS) estimation, with Equation 1 representing the second stage (with actual treatment, N_{ip} replaced by first-stage predicted treatment, $\widehat{N_{ip}}$) and first stages given by:

$$N_{ip} = \mathbf{X}_{ip}\boldsymbol{\pi}_0 + \pi_1 Z_{ip} + \nu_{ip} \tag{2}$$

where Z_{ip} is either of the instruments and ν_{ip} is the error.

The resulting estimates of τ^{IE} and τ^{AV} are LATE's among their respective compliant subpopulations. Given the variation that produces these localized experiments stems from big-picture homeless strategy, these instruments isolate treatment effects among families, who as a logical matter, face augmented barriers to local placement: they are treated only when the policy environment makes doing so especially easy. If, as might be anticipated, the responses of these marginally-treated homeless families are distinct from the average responses OLS identifies, it is of considerable interest to understand who these families are. Put differently, carefully chosen instruments—i.e., policy variables that influence treatment participation margins—can identify treatment effects that are policy relevant in the sense of Heckman and Vytlacil (1999, 2001, 2005, 2007).

Accordingly, I supplement my IV analysis with additional exercises characterizing compliers. While it is fundamentally impossible to identify individual compliers, it is possible to estimate their average characteristics. Angrist and Pischke (2008) show how to do this in the canonical binary instrument case; Dahl, Kostøl and Mogstad (2014) and Dobbie, Goldin and Yang (2018) implement an analogous procedure for continuous instruments. In Cassidy (2020), I extend this work to incorporate explicit hypothesis tests and continuous characteristics. I follow the same procedure here[49].

5.3 Regression Discontinuity: A Boost at School-Starting

A complementary identification strategy exploits policy rules native to the neighborhood placement policy itself. The policy is, expressly, an educational policy: the explicit goal is to place families near their youngest children's *schools*[50]. This lends itself to a regression

[49]In brief, this algorithm uses first-stage regressions and convenient conditional probability equivalences to estimate the relative prevalences of traits in the compliant subpopulation; standard errors are calculated through bootstrap resampling. Details are provided in the empirical appendix of Cassidy (2020).

[50]Most students in NYC attend their residentially-zoned school, so placement near a youngest child's school usually means older siblings are near their schools as well.

discontinuity design[51]. Families whose oldest children are younger than school age have a less compelling case for local placement than do those with school-age children. While DHS seeks to place all families in their origin boroughs, those with student members get priority.

My RD setup is both discrete and fuzzy, which introduces several non-standard issues[52]. My running variable is the potential grade attained by a family's oldest child during the year of shelter entry: $A_{ip} = \lfloor \frac{EOY - DOB}{365.25} - 5 \rfloor$, where EOY is December 31 of the shelter entry year, DOB is date of birth, and the L-brackets indicate the floor operator. In, NYC, children are eligible for, and required to, attend kindergarten in the calendar years they turn five, so this assignment variable gives families' oldest children's potential grades, normalized so that zero is kindergarten. Policy dictates this running variable be discrete: age matters in years. There are 16 support points, $A_{ip} \in \{-3, -2, \ldots, 11, 12\}$[53].

Because having a school-age child increases the chances of local placement but does not guarantee it, my RD is fuzzy. What changes sharply at the school starting threshold is treatment assignment, not treatment status. It follows that school-age threshold crossing, $T_{ip} = \mathbf{1}\{A_{ip} \geq 0\}$, is an instrument for local placement.

Discrete fuzziness dictates my RD analysis reduces to standard IV (Angrist and Pischke, 2008). Traditional RD concerns—local polynomial choice and bandwidth selection—are simplified. I estimate two categories of models, which I refer to as "Wald" and "Linear." The general form of my Wald equation is

$$\begin{aligned} N_{ip} &= \mathbf{X}_{ip}\boldsymbol{\pi}_0 + \pi_1 T_{ip} + \nu_{ip} \implies \widehat{N}_{ip} && \text{(first stage)} \\ Y_{ip} &= \mathbf{X}_{ip}\boldsymbol{\pi}_1 + \tau^{RDW}\widehat{N}_{ip} + \varepsilon_{ip} && \text{(second stage)} \end{aligned} \tag{3}$$

The Wald setup is based on local randomization approach to RD inference (Cattaneo, Idrobo and Titiunik, 2018). The key assumption is that treatment assignment is as-good-as-random in some neighborhood of the assignment cutoff. Rather than make any assumptions about functional forms in the neighborhood of the cutoff, I simply pool the running variable for a limited set of support points at or near the threshold.

I vary this model across three dimensions: bandwidth, threshold, and covariates. For bandwidths, I use both the narrowest possible comparison, $A_{ip} \in \{-1, 0\}$, as well as a "wide Wald" frame expanded to two support points on either side of the threshold. Second, to address variability in school-starting age (discussed below), I variously include and exclude

[51] For details on RD, see, e.g., Hahn, Todd and Van der Klaauw (2001); Imbens and Lemieux (2008); Lee and Lemieux (2010); Cattaneo, Idrobo and Titiunik (2018, 2017)

[52] See Kolesár and Rothe (2018); Lee and Card (2008); Dong (2015); Frandsen (2017).

[53] I exclude $A_{ip} = \{-5, -4\}$ because families who enter shelter during children's birth years or soon thereafter have idiosyncratic outcomes.

families at the $A_{ip} = 0$ threshold. Exclusion yields a potentially sharper comparison, at the risk of being less representative. Finally, I present estimates both with and without Main covariates, with the following adjustment. My running variable is highly collinear with family size, number of children under 18, and head of household's age, so I replace the continuous measures with indicators for whether a family is above-median in these characteristics; I refer to this modified set as "Main RD" covariates.

More common than local randomization, RD proceeds from continuity assumptions: namely, that conditional expectations of treatment and outcomes, as functions of the running variable, are smooth on either side of the cutoff, with any discontinuity in extrapolated intercepts attributed to the effect of threshold-crossing (Cattaneo, Idrobo and Titiunik, 2017). My "Linear" models are rooted in this framework. I allow the slopes to differ on either side of the threshold, estimating the following set of equations by 2SLS:

$$\begin{aligned}
N_{ip} &= \mathbf{X}_{ip}\boldsymbol{\pi}_{10} + \pi_{11}T_{ip} + \pi_{12}A_{ip} \\
&\quad + \pi_{13}(A_{ip} \times T_{ip}) + \nu_{1ip} \implies \widehat{N}_{ip} && \text{(first stage)} \\
N_{ip} \times A_{ip} &= \mathbf{X}_{ip}\boldsymbol{\pi}_{20} + \pi_{21}T_{ip} + \pi_{22}A_{ip} \\
&\quad + \pi_{23}(A_{ip} \times T_{ip}) + \nu_{2i} \implies \widehat{N_{ip} \times A}_{ip} && \text{(first stage)} \\
Y_{ip} &= \mathbf{X}_{ip}\boldsymbol{\pi}_{30} + \tau^{RDL}\widehat{N}_{ip} + \pi_{32}A_{ip} \\
&\quad + \pi_{33}(\widehat{A_{ip} \times N}_{ip}) + \varepsilon_{ip} && \text{(second stage)} \qquad (4)
\end{aligned}$$

Given the normalization of the running variable, τ^{RDL} gives the estimated treatment effect at the threshold. As with the Wald estimates, I present several specifications, estimating the model (a) for global ($[-3, 12]$) and local ($[-3, 3]$) bandwidths, (b) including and excluding the threshold, and (c) with and without Main RD covariates.

For both Wald and Linear RD inference, I continue to cluster standard errors at the family group level. Following Lee and Card (2008), conventional practice for discrete RD has been to cluster on the running variable. However, recent research by Kolesár and Rothe (2018) demonstrates that these standard errors can be substantially too small, especially when, as here, there is a limited number of support points[54]. Since their results show traditional heteroskedasticity-robust standard errors are about as good as the more elaborate bias-corrected variants they propose, I stick with family group clustering, which, in any event, is standard in IV estimation and thus ensures comparability with my non-RD IV results.

For my RD to be valid, it must satisfy standard IV assumptions. Discontinuous treatment

[54] In related work, Dong (2015) offers corrections when the running variable is a discretized version of a continuous variable. Though my running variable falls in this category, I do not pursue it here, as the discrete age is the policy-relevant attribute.

probabilities at the threshold (i.e., first-stage relevance) is empirically clear and monotonicity is uncontroversial. The exclusion restriction is the highest hurdle. It must be that school starting affects potential outcomes only through its influence on treatment probabilities. In the homeless shelter context, preferential placements based on school enrollment make shelter assignments a major channel through which school-agedness effects are transmitted. But having a child start school frees up time that would otherwise be spent on child care for work and leisure. From a time allocation perspective, one would expect families with kindergärtners would have higher rates of employment and shorter shelter stays. On the other hand, stays could lengthen if desires to not disrupt school motivate families to delay move-outs.

Another challenge for validity is that school-starting age is itself fuzzy. Although most children attend kindergarten during their age-five years, parents may optionally enroll their children in prekindergarten at age four or defer school-starting until first grade at age six[55]. In addition, families may enter shelter at any time during their children's school-starting years, including prior to school enrollment (i.e., January–June). Setting age five (i.e., potential grade zero) as the strict treatment assignment threshold will thus impart some degree of misclassification. Fortunately, this is a minor concern. As the graphical evidence presented in Section 6 demonstrates, age five is the empirically obvious discontinuity point: while the probability of treatment rises about two percentage points per year up until age four, it gets a five percentage point bump at age five (from 46.9 percent to 52.0 percent). Part of the reason for the sharp divide is that my running variable, which is based on calendar year, captures four-year-olds in the second halves of their pre-K-eligible school years as among those assigned to treatment. My threshold-omitting and linear specifications, which are less sensitive to blurry treatment assignment, offer even sharper contrasts[56].

5.4 Family Fixed Effects: Multi-Spell Counterfactuals

My third identification strategy relies on the panel nature of my data. Repeat spells of homelessness are not uncommon. A fifth (10,390) of families in my sample have multiple stays during my study period (see Table A.2). When these families' treatment statuses vary across these stays, comparing own outcomes when placed locally and distantly is an

[55]Based on data from Cassidy (2020), I estimate at least 93 percent of homeless children start school by kindergarten; of these, slightly more than half attend pre-K. The City's introduction of universal pre-K in 2014 guaranteed all four-year-olds public pre-kindergarten spots. Prior to 2014, only a quarter of four-year-olds attended full-day public pre-K (NYC Mayor's Office, 2014).

[56]In principle, school-starting fuzziness could be remedied with data on children's actual enrollment statuses, which I lack due to confidentiality restrictions. However, actual enrollment status is potentially less desirable as an instrument, as school-starting is subject to parental choice endogeneity.

exacting way to estimate treatment effects. Implementing my family fixed effects estimator is a straightforward modification of Equation 1 to include individual student dummies, α_i. For family i in shelter spell p,

$$Y_{ip} = \alpha_i + \tau^{FE} N_{ip} + \mathbf{X}_{ip}\boldsymbol{\beta} + \varepsilon_{ip} \tag{5}$$

I continue to cluster standard errors at the family group level.

Consistency relies upon the assumption of no spell-varying unobservables. This assumption is strengthened by the presence of administrative covariates capturing broad classes of cross-spell variation. In addition, the underlying quasi-randomness of shelter scarcity continues to apply to each spell.

Prior research suggests homeless spells among low-income families are largely based on luck (O'Flaherty, 2010). It follows that those with multiple bad hands are representative of homeless families in general. At the least, their findings generalize to the considerable subsample of multi-spell families.

6 Results

6.1 Descriptives and Randomization Check

My first empirical task is to assess the plausibility of the natural experiment assumption. Is shelter assignment truly determined by a scarcity-based queuing?

Table 2 formally tests this proposition, while also descriptively summarizing the Full sample. The randomization check consists of separate bivariate regressions of baseline covariates and pre-shelter outcomes on an indicator for in-borough placement. The difference between treated (in-borough) and untreated (out-of-borough) families is the coefficient on treatment. If placements are truly random, these characteristics should be approximately balanced.

Due to the large sample size, group contrasts are often statistically significant, but they are rarely economically meaningful. Families placed in- and out-of-borough are virtually identical in family composition and education, as well as pre-shelter public benefit use, employment, and earnings.

The big differences are innocuous and expected. There is systematic variation in treatment probability by year, month, and borough. Shelter is relatively more abundant in the early years of my sample (when the homeless population is smaller), during the early months of the year (when fewer families enter shelter), and in the Bronx (where a plurality of shelters are located). Along related lines, treated families are more likely to be placed in cluster units (which are more common in the Bronx and earlier in the sample), while their untreated counterparts are more likely to be assigned to commercial hotels (which are more common in the other boroughs and later in the sample).

Other placement criteria matter, too. Due to safety concerns, families experiencing domestic violence are considerably less likely to be treated, accounting for 22 percent of in-borough placements but 37 percent of out-of-borough ones. Conversely, evictions are more common in-borough. Families with health limitations are also more challenging to place: 32 percent of out-of-borough families have health issues, compared with 28 percent of in-borough ones. In-borough families have older oldest children, averaging third grade, versus the out-of-borough average of second. Family heads are older, too.

Overall, the data supports the administrative impression that shelter placements depend upon availability, conditioned on placement criteria.

6.2 OLS Results

Tables 3A and 3B present my main OLS results. Given the evidence for conditional random assignment, these are my preferred ATE estimates. Each cell gives the coefficient on in-borough placement from a separate regression. Outcomes are listed in rows and organized into three panels. Panel A in Table 3A analyzes stays and returns—the most salient outcomes in the homeless services domain. Table 3B is split into two panels: year post-entry outcomes (B1), which refer to the year following a family's shelter entry (and is typically, but not always, spent in shelter), and year post-exit outcomes (B2), which refer to the year following shelter exit (and is typically, though not always, spent out of shelter). Column 1 gives outcome means. Columns 2–5 present sequentially more stringent covariates for the Full sample. Columns 6 (Non-Domestic-Violence) and 7 (Pre-2015) consider alternative samples for robustness. My preferred estimates are those in Column 4, which include Main covariates for the Full sample. Family-group clustered standard errors are given in parentheses. Sample sizes are given in braces under the first outcome in each panel, as well as for subsequent within-panel outcomes where the sample size differs from the first due to censoring.

As would be expected under random assignment, covariates beyond placement factors make little difference in the results. Focusing on Panel A's Main estimates (Col 4), families assigned in-borough stay 12.7 percent longer than those placed out-of-borough. With lengths of stay averaging 424 days, this implies in-borough families remain in shelter 54 days longer, though, the log specification acknowledges these effects may be non-linear. In-borough families are also 1.8 pp (4.6 percent) more likely to exit with a rental subsidy. They do not appear any more likely to return to shelter.

Panel B1 (Table 3B) shows that, during their years of shelter entry, in-borough families are 1.1 pp (1.4 percent) more likely to receive Cash Assistance. They are also 1.0 pp (2.1 percent) more likely to be employed and have 9.9 percent higher quarterly earnings. It is not clear whether the labor boost is due to preserving existing employment relationships or through new opportunities fostered by retained social ties. There is no impact on Food Stamps, likely because almost all homeless families receive it. Panel B2 illustrates that elevated Cash Assistance recipiency continues in the year post-shelter exit, by 1.7 pp (2.3 percent). During this year, the benefits connection extends to Food Stamps as well, by 0.8 pp (0.9 percent). But employment effects disappear.

This pattern of outcomes is consistent with the search effort model of shelter behavior. Homeless families respond to program incentives by allocating effort to their highest-value priorities. In-borough placement is preferred, so families stay longer and require additional impetus—rental assistance—to leave. Time otherwise spent on housing search is instead allocated to labor and consumption[57].

These findings remain consistent in my Shelter specification (Col 5), which controls for provider quality, as well as in the Non-DV (Col 6) and Pre-2015 (Col 7) samples, suggesting neither eligibility reasons nor censoring issues are driving my results.

Tables 4A and 4B present additional robustness checks, examining the same outcomes for treatment defined as school district placement and school-shelter distance, in miles, controlling for Main covariates. School district treatment (Col 1) confirms my Full sample results for length of stay (8.5 percent longer), entry-year employment (+1.8 pp), and entry-year earnings (+13 percent). However, other results are near zero or imprecise, likely for two reasons. First, only a small minority of families are placed in their school districts. Second, the stakes are higher for borough treatment: untreated families by the school district standard can still be quite close to their prior addresses. Being very close to home may be more important for jobs than it is for other outcomes.

Distance treatment broadly confirms my main findings, demonstrating that genuine proximity effects—rather than borough quirks—are at play. The Full sample (Col 4) results show that families stay 1.4 percent longer for every mile they are placed closer to their prior residences. At the average borough treatment distance gap of 6.6 miles, this translates to 9.4 percent longer stays. The probability of subsidized exit increases by 0.26 pp per mile closer

[57]One concern with this behavioral interpretation is that City rental assistance policy could be driving length of stay. If the City prioritized out-of-borough families for subsidies, longer stays for in-borough families would be an artifact of subsidy queuing. In Appendix C and Table A.6, I provide evidence that this is not the case: the effect of in-borough placement on length of stay is, if anything, strengthened when accounting for subsidies. Table A.7 confirms this is also true of my IV results.

to school, while the likelihood of Cash Assistance receipt increases 0.15 pp/mile post-entry and 0.16 pp/mile post-exit. Entry-year employment increases by 0.20 pp/mile closer and earnings by 1.6 percent/mile[58].

6.3 IV Results

Although I believe my OLS results credibly describe average policy responses in my quasi-experimental setting, prudent skepticism nevertheless dictates—and policy exogeneity permits—alternative identification strategies. Tables 5A and 5B present my main policy IV results. Similar in organization to Tables 3A and 3B, the first three columns assess the ineligibility rate instrument while the latter three analyze the aversion ratio.

Both instruments are very strong. First-stage F-stats, given in brackets (for the first outcome in each panel, as well as for subsequent outcomes with censored samples), are consistently above 20 for the ineligibility rate and double that for the aversion ratio. As expected, policy strictness increases the likelihood of local placement. A 10 pp increase in the ineligibility rate increases the chances of in-borough placement by 3.0 pp (Col 2), while an additional averted stay per unit entrant raises treatment probability by 6.1 pp (Col 5).

Length of stay continues to exhibit the most striking findings. LATE's for compliers are in the direction of OLS ATE's but an order of magnitude larger (Panel A). Per my Main specification (the point estimates for the Placement and Shelter specifications are similarly precise and slightly smaller in magnitude), families placed in-borough when the ineligibility rate is high but not otherwise stay four times longer (Col 2). Aversion ratio compliers (Col 5) stay 2.6 times longer when placed locally. Ineligibility rate compliers are also 29 pp more likely to return to shelter. The largest departure from OLS is that policy compliers are substantially less likely to exit with a subsidy: by 79 pp for the ineligibility rate and by 33 pp for the aversion ratio.

Compliers' use of other public benefits (Panels B1 and B2) are also more strongly influenced by proximity than homeless families overall. Continuing to focus on Main covariate specifications (Cols 2 and 5), ineligibility rate compliers are 65 pp more likely to receive Cash Assistance during their shelter entry years, and 43 pp more likely to receive it post-exit. LATE's for aversion ratio compliers are slightly smaller—34 pp entry year Cash Assistance, 27 pp exit year Cash Assistance—but still huge. As with OLS, there appears to be little effect on compliers' use of Food Stamps either during or after shelter. Unlike OLS, labor market impacts for compliers arise after shelter. There are no statistically significant effects for either instrument during the year post-entry. Post-exit, however, ineligibility rate compli-

[58]Table A.8 repeats Tables 3A and 3B for several alternative outcome definitions.

ers are 40 pp more likely to be employed. Aversion ratio compliers have a 34 pp employment boost—and earn seven times more.

These coefficients are large, but not implausible. Outcomes among homeless families have wide variation. A 400 percent increase in length of stay takes families from the median (294 days) to about the 95th percentile (1,246 days); the fifth percentile is just 20 days (see Figure A.6). Similarly, only a third of families are on Cash Assistance at shelter entry and just 43 percent work in the prior year, so the room for impact is large.

What's more, compliers—-who are placed in-borough *only* when policy makes it easy to do so—are families with considerable barriers to local placement. These constraints, discussed below, may also make it more difficult to find permanent housing, as well as generate inertial incentives to stick with in-borough shelter apartments that are nontrivial to obtain. Consequently, length of stay increases, allowing more time for other treatment effects to percolate. About 8 percent of my Full sample are ineligibility rate compliers and 10 percent comply with the aversion ratio (see Tables A.9 and A.10).

Tables 6A and 6B compare the average characteristics of ineligibility rate compliers with non-compliers, using the Dahl, Kostøl and Mogstad (2014) and Dobbie, Goldin and Yang (2018) procedure with a modified first-stage controlling for time trends and seasonality[59]. The most notable contrast is borough of origin. 57 percent of compliers are from the Bronx, compared with 39 percent of non-compliers[60]. Compliers also tend to be medium-large families: 39 percent have four or five members, compared with 28 percent of non-compliers. Other comparisons are imprecisely estimated. It should also be noted that these complier characteristics are indicative but not unqualified: majorities of large, young, and Bronx families are non-compliers, after all, so unobservables and characteristic interactions are clearly implicated[61].

Aversion ratio compliers (Tables 7A and 7B), like ineligibility rate ones, are more likely than non-compliers to originate from the Bronx (55 percent vs. 39 percent). The family size contrast loses statistical precision, though the point estimate (+6 pp for family size of 4–5) is similar. What becomes more notable is Cash Assistance receipt. Just 23 percent of aversion ratio compliers are on CA at shelter entry, compared with 37 percent of non-compliers[62].

[59]See Tables A.22 and A.23 for comparisons of additional characteristics.

[60]Compliers are also less likely to be African-American (43 percent vs. 57 percent) or sheltered in commercial hotels (8 percent vs. 29 percent), though these contrasts are likely explained by borough. Only 45 percent of Bronx entrants are Black, versus 55 percent of shelter entrants overall; likewise, just 21 percent of Bronx placements are in commercial hotels.

[61]Differences between this depiction of ineligibility rate compliers and that discussed in Cassidy (2020) are likely due to the facts that the latter (implicitly) weights results at the child level, includes only school-age children, and covers fewer years. In addition, that paper defines borough of origin in terms of school address.

[62]Further, 35 percent have health limitations, compared with 29 percent of non-compliers; while this contrast narrowly misses statistical significance, it is indicative of the finding in Cassidy (2020), where the

Large families from the Bronx disproportionately benefit when eligibility policy gets tighter or move-outs more common. The Bronx is where 41 percent of homeless families originate—by far the most of any borough—and also where the most out-of-borough families are placed (29 percent). Not uncoincidentally, PATH, the City's central intake center for homeless families, is also located there. When eligibility gets strict, applications become more labor-intensive; Bronx families have easier access, gaining an advantage as the out-of-borough flow slows. Large families also benefit from less congestion. The bigger a family, the harder is it to find suitable units; less competition improves the odds.

It is reasonable that large Bronx families also be especially responsive to local placement. The Bronx is small, isolated, and poor (U.S. Census Bureau, 2018), so treatment is more meaningful. In-borough placements are closer and out-of-borough ones further than non-Bronx averages. Competition for high-quality, affordable housing is fierce. Bronx families, especially large ones, fortunate to secure local placements thus have less incentive to leave.

At the same time, aversion ratio LATE's are generally 50–60 percent the magnitudes of their ineligibility rate counterparts. The difference in pre-shelter CA receipt may help explain why. As reflected by their lower reliance on public benefits—as well as large, precise post-shelter employment responses—aversion compliers would seem to be drawn from the higher end of the self-sufficiency spectrum.

A conservative perspective suggests interpreting these IV results as upper bounds. Both instruments are based on time variation and may pick up the effects of complementary policies (e.g., improved shelter quality). In my main results, I control for macro patterns with a year cubic. Table A.15 and Figures A.3 and A.4 detail a time trend sensitivity analysis. The OLS results are little changed. Sufficiently flexible trends absorb much variation in IV reduced forms, but robust first stages suggest overfitting rather than exclusion restriction violations is to blame: to the extent time trends capture correlated policy changes, these correlated changes appear small and eligibility policy remains independently informative[63].

For the skeptical reader inclined to think in terms of homogeneous effects and endogeneity, my IV results suggest OLS, if anything, is understating true policy impacts. But heterogeneity seems the more parsimonious story consistent with facts.

---

unit of complier comparison is school-age children. Aversion ratio compliers are also less likely to be in commercial hotels (-16 pp) or Black (-9 pp), with the latter marginally insignificant.

[63] Additional robustness checks for the ineligibility and aversion instruments are detailed in Tables A.13 and A.14, respectively. My main results are confirmed.

6.4 RD Results

Having a school-age child is a third instrument, with its own population of compliers: families placed locally only when they have school-age children. Figures 2–4 show how treatment and outcomes vary according to the running variable, oldest child's (potential) grade. Each graph plots mean outcomes and 95 percent confidence intervals by grade, along with linear trends fit separately on either side of the threshold. Left of the threshold, the regression is fit on the [-3,-1] interval and extrapolated from -5 to 0; the above-threshold regression is fit on the full [0,12] interval[64].

The top left panel of Figure 2 shows the fuzzy RD first-stage is strong. Although the probability of in-borough placement increases at young ages, there is an unmistakable boost when families' oldest children reach school age. Families whose oldest children are six are about 8 pp (17 percent) more likely to be placed in-borough than those whose oldest are four. Treatment probabilities remain basically flat at older ages, though there may be a slight bump around middle school starting (grade six)[65]. Length of stay exhibits an even starker discontinuity at school starting (Figure 2, top right). Exits and returns do not display decisive breaks (bottom panels) .

Figures 3 and 4 show entry- and exit-year benefit and employment outcomes, respectively. These results are, in general, nosier and treatment effects more muted. Cash Assistance displays the clearest discontinuity around school starting, with notable increases during the kindergarten ($A_{ip} = 0$) and first-grade ($A_{ip} = 1$) years, both during and following shelter (top left panels). Food Stamps appear unrelated to school-starting (top rights). Labor market outcomes are more nuanced (bottom panels). During the year of shelter entry, employment and earnings drop noticeably among families whose oldest children are in first or second grade, but hold steady, or even slightly increase, among those with kindergarten-age children. Post-shelter, there is slightly stronger evidence of an adverse labor market impact, especially with earnings, though it is difficult to disentangle discontinuities from general patterns of less employment among those who enter shelter with older children.

Tables 8A and 8B formalize the RD analysis, confirming the visual impression. As before, results are grouped into three panels, with each row considering a separate outcome. Column 1 gives Wald estimates for immediately adjacent threshold points ($A_{ip} = \{-1, 0\}$), while Column 2 excludes the threshold in assessing a symmetric two-year window ($A_{ip} = \{-2, -1, 1, 2\}$). Columns 3 and 4 assess global linear fits, the latter controlling for Main RD

[64]Negative "grades" should be interpreted as years before conventional school starting age. I exclude -5 and -4 in fitting the below-threshold regression due to unrepresentative patterns among families with very young oldest children.

[65]Figure A.8 shows an analogous pattern holds for distance treatment.

covariates.

Families whose treatment status is affected by having a school-age child stay about 3–7 times longer when placed in-borough (Table 8A)[66]. They are 35–66 pp more likely to leave shelter with a subsidy. But there is little evidence of impact on shelter returns.

There are few clear entry-year impacts on benefits and employment (Table 8B, Panel B1). The exception, as might be anticipated, is Cash Assistance, which has generally large positive coefficients, precisely estimated in the covariate-adjusted global linear specification (Col 4), suggesting a 18 pp increase in the probability of Cash Assistance receipt among compliers. Food Stamps and employment effects are unclear, though the balance of evidence for the latter is suggestive of mild negative impacts.

Exit-year effects are generally sharper (Table 8B, Panel B2). Cash Assistance is again the most striking result, with compliers 14–40 pp more likely to receive it, significant in all specifications. At the same time, local placement appears to adversely impact compliers' post-exit labor market outcomes. Point estimates for both employment and earnings are uniformly negative, though statistically significant only in the highly-powered wide Wald case (Col 2; -29 pp employment decrease; 4.5 times fewer earnings). Food Stamps impacts remain difficult to discern[67].

Correct inferences depend on whether families who enter shelter with young oldest children are suitable counterfactuals for those with school-age ones. Families congregating on either side of the threshold would be evidence of deliberate sorting that would invalidate RD identification. The histogram in Figure 5 demonstrates this is not the case: the frequency of shelter entry is smooth around the treatment threshold. The formal Frandsen (2017) test for the manipulation of a discrete running variable confirms this impression, delivering a maximum p-value of 0.832, which cannot nearly reject the null of no sorting.

A second implication of random assignment is that families below and above the treatment threshold be similar in baseline covariates and pre-shelter outcomes. To assess this proposition, Figures 6–8 repeat the RD plots for these characteristics, while Tables A.21A–A.21B provide the formal regression analysis[68]. The presence of threshold-crossing induced treatment effects for any of these "outcomes" is evidence that the RD independence and exclusion assumptions may be violated.

There are no discontinuities for most variables, including pre-shelter public benefit use

[66] To see this, note that $e^{1.065} = 2.9$ and $e^{1.986} = 7.3$.

[67] Tables A.16 and A.17 provide additional Wald and Linear specification permutations, respectively. Table A.18 reproduces the RD analysis for my three alternative samples. Table A.19 replicates the RD analysis distance treatment across all four samples. Table A.20 repeats Tables 8A and 8B for an alternative running variable: "potential grade" defined based on school years, starting in July and ending in June. The main conclusions remain unchanged.

[68] Figures A.11–A.13 give the three-year window versions.

and labor market outcomes, though employment and earnings peak among families whose oldest children are five. On the other hand, year of shelter entry (families with older oldest children enter in later years), housing conditions as an eligibility reason (less likely with school-age children), and education (those with school-age children are more highly educated) do have discontinuities at the threshold[69]. Overall, families around the school-starting threshold are comparable; most differences are expected.

A perhaps more important caution relates to representativeness: I estimate school-starting compliers constitute about one percent of my sample, or about a tenth the size of my IV complier populations. Nevertheless, school-starting families are an important subpopulation in their own right[70].

6.5 Family FE Results

My final identification strategy capitalizes on a different sort of natural experiment: multiple homeless spells. Tables 9A and 9B summarize the analysis. The first four columns assess the Full sample. Columns 5 and 6 consider robustness-check subsamples. The results are virtually identical to OLS; if anything, they slightly strengthen key findings. Per my Full sample Main specification (Col 3), families stay 17 percent longer when placed in-borough. Public benefit use is greater as well. They are 2.6 pp more likely to exit with a subsidy and 1.6–1.7 pp more likely to receive Cash Assistance during and after shelter. Entry-year employment increases by 1.7 pp and quarterly earnings by 15 percent. There is no evidence of impacts for Food Stamps or post-shelter labor market outcomes. The length of stay, subsidized, and entry-year Cash Assistance results hold for both alternative subsamples. The entry-year earnings finding holds for the Pre-2015 sample and the exit-year Cash Assistance result holds for the Non-DV sample. Other subsample point estimates are in the expected directions.

7 Conclusion

Homeless families placed in shelters in their neighborhoods of origin remain in shelter longer and are better connected to public benefits. Per the natural experiment of shelter scarcity—which justifies OLS identification and facilitates family fixed effects as well—*average* families

[69]In addition, there are threshold kinks in shelter locations, but these are expected given most homeless families originate from the Bronx and Brooklyn and those with school-age children are prioritized for in-borough placement. Boroughs of origin show no such patterns.

[70]Table A.25 provides a formal complier characterization exercise, finding compliers disproportionately have Bronx and Brooklyn origin, Tier II placements, fewer members, and younger heads.

stay 13–17 percent longer when assigned in-borough. They are about 5 percent more likely to exit shelter with a rental subsidy, and have 2 percent greater propensities to receive Cash Assistance, both during and after shelter. They also work more, with 10–15 percent higher earnings during the year of shelter entry when placed locally, though labor market effects attenuate post-shelter.

These are meaningful impacts. Yet they pale in comparison to effects among *marginally-treated* families—those who, due to such factors as geography, composition, or children's ages, tend to secure in-borough placements only when conditions are favorable. Both policy (IV) and school-starting (RD) compliers stay on the order of four times longer when placed in-borough. Both are overwhelmingly—by roughly 40 pp—more likely to receive Cash Assistance, with policy compliers having more pronounced effects during shelter and school-starting compliers exhibiting greater returns following it. Similarly large, but divergent, labor market impacts arise post-shelter, with policy compliers seeing 30-pp boosts in employment and school-starting compliers experiencing equally pronounced declines.

These results complement those in Cassidy (2020), where I find that homeless students placed in shelters in their school boroughs have markedly better attendance, performance, and stability. As with their families as a whole, students with especially challenging placement limitations exhibit greater policy responsiveness.

The challenge for policymakers is partly philosophical. The current policy objective is to place all families locally, to the extent constraints allow. But other objectives are possible. For example, if the goal is to minimize shelter use, then policy designs that make program participation less pleasant (such as distant placements) are likely to be effective. On the other hand, if the aim is to maximize the well-being of participants while they are participating, then loosening resource constraints through benefit enhancements (such as local placements) is preferable. Of course, long-term consequences matter, too. While this study is unable to assess such outcomes, the findings of generally smaller differences between treated and untreated families post-shelter, combined with the empirical regularity that most homeless families do not become long-term homeless, suggest modest increases in benefit generosity are unlike to be harmful.

Given finite resources, some families will inevitably be served suboptimally. In this context, my results suggest distinct priorities for differentially-situated groups is desirable. If locally-placed families are more apt to work, but less likely to seek housing, they should be targeted for supplemental housing search assistance. Correspondingly, distantly-placed families may have greater difficultly forging labor market ties; they should be prioritized for job training services and transit subsidies. In general, supplementary services should complement families' comparative advantages in manners compatible with their incentives.

If all homeless families were the same, there would not be much more to the story. But the theme of heterogeneity underscores a more primitive point: the potential gains from better targeting local placements. The most immediate question is not whether \$10,000 is the right price to pay for, on average, 10 percent gains in earnings and school attendance, but instead how those costly shelter slots can be more efficiently allocated to the families poised to benefit the most. I find that difficult-to-place families are particularly sensitive to their shelter assignments; this "resistance" to treatment is partly predictable from administrative observables, including families' aptitudes for navigating the application process. Screening practices should be augmented to better identify high responders. Counterintuitively, the families perceived to be the most challenging to place proximately should have their slots prospectively reserved. Services better tailored to family needs should generate surpluses that can be used to compensate families given less desirable assignments.

At the core of my study is a natural experiment. Shelter assignment location is essentially random. It should not be.

8 References

Allgood, Sam, and Ronald S Warren. 2003. "The Duration of Homelessness: Evidence from a National Survey." *Journal of Housing Economics*, 12(4): 273–290.

Allgood, Sam, Myra L Moore, and Ronald S Warren. 1997. "The Duration of Sheltered Homelessness in a Small City." *Journal of Housing Economics*, 6(1): 60–80.

Andersson, Fredrik, John C Haltiwanger, Mark J Kutzbach, Giordano E Palloni, Henry O Pollakowski, and Daniel H Weinberg. 2016. "Childhood Housing and Adult Earnings: A Between-siblings Analysis of Housing Vouchers and Public Housing." National Bureau of Economic Research.

Angrist, Joshua D, and Jorn-Steffen Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton university press.

Angrist, Joshua D, Guido W Imbens, and Donald B Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association*, 91(434): 444–455.

Besley, Timothy, and Stephen Coate. 1992. "Workfare Versus Welfare: Incentive Arguments for Work Requirements in Poverty-alleviation Programs." *The American Economic Review*, 82(1): 249–261.

Besley, Timothy, and Stephen Coate. 1995. "The Design of Income Maintenance Programmes." *The Review of Economic Studies*, 62(2): 187–221.

Buckner, John C. 2008. "Understanding the Impact of Homelessness on Children: Challenges and Future Research Directions." *American Behavioral Scientist*, 51(6): 721–736.

Cahuc, Pierre, Stéphane Carcillo, and André Zylberberg. 2014. *Labor Economics.* MIT Press.

Cassidy, Michael. 2020. "A Closer Look: Proximity Boosts Homeless Student Performance in New York City." *IZA Discussion Papers.*

Cattaneo, Matias D, Nicolás Idrobo, and Rocıo Titiunik. 2017. "A Practical Introduction to Regression Discontinuity Designs: Part I." *Cambridge Elements: Quantitative and Computational Methods for Social Science.*

Cattaneo, Matias D, Nicolás Idrobo, and Rocıo Titiunik. 2018. "A Practical Introduction to Regression Discontinuity Designs: Part II." *Cambridge Elements: Quantitative and Computational Methods for Social Science.*

Chetty, Raj. 2008. "Moral Hazard Versus Liquidity and Optimal Unemployment Insurance." *Journal of Political Economy*, 116(2): 173–234.

Chetty, Raj, and Amy Finkelstein. 2013. "Social Insurance: Connecting Theory to Data." *Handbook of Public Economics*, 5: 111–193.

Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz. 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *The American Economic Review*, 106(4): 855–902.

Coalition for the Homeless. 2019. "New York City Homeless Municipal Shelter Population, 1983-Present."

Cobb-Clark, Deborah A, and Anna Zhu. 2017. "Childhood Homelessness and Adult Employment: The Role of Education, Incarceration, and Welfare Receipt." *Journal of Population Economics*, 30(3): 893–924.

Cobb-Clark, Deborah A, Nicolas Herault, Rosanna Scutella, and Yi-Ping Tseng. 2016. "A Journey Home: What Drives How Long People Are Homeless?" *Journal of Urban Economics*, 91: 57–72.

Cohen, Elissa, Sarah Minton Megan Thompson Elizabeth Crowe, and Linda Giannarelli. 2016. "Welfare Rules Databook: State TANF Policies as of July 2015, OPRE Report 2016-67."

Collinson, Robert, and Davin Reed. 2018. "The Effects of Evictions on Low-income Households." Working paper, New York University, Wagner School.

Corinth, Kevin. 2017. "The Impact of Permanent Supportive Housing on Homeless Populations." *Journal of Housing Economics*, 35: 69–84.

Cragg, Michael, and Brendan O'Flaherty. 1999. "Do Homeless Shelter Conditions Determine Shelter Population? the Case of the Dinkins Deluge." *Journal of Urban Economics*, 46(3): 377–415.

Culhane, Dennis P., Stephen Metraux, Jung Min Park, Maryanne Schretzman, and Jesse Valente. 2007. "Testing a Typology of Family Homelessness Based on Patterns of Public Shelter Utilization in Four U.s. Jurisdictions: Implications for Policy and Program Planning." *Housing Policy Debate*, 18(1): 1–28.

Curtis, Marah A, Hope Corman, Kelly Noonan, and Nancy E Reichman. 2013. "Life Shocks and Homelessness." *Demography*, 50(6): 2227–2253.

Dahl, Gordon B, Andreas Ravndal Kostøl, and Magne Mogstad. 2014. "Family Welfare Cultures." *The Quarterly Journal of Economics*, 129(4): 1711–1752.

Dobbie, Will, Jacob Goldin, and Crystal S Yang. 2018. "The Effects of Pretrial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges." *American Economic Review*, 108(2): 201–40.

Dong, Yingying. 2015. "Regression Discontinuity Applications with Rounding Errors in the Running Variable." *Journal of Applied Econometrics*, 30(3): 422–446.

Eckstein, Zvi, and Gerard J Van den Berg. 2007. "Empirical Labor Search: A Survey." *Journal of Econometrics*, 136(2): 531–564.

Edwards, Aaron. 2012. "New York Acts Quickly Amid Sharp Rise in Homelessness." *New York Times*.

Eide, Stephen. 2018. "Benchmarking Homeless Shelter Performance: A Proposal for Easing America's Homeless Crisis." Manhattan Institute.

Ellen, Ingrid Gould, and Brendan O'Flaherty. 2010. *How to House the Homeless.* Russell Sage Foundation.

Evans, William N, David C Philips, and Krista J Ruffini. 2019. "Reducing and Preventing Homelessness: A Review of the Evidence and Charting a Research Agenda." National Bureau of Economic Research.

Evans, William N, James X Sullivan, and Melanie Wallskog. 2016. "The Impact of Homelessness Prevention Programs on Homelessness." *Science*, 353(6300): 694–699.

Fermino, Jennifer. 2016*a*. "Loose shelter requirements led city to house homeless families that may have had elsewhere to stay." *New York Daily News*.

Fermino, Jennifer. 2016*b*. "Shelter jam tied to regs." *New York Daily News*.

Frandsen, Brigham R. 2017. "Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design When the Running Variable Is Discrete." In *Regression Discontinuity Designs: Theory and Applications*. 281–315. Emerald Publishing Limited.

Galiani, Sebastian, Alvin Murphy, and Juan Pantano. 2015. "Estimating Neighborhood Choice Models: Lessons from a Housing Assistance Experiment." *The American Economic Review*, 105(11): 3385–3415.

Glomm, Gerhard, and Andrew John. 2002. "Homelessness and Labor Markets." *Regional Science and Urban Economics*, 32(5): 591–606.

Goodman, Sarena, Peter Messeri, and Brendan O'Flaherty. 2014. "How Effective Homelessness Prevention Impacts the Length of Shelter Spells." *Journal of Housing Economics*, 23: 55–62.

Goodman, Sarena, Peter Messeri, and Brendan O'Flaherty. 2016. "Homelessness Prevention in New York City: On Average, It Works." *Journal of Housing Economics*, 31: 14–34.

Gould, Thomas E, and Arthur R Williams. 2010. "Family Homelessness: An Investigation of Structural Effects." *Journal of Human Behavior in the Social Environment*, 20(2): 170–192.

Gubits, Daniel, Marybeth Shinn, Michelle Wood, Stephen Bell, Samuel Dastrup, Claudia Solari, Scott Brown, Debi McInnis, Tom McCall, and Utsav Kattel. 2016. "Family Options Study: 3-Year Impacts of Housing and Services Interventions for Homeless Families." *Available at Ssrn 3055295.*

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw. 2001. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica*, 69(1): 201–209.

Harris, Elizabeth A. 2016. "Under New Policy for Homeless Families, Children Can Miss Less School." *New York Times.*

Heckman, James J. 2010. "Building Bridges between Structural and Program Evaluation Approaches to Evaluating Policy." *Journal of Economic Literature*, 48(2): 356–98.

Heckman, James J, and Edward J Vytlacil. 1999. "Local Instrumental Variables and Latent Variable Models for Identifying and Bounding Treatment Effects." *Proceedings of the National Academy of Sciences*, 96(8): 4730–4734.

Heckman, James J, and Edward J Vytlacil. 2007. "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast Their Effects in New Environments." *Handbook of Econometrics*, 6: 4875–5143.

Heckman, James J, and Edward Vytlacil. 2001. "Policy-Relevant Treatment Effects." *American Economic Review*, 91(2): 107–111.

Heckman, James J, and Edward Vytlacil. 2005. "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica*, 73(3): 669–738.

Heckman, James J, Sergio Urzua, and Edward Vytlacil. 2006. "Understanding Instrumental Variables in Models with Essential Heterogeneity." *The Review of Economics and Statistics*, 88(3): 389–432.

Imbens, Guido W, and Joshua D Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica*, 62(2): 467–475.

Imbens, Guido W, and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics*, 142(2): 615–635.

Jacob, Brian A. 2004. "Public Housing, Housing Vouchers, and Student Achievement: Evidence from Public Housing Demolitions in Chicago." *The American Economic Review*, 94(1): 233–258.

Jacob, Brian A, and Jens Ludwig. 2012. "The Effects of Housing Assistance on Labor Supply: Evidence from a Voucher Lottery." *The American Economic Review*, 102(1): 272–304.

Jacob, Brian A, Max Kapustin, and Jens Ludwig. 2015. "The Impact of Housing Assistance on Child Outcomes: Evidence from a Randomized Housing Lottery." *The Quarterly Journal of Economics*, 130(1): 465–506.

Jorgensen, Jillian. 2017. "More homeless families denied at shelters after NYC toughens eligibility." *New York Daily News*.

Katz, Alyssa. 2015. "BILL MADE HIS BED: How de Blasio's decisions on homelessness contributed to the crisis he now decries." *New York Daily News*.

Katz, Lawrence F, and Bruce D Meyer. 1990. "The Impact of the Potential Duration of Unemployment Benefits on the Duration of Unemployment." *Journal of Public Economics*, 41(1): 45–72.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz. 2007. “Experimental Analysis of Neighborhood Effects.” *Econometrica*, 75(1): 83–119.

Kolesár, Michal, and Christoph Rothe. 2018. “Inference in Regression Discontinuity Designs with a Discrete Running Variable.” *American Economic Review*, 108(8): 2277–2304.

Lalive, Rafael, Jan Van Ours, and Josef Zweimüller. 2006. “How Changes in Financial Incentives Affect the Duration of Unemployment.” *The Review of Economic Studies*, 73(4): 1009–1038.

Lee, David S, and David Card. 2008. “Regression Discontinuity Inference with Specification Error.” *Journal of Econometrics*, 142(2): 655–674.

Lee, David S., and Thomas Lemieux. 2010. “Regression Discontinuity Designs in Economics.” *Journal of Economic Literature*, 48(2): 281–355.

Ludwig, Jens, Greg J Duncan, Lisa A Gennetian, Lawrence F Katz, Ronald C Kessler, Jeffrey R Kling, and Lisa Sanbonmatsu. 2012. “Neighborhood Effects on the Long-term Well-being of Low-income Adults.” *Science*, 337(6101): 1505–1510.

Ludwig, Jens, Greg J Duncan, Lisa A Gennetian, Lawrence F Katz, Ronald C Kessler, Jeffrey R Kling, and Lisa Sanbonmatsu. 2013. “Long-term Neighborhood Effects on Low-income Families: Evidence from Moving to Opportunity.” *The American Economic Review*, 103(3): 226–231.

Ludwig, Jens, Jeffrey B Liebman, Jeffrey R Kling, Greg J Duncan, Lawrence F Katz, Ronald C Kessler, and Lisa Sanbonmatsu. 2008. “What Can We Learn about Neighborhood Effects from the Moving to Opportunity Experiment?” *American Journal of Sociology*, 114(1): 144–188.

McCall, John Joseph. 1970. “Economics of Information and Job Search.” *The Quarterly Journal of Economics*, 113–126.

Miller, Peter M. 2011. “A Critical Analysis of the Research on Student Homelessness.” *Review of Educational Research*, 81(3): 308–337.

Mortensen, Dale T, and Christopher A Pissarides. 1999. “New Developments in Models of Search in the Labor Market.” *Handbook of Labor Economics*, 3: 2567–2627.

New York City Mayor’s Office of Operations. 2019. “Fiscal 2019 Mayor’s Management Report.”

New York Daily News Editorial. 2014. "Bill homes in." *New York Daily News*.

New York State Office of Temporary and Disability Assistance. 2015*a*. "Administrative Directive 15-ADM-06."

New York State Office of Temporary and Disability Assistance. 2015*b*. "New York State Plan and Executive Certification: Administration of the Block Grant for Temporary Assistance for Needy Families, 2015–2017."

New York State Office of Temporary and Disability Assistance. 2016*a*. "Administrative Directive 16-ADM-11."

New York State Office of Temporary and Disability Assistance. 2016*b*. "Temporary Assistance Source Book."

New York State Office of Temporary and Disability Assistance. 2017. "Temporary Assistance."

New York State Office of Temporary and Disability Assistance. 2019. "Supplemental Nutrition Assistance Program (SNAP)."

Nichols, Albert L, and Richard J Zeckhauser. 1982. "Targeting Transfers through Restrictions on Recipients." *The American Economic Review*, 72(2): 372–377.

NYC Department of Homeless Services. 2019*a*. "About DHS."

NYC Department of Homeless Services. 2019*b*. "Daily Report, October 21, 2019."

NYC Department of Homeless Services. 2019*c*. "Families with Children."

NYC Department of Homeless Services. 2019*d*. "Stats and Reports."

NYC Human Resources Administration. 2019. "HRA Monthly Fact Sheet, August 2019."

NYC Independent Budget Office. 2014. "The Rising Number of Homeless Families in NYC, 2002–2012: A Look at Why Families Were Granted Shelter, the Housing They Had Lived in and Where They Came From."

NYC Mayor's Office. 2014. "Ready to Launch: New York City's Implementation Plan for Free, High-Quality, Full-Day Universal Pre-Kindergarten."

NYC Mayor's Office. 2017. "Turning the Tide on Homeless in New York City."

NYC Mayor's Office of Operations. 2012. "Mayor's Management Report, September 2012."

NYC Mayor's Office of Operations. 2017. "Mayor's Management Report, September 2017."

NYC Mayor's Office of Operations. 2018. "Mayor's Management Report, September 2018."

NYC Office of Management and Budget. 2019. "February 2019 Financial Plan: Budget Function Analysis."

NYU Furman Center. 2016. "State of New York City's Housing and Neighborhoods in 2016."

O'Flaherty, Brendan. 1995. "An Economic Theory of Homelessness and Housing." *Journal of Housing Economics*, 4(1): 13–49.

O'Flaherty, Brendan. 2004. "Wrong Person and Wrong Place: For Homelessness, the Conjunction Is What Matters." *Journal of Housing Economics*, 13(1): 1–15.

O'Flaherty, Brendan. 2009. "When Should Homeless Families Get Subsidized Apartments? A Theoretical Inquiry." *Journal of Housing Economics*, 18(2): 69–80.

O'Flaherty, Brendan. 2010. "Homelessness As Bad Luck: Implications for Research and Policy." *How to House the Homeless.new York: Russell Sage Foundation*, 143–182.

O'Flaherty, Brendan. 2019. "Homelessness research: A guide for economists (and friends)." *Journal of Housing Economics.*

O'Flaherty, Brendan, and Ting Wu. 2006. "Fewer Subsidized Exits and a Recession: How New York City's Family Homeless Shelter Population Became Immense." *Journal of Housing Economics*, 15(2): 99–125.

O'Flaherty, Brendan, Rosanna Scutella, and Yi-Ping Tseng. 2018*a*. "Private Information, Exits from Homelessness, and Better Ways to Operate Rehousing Programs." *Journal of Housing Economics*, 41: 93–105.

O'Flaherty, Brendan, Rosanna Scutella, and Yi-Ping Tseng. 2018*b*. "Using Private Information to Predict Homelessness Entries: Evidence and Prospects." *Housing Policy Debate*, 28(3): 368–392.

Oreopoulos, Philip. 2003. "The Long-run Consequences of Living in a Poor Neighborhood." *The Quarterly Journal of Economics*, 118(4): 1533–1575.

Pissarides, Christopher A. 2000. *Equilibrium Unemployment Theory.* MIT Press.

Routhier, Giselle. 2017*a*. "Family Homelessness in NYC: City and State Must Meet Unprecedented Scale of Crisis with Proven Solutions." Coalition for the Homeless.

Routhier, Giselle. 2017*b*. "State of the Homeless 2017." Coalition for the Homeless.

Samuels, Judith, Marybeth Shinn, and John C Buckner. 2010. "Homeless Children: Update on Research, Policy, Programs, and Opportunities." *Washington, Dc: Office of the Assistant Secretary for Planning and Evaluation, Us Department of Health and Human Services.*

Sanbonmatsu, Lisa, Jens Ludwig, Lawrence F Katz, Lisa A Gennetian, Greg J Duncan, Ronald C Kessler, Emma Adam, Thomas W McDade, and Stacy Tessler Lindau. 2011. "Moving to Opportunity for Fair Housing Demonstration Program–Final Impacts Evaluation."

Secret, Mosi. 2011. "Clock Ticks for a Key Homeless Program." *New York Times.*

Shinn, M., B. C. Weitzman, D. Stojanovic, J. R. Knickman, L. Jimenez, L. Duchon, S. James, and D. H. Krantz. 1998. "Predictors of homelessness among families in New York City: from shelter request to housing stability." *American Journal of Public Health*, 88(11): 1651–1657.

Spinnewijn, Johannes. 2013. "Training and Search during Unemployment." *Journal of Public Economics*, 99: 49–65.

Stewart, Nikita. 2017. "Fight Looms as Bill de Blasio Plans to Seek 90 New Homeless Shelters." *New York Times.*

Topa, Giorgio, Yves Zenou, et al. 2015. "Neighborhood and Network Effects." *Handbook of Regional and Urban Economics*, 5: 561–624.

University of Michigan Law School. 2017. "Case Profile: McCain v. Koch."

U.S. Census Bureau. 2018. "QuickFacts."

U.S. Department of Housing and Urban Development. 2018. "Part 1: Point-in-Time Estimates of Homelessness." *The 2018 Annual Homeless Assessment Report (AHAR) to Congress.*

Vytlacil, Edward. 2002. "Independence, Monotonicity, and Latent Index Models: An Equivalence Result." *Econometrica*, 70(1): 331–341.

Wasi, Nada, Aaron Flaaen, et al. 2015. "Record Linkage Using Stata: Pre-processing, Linking and Reviewing Utilities." *Stata Journal*, 15(3): 672–697.

Wood, Michelle, Jennifer Turnham, and Gregory Mills. 2008. "Housing affordability and family well-being: Results from the housing voucher evaluation." *Housing Policy Debate*, 19(2): 367–412.

9 Tables

Table 1: Data and Sample Overview: Eligible NYC DHS Family Shelter Entrants, 2010–2016

Family Spells	Count	Percent
All	68,584	1.00
NYC Entrants	61,406	0.90
Full Sample with Borough Treatment Status	61,120	0.89
Full Sample with Treatment and Running Variable[a]	59,253	0.86
Non-DV Sample	43,235	0.63
Pre-2015 Sample	41,717	0.61
One School-Age Child Sample	40,779	0.59

Unit of observation is family shelter spell. Data from NYC administrative records, as described in text. Indentation indicates cumulative refinement.

[a] Running variable is oldest child's potential grade level for children under 18 years of age. Families whose oldest children are 19–21 years are excluded.

Table 2: Descriptives and Random Assignment

Variable	Overall Mean	Overall SD	Randomization Check Out-of-Boro	Randomization Check In-Boro	Randomization Check Diff.
Year Entered Shelter	2013.01	2.07	2013.38	2012.65	-0.72**
Month Entered Shelter	6.52	3.40	6.78	6.28	-0.50**
Manhattan Origin	0.12	0.33	0.16	0.09	-0.07**
Bronx Origin	0.41	0.49	0.33	0.49	0.16**
Brooklyn Origin	0.32	0.47	0.31	0.32	0.01**
Queens Origin	0.12	0.33	0.15	0.10	-0.06**
Staten Island Origin	0.03	0.16	0.05	0.01	-0.04**
Family Size	3.35	1.39	3.34	3.36	0.02*
Family Members Under 18	1.97	1.19	1.95	1.99	0.04**
Oldest Child's Grade	2.57	5.32	1.95	3.18	1.23**
Health Issue Present	0.30	0.46	0.32	0.28	-0.04**
Eligibility: Eviction	0.33	0.47	0.28	0.39	0.10**
Eligibility: Overcrowding	0.18	0.38	0.17	0.19	0.02**
Eligibility: Conditions	0.08	0.28	0.08	0.09	0.01**
Eligibility: Domestic Violence	0.30	0.46	0.37	0.22	-0.15**
Eligibility: Other	0.11	0.31	0.10	0.11	0.01**
Female	0.92	0.28	0.92	0.91	-0.01**
Age	31.54	8.86	30.94	32.13	1.20**
Partner/Spouse Present	0.26	0.44	0.27	0.24	-0.03**
Pregnant	0.07	0.25	0.07	0.06	-0.01**
Black	0.56	0.50	0.57	0.55	-0.02**
White	0.03	0.16	0.03	0.02	-0.01**
Hispanic	0.38	0.48	0.36	0.39	0.03**
No Degree	0.57	0.50	0.56	0.58	0.01**
High School Grad	0.32	0.47	0.32	0.32	-0.01*
Some College or More	0.05	0.22	0.05	0.05	-0.00
Unknown Education	0.06	0.24	0.06	0.06	-0.00
On Cash Assistance	0.35	0.48	0.36	0.35	-0.01**
On Food Stamps	0.73	0.44	0.73	0.73	0.00
Employed Year Pre	0.43	0.50	0.44	0.43	-0.01**
Log AQ Earnings Year Pre	3.01	3.58	3.02	2.99	-0.03
Tier II Shelter	0.55	0.50	0.55	0.55	0.01**
Commercial Hotel	0.28	0.45	0.30	0.25	-0.05**
Family Cluster Unit	0.16	0.37	0.14	0.19	0.05**
Manhattan Shelter	0.18	0.39	0.27	0.09	-0.18**
Bronx Shelter	0.39	0.49	0.29	0.49	0.20**
Brooklyn Shelter	0.27	0.44	0.22	0.32	0.11**
Queens Shelter	0.15	0.36	0.21	0.10	-0.11**
Staten Island Shelter	0.01	0.09	0.01	0.01	-0.01**
School District Placement	0.10	0.30	0.00	0.19	0.19**
Placement Distance (miles)	5.89	4.65	9.27	2.66	-6.61**
Borough Placement	0.51	0.50	0.00	1.00	1.00

Treatment defined as placed in-borough. Group contrasts obtained from separate bivariate OLS regressions of each characteristic on treatment indicator. Differences between in-borough and out-of-borough means are coefficients on treatment indicator. Standard errors clustered at the family group level. Unit of observation is family-spell. Full Sample: 59,253 observations. See Appendix for additional covariates. * $p < 0.10$, ** $p < 0.05$

Table 3A: OLS Main Results

Outcome	Outcome Mean (1)	Base (2)	Placement (3)	Main (4)	Shelter (5)	Main (6)	Main (7)
	Full Sample					Non-DV	Pre-2015
A. Stays and Returns							
Log Length of Stay	5.501	0.139**	0.107**	0.120**	0.115**	0.085**	0.125**
	(1.241)	(0.010)	(0.010)	(0.011)	(0.011)	(0.012)	(0.013)
	{59,253}	{59,253}	{59,253}	{59,253}	{59,247}	{41,744}	{41,717}
Subsidized Exit	0.392	0.007*	0.021**	0.018**	0.017**	0.017**	0.016**
	(0.488)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
	{57,962}	{57,962}	{57,962}	{57,962}	{57,954}	{40,766}	{41,420}
Returned to Shelter	0.151	-0.025**	-0.005	-0.005	-0.004	-0.000	-0.007*
	(0.358)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
	{52,274}	{52,274}	{52,274}	{52,274}	{52,271}	{36,768}	{40,552}
Placement Controls		No	Yes	Yes	Yes	Yes	Yes
Family & Shelter Controls		No	No	Yes	Yes	Yes	Yes
Shelter Fixed Effects		No	No	No	Yes	No	No

Each cell reports the coefficient on in-borough shelter placement from a separate OLS regression of the row-delineated outcome on the treatment indicator, controlling for the column-enumerated covariates. Placement covariates are dummies for shelter entry month, borough of origin, health issue, and eligibility reason, as well as a cubic polynomial in year of shelter entry and linear controls for family size, number of family members under 18, and oldest child's grade. Main covariates are placement covariates plus family and shelter covariates. Family covariates are dummies for head gender, race, partner presence, education category, Cash Assistance receipt, and Food Stamps receipt, as well continuous controls for head age and log average quarterly earnings. Shelter covariates are dummies for shelter type and shelter borough. All covariates are defined at shelter entry or as near as possible. Supercolumns give samples. Standard errors clustered at family group level in parentheses. Number of observations given in braces below first outcome in each panel, as well as for any subsequent outcome where the sample size differs from the first due to censoring. * $p < 0.10$, ** $p < 0.05$

Table 3B: OLS Main Results

Outcome	Outcome Mean (1)	Base (2)	Placement (3)	Main (4)	Shelter (5)	Non-DV Main (6)	Pre-2015 Main (7)
	Full Sample					Non-DV	Pre-2015
B1. Year Post-Entry Outcomes							
Cash Assistance	0.782	0.019**	0.015**	0.011**	0.010**	0.011**	0.013**
	(0.413)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)
	{59,253}	{59,253}	{59,253}	{59,253}	{59,247}	{41,744}	{41,717}
Food Stamps	0.896	0.010**	0.006**	0.003	0.003	-0.000	0.003
	(0.306)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Employed	0.479	0.006	0.012**	0.010**	0.010**	0.009*	0.010**
	(0.500)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Log Avg. Quarterly Earnings	3.377	0.088**	0.108**	0.094**	0.086**	0.087**	0.085**
	(3.679)	(0.031)	(0.032)	(0.028)	(0.028)	(0.033)	(0.033)
B2. Year Post-Exit Outcomes							
Cash Assistance	0.738	0.014**	0.019**	0.017**	0.016**	0.021**	0.016**
	(0.440)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.004)
	{48,082}	{48,082}	{48,082}	{48,082}	{48,076}	{33,761}	{39,974}
Food Stamps	0.884	0.010**	0.011**	0.008**	0.008**	0.003	0.008**
	(0.321)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Employed	0.455	0.005	0.009*	0.003	0.002	0.005	0.003
	(0.498)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)
Log Avg. Quarterly Earnings	3.268	0.094**	0.084**	0.043	0.036	0.050	0.041
	(3.732)	(0.034)	(0.036)	(0.033)	(0.033)	(0.040)	(0.036)
Placement Controls		No	Yes	Yes	Yes	Yes	Yes
Family & Shelter Controls		No	No	Yes	Yes	Yes	Yes
Shelter Fixed Effects		No	No	No	Yes	No	No

Each cell reports the coefficient on in-borough shelter placement from a separate OLS regression of the row-delineated outcome on the treatment indicator, controlling for the column-enumerated covariates. Placement covariates are dummies for shelter entry month, borough of origin, health issue, and eligibility reason, as well as a cubic polynomial in year of shelter entry and linear controls for family size, number of family members under 18, and oldest child's grade. Main covariates are placement covariates plus family and shelter covariates. Family covariates are dummies for head gender, race, partner presence, education category, Cash Assistance receipt, and Food Stamps receipt, as well continuous controls for head age and log average quarterly earnings. Shelter covariates are dummies for shelter type and shelter borough. All covariates are defined at shelter entry or as near as possible. Supercolumns give samples. Standard errors clustered at family group level in parentheses. Number of observations given in braces below first outcome in each panel, as well as for any subsequent outcome where the sample size differs from the first due to censoring. * $p < 0.10$, ** $p < 0.05$

Table 4A: OLS Robustness

	School District			Distance		
	Full	Non-DV	Pre-2015	Full	Non-DV	Pre-2015
	(1)	(2)	(3)	(4)	(5)	(6)
A. Stays and Returns						
Log Length of Stay	0.0814**	0.0559**	0.0757**	-0.0143**	-0.0108**	-0.0141**
	(0.0162)	(0.0169)	(0.0193)	(0.0012)	(0.0013)	(0.0016)
	{54,306}	{38,587}	{38,053}	{54,306}	{38,587}	{38,053}
Subsidized Exit	0.0009	-0.0036	0.0011	-0.0026**	-0.0023**	-0.0019**
	(0.0068)	(0.0077)	(0.0078)	(0.0005)	(0.0006)	(0.0006)
	{53,121}	{37,687}	{37,789}	{53,121}	{37,687}	{37,789}
Returned to Shelter	-0.0066	-0.0037	-0.0105*	0.0004	0.0002	0.0005
	(0.0054)	(0.0059)	(0.0059)	(0.0004)	(0.0005)	(0.0005)
	{47,858}	{33,963}	{36,991}	{47,858}	{33,963}	{36,991}
Placement Controls	Yes	Yes	Yes	Yes	Yes	Yes
Family & Shelter Controls	Yes	Yes	Yes	Yes	Yes	Yes
Shelter FE	No	No	No	No	No	No

Each cell reports the coefficient on local shelter placement from a separate OLS regression of the row-delineated outcome on the treatment, controlling for Main covariates, described in Table 3A. Columns give samples; super-columns give treatment definitions. Standard errors clustered at family group level in parentheses. Number of observations given in braces below first outcome in each panel, as well as for any subsequent outcome where the sample size differs due to censoring. * $p < 0.10$, ** $p < 0.05$

Table 4B: OLS Robustness

	School District			Distance		
	Full	Non-DV	Pre-2015	Full	Non-DV	Pre-2015
	(1)	(2)	(3)	(4)	(5)	(6)
B1. Year Post-Entry Outcomes						
Cash Assistance	0.0013	-0.0031	0.0014	-0.0015**	-0.0015**	-0.0014**
	(0.0051)	(0.0057)	(0.0060)	(0.0004)	(0.0004)	(0.0004)
	{54,306}	{38,587}	{38,053}	{54,306}	{38,587}	{38,053}
Food Stamps	0.0042	-0.0012	0.0046	-0.0005**	-0.0001	-0.0003
	(0.0032)	(0.0036)	(0.0036)	(0.0002)	(0.0003)	(0.0003)
Employed	0.0181**	0.0129*	0.0146*	-0.0020**	-0.0018**	-0.0015**
	(0.0063)	(0.0070)	(0.0075)	(0.0004)	(0.0005)	(0.0005)
Log Avg. Quarterly Earnings	0.1262**	0.0945*	0.0895*	-0.0164**	-0.0150**	-0.0117**
	(0.0451)	(0.0503)	(0.0535)	(0.0031)	(0.0037)	(0.0038)
B2. Year Post-Exit Outcomes						
Cash Assistance	-0.0047	-0.0053	-0.0066	-0.0016**	-0.0019**	-0.0016**
	(0.0065)	(0.0074)	(0.0072)	(0.0005)	(0.0006)	(0.0005)
	{43,981}	{31,172}	{36,453}	{43,981}	{31,172}	{36,453}
Food Stamps	0.0018	-0.0048	0.0005	-0.0002	0.0003	-0.0001
	(0.0043)	(0.0049)	(0.0046)	(0.0003)	(0.0004)	(0.0003)
Employed	-0.0081	-0.0092	-0.0102	-0.0006	-0.0005	-0.0007
	(0.0073)	(0.0081)	(0.0080)	(0.0005)	(0.0006)	(0.0006)
Log Avg. Quarterly Earnings	-0.0441	-0.0556	-0.0617	-0.0060	-0.0040	-0.0055
	(0.0536)	(0.0601)	(0.0586)	(0.0037)	(0.0045)	(0.0042)
Placement Controls	Yes	Yes	Yes	Yes	Yes	Yes
Family & Shelter Controls	Yes	Yes	Yes	Yes	Yes	Yes
Shelter FE	No	No	No	No	No	No

Each cell reports the coefficient on local shelter placement from a separate OLS regression of the row-delineated outcome on the treatment, controlling for Main covariates, described in Table 3A. Columns give samples; supercolumns give treatment definitions. Standard errors clustered at family group level in parentheses. Number of observations given in braces below first outcome in each panel, as well as for any subsequent outcome where the sample size differs due to censoring. * $p < 0.10$, ** $p < 0.05$

Table 5A: IV Main Results

	Ineligibility Rate			Aversion Ratio		
Outcome	Placement (1)	Main (2)	Shelter (3)	Placement (4)	Main (5)	Shelter (6)
A. Stays and Returns						
Log Length of Stay	1.121**	1.367**	1.151**	0.781**	0.946**	0.765**
	(0.403)	(0.527)	(0.471)	(0.282)	(0.342)	(0.331)
	[42.9]	[28.8]	[33.3]	[80.7]	[60.8]	[61.0]
Subsidized Exit	-0.581**	-0.789**	-0.664**	-0.244**	-0.331**	-0.291**
	(0.186)	(0.257)	(0.224)	(0.120)	(0.147)	(0.145)
	[39.7]	[26.2]	[30.6]	[75.4]	[55.8]	[55.8]
Returned to Shelter	0.219*	0.287*	0.272*	0.058	0.088	0.093
	(0.130)	(0.166)	(0.156)	(0.088)	(0.104)	(0.106)
	[36.0]	[25.2]	[28.0]	[71.5]	[55.7]	[54.2]
First Stage Instrument Coefficient	0.387**	0.303**	0.330**	0.073**	0.061**	0.062**
	(0.059)	(0.056)	(0.057)	(0.008)	(0.008)	(0.008)
Placement Controls	Yes	Yes	Yes	Yes	Yes	Yes
Family & Shelter Controls	No	Yes	Yes	No	Yes	Yes
Shelter FE	No	No	Yes	No	No	Yes

Each cell reports the coefficient on in-borough shelter placement from a separate 2SLS regression of the row-delineated outcome on the treatment indicator, controlling for the column-enumerated covariates, described in Table 3A. Instruments are indicated by supercolumns. Standard errors clustered at family group level in parentheses. First-stage F-stats given in brackets below first outcome in each panel, as well as for any subsequent outcome where the sample size differs due to censoring. All results are for Full sample; number of observations given in Tables 3A and 3B. * $p < 0.10$, ** $p < 0.05$

Table 5B: IV Main Results

	Ineligibility Rate			Aversion Ratio		
	Placement	Main	Shelter	Placement	Main	Shelter
Outcome	(1)	(2)	(3)	(4)	(5)	(6)
B1. Year Post-Entry Outcomes						
Cash Assistance	0.529**	0.651**	0.579**	0.292**	0.338**	0.317**
	(0.152)	(0.183)	(0.162)	(0.101)	(0.105)	(0.104)
	[42.9]	[28.8]	[33.3]	[80.7]	[60.8]	[61.0]
Food Stamps	-0.137	-0.142	-0.095	-0.088	-0.100	-0.069
	(0.100)	(0.093)	(0.085)	(0.073)	(0.064)	(0.063)
Employed	-0.101	-0.020	-0.022	0.066	0.116	0.102
	(0.157)	(0.171)	(0.159)	(0.115)	(0.118)	(0.118)
Log Avg. Quarterly Earnings	0.264	1.245	1.035	0.650	1.085	0.903
	(1.152)	(1.243)	(1.148)	(0.847)	(0.851)	(0.846)
B2. Year Post-Exit Outcomes						
Cash Assistance	0.394**	0.428**	0.420**	0.267**	0.265**	0.288**
	(0.189)	(0.210)	(0.195)	(0.126)	(0.129)	(0.132)
	[27.4]	[20.3]	[23.2]	[56.6]	[46.4]	[45.4]
Food Stamps	-0.023	-0.064	-0.051	0.048	0.023	0.040
	(0.130)	(0.130)	(0.120)	(0.091)	(0.086)	(0.087)
Employed	0.386*	0.397*	0.363*	0.395**	0.338**	0.330**
	(0.211)	(0.232)	(0.214)	(0.147)	(0.149)	(0.150)
Log Avg. Quarterly Earnings	2.515	2.508	2.317	2.591**	2.035*	2.003*
	(1.562)	(1.673)	(1.551)	(1.093)	(1.078)	(1.090)
First Stage Instrument Coefficient	0.387**	0.303**	0.330**	0.073**	0.061**	0.062**
	(0.059)	(0.056)	(0.057)	(0.008)	(0.008)	(0.008)
Placement Controls	Yes	Yes	Yes	Yes	Yes	Yes
Family & Shelter Controls	No	Yes	Yes	No	Yes	Yes
Shelter FE	No	No	Yes	No	No	Yes

Each cell reports the coefficient on in-borough shelter placement from a separate 2SLS regression of the row-delineated outcome on the treatment indicator, controlling for the column-enumerated covariates, described in Table 3A. Instruments are indicated by supercolumns. Standard errors clustered at family group level in parentheses. First-stage F-stats given in brackets below first outcome in each panel, as well as for any subsequent outcome where the sample size differs due to censoring. All results are for Full sample; number of observations given in Tables 3A and 3B. * $p < 0.10$, ** $p < 0.05$

Table 6A: Complier Characteristics: Ineligibility Rate Instrument

	Compliers	Non-Compliers	Diff.
Manhattan Origin	0.00	0.14	-0.13
	(0.003)	(0.000)	[-2.56]
Bronx Origin	0.57	0.39	0.18
	(0.006)	(0.000)	[2.28]
Brooklyn Origin	0.25	0.32	-0.07
	(0.005)	(0.000)	[-0.99]
Queens Origin	0.10	0.13	-0.02
	(0.003)	(0.000)	[-0.45]
Staten Island Origin	0.02	0.03	-0.00
	(0.000)	(0.000)	[-0.33]
Health Issue Present	0.33	0.30	0.04
	(0.004)	(0.000)	[0.61]
Eligibility: Eviction	0.29	0.34	-0.05
	(0.005)	(0.000)	[-0.67]
Eligibility: Domestic Violence	0.30	0.30	0.00
	(0.004)	(0.000)	[0.01]
Female	0.97	0.91	0.06
	(0.002)	(0.000)	[1.26]
Partner/Spouse Present	0.31	0.25	0.06
	(0.004)	(0.000)	[0.99]
Black	0.43	0.57	-0.14
	(0.006)	(0.000)	[-1.79]
Hispanic	0.46	0.37	0.09
	(0.006)	(0.000)	[1.18]
White	0.06	0.02	0.04
	(0.001)	(0.000)	[1.57]
No Degree	0.61	0.57	0.05
	(0.005)	(0.000)	[0.67]
High School Grad	0.30	0.32	-0.02
	(0.005)	(0.000)	[-0.29]
Some College or More	0.06	0.05	0.01
	(0.001)	(0.000)	[0.21]

Full sample. Treatment is in-borough placement. Instrument is 15-day moving average of the initial ineligibility rate for 30-day application period. Compliers are families placed in-borough when the ineligibility rate is high, but not otherwise. Non-compliers consist of always-takers and never-takers. Compiler and non-complier characteristics, adjusted for year and month of shelter entry, are estimated from the algorithm described in Cassidy (2020). Standard errors (in parentheses) and differences in means (with t-stats in brackets) are calculated from 200 bootstrap replications, clustering by family group.

Table 6B: Complier Characteristics: Ineligibility Rate Instrument

	Compliers	Non-Compliers	Diff.
On Cash Assistance	0.30	0.36	-0.06
	(0.005)	(0.000)	[-0.78]
On Food Stamps	0.75	0.73	0.02
	(0.006)	(0.000)	[0.29]
Employed Year Pre	0.39	0.44	-0.05
	(0.005)	(0.000)	[-0.67]
Tier II Shelter	0.63	0.54	0.08
	(0.004)	(0.000)	[1.27]
Commercial Hotel	0.08	0.29	-0.21
	(0.005)	(0.000)	[-3.08]
Family Cluster Unit	0.19	0.16	0.02
	(0.003)	(0.000)	[0.46]
Family Size 1–3	0.54	0.64	-0.11
	(0.006)	(0.000)	[-1.43]
Family Size 4–5	0.39	0.28	0.11
	(0.004)	(0.000)	[1.68]
Family Size 6+	0.07	0.08	-0.01
	(0.002)	(0.000)	[-0.27]
Age	31.69	31.53	0.16
	(1.350)	(0.013)	[0.14]
Log AQ Earnings Year Pre	2.73	3.03	-0.30
	(0.259)	(0.002)	[-0.60]

Full sample. Treatment is in-borough placement. Instrument is 15-day moving average of the initial ineligibility rate for 30-day application period. Compliers are families placed in-borough when the ineligibility rate is high, but not otherwise. Non-compliers consist of always-takers and never-takers. Compiler and non-complier characteristics, adjusted for year and month of shelter entry, are estimated from the algorithm described in Cassidy (2020). Standard errors (in parentheses) and differences in means (with t-stats in brackets) are calculated from 200 bootstrap replications, clustering by family group.

Table 7A: Complier Characteristics: Aversion Ratio Instrument

	Compliers	Non-Compliers	Diff.
Manhattan Origin	0.09	0.13	-0.04
	(0.001)	(0.000)	[-1.18]
Bronx Origin	0.55	0.39	0.16
	(0.004)	(0.000)	[2.67]
Brooklyn Origin	0.20	0.33	-0.13
	(0.003)	(0.000)	[-2.29]
Queens Origin	0.12	0.13	-0.01
	(0.002)	(0.000)	[-0.17]
Staten Island Origin	0.03	0.03	0.01
	(0.000)	(0.000)	[0.51]
Health Issue Present	0.35	0.29	0.06
	(0.002)	(0.000)	[1.27]
Eligibility: Eviction	0.33	0.34	-0.01
	(0.003)	(0.000)	[-0.13]
Eligibility: Domestic Violence	0.27	0.30	-0.03
	(0.002)	(0.000)	[-0.59]
Female	0.94	0.91	0.03
	(0.001)	(0.000)	[0.82]
Partner/Spouse Present	0.28	0.25	0.03
	(0.002)	(0.000)	[0.61]
Black	0.47	0.57	-0.09
	(0.004)	(0.000)	[-1.51]
Hispanic	0.43	0.37	0.06
	(0.003)	(0.000)	[0.98]
White	0.05	0.02	0.03
	(0.000)	(0.000)	[1.46]
No Degree	0.59	0.57	0.02
	(0.003)	(0.000)	[0.41]
High School Grad	0.31	0.32	-0.01
	(0.002)	(0.000)	[-0.21]
Some College or More	0.07	0.05	0.02
	(0.001)	(0.000)	[0.76]

Full sample. Treatment is in-borough placement. Instrument is 15-day moving average of the aversion ratio. Compliers are families placed in-borough when the aversion ratio is high, but not otherwise. Non-compliers consist of always-takers and never-takers. Compiler and non-complier characteristics, adjusted for year and month of shelter entry, are estimated from the algorithm described in Cassidy (2020). Standard errors (in parentheses) and differences in means (with t-stats in brackets) are calculated from 200 bootstrap replications, clustering by family group.

Table 7B: Complier Characteristics: Aversion Ratio Instrument

	Compliers	Non-Compliers	Diff.
On Cash Assistance	0.23	0.37	-0.14
	(0.003)	(0.000)	[-2.43]
On Food Stamps	0.70	0.74	-0.04
	(0.003)	(0.000)	[-0.63]
Employed Year Pre	0.39	0.44	-0.05
	(0.003)	(0.000)	[-0.91]
Tier II Shelter	0.59	0.55	0.04
	(0.003)	(0.000)	[0.82]
Commercial Hotel	0.14	0.29	-0.16
	(0.003)	(0.000)	[-2.83]
Family Cluster Unit	0.19	0.16	0.02
	(0.001)	(0.000)	[0.63]
Family Size 1–3	0.60	0.64	-0.04
	(0.003)	(0.000)	[-0.76]
Family Size 4–5	0.35	0.28	0.06
	(0.003)	(0.000)	[1.26]
Family Size 6+	0.05	0.08	-0.02
	(0.001)	(0.000)	[-0.82]
Age	32.20	31.47	0.73
	(0.949)	(0.014)	[0.74]
Log AQ Earnings Year Pre	2.76	3.03	-0.27
	(0.147)	(0.002)	[-0.71]

Full sample. Treatment is in-borough placement. Instrument is 15-day moving average of the aversion ratio. Compliers are families placed in-borough when the aversion ratio is high, but not otherwise. Non-compliers consist of always-takers and never-takers. Compiler and non-complier characteristics, adjusted for year and month of shelter entry, are estimated from the algorithm described in Cassidy (2020). Standard errors (in parentheses) and differences in means (with t-stats in brackets) are calculated from 200 bootstrap replications, clustering by family group.

Table 8A: Regression Discontinuity Main Results

	(1)	(2)	(3)	(4)
A. Stays and Returns				
Log Length of Stay	1.986**	1.612**	1.357**	1.065**
	(0.705)	(0.271)	(0.436)	(0.331)
	{7,679}	{14,925}	{50,480}	{50,480}
Subsidized Exit	0.353*	0.661**	0.622**	0.370**
	(0.211)	(0.106)	(0.171)	(0.126)
	{7,548}	{14,642}	{49,334}	{49,334}
Returned to Shelter	-0.067	-0.247**	0.013	-0.042
	(0.153)	(0.075)	(0.120)	(0.101)
	{6,798}	{13,268}	{44,574}	{44,574}
First Stage	0.051**	0.089**	0.051**	0.058**
	(0.011)	(0.008)	(0.013)	(0.012)
	[20.4]	[117.8]	[89.6]	[104.1]
Order	Wald	Wald	Linear	Linear
Bandwidth	{-1,0}	{-2,-1,1,2}	[-3,12]	[-3,12]
Threshold	Yes	No	Yes	Yes
Covariates	No	No	No	Yes

The table presents fuzzy regression discontinuity analysis using families' oldest children's potential grades (end-of-calendar-year age year minus five) as the running variable. Each cell reports the coefficient on in-borough shelter placement from a separate 2SLS regression of the row-delineated outcome on the treatment indicator, using as the instrument an indicator for whether a family's oldest child's potential grade is zero or greater. Columns 1 and 2 give Wald estimates pooling the running variable for the given bandwidth; coefficients are thus instrumented mean comparisons between families without and with school-aged children. Columns 3 and 4 fit linear regressions on the running variable for the given bandwidths, allowing for different slopes on either side of the threshold; the coefficients are the differences in intercepts at the threshold. Column 4 controls for Main RD covariates. Standard errors clustered at family group level in parentheses. Number of observations given in braces below first outcome in each panel, as well as for any subsequent outcome where the sample size differs due to censoring. First-stage given for in-borough placement indicator. First-stage F-stat, in brackets, given for log length of stay regressions. * $p < 0.10$, ** $p < 0.05$

Table 8B: Regression Discontinuity Main Results

	(1)	(2)	(3)	(4)
B1. Year Post-Entry Outcomes				
Cash Assistance	0.223	0.025	0.170	0.183**
	(0.188)	(0.076)	(0.128)	(0.092)
	{7,679}	{14,925}	{50,480}	{50,480}
Food Stamps	0.070	-0.137**	-0.037	0.009
	(0.130)	(0.056)	(0.090)	(0.055)
Employed	0.001	-0.268**	-0.123	-0.081
	(0.223)	(0.098)	(0.156)	(0.114)
Log Avg. Quarterly Earnings	0.881	-1.131	-0.568	-0.277
	(1.623)	(0.690)	(1.124)	(0.815)
B2. Year Post-Exit Outcomes				
Cash Assistance	0.403**	0.138*	0.398**	0.347**
	(0.191)	(0.084)	(0.152)	(0.120)
	{6,295}	{12,246}	{41,110}	{41,110}
Food Stamps	0.212	-0.107*	0.091	0.071
	(0.130)	(0.059)	(0.100)	(0.073)
Employed	-0.147	-0.287**	-0.219	-0.135
	(0.203)	(0.099)	(0.162)	(0.128)
Log Avg. Quarterly Earnings	-0.901	-1.533**	-1.606	-0.909
	(1.485)	(0.714)	(1.192)	(0.935)
	{6,295}	{12,246}	{41,110}	{41,110}
First Stage	0.051**	0.089**	0.051**	0.058**
	(0.011)	(0.008)	(0.013)	(0.012)
	[20.4]	[117.8]	[89.6]	[104.1]
Order	Wald	Wald	Linear	Linear
Bandwidth	{-1,0}	{-2,-1,1,2}	[-3,12]	[-3,12]
Threshold	Yes	No	Yes	Yes
Covariates	No	No	No	Yes

The table presents fuzzy regression discontinuity analysis using families' oldest children's potential grades (end-of-calendar-year age year minus five) as the running variable. Each cell reports the coefficient on in-borough shelter placement from a separate 2SLS regression of the row-delineated outcome on the treatment indicator, using as the instrument an indicator for whether a family's oldest child's potential grade is zero or greater. Columns 1 and 2 give Wald estimates pooling the running variable for the given bandwidth; coefficients are thus instrumented mean comparisons between families without and with school-aged children. Columns 3 and 4 fit linear regressions on the running variable for the given bandwidths, allowing for different slopes on either side of the threshold; the coefficients are the differences in intercepts at the threshold. Column 4 controls for Main RD covariates. Standard errors clustered at family group level in parentheses. Number of observations given in braces below first outcome in each panel, as well as for any subsequent outcome where the sample size differs due to censoring. First-stage given for in-borough placement indicator. First-stage F-stat, in brackets, given for log length of stay regressions. * $p < 0.10$, ** $p < 0.05$

Table 9A: Family Fixed Effects Results

	Full Sample				Non-DV	Pre-2015
	Base	Placement	Main	Shelter	Main	Main
Outcome	(1)	(2)	(3)	(4)	(5)	(6)
A. Stays and Returns						
Log Length of Stay	0.091**	0.156**	0.158**	0.149**	0.093**	0.133**
	(0.024)	(0.024)	(0.025)	(0.025)	(0.030)	(0.034)
	{20,149}	{20,149}	{20,149}	{20,125}	{11,134}	{12,570}
Subsidized Exit	-0.020**	0.029**	0.026**	0.024**	0.023*	0.033**
	(0.009)	(0.008)	(0.009)	(0.009)	(0.012)	(0.011)
	{19,659}	{19,659}	{19,659}	{19,633}	{10,850}	{12,467}
Returned to Shelter	0.011	-0.013	-0.007	-0.004	-0.001	-0.005
	(0.011)	(0.011)	(0.011)	(0.011)	(0.015)	(0.014)
	{17,464}	{17,464}	{17,464}	{17,444}	{9,597}	{12,089}
Placement Controls	No	Yes	Yes	Yes	Yes	Yes
Family & Shelter Controls	No	No	Yes	Yes	Yes	Yes
Shelter FE	No	No	No	Yes	No	No

Each cell reports the coefficient on in-borough shelter placement from a separate OLS regression of the row-delineated outcome on the treatment indicator, controlling for family fixed effects. Standard errors clustered at family group level in parentheses. Number of observations given in braces below first outcome in each panel, as well as for any subsequent outcome where the sample size differs from the first due to censoring. * $p < 0.10$, ** $p < 0.05$

Table 9B: Family Fixed Effects Results

	Full Sample				Non-DV	Pre-2015
	Base	Placement	Main	Shelter	Main	Main
Outcome	(1)	(2)	(3)	(4)	(5)	(6)
B1. Year Post-Entry Outcomes						
Cash Assistance	0.017**	0.016**	0.017**	0.019**	0.016*	0.013*
	(0.006)	(0.006)	(0.006)	(0.006)	(0.009)	(0.008)
	{20,149}	{20,149}	{20,149}	{20,125}	{11,134}	{12,570}
Food Stamps	-0.001	-0.002	-0.002	-0.001	-0.001	-0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
Employed	0.010	0.014*	0.017**	0.016**	0.010	0.013
	(0.007)	(0.007)	(0.008)	(0.008)	(0.011)	(0.010)
Log Avg. Quarterly Earnings	0.062	0.109**	0.137**	0.129**	0.055	0.119*
	(0.049)	(0.050)	(0.052)	(0.053)	(0.073)	(0.066)
B2. Year Post-Exit Outcomes						
Cash Assistance	0.013*	0.013*	0.016**	0.016**	0.026**	0.012
	(0.007)	(0.007)	(0.008)	(0.008)	(0.011)	(0.009)
	{15,585}	{15,585}	{15,585}	{15,569}	{8,498}	{11,820}
Food Stamps	0.006*	0.004	0.006	0.006	0.009	0.002
	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.004)
Employed	-0.005	-0.004	0.001	0.004	0.005	-0.000
	(0.008)	(0.009)	(0.009)	(0.009)	(0.013)	(0.011)
Log Avg. Quarterly Earnings	-0.037	-0.002	0.025	0.047	0.013	0.027
	(0.057)	(0.059)	(0.061)	(0.063)	(0.088)	(0.071)
Placement Controls	No	Yes	Yes	Yes	Yes	Yes
Family & Shelter Controls	No	No	Yes	Yes	Yes	Yes
Shelter FE	No	No	No	Yes	No	No

Each cell reports the coefficient on in-borough shelter placement from a separate OLS regression of the row-delineated outcome on the treatment indicator, controlling for family fixed effects. Standard errors clustered at family group level in parentheses. Number of observations given in braces below first outcome in each panel, as well as for any subsequent outcome where the sample size differs from the first due to censoring. * $p < 0.10$, ** $p < 0.05$

10 Figures

Figure 1: Policy Instruments Time Series

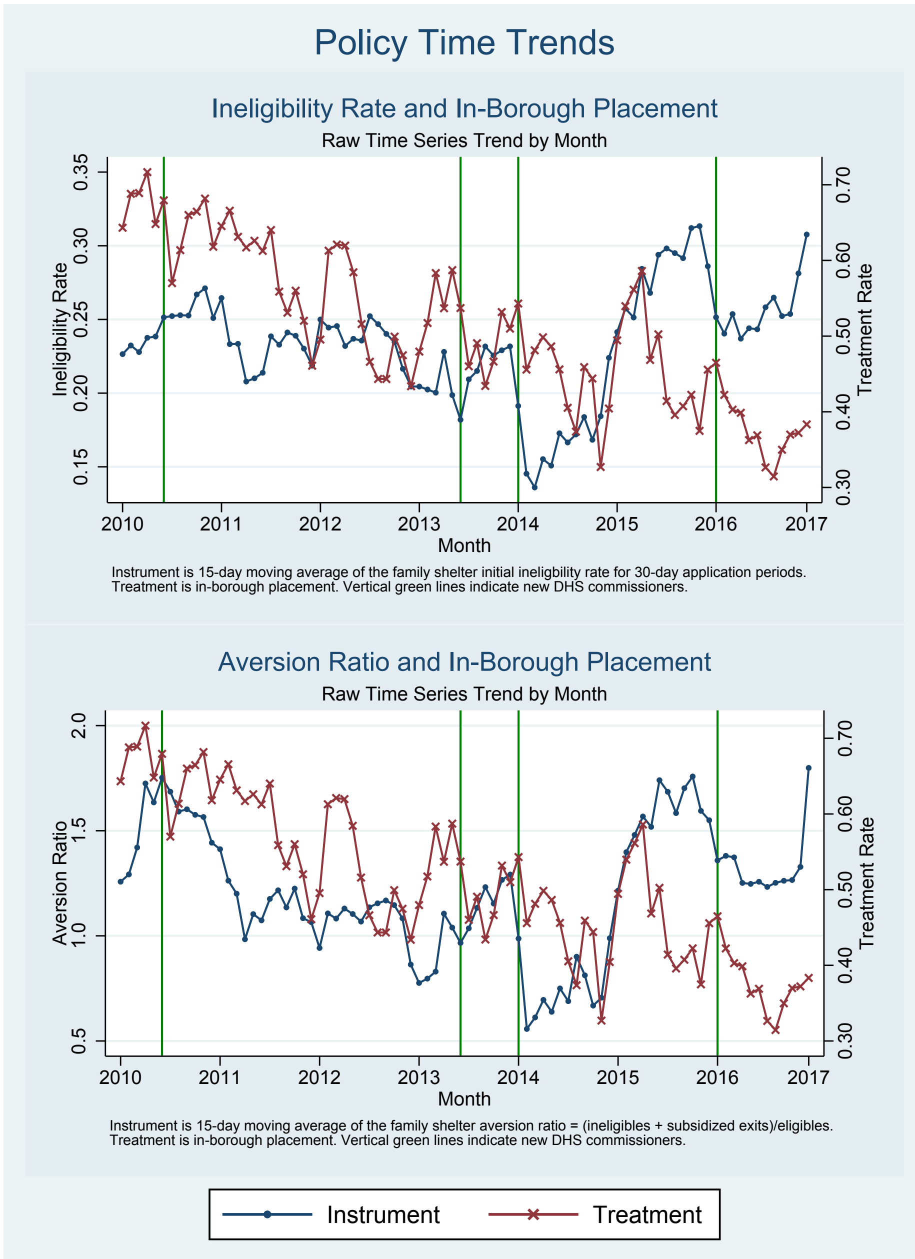


Figure 2: Regression Discontinuity Treatment, Stays, and Returns

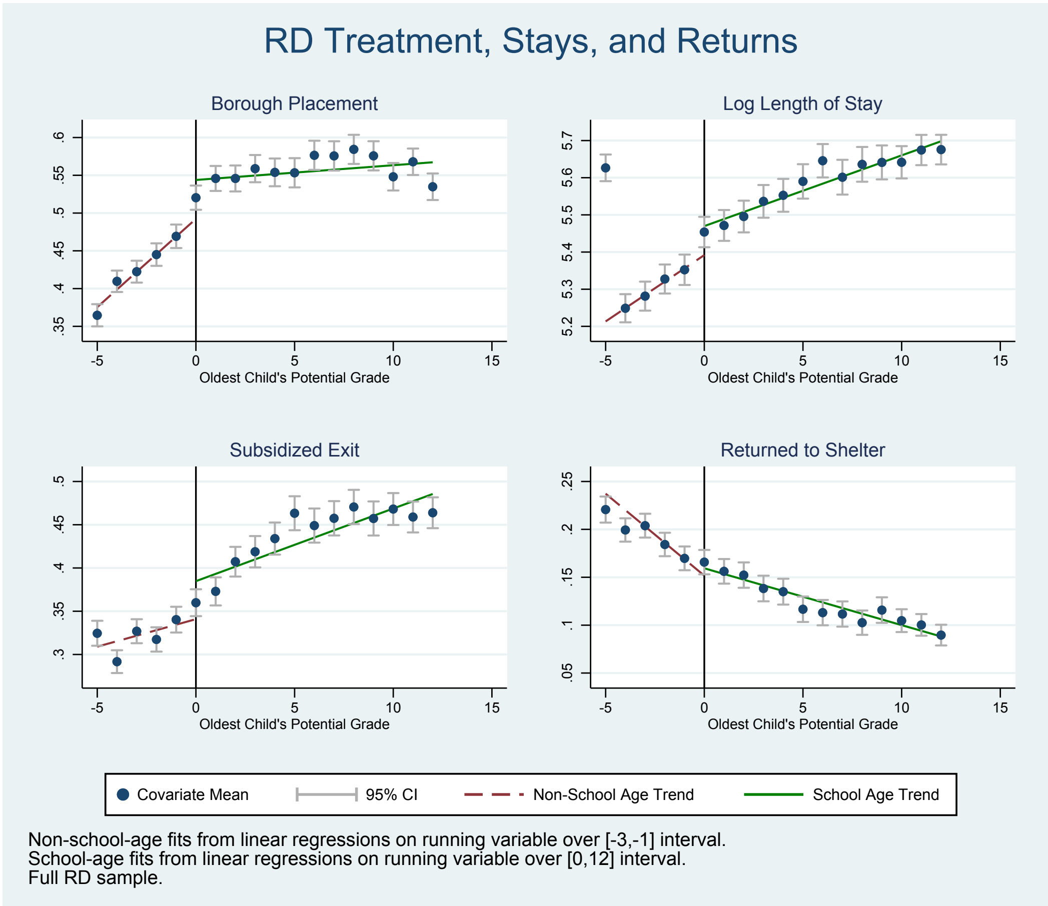


Figure 3: Regression Discontinuity Entry Year Outcomes

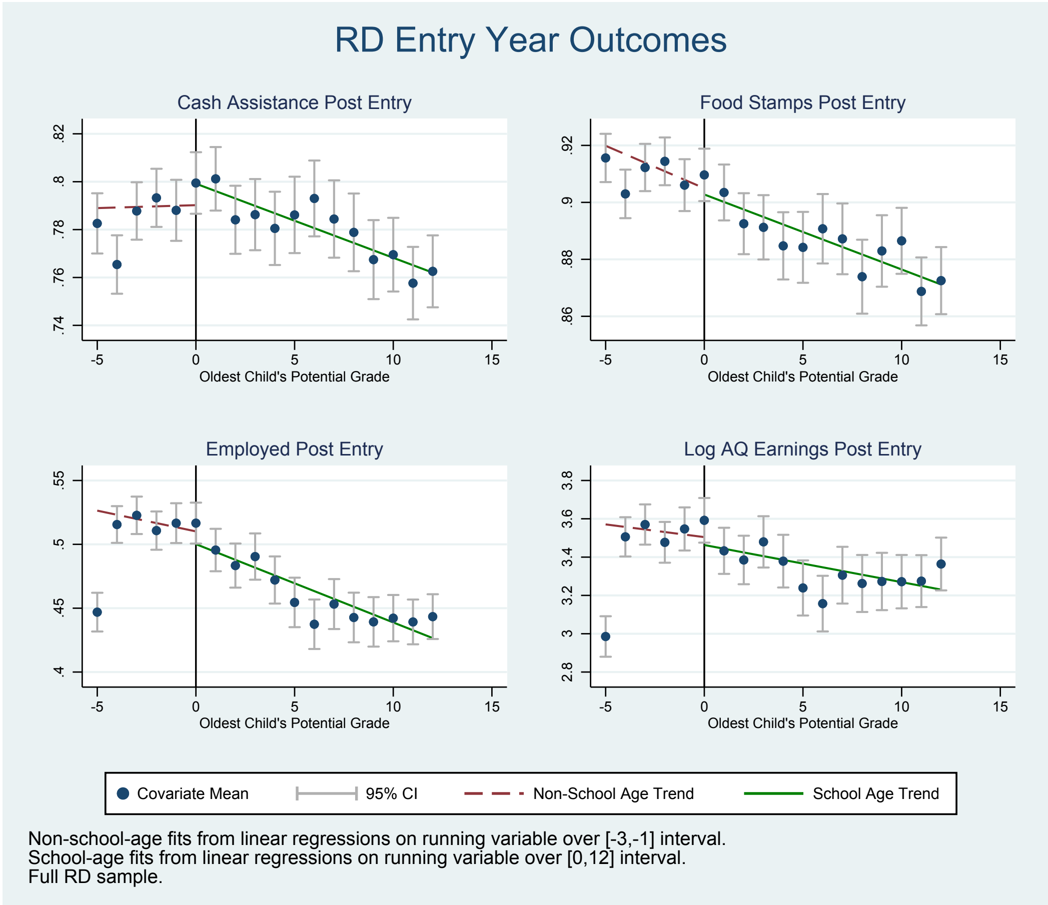


Figure 4: Regression Discontinuity Exit Year Outcomes

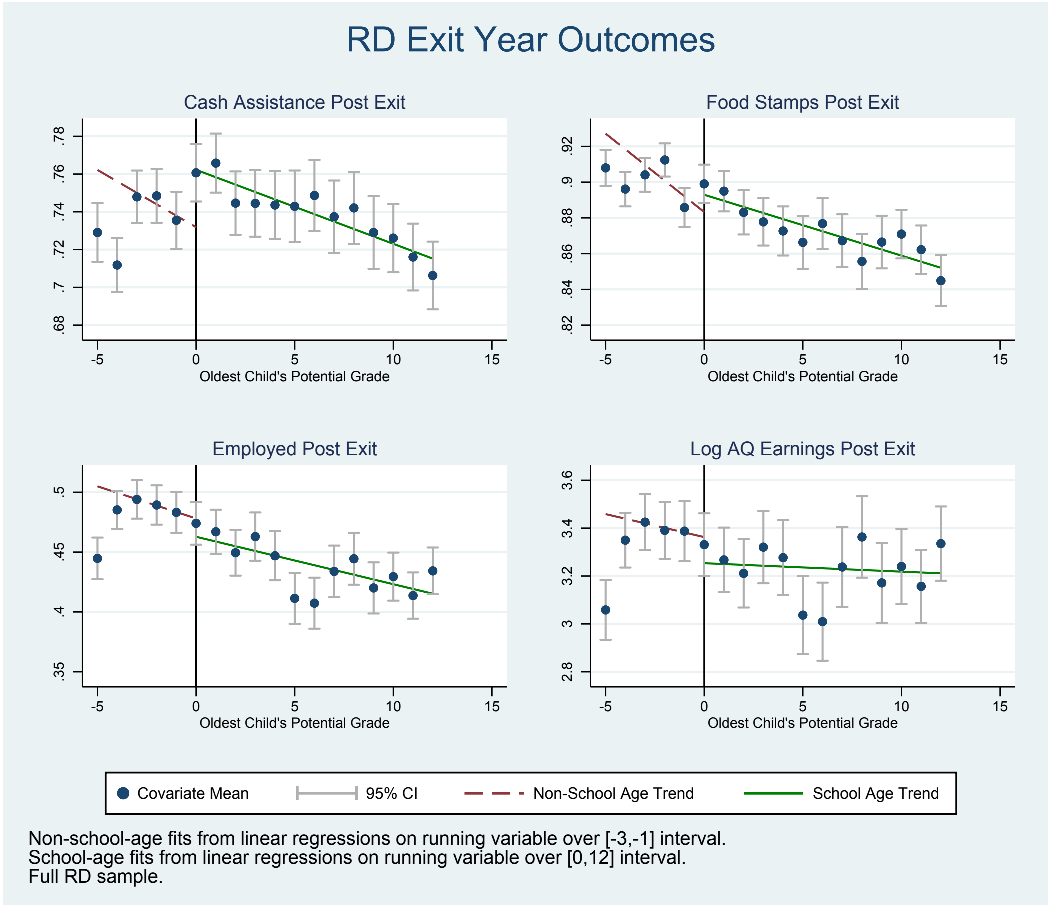


Figure 5: Density of Assignment Variable

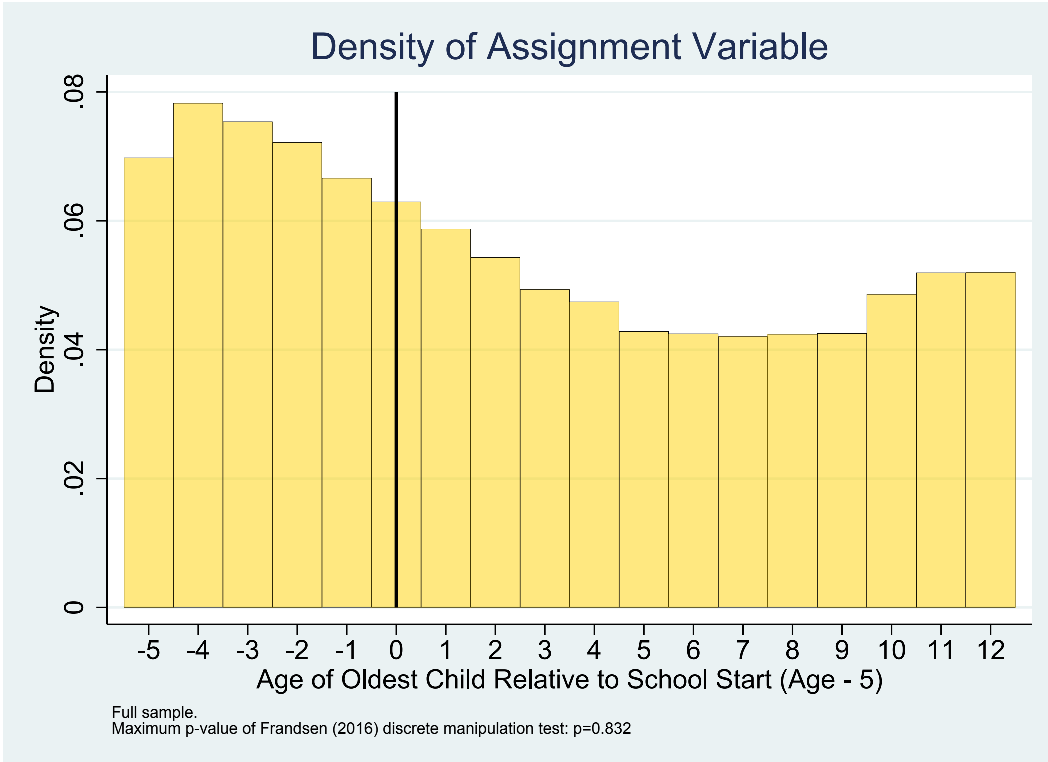


Figure 6: Regression Discontinuity Baseline Covariates

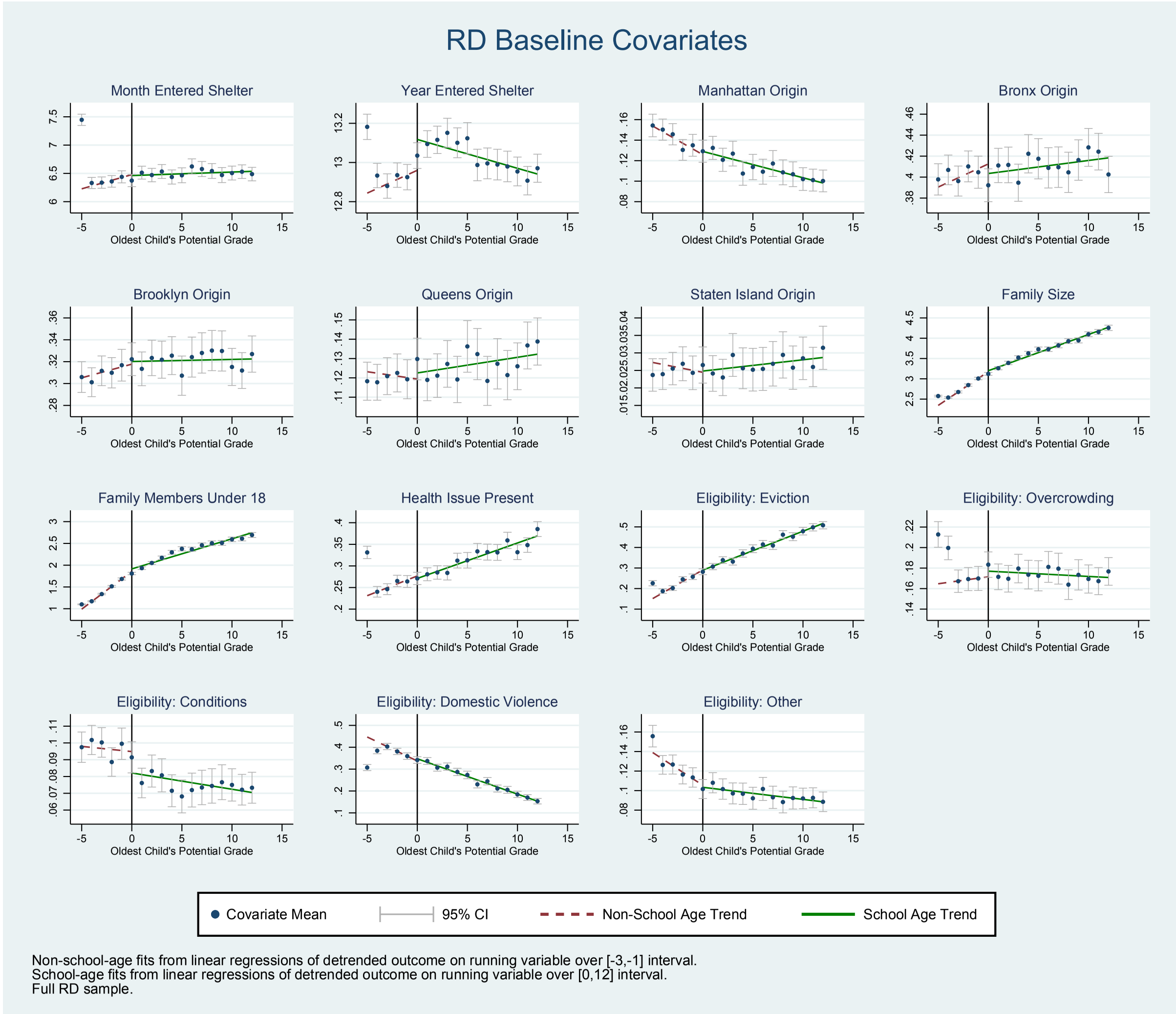
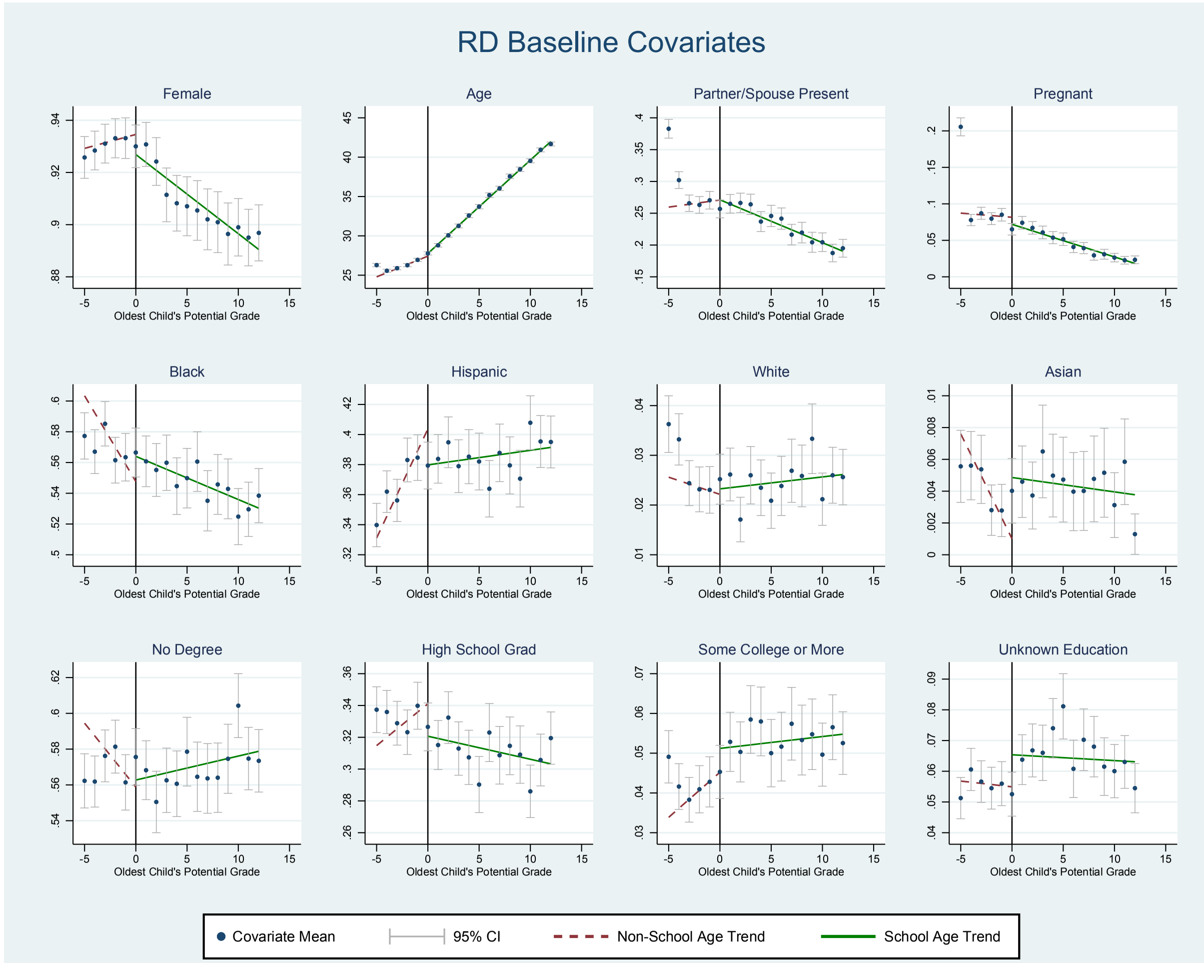
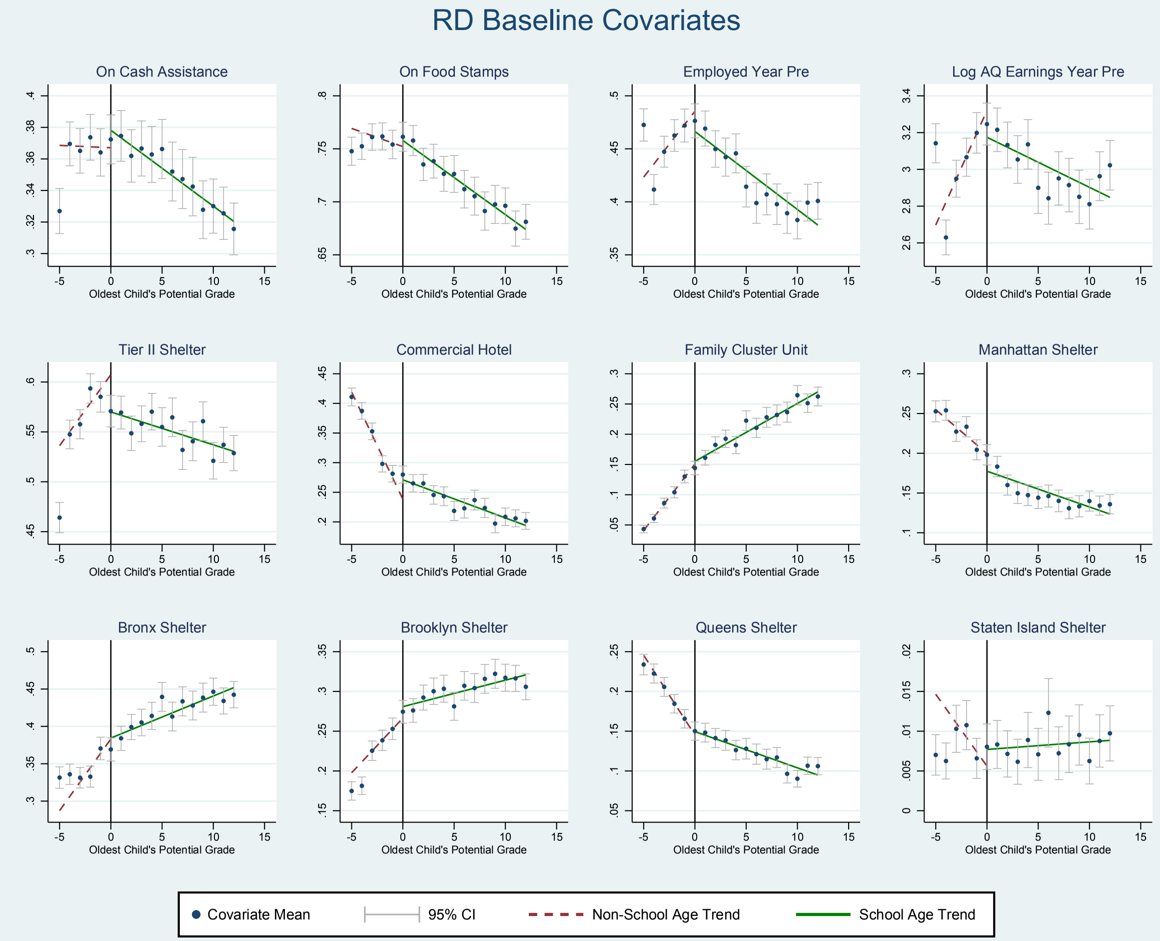


Figure 7: Regression Discontinuity Baseline Covariates



Non-school-age fits from linear regressions of detrended outcome on running variable over [-3,-1] interval.
School-age fits from linear regressions of detrended outcome on running variable over [0,12] interval.
Full RD sample.

Figure 8: Regression Discontinuity Baseline Covariates



Non-school-age fits from linear regressions of detrended outcome on running variable over [-3,-1] interval.
School-age fits from linear regressions of detrended outcome on running variable over [0,12] interval.
Full RD sample.

Supplemental Appendices to "Short Moves and Long Stays: Homeless Family Responses to Exogenous Shelter Assignments in New York City"

A Data Appendix

Unless otherwise noted, data management activities are carried out using `Stata 16`. For certain tasks where `R` has a comparative advantage, I use it instead and make note.

A.1 Data Sources

My data consist of administrative records matched across several City agencies. The core data source is the Department of Homeless Services' (DHS) Client Assistance and Rehousing Enterprise System (CARES), which is the City's management information system of record for homeless families. CARES is designed to accommodate all aspects of homeless services provision and program management. At the front-end, CARES consists of a graphical user interface software application that allows both City staff and contracted service providers to enter, update, and view client information in accordance with role-based access privileges. Behind the scenes is an elaborate relational database where records are stored. While the primary purpose of CARES is prosaic—to permit efficient administration of homeless services—the system also includes fairly robust (if sometime convoluted) reporting capabilities to facilitate program evaluation and statistical reporting.

My sample consists of all eligible family shelter applications from January 1, 2010 to December 31, 2016[1]. I focus on these years because this is the period in which shelter capacity constraints have been the most binding, and thus where the case for random neighborhood assignment is the strongest. In addition, the CARES system came online during 2012; prior to that, DHS relied on less robust information technologies.

CARES is a comprehensive system, encompassing virtually all aspects of family homelessness, from application through case management[2]. Data in CARES is collected from two

[1]Specifically, it consists of all families who both began their application and their shelter stay between 1/1/10 and 12/31/16.

[2]CARES is similarly used to manage single adult homelessness, but as that is not the focus of this study, I do not discuss it here.

main sources. The first is the Temporary Housing Assistance (THA) application, which all families requesting shelter are required to fill out at intake[3]. The THA consists of information pertinent to the eligibility determination and placement decisions made by DHS staff. In addition to basic identifying information for all family members at the time of application (e.g., name, date of birth, Social Security Number) and their relationships, it also contains demographic attributes (e.g., sex, race, ethnicity, pregnancy status) as well as the family's address of origin and reason for applying for homeless assistance. As might be expected, CARES also records information relevant to the application process itself, including application type, eligibility determination outcome and official eligibility reason, diversion efforts, and dates of application and adjudication.

The second main CARES data domain relevant to this paper is known as the Lodge History (Lodge), which, as the name suggests, tracks families' experiences in shelter. Unlike the THA, it is not a form, but rather a query culling key stay-related data from multiple tables (and which are collected at various points during a family's time in shelter). It records the facility, building, and unit into which a family is placed, and for what dates they resided there[4]. It is not uncommon for families to change facilities or units during a shelter stay; correspondingly, the system tracks all of the ins and outs. When families leave shelter, the Lodge component of CARES records the date, as well as the type of exit and destination address (if known)[5].

The distinction between the THA and Lodge is somewhat artificial, as CARES is an integrated application used across multiple DHS administrative units (including eligibility and placement staff) and providers. Thus, families' information can continually be updated or augmented; indeed, the source of a particular data field is sometimes categorized by the main data tables upon which a particular query relies, rather than the point at which the data was collected. Another distinguishing feature is who does the data entry: THA information is entered by frontline DHS staff, while Lodge data may be entered by DHS staff or providers.

One example illustrative of the complexities of data collection in CARES is families' health status. Medical and mental health information relevant to shelter placements is collected via several standard assessments which may take place at various points during a family's shelter stay, beginning at intake. Consequently, DHS' health query comprises data

[3]While NYC has a right to shelter, families must be deemed eligible, in the sense that they are bona-fide homeless with no other place to go.

[4]Some facilities consist of multiple buildings. In the case of cluster units—apartments scattered across otherwise private residences—these buildings may not even be in the same borough. Thus, facility alone, which is more of a synonym for "provider contract," is not sufficient to identify shelter location.

[5]Families are not required to report their exit to DHS.

from both the THA and the Lodge History.

To summarize, CARES client data may be categorized along several non-mutually exclusive dimensions: transactional source, point of collection, user role, topical content (as organized by relational database tables), or the query that extracts it. For purposes of this analysis, I typically classify CARES client data as coming from the THA or Lodge, depending on whether the data is primarily collected at intake (THA) or during a shelter stay (Lodge). Strictly speaking, this may be an oversimplification, but it is one that useful for organizing data concepts.

Though the focus of CARES, first and foremost, is on clients, DHS families need places to go. Consequently, CARES also functions as an inventory management system, allowing staff to track the capacity and occupancy of all homeless shelter units within DHS' purview. These include, in addition to traditional Tier II shelters (these are apartment buildings officially designated as shelters), "cluster" units scattered among private apartments, contracted hotels, and commercial hotels. While the City owns and operates some shelters directly, the majority are under contract with non-profit service providers. This facility management aspect of CARES is critical to the ability of staff to place clients in suitable situations[6].

Correspondingly, the third CARES-based data source for this paper is DHS' facilities query. It includes daily capacity and occupancy for each facility and building within DHS' portfolio, along with addresses and unique identifiers.

Client data from CARES constitutes the core data for this paper. However, it is hardly the case that all information relevant to assessing homeless services is maintained by DHS alone. Indeed, the vast majority of the City's social services and poverty alleviation programs are the domain of the Department of Social Services (DSS). Also known as the Human Resources Administration (HRA), DSS is NYC's officially designated local social service agency[7]. It bears responsibility for administering virtually all of the programs associated with the social safety net, notably: Temporary Assistance for Needy Families (TANF) and its NYS counterpart for single adults, Safety Net Assistance (SNA); the Supplemental Nutrition Assistance Program (SNAP, formerly know as Food Stamps); and Medicaid[8].

Data on public benefit use is maintained in HRA's Welfare Management System (WMS), which is the NYS information system of record for cash assistance (TANF/SNA) and SNAP.

[6] As the facilities management component of CARES is not as well developed as the client management part, DHS also relies on several other information systems to manage facilities.

[7] In this paper, I use "DSS" and "HRA" interchangeably when referring to the agency.

[8] In fact, the relationship between DHS and DSS is complicated and dynamic, largely for reasons having to do with the challenges of family homelessness. DHS was originally part of DSS, until it was spun off as an independent agency in 1993. However, in 2016, Mayor de Blasio again consolidated DHS under the DSS umbrella, managed by a single commissioner, Steve Banks. Nevertheless, it remains conventional to refer to the departments as distinct.

Reporting from WMS is conducted through an analytically-oriented front-end application, the Electronic Data Warehouse (EDW). For this study, HRA provided data for all individuals who interacted with CA from 2001–2016 and SNAP from 2004 to 2016, as well as the type of assistance received and the associated dates of receipt[9]. Linking information on patterns public benefit use to family shelter stays is critical for understanding how shelter services impact other economic outcomes.

Of course, the ultimate ambition of most government-administered human service programs, from homeless services to poverty assistance, is employment and earned income. Accordingly, a rigorous evaluation of family homelessness policy must include an accounting of labor market outcomes. To that end, the New York State Department of Labor (DOL) has provided quarterly employment and earnings data for all DHS family shelter clients whose Social Security Numbers match DOL records. This labor data spans the first quarter of 2004 to the first quarter of 2017.

An earlier version of this paper, completed in November 2017, was based on DHS data through 2016. Subsequently, in early 2018, DHS data on stays of 2010–2016 family shelter entrants was provided and used to match the 2001–2016 CA and FS records. A second DHS data update, in May 2019, revised length of stay, exits, and returns data for the 2010–2016 DHS families cohort through May 2019. However, no additional match with CA, FS, or DOL records was made. Due to improved data quality, the May 2019 DHS data update also revised pre-2017 shelter exit dates for a small number of spells, with implications for pre-existing public benefit data matches. 79 spells previously marked as incomplete ended prior to 2017, and thus should have CA/FS data, while 7 spells erroneously marked as complete—and thus with CA/FS outcome data—are included in the sample. These observations have no meaningful impact on the results.

A.2 Querying

CARES is an ambitious and detailed information system, customized for DHS' unique needs with many features, user levels, and purposes. Although it was designed, in part, with reporting and analysis in mind, its underlying complexity—literally thousands of relational database tables—means that extracting information often requires a bit of programming gymnastics. In addition to user-entered data, CARES automatically generates several fields, including unique identifiers for individuals, families, and cases, as well as the dates on which transactions (e.g., application approval, moves, case closing) take place. Such automation simplifies data entry and facilitates reporting.

[9]Several demographic variables are present as well, including race and education. These fields can be used as a robustness check on CARES data (or as an IV for measurement error).

The majority of CARES statistical reporting is conducted by means of standard "stock" queries, including the THA and Lodge data discussed above. The underlying SQL code is written and maintained by staff in DHS' Management Information Systems (MIS) and Policy & Planning (PP) units, as well as by CIDI staff. A common extension is joining the results of several queries through unique identifiers. However, in the case of several fields crucial to this study—target schools, shelter building ID's and addresses, race, health status—DHS had to customize existing queries to include additional fields.

To be precise, the DHS data in this paper come from six separate CARES queries:

1. **Standard THA**: described above.

2. **Standard Lodge**: described above.

3. **THA supplemented with target school**: Given DHS' school-based placement policy, caseworkers collect information on youngest child's school. However, this field is sparsely and irregularly populated.

4. **Lodge supplemented with race and shelter building ID**: Standard queries lack building identifiers and the race category variable.

5. **Facilities**: provides daily shelter capacity and occupancy at the facility-building level, along with addresses.

6. **Health**: contains information on family members' medical and mental health (including substance abuse), which may pertain to shelter placement decisions. Health assessments may occur both at intake and during shelter stay.

A.3 Structure of the Data

A.3.1 The Core DHS Data

As described above, the foundational data for this paper consist of a joined standard THA-Lodge query encompassing all eligible families with children who applied for shelter and began their stays between 1/1/2010 and 12/31/2016. The raw data are at the individual-bed stay level: that is, there is one record corresponding to each shelter unit assignment for each individual—437,337 observations in all.

Key variables in the foundational data include: unique family and individual identifiers (including system generated ID's as well as name, date of birth, and SSN); application attributes (e.g., type of application, client-provided homelessness reason, officially determined eligibility reason, address of origin, and key dates in the application process); basic personal characteristics (e.g., sex, household relationships, ethnicity, a pregnancy indicator); and shelter stay characteristics (facility, facility type, dates of stay). The majority of these variables are self-reported by the (prospective) clients; exceptions are staff-designated fields, such as official eligibility reason. However, all information is entered into CARES by caseworkers, providing a measure of validation and error-checking. Of note, this data entry process also provides rationale for asserting that, to the extent errors occur in the data, mismeasurement is of the classical variety.

To this foundational data is appended THA-based target school information and Lodge-based building ID and race category. None of these variables are present in the standard queries. Target school gives the name and code (or sometimes the address) of the youngest child's school, which provides the target shelter neighborhood. Unfortunately, this variable is populated irregularly. Race is self-reported based on standard categories (e.g., White, Black, Asian); note that Hispanic/Latino identity is recorded by the separate ethnicity variable. Building ID gives the precise building where a family is placed within a facility. In CARES nomenclature, "facility" is a loose term, referring more to a distinct provider contract than to a particular location. For example, buildings within cluster facilities may be spread widely across neighborhoods—in some cases, even across different boroughs.

Once a building ID for each family is established, records are linked to the facilities query in order to append data on shelter address (as well as such things as facility and building name).

As a final preliminary step, records are matched to the standalone health query. This provides information on all family members' physical and mental health, including such things a mobility limitations and medical device usage, which in part determine which shelters can suitably accommodate families with special needs.

These queries are linked together based on several identifier fields. Depending on the queries involved, uniquely identifying records may require using several ID fields simultaneously. Together, I refer to the aggregately joined DHS data as the "Core DHS" data.

A.3.2 DSS/HRA Data

On a parallel track, HRA benefits data are processed into a form suitable for linkage to the Core DHS data. Raw HRA data consists of individual-case status level records. There are separate files for each program (CA and SNAP) and each year (2001–2016 for CA and 2004–2016 for SNAP). That is, for each program and each year, a file consists of every case status (applying, active, single issue, sanctioned, closed, denied) each individual had during that year and the corresponding dates. These files also include personal identifiers (name, SSN, DOB, WMS ID, case number) as well as demographic information (e.g., sex, race, education level). Separate years are necessary as the files are very large, containing potentially millions of records.

Variable fields are first cleaned and standardized along the lines described for the DHS data below. Relevant analytical variables, such as length of benefit receipt and benefit indicators, are defined. At the same time, irrelevant variables are dropped, as are individuals too young to be heads of household.

The individual years of data are then appended together into a single file for each program (CA and SNAP) and collapsed to a single summary observation for each unique individual, as indicated by SSN[10]. This process reduces the resulting files—one for CA and one for SNAP—to manageable sizes for purposes of linking to the DHS Core data. As described below, the actual linkage of HRA and Core DHS data occurs only after the Core DHS data is cleaned and collapsed. This sequencing is practical: the linking process relies on probabilistic matching, which can only be accomplished in reasonable time if the number of records is modest.

A.3.3 DOL Data

DOL data consists of quarterly earnings and industry[11] for each individual in the DHS Core data with a matching Social Security Number. That is, in contrast to the DHS–HRA data match, the DHS–DOL match, discussed below, is entirely deterministic, requiring exact SSN matches. Observations are at the individual-quarter level.

Processing the DOL data consists of several steps. First, nominal dollars are converted to real fourth quarter (Q4) 2016 dollars, using the Consumer Price Index (CPI) for All Urban Consumers. In addition, industry codes are summarized in terms of NAICS sectors[12]. Then (and in reference to DHS family-episodes), data are aggregated over the appropriate analytical time periods—the year prior to shelter entry, the year following shelter entry, and the year post-shelter exit. For each of these periods, I define an indicator for employment, a count of quarters worked, and a sum of earnings. Finally, I calculate average quarterly earnings (always dividing by the minimum of four quarters or the number of quarters maximally observed in the given period, regardless of whether an individual was employed). For analytical purposes I add one to this total and take the natural logarithm, thus arriving at measures of log average quarterly earnings for the three periods of interest, and without excluding individuals with zero earnings.

[10]Neither WMS ID nor case number uniquely identify records; moreover, SSN provides a common link to DHS data.

[11]Industry is described by standard North American Industry Classification System (NAICS) codes.

[12]However, I exclude sector covariates from earnings analysis due to the possible simultaneous determination of industry and wages.

A.4 Geocoding and Linking

A.4.1 Preprocessing

Having constructed the DHS portion of the Analytical dataset, two major data management steps remain: linking records across agencies and geocoding. Each is described in its own section below.

To carry out either task with maximal effectiveness, however, first requires cleaning and standardizing the variables implicated. This turns out to be a not inconsiderable challenge.

Geocoding software generally requires addresses to be inputed in standardized format—with, for example, street address, city, and zip codes in stored in separate fields—and largely error free (some software is better than others at discerning near matches). In other words, address data requires some of the highest accuracy of any field to be useful; if it contains errors, the software is unable to code addresses correctly. Ironically, addresses tend to be one of the most error-prone fields in DHS data. Common mistakes include misspelled street names, erroneous zip codes, addresses out of the valid range for a street, and boroughs inconsistent with street names. Particularly problematic are hyphenated addresses and prefixed street names (e.g., East or West). In addition, some entries erroneously merge separate fields (e.g., a street address containing an apartment number).

To address these address issues, I wrote a simple `R` script that corrects the most glaring mistakes. The program takes as its input the list of addresses from my Analytical Stata dataset. It parses addresses into conceptually distinct elements (address number, street, borough, city, state, and zip code). Then, using regular expressions and other string functions, it corrects the most common spelling, punctuation, grammatical, and notational mistakes, resulting in a list of mostly standardized addresses. Finally, using string distance algorithms, it compares street names to an official registry, replacing likely mistakes with their closest valid substitutes. These cleaned and standardized addresses are then inputed to geocoding software, with better success than the raw data.

The second place cleaning and standardization arises is with linking administrative records across agencies. The City does not, in general, have unique cross-agency identifiers for clients who interact with multiple departments. What's more, the standard individual identifier—Social Security Number—is error prone and often missing, either because clients' forget them or never had them. Thus, to achieve the highest possible matching rate between DHS and HRA data—the absence of evidence is not evidence of absence, after all—requires use of probabilistic linkage techniques.

Because probabilistic linkage typically relies on string comparison metrics, the success of the process will only be as good as the quality of the underlying data. Thus, I make simple alterations to improve the data quality of matching fields—first name, last name, date of birth, and SSN. Adjustments include: adding leading zeros to erroneously front-truncated SSN's, ensuring all names are fully uppercase, and arranging dates in standard formats.

A.4.2 Geocoding

Broadly, geocoding is the process of assigning standardized geographical coordinates or categories to addresses, areas, or other spatial positions—in essence, a systematic way of locating places on a map. In the case of administrative records, it entails iterating multiple rounds with specialized software packages.

The first step, as described in the previous section, is to clean and standardize the raw address data queried from CARES. This consists of parsing the data into its topically distinct subcomponents—address number, street name, city (borough), state, and zip—and making several simple cosmetic adjustments, such as removing extraneous punctuation and spaces and enforcing uniform capitalization. This is necessary because geocoding software can be quite literal its interpretation, demanding punctilious formatting and offering scant ability to make approximate matches.

The client address of origin variables from the Stata dataset are then exported to a Microsoft Excel file, which serves as the input to my geocoding software of choice, Geosupport Desktop Edition, version 17.1, which is a highly customized geocoding application for addresses in New York City published by the NYC Department of City Planning (DCP). Usually referred to by its acronym, GBAT, Geosupport Desktop Edition is a publicly available graphical front-end to the comprehensive Geosupport System mainframe application designed and maintained by DCP.

Taking as inputs address number, street name, and borough (or zip), GBAT can return a wide array of geographical classifiers. For purposes of this study, I emphasize several important neighborhood classifications: borough (boro), school district (SD), community district (CD), Census tract (CT), and neighborhood tabulation area (NTA).

I also output spatial X-Y coordinates for each address. GBAT uses the State Plane Coordinate (SPC) system, which approximates the Earth's surface as being flat within relatively confined geographic areas. According SPC, NYC falls in the New York-Long Island zone (NAD 83). With the origin of this zone set to the extreme Southwest, all NYC locations receive positive Cartesian coordinates, with X indicating East and Y indicating North. Units are in feet. Thus, SPC makes it simple to calculate the Cartesian distance between two addresses (NYC Department of City Planning, 2017).

GBAT returns an updated Excel file appended with the geocoded fields, which is straightforward to merge back into the original Stata dataset using unique record identifiers. (Recall there is one record per family-episode.)

Approximately 20 percent of addresses fail to geocode in the first round. For about half of these, this is appropriate: the addresses are outside NYC, as a nontrivial share of the family shelter population arrives from other cities and states (though some of these families may have prior ties to NYC).

The other half of geocoding failures are attributable to frequent errors in the raw DHS data. To remedy such mismeasurement, I import the list of failed addresses into `R` and implement the address cleaning program discussed in the previous section. This code corrects common data entry errors, such as misspellings and inconsistent use of directional prefixes. I then export the results to a second Excel file and repeat the GBAT geocoding process. This improves the success rate somewhat.

Overall, 57,500 of 70,000 client address observations code successfully. Of the remainder, 7,300 are out-of-towners. 5,200 fail to geocode. Future work will entail investigating the reasons for these failures and writing code to improve the success rate. In other words, the iterative data cleaning-geocoding process will repeat several more cycles.

Of course, addresses of origin are only half the story, and I repeat the geocoding process for shelter building addresses. As these addresses are maintained by DHS staff, the success rate is quite high.

Finally, with all geocoding data merged back into the Analytical dataset, I use the geocoded neighborhoods to classify families assigned to shelters in their neighborhoods of origin and those placed in distant neighborhoods. Given the fluid definition of neighborhood, I use the full set of potential categories: borough, SD, CD, CT, NTA, and zip. Spatial coordinates also permit a continuous proximity metric.

Future work may also involve geocoding exit addresses in those cases where these addresses are known.

A.4.3 Record Linkage

In the presence of common individual identifiers, linking records from disparate databases is simple and fast. Unfortunately, DHS family and individual ID's are not the same as those used by HRA in the administration of CA and SNAP, complicating the task of discerning patterns of public benefit use among homeless families.

In principle, Social Security numbers should serve as a cross-agency link, but in practice SSNs are frequently entered erroneously or missing. Thus, it is necessary to rely of probabilistic, or stochastic, linking methods. Also know as "fuzzy matching," there are several probabilistic linkage techniques common in the computer science and statistics literatures, most of which entail the use of string comparison metrics and are based on the pioneering work of Fellegi and Sunter (1969).

Though the mathematics can get complicated, the basic idea is to compare all possible pairs of records in each data set and assess their similarity—for instance, by counting the number of changes (insertions, deletions, and substitutions) to one string necessary to arrive at the other (the Levenshtein distance), or by considering the number of shared character sequences of a given length (q-grams). Patterns of matches among the compared fields are fed into an maximum likelihood type algorithm in order to categorize probable matches and non-matches, with probability thresholds set to distinguish true matches. Though sophisticated, these techniques also review considerable clerical review and judgment calls.

In this study, I primarily rely upon the user-written `reclink2` Stata command, which utilizes a bigram (two-character) string comparator and achieves success rates on the order of 97 percent(Wasi, Flaaen et al., 2015). In some cases, I also rely upon the `R` packages `RecordLinkage` and `stringdist` (Sariyar and Borg, 2010; Borg and Sariyar, 2016; van der Loo, 2014)[13]. I match on four variables: SSN, first name, last name, and date of birth (as a six-digit string with two-digit day, month, and year).

Besides distinguishing between true matches on the one hand and false positives and false negatives on the other, the other major challenge of probabilistic record linkage is computational efficiency. Comparing datasets of size m and n requires $m \times n$ computations, which become unmanageably slow on computers with conventional memory capabilities, given the millions of records involved.

I employ several strategies to improve the speed of computation. First, I reduce the linking datasets to the minimal useful record sets. In the case of the core Analytical dataset, this means running the match after collapsing the data to one observation per family-episode (so that the match occurs based on household head only). For the HRA data, this entails dropping all observations with a date of birth such that they would not be 16 years of age by the end of the sample period (New York requires individuals to be 16 in order to be a CA or SNAP head of household), as well as collapsing to a unique observation for each SSN.

However, there are still in excess of 2 million CA observations and 3 million SNAP observations that must be matched with the 68,079 DHS family observations. Exact matches—where all four fields perfectly correspond—reduce the workload greatly. About 57,000 DHS observations are perfect matches, removing these from subsequent computation. In addition, as is conventional, I employ a "blocking" strategy on all four linking variables, which means than only pairs with an exact match on at least one of these fields is considered, significantly reducing the number of comparisons. Finally, I match on CA first and then take only the

[13]The help files and associated journal articles documenting these commands have also been invaluable resources in learning about the techniques, as described above.

remaining non-matches to the larger SNAP data; this is possible because HRA maintains common identifiers across the programs it administers.

Erring modestly on the side of false positives, I successfully match about 67,600 of the 70,000 DHS families to HRA—in line with what would be expected about homeless family participation of public benefit programs.

Currently, I am able to identify whether a family received CA or SNAP, during which years, and their lifetime lengths of benefit receipt. The next steps in this process are to use the unique identifiers—which obviate the need for future fuzzy matching—to link DHS data to the uncollapsed HRA data sets, in order to distinguish between benefit receipt occurring before, during, and after shelter episodes. This is a data-intensive task, since it entails unique start and end dates for each family (rather than simple year indicators). However, since my DHS and HRA data are now deterministically linked, it should be computationally feasible.

The linking process for the DOL data is simplified by an administrative constraint: because DOL conducts strictly deterministic SSN matches with DHS data, my DOL data sample consists only of successfully matched SSN's present in the DHS Core data.

A.5 Defining Analytical Variables

Having pre-processed each data set—DHS, HRA, and DOL—and defined data linkage rules, what remains is to use the raw data to construct variables that are most appropriate for analytical purposes. These variables include both covariates to be used as controls (e.g., earnings and benefit use pre-shelter) as well as outcomes (e.g., earnings and benefit use post-shelter). Creating these variables is not a simple task, either conceptually or logistically.

The complexity arises from the flow nature of the data sample: I do not observe all families for the same length of time. This is true not only of the core DHS data in isolation—obviously families who enter shelter in 2016 have less potential observation time than those entering in 2010—but, in fact, it is doubly true of the matched HRA and DOL data: families who enter shelter earlier in my sample have less potential observation time pre-shelter and more potential observation time post-shelter. As a result, raw comparisons of earnings, employment, or benefit use can be misleading—biased as an artifact of the sampling scheme.

To best put families on an equal footing for purposes of benefit and employment analysis, I take the approach of focusing three one-year windows: the year (or, as necessary, four quarters) prior to shelter entry, the year following shelter entry, and the year following shelter exit. (When quarters are the unit of time, all such periods are defined as excluding the quarter of transition and inclusive of the following four quarters. When days are the time unit, periods begin on the day of transition and extend for the the next 365 days, inclusive.)

Because observations can still be censored within these year intervals, my second normalization is use indicator or rate variables. Specifically, for benefit use and employment, I prioritize binary indicators (e.g., a dummy for employment or CA receipt) or fractional responses, with denominators set to the minimum of a year or the length of observation before censoring (e.g., percent of quarters employed or percent of days active on CA). For earnings, I focus on average real quarterly earnings, where the denominator is the minimum of four quarters or the number of quarters before censoring. In addition, I count all quarters, whether or not employed, so this measure is not conditional upon working.

A second complexity is that some families are observed for more than one episode during the sample period, necessitating separate computation of these analytical variables for each episode, which, for technical reasons, requires considerable care, as well as iterating the variable definition code for each episode instance. For purposes of variable definition, my general approach is to treat each episode as independent. This means that certain components of the raw data can overlap episodes. For example, if a family reenters shelter within six months of exiting, the subsequent six months of earnings will count as post-exit earnings for the first episode and post-entry earnings for the second episode.

A.5.1 DHS Analytical Data: Reshaping and Conceptualizing

Returning to the Core DHS data, the centerpiece of the analysis, the first step in creating the final "Analytical" dataset is to organize and restructure the raw data. The raw individual-bed date file structure is too detailed to be analytically tractable, so the basic idea is to collapse records into a single observation for each family and shelter episode. As described below, this data management process consists of four key activities: reshaping, deduplicating, defining, and recoding.

To do so is not necessarily straightforward, as it requires defining the key concept of shelter "episode." Conceptually, an *episode* is a discrete stay in shelter. However, it is common in the family shelter system that families enter and exit multiple times in close proximity—a few days in and a few days out—as they shuttle between shelter apartments, family, and friends. Brief hiatuses are not true exits. Conventionally, DHS defines the true end of a shelter episode as one in which a family does not return for at least 30 days; thus, any return within 30 days is considered to be part of the same episode.

I adopt the same 30-day standard for defining episodes in this paper. However, this notion does not have an analogue in CARES; case numbers, which are probably the closest proxy, are not defined by gaps in stays but by applications and case composition.

Thus, it is necessary to define an episode "by hand." To do so, I order observations by family ID (which uniquely identify families)[14]. A further complexity in this regard is that, in the raw data, there are potentially multiple observations for each individual in each family and, moreover, family composition can change during the course of a stay as members enter and leave[15]. This creates complex patterns of overlapping and interweaving shelter unit stays for families; recall that each move within the shelter system—it is common for families to move to different units within a building or to different facilities altogether—triggers a new observation in the raw data. What's more, data for certain fields are occasionally missing, which complicates accurate ordering of the data.

To deal with these complications, I take the following approach in defining episodes. First, I drop any observations with irredeemably missing data (e.g., lack all key identifiers), about 10,000 observations in all (a trivial fraction of the data). I then define the start date of an observation as the "bed start" date for that record (in DHS terminology, "bed start" means beginning of stay in a particular unit), or, if this is missing, as the application date. The corresponding observation end date is the "bed end" date for the record, or, if it is missing, the exit date. I then order the observations by date within each unique family ID. Note that in this setup, observations for each individual in the family are not sequential; the continuity of a family-episode is defined by the continued (without > 30-day gaps) presence of *any* family member, not dependent upon particular family members. I then calculate the gap between the beginning of one observation and the end of its predecessor. Any gap greater than 30 days defines a new episode for that family. Episode start date is defined as the minimum (first) observed date for the family, while episode end date is defined as the maximum (most recent) observation date.

Corresponding to the concept of episode are measures of length of stay (LOS), the proximate outcome of utmost importance to City policymakers. While there is not official LOS metric (and specifically none recorded in the data), DHS maintains two standard concepts.

The most straightforward is *system* length of stay, which is simply defined as the difference, in days, between the family's episode end date and start date. It does not exclude any gaps in stay that might occur if a family leaves temporarily and returns within 30 days. A somewhat more refined concept is *shelter* length of stay, which does deduct shelter occupancy gaps from the total. In practice, the both concepts yield similar results, so for simplicity I favor the system LOS measure. Note that many episodes are censored in the sense of stays not completed during the sample period. Such observations are tracked with a censoring indicator and assigned a LOS based on the latest observed bed end date of 1/1/2017.

[14]Note that an individual may be part of more than one family, e.g., in the case of child that has her own child and subsequently becomes a head of household.

[15]It is not uncommon, for instance, for older children to come and go during a parents' stay in shelter, spending the interludes with relatives.

Having defined a coherent concept of episode, I collapse observations into the desired single observation per family-episode structure. From a data management perspective, this is classified as deduplication: creating unique records at the desired unit of analysis.

Other data management tasks are of the more routine variety, and include the following:

- **Converting variables to formats suitable for analysis**: Many variables are initially stored as strings and must be converted to factors or continuous variables. In addition, dates (also strings) must be converted to analytical date formats.

- **Recoding overly-detailed categorical variables**: Some fields, such as eligibility reason and exit reason, contain a multitude of nuanced codes that can more helpfully be classified in fewer broader categories.

- **Defining derivative variables**: Some variables must be transformed for purposes of analysis. For example, age is more useful than date of birth. Other examples include indicators for year of entry, quarter of entry, incomplete episodes, originating from outside NYC, and having a school age child.

When all is said and done, there is one unique record for each family-shelter episode (some families enter and leave shelter multiple times). The raw data consists of 70,632 family-episodes. 2,553 were dropped due to decisively missing data (e.g., family ID, entry dates, no children present), leaving 68,079 observations in my complete Analytical dataset. However, for two reasons my effective Analytical sample is smaller. 7,099 families originate from outside NYC, leaving 60,980 family-episodes relevant for assessing neighborhood effects (non-NYC families cannot be placed in their home neighborhood). However, 8,008 NYC family-episodes were unable to be geocoded, due to missing or erroneous origin or shelter address. Thus, what I refer to as my "full sample" consists of 52,972 family-episodes, which both originate in NYC and are not missing any defining data.

In addition to a family identifiers, key variables of DHS origin include household demographics (age, sex, race); household composition (household size, number of children, number of adults, ages, and relationship descriptors); address of origin; and homelessness episode attributes (reason found eligible (e.g., eviction, overcrowding, domestic violence), shelter ID, shelter address, shelter type (Tier II, cluster, contracted hotel, commercial hotel), shelter entry date, shelter exit date, exit type (subsidized, unsubsidized, type of subsidy), exit destination type and address).

A.5.2 DSS/HRA and DOL Analytical Data: Reshaping and Conceptualizing

For both the HRA and DOL data, I only retain analytical information only for family heads for computational simplicity. In practice, this is not likely to significantly impact the results, as most families are headed by a single adult, upon who the family depends for both employment and benefits access. Moreover, of necessity, many family covariates, such as race and age, are defined in terms of the household head, so this is consistent with my general approach to defining family attributes.

From a technical standpoint, constructing analytical variables from the HRA and DOL data require four steps. First, using only key individual identifiers (like SSN and name), I create the DHS-HRA and DHS-DOL linkage keys (as described above). Second, I use these keys to respectively merge DHS family-episodes and associate key attributes (like start and end dates) into each of the HRA and DOL datasets. Third, I create the pre/during/post-shelter analytical variables of interest in each dataset. If necessary, I collapse the data so as to maintain a unique observation for each family-episode. Finally, I merge the results back to the DHS Core data, such that my main dataset is neatly appended with the necessary HRA and DOL analytical variables.

In the following section, I outline the basic principles and assumptions used in constructing the key analytical variables. I then describe these variables, organized by source, beginning with those derived from DHS data, followed by HRA and DOL.

A.6 Basic Principles for Analytical Variables

From an econometric standpoint, my population of interest is the universe of potential entrants to NYC family shelter. Viewed from this perspective, my (raw) sample consists of all families who applied for and were found eligible for NYC family shelter from 2010 to 2016.

In the ideal world, I would fully observe all families in my sample, with complete, accurate data on all characteristics of interest, including uncensored lengths of stay and post-shelter outcomes.

In practice, of course, this is impossible. The recency of the data combined with flow sampling guarantees right-censoring; moreover, the censoring point will be variable, with families who entered shelter more recently more likely to be censored.

While I could focus on earlier entrants, there are several strong reasons for not doing so. DHS' information systems underwent a major overhaul in 2011–2012, and the more recent data is higher quality. What's more, shelter capacity has gotten tighter over time, which makes the natural experiment assumption more viable in recent years. Finally, recency means relevance, and all else equal it is of greatest policy interest to characterize the situation today.

But the data is imperfect in other ways, too. While administrative data carries with it the legitimacy of official records, errors remain. In particular, key variables, such as client addresses, can be missing or mistaken. Identifiers can be miscoded or absent as well, and match rates are not 100 percent.

Dealing with these inevitable imperfections means making assumptions. Most important are the following four.

First, I assume censoring is noninformative. That is, conditional on what I can observe, length of stay is independent of censoring time. This is plausible since censoring is an artifact of my flow sampling scheme. Of course, for any given shelter entry date, families that stay longer are more likely to be censored; for purposes of estimating the causal effect of local placement, independent censoring means I must be able to assume uncensored observations are representative of censored ones. In other words, there is no unobservable that is systematically related to both treatment status and censoring.

In some cases, I also make the related assumption that, "selected" observations—families for whom post-shelter outcomes are fully observed because their shelter stays ended early enough relative to the censoring date in my sample—are are representative of those for whom outcomes are unavailable. But I also pursue estimations strategies that allow me to weaken this assumption.

Second, I assume missing data is noninformative. Since missing data can arise in my sample either because a field is missing or because of a non-match, this assumption actually nested two subparts. On one hand, I assume that when fields are missing or miscoded, such errors happen at random—or at least for reasons unrelated to treatment status. On the other, I assume that a non-linkage between DHS and HRA/DOL data consists a true non-match: these families are truly not receiving benefits or not working. Or, at the least, if a false negative occurs (due to, for instance, erroneous SSN), it is at random conditional on observables and not systematically related to treatment status. This assumption is strengthened by the fact that the data is entered by case workers, who both serve as a quality control and a potential source of errors; in either case, the point is that the flawed data is not systematically attributable to family unobservables.

Third, and along related lines, to avoid incidentally truncating the analytical sample, where defensible I code potentially missing data as zero for binary indicators and continuous variables, and as an "unknown" category for categorical variables. This arises in two types of cases. In the first type of case, as with the indicator for health issues, missing values are interpreted as indicative of true absences. Health is an important criterion in shelter placement decisions, and thus families not receiving such a screening are assumed not to have significant limitations. Similarly, a non-link to CA data is interpreted as truly not being on CA. While these assumptions are surely violated in some cases, it is reasonable that they hold on average—and average marginal effects is typically what I am interested in measuring.

The second type of case arises when I introduce covariates to control for potentially confounding influences—but not with the goal of interpreting these covariate coefficients causally. Prominent cases are race and education. Some families do not report their race or have missing education data. I wish to control for race and education when estimating treatment effects, but I do not want to exclude the (small) subsets of families from whom such information is unavailable. Group such families into an "unknown" category is a compromise. While this complicates interpretation of race and education coefficients due to the potential heterogeneity within these groups, these are not the coefficients I care about. What's more, if such data is missing at random, then these categories approximate a group with average characteristics (which is somewhat interpretable). At the other extreme, if data is not unknown at random, unknowingness can itself be informative. As a matter of practice, my results do not much change whether I omit missing data or code it as unknown.

My fourth and final data assumption is to treat family-episodes as independent events, with the exception of clustering standard errors at the family group level. While the data are clearly not completely independent and identically distributed (iid), as an approximation it is not so bad, and it simplifies the analysis. For one thing, over two-thirds of families in the data are present for only one episode. For another, prior research (O'Flaherty, 2010) has demonstrated family homelessness is largely a matter of bad luck—and so the event of becoming homeless, even among those with a history of homelessness, is driven in part by factors beyond a family's control. This, combined with adjusting standard errors appropriately, accounts for arbitrary within family-group correlation of unobservables.

I do, however, explore the robustness of this assumption using several strategies. First, I re-estimate important results keeping only the first episode for each family-group, which leaves the results unchanged. (On the other hand, doing this is undesirable as a control for prior shelter experience, as some families may have had shelter episodes before my sample period began.) Second, at the other extreme, I estimate a family fixed effects specification (which includes families with two or more episodes), and also find my main results to be unchanged.

Having made the necessary assumptions about the data generating process, I adhere to two general rules when defining analytical variables. Note that I use the term "analytical variable" to distinguish variables I create for purposes of analysis from "raw" variables present in the original administrative data. Unless otherwise noted, "variable" used without a qualifier refers to analytical variables, since almost all fields requiring some degree of editing to be suitable for econometric analysis.

The first rule is to define variables at the time of shelter entry. This is sensible because, at least for the DHS data, this is the point at which the data is actually collected. Further, it puts all families on equal footing in terms of their shelter experiences. Finally, for factors where endogeneity might be a concern, it is the point at which conditions are most plausibly exogenous. (For example, initial shelter placement is likely to be more exogenous than subsequent moves to other facilities.) Implicit in this setup is the assumption that variables are time-invariant. As a first approximation, this is probably sufficient. Although family circumstances change (e.g., the birth of a child), most shelter stays are less than two years long, a relatively brief window for evolution. As with most rules, there are a few exceptions to this edict, which I discuss below.

The second rule consists of a two-level hierarchy for assigning characteristics to families. For "compilable" characteristics which are shared by all family members, like shelter assignment or eligibility reason, I do the obvious thing and assign that value upon shelter entry to the family. For compilable characteristics which can be sensibly aggregated across family members (e.g., family size or number of children), I violate the "at-entry" rule and assign the family its maximum (or total, as the case may be) for the episode. For example, family size is defined as the total unique number of family members present during a shelter episode, whether or not initially present. It is relatively common for both children and adults to come and go during the course of a shelter stay (spending interims with relatives or friends). Thus, fully accounting for all family members, rather than just those present on day one, seems more sensible. Econometric considerations guide these choices. For example, *maximum* household size likely best reflects a family's true resource constraints and opportunities, while *initial* shelter assignment is more plausibly exogenous than subsequent moves, which a family may have a stronger role in directing

The second level of family characteristics consists of what I refer to as "uncompilable" characteristics. These are attributes that have no simple aggregate (at least insofar as econometric meaningfulness is concerned), such as age, sex, and race. Rather than try to create summary measures of questionable import (e.g., average age), I instead define these characteristics in terms of the (initial) head of family, on the basis the family head exerts the greater influence on outcomes—especially given that the typical homeless family is consists of a single mother with young children.

With these guiding principles in mind, I now turn to definitions of key concepts and variables. I highlight only the most important variables used in the analysis. For a complete listing of variables and descriptive statistics, refer to the tables at the end of the document.

The following sections categorize variables based on their role in the analysis: outcomes, treatments, or explanatory covariates. In the presentation, I emphasize key assumptions, missing data issues, and resolving potential ambiguities.

A.6.1 Covariates

Most of my explanatory variables consist of family characteristics. Female is a dummy that is equal to one for female head of family and zero otherwise. Age is a continuous measure of the duration between the head's date of birth and shelter entry date. Race consists of six mutually exclusive categories: White, Black, Hispanic, Asian, Other, and Unknown (if race is refused or missing). Partner present is a dummy equal to one if the head's significant other is present in shelter, whether or not such a partner is a married spouse. Family size is a count of unique individuals present at any time during a shelter stay. Children (under 21 year of age) and dependents (which may include adults) are similarly defined. Pregnancy is a dummy equal to one if the family indicates a pregnant member at shelter entry, and zero otherwise. School age is a dummy equal to one if there is a family member present between the ages of five and 21 (inclusive) prior to 2014, and between four and 21 from 2014 on (the year universal pre-k (UPK) began in the City). Health issue is a dummy based on screenings performed by DHS and providers both at intake and during shelter stays. It equals one if any family member has a medical, mental health, or substance abuse issue (each consisting of multiple subcategories). Education consists of four mutually exclusive categories: no degree (less than high school), high school graduate, some college or more, and unknown. On Cash Assistance and On Food Stamps are dummies equal to one if a family has an active benefit case in the respective program at the time of shelter entry. Log average quarterly earnings in the year prior to shelter entry is exactly what it sounds like; it factors in all quarters, whether or not a family is working (and I add one to each family's earnings before taking the log, to avoid omitting these families).

The next important category of controls are shelter covariates: variables related to a family's shelter episode. These include categorical variables for primary (official) shelter eligibility reason (8 categories: eviction, overcrowding, conditions, domestic violence, child welfare, existing case, discharge, and other) and facility type (4 categories: Tier II shelter, commercial hotel, cluster unit, or other). I also include a dummy for whether a family receives diversion services designed to prevent homelessness.

All main regression specifications also include fixed effects (dummies) for year of shelter entry, quarter of shelter entry, borough of origin, and shelter borough. Some specifications also include borough-year fixed effects, which are interaction dummies for year and origin borough and year and shelter borough. To control for unobservable facility and provider quality, some specifications additionally feature facility fixed effects (264 dummies). In all cases with dummies, categorical variables, and fixed effects, a base category is dropped in estimation to avoid multicollinearity in the presence of a constant term.

A.6.2 Treatments

I use several definitions of treatment. Key to treatment definitions are the address data maintained by DHS. To be part of the analytical sample, families must have valid, non-missing, geocodable addresses, both of origin and of shelter. Origin address is defined as the family's "last known address" reported to DHS. Note that a small share of families (less than 4%) report other shelters as their prior address. In light of this, and given that unstably housed family may move frequently, it is best to interpret origin addresses as a place where families have some preexisting community ties.

In my main analysis, treatment is defined as a family being placed in its borough of origin. New York City consists of five boroughs, or counties, Manhattan, the Bronx, Brooklyn, Queens, and Staten Island, ranging in size from about half a million persons in Staten Island to 2.5 million in Brooklyn. Clearly, referring to geographies of such breadth does not quite comport with the conventional definition of a neighborhood. Nevertheless, as geographically contiguous entities with legally designated boundaries, distinct identities, and palpable intra-borough affinities, NYC's five counties do embody many of the characteristics associated with small communities. Boroughs are also appealing as a neighborhood definition from the standpoint of treatment balance: about half of homeless families in my sample are placed in shelters in their home boroughs and half in other boroughs.

Alternatively, I also define neighborhoods in terms of the City's 32 school districts, which are administrative boundaries for the public school system. These are the next largest geographies for which data is readily available; about 9% of my sample is placed in their school districts of origin. Smaller units of geography, such as Community Districts or Census Tracts, do not have sufficient local placements to permit precise analysis.

Finally, I consider a continuous measure of treatment that measures the distance, in miles, between a family's last known address and its shelter address. This is based on Cartesian geospatial coordinates produced by GBAT. It is straightforward to calculate the Euclidean straight line distance between pairs of addresses and convert the units to miles.

A.6.3 Outcomes

I consider a range of outcomes. The most salient one, and the one I feature most prominently, is length of stay (LOS). This is a measure, in days, of the elapsed time between a family's entry into shelter and its exit. In particular, I prioritize a "system" LOS concept, which counts gaps in stay towards the total, so long as these gaps are 30 days or fewer; it is not uncommon for families to leave shelter for a few days, then return. Out-of-shelter gaps longer than 30 days are considered true exits; subsequent returns are consider new episodes. Incidentally, this is how another outcome I consider, returns to shelter within a year of exit, is computed. Subsidized exits from shelter are those in which the family receives any form of rental assistance. This encompasses a variety of programs, which typically offer time-limited benefits that partially offset housing costs so long as the family meets eligibility criteria. An alternative duration measure, "shelter" LOS, excludes the interludes from the count. In practice, both measures produce similar results, so I use the shelter concept, because it is simpler.

I also consider two other primary categories of outcomes: public benefit use and labor market results. My public benefit use data comes from HRA and consists of indicators and durations of families' receipt of Cash Assistance and Food Stamps. I focus on two period: the year post-shelter entry and the year post-shelter exit. While I have durations of active receipt, for simplicity I prioritize dummies indicating active program status at any time during these periods.

My labor market data derives from DOL. Again focusing the year post-entry and year post-exit, I construct indicators for positive earnings during any quarter in those years as my measure of employment. Correspondingly, my measure of earnings is log average quarterly earnings. Average quarterly earnings themselves are in real 2016 dollars, are inclusive of all quarters, whether working or not, and have one dollar added to them for each family, so as not to incidentally drop observations when taking logs.

B Policy Background: Family Homelessness in NYC

Neither homelessness nor poverty are foreign to municipalities anywhere in the United States, but nowhere is the intersection of these issues thrown into starker resolution than it is in New York City. Complicating understanding of homelessness—and perhaps, in part, explaining its absence from economists' agenda—is a fundamental misconception about *who* the homeless really are. While disheveled shopping carts, cardboard tatters, and infelicitous hygiene pervade the popular consciousness, it is actually the case that some 200,000 of the 550,000 Americans who are homeless each day are *families with children* (National Alliance to End Homelessness, 2016; Khadduri and Culhane, 2016).

Unlike their single adult counterparts, this misbranded cohort—"unhoused" African-American and Hispanic mothers and young children is more accurate—suffers not, primarily, from substance abuse and mental illness, but from poverty. Aside from bad luck—often in form of unexpected income loss, health crisis, or domestic strife—these families are otherwise mostly indistinguishable from the marginally housed poor at large, not in the least in that the “shelters” in which they are placed frequently resemble the momentarily unaffordable apartments from whence they came (O’Flaherty, 2010; Culhane et al., 2007; Shinn et al., 1998; Curtis et al., 2013).

For them, homelessness is a temporary condition, not an immutable characteristic—a particularly acute form of poverty manifested in the deprivation of a fundamental element of the consumption bundle (Mullainathan and Shafir, 2013; Desmond, 2016). Getting these families back on their feet fast—or preventing their displacement in the first place—is thus an important policy goal.

The task in an exceedingly difficult one. Since 1994, New York’s homeless census has nearly tripled, from 24,000 to 60,000 in 2017. More than two-thirds of these are people in families; fully 23,000 are children (NYC Department of Homeless Services, 2019*b*). Indeed, NYC accounts for about a fifth of all homeless families in the U.S (NYC Department of Homeless Services, 2019*e*; National Alliance to End Homelessness, 2016).

Family homelessness is particularly pronounced in New York City for two reasons. First, unique among municipalities in the U.S., NYC has a legal right to shelter, the consequence of a series of consent decrees in the 1980s (NYC Independent Budget Office, 2014). The City is legally obligated to provide emergency accommodations to any family able to demonstrate it has no suitable alternative.

This legal mandate has evolved over time as settlements worked their ways through the courts; the right originated from a class action, *McCain v. Koch*, brought by the Legal Aid Society in 1983, in which New York State Court held the City and State were required to provide homeless families with emergency housing under the State Constitution and Social Services Law (NYC Independent Budget Office, 2014; University of Michigan Law School, 2017; Kaufman and Chen, 2008) [16].

Derivative cases during the ensuing decades established standards for temporary shelter, as homeless services where governed by a mix of executive policymaking and judicial edict. A formal settlement was not reached until 2008, in the form of *Boston v. City of New York*, whereupon the Bloomberg administration, Legal Aid, and the courts came to agreement on appropriate eligibility determination and shelter management standards (NYC Independent Budget Office, 2014; University of Michigan Law School, 2017; Kaufman and Chen, 2008). These mandates mean that NYC faces a steady inflow of homeless families in ways that other

[16] A 1981 predecessor case, also brought by Legal Aid, *Callahan v. Cary*, introduced the right to shelter for single adults. (NYC Independent Budget Office, 2014)

cities do not; indeed, a tenth of family shelter entrants report most recent prior addresses that are outside of the City.

Further complicating matters is NYC's notoriously competitive real estate market. New York is a city of renters, with over two-thirds of households renting their residences, nearly double the national average. In the decade ending in 2015, median rent in NYC grew three times the pace of median incomes (18.3% versus 6.6%). Vacancy rates are consistently below 4% (NYU Furman Center, 2016). According the City, demand for affordable apartments exceeds supply by a factor of two; approximately half of renters in the City as rent-burdened, defined as allocating more than 30% of household income to rent (NYC Mayor's Office, 2017). The situation is especially severe for the lowest income families most at-risk for homelessness. Nine in ten households with income below 30% of the area median spent upwards of 30% of their income on rent (NYU Furman Center, 2016).

Expensive housing, paired with poverty's relentless vicissitudes and a legal escape valve, make NYC's steady rise in homelessness none too surprising. The City has had to expand shelter apace. In 2016, shelter vacancy rates were easily below 1%, even as commercial hotels were brought into the mix to fill gaps (NYC Mayor's Office, 2017). Yet adding capacity is a Sisyphean struggle of its own, with proposals for new shelters frequently greeted by virulent community opposition (Stewart, 2017). Homeless service provision is thus forced to strike a delicate compromise between policy ideals and political realities, an important constraint on optimal implementation.

Responsibility for managing shelters and supports for homeless families and individuals falls primarily to the Department of Homeless Services (DHS), a Mayoral agency under the purview of the larger City's Department of Social Services (DSS), which is the City's officially designated local social service agency. Families apply for shelter at at a central intake center in the Bronx, known as PATH (Prevention Assistance and Temporary Housing). There, they are screened by HRA caseworkers for prevention services, including eligibility for temporary rental assistance and anti-eviction legal services, as well as for domestic violence. If alternative housing remedies are unavailable, families apply for shelter apartments, which requires, among other things proper, identification and detailed housing histories. Families are given temporary (generally about 10-day) accommodations while DHS investigative staff assesses eligibility. Families deemed eligible are then given formal shelter assignments by dedicated placement staff, who consider such criteria as family size, health issues, safety, and proximity to children's schools. Often the preliminary and formal shelter assignments are the same[17]. It is this group of families (those deemed eligible) and this placement step

[17]Details are based on NYC Department of Homeless Services (2019*c*); NYC Independent Budget Office (2014) as well as author's conversations with City officials.

(initial formal shelter assignment) that constitute my sample and treatment.

NYC's family shelter system is vast and complex. As of November 2016, the City's shelter portfolio consisted of 169 traditional Tier II shelters (housing 8,617 families and 26,225 individuals), 276 cluster apartments scattered in otherwise private buildings (3,045 familes ; 11,067 individuals), and 68 commercial hotels (2,057 families; 5,798 individuals) (NYC Mayor's Office, 2017).

It is also expensive. In 2017, the average cost of sheltering one family for one night (inclusive of rent and services) was $171. Overall, DHS' budget, inclusive of management operations, is $1.8 billion—and this does not include welfare benefits administered by other agencies (NYC Mayor's Office of Operations, 2017).

Also of note is how services are carried out. While DHS does operate some shelters directly, most homeless services provision is carried out through contracts with community-based non-profit organizations who operate shelters. A case in point: 82% of DHS' budget consists of such contracts. This service arrangement is not unique to homeless services; most social service programs in the City are administered this way (NYC Mayor's Office of Operations, 2017).

Given homelessness' stubborn rise, my sample period, 2010–2016 has been a time of flux for homeless policy in New York City. The sample begins in the aftermath of the Great Recession and concludes at a time when the economy had regained nearly full strength. Michael Bloomberg's mayoralty spanned the first four years, while Bill de Blasio's tenure began in 2014. Developments at the State and Federal levels—both critical funding sources—has also played a leading role.

Throughout this period, a pillar of the City's homelessness strategy has been community continuity. To the extent capacity allows, the City endeavors to place families in their neighborhoods of origin. Predicated on the goal of keeping children in their home schools', the policy reflects a more general premise—that families are better positioned to expeditiously return to permanent housing when they remain connected to their support networks, including relatives, friends, and places of work and worship (NYC Mayor's Office, 2017). Since at least 1997, the city has monitored the share of families placed in shelters according to their youngest child's school as a DHS performance indicator. By this measure, 84 percent of families where successfully placed in their home neighborhoods as of 2010. However, capacity constraints have become increasingly binding as the shelter population has grown. By 2017, this share of families placed in proximity to their children's schools had dropped to 50 percent (NYC Mayor's Office of Operations, 2002, 2012, 2017).

While a full accounting homeless policy developments is beyond the scope of this paper, a brief discussion of its contours provide context. Core elements of the City's strategy to reduce homelessness include prevention, affordable housing, and rental assistance[18].

Homebase, the City's signature homeless prevention program, offers families at risk for homelessness a panoply of supports, ranging from case management and counseling to benefits assistance and referrals. Instituted in 2004, as of 2016 it serves 25,000 families a year. Academic research finds it to be effective in forestalling shelter entries. Further, as of 2017, the City spends upward of $62 million a year on anti-eviction legal services, which helped to avoid about 20,000 evictions per year. Similarly, emergency rental assistance, typically for families in arrears, stop temporary difficulties from ballooning. From 2014–2016, the City allocated $551 million to assisting 161,000 such households. In terms of affordable housing, the de Blasio administration has pledged to create of preserve 200,000 units, of which 62,000 were financed as of 2016. Strategies include zoning regulations, tax credits, and capital funding.

For families in shelter, rental assistance is frequently a catalyst for returns to permanent housing. Advantage, the most prominent Bloomberg-era program, provided some 25,000 formerly homeless families with two years of subsidized housing. At its peak it cost $207 million, but ended in controversial fashion in 2011 when the State withdrew funding. In its place has come Living in Communities (LINC), launched by the de Blasio administration in 2014. LINC, a collection of six programs targeting families meeting various criteria—involving such things as employment, age, or domestic violence status—offers time-limited (usually 2–5 years) rental assistance to families meeting income standards (usually below 200% of the federal poverty level) and minimum shelter stays (usually 90 days). Along related lines, CityFEPS provides families who have been evicted or are at-risk for losing their homeless with an "eviction prevention supplement." Both programs require families to contribute 30% of their income towards rent; the subsidy covers the remainder, up to a maximum of $1,515 for a family of three. From 2014 to 2016, the programs combined to serve more than 26,000 people. Subsequent to my study period, LINC and CityFEPS have been replaced by CityFHEPS (NYC Human Resources Administration, 2019*a*). Traditional federal programs, including Section 8 Housing Choice Vouchers and public housing also play a role, though both have been limited by funding constraints and long waiting lists in recent years. There are also some smaller programs.

Of course, homeless services is but one—albeit highly visible—component of NYC's safety net for low-income families. The City's Human Resources Administration (HRA)—which along with DHS comprises the the Department of Social Services (DSS)—oversees the na-

[18] The following discussion of prevention, affordable housing, and rental assistance is primarily based NYC Mayor's Office (2017) and discussions with City officials. Additional details are provided by: NYC Department of Homeless Services (2019*d*); NYC Independent Budget Office (2011, 2014)

tion's largest apparatus for administering poverty alleviation programs[19]. Notable in HRA's portfolio are the "big three" social benefit programs: Cash Assistance (CA), the Supplemental Nutrition Assistance Program (SNAP, formerly know as Food Stamps), and Medicaid. Because Cash Assistance and Food Stamps figure prominently in the analysis—and also because they help to characterize the poverty that homeless families face—a bit of background is helpful.

Cash Assistance (CA) consists of Temporary Assistance for Needy Families (TANF; which, in New York, is referred to as FA, or Family Assistance) and its State counterpart for single adults and time-limited families, Safety Net Assistance (SNA). Sometimes described as "public assistance" or "welfare," CA provides unrestricted monetary transfers to poor individuals and families. As such, CA can be thought of as the present-day version of the classic poverty alleviation program. Since welfare reform of the 1990s, work requirements have been the centerpiece of CA. In order to maintain eligibility, able recipients must be engaged in 30 hours of employment activities per week, which can include such things as training, education, and job search. (In practice, exemptions are common and sanctions may be unevenly enforced.) Eligibility for CA is limited to the very poorest. In New York, maximum monthly income at initial eligibility is $879 per month for a family of three. Benefits are similarly tight, topping out at $789 a month for a three-person family. Together, these strict requirements, and well as the need for periodic recertification, means benefits can frequently lapse as families are sanctioned. 358,000 New York City residents were actively receiving CA as of August 2017 (Cohen and Giannarelli, 2016; New York State Office of Temporary and Disability Assistance, 2016, 2015, 2017; NYC Human Resources Administration, 2019*b*).

Food Stamps (FS), officially known as SNAP, provides low-income families with categorical dollars each month that must be spent on food. Its eligibility standards are less strict than CA; correspondingly, its caseloads are much larger. Income is the primary criterion; as of 2017, a family of three with earned income could quality so long as household income was $30,636 or less ($2,500/month). (For such families without earnings, the eligibility standard was $26,556.) While some able-bodied adults without dependents may be required to work, such requirements are typically mild and unevenly enforced. Benefits, like eligibility, is based on a formula determined by family size. In 2017, a family of three receives $504 monthly. 1.7 million NYC residents received SNAP as of August 2017 (New York State Office of Temporary and Disability Assistance, 2019; NYC Human Resources Administration,

[19]In fact, the relationship between DHS and HRA is complicated and dynamic, largely for reasons having to do with the challenges of family homelessness. DHS was originally part of HRA, until it was spun off as an independent agency in 1993. However, in 2016, Mayor de Blasio again consolidated DHS and HRA under the DSS umbrella, managed by a single commissioner, Steve Banks. Nevertheless, it remains conventional to refer to the departments as distinct. See NYC Department of Homeless Services (2019*a*) for more detail.

2019*b*).

C Supplementary Analysis

C.1 Subsidies and Length of Stay

My main empirical result is that in-borough families remain in shelter significantly longer than those placed out-of-borough. They are also more likely to exit shelter with rental assistance. I interpret these facts through the lens of my search effort model. Families prefer local placements so they stay longer and either (a) need increased incentive to leave, or (b) are willing to tolerate longer stays to access subsidies[20].

One concern with this interpretation is endogenous subsidy allocation: City subsidy policy could be driving length of stay. If, for example, the City prioritized out-of-borough families for rental assistance, longer stays for in-borough families would be an artifact of subsidy queuing. In this section, I provide evidence that this is not the case.

The first observation is that, per my OLS results with subsidy as the dependent variable, in-borough families are about 5 percent more likely to exit with a subsidy. So, a stylized fact is that the City is more apt to allocate subsidies to in-borough families.

Question two is the effect. Naively, the effect of subsidies on LOS is ambiguous. If subsides hasten exits, LOS is reduced; if families remain in shelter longer waiting for subsidies, LOS would increase. In the former case (subsides shorten stays), any subsidy-based endogeneity biases the LOS effect downward, since in-borough families are receiving more of them. In the latter case (subsides lengthen stays), the endogeneity concern would necessarily be that the City is reserving more subsidies for in-borough families, which may be the case, given in-borough families' greater likelihood of subsidy receipt. But this would seem unlikely, since out-of-borough families are theoretically facing a harder time in shelter.

A third—perhaps most realistic—possibility is that the City simply forces in-borough families to wait longer for subsidies than out-of-borough ones. But, under this hypothesis, it would seem odd that in-borough families are more likely to receive subsidies. If out-of-borough families get priority for subsides—and theoretically have greater incentive to use them—why are they less likely to exit shelter with subsidies?

Tables A.6 and A.7 provide evidence assessing these possibilities. The former is for OLS; the latter for IV. They follow the same setup. There are 6 columns. The first four consider the full 2010–2016 period; the last two are limited to the April 2011 to December 2013

[20] I define subsidies broadly, as including Advantage, LINC, NYCHA public housing, Section 8, FEPS, rental assistance one-shots, and similar programs. The availability of these subsidies vary widely over time, given frequent policy changes.

period, when subsidies for homeless families were quite scarce (and thus serve as a test of length of stay in a "no subsidy" environment). Col 1 repeats my main analysis. Col 2 is limited to unsubsidized exits. Col 3 is limited to subsidized exits. Col 4 includes an interaction between subsidized exits and treatment (as well as the main effect). Col 5 is limited to families entering shelter between 4/2011 and 12/2013. Col 6 is limited to families both entering and exiting within that period.

In brief, the findings are that the effect of in-borough placement on length of stay is, if anything, strengthened when accounting for subsidies. The Col 5 results for both OLS and IV are larger than their Col 1 counterparts. The Col 6 result is smaller given the exclusion long stayers, but the LOS effect is still significant (in any case, selecting a sample in this fashion is problematic). Further, Cols 2–4 show that basically all of the effect of in-borough placement is on unsubsidized families. To be precise, unsubsidized in-borough families stay longer than unsubsidized out-of-borough ones, but subsidized in-borough families stay about the same as, or shorter than, subsidized out-of-borough ones. Put slightly differently, subsidy increases LOS, but less for in-borough families.

D References

Borg, Andreas, and Murat Sariyar. 2016. "Recordlinkage: Record Linkage in R." R package version 0.4-10.

Cassidy, Michael. 2020. "A Closer Look: Proximity Boosts Homeless Student Performance in New York City." *IZA Discussion Papers*.

Cohen, Elissa, Sarah Minton Megan Thompson Elizabeth Crowe, and Linda Giannarelli. 2016. "Welfare Rules Databook: State TANF Policies as of July 2015, OPRE Report 2016-67."

Culhane, Dennis P., Stephen Metraux, Jung Min Park, Maryanne Schretzman, and Jesse Valente. 2007. "Testing a Typology of Family Homelessness Based on Patterns of Public Shelter Utilization in Four U.s. Jurisdictions: Implications for Policy and Program Planning." *Housing Policy Debate*, 18(1): 1–28.

Curtis, Marah A, Hope Corman, Kelly Noonan, and Nancy E Reichman. 2013. "Life Shocks and Homelessness." *Demography*, 50(6): 2227–2253.

Desmond, Matthew. 2016. *Evicted: Poverty and Profit in the American City.* Broadway Books.

Fellegi, Ivan P, and Alan B Sunter. 1969. "A Theory for Record Linkage." *Journal of the American Statistical Association*, 64(328): 1183–1210.

Kaufman, Leslie, and David Chen. 2008. "City Reaches Deal on Shelter for Homeless." *New York Times*.

Khadduri, Jill, and Dennis Culhane. 2016. "The 2016 Annual Homeless Assessment Report (AHAR) to Congress."

Mullainathan, Sendhil, and Eldar Shafir. 2013. *Scarcity: Why Having Too Little Means so Much.* Macmillan.

National Alliance to End Homelessness. 2016. "The State of Homelessness in America 2016."

New York State Office of Temporary and Disability Assistance. 2015. "New York State Plan and Executive Certification: Administration of the Block Grant for Temporary Assistance for Needy Families, 2015–2017."

New York State Office of Temporary and Disability Assistance. 2016. "Temporary Assistance Source Book."

New York State Office of Temporary and Disability Assistance. 2017. "Temporary Assistance."

New York State Office of Temporary and Disability Assistance. 2019. "Supplemental Nutrition Assistance Program (SNAP)."

NYC Department of City Planning. 2017. "Geosupport System: User Programming Guide, Software Version 17.1."

NYC Department of Homeless Services. 2019*a*. "About DHS."

NYC Department of Homeless Services. 2019*b*. "Daily Report, October 21, 2019."

NYC Department of Homeless Services. 2019*c*. "Families with Children."

NYC Department of Homeless Services. 2019*d*. "Rental Assistance."

NYC Department of Homeless Services. 2019*e*. "Stats and Reports."

NYC Human Resources Administration. 2019*a*. "CITYFEPS."

NYC Human Resources Administration. 2019*b*. "HRA Monthly Fact Sheet, August 2019."

NYC Independent Budget Office. 2011. "Analysis of the Mayor's Preliminary Budget for 2012."

NYC Independent Budget Office. 2014. "The Rising Number of Homeless Families in NYC, 2002–2012: A Look at Why Families Were Granted Shelter, the Housing They Had Lived in and Where They Came From."

NYC Mayor's Office. 2017. "Turning the Tide on Homeless in New York City."

NYC Mayor's Office of Operations. 2002. "Mayor's Management Report, Fiscal 2002."

NYC Mayor's Office of Operations. 2012. "Mayor's Management Report, September 2012."

NYC Mayor's Office of Operations. 2017. "Mayor's Management Report, September 2017."

NYU Furman Center. 2016. "State of New York City's Housing and Neighborhoods in 2016."

O'Flaherty, Brendan. 2010. "Homelessness As Bad Luck: Implications for Research and Policy." *How to House the Homeless.new York: Russell Sage Foundation*, 143–182.

Sariyar, Murat, and Andreas Borg. 2010. "The Recordlinkage Package: Detecting Errors in Data." *R Journal*, 2(2).

Shinn, M., B. C. Weitzman, D. Stojanovic, J. R. Knickman, L. Jimenez, L. Duchon, S. James, and D. H. Krantz. 1998. "Predictors of homelessness among families in New York City: from shelter request to housing stability." *American Journal of Public Health*, 88(11): 1651–1657.

Stewart, Nikita. 2017. "Fight Looms as Bill de Blasio Plans to Seek 90 New Homeless Shelters." *New York Times.*

University of Michigan Law School. 2017. "Case Profile: McCain v. Koch."

van der Loo, M.P.J. 2014. "The stringdist Package for Approximate String Matching." *The R Journal*, 6: 111–122.

Wasi, Nada, Aaron Flaaen, et al. 2015. "Record Linkage Using Stata: Pre-processing, Linking and Reviewing Utilities." *Stata Journal*, 15(3): 672–697.

E Supplementary Tables

Table A.1: Summary of Key Variables by Shelter Entry Year

	Year of Shelter Entry							
	2010	2011	2012	2013	2014	2015	2016	Total
A. Shelter Entry Characteristics								
Families Entering	9,911	7,475	7,937	7,642	8,752	8,161	9,375	59,253
Individuals Entering	31,789	25,219	27,873	26,619	29,610	27,264	30,370	198,744
Borough Placement	0.66	0.59	0.51	0.51	0.44	0.46	0.38	0.51
Placement Distance (miles)	4.68	5.21	5.88	5.80	6.43	6.37	6.91	5.89
Ineligibility Rate	0.25	0.23	0.24	0.21	0.17	0.28	0.26	0.23
Aversion Ratio	1.53	1.17	1.09	1.05	0.75	1.58	1.32	1.22
Occupancy Rate	0.89	0.90	0.95	0.96	0.96	0.97	0.96	0.94
B. Stays and Returns								
Length of Stay	365.1	441.2	451.9	436.2	436.3	438.1	417.3	424.3
Subsidized Exit	0.34	0.14	0.26	0.39	0.50	0.56	0.53	0.39
Returned to Shelter	0.12	0.15	0.17	0.16	0.13	0.15	0.20	0.15
C. Year Post-Shelter Entry								
Cash Assistance	0.77	0.80	0.81	0.79	0.79	0.81	0.72	0.78
Food Stamps	0.91	0.90	0.91	0.91	0.91	0.89	0.85	0.90
Employed	0.47	0.44	0.44	0.46	0.52	0.53	0.48	0.48
Avg. Quarterly Earnings	1094.6	1015.1	958.1	1045.2	1232.8	1416.2	1500.6	1188.9
D. Year Post-Shelter Exit								
Cash Assistance	0.72	0.72	0.74	0.74	0.77	0.77	0.68	0.74
Food Stamps	0.89	0.88	0.89	0.89	0.89	0.88	0.85	0.88
Employed	0.44	0.43	0.44	0.47	0.49	0.49	0.40	0.45
Avg. Quarterly Earnings	1219.9	1169.8	1175.5	1322.3	1476.9	1550.0	1342.3	1306.2
E. Censoring								
Family Spell	0.00	0.00	0.00	0.01	0.02	0.04	0.07	0.02
Full Year Post-Spell	0.00	0.00	0.01	0.02	0.04	0.08	0.21	0.05
CA/FS Year Post-Entry	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.16
CA/FS Year Post-Exit	0.01	0.03	0.07	0.13	0.31	0.72	1.00	0.34
Labor Year Post-Exit	0.01	0.02	0.06	0.11	0.24	0.61	0.98	0.30

Includes only family shelter entrants originating from NYC. Unit of observation is family-spell. Families and individual entering are counts; all other statistics are family-spell means.

Table A.2: Families by Number of Spells

Homeless Spells	# of Families	Percent
1	37,587	78.3
2	8,015	16.7
3	1,831	3.8
4+	544	1.1
Total	47,977	100.0

Includes only family shelter entrants originating from NYC.

Table A.3A: Descriptives and Random Assignment

	Overall			Randomization Check		
Variable	N	Mean	SD	Out-of-Boro	In-Boro	Diff.
Year Entered Shelter	59,253	2013.01	2.07	2013.38	2012.65	-0.72**
Month Entered Shelter	59,253	6.52	3.40	6.78	6.28	-0.50**
Q1 Entry	59,253	0.25	0.43	0.22	0.27	0.05**
Q2 Entry	59,253	0.23	0.42	0.22	0.25	0.03**
Q3 Entry	59,253	0.28	0.45	0.31	0.26	-0.05**
Q4 Entry	59,253	0.24	0.42	0.25	0.22	-0.03**
Manhattan Origin	59,253	0.12	0.33	0.16	0.09	-0.07**
Bronx Origin	59,253	0.41	0.49	0.33	0.49	0.16**
Brooklyn Origin	59,253	0.32	0.47	0.31	0.32	0.01**
Queens Origin	59,253	0.12	0.33	0.15	0.10	-0.06**
Staten Island Origin	59,253	0.03	0.16	0.05	0.01	-0.04**
Family Size	59,253	3.35	1.39	3.34	3.36	0.02*
Family Members Under 18	59,253	1.97	1.19	1.95	1.99	0.04**
Oldest Child's Grade	59,253	2.57	5.32	1.95	3.18	1.23**
Health Issue Present	59,253	0.30	0.46	0.32	0.28	-0.04**
Eligibility: Eviction	59,253	0.33	0.47	0.28	0.39	0.10**
Eligibility: Overcrowding	59,253	0.18	0.38	0.17	0.19	0.02**
Eligibility: Conditions	59,253	0.08	0.28	0.08	0.09	0.01**
Eligibility: Domestic Violence	59,253	0.30	0.46	0.37	0.22	-0.15**
Eligibility: Other	59,253	0.11	0.31	0.10	0.11	0.01**
Eligibility: Unknown	59,253	0.00	0.01	0.00	0.00	0.00
Female	59,253	0.92	0.28	0.92	0.91	-0.01**
Age	59,253	31.54	8.86	30.94	32.13	1.20**
Partner/Spouse Present	59,253	0.26	0.44	0.27	0.24	-0.03**
Pregnant	59,253	0.07	0.25	0.07	0.06	-0.01**

Treatment defined as placed in-borough. Group contrasts obtained from separate bivariate OLS regressions of each characteristic on treatment indicator. Differences between in-borough and out-of-borough means are coefficients on treatment indicator. Standard errors clustered at the family group level. Unit of observation is family-spell. Full sample. * $p < 0.10$, ** $p < 0.05$

Table A.3B: Descriptives and Random Assignment

Variable	Overall N	Overall Mean	Overall SD	Randomization Check Out-of-Boro	Randomization Check In-Boro	Randomization Check Diff.
Black	59,253	0.56	0.50	0.57	0.55	-0.02**
White	59,253	0.03	0.16	0.03	0.02	-0.01**
Hispanic	59,253	0.38	0.48	0.36	0.39	0.03**
Asian	59,253	0.00	0.07	0.00	0.00	-0.00
Other Race	59,253	0.00	0.06	0.00	0.00	-0.00
Unknown Race	59,253	0.03	0.17	0.03	0.03	-0.00*
No Degree	59,253	0.57	0.50	0.56	0.58	0.01**
High School Grad	59,253	0.32	0.47	0.32	0.32	-0.01*
Some College or More	59,253	0.05	0.22	0.05	0.05	-0.00
Unknown Education	59,253	0.06	0.24	0.06	0.06	-0.00
On Cash Assistance	59,253	0.35	0.48	0.36	0.35	-0.01**
On Food Stamps	59,253	0.73	0.44	0.73	0.73	0.00
Employed Year Pre	59,253	0.43	0.50	0.44	0.43	-0.01**
Log AQ Earnings Year Pre	59,253	3.01	3.58	3.02	2.99	-0.03
Tier II Shelter	59,253	0.55	0.50	0.55	0.55	0.01**
Commercial Hotel	59,253	0.28	0.45	0.30	0.25	-0.05**
Family Cluster Unit	59,253	0.16	0.37	0.14	0.19	0.05**
Other Facility	59,253	0.01	0.10	0.01	0.01	-0.01**
Manhattan Shelter	59,253	0.18	0.39	0.27	0.09	-0.18**
Bronx Shelter	59,253	0.39	0.49	0.29	0.49	0.20**
Brooklyn Shelter	59,253	0.27	0.44	0.22	0.32	0.11**
Queens Shelter	59,253	0.15	0.36	0.21	0.10	-0.11**
Staten Island Shelter	59,253	0.01	0.09	0.01	0.01	-0.01**
School District Placement	54,306	0.10	0.30	0.00	0.19	0.19**
Placement Distance (miles)	54,306	5.89	4.65	9.27	2.66	-6.61**
Borough Placement	59,253	0.51	0.50	0.00	1.00	1.00

Treatment defined as placed in-borough. Group contrasts obtained from separate bivariate OLS regressions of each characteristic on treatment indicator. Differences between in-borough and out-of-borough means are coefficients on treatment indicator. Standard errors clustered at the family group level. Unit of observation is family-spell. Full sample. * $p < 0.10$, ** $p < 0.05$

Table A.4: Descriptives and Random Assignment

	Overall			Randomization Check		
Variable	N	Mean	SD	Out-of-Boro	In-Boro	Diff.
Jan Entry	59,253	0.09	0.29	0.08	0.10	0.01**
Feb Entry	59,253	0.08	0.26	0.07	0.08	0.02**
Mar Entry	59,253	0.08	0.27	0.07	0.09	0.02**
Apr Entry	59,253	0.08	0.27	0.07	0.08	0.02**
May Entry	59,253	0.08	0.27	0.07	0.08	0.01**
Jun Entry	59,253	0.08	0.27	0.08	0.08	0.01**
Jul Entry	59,253	0.09	0.28	0.10	0.08	-0.02**
Aug Entry	59,253	0.10	0.30	0.11	0.09	-0.02**
Sep Entry	59,253	0.10	0.30	0.10	0.09	-0.02**
Oct Entry	59,253	0.09	0.28	0.09	0.08	-0.01**
Nov Entry	59,253	0.08	0.27	0.08	0.07	-0.01**
Dec Entry	59,253	0.07	0.26	0.08	0.07	-0.01**
2010 Entry	59,253	0.17	0.37	0.12	0.22	0.10**
2011 Entry	59,253	0.13	0.33	0.10	0.15	0.04**
2012 Entry	59,253	0.13	0.34	0.13	0.14	0.00

Treatment defined as placed in-borough. Group contrasts obtained from separate bivariate OLS regressions of each characteristic on treatment indicator. Differences between in-borough and out-of-borough means are coefficients on treatment indicator. Standard errors clustered at the family group level. Unit of observation is family-spell. Sample is all NYC family shelter entrants from 2010–2016 with non-missing origin and shelter boroughs. * $p < 0.10$, ** $p < 0.05$

Table A.5: Descriptives and Random Assignment

	Overall			Randomization Check		
Variable	N	Mean	SD	Out-of-Boro	In-Boro	Diff.
Log Length of Stay	59,253	5.50	1.24	5.43	5.57	0.14**
Log Shelter LOS (Excl. Gaps)	59,253	5.50	1.24	5.42	5.57	0.14**
Length of Stay (Days)	59,253	424.33	406.67	410.96	437.35	26.40**
Log LOS (2017)	59,253	5.48	1.21	5.40	5.55	0.15**
Subsidized Exit	57,962	0.39	0.49	0.39	0.40	0.01*
Unsubsidized Exit	57,962	0.60	0.49	0.60	0.60	-0.00
Returned to Shelter (One Year)	52,274	0.15	0.36	0.16	0.14	-0.03**
Cash Assistance Post Entry	59,253	0.78	0.41	0.77	0.79	0.02**
CA Post Entry Percent	59,253	0.62	0.41	0.61	0.64	0.02**
Food Stamps Post Entry	59,253	0.90	0.31	0.89	0.90	0.01**
FS Post Entry Percent	59,253	0.82	0.34	0.80	0.83	0.03**
Employed Post Entry	59,253	0.48	0.50	0.48	0.48	0.01
Empl. Post Entry Percent	59,253	0.34	0.41	0.33	0.34	0.01**
Log AQ Earnings Post Entry	59,253	3.38	3.68	3.33	3.42	0.09**
AQ Earnings Post Entry	59,253	1188.87	2274.61	1153.03	1223.76	70.73**
Cash Assistance Post Exit	48,082	0.74	0.44	0.73	0.74	0.01**
CA Post Exit Percent	59,253	0.41	0.44	0.39	0.42	0.03**
Food Stamps Post Exit	48,082	0.88	0.32	0.88	0.89	0.01**
FS Post Exit Percent	59,253	0.60	0.45	0.57	0.62	0.05**
Employed Post Exit	48,082	0.45	0.50	0.45	0.46	0.01
Empl. Post Exit Percent	59,253	0.27	0.40	0.26	0.29	0.03**
Log AQ Earnings Post Exit	48,082	3.27	3.73	3.22	3.31	0.09**
AQ Earnings Post Exit	48,082	1306.24	2515.85	1247.82	1358.75	110.93**

Treatment defined as placed in-borough. Group contrasts obtained from separate bivariate OLS regressions of each characteristic on treatment indicator. Differences between in-borough and out-of-borough means are coefficients on treatment indicator. Standard errors clustered at the family group level. Unit of observation is family-spell. Sample is all NYC family shelter entrants from 2010–2016 with non-missing origin and shelter boroughs. * $p < 0.10$, ** $p < 0.05$

Table A.6: Subsidized Exits and Length of Stay: OLS Results

	Full Period (2010-2016)				Apr. 2011-Dec. 2013	
	All Families (1)	Unsubs-idized (2)	Subsid-ized (3)	Interac-tion (4)	Entry (5)	Entry & Exit (6)
Borough Placement	0.120**	0.152**	0.020*	0.218**	0.132**	0.066**
	(0.011)	(0.014)	(0.011)	(0.014)	(0.020)	(0.020)
Subsidized Exit				1.024**		
				(0.013)		
Borough Placement × Subsidized Exit				-0.282**		
				(0.017)		
Obs.	59,253	35,260	22,702	57,962	20,918	4,852

Each column is a separate regression of log length of stay on an indicator for in-borough placement, Main covariates, and addressing subsidized exits in the column-enumerated manner. Subsidized exits include any sort of rental assistance (e.g., Advantage, LINC, NYCHA, Section 8, FEPS, one-shots). Base sample is Full sample. Col 1 repeats results from main text. Col 2 is limited to families with unsubsidized exits only. Col 3 is limited to families with subsidized exits only. Col 4 includes an indicator for subsidized exit and its interaction with treatment. Col 5 is limited to families entering shelter 4/1/2011–12/31/2013.Col 6 is limited to families entering and exiting shelter 4/1/2011–12/31/2013. Standard errors clustered at family group level in parentheses. * $p < 0.10$, ** $p < 0.05$

Table A.7: Subsidized Exits and Length of Stay: Aversion Ratio IV Results

	Full Period (2010-2016)				Apr. 2011-Dec. 2013	
	All Families (1)	Unsubs-idized (2)	Subsid-ized (3)	Interac-tion (4)	Entry (5)	Entry & Exit (6)
Borough Placement	0.95**	2.63**	0.59	6.02**	1.38**	-10.07
	(0.34)	(0.64)	(0.37)	(0.93)	(0.71)	(6.63)
Subsidized Exit				7.13**		
				(0.69)		
Borough Placement $\times$ Subsidized Exit				-12.32**		
				(1.35)		
Obs.	59,253	35,260	22,702	57,962	20,918	4,852

Each column is a separate regression of log length of stay on an indicator for in-borough placement, Main covariates, and addressing subsidized exits in the column-enumerated manner. Subsidized exits include any sort of rental assistance (e.g., Advantage, LINC, NYCHA, Section 8, FEPS, one-shots). Base sample is Full sample. Col 1 repeats results from main text. Col 2 is limited to families with unsubsidized exits only. Col 3 is limited to families with subsidized exits only. Col 4 includes an indicator for subsidized exit and its interaction with treatment. Col 5 is limited to families entering shelter 4/1/2011–12/31/2013.Col 6 is limited to families entering and exiting shelter 4/1/2011–12/31/2013. Standard errors clustered at family group level in parentheses. * $p < 0.10$, ** $p < 0.05$

Table A.8: OLS Outcome Robustness

	Full Sample					Non-DV	Pre-2015
Outcome	Outcome Mean (1)	Raw (2)	Placement (3)	Main (4)	Shelter (5)	Main (6)	Main (7)
A. Stays and Returns							
Log LOS (excl. gaps)	5.496** (1.243) {59,253}	0.141** (0.010) {59,253}	0.108** (0.010) {59,253}	0.120** (0.011) {59,253}	0.115** (0.011) {59,247}	0.086** (0.012) {41,744}	0.126** (0.013) {41,717}
Length of Stay (days)	424.333** (406.668) {59,253}	26.397** (3.334) {59,253}	17.587** (3.417) {59,253}	23.090** (3.544) {59,253}	22.341** (3.541) {59,247}	19.767** (4.430) {41,744}	24.741** (4.446) {41,717}
Log LOS (as of 2017)	5.476** (1.215) {59,253}	0.146** (0.010) {59,253}	0.105** (0.010) {59,253}	0.117** (0.011) {59,253}	0.113** (0.011) {59,247}	0.083** (0.011) {41,744}	0.125** (0.013) {41,717}
Unsubsidized Exit	0.600** (0.490) {57,962}	-0.004 (0.004) {57,962}	-0.020** (0.004) {57,962}	-0.017** (0.004) {57,962}	-0.016** (0.004) {57,954}	-0.016** (0.005) {40,766}	-0.015** (0.005) {41,420}
B. Year Post-Shelter Entry							
CA Percent of Year	0.624** (0.410) {59,253}	0.024** (0.003) {59,253}	0.009** (0.004) {59,253}	0.004 (0.003) {59,253}	0.004 (0.003) {59,247}	0.006 (0.004) {41,744}	0.006 (0.004) {41,717}
FS Percent of Year	0.815** (0.342) {59,253}	0.032** (0.003) {59,253}	0.004 (0.003) {59,253}	-0.001 (0.002) {59,253}	-0.000 (0.002) {59,247}	-0.006** (0.003) {41,744}	0.004 (0.002) {41,717}
Employed: Quarterly Proportion	0.337** (0.406) {59,253}	0.010** (0.003) {59,253}	0.014** (0.003) {59,253}	0.012** (0.003) {59,253}	0.011** (0.003) {59,247}	0.011** (0.004) {41,744}	0.011** (0.004) {41,717}
Avg. Quarterly Earnings	1188.870** (2274.606) {59,253}	70.734** (18.974) {59,253}	36.204* (19.470) {59,253}	27.902 (17.893) {59,253}	23.739 (18.000) {59,247}	29.523 (22.207) {41,744}	22.168 (20.448) {41,717}
C. Year Post-Shelter Exit							
CA Percent of Year	0.407** (0.435) {59,253}	0.027** (0.004) {59,253}	-0.002 (0.004) {59,253}	-0.006* (0.004) {59,253}	-0.005 (0.004) {59,247}	-0.002 (0.004) {41,744}	-0.001 (0.004) {41,717}
FS Percent of Year	0.596** (0.451) {59,253}	0.053** (0.004) {59,253}	-0.004 (0.003) {59,253}	-0.010** (0.003) {59,253}	-0.008** (0.003) {59,247}	-0.013** (0.004) {41,744}	-0.003 (0.004) {41,717}
Employed: Quarterly Proportion	0.270** (0.397) {59,253}	0.031** (0.003) {59,253}	0.007** (0.003) {59,253}	0.004 (0.003) {59,253}	0.004 (0.003) {59,247}	0.006 (0.004) {41,744}	0.005 (0.004) {41,717}
Avg. Quarterly Earnings	1306.237** (2515.846) {48,082}	110.929** (23.163) {48,082}	48.035** (24.293) {48,082}	33.994 (23.227) {48,082}	28.066 (23.470) {48,076}	38.382 (29.069) {33,761}	24.918 (25.269) {39,974}
Time Control		None	$Year^3$	$Year^3$	$Year^3$	$Year^3$	$Year^3$
Placement Controls		No	Yes	Yes	Yes	Yes	Yes
Family & Shelter Controls		No	No	Yes	Yes	Yes	Yes
Shelter FE		No	No	No	Yes	No	No

Each cell reports the coefficient on in-borough shelter placement from a separate OLS regression of the row-delineated outcome on the treatment indicator, controlling for the column-enumerated covariates. Supercolumns give samples. Standard errors clustered at family group level in parentheses. Number of observations given in braces. * $p < 0.10$, ** $p < 0.05$

Table A.9: Compliance Type Shares: Ineligibility Rate Instrument

	1%	1.5%	2%
Compliers	0.08	0.08	0.07
Always-Takers	0.64	0.64	0.64
Never-Takers	0.28	0.28	0.28

Main sample. Results from linear first-stage, controlling for year and month of shelter entry. Percentages in second row refer to percentiles used as thresholds to define low and high instrument values. See Cassidy (2020) for estimation method details.

Table A.10: Compliance Type Shares: Aversion Ratio

	1%	1.5%	2%
Compliers	0.10	0.10	0.09
Always-Takers	0.62	0.62	0.62
Never-Takers	0.28	0.28	0.28

Main sample. Results from linear first-stage, controlling for year and month of shelter entry. Percentages in second row refer to percentiles used as thresholds to define low and high instrument values. See Cassidy (2020) for estimation method details.

Table A.11: Complier Characteristics: Ineligibility Rate Instrument

	Compliers	Non-Compliers	Diff.
Manhattan Origin	0.00	0.14	-0.13
	(0.003)	(0.000)	[-2.56]
Bronx Origin	0.57	0.39	0.18
	(0.006)	(0.000)	[2.28]
Brooklyn Origin	0.25	0.32	-0.07
	(0.005)	(0.000)	[-0.99]
Queens Origin	0.10	0.13	-0.02
	(0.003)	(0.000)	[-0.45]
Staten Island Origin	0.02	0.03	-0.00
	(0.000)	(0.000)	[-0.33]
Health Issue Present	0.33	0.30	0.04
	(0.004)	(0.000)	[0.61]
Eligibility: Eviction	0.29	0.34	-0.05
	(0.005)	(0.000)	[-0.67]
Eligibility: Overcrowding	0.16	0.18	-0.02
	(0.003)	(0.000)	[-0.34]
Eligibility: Conditions	0.11	0.08	0.03
	(0.002)	(0.000)	[0.67]
Eligibility: Domestic Violence	0.30	0.30	0.00
	(0.004)	(0.000)	[0.01]
Eligibility: Other	0.08	0.11	-0.03
	(0.003)	(0.000)	[-0.67]
Female	0.97	0.91	0.06
	(0.002)	(0.000)	[1.26]
Partner/Spouse Present	0.31	0.25	0.06
	(0.004)	(0.000)	[0.99]
Pregnant	0.04	0.07	-0.03
	(0.001)	(0.000)	[-0.86]
Black	0.43	0.57	-0.14
	(0.006)	(0.000)	[-1.79]
Hispanic	0.46	0.37	0.09
	(0.006)	(0.000)	[1.18]
White	0.06	0.02	0.04
	(0.001)	(0.000)	[1.57]
No Degree	0.61	0.57	0.05
	(0.005)	(0.000)	[0.67]
High School Grad	0.30	0.32	-0.02
	(0.005)	(0.000)	[-0.29]
Some College or More	0.06	0.05	0.01
	(0.001)	(0.000)	[0.21]
Unknown Education	0.02	0.07	-0.05
	(0.001)	(0.000)	[-1.27]
On Cash Assistance	0.30	0.36	-0.06
	(0.005)	(0.000)	[-0.78]
On Food Stamps	0.75	0.73	0.02
	(0.006)	(0.000)	[0.29]
Employed Year Pre	0.39	0.44	-0.05
	(0.005)	(0.000)	[-0.67]
Tier II Shelter	0.63	0.54	0.08
	(0.004)	(0.000)	[1.27]
Commercial Hotel	0.08	0.29	-0.21
	(0.005)	(0.000)	[-3.08]
Family Cluster Unit	0.19	0.16	0.02
	(0.003)	(0.000)	[0.46]
Family Size 1–3	0.54	0.64	-0.11
	(0.006)	(0.000)	[-1.43]
Family Size 4–5	0.39	0.28	0.11
	(0.004)	(0.000)	[1.68]
Family Size 6+	0.07	0.08	-0.01
	(0.002)	(0.000)	[-0.27]
Age	31.69	31.53	0.16
	(1.350)	(0.013)	[0.14]
Log AQ Earnings Year Pre	2.73	3.03	-0.30
	(0.259)	(0.002)	[-0.60]

Main sample. Treatment is in-borough placement. Instrument is 15-day moving average of the initial ineligibility rate for 30-day application period. Compliers are families placed in-borough when the ineligibility rate is high, but not otherwise. Non-compliers consist of always-takers and never-takers. Compiler and non-complier characteristics, adjusted for year and month of shelter entry, are estimated from the algorithm described in Cassidy (2020). Standard errors (in parentheses) and differences in means (with t-stats in brackets) are calculated from 200 bootstrap replications, clustering by family group.

Table A.12: Complier Characteristics: Aversion Ratio Instrument

	Compliers	Non-Compliers	Diff.
Manhattan Origin	0.09	0.13	-0.04
	(0.001)	(0.000)	[-1.18]
Bronx Origin	0.55	0.39	0.16
	(0.004)	(0.000)	[2.67]
Brooklyn Origin	0.20	0.33	-0.13
	(0.003)	(0.000)	[-2.29]
Queens Origin	0.12	0.13	-0.01
	(0.002)	(0.000)	[-0.17]
Staten Island Origin	0.03	0.03	0.01
	(0.000)	(0.000)	[0.51]
Health Issue Present	0.35	0.29	0.06
	(0.002)	(0.000)	[1.27]
Eligibility: Eviction	0.33	0.34	-0.01
	(0.003)	(0.000)	[-0.13]
Eligibility: Overcrowding	0.15	0.18	-0.03
	(0.002)	(0.000)	[-0.61]
Eligibility: Conditions	0.11	0.08	0.02
	(0.001)	(0.000)	[0.73]
Eligibility: Domestic Violence	0.27	0.30	-0.03
	(0.002)	(0.000)	[-0.59]
Eligibility: Other	0.10	0.11	-0.01
	(0.001)	(0.000)	[-0.28]
Female	0.94	0.91	0.03
	(0.001)	(0.000)	[0.82]
Partner/Spouse Present	0.28	0.25	0.03
	(0.002)	(0.000)	[0.61]
Pregnant	0.05	0.07	-0.02
	(0.001)	(0.000)	[-0.66]
Black	0.47	0.57	-0.09
	(0.004)	(0.000)	[-1.51]
Hispanic	0.43	0.37	0.06
	(0.003)	(0.000)	[0.98]
White	0.05	0.02	0.03
	(0.000)	(0.000)	[1.46]
No Degree	0.59	0.57	0.02
	(0.003)	(0.000)	[0.41]
High School Grad	0.31	0.32	-0.01
	(0.002)	(0.000)	[-0.21]
Some College or More	0.07	0.05	0.02
	(0.001)	(0.000)	[0.76]
Unknown Education	0.02	0.07	-0.04
	(0.001)	(0.000)	[-1.47]
On Cash Assistance	0.23	0.37	-0.14
	(0.003)	(0.000)	[-2.43]
On Food Stamps	0.70	0.74	-0.04
	(0.003)	(0.000)	[-0.63]
Employed Year Pre	0.39	0.44	-0.05
	(0.003)	(0.000)	[-0.91]
Tier II Shelter	0.59	0.55	0.04
	(0.003)	(0.000)	[0.82]
Commercial Hotel	0.14	0.29	-0.16
	(0.003)	(0.000)	[-2.83]
Family Cluster Unit	0.19	0.16	0.02
	(0.001)	(0.000)	[0.63]
Family Size 1–3	0.60	0.64	-0.04
	(0.003)	(0.000)	[-0.76]
Family Size 4–5	0.35	0.28	0.06
	(0.003)	(0.000)	[1.26]
Family Size 6+	0.05	0.08	-0.02
	(0.001)	(0.000)	[-0.82]
Age	32.20	31.47	0.73
	(0.949)	(0.014)	[0.74]
Log AQ Earnings Year Pre	2.76	3.03	-0.27
	(0.147)	(0.002)	[-0.71]

Main sample. Treatment is in-borough placement. Instrument is 15-day moving average of the aversion ratio. Compliers are families placed in-borough when the aversion ratio is high, but not otherwise. Non-compliers consist of always-takers and never-takers. Compiler and non-complier characteristics, adjusted for year and month of shelter entry, are estimated from the algorithm described in Cassidy (2020). Standard errors (in parentheses) and differences in means (with t-stats in brackets) are calculated from 200 bootstrap replications, clustering by family group.

Table A.13: IV Robustness: Ineligibility Rate

	Borough			School District			Distance		
	Full	Non-DV	Pre-2015	Full	Non-DV	Pre-2015	Full	Non-DV	Pre-2015
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Stays and Returns									
Log Length of Stay	1.367**	0.971*	5.083*	4.650**	2.205*	7.198**	-0.203**	-0.166*	-0.497**
	(0.527)	(0.573)	(2.820)	(2.084)	(1.201)	(3.135)	(0.078)	(0.097)	(0.213)
	[28.8]	[18.3]	[4.0]	[8.7]	[11.7]	[6.9]	[17.1]	[8.6]	[7.2]
Subsidized Exit	-0.789**	-1.145**	-2.513*	-2.070**	-1.998**	-2.244**	0.094**	0.161**	0.173**
	(0.257)	(0.380)	(1.489)	(0.885)	(0.721)	(1.017)	(0.037)	(0.071)	(0.084)
	[26.2]	[16.8]	[3.3]	[9.2]	[12.7]	[7.2]	[15.6]	[7.7]	[6.0]
Returned to Shelter	0.287*	0.362*	-0.301	1.287	1.039	-0.201	-0.039	-0.066	0.017
	(0.166)	(0.210)	(0.405)	(0.879)	(0.676)	(0.397)	(0.025)	(0.047)	(0.032)
	[25.2]	[15.3]	[4.3]	[4.3]	[5.3]	[9.2]	[13.4]	[4.7]	[7.8]
B. Year Post-Shelter Entry									
Cash Assistance	0.651**	0.699**	0.356	1.498**	1.051**	0.688	-0.085**	-0.107**	-0.048
	(0.183)	(0.238)	(0.415)	(0.654)	(0.457)	(0.519)	(0.027)	(0.045)	(0.036)
	[28.8]	[18.3]	[4.0]	[8.7]	[11.7]	[6.9]	[17.1]	[8.6]	[7.2]
Food Stamps	-0.142	-0.199*	-0.107	-0.546*	-0.456*	-0.065	0.017	0.029	0.005
	(0.093)	(0.120)	(0.247)	(0.323)	(0.251)	(0.279)	(0.013)	(0.020)	(0.020)
	[28.8]	[18.3]	[4.0]	[8.7]	[11.7]	[6.9]	[17.1]	[8.6]	[7.2]
Employed	-0.020	0.012	0.037	-0.161	-0.050	0.226	0.001	-0.001	-0.015
	(0.171)	(0.213)	(0.474)	(0.511)	(0.410)	(0.560)	(0.023)	(0.032)	(0.040)
	[28.8]	[18.3]	[4.0]	[8.7]	[11.7]	[6.9]	[17.1]	[8.6]	[7.2]
Log Avg. Quarterly Earnings	1.245	1.020	0.606	2.445	1.437	1.817	-0.155	-0.149	-0.118
	(1.243)	(1.553)	(3.354)	(3.747)	(2.997)	(3.973)	(0.169)	(0.240)	(0.280)
	[28.8]	[18.3]	[4.0]	[8.7]	[11.7]	[6.9]	[17.1]	[8.6]	[7.2]
C. Year Post-Shelter Exit									
Cash Assistance	0.428**	0.227	0.283	1.289*	0.487	0.598	-0.059**	-0.048	-0.045
	(0.210)	(0.236)	(0.405)	(0.763)	(0.529)	(0.482)	(0.030)	(0.047)	(0.036)
	[20.3]	[14.1]	[5.2]	[6.2]	[7.1]	[9.3]	[12.5]	[4.5]	[8.9]
Food Stamps	-0.064	-0.187	-0.197	-0.102	-0.416	-0.122	0.001	0.031	0.008
	(0.130)	(0.165)	(0.271)	(0.381)	(0.372)	(0.293)	(0.017)	(0.032)	(0.022)
	[20.3]	[14.1]	[5.2]	[6.2]	[7.1]	[9.3]	[12.5]	[4.5]	[8.9]
Employed	0.397*	0.405	0.448	1.422*	1.002	0.632	-0.066*	-0.091	-0.047
	(0.232)	(0.279)	(0.479)	(0.862)	(0.675)	(0.537)	(0.034)	(0.064)	(0.040)
	[20.3]	[14.1]	[5.2]	[6.2]	[7.1]	[9.3]	[12.5]	[4.5]	[8.9]
Log Avg. Quarterly Earnings	2.508	2.133	1.998	9.156	5.547	2.754	-0.419*	-0.509	-0.206
	(1.673)	(1.988)	(3.298)	(5.991)	(4.640)	(3.726)	(0.240)	(0.427)	(0.277)
	[20.3]	[14.1]	[5.2]	[6.2]	[7.1]	[9.3]	[12.5]	[4.5]	[8.9]
Time Control	Year^3	Year^3	Year^3	Year^3	Year^3	Year^3	Year^3	Year^3	Year^3
Placement Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Family & Shelter Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Shelter FE	No	No	No	No	No	No	No	No	No

Each cell reports the coefficient on local shelter placement from a separate 2SLS regression of the row-delineated outcome on the treatment using the ineligibility rate as the instrument and controlling for Main covariates. Columns give samples; supercolumns give treatment definitions. Standard errors clustered at family group level in parentheses. First-stage F-stats in brackets. * $p < 0.10$, ** $p < 0.05$

Table A.14: IV Robustness: Aversion Ratio

	Borough			School District			Distance		
	Full	Non-DV	Pre-2015	Full	Non-DV	Pre-2015	Full	Non-DV	Pre-2015
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A. Stays and Returns									
Log Length of Stay	0.946**	0.531	2.110**	2.930**	1.275	4.479**	-0.120**	-0.071	-0.205**
	(0.342)	(0.357)	(0.634)	(1.109)	(0.806)	(1.547)	(0.043)	(0.048)	(0.058)
	[60.8]	[42.9]	[27.5]	[20.3]	[21.4]	[14.6]	[45.2]	[26.7]	[38.1]
Subsidized Exit	-0.331**	-0.527**	-0.483**	-0.874**	-1.078**	-0.836*	0.034*	0.061**	0.035*
	(0.147)	(0.190)	(0.219)	(0.427)	(0.416)	(0.445)	(0.018)	(0.026)	(0.020)
	[55.8]	[38.8]	[26.3]	[20.3]	[22.1]	[14.7]	[41.8]	[23.8]	[36.4]
Returned to Shelter	0.088	0.098	-0.337**	0.276	0.240	-0.702**	-0.006	-0.009	0.038**
	(0.104)	(0.122)	(0.162)	(0.372)	(0.345)	(0.330)	(0.013)	(0.019)	(0.016)
	[55.7]	[37.7]	[27.1]	[12.2]	[11.7]	[16.2]	[37.2]	[17.4]	[37.8]
B. Year Post-Shelter Entry									
Cash Assistance	0.338**	0.318**	0.015	0.600**	0.384	0.123	-0.041**	-0.042**	-0.006
	(0.105)	(0.125)	(0.149)	(0.304)	(0.268)	(0.304)	(0.013)	(0.017)	(0.014)
	[60.8]	[42.9]	[27.5]	[20.3]	[21.4]	[14.6]	[45.2]	[26.7]	[38.1]
Food Stamps	-0.100	-0.133*	-0.028	-0.374*	-0.336*	0.071	0.009	0.015	-0.003
	(0.064)	(0.076)	(0.095)	(0.195)	(0.176)	(0.189)	(0.008)	(0.010)	(0.009)
	[60.8]	[42.9]	[27.5]	[20.3]	[21.4]	[14.6]	[45.2]	[26.7]	[38.1]
Employed	0.116	0.164	0.284	0.190	0.279	0.576	-0.013	-0.021	-0.027
	(0.118)	(0.141)	(0.190)	(0.334)	(0.308)	(0.396)	(0.014)	(0.019)	(0.018)
	[60.8]	[42.9]	[27.5]	[20.3]	[21.4]	[14.6]	[45.2]	[26.7]	[38.1]
Log Avg. Quarterly Earnings	1.085	1.258	1.101	2.085	2.354	2.541	-0.126	-0.172	-0.121
	(0.851)	(1.019)	(1.310)	(2.424)	(2.245)	(2.686)	(0.104)	(0.138)	(0.126)
	[60.8]	[42.9]	[27.5]	[20.3]	[21.4]	[14.6]	[45.2]	[26.7]	[38.1]
C. Year Post-Shelter Exit									
Cash Assistance	0.265**	0.087	0.102	0.789*	0.192	0.285	-0.033**	-0.018	-0.015
	(0.129)	(0.148)	(0.171)	(0.447)	(0.392)	(0.331)	(0.016)	(0.024)	(0.017)
	[46.4]	[34.2]	[27.2]	[12.5]	[11.4]	[16.2]	[33.5]	[15.1]	[37.1]
Food Stamps	0.023	-0.075	0.003	0.089	-0.214	0.018	-0.007	0.009	-0.004
	(0.086)	(0.103)	(0.114)	(0.267)	(0.275)	(0.214)	(0.011)	(0.016)	(0.011)
	[46.4]	[34.2]	[27.2]	[12.5]	[11.4]	[16.2]	[33.5]	[15.1]	[37.1]
Employed	0.338**	0.352**	0.268	1.138**	0.930*	0.585	-0.047**	-0.060**	-0.028
	(0.149)	(0.174)	(0.197)	(0.545)	(0.509)	(0.387)	(0.019)	(0.030)	(0.019)
	[46.4]	[34.2]	[27.2]	[12.5]	[11.4]	[16.2]	[33.5]	[15.1]	[37.1]
Log Avg. Quarterly Earnings	2.035*	1.975	1.221	7.314*	5.541	3.126	-0.295**	-0.358*	-0.143
	(1.078)	(1.261)	(1.406)	(3.850)	(3.590)	(2.739)	(0.138)	(0.215)	(0.137)
	[46.4]	[34.2]	[27.2]	[12.5]	[11.4]	[16.2]	[33.5]	[15.1]	[37.1]
Time Control	$Year^3$	$Year^3$	$Year^3$	$Year^3$	$Year^3$	$Year^3$	$Year^3$	$Year^3$	$Year^3$
Placement Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Family & Shelter Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Shelter FE	No	No	No	No	No	No	No	No	No

Each cell reports the coefficient on local shelter placement from a separate 2SLS regression of the row-delineated outcome on the treatment using the aversion ratio as the instrument and controlling for Main covariates. Columns give samples; supercolumns give treatment definitions. Standard errors clustered at family group level in parentheses. First-stage F-stats in brackets. * $p < 0.10$, ** $p < 0.05$

Table A.15: Time Trend Robustness

	OLS				Ineligibility Rate IV				Aversion Ratio IV			
Outcome	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. Stays and Returns												
Log Length of Stay	0.121**	0.119**	0.120**	0.118**	1.704**	3.941**	1.943**	0.093	1.298**	1.787**	1.284**	0.108
	(0.011)	(0.011)	(0.011)	(0.011)	(0.438)	(1.815)	(0.652)	(0.352)	(0.275)	(0.542)	(0.318)	(0.274)
	{59,253}	{59,253}	{59,253}	{59,253}	[47.0]	[6.5]	[23.2]	[52.8]	[106.2]	[32.3]	[76.8]	[85.8]
Subsidized Exit	0.024**	0.019**	0.021**	0.021**	0.110	-1.301*	-2.675**	0.338**	1.446**	-0.207	0.057	0.227*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.159)	(0.712)	(0.635)	(0.154)	(0.176)	(0.196)	(0.127)	(0.118)
	{57,962}	{57,962}	{57,962}	{57,962}	[44.7]	[5.5]	[20.8]	[48.9]	[101.7]	[28.3]	[70.5]	[79.3]
Returned to Shelter	-0.006*	-0.005	-0.006*	-0.005	0.347**	-0.047	0.772**	0.208*	-0.147*	-0.234	-0.064	0.156*
	(0.003)	(0.003)	(0.003)	(0.003)	(0.141)	(0.294)	(0.258)	(0.120)	(0.083)	(0.143)	(0.093)	(0.091)
	{52,274}	{52,274}	{52,274}	{52,274}	[36.9]	[7.2]	[18.2]	[46.7]	[91.2]	[33.1]	[69.2]	[79.2]
B. Year Post-Shelter Entry												
Cash Assistance	0.011**	0.011**	0.012**	0.012**	-0.361**	-1.348**	0.954**	-0.387**	-0.346**	-0.572**	0.642**	-0.380**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.124)	(0.618)	(0.250)	(0.117)	(0.083)	(0.174)	(0.112)	(0.091)
	{59,253}	{59,253}	{59,253}	{59,253}	[47.0]	[6.5]	[23.2]	[52.8]	[106.2]	[32.3]	[76.8]	[85.8]
Food Stamps	0.003*	0.003	0.004*	0.004*	-0.257**	-1.015**	-0.048	-0.473**	-0.120**	-0.372**	0.101*	-0.352**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.080)	(0.452)	(0.099)	(0.095)	(0.048)	(0.115)	(0.056)	(0.066)
	{59,253}	{59,253}	{59,253}	{59,253}	[47.0]	[6.5]	[23.2]	[52.8]	[106.2]	[32.3]	[76.8]	[85.8]
Employed	0.010**	0.010**	0.011**	0.011**	-0.877**	-0.441	-0.087	-0.276**	-0.384**	0.168	0.439**	-0.207**
	(0.004)	(0.004)	(0.004)	(0.004)	(0.187)	(0.407)	(0.193)	(0.132)	(0.099)	(0.167)	(0.115)	(0.101)
	{59,253}	{59,253}	{59,253}	{59,253}	[47.0]	[6.5]	[23.2]	[52.8]	[106.2]	[32.3]	[76.8]	[85.8]
Log Avg. Quarterly Earnings	0.095**	0.093**	0.100**	0.097**	-3.259**	-1.624	0.975	-1.473	-1.096*	0.343	3.084**	-1.412*
	(0.028)	(0.028)	(0.028)	(0.028)	(1.081)	(2.702)	(1.383)	(0.929)	(0.658)	(1.186)	(0.825)	(0.725)
	{59,253}	{59,253}	{59,253}	{59,253}	[47.0]	[6.5]	[23.2]	[52.8]	[106.2]	[32.3]	[76.8]	[85.8]
C. Year Post-Shelter Exit												
Cash Assistance	0.017**	0.017**	0.017**	0.017**	-0.441**	0.240	-0.098	-0.062	-0.193*	0.271	0.159	-0.075
	(0.004)	(0.004)	(0.004)	(0.004)	(0.200)	(0.406)	(0.235)	(0.160)	(0.109)	(0.182)	(0.111)	(0.117)
	{48,082}	{48,082}	{48,082}	{48,082}	[23.1]	[4.9]	[13.3]	[28.8]	[66.4]	[25.3]	[60.0]	[53.1]
Food Stamps	0.009**	0.009**	0.009**	0.009**	-0.131	-0.090	-0.188	-0.256**	0.021	0.139	0.035	-0.112
	(0.003)	(0.003)	(0.003)	(0.003)	(0.124)	(0.271)	(0.166)	(0.121)	(0.071)	(0.122)	(0.075)	(0.084)
	{48,082}	{48,082}	{48,082}	{48,082}	[23.1]	[4.9]	[13.3]	[28.8]	[66.4]	[25.3]	[60.0]	[53.1]
Employed	0.003	0.003	0.004	0.003	-1.171**	0.477	-0.018	0.322*	-0.716**	0.375*	0.326**	0.171
	(0.004)	(0.004)	(0.004)	(0.004)	(0.314)	(0.493)	(0.266)	(0.189)	(0.148)	(0.209)	(0.130)	(0.132)
	{48,082}	{48,082}	{48,082}	{48,082}	[23.1]	[4.9]	[13.3]	[28.8]	[66.4]	[25.3]	[60.0]	[53.1]
Log Avg. Quarterly Earnings	0.042	0.042	0.047	0.040	-7.755**	2.137	-0.792	2.054	-4.504**	1.855	1.978**	1.029
	(0.033)	(0.033)	(0.033)	(0.033)	(2.186)	(3.385)	(1.970)	(1.371)	(1.039)	(1.477)	(0.945)	(0.973)
	{48,082}	{48,082}	{48,082}	{48,082}	[23.1]	[4.9]	[13.3]	[28.8]	[66.4]	[25.3]	[60.0]	[53.1]
Time Control	Year Linear	Year Dummies	3-Knot Month^2 Spline	7-Knot Month^3 Spline	Year Linear	Year Dummies	3-Knot Month^2 Spline	7-Knot Month^3 Spline	Year Linear	Year Dummies	3-Knot Month^2 Spline	7-Knot Month^3 Spline
Main Covariates	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Each cell reports the coefficient on in-borough shelter placement from a separate regressions of the row-delineated outcome on the treatment indicator using the supercolumn-enumerated method, controlling for Main covariates. Columns give alternative time trend controls. Standard errors clustered at family group level in parentheses. Number of observations given in braces. First-stage F-stats in brackets. * $p < 0.10$, ** $p < 0.05$

Table A.16: Regression Discontinuity Main Results: Wald Estimates

	No Controls				Controls			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Stays and Returns								
Log Length of Stay	1.986**	1.619**	1.557**	1.612**	1.205**	0.732**	0.705*	0.467
	(0.705)	(0.311)	(0.442)	(0.271)	(0.569)	(0.299)	(0.388)	(0.284)
	{7,679}	{15,436}	{7,430}	{14,925}	{7,679}	{15,436}	{7,430}	{14,925}
Subsidized Exit	0.353*	0.473**	0.406**	0.661**	0.128	0.231**	0.170	0.363**
	(0.211)	(0.109)	(0.152)	(0.106)	(0.184)	(0.108)	(0.141)	(0.111)
	{7,548}	{15,156}	{7,299}	{14,642}	{7,548}	{15,156}	{7,299}	{14,642}
Returned to Shelter	-0.067	-0.199**	-0.167	-0.247**	-0.060	-0.167*	-0.172	-0.220**
	(0.153)	(0.083)	(0.117)	(0.075)	(0.152)	(0.095)	(0.122)	(0.094)
	{6,798}	{13,725}	{6,590}	{13,268}	{6,798}	{13,725}	{6,590}	{13,268}
B. Year Post-Shelter Entry								
Cash Assistance	0.223	0.126	0.172	0.025	0.131	0.146*	0.248**	0.162**
	(0.188)	(0.087)	(0.126)	(0.076)	(0.151)	(0.085)	(0.112)	(0.082)
	{7,679}	{15,436}	{7,430}	{14,925}	{7,679}	{15,436}	{7,430}	{14,925}
Food Stamps	0.070	-0.049	-0.034	-0.137**	0.012	-0.002	0.052	0.018
	(0.130)	(0.062)	(0.089)	(0.056)	(0.089)	(0.050)	(0.065)	(0.049)
	{7,679}	{15,436}	{7,430}	{14,925}	{7,679}	{15,436}	{7,430}	{14,925}
Employed	0.001	-0.094	-0.275*	-0.268**	0.000	-0.027	-0.108	-0.083
	(0.223)	(0.108)	(0.159)	(0.098)	(0.189)	(0.106)	(0.138)	(0.102)
	{7,679}	{15,436}	{7,430}	{14,925}	{7,679}	{15,436}	{7,430}	{14,925}
Log Avg. Quarterly Earnings	0.881	0.059	-1.491	-1.131	0.567	0.306	-0.560	-0.100
	(1.623)	(0.776)	(1.130)	(0.690)	(1.333)	(0.745)	(0.970)	(0.718)
	{7,679}	{15,436}	{7,430}	{14,925}	{7,679}	{15,436}	{7,430}	{14,925}
C. Year Post-Shelter Exit								
Cash Assistance	0.403**	0.250**	0.354**	0.138*	0.303*	0.255**	0.373**	0.231**
	(0.191)	(0.096)	(0.140)	(0.084)	(0.172)	(0.103)	(0.135)	(0.099)
	{6,295}	{12,675}	{6,092}	{12,246}	{6,295}	{12,675}	{6,092}	{12,246}
Food Stamps	0.212	-0.031	0.107	-0.107*	0.130	0.008	0.157*	0.021
	(0.130)	(0.065)	(0.094)	(0.059)	(0.106)	(0.062)	(0.082)	(0.061)
	{6,295}	{12,675}	{6,092}	{12,246}	{6,295}	{12,675}	{6,092}	{12,246}
Employed	-0.147	-0.189*	-0.189	-0.287**	-0.170	-0.088	-0.009	-0.092
	(0.203)	(0.109)	(0.153)	(0.099)	(0.190)	(0.114)	(0.143)	(0.110)
	{6,295}	{12,675}	{6,092}	{12,246}	{6,295}	{12,675}	{6,092}	{12,246}
Log Avg. Quarterly Earnings	-0.901	-1.063	-1.404	-1.533**	-1.241	-0.603	-0.305	-0.458
	(1.485)	(0.793)	(1.126)	(0.714)	(1.372)	(0.826)	(1.039)	(0.798)
	{6,295}	{12,675}	{6,092}	{12,246}	{6,295}	{12,675}	{6,092}	{12,246}
First Stage	0.051**	0.076**	0.077**	0.089**	0.053**	0.068**	0.075**	0.073**
	(0.011)	(0.008)	(0.012)	(0.008)	(0.010)	(0.007)	(0.011)	(0.008)
	[20.4]	[90.3]	[44.1]	[117.8]	[25.9]	[83.8]	[50.0]	[89.8]
Bandwidth	{-1,0}	[-2,1]	{-1,1}	{-2,-1,1,2}	{-1,0}	[-2,1]	{-1,1}	{-2,-1,1,2}
Covariates	No	No	No	No	Yes	Yes	Yes	Yes

This table presents a more comprehensive set of Wald fuzzy regression discontinuity estimates. Each cell reports the coefficient on in-borough shelter placement from a separate 2SLS regression of the row-delineated outcome on the treatment indicator, using as the instrument an indicator for whether a family's oldest child's potential grade (end-of-calendar-year age year minus five) is zero or greater. Wald estimates pool the running variable for the given bandwidth; coefficients are thus instrumented mean comparisons between families without and with school-aged children. The first four columns have no covariates. The last four control for RD Main covariates. Standard errors clustered at family group level in parentheses. Number of observations given in braces. First-stage given for in-borough placement indicator. First-stage F-stat, in brackets, given for log length of stay regressions. * $p < 0.10$, ** $p < 0.05$

Table A.17: Regression Discontinuity Main Results: Linear Estimates

	No Controls					Controls				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A. Stays and Returns										
Log Length of Stay	1.611	0.910	2.075	1.357**	1.281**	1.505*	0.917	1.885*	1.065**	0.918**
	(0.993)	(0.959)	(1.340)	(0.436)	(0.354)	(0.827)	(0.855)	(1.042)	(0.331)	(0.285)
	{26,046}	{22,316}	{19,641}	{50,480}	{55,118}	{26,046}	{22,316}	{19,641}	{50,480}	{55,118}
Subsidized Exit	0.247	0.121	0.261	0.622**	0.479**	0.160	0.092	0.159	0.370**	0.256**
	(0.299)	(0.333)	(0.365)	(0.171)	(0.131)	(0.255)	(0.302)	(0.297)	(0.126)	(0.104)
	{25,543}	{21,886}	{19,284}	{49,334}	{53,907}	{25,543}	{21,886}	{19,284}	{49,334}	{53,907}
Returned to Shelter	0.226	0.204	0.212	0.013	-0.084	0.187	0.131	0.193	-0.042	-0.107
	(0.230)	(0.265)	(0.277)	(0.120)	(0.097)	(0.212)	(0.252)	(0.246)	(0.101)	(0.089)
	{23,141}	{19,860}	{17,508}	{44,574}	{48,712}	{23,141}	{19,860}	{17,508}	{44,574}	{48,712}
B. Year Post-Shelter Entry										
Cash Assistance	0.361	0.363	0.400	0.170	0.059	0.242	0.333	0.215	0.183**	0.104
	(0.301)	(0.318)	(0.381)	(0.128)	(0.107)	(0.216)	(0.253)	(0.254)	(0.092)	(0.079)
	{26,046}	{22,316}	{19,641}	{50,480}	{55,118}	{26,046}	{22,316}	{19,641}	{50,480}	{55,118}
Food Stamps	0.157	0.086	0.228	-0.037	-0.095	0.033	0.061	0.029	0.009	-0.018
	(0.203)	(0.214)	(0.261)	(0.090)	(0.077)	(0.125)	(0.143)	(0.146)	(0.055)	(0.048)
	{26,046}	{22,316}	{19,641}	{50,480}	{55,118}	{26,046}	{22,316}	{19,641}	{50,480}	{55,118}
Employed	0.090	-0.293	0.403	-0.123	-0.145	-0.019	-0.341	0.238	-0.081	-0.070
	(0.340)	(0.380)	(0.455)	(0.156)	(0.131)	(0.265)	(0.312)	(0.319)	(0.114)	(0.101)
	{26,046}	{22,316}	{19,641}	{50,480}	{55,118}	{26,046}	{22,316}	{19,641}	{50,480}	{55,118}
Log Avg. Quarterly Earnings	1.169	-2.488	4.022	-0.568	-0.747	0.374	-2.688	2.673	-0.277	-0.213
	(2.476)	(2.784)	(3.483)	(1.124)	(0.940)	(1.868)	(2.244)	(2.335)	(0.815)	(0.723)
	{26,046}	{22,316}	{19,641}	{50,480}	{55,118}	{26,046}	{22,316}	{19,641}	{50,480}	{55,118}
C. Year Post-Shelter Exit										
Cash Assistance	0.650**	0.672*	0.670*	0.398**	0.173	0.557**	0.666**	0.505*	0.347**	0.206**
	(0.301)	(0.349)	(0.378)	(0.152)	(0.119)	(0.253)	(0.312)	(0.294)	(0.120)	(0.099)
	{21,348}	{18,327}	{16,182}	{41,110}	{44,941}	{21,348}	{18,327}	{16,182}	{41,110}	{44,941}
Food Stamps	0.322*	0.299	0.346	0.091	-0.016	0.183	0.264	0.130	0.071	0.029
	(0.193)	(0.218)	(0.242)	(0.100)	(0.081)	(0.144)	(0.177)	(0.166)	(0.073)	(0.062)
	{21,348}	{18,327}	{16,182}	{41,110}	{44,941}	{21,348}	{18,327}	{16,182}	{41,110}	{44,941}
Employed	-0.132	-0.222	-0.017	-0.219	-0.268**	-0.178	-0.236	-0.056	-0.135	-0.134
	(0.283)	(0.322)	(0.349)	(0.162)	(0.135)	(0.255)	(0.292)	(0.298)	(0.128)	(0.113)
	{21,348}	{18,327}	{16,182}	{41,110}	{44,941}	{21,348}	{18,327}	{16,182}	{41,110}	{44,941}
Log Avg. Quarterly Earnings	-1.315	-2.477	-0.269	-1.606	-1.898*	-1.561	-2.406	-0.547	-0.909	-0.866
	(2.095)	(2.444)	(2.566)	(1.192)	(0.992)	(1.873)	(2.171)	(2.162)	(0.935)	(0.826)
	{21,348}	{18,327}	{16,182}	{41,110}	{44,941}	{21,348}	{18,327}	{16,182}	{41,110}	{44,941}
First Stage	0.033**	0.045**	0.032**	0.051**	0.057**	0.039**	0.050**	0.040**	0.058**	0.063**
	(0.014)	(0.018)	(0.014)	(0.013)	(0.011)	(0.013)	(0.016)	(0.013)	(0.012)	(0.010)
	[7.4]	[4.9]	[4.3]	[89.6]	[109.1]	[8.6]	[5.6]	[6.8]	[104.1]	[120.7]
Bandwidth	[-3,3]	[-3,3]	[-3,2]	[-3,12]	[-4,12]	[-3,3]	[-3,3]	[-3,2]	[-3,12]	[-4,12]
Includes Threshold	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
Covariates	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes

This table presents a more comprehensive set of linear fuzzy regression discontinuity estimates. Each cell reports the coefficient on in-borough shelter placement from a separate 2SLS regression of the row-delineated outcome on the running variable (oldest child's potential grade; i.e., end-of-calendar-year age year minus five), the treatment indicator, and treatment interacted with the running variable, so as to allow for different slopes on either side of the threshold (school starting; i.e., potential grade zero). The instrument an indicator for whether a family's oldest child's potential grade is zero or greater; the interaction term is also instrumented. Reported coefficients are thus the difference in intercepts at the threshold. The first four columns have no covariates. The last four control for RD Main covariates. Standard errors clustered at family group level in parentheses. Number of observations given in braces. First-stage given for in-borough placement indicator. First-stage F-stat, in brackets, given for log length of stay regressions. * $p < 0.10$, ** $p < 0.05$

Table A.18: Regression Discontinuity Robustness: Alternative Samples for Borough Treatment

	Non-DV Sample				Pre-2015 Sample				One School Age Sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
A. Stays and Returns												
Log Length of Stay	1.070*	0.968**	0.107	0.590**	1.788**	1.426**	0.197	0.675**	1.588**	1.331**	0.645**	1.172**
	(0.632)	(0.265)	(0.267)	(0.192)	(0.771)	(0.284)	(0.322)	(0.319)	(0.620)	(0.264)	(0.277)	(0.395)
	{4,986}	{9,718}	{9,718}	{36,023}	{5,435}	{10,541}	{10,541}	{35,758}	{7,441}	{13,212}	{13,212}	{32,006}
Subsidized Exit	0.203	0.561**	0.263**	0.252**	0.236	0.513**	0.310**	0.255**	0.255	0.471**	0.290**	0.451**
	(0.246)	(0.123)	(0.119)	(0.085)	(0.229)	(0.103)	(0.121)	(0.115)	(0.196)	(0.097)	(0.104)	(0.148)
	{4,889}	{9,508}	{9,508}	{35,155}	{5,411}	{10,479}	{10,479}	{35,498}	{7,321}	{12,969}	{12,969}	{31,310}
Returned to Shelter	0.027	-0.150*	-0.118	-0.059	-0.124	-0.165**	-0.174*	-0.037	-0.070	-0.239**	-0.177**	0.024
	(0.173)	(0.081)	(0.092)	(0.070)	(0.182)	(0.077)	(0.101)	(0.097)	(0.148)	(0.073)	(0.088)	(0.124)
	{4,419}	{8,595}	{8,595}	{31,764}	{5,305}	{10,260}	{10,260}	{34,744}	{6,607}	{11,787}	{11,787}	{28,404}
B. Year Post-Shelter Entry												
Cash Assistance	0.437*	0.031	0.114	0.178**	0.230	0.055	0.145	0.102	0.213	0.071	0.155**	0.259**
	(0.234)	(0.089)	(0.086)	(0.062)	(0.208)	(0.080)	(0.092)	(0.088)	(0.177)	(0.077)	(0.079)	(0.110)
	{4,986}	{9,718}	{9,718}	{36,023}	{5,435}	{10,541}	{10,541}	{35,758}	{7,441}	{13,212}	{13,212}	{32,006}
Food Stamps	0.119	-0.152**	-0.033	-0.019	0.110	-0.100*	-0.026	-0.041	0.069	-0.064	0.053	0.038
	(0.143)	(0.064)	(0.051)	(0.037)	(0.141)	(0.056)	(0.054)	(0.053)	(0.122)	(0.055)	(0.048)	(0.064)
	{4,986}	{9,718}	{9,718}	{36,023}	{5,435}	{10,541}	{10,541}	{35,758}	{7,441}	{13,212}	{13,212}	{32,006}
Employed	0.127	-0.221*	-0.026	-0.090	-0.150	-0.342**	-0.118	-0.149	0.051	-0.008	-0.007	0.083
	(0.262)	(0.113)	(0.107)	(0.077)	(0.252)	(0.105)	(0.116)	(0.113)	(0.210)	(0.095)	(0.098)	(0.133)
	{4,986}	{9,718}	{9,718}	{36,023}	{5,435}	{10,541}	{10,541}	{35,758}	{7,441}	{13,212}	{13,212}	{32,006}
Log Avg. Quarterly Earnings	1.806	-1.032	0.081	-0.619	-0.116	-2.008**	-0.616	-0.745	1.257	0.697	0.265	0.858
	(1.957)	(0.811)	(0.760)	(0.555)	(1.774)	(0.735)	(0.808)	(0.802)	(1.539)	(0.693)	(0.694)	(0.949)
	{4,986}	{9,718}	{9,718}	{36,023}	{5,435}	{10,541}	{10,541}	{35,758}	{7,441}	{13,212}	{13,212}	{32,006}
C. Year Post-Shelter Exit												
Cash Assistance	0.561**	0.141	0.179*	0.260**	0.364*	0.093	0.157	0.226**	0.385**	0.173**	0.210**	0.478**
	(0.230)	(0.096)	(0.102)	(0.077)	(0.218)	(0.087)	(0.105)	(0.105)	(0.185)	(0.084)	(0.096)	(0.154)
	{4,091}	{7,912}	{7,912}	{29,238}	{5,234}	{10,112}	{10,112}	{34,229}	{6,144}	{10,882}	{10,882}	{26,170}
Food Stamps	0.225	-0.088	0.016	0.021	0.115	-0.130**	-0.058	-0.029	0.209*	-0.030	0.066	0.126
	(0.145)	(0.065)	(0.063)	(0.047)	(0.146)	(0.062)	(0.065)	(0.065)	(0.127)	(0.058)	(0.059)	(0.090)
	{4,091}	{7,912}	{7,912}	{29,238}	{5,234}	{10,112}	{10,112}	{34,229}	{6,144}	{10,882}	{10,882}	{26,170}
Employed	-0.020	-0.287**	-0.075	-0.168**	-0.275	-0.281**	-0.074	-0.172	-0.097	-0.095	-0.048	0.028
	(0.230)	(0.113)	(0.113)	(0.083)	(0.243)	(0.104)	(0.118)	(0.117)	(0.197)	(0.095)	(0.107)	(0.156)
	{4,091}	{7,912}	{7,912}	{29,238}	{5,234}	{10,112}	{10,112}	{34,229}	{6,144}	{10,882}	{10,882}	{26,170}
Log Avg. Quarterly Earnings	0.159	-1.582*	-0.271	-1.215**	-1.450	-1.694**	-0.518	-1.211	-0.616	-0.309	-0.345	0.056
	(1.698)	(0.825)	(0.824)	(0.607)	(1.744)	(0.747)	(0.852)	(0.858)	(1.444)	(0.699)	(0.777)	(1.140)
	{4,091}	{7,912}	{7,912}	{29,238}	{5,234}	{10,112}	{10,112}	{34,229}	{6,144}	{10,882}	{10,882}	{26,170}
First Stage	0.054**	0.093**	0.084**	0.064**	0.055**	0.101**	0.078**	0.069**	0.055**	0.096**	0.081**	0.061**
	(0.014)	(0.010)	(0.009)	(0.015)	(0.013)	(0.010)	(0.009)	(0.014)	(0.012)	(0.009)	(0.008)	(0.012)
	[15.3]	[85.4]	[79.8]	[94.6]	[16.6]	[106.1]	[73.4]	[86.1]	[23.0]	[115.9]	[97.7]	[76.3]
Order	Wald	Wald	Wald	Linear	Wald	Wald	Wald	Linear	Wald	Wald	Wald	Linear
Bandwidth	{-1,0}	{-2,-1,1,2}	{-2,-1,1,2}	[-3,12]	{-1,0}	{-2,-1,1,2}	{-2,-1,1,2}	[-3,12]	{-1,0}	{-2,-1,1,2}	{-2,-1,1,2}	[-3,12]
Threshold	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes
Covariates	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes

This table extends the fuzzy regression discontinuity analysis for three alternative samples, given in supercolumns. The Non-DV sample consists of families eligible for shelter for reasons other than domestic violence. The Pre-2015 sample consists of families entering shelter before 2014. The One School Age Sample consists of families with a single school-aged child. Each cell reports the coefficient on in-borough shelter placement from a separate 2SLS regression of the row-delineated outcome on the treatment indicator, using as the instrument an indicator for whether a family's oldest child's potential grade (end-of-calendar-year age year minus five) is zero (i.e., in-school) or greater. The first three columns for each sample present Wald estimates for varying bandwidths, while the fourth fits a linear regression on the running variable for the given bandwidths, allowing for different slopes on either side of the threshold. The first two columns for each sample have no covariates; the last two control for Main RD covariates. Standard errors clustered at family group level in parentheses. Number of observations given in braces. First-stage given for in-borough placement indicator. First-stage F-stat, in brackets, given for log length of stay regressions. * $p < 0.10$, ** $p < 0.05$

Table A.19: Regression Discontinuity Robustness: Distance Treatment

	Full Sample				Non-DV Sample				Pre-2015 Sample				One School Age Sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
A. Stays and Returns																
Log Length of Stay	-0.266**	-0.199**	-0.054	-0.114**	-0.141	-0.121**	-0.014	-0.075**	-0.270*	-0.176**	-0.025	-0.082**	-0.195**	-0.160**	-0.081**	-0.138**
	(0.118)	(0.037)	(0.033)	(0.032)	(0.101)	(0.035)	(0.032)	(0.022)	(0.151)	(0.039)	(0.038)	(0.040)	(0.092)	(0.033)	(0.033)	(0.055)
	{7,028}	{13,639}	{13,639}	{46,282}	{4,611}	{8,928}	{8,928}	{33,311}	{4,938}	{9,577}	{9,577}	{32,635}	{6,807}	{12,073}	{12,073}	{29,280}
Subsidized Exit	-0.057	-0.084**	-0.043**	-0.041**	-0.025	-0.072**	-0.035**	-0.030**	-0.039	-0.065**	-0.039**	-0.034**	-0.041	-0.058**	-0.038**	-0.064**
	(0.036)	(0.015)	(0.013)	(0.013)	(0.039)	(0.017)	(0.014)	(0.010)	(0.042)	(0.014)	(0.015)	(0.015)	(0.030)	(0.013)	(0.013)	(0.022)
	{6,910}	{13,377}	{13,377}	{45,228}	{4,523}	{8,735}	{8,735}	{32,511}	{4,918}	{9,524}	{9,524}	{32,405}	{6,699}	{11,846}	{11,846}	{28,644}
Returned to Shelter	0.002	0.034**	0.026**	0.011	-0.020	0.019*	0.012	0.009	0.005	0.022**	0.021*	0.008	0.004	0.031**	0.021**	-0.006
	(0.027)	(0.010)	(0.011)	(0.011)	(0.032)	(0.011)	(0.011)	(0.009)	(0.034)	(0.010)	(0.012)	(0.012)	(0.024)	(0.010)	(0.011)	(0.019)
	{6,211}	{12,114}	{12,114}	{40,829}	{4,078}	{7,892}	{7,892}	{29,354}	{4,819}	{9,319}	{9,319}	{31,707}	{6,034}	{10,755}	{10,755}	{25,947}
B. Year Post-Shelter Entry																
Cash Assistance	-0.035	-0.003	-0.021**	-0.019**	-0.067*	-0.007	-0.017*	-0.022**	-0.041	-0.006	-0.018*	-0.012	-0.033	-0.008	-0.020**	-0.035**
	(0.030)	(0.010)	(0.010)	(0.009)	(0.040)	(0.012)	(0.010)	(0.007)	(0.038)	(0.011)	(0.011)	(0.011)	(0.026)	(0.009)	(0.009)	(0.016)
	{7,028}	{13,639}	{13,639}	{46,282}	{4,611}	{8,928}	{8,928}	{33,311}	{4,938}	{9,577}	{9,577}	{32,635}	{6,807}	{12,073}	{12,073}	{29,280}
Food Stamps	-0.009	0.017**	-0.002	0.002	-0.014	0.018**	0.004	0.002	-0.017	0.014*	0.003	0.008	-0.009	0.008	-0.006	-0.006
	(0.020)	(0.008)	(0.006)	(0.005)	(0.022)	(0.008)	(0.006)	(0.004)	(0.025)	(0.008)	(0.006)	(0.007)	(0.017)	(0.007)	(0.006)	(0.009)
	{7,028}	{13,639}	{13,639}	{46,282}	{4,611}	{8,928}	{8,928}	{33,311}	{4,938}	{9,577}	{9,577}	{32,635}	{6,807}	{12,073}	{12,073}	{29,280}
Employed	0.002	0.032**	0.008	0.015	-0.023	0.028*	0.007	0.016*	0.032	0.040**	0.011	0.022	-0.006	0.000	-0.000	-0.010
	(0.034)	(0.013)	(0.012)	(0.011)	(0.041)	(0.015)	(0.013)	(0.009)	(0.045)	(0.014)	(0.014)	(0.014)	(0.030)	(0.012)	(0.012)	(0.018)
	{7,028}	{13,639}	{13,639}	{46,282}	{4,611}	{8,928}	{8,928}	{33,311}	{4,938}	{9,577}	{9,577}	{32,635}	{6,807}	{12,073}	{12,073}	{29,280}
Log Avg. Quarterly Earnings	-0.142	0.127	-0.002	0.070	-0.338	0.124	0.008	0.107*	0.048	0.223**	0.040	0.107	-0.186	-0.094	-0.042	-0.126
	(0.250)	(0.091)	(0.084)	(0.081)	(0.317)	(0.106)	(0.091)	(0.063)	(0.310)	(0.097)	(0.096)	(0.101)	(0.224)	(0.086)	(0.082)	(0.132)
	{7,028}	{13,639}	{13,639}	{46,282}	{4,611}	{8,928}	{8,928}	{33,311}	{4,938}	{9,577}	{9,577}	{32,635}	{6,807}	{12,073}	{12,073}	{29,280}
C. Year Post-Shelter Exit																
Cash Assistance	-0.073*	-0.015	-0.026**	-0.039**	-0.091**	-0.017	-0.020*	-0.027**	-0.069	-0.009	-0.017	-0.026*	-0.065**	-0.019*	-0.025**	-0.074**
	(0.038)	(0.011)	(0.012)	(0.014)	(0.046)	(0.013)	(0.012)	(0.009)	(0.046)	(0.012)	(0.013)	(0.013)	(0.033)	(0.011)	(0.012)	(0.028)
	{5,743}	{11,166}	{11,166}	{37,616}	{3,774}	{7,254}	{7,254}	{26,999}	{4,751}	{9,184}	{9,184}	{31,229}	{5,603}	{9,913}	{9,913}	{23,882}
Food Stamps	-0.034	0.015*	-0.003	-0.006	-0.036	0.011	-0.003	-0.002	-0.016	0.019**	0.007	0.007	-0.032	0.004	-0.008	-0.022
	(0.024)	(0.008)	(0.007)	(0.009)	(0.026)	(0.009)	(0.008)	(0.006)	(0.027)	(0.008)	(0.008)	(0.008)	(0.021)	(0.007)	(0.007)	(0.015)
	{5,743}	{11,166}	{11,166}	{37,616}	{3,774}	{7,254}	{7,254}	{26,999}	{4,751}	{9,184}	{9,184}	{31,229}	{5,603}	{9,913}	{9,913}	{23,882}
Employed	0.019	0.033**	0.006	0.019	-0.006	0.035**	0.009	0.020**	0.048	0.032**	0.003	0.023	0.011	0.007	0.001	-0.011
	(0.035)	(0.014)	(0.013)	(0.015)	(0.039)	(0.015)	(0.014)	(0.010)	(0.047)	(0.014)	(0.014)	(0.015)	(0.031)	(0.012)	(0.013)	(0.026)
	{5,743}	{11,166}	{11,166}	{37,616}	{3,774}	{7,254}	{7,254}	{26,999}	{4,751}	{9,184}	{9,184}	{31,229}	{5,603}	{9,913}	{9,913}	{23,882}
Log Avg. Quarterly Earnings	0.102	0.163*	0.021	0.124	-0.091	0.185*	0.033	0.152**	0.233	0.176*	0.013	0.159	0.058	-0.000	0.004	-0.075
	(0.255)	(0.098)	(0.097)	(0.110)	(0.291)	(0.110)	(0.100)	(0.073)	(0.329)	(0.101)	(0.104)	(0.109)	(0.231)	(0.090)	(0.095)	(0.186)
	{5,743}	{11,166}	{11,166}	{37,616}	{3,774}	{7,254}	{7,254}	{26,999}	{4,751}	{9,184}	{9,184}	{31,229}	{5,603}	{9,913}	{9,913}	{23,882}
First Stage	-0.350**	-0.706**	-0.645**	-0.520**	-0.368**	-0.745**	-0.734**	-0.545**	-0.329**	-0.796**	-0.680**	-0.553**	-0.402**	-0.813**	-0.718**	-0.504**
	(0.110)	(0.080)	(0.074)	(0.114)	(0.134)	(0.095)	(0.089)	(0.141)	(0.129)	(0.092)	(0.086)	(0.132)	(0.112)	(0.086)	(0.079)	(0.117)
	[10.1]	[78.7]	[75.5]	[80.2]	[7.5]	[61.0]	[67.7]	[74.1]	[6.5]	[74.4]	[63.2]	[56.9]	[12.9]	[89.7]	[82.3]	[49.1]
Order	Wald	Wald	Wald	Linear	Wald	Wald	Wald	Linear	Wald	Wald	Wald	Linear	Wald	Wald	Wald	Linear
Bandwidth	{-1,0}	{-2,-1,1,2}	{-2,-1,1,2}	[-3,12]	{-1,0}	{-2,-1,1,2}	{-2,-1,1,2}	[-3,12]	{-1,0}	{-2,-1,1,2}	{-2,-1,1,2}	[-3,12]	{-1,0}	{-2,-1,1,2}	{-2,-1,1,2}	[-3,12]
Threshold	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes
Covariates	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes

This table extends the fuzzy regression discontinuity analysis for distance treatment, measured in miles between origin and shelter address. Supercolumns give samples. Each cell reports the coefficient on placement distance from a separate 2SLS regression of the row-delineated outcome on distance treatment, using as the instrument an indicator for whether a family's oldest child's potential grade (end-of-calendar-year age year minus five) is zero (i.e., in-school) or greater. The first three columns for each sample present Wald estimates for varying bandwidths, while the fourth fits a linear regression on the running variable for the given bandwidths, allowing for different slopes on either side of the threshold. The first two columns for each sample have no covariates; the last two control for Main RD covariates. Standard errors clustered at family group level in parentheses. Number of observations given in braces. First-stage given for distance treatment. First-stage F-stat, in brackets, given for log length of stay regressions.
* $p < 0.10$, ** $p < 0.05$

Table A.20: Regression Discontinuity Robustness: School Year Running Variable Definition

	No Controls				Controls			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Stays and Returns								
Log Length of Stay	1.998**	1.419**	0.979	0.874**	0.975	-0.061	0.675	0.670**
	(0.900)	(0.323)	(1.395)	(0.324)	(0.784)	(0.363)	(1.289)	(0.321)
	{7,384}	{14,306}	{25,160}	{48,230}	{7,384}	{14,306}	{25,160}	{48,230}
Subsidized Exit	0.462	0.666**	-0.188	0.521**	0.183	0.254*	-0.319	0.323**
	(0.282)	(0.134)	(0.450)	(0.138)	(0.272)	(0.141)	(0.456)	(0.129)
	{7,240}	{14,027}	{24,651}	{47,108}	{7,240}	{14,027}	{24,651}	{47,108}
Returned to Shelter	-0.303	-0.240**	0.102	-0.112	-0.285	-0.217*	0.011	-0.147
	(0.208)	(0.095)	(0.340)	(0.100)	(0.214)	(0.123)	(0.321)	(0.100)
	{6,487}	{12,676}	{22,300}	{42,506}	{6,487}	{12,676}	{22,300}	{42,506}
B. Year Post-Shelter Entry Outcomes								
Cash Assistance	0.235	0.057	0.587	0.212**	0.210	0.250**	0.375	0.207**
	(0.242)	(0.094)	(0.509)	(0.106)	(0.223)	(0.109)	(0.395)	(0.096)
	{7,384}	{14,306}	{25,160}	{48,230}	{7,384}	{14,306}	{25,160}	{48,230}
Food Stamps	0.032	-0.138*	0.325	-0.007	0.085	0.055	0.159	0.048
	(0.169)	(0.070)	(0.338)	(0.075)	(0.131)	(0.064)	(0.228)	(0.060)
Employed	-0.264	-0.278**	0.099	-0.129	-0.149	0.027	-0.064	-0.118
	(0.301)	(0.121)	(0.525)	(0.127)	(0.278)	(0.131)	(0.463)	(0.117)
Log Avg. Quarterly Earnings	-1.614	-1.113	-0.358	-1.039	-1.217	0.663	-1.465	-0.909
	(2.155)	(0.855)	(3.783)	(0.924)	(1.972)	(0.931)	(3.344)	(0.824)
C. Year Post-Shelter Exit Outcomes								
Cash Assistance	0.164	0.180*	0.688	0.333**	0.161	0.322**	0.591	0.298**
	(0.233)	(0.106)	(0.486)	(0.137)	(0.232)	(0.132)	(0.421)	(0.125)
	{5,986}	{11,666}	{20,531}	{39,145}	{5,986}	{11,666}	{20,531}	{39,145}
Food Stamps	-0.166	-0.126*	0.127	-0.031	-0.128	0.060	0.006	0.002
	(0.170)	(0.076)	(0.280)	(0.089)	(0.148)	(0.080)	(0.224)	(0.073)
Employed	-0.356	-0.293**	-0.232	-0.271*	-0.207	0.029	-0.244	-0.191
	(0.283)	(0.125)	(0.464)	(0.149)	(0.270)	(0.142)	(0.428)	(0.133)
Log Avg. Quarterly Earnings	-1.962	-1.439	-2.227	-1.787	-1.129	0.468	-2.179	-1.087
	(2.023)	(0.898)	(3.475)	(1.088)	(1.934)	(1.030)	(3.152)	(0.964)
	{5,986}	{11,666}	{20,531}	{39,145}	{5,986}	{11,666}	{20,531}	{39,145}
First Stage	0.040**	0.074**	0.013	0.031**	0.037**	0.058**	0.013	0.029**
	(0.012)	(0.008)	(0.014)	(0.013)	(0.011)	(0.008)	(0.013)	(0.012)
	[12.0]	[76.7]	[8.3]	[88.3]	[12.3]	[54.1]	[9.2]	[95.8]
Order	Wald	Wald	Linear	Linear	Wald	Wald	Linear	Linear
Bandwidth	{-1,0}	{-2,-1,1,2}	[-3,3]	[-3,12]	{-1,0}	{-2,-1,1,2}	[-3,3]	[-3,12]
Threshold	Yes	No	Yes	Yes	Yes	No	Yes	Yes
Covariates	No	No	No	No	Yes	Yes	Yes	Yes

The table presents fuzzy regression discontinuity analysis using families' oldest children's potential grade levels, adjusted for timing of shelter entry relative to the school year as the running variable (i.e., end-of-calendar-year age year minus five for July-December shelter entrants and end-of-calendar-year age year minus six for January-June shelter entrants). Each cell reports the coefficient on in-borough shelter placement from a separate 2SLS regression of the row-delineated outcome on the treatment indicator, using as the instrument an indicator for whether a family's oldest child's potential grade is zero or greater. Columns 1, 2, 5, and 6 give Wald estimates pooling the running variable for the given bandwidth; coefficients are thus instrumented mean comparisons between families without and with school-aged children. Columns 3, 4, 7, and 8 fit linear regressions on the running variable for the given bandwidths, allowing for different slopes on either side of the threshold; the coefficients are the difference in intercepts at the threshold. Columns 1, 3, 5, and 7 include the threshold in the analysis; Columns 2, 4, 6, and 8 exclude it. The last four columns control for Main RD covariates. Standard errors clustered at family group level in parentheses. Number of observations given in braces below first outcome in each panel, as well as for any subsequent outcome where the sample size differs due to censoring. First-stage given for in-borough placement indicator. First-stage F-stat, in brackets, given for log length of stay regressions. * $p < 0.10$, ** $p < 0.05$

Table A.21A: Regression Discontinuity Baseline Covariates

	Wald				Linear	
	(1)	(2)	(3)	(4)	(5)	(6)
Month Entered Shelter	-1.284	0.965	0.656	1.068*	-1.870	0.188
	(1.506)	(1.049)	(0.628)	(0.642)	(2.321)	(1.040)
Year Entered Shelter	2.147*	2.206**	1.862**	1.950**	2.003	2.778**
	(1.108)	(0.770)	(0.462)	(0.470)	(1.643)	(0.837)
Manhattan Origin	-0.112	-0.033	-0.062	-0.065	0.071	0.018
	(0.152)	(0.103)	(0.063)	(0.063)	(0.233)	(0.106)
Bronx Origin	-0.243	0.084	-0.038	0.043	-0.283	-0.107
	(0.230)	(0.148)	(0.092)	(0.092)	(0.352)	(0.155)
Brooklyn Origin	0.107	-0.044	0.081	0.056	0.056	0.060
	(0.208)	(0.141)	(0.087)	(0.087)	(0.313)	(0.144)
Queens Origin	0.205	-0.004	0.031	-0.011	0.159	0.032
	(0.156)	(0.098)	(0.061)	(0.061)	(0.231)	(0.101)
Staten Island Origin	0.044	-0.002	-0.013	-0.023	-0.004	-0.003
	(0.073)	(0.046)	(0.029)	(0.029)	(0.107)	(0.048)
Family Size	2.227**	3.311**	4.096**	4.487**	-0.949	1.140**
	(0.685)	(0.599)	(0.425)	(0.463)	(0.759)	(0.402)
Family Members Under 18	2.441**	3.240**	4.105**	4.412**	-0.672	1.901**
	(0.660)	(0.556)	(0.408)	(0.439)	(0.597)	(0.434)
Health Issue Present	0.135	0.216	0.173**	0.202**	-0.033	-0.078
	(0.200)	(0.140)	(0.085)	(0.087)	(0.299)	(0.136)
Eligibility: Eviction	0.474**	0.647**	0.709**	0.798**	0.010	0.074
	(0.216)	(0.161)	(0.100)	(0.106)	(0.297)	(0.136)
Eligibility: Overcrowding	0.263	0.021	0.069	0.011	0.219	0.103
	(0.178)	(0.115)	(0.070)	(0.070)	(0.264)	(0.116)
Eligibility: Conditions	-0.158	-0.305**	-0.126**	-0.160**	-0.184	-0.190**
	(0.137)	(0.098)	(0.055)	(0.055)	(0.208)	(0.095)
Eligibility: Domestic Violence	-0.348	-0.292**	-0.515**	-0.539**	0.025	0.098
	(0.216)	(0.145)	(0.094)	(0.094)	(0.325)	(0.151)
Eligibility: Other	-0.231	-0.071	-0.141**	-0.113*	-0.074	-0.086
	(0.150)	(0.096)	(0.060)	(0.059)	(0.215)	(0.099)
Female	-0.061	-0.031	-0.058	-0.062	0.047	-0.076
	(0.113)	(0.076)	(0.048)	(0.048)	(0.171)	(0.081)
Age	16.141**	24.170**	27.804**	31.521**	0.138	-1.438
	(4.616)	(4.190)	(2.839)	(3.173)	(4.779)	(2.547)
Partner/Spouse Present	-0.266	-0.073	-0.053	-0.012	-0.307	0.094
	(0.205)	(0.135)	(0.082)	(0.082)	(0.313)	(0.138)
Pregnant	-0.390**	-0.143*	-0.169**	-0.130**	-0.314	-0.114
	(0.144)	(0.084)	(0.052)	(0.051)	(0.205)	(0.086)
Obs.	7,679	7,430	18,655	14,925	26,046	50,480
Order	Wald	Wald	Wald	Wald	Linear	Linear
Bandwidth	{-1,0}	{-1,1}	[-2,2]	{-2,-1,1,2}	[-3,3]	[-3,12]
Threshold	Yes	No	Yes	No	Yes	Yes
Covariates	No	No	No	No	No	No

This table assesses the plausibility of the fuzzy regression discontinuity design by checking whether baseline covariates are similar on both sides of the treatment threshold (oldest child of school-starting age). Each cell reports the coefficient on in-borough shelter placement from a separate 2SLS regression of the row-delineated characteristic on the treatment indicator, using as the instrument an indicator for whether a family's oldest child's potential grade (end-of-calendar-year age year minus five) is zero or greater. The first four columns present Wald estimates (pooled instrumented mean comparisons), while the last two present linear estimates, allowing for different slopes on either side of the threshold. Within these groups, columns vary by bandwidth and whether the threshold itself is included. Standard errors clustered at family group level in parentheses. * $p < 0.10$, ** $p < 0.05$

Table A.21B: Regression Discontinuity Baseline Covariates

	Wald				Linear	
	(1)	(2)	(3)	(4)	(5)	(6)
Black	0.060	-0.035	-0.017	-0.049	0.344	0.165
	(0.219)	(0.150)	(0.093)	(0.093)	(0.347)	(0.157)
Hispanic	-0.104	-0.011	0.022	0.059	-0.405	-0.228
	(0.215)	(0.147)	(0.090)	(0.091)	(0.348)	(0.159)
White	0.042	0.040	-0.001	-0.015	0.071	0.007
	(0.068)	(0.047)	(0.028)	(0.027)	(0.104)	(0.046)
Asian	0.024	0.024	0.017	0.016	0.050	0.046**
	(0.026)	(0.020)	(0.012)	(0.012)	(0.044)	(0.020)
No Degree	0.278	0.089	-0.080	-0.136	0.365	-0.025
	(0.226)	(0.150)	(0.093)	(0.094)	(0.350)	(0.154)
High School Grad	-0.260	-0.322**	-0.083	-0.087	-0.331	-0.243
	(0.215)	(0.150)	(0.088)	(0.089)	(0.329)	(0.151)
Some College or More	0.049	0.131*	0.094**	0.110**	0.010	0.111*
	(0.093)	(0.068)	(0.040)	(0.042)	(0.140)	(0.067)
Unknown Education	-0.067	0.102	0.069	0.112**	-0.043	0.157**
	(0.102)	(0.073)	(0.043)	(0.045)	(0.156)	(0.075)
On Cash Assistance	0.161	0.137	0.010	-0.007	0.199	0.215
	(0.221)	(0.150)	(0.090)	(0.091)	(0.337)	(0.154)
On Food Stamps	0.142	0.048	-0.074	-0.124	0.331	0.123
	(0.193)	(0.131)	(0.080)	(0.082)	(0.311)	(0.134)
Employed Year Pre	0.086	-0.038	-0.017	-0.081	0.087	-0.092
	(0.225)	(0.153)	(0.094)	(0.095)	(0.351)	(0.164)
Log AQ Earnings Year Pre	0.931	0.215	0.887	0.513	0.662	-0.689
	(1.597)	(1.090)	(0.670)	(0.674)	(2.487)	(1.167)
Tier II Shelter	-0.282	-0.207	-0.326**	-0.338**	-0.720*	-0.392**
	(0.229)	(0.154)	(0.096)	(0.097)	(0.401)	(0.172)
Commercial Hotel	-0.028	-0.211	-0.245**	-0.279**	0.767**	0.199
	(0.201)	(0.137)	(0.085)	(0.085)	(0.390)	(0.161)
Family Cluster Unit	0.275*	0.405**	0.563**	0.614**	-0.085	0.195*
	(0.162)	(0.120)	(0.079)	(0.084)	(0.233)	(0.110)
Mahattan Shelter	-0.122	-0.275**	-0.475**	-0.531**	0.031	-0.404**
	(0.174)	(0.117)	(0.076)	(0.077)	(0.275)	(0.123)
Bronx Shelter	-0.026	0.178	0.406**	0.454**	-0.248	0.156
	(0.216)	(0.143)	(0.089)	(0.090)	(0.348)	(0.147)
Brooklyn Shelter	0.421**	0.302**	0.438**	0.431**	0.196	0.324**
	(0.206)	(0.136)	(0.085)	(0.086)	(0.295)	(0.137)
Queens Shelter	-0.303*	-0.227**	-0.358**	-0.342**	-0.040	-0.089
	(0.167)	(0.110)	(0.071)	(0.070)	(0.249)	(0.116)
Staten Island Shelter	0.029	0.023	-0.011	-0.011	0.060	0.013
	(0.039)	(0.027)	(0.017)	(0.017)	(0.063)	(0.027)
Obs.	7,679	7,430	18,655	14,925	26,046	50,480
Order	Wald	Wald	Wald	Wald	Linear	Linear
Bandwidth	{-1,0}	{-1,1}	[-2,2]	{-2,-1,1,2}	[-3,3]	[-3,12]
Threshold	Yes	No	Yes	No	Yes	Yes
Covariates	No	No	No	No	No	No

This table assesses the plausibility of the fuzzy regression discontinuity design by checking whether baseline covariates are similar on both sides of the treatment threshold (oldest child of school-starting age). Each cell reports the coefficient on in-borough shelter placement from a separate 2SLS regression of the row-delineated characteristic on the treatment indicator, using as the instrument an indicator for whether a family's oldest child's potential grade (end-of-calendar-year age year minus five) is zero or greater. The first four columns present Wald estimates (pooled instrumented mean comparisons), while the last two present linear estimates, allowing for different slopes on either side of the threshold. Within these groups, columns vary by bandwidth and whether the threshold itself is included. Standard errors clustered at family group level in parentheses. * $p < 0.10$, ** $p < 0.05$

Table A.22: Complier Characteristics: Ineligibility Rate Instrument

	Compliers	Non-Compliers	Diff.
Manhattan Origin	0.00	0.14	-0.13
	(0.003)	(0.000)	[-2.56]
Bronx Origin	0.57	0.39	0.18
	(0.006)	(0.000)	[2.28]
Brooklyn Origin	0.25	0.32	-0.07
	(0.005)	(0.000)	[-0.99]
Queens Origin	0.10	0.13	-0.02
	(0.003)	(0.000)	[-0.45]
Staten Island Origin	0.02	0.03	-0.00
	(0.000)	(0.000)	[-0.33]
Health Issue Present	0.33	0.30	0.04
	(0.004)	(0.000)	[0.61]
Eligibility: Eviction	0.29	0.34	-0.05
	(0.005)	(0.000)	[-0.67]
Eligibility: Overcrowding	0.16	0.18	-0.02
	(0.003)	(0.000)	[-0.34]
Eligibility: Conditions	0.11	0.08	0.03
	(0.002)	(0.000)	[0.67]
Eligibility: Domestic Violence	0.30	0.30	0.00
	(0.004)	(0.000)	[0.01]
Eligibility: Other	0.08	0.11	-0.03
	(0.003)	(0.000)	[-0.67]
Female	0.97	0.91	0.06
	(0.002)	(0.000)	[1.26]
Partner/Spouse Present	0.31	0.25	0.06
	(0.004)	(0.000)	[0.99]
Pregnant	0.04	0.07	-0.03
	(0.001)	(0.000)	[-0.86]
Black	0.43	0.57	-0.14
	(0.006)	(0.000)	[-1.79]
Hispanic	0.46	0.37	0.09
	(0.006)	(0.000)	[1.18]
White	0.06	0.02	0.04
	(0.001)	(0.000)	[1.57]
No Degree	0.61	0.57	0.05
	(0.005)	(0.000)	[0.67]
High School Grad	0.30	0.32	-0.02
	(0.005)	(0.000)	[-0.29]
Some College or More	0.06	0.05	0.01
	(0.001)	(0.000)	[0.21]
Unknown Education	0.02	0.07	-0.05
	(0.001)	(0.000)	[-1.27]
On Cash Assistance	0.30	0.36	-0.06
	(0.005)	(0.000)	[-0.78]
On Food Stamps	0.75	0.73	0.02
	(0.006)	(0.000)	[0.29]
Employed Year Pre	0.39	0.44	-0.05
	(0.005)	(0.000)	[-0.67]
Tier II Shelter	0.63	0.54	0.08
	(0.004)	(0.000)	[1.27]
Commercial Hotel	0.08	0.29	-0.21
	(0.005)	(0.000)	[-3.08]
Family Cluster Unit	0.19	0.16	0.02
	(0.003)	(0.000)	[0.46]
Family Size 1–3	0.54	0.64	-0.11
	(0.006)	(0.000)	[-1.43]
Family Size 4–5	0.39	0.28	0.11
	(0.004)	(0.000)	[1.68]
Family Size 6+	0.07	0.08	-0.01
	(0.002)	(0.000)	[-0.27]
Age	31.69	31.53	0.16
	(1.350)	(0.013)	[0.14]
Log AQ Earnings Year Pre	2.73	3.03	-0.30
	(0.259)	(0.002)	[-0.60]

Main sample. Treatment is in-borough placement. Instrument is 15-day moving average of the initial ineligibility rate for 30-day application period. Compliers are families placed in-borough when the ineligibility rate is high, but not otherwise. Non-compliers consist of always-takers and never-takers. Compiler and non-complier characteristics, adjusted for year and month of shelter entry, are estimated from the algorithm described in Cassidy (2020). Standard errors (in parentheses) and differences in means (with t-stats in brackets) are calculated from 200 bootstrap replications, clustering by family group.

Table A.23: Complier Characteristics: Aversion Ratio Instrument

	Compliers	Non-Compliers	Diff.
Manhattan Origin	0.09	0.13	-0.04
	(0.001)	(0.000)	[-1.18]
Bronx Origin	0.55	0.39	0.16
	(0.004)	(0.000)	[2.67]
Brooklyn Origin	0.20	0.33	-0.13
	(0.003)	(0.000)	[-2.29]
Queens Origin	0.12	0.13	-0.01
	(0.002)	(0.000)	[-0.17]
Staten Island Origin	0.03	0.03	0.01
	(0.000)	(0.000)	[0.51]
Health Issue Present	0.35	0.29	0.06
	(0.002)	(0.000)	[1.27]
Eligibility: Eviction	0.33	0.34	-0.01
	(0.003)	(0.000)	[-0.13]
Eligibility: Overcrowding	0.15	0.18	-0.03
	(0.002)	(0.000)	[-0.61]
Eligibility: Conditions	0.11	0.08	0.02
	(0.001)	(0.000)	[0.73]
Eligibility: Domestic Violence	0.27	0.30	-0.03
	(0.002)	(0.000)	[-0.59]
Eligibility: Other	0.10	0.11	-0.01
	(0.001)	(0.000)	[-0.28]
Female	0.94	0.91	0.03
	(0.001)	(0.000)	[0.82]
Partner/Spouse Present	0.28	0.25	0.03
	(0.002)	(0.000)	[0.61]
Pregnant	0.05	0.07	-0.02
	(0.001)	(0.000)	[-0.66]
Black	0.47	0.57	-0.09
	(0.004)	(0.000)	[-1.51]
Hispanic	0.43	0.37	0.06
	(0.003)	(0.000)	[0.98]
White	0.05	0.02	0.03
	(0.000)	(0.000)	[1.46]
No Degree	0.59	0.57	0.02
	(0.003)	(0.000)	[0.41]
High School Grad	0.31	0.32	-0.01
	(0.002)	(0.000)	[-0.21]
Some College or More	0.07	0.05	0.02
	(0.001)	(0.000)	[0.76]
Unknown Education	0.02	0.07	-0.04
	(0.001)	(0.000)	[-1.47]
On Cash Assistance	0.23	0.37	-0.14
	(0.003)	(0.000)	[-2.43]
On Food Stamps	0.70	0.74	-0.04
	(0.003)	(0.000)	[-0.63]
Employed Year Pre	0.39	0.44	-0.05
	(0.003)	(0.000)	[-0.91]
Tier II Shelter	0.59	0.55	0.04
	(0.003)	(0.000)	[0.82]
Commercial Hotel	0.14	0.29	-0.16
	(0.003)	(0.000)	[-2.83]
Family Cluster Unit	0.19	0.16	0.02
	(0.001)	(0.000)	[0.63]
Family Size 1–3	0.60	0.64	-0.04
	(0.003)	(0.000)	[-0.76]
Family Size 4–5	0.35	0.28	0.06
	(0.003)	(0.000)	[1.26]
Family Size 6+	0.05	0.08	-0.02
	(0.001)	(0.000)	[-0.82]
Age	32.20	31.47	0.73
	(0.949)	(0.014)	[0.74]
Log AQ Earnings Year Pre	2.76	3.03	-0.27
	(0.147)	(0.002)	[-0.71]

Main sample. Treatment is in-borough placement. Instrument is 15-day moving average of the aversion ratio. Compliers are families placed in-borough when the aversion ratio is high, but not otherwise. Non-compliers consist of always-takers and never-takers. Compiler and non-complier characteristics, adjusted for year and month of shelter entry, are estimated from the algorithm described in Cassidy (2020). Standard errors (in parentheses) and differences in means (with t-stats in brackets) are calculated from 200 bootstrap replications, clustering by family group.

Table A.24: Compliance Type Shares: Regression Discontinuity

	1%	1.5%	2%
Compliers	0.01	0.01	0.01
Always-Takers	0.67	0.67	0.67
Never-Takers	0.33	0.33	0.33

Main sample. Results from linear first-stage, controlling for year and month of shelter entry. Percentages in second row refer to percentiles used as thresholds to define low and high instrument values. See Cassidy (2020) for estimation method details.

Table A.25: Complier Characteristics: Regression Discontinuity

	Compliers	Non-Compliers	Diff.
Manhattan Origin	-0.03	0.13	-0.16
	(0.000)	(0.000)	[-13.42]
Bronx Origin	0.56	0.41	0.16
	(0.000)	(0.000)	[8.37]
Brooklyn Origin	0.50	0.31	0.19
	(0.000)	(0.000)	[9.94]
Queens Origin	-0.07	0.13	-0.20
	(0.000)	(0.000)	[-14.23]
Staten Island Origin	0.01	0.03	-0.02
	(0.000)	(0.000)	[-4.54]
Health Issue Present	0.28	0.30	-0.02
	(0.000)	(0.000)	[-1.69]
Eligibility: Eviction	0.26	0.34	-0.08
	(0.000)	(0.000)	[-4.89]
Eligibility: Overcrowding	0.16	0.18	-0.02
	(0.000)	(0.000)	[-1.38]
Eligibility: Conditions	0.07	0.08	-0.01
	(0.000)	(0.000)	[-1.66]
Eligibility: Domestic Violence	0.20	0.30	-0.10
	(0.000)	(0.000)	[-6.34]
Eligibility: Other	0.09	0.11	-0.02
	(0.000)	(0.000)	[-1.85]
Female	0.91	0.92	-0.01
	(0.000)	(0.000)	[-0.98]
Partner/Spouse Present	0.27	0.26	0.01
	(0.000)	(0.000)	[0.86]
Pregnant	0.08	0.07	0.01
	(0.000)	(0.000)	[0.94]
Black	0.58	0.56	0.02
	(0.000)	(0.000)	[1.27]
Hispanic	0.36	0.38	-0.02
	(0.000)	(0.000)	[-1.08]
White	0.02	0.03	-0.00
	(0.000)	(0.000)	[-0.83]
No Degree	0.59	0.57	0.02
	(0.000)	(0.000)	[0.88]
High School Grad	0.32	0.32	-0.00
	(0.000)	(0.000)	[-0.22]
Some College or More	0.06	0.05	0.01
	(0.000)	(0.000)	[1.14]
Unknown Education	0.04	0.06	-0.02
	(0.000)	(0.000)	[-2.09]
On Cash Assistance	0.39	0.35	0.03
	(0.000)	(0.000)	[1.90]
On Food Stamps	0.77	0.73	0.04
	(0.000)	(0.000)	[2.39]
Employed Year Pre	0.46	0.43	0.02
	(0.000)	(0.000)	[1.16]
Tier II Shelter	0.64	0.55	0.09
	(0.000)	(0.000)	[4.69]
Commercial Hotel	0.16	0.28	-0.11
	(0.000)	(0.000)	[-7.49]
Family Cluster Unit	0.08	0.17	-0.09
	(0.000)	(0.000)	[-5.68]
Family Size 1–3	0.86	0.63	0.23
	(0.000)	(0.000)	[10.71]
Family Size 4–5	0.20	0.29	-0.09
	(0.000)	(0.000)	[-5.05]
Family Size 6+	-0.00	0.08	-0.08
	(0.000)	(0.000)	[-5.23]
Age	27.87	32.38	-4.51
	(0.118)	(0.008)	[-12.74]
Log AQ Earnings Year Pre	3.06	3.00	0.06
	(0.017)	(0.001)	[0.45]

Main sample. Treatment is in-borough placement. Instrument is an indicator for whether a family's oldest child's potential grade is zero (kindergarten) or greater. Compliers are families placed in-borough when they have school-aged children, but not otherwise. Non-compliers consist of always-takers and never-takers. Compiler and non-complier characteristics, adjusted for year and month of shelter entry, are estimated from the algorithm described in Cassidy (2020). Standard errors (in parentheses) and differences in means (with t-stats in brackets) are calculated from 200 bootstrap replications, clustering by family group.

F Supplementary Figures

Figure A.1: Policy Instrument Time Series: Seasonally Detrended

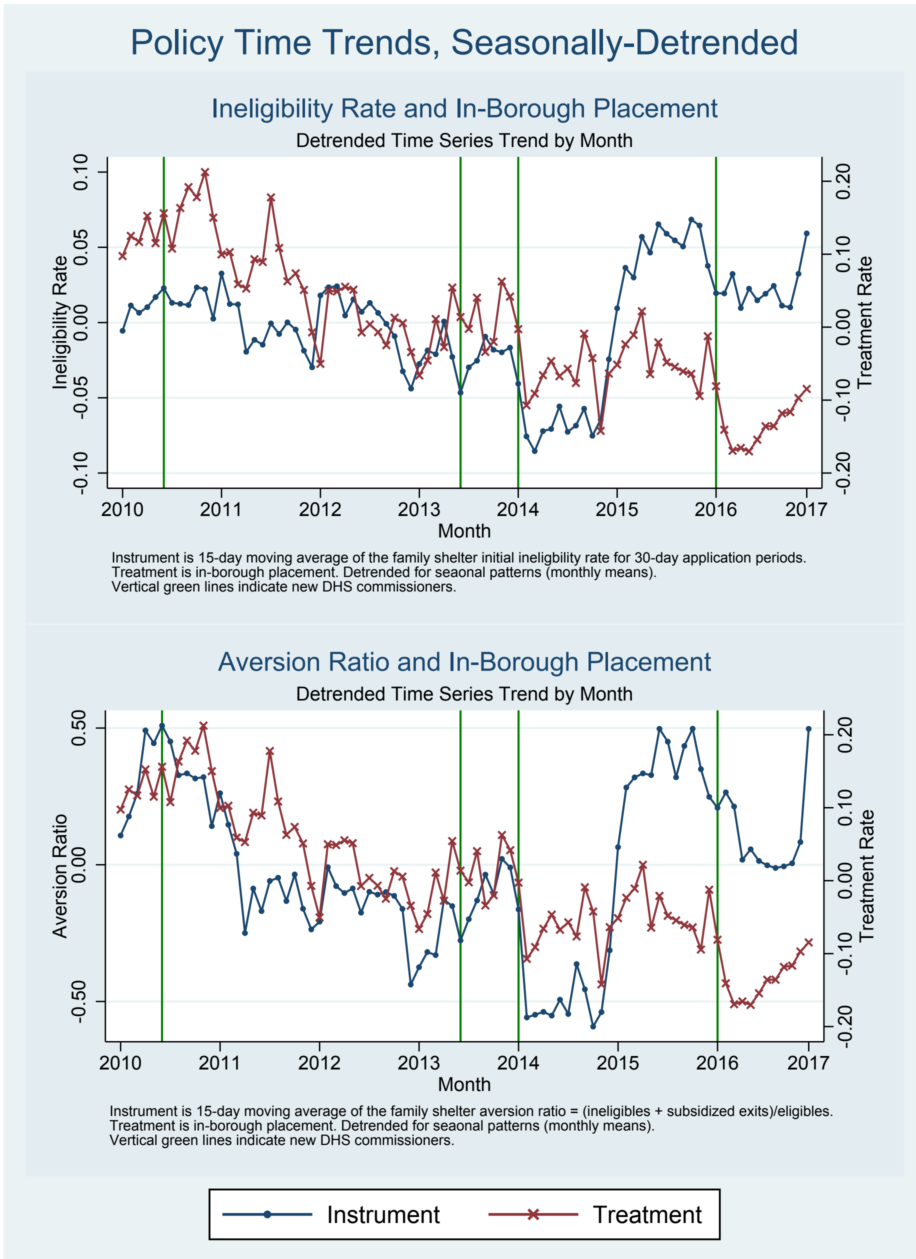


Figure A.2: Treatment, Length of Stay, and Policy Time Trends

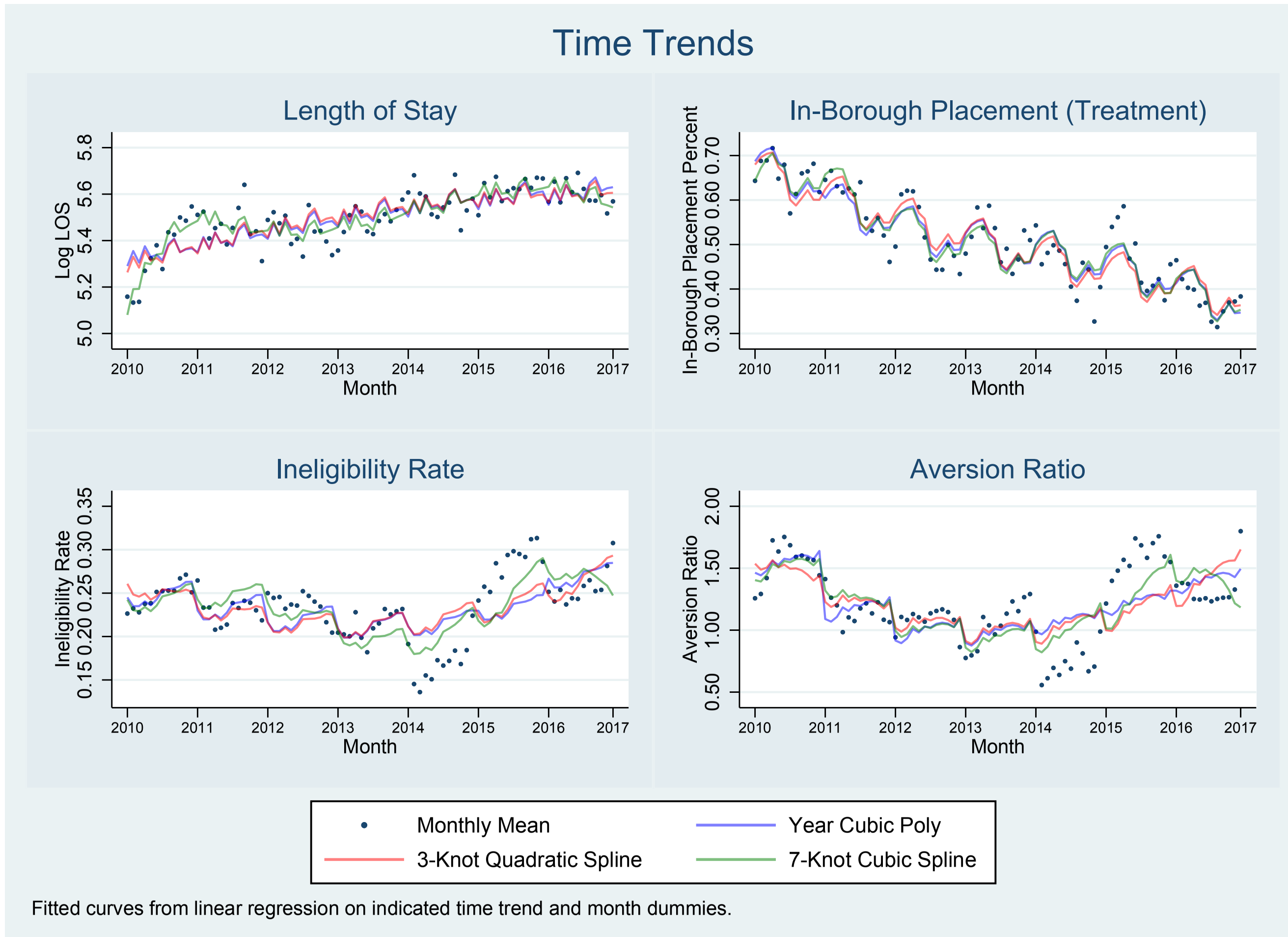


Figure A.3: Instrument First Stages and Time Trends

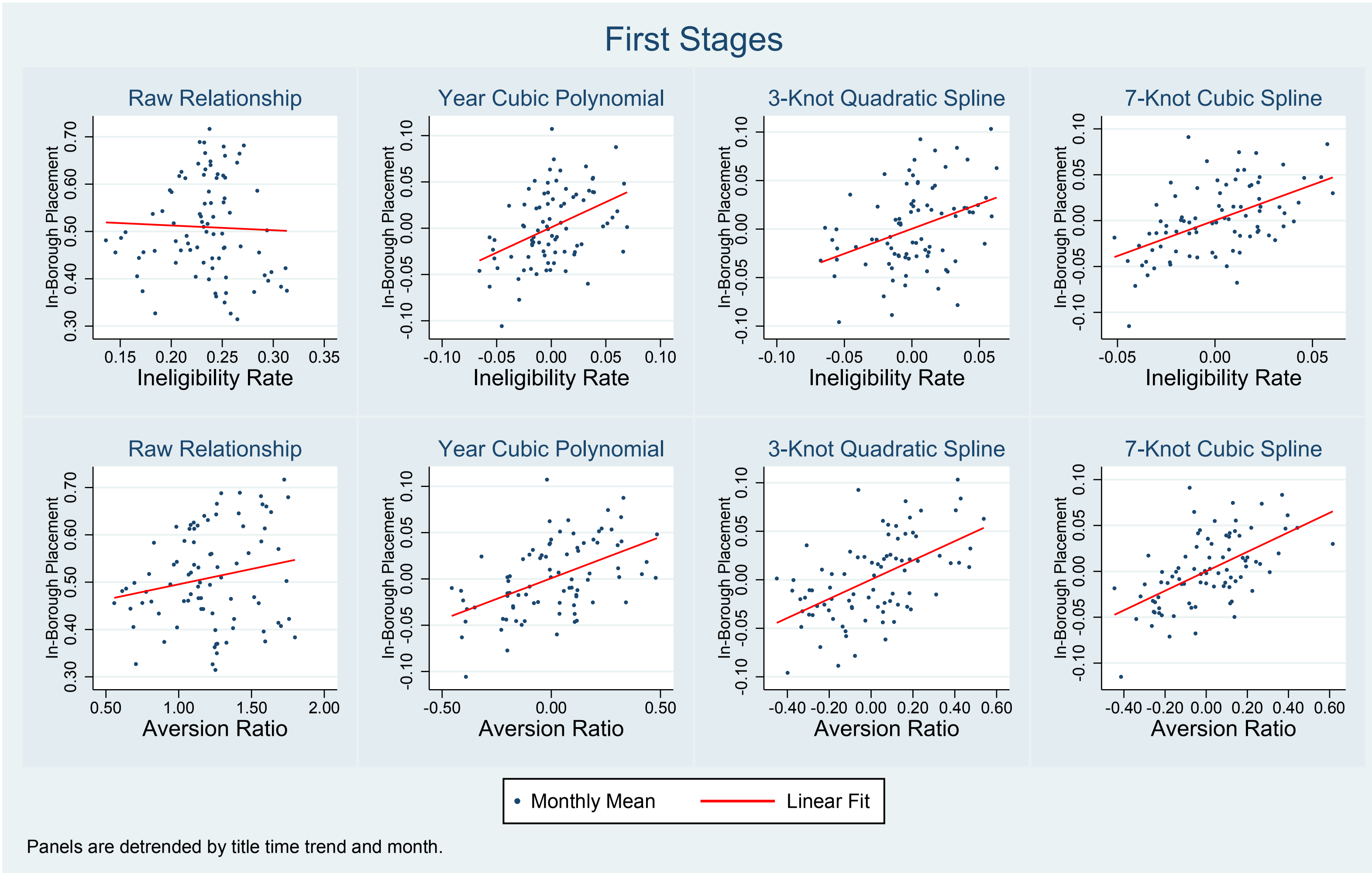
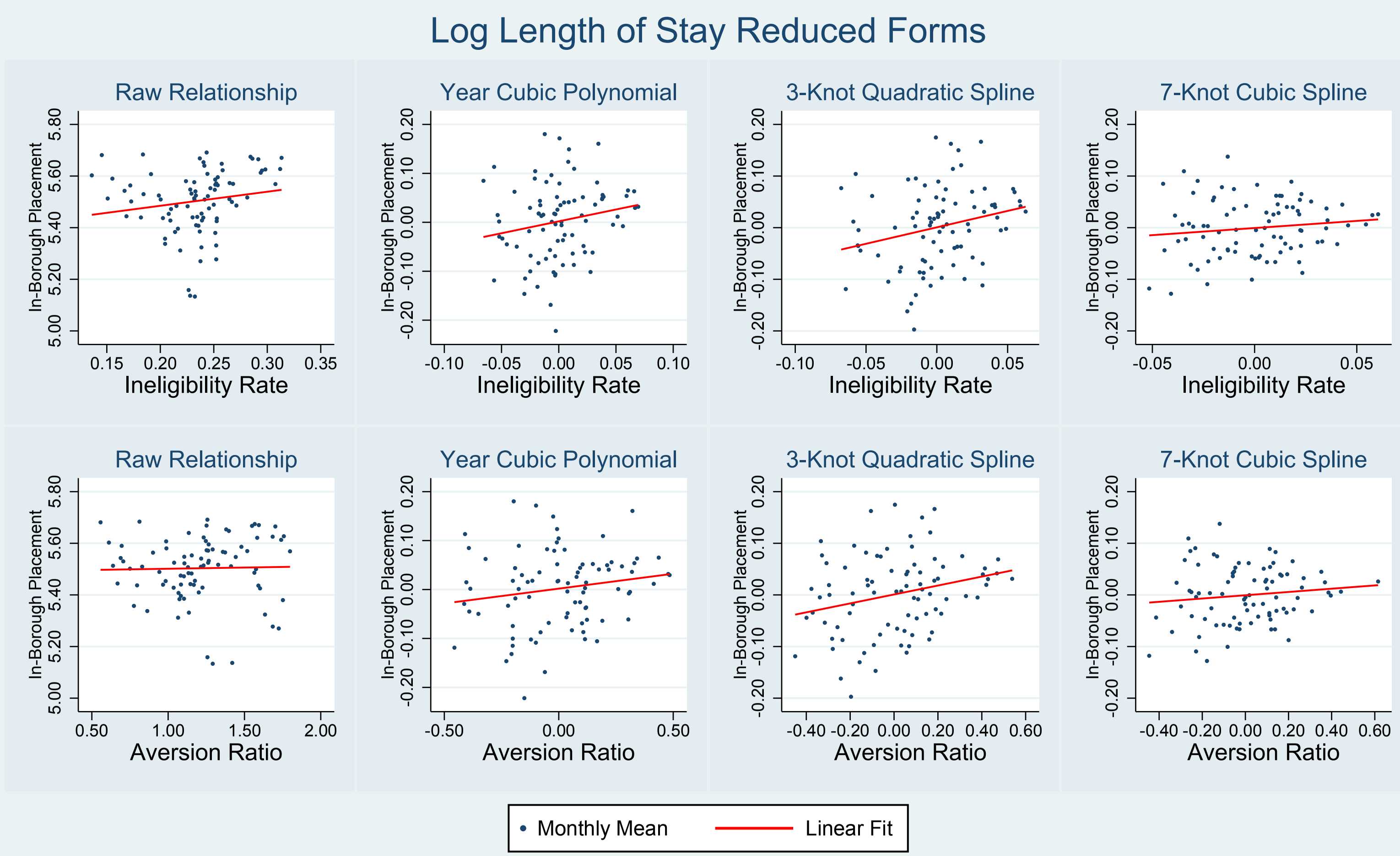


Figure A.4: Instrument Length of Stay Reduced Form and Time Trends



Panels are detrended by title time trend and month.

Figure A.5: Randomization Check

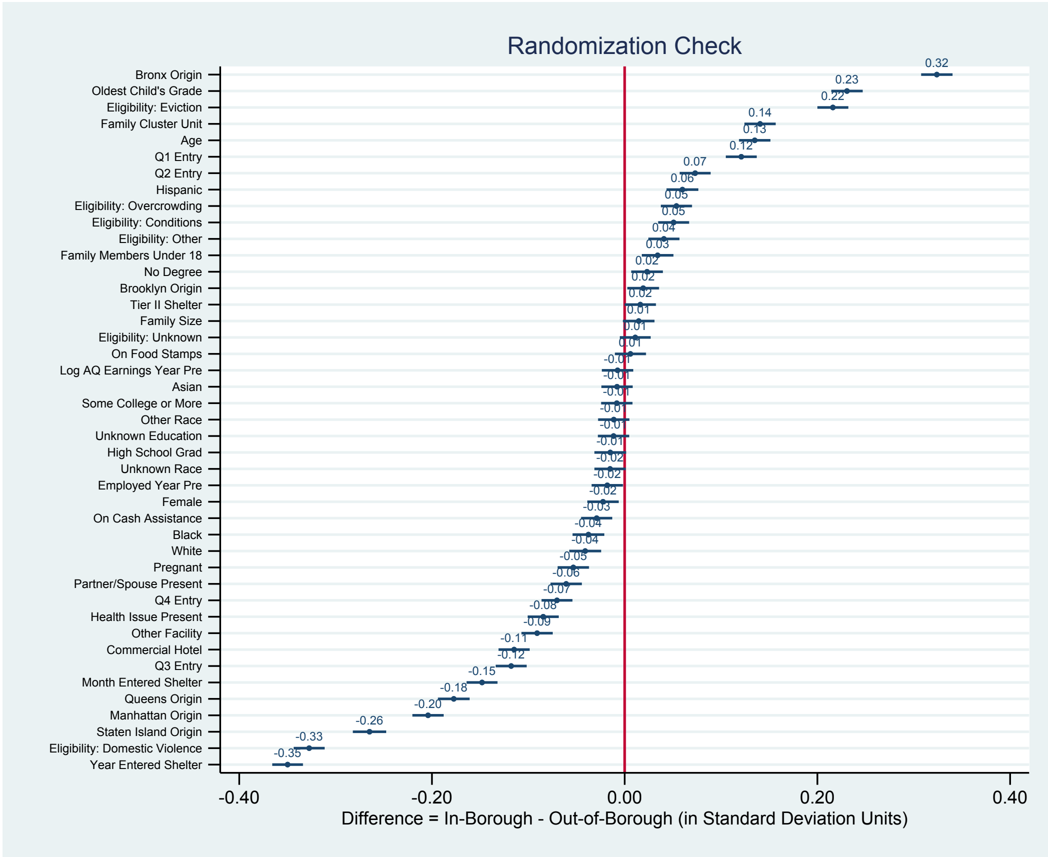


Figure A.6: Length of Stay Density

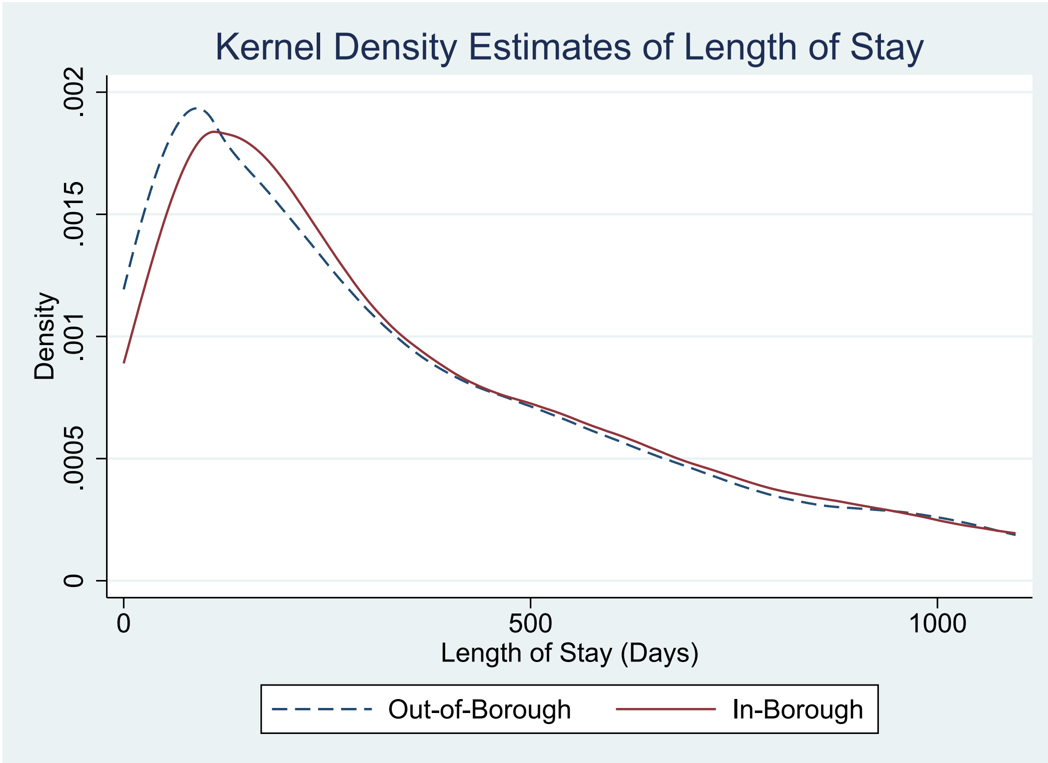


Figure A.7: Log Length of Stay Density

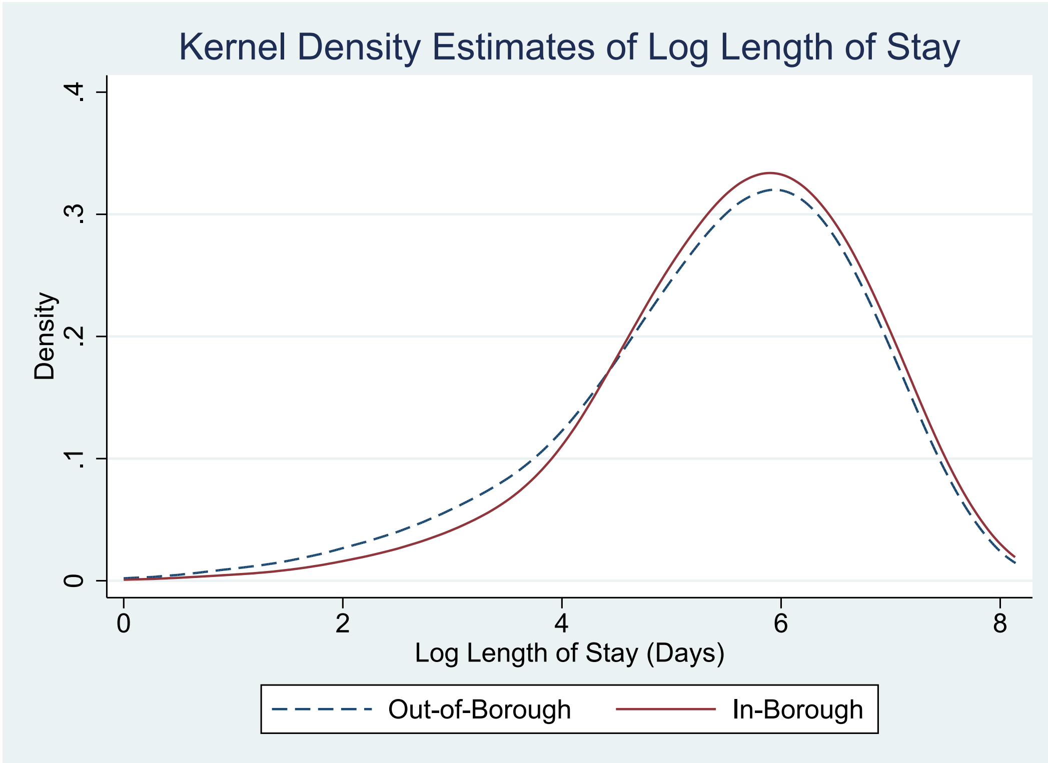


Figure A.8: Regression Discontinuity First Stages

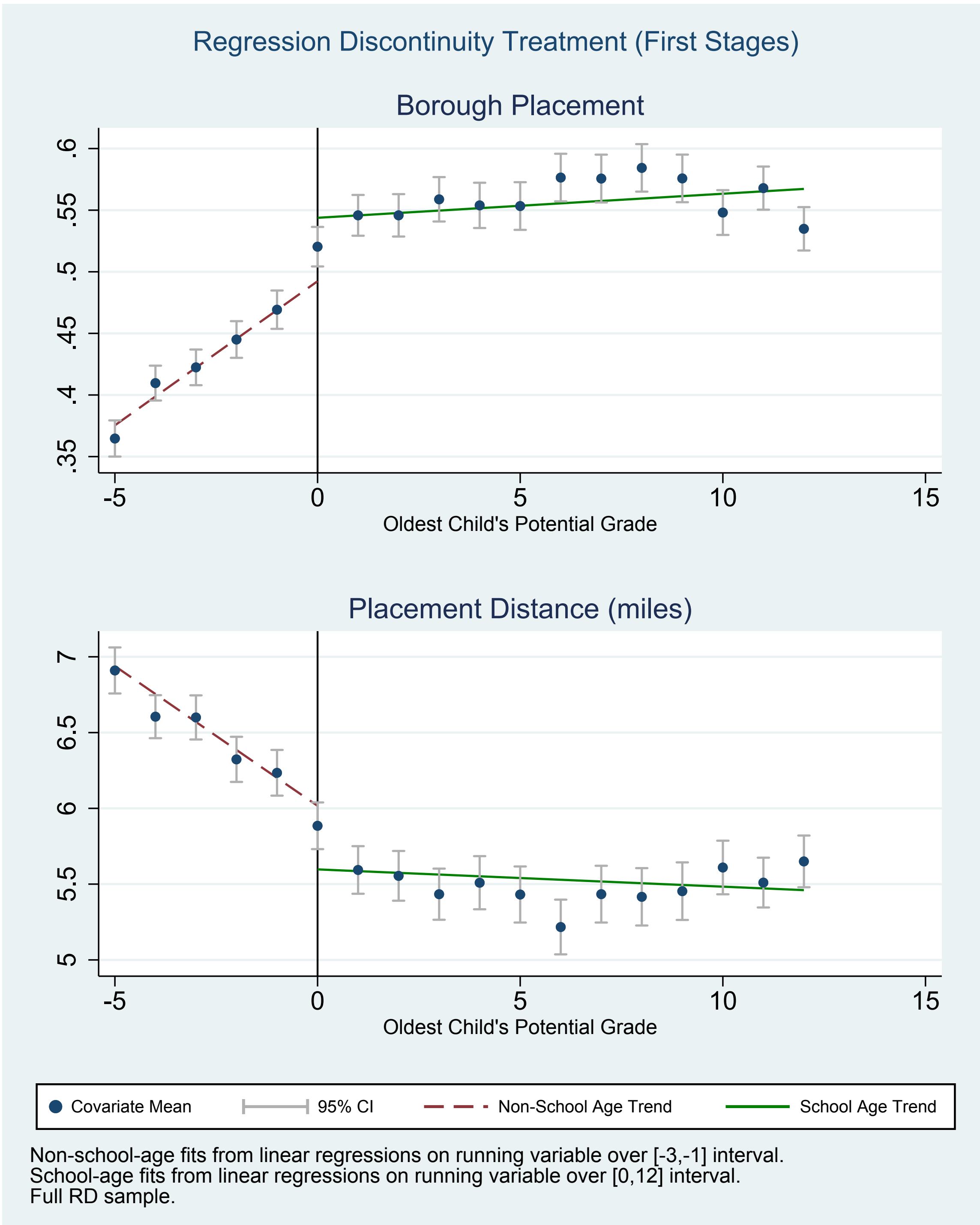
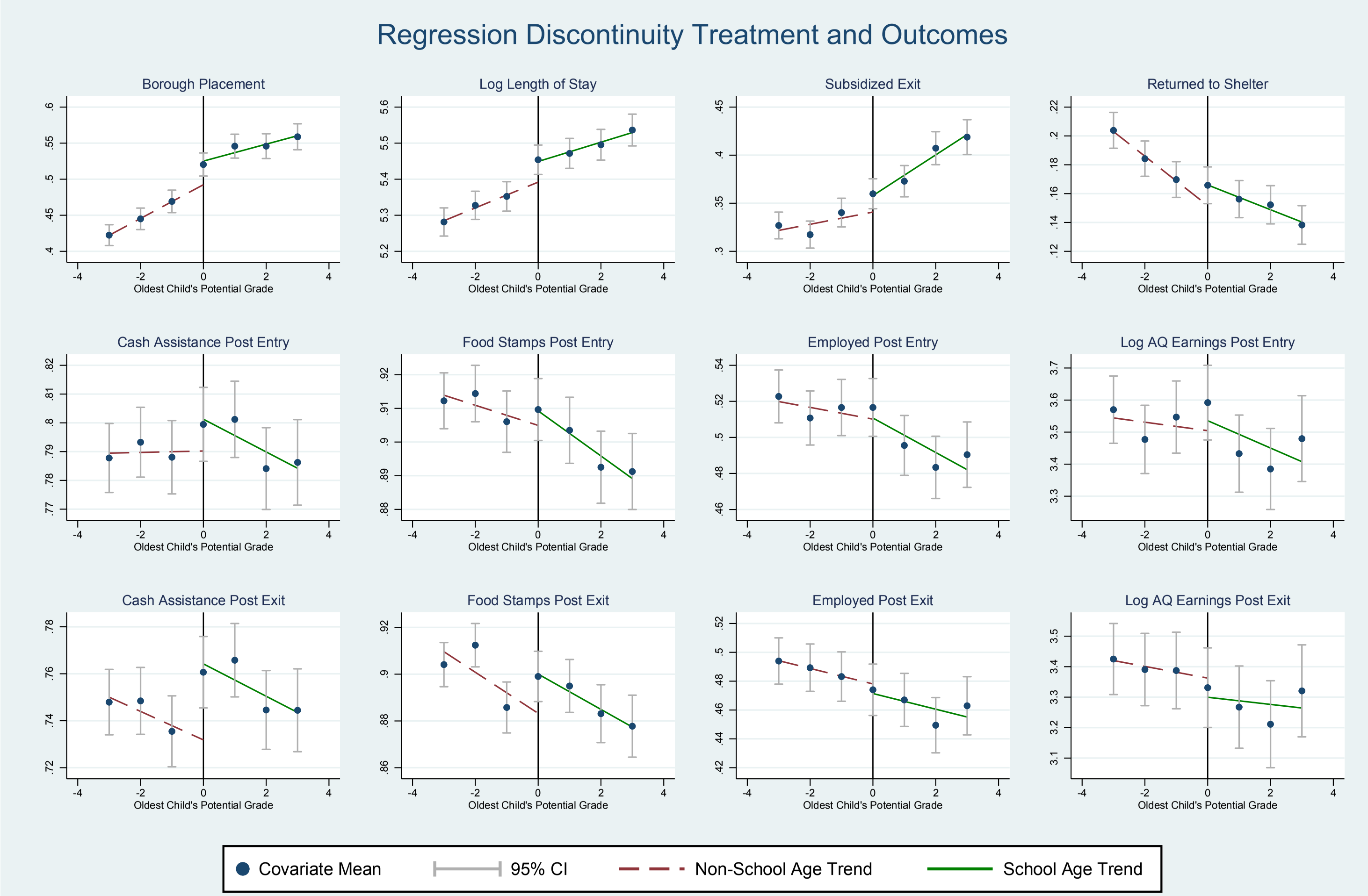
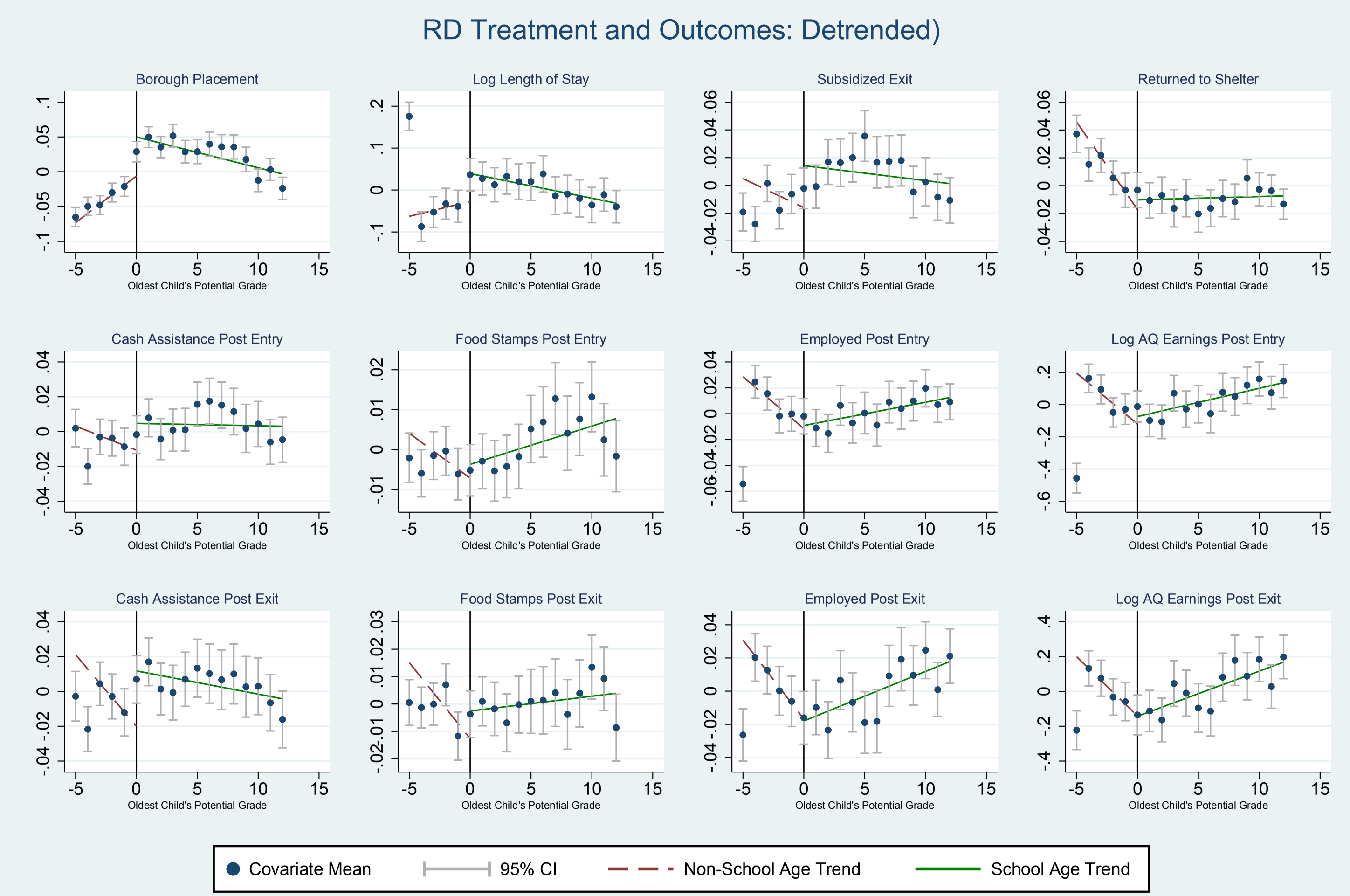


Figure A.9: Regression Discontinuity Treatment and Outcomes



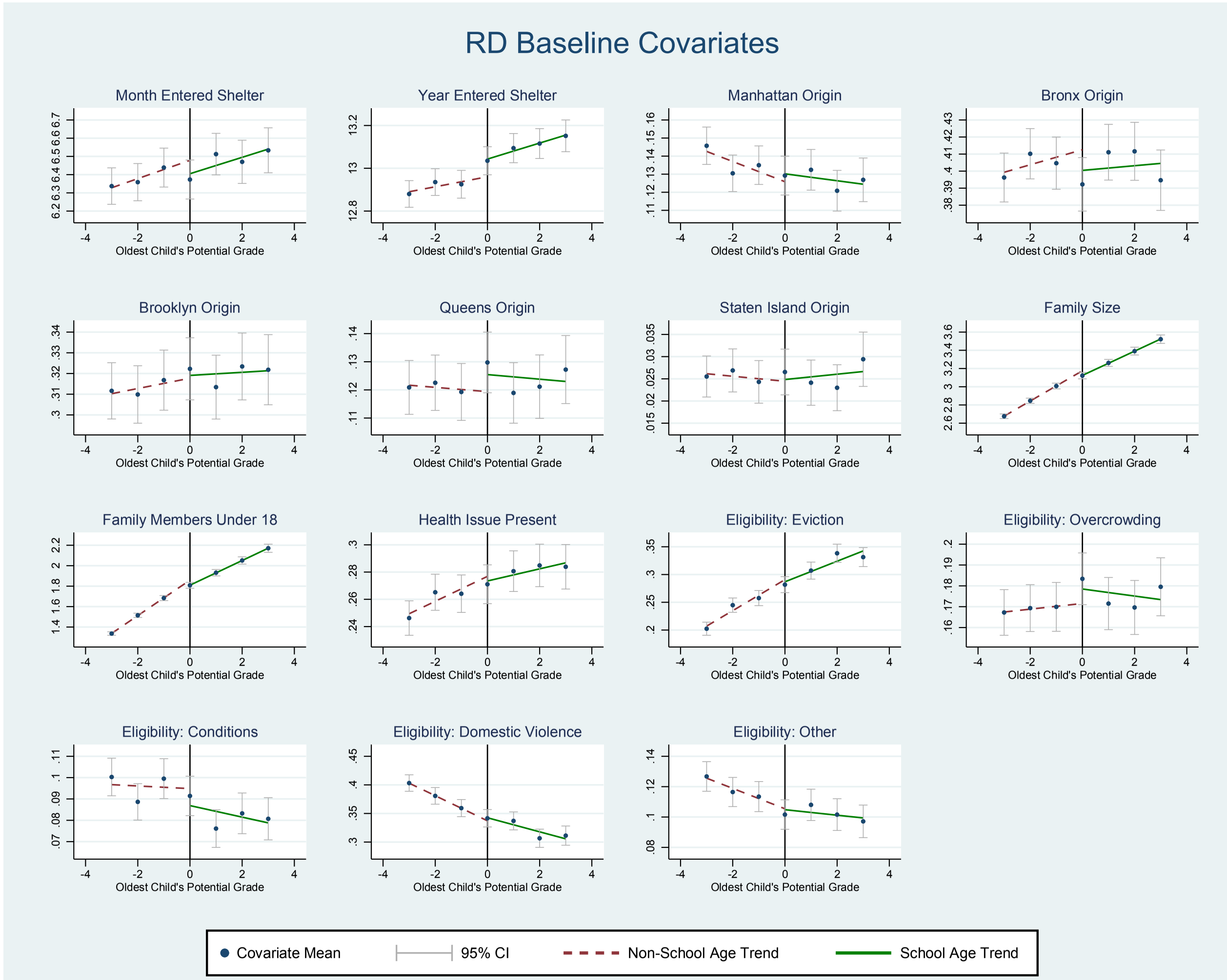
Non-school-age fits from linear regressions on running variable over [-3,-1] interval.
School-age fits from linear regressions on running variable over [0,3] interval.
Full RD sample.

Figure A.10: Regression Discontinuity Treatment and Outcomes: Detrended



Non-school-age fits from linear regressions of detrended outcome on running variable over [-3,-1] interval.
School-age fits from linear regressions of detrended outcome on running variable over [0,12] interval.
Detrended outcomes are residuals from a linear regression on Main RD covariates.
Full RD sample.

Figure A.11: Regression Discontinuity Baseline Covariates



Non-school-age fits from linear regressions of detrended outcome on running variable over [-3,-1] interval.
School-age fits from linear regressions of detrended outcome on running variable over [0,3] interval.
Full RD sample.

Figure A.12: Regression Discontinuity Baseline Covariates

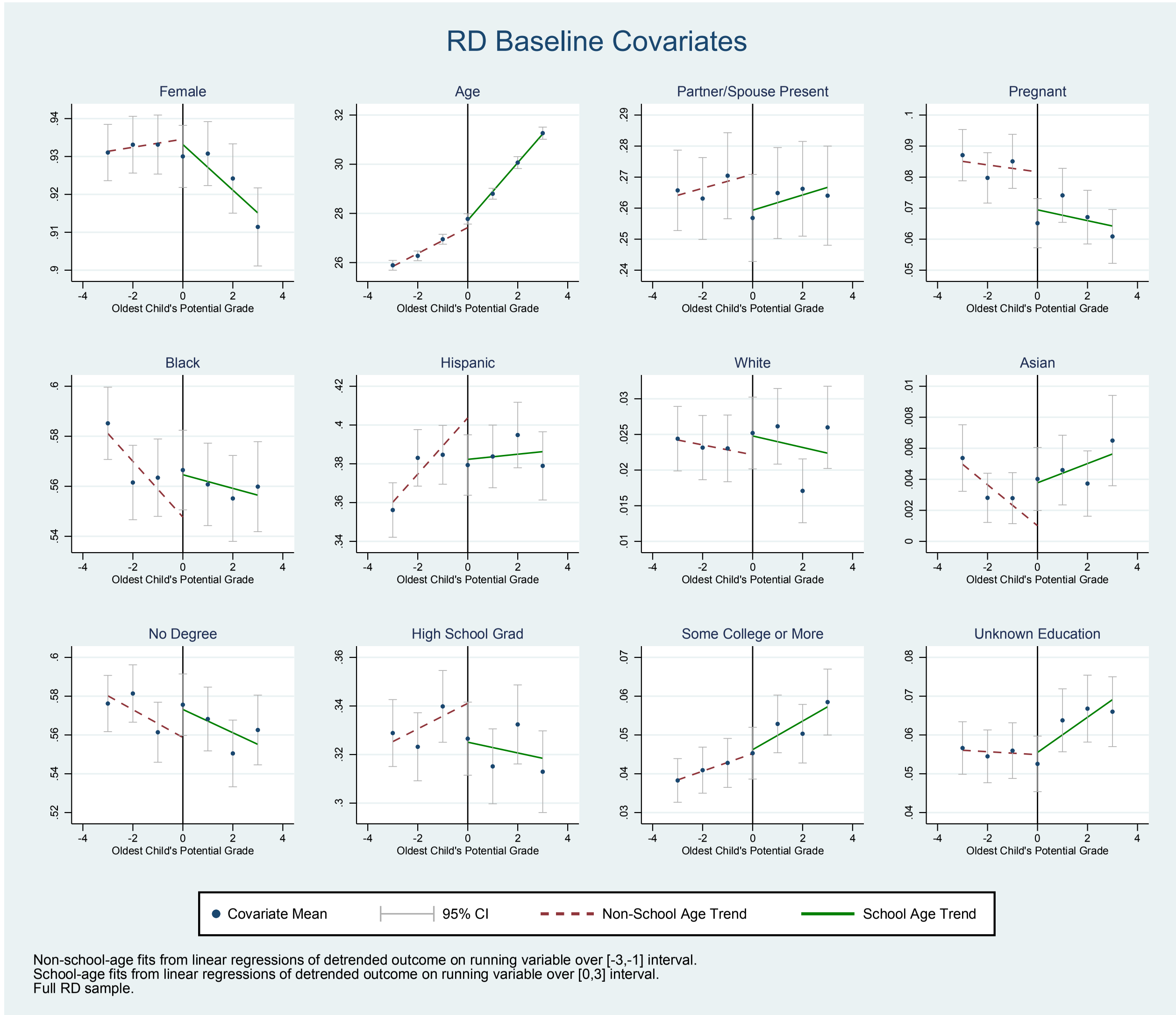
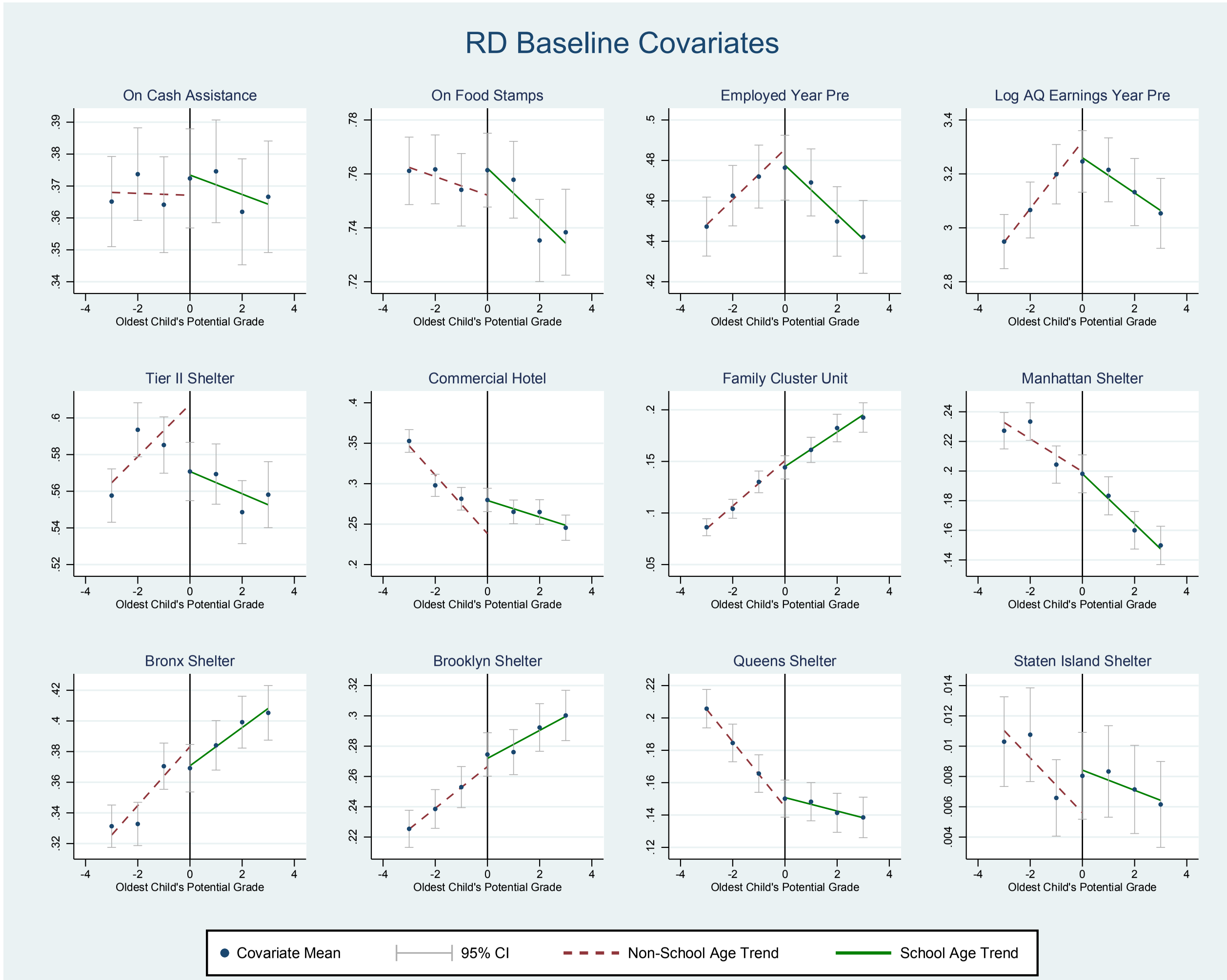


Figure A.13: Regression Discontinuity Baseline Covariates



Non-school-age fits from linear regressions of detrended outcome on running variable over [-3,-1] interval.
School-age fits from linear regressions of detrended outcome on running variable over [0,3] interval.
Full RD sample.