The Childcare Effect on Homelessness*

Christopher Babcock†

March 22, 2023

Most recent version available at:

https://christophermbabcock.com/babcockmaeconthesis/

Abstract

A difference-in-difference model is used to test if expansion of early education programs in New York City reduces population of homeless families with children in the affected age range. Results indicate that there is most likely a very minimal homeless reduction effect through introduction or expansion of early childhood education programs. These education policies should continue to be targeted to those they directly benefit (children themselves) and do not have a major secondary benefit on homeless reduction through increased employment opportunities for parents of affected children.

*I am grateful to Dan O'Flaherty for his guidance and support. He is truly the expert on the topic of homelessness. A thank you to Mike Cassidy for his mentorship on data collection and expertise navigating New York City's government publications and data portals. I'd also like to thank Shane Bordeau and Bernard Salanie for their counsel and recommendations throughout the M.A. Economics program at Columbia.

†Department of Economics, Columbia University. cb3724@columbia.edu

Contents

1 Introduction

On an average night in 2022, more than 580,000 individuals are homeless in the United States (U.S. Department of Housing and Urban Development 2022b). As the U.S. moved from an agrarian-dominated to an industrial, city-oriented society following the Civil War, the homeless segment of the country was born. First, as migrant workers or hobos romanticized in novels such as *The Grapes of Wrath*, "tramps" found seasonal work and traveled the country by rail (Jones 2015). Homeless levels peaked during the Great Depression, but with the New Deal legislation in the 1930s, sufficient affordable housing became available and homeless levels dropped. In the 1970's median home and rent, prices began increasing. By 1980, the annual increase in median rent prices was at an all-time high above 10% (iPropertyManagement 2022). High housing costs coupled with stagnating wage growth, (National Alliance to End Homelessness 2020) along with other compounding factors such as the gentrification of cities, the emergence of the AIDS epidemic, and a decrease in the U.S. Department of Housing and Urban Development (HUD) budget, ushered in the modern era of homelessness (National Academies of Sciences 2018).

1.1 Homelessness in New York City

Approaching 9 million residents, New York City, is the largest city in the United States. New York City presents a rich environment to study the homeless epidemic due to its sheer number of homeless, policies, and record keeping. The city has a disproportionate amount of homeless individuals relative to the rest of the country. Approximately 11% of the U.S. homeless singles and 30% of homeless families reside in New York City, while only 2.5% of the nation's population call the city home (NYC Department of Homeless Services 2022) (U.S. Department of Housing and Urban Development 2022b).

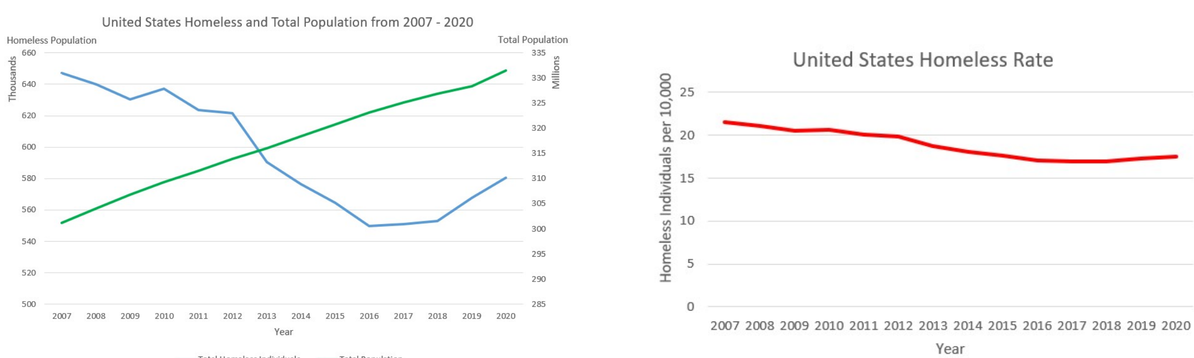
1.2 Hypothesis

According to a *NYC's Department of Homeless Services* (DHS) March 2023 daily homeless census, individuals in families with children account for over 45,000 of the city's 73,000 homeless population. Families with children amount to over 60% of the homeless individuals in the NYC shelter system. DHS organizes the ages of the individuals into eight groups. The largest group, children age 0-5, account for over 15% of the total homeless population. With the reduction of homeless families with young children in mind, this study hypothesizes that subsidized or publicly available childcare would allow low-income parents to focus more time on gaining or being employed. Greater employment opportunities may result in a decrease in the number of families with small children with respect to other homeless groups. The introduction of universal pre-kindergarten for 4-year-olds in 2014 and universal 3-K for 3-year-olds in 2019 in New York City provides a treatment effect for testing.

2 Literature Review

2.1 Overview of Modern Homelessness in the United States

HUD, the federal government agency responsible for homeless program funding, divides the country into 388 geographic districts called Continuums of Care (CoC) to administer care, utilize federal funding, and collect data. HUD uses a database called Homeless Management Information System (HMIS) to store individualized data on current or at-risk homeless persons reported through each CoC. Each CoC is required to conduct point-in-time (PIT) counts on a single night during the last 10 days in January of every odd-numbered year (U.S. Department of Housing and Urban Development 2022a). Most often with the help of volunteers who comb city streets, both sheltered and unsheltered homeless are counted. This gives a census of the entire nation's homeless population. Many CoCs also publish monthly or even daily counts for their sheltered individuals and families.



(a) US Homeless and Total Population

(b) US Homeless Rate

Figure 1: Macro Graphs of United States Homeless Population

HUD has published detailed yearly PIT counts since 2007 that include data from each CoC on various categorizations of a homeless population, such as race, gender, age, family status, etc. The most common way to measure homelessness is the percentage of homeless individuals in the overall population. Both the overall number of homeless and the homeless rate decreased from 2007 to 2018 from 21.4 homeless persons per 10,000 to 16.9. The absolute number of homeless individuals decreased as well during this period from 647,248 to 552,830, while the population of the United States was increasing from 301 million to 328 million. Since 2018, the homeless rate has risen back to 17.5 homeless persons per 10,000 (U.S. Department of Housing and Urban Development 2022b).

2.2 Proposed Solutions

Similar to younger siblings getting the used clothes and toys of their older family members, poor get their housing as hand-me-downs from the middle class (O'Flaherty 2019b). As middle class housing shrinks, a scarcity is created and homelessness rises. Hanratty 2017 supports this theory, determining that the most significant factors in determining area homelessness are a high percentage of property renters vs. property owners, a high poverty rate, and a high median monthly rent in that order. Piña and Pirog 2019 uses a three-tiered system to describe homeless prevention policy. First, ensure there is enough inventory and access to affordable housing for the poor. Second, for individuals which housing loss is imminent or just took place, prevent new entrants to homelessness through short-term rental assistance or rapid rehousing. Third, for individuals who are unable to secure housing, emergency solutions such as homeless shelters and transitional housing help people get back to a more permanent solution.

2.3 History of New York City Homelessness

One reason for the vast amount of homeless in New York City can be traced back to the reduction in single-room occupancy (SRO) housing. SRO units are generally cheaper bedrooms with shared kitchen and bathroom facilities for its residents. Today they are heavily regulated. Companies such as *Outpost Club* operate a few dozen around the city that boast modern kitchens and baths on each floor but allows individual rents to decrease, as they can fit more people in the building per square foot. Similar to "dorm-style" living, SROs provided ample affordable housing. In 1960 the city held 129,000 SRO units, but the city started levying restrictions. First, families were prohibited from renting SROs. Over the next two decades, changes in building regulations and tax codes disincentivized the construction and renovation of SROs. By 1979, only 23,000 SROs remained (Coalition for the Homeless n.d.).

In the 1950s, New York State also enacted policies deinstitutionalizing mental hospitals. This change was due to a myriad of factors, including improved medications and counseling, changing attitudes toward psychiatric centers, and cases of patient abuse. Between 1965 and 1980, there was a 68% decline in inpatient mental facilities (Coalition for the Homeless n.d.). The influx of these formerly discharged patients and declining low-cost housing options caused an overflow of housing demand with a dwindling supply. Some of these people had nowhere else to turn, and homelessness began to rise.

As the homeless population in New York City began to rise in the late 1970s, it became increasingly common to see people sleeping in subway stations, sidewalks, parks, and other public places. In 1979, right-to-shelter was established in the city as a result of *Callahan v. Carey*. This landmark decree involved attorney Robert Hayes, a recent graduate of New York University Law School, working his first trial case. With the help of Kim Hopper and Ellen Baxter, two Columbia University students studying homeless anthropology, Hayes filed a class action lawsuit against the current New York State governor, Hugh Carey (Frazier 2013). The principal plaintiff, Robert Callahan, was a homeless alcoholic and former cook. On December 5, 1979, as the cold weather was setting in, Hayes won his first case, obligating the city to provide shelter to all homeless men. Three years later, *Eldredge v. Koch* extended right-to-shelter from *Callahan* to women. And in 1986, *McCain v. Koch* established right-to-shelter for homeless families (Hayes 1987). With its right-to-shelter precedents, New York City is unique among municipalities. It is one of three cities (Washington D.C. and Salt Lake City) that have right-to-shelter for individuals and one of five areas (Washington D.C., Salt Lake City, Minneapolis, and the state of Massachusetts) that have right-to-shelter for homeless families (Hanratty 2017).

2.4 Motivation

Mike Cassidy's job market papers (Cassidy 2020a) and (Cassidy 2020b) focus on the effects of proximity of initial homeless family shelter placement from their origin neighborhood in New York City. First he argues that "proximity augments homeless students' education outcomes," including increased attendance and test scores. In the second he argues that proximity increases shelter length of stay, employment, and public assistance used. Combining ideas from these two papers, I hypothesized there might be a correlation between the populations of homeless groups and the disparate benefits certain groups were provided. If correct, this would provide treated and untreated groups from which a treatment effect could be ascertained. Family homeless shelters in New York City are not homogeneous; some offer services like medical and childcare onsite, while others do not. Another treatment effect consists of splitting the homeless into categories of those that received some type of assistance such as cash assistance, acceptance into a homeless prevention program, housing vouchers, or childcare vouchers. Parlaying the potential childcare effect on the population of homeless families, I decided to investigate a novel idea: if educational expansion by a municipality into earlier age groups leads to a decrease in the homeless family population or new entrants to homelessness with children in this age range.

Expanding enrollment to younger ages in the public school system than the current nationally mandated first grade (six years old) offers a form of childcare. The typical makeup of a homeless family is a single mother with children (Malik 2018). In the national political arena, universal pre-school has become a talking point for many candidates. Its main benefit is targeting educational attainment and employment outcomes for the child. A secondary benefit, I hypothesize, is this increases employment opportunities for the parents. Similar to any other form of childcare, a child in school at an earlier age allows an unemployed or partially employed parent the opportunity to focus more of their time and energy on securing employment. Greater employment outcomes may equate to fewer families becoming homeless. New York City is ripe for testing this theory. Not only was Pre-Kindergarten, Pre-K, or 4K (4-year-olds), as I will reference it in this paper, expanded in 2014, but a universal 3K (3-year-olds) program began in 2019, and its rollout is nearing completion across the city's school districts.

3 Data

3.1 Early Childhood Education in New York City

Fig. 2 shows a timeline of the expansion of 3K and 4K into the city's schools that is referenced on multiple city websites and by many reputable publications such as *The New York Times*. This figure is a bit confusing. The 19,000 Pre-K students referenced in the timeline prior to 2014 when Mayor Bill de Blasio implemented Universal Pre-Kindergarten (UPK), are the amount of 4K students enrolled at district schools. District schools are one of four 4K settings. The other three settings are Early Education Centers (NYCEEC), Pre-K Centers, and Family Child Care Home Based Programs (NYC Department of Education 2022a). Approximately 30,000 more students were enrolled in these other three 4K settings in 2014, which at the time were privately financed. In UPK's first year, de Blasio's initiative did increase 4K enrollment by more than 10,000 students in the district schools while also subsidizing the students in the other three settings. It also increased 4K enrollment rates as a percentage that attended kindergarten the following year from 65% to 90% in only two school years. The timeline's post-2014 numbers include children enrolled in all four settings, not simply district schools like the 2013 numbers display.

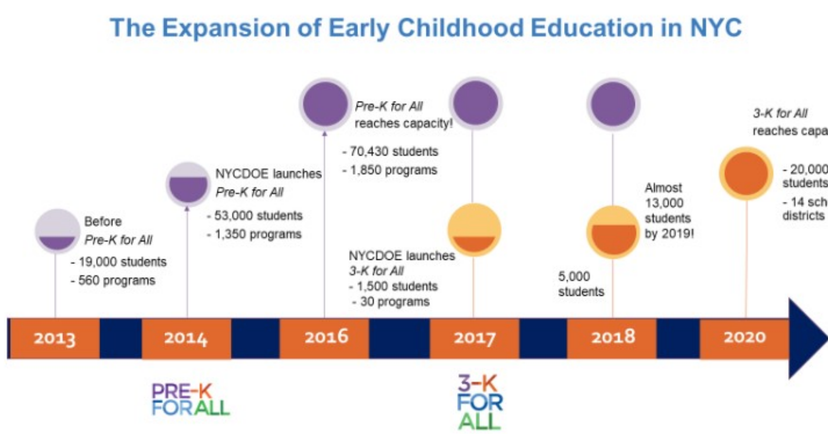


Figure 2: Early Childhood Education Expansion Timeline

Throughout his campaign, de Blasio touted UPK as a way not only to educate the young but give families a chance to defray childcare costs and increase their income through increased employment opportunities that come when a child is in school. To this end, parents could apply for two types of seats: full-day school year seats (6 hours and 20 minutes from September to June) or extended-day year-round seats (10 hours year-round) to ensure parents could maintain employment.

I obtained Pre-K enrollment data through publicly available data that the city publishes on *NYC Open Data*. School enrollment data is a snapshot taken of every school on October 31st of each school year. Every few years the city publishes a compilation of three to four years of school-specific enrollment demographic data. Schools are organized by District Borough Number (DBN). Districts are a two-digit number. Boroughs are a single character where K = Brooklyn, M = Manhattan, Q = Queens, R = Staten Island, X = Bronx. Finally, the last three digits of a

Figure 3: New York City Total School Enrollment By Grade for School Years from 2011 to 2022

Year	Total Enrollment	3-K Fraction	3-K Treatment	4-K Fraction	4-K Treatment	Grade 3K	Grade 4K (Half Day & Full Day)	Grade K	Grade 1
2011-12	1,025,221	0.00	0.00	0.26	0.00		22,208	81,031	83,033
2012-13	1,042,259	0.00	0.00	0.26	0.00		22,296	84,623	85,954
2013-14	1,104,479	0.00	0.00	0.65	0.00		55,734	85,375	89,630
2014-15	1,122,783	0.00	0.00	0.79	0.55		66,403	85,775	88,704
2015-16	1,133,963	0.00	0.00	0.87	0.88		71,847	84,386	87,804
2016-17	1,141,232	0.00	0.00	0.89	0.96		72,553	82,517	86,008
2017-18	1,135,334	0.01	0.02	0.90	1.00	824	70,704	81,588	83,578
2018-19	1,126,501	0.05	0.07	0.90	1.00	3,314	70,590	78,627	81,908
2019-20	1,131,868	0.29	0.42	0.97	1.00	17,586	69,894	78,587	79,171
2020-21	1,094,138	0.27	0.39	0.88	0.92	15,480	60,501	72,265	76,037
2021-22	1,058,888	0.61	0.87	0.83	0.73	34,740	57,315	68,828	70,546
		70% = fully treated		90% = fully treated					

DBN is the school-specific number within a certain district and borough. More recent years include more demographic characteristics than older years such as number of Native American students or economic need data. After merging all years into one dataset, I decomposed the DBN to calculate borough and citywide demographics of students in each grade for each school year from 2011 to 2022.

Fig. 3 shows the partial citywide table for enrollment for each school year. The columns in blue are the raw data from the dataset. The beige columns are the ones I created. Light beige shows the fraction of students that are enrolled as a percentage of that cohort in the advancing grade the following year. For example, 0.89 is circled in green in the "4-K Fraction" column for the 2016-17 school year. This is obtained by dividing the 2016-17 4K enrollment by the 2017-18 Kindergarten enrollment or $\frac{72{,}553}{81{,}558}$. This enables one to calculate the percentage of students enrolled in a certain grade as a percentage of that same group of students the following year. The two fractions in red at the bottom of the table are such because there is no next year to calculate from, so the current year's enrollment is used instead. The two red numbers for 4K enrollment in 2011-12 and 2012-13 are as described earlier - these numbers just report the 4K students enrolled at the district schools, whereas the following years starting in 2013-14, report students in all four 4K settings, not just in district schools. The dark beige columns depict a treatment effect percentage. Initially, in my calculations, I used a continuous treatment variable as depicted here. It appeared that even when

3K was available to anyone only 70% of students who would enroll in 4K the following year used the program. Thus, a 70% 3K usage rate was considered fully treated. I handled 4K differently because enrollment was already at about 65% even before de Blasio implemented UPK. Thus 65% 4K enrollment was the lower bound or no treatment up to 90% enrollment for full treatment. As I continued my research, I discovered that the rollout was done in poorer districts first and finished with higher-income districts. Thus my data analysis switched to a simpler binary value method of using 0 or 1 after a date cutoff instead of using a continuous treatment variable in my initial regressions.

3.2 New York City Homelessness

NYC Department of Homeless Services (DHS) posts "DHS Data Dashboard Tables" on their website from fiscal years 2012 to 2022. A one-year lag must be used to compare school enrollment data to DHS data. Fiscal years are measured from July 1, FY-1 to June 30, FY, where school years are measured from July 1, SY to June 30, SY+1. DHS Data Dashboard Tables were combined and cleaned into two databases, one which includes monthly data and the other yearly averages or totals. Both are organized into three major homeless category demographics: families with children, adult families, and single adults. At least over the past ten years, the average number of individuals inside each of these units has remained constant and will be treated as a fixed number for my analysis. These numbers are as follows: 3.1 individuals in a typical family with children, 2.1 individuals in a typical adult family, and 1 individual for all single adults. Most families with children include just a single adult, usually a mother with two children. Families with children include any family group with an individual below the age of 18 or a single pregnant mother. Adult families include all units greater than one individual in which all individuals are greater than age 18. Each month of the data set contains shelter population by age groups (0-5, 6-13, 14-17, 18-20, 21-29, 30-44, 45-64, and 65+) for the three demographic categories along with supplemental racial demographics. A majority of the data set is available for all 132 months. However, DHS did not report the same data every period. Other data reported for the three demographic categories include shelter intakes by borough for the first four years, shelter exits for the first six years, and placements out of shelter for the final six years.

Another source for my analysis was the Mayor's Management Report (MMR) published by the *NYC Mayor's Office of Operations*. It contains both budget information and valuable metrics on many programs including a very in-depth section on DHS. All of the reports since 1997 are published online at NYC Mayor's Office of Operations n.d. While the MMRs only have fiscal year data, unlike the monthly Data Dashboard, the MMRs have more useful metrics of the three categories of homeless, shelter exit types (such as subsidized or unsubsidized), and the average length of stay of each group every year, which DHS does not publish.

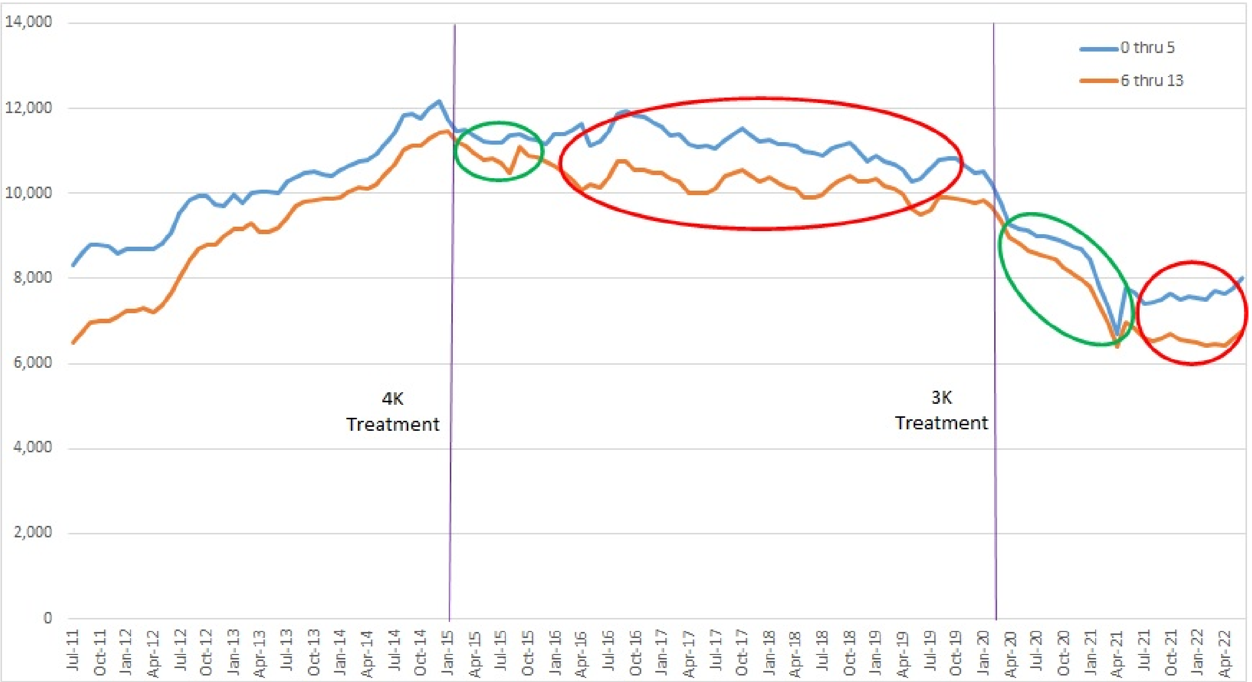
Fairly quickly, I concluded my results would not be very powerful without individual-level homeless data - specifically individual intake and applicant data. Because the publicly available data only has bins of individuals' ages (for example homeless children ages 0-5), I cannot link a treated or untreated family together. I cannot link siblings and parents of a family with a 4-year-old. Similarly, I cannot link an untreated family together (for example a family with a single 2-year-old). Without individualized data, I can only suppose that the group of 0-5 years is more treated than those aged 6-13. This latter group could have siblings that are age 4 (or age 3 for the 3K treatment). Because of this limitation, I manipulate the treated group in different regressions. Sometimes all of the age groups that are in the category "families with children" are treated as they could have a 3 or 4-year-old in their family while age groups from "adult families" are considered untreated since they cannot have a 3 or 4-year-old child. In other regressions, I take the 0-5 years as the treated and 6-13 as the untreated group simply because of an assumption that more of the 0-5 group should be treated than the 6-13 group.

Because of these data limitations, I began reaching out to the city in early October and filed a formal Freedom of Information Law (FOIL) request with DHS to obtain de-identified individualized intake and applicant data for the homeless shelter system between 2011 and 2022. Expect an update to this working paper in the summer of 2023 with the individualized data.

4 Model and Estimation Approach

Fig. 4 depicts the average daily shelter population of two age groups, 0-5 and 6-13, from July 2011 to June 2022. In this setup, the blue line, representing the 0-5 age bracket, is the treated group and the 6-13 age bracket in orange would be the untreated group. The purple vertical lines represent 4K and 3K treatment dates. Any treated group to the right of these lines has the

Figure 4: Average Daily Population of Homeless Individuals in Two Different Groups



respective treatment.

Looking just at green ovals on fig. 4 show what looks like a treatment effect as the two populations come closer together for the first year after treatment. After about a year, however, the red ovals display a discouraging result - either that the treatment wears off or there was no treatment to begin with.

One possible explanation of a non-persistent treatment effect is that U4K and U3K were implemented in low-income areas of the city. This gave primarily low-income mothers an advantage in the labor market where they could now seek employment if they were unable to before because they were caring for their youngest child. After the Pre-K programs are more established around the city in the second year, there is more competition for employment among this group. After year two, women with higher incomes and higher education are able to enter the workforce. I am not convinced that this explanation is correct, nor do I have data to substantiate this claim, but it is plausible. Regardless, in this case, the treatment effect does not persist.

Looking closely at fig. 4 reveals that both treatments are lagged from the dates the universal Pre-K programs took effect. Lagging the treatment effect should better account for the long and increasing length of stay of the homeless once they are admitted into the shelter system. When UPK was implemented through 4K and 3K programs, the average length of stay for homeless families with children was approximately 15 months. Even if the UPK programs caused a decrease in the intakes of homeless families, because of the length of stays of the current families already in the shelter system, changes to the shelter population would lag behind intake changes. This is discussed further in section 6.1.

Unfortunately, the publicly available macro data sorted by age groups lists the average daily shelter population and not shelter intakes. Yearly intakes for each category are published but not for the various age groups within the demographic categories. My initial regressions used *shelter population* as the dependent variable, but I also duplicated all regressions to simulate a flow by using the *change in monthly population* as a proxy for intake.

The initial population model is a difference-in-difference regression with an U4K or U3K timed treatment effect. The yearly model estimates the homeless population of a given category (family with children, adult family, single adult) and age group i, and year t,

$$\text{Shelter Population}_{it} = \beta_0 + \beta_1 P_t + \beta_2 T_i + \boldsymbol{\beta_3} \left(P_t \cdot T_i\right) + \beta_4 E_{ct} + \gamma_t + \epsilon_{it}$$

P : binary dummy variable for a time periods post-4K (or 3K) treatment

T : binary dummy variable for treated group

E : shelter exits in a specific homeless category c in year t

γ : year fixed effects

ϵ : error at time t for group i

β_1 : represents the difference between the post-treatment period and pre-treatment period for the untreated group

β_2 : represents the difference in the pre-treatment period between the treated and untreated groups

$\boldsymbol{\beta_3}$: represents the difference-in-differences between the treated and untreated, pre and post-treatment

β_4 : coefficient multiplied by the number of shelter exits from the category that i belongs

In this setup we look for β_3 coefficient to be negative and significant; in other words, there is a decrease in shelter population due to the treatment of UPK. As I proceed with variations on this regression, I include category fixed effects, changed P from a continuous to binary variable, and shifted P along the time continuum to study how the model was affected as the treatment month shifted.

The monthly regressions did not include year fixed effects but rather a time variable to remove any time trend.

The flow regression setup is,

$$\Delta ShelterPop_{it,Treated} = \beta_0 + \boldsymbol{\beta_1} P_t + \beta_2 \Delta ShelterPop_{it,Untreated} + \beta_3(Time) + \epsilon_{it}$$

5 Results

All variables are binary unless otherwise noted. The "month variable" is a linear variable that assigns 1 to the fist time period through n in the last time period. It is used to detrend the data.

5.1 4K Treatment

This section includes regressions run at the monthly level only. Appendix A.1 contains yearly regression results for the 4K treatment effect. Monthly 4K regressions were run across the 36 time periods from January 2013 - December 2015. 4K treatment was placed on and after January 2015, about four months after U4K was implemented.

5.1.1 Comparison Between Homeless Children in Families from 0-5 and 6-13

The regression in table 1 compares the groups of homeless children ages 0-5 and 6-13. 0-5 becomes the treated group, and 6-13 the untreated group. While it is likely only a small portion of the 0-5 group would be treated (all of the 4-year-olds), this group is directly affected. The 6-13 group is only indirectly affected since they would need a 4-year-old sibling. Besides the time-detrending month variable, this regression follows a standard difference-in-difference model. Where the 4K treatment timing and 4K treatment group (age 0-5) intersect, is the key variable. This shows a treatment effect at the 10% significance level of a reduction in the shelter population of 303 individuals in the treated group, post-treatment period.

The regression in table 2 is a monthly flow regression to simulate intake or change in the shelter

Table 1: Monthly DiD Comparison Between Sheltered Homeless Children in NYC Ages 0-5 and 6-13 from 2013-2015

	(1)
VARIABLES	Avg Shelter Population
4K Treatment	-641.53****
	(164.41)
Age Group 0-5	716.50****
	(94.87)
4K & 0-5	**-303.42****
	(164.31)
Month Variable	84.82****
	(6.46)
Constant	-44,877.07****
	(4,184.92)
Observations	72
R-squared	0.83

Standard errors in parentheses

**** p<0.01, *** p<0.05, ** p<0.10, * p<0.15

Table 2: Monthly Flow Comparison Between Sheltered Homeless Children in NYC Ages 0-5 and 6-13 from 2013-2015

	(1)	(2)
VARIABLES	$\Delta ShelterPopulation_{Treated}$	$\Delta ShelterPopulation_{Treated}$
4K Treatment	**-131.08*** **	**-232.81*** **
	(53.27)	(89.74)
$\Delta ShelterPopulation_{Untreated}$	0.36***	0.32***
	(0.15)	(0.15)
Month Variable		5.29
		(3.79)
Constant	66.99***	4.95
	(31.40)	(54.10)
Observations	36	36
R-squared	0.40	0.43

Standard errors in parentheses

**** p<0.01, *** p<0.05, ** p<0.10, * p<0.15

population. This shows a treatment effect of a reduction of 131 individuals (those in the 0-5 group) after the 4K treatment at the 5% significance level. Once the detrending time variable (Month) is added to this regression, it increases the weight of the 4K treatment to a decrease of 233 individuals in the treated group post-treatment.

5.1.2 Comparison Between Homeless Adults in Families with and without Children

Table 3: Monthly Flow Comparison Between Sheltered Homeless Families with and without Children in NYC Ages 21-29 and 30-44 from 2012-2015

	(1)	(2)
VARIABLES	$\Delta ShelterPopulation_{Treated}$	$\Delta ShelterPopulation_{Treated}$
4K Treatment	**-54.21**	**-89.51**
	(101.51)	(178.30)
$\Delta ShelterPopulation_{Untreated}$	1.89	1.92
	(2.28)	(2.32)
Month Variable		1.97
		(8.10)
Constant	41.31	16.39
	(64.25)	(121.65)
Observations	36	36
R-squared	0.03	0.03

Standard errors in parentheses
**** p<0.01, *** p<0.05, ** p<0.10, * p<0.15

The regression in table 3 compares groups of adults with indirect 4K effects (families with children) and groups with no 4K effects (adult families). The treated and untreated groups are the same age (21-29 and 30-44). These groups were chosen because, among those that have children, these are the age ranges that are most likely to have a 4-year-old child. This shows the same model as the previous table 2 above. With treated and untreated groups split differently in this regression, the treatment effects are not as strong. They are also not significant at even the 15% level. Looking at the bold numbers in table 3, it can be seen that 0 is inside the standard deviation for both tests. While this does show a negative treatment as expected, working with this subset of data produces a possibility that there is no treatment effect.

5.1.3 Comparison Between Various Age Groups in Families with and without Children

Table 4: Monthly DiD Comparison Between Sheltered Homeless Families in NYC from 2013-2015

VARIABLES	(1) Avg Shelter Population	(2) Avg Shelter Population
4K Treatment	1,623.00****	1,789.13****
	(327.16)	(342.66)
Families w/ Children	4,886.15****	8,433.18****
	(750.28)	(199.88)
Families w/ Children & 4K	**-1,303.32******	**-888.86*****
	(415.49)	(428.32)
Shelter Exits	5.59****	
	(1.14)	
Age Groups		
6-13	1,424.33****	1,607.95****
	(289.42)	(302.20)
21-29 w/ Children	-1,046.67****	-771.24****
	(207.08)	(209.83)
30-44 w/ Children	-763.95****	-488.52***
	(207.08)	(209.83)
Observations	216	216
R-squared	0.97	0.96

Standard errors in parentheses
**** p<0.01, *** p<0.05, ** p<0.10, * p<0.15

The regression in table 4 is similar to table 1 except that it contains age groups and shelter exits in the first regression. Shelter exits are the number of people from that particular category (families with children or adult families) that exit the shelter that month. The shelter exit coefficient is 5.59. The treated group is four of the age groups in families with children (0-5, 6-13, 21-29, and 30-44). The untreated group are adult families ages 21-29 and 30-44. The 4K treatment variable is strong and significant in both of these cases.

Table 5 shows a basic flow regression for the same treated group as table 4. Again, as with many of the flow regressions, the treatment variable is insignificant and R^2 value is extremely low. This is would be a poor model.

Table 5: Monthly Flow Comparison Between Sheltered Homeless Families in NYC from 2013-2015

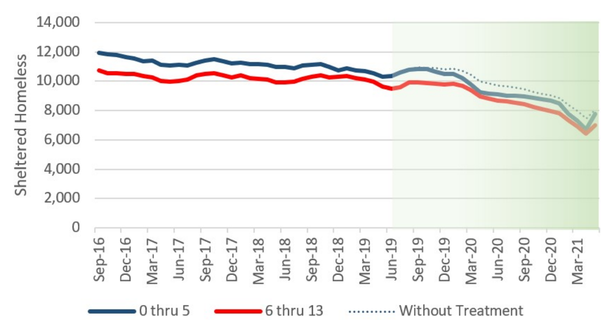
VARIABLES	(1) $\Delta ShelterPopulation_{Treated}$	(2) $\Delta ShelterPopulation_{Treated}$
4K Treatment	**-70.15**	**-69.20**
	(98.42)	(173.47)
$\Delta ShelterPopulation_{Untreated}$	0.21	0.21
	(2.20)	(2.24)
Month Variable		-0.05
		(7.88)
Constant	61.09	61.75
	(57.80)	(115.35)
Observations	36	36
R-squared	0.02	0.02

Standard errors in parentheses
**** p<0.01, *** p<0.05, ** p<0.10, * p<0.15

5.2 3K Treatment

Monthly 3K regressions were run across the 27 time periods from January 2019 - March 2021. 3K treatment was placed on and after March 2020, about six months after U3K was implemented. In general the 3K affects were stronger and more significant than the 4K effects.

Figure 5: Implementation of 3K and a Counterfactual



5.2.1 Comparison Between Homeless Children in Families from 0-5 and 6-13

Table 6 returns to a treated group of 0-5 and untreated group of 6-13. There is a significant reduction at the 5% level of 394 individuals from the treatment. The month variable being negative shows that during this period, shelter population is decreasing.

Table 6: Monthly DiD Comparison Between Sheltered Homeless Children in NYC Ages 0-5 and 6-13 from 2019-2021

VARIABLES	(1) Avg Shelter Population
3K Treatment	-1,189.57****
	(168.73)
Age Group 0-5	864.58****
	(94.38)
3K & 0-5	**-393.72*** **
	(181.90)
Month Variable	-28.72****
	(4.20)
Constant	30,273.87****
	(2,951.41)
Observations	104
R-squared	0.87

Standard errors in parentheses
**** p<0.01, *** p<0.05, ** p<0.10, * p<0.15

5.2.2 Comparison Between Homeless Adults in Families with and without Children

Interestingly, I was not able to find table 7's counterpart on the 4K level that would give a significant result for 4K treatment. This table gives the most significant result of all regressions thus far and strongly supports a 3K treatment effect by comparing the same age groups from 21-44 of those with and without children. A flow regression with these same parameters yielded a decreasing flow treatment effect but the results were not significant at the 15% level.

5.2.3 Comparison Between Various Age Groups in Families with and without Children

Table 8 gives the most convincing argument for the treatment effect. It has the most observations, a high R^2, and in every regression, the treatment variable is significant. The treatment variables also greatly reduce the homeless population of the treated group. Placements are a new variable in this regression. In the later years of my data set, the DOE switched to placements versus exists. Placements imply that the shelter system is assisting the family to leave the shelter and is a subset of exits in which a family could be assisted or leave on their own. In the second

Table 7: Monthly DiD Comparison Between Sheltered Homeless Families with and without Children in NYC Ages 21-29 and 30-44 from 2019-2021

VARIABLES	(1) Avg Shelter Population
3K Treatment	17.36
	(82.43)
Families w/ Children	6,063.05****
	(46.11)
Families w/ Children & 4K	**-1,213.95******
	(88.86)
Month Variable	-13.74****
	(2.05)
Age Group 30-44	296.72****
	(39.41)
Constant	10,925.32****
	(1,441.93)
Observations	208
R-squared	0.99

Standard errors in parentheses
**** p<0.01, *** p<0.05, ** p<0.10, * p<0.15

Table 8: Monthly DiD Comparison Between Sheltered Homeless Families in NYC from 2019-2021

VARIABLES	(1) Avg Shelter Population	(2) Avg Shelter Population	(3) Avg Shelter Population
3K Treatment	145.37*	147.22*	-339.80****
	(95.73)	(96.02)	(88.46)
Families w/ Children	6,502.17****	6,141.05****	6,141.05****
	(217.89)	(55.93)	(62.69)
Families w/ Children & 3K	**-1,590.10******	**-1,503.66******	**-1,503.66******
	(108.74)	(96.65)	(108.34)
Placements	-0.52**		
	(0.30)		
Month Variable	-18.95****	-18.73****	
	(2.10)	(2.10)	
Age Groups			
6-13	-758.58****	-758.58****	-758.58****
	(69.79)	(70.01)	(78.48)
21-29 w/ Children	-3,414.69****	-3,414.69****	-3,414.69****
	(65.28)	(65.49)	(73.41)
30-44 w/ Children	-3,117.97****	-3,117.97****	-3,117.97****
	(65.28)	(65.49)	(73.41)
Constant	18,028.86****	17,848.80****	4,689.84****
	(1,479.00)	(1,479.97)	(82.01)
Observations	312	312	312
R-squared	0.99	0.99	0.99

Standard errors in parentheses
**** p<0.01, *** p<0.05, ** p<0.10, * p<0.15

regression, placements were removed, and in the third, the time-month variable was also removed. The treatment variable stays significant, which is a good sign.

Table 9: Monthly Flow Comparison Between Sheltered Homeless Families in NYC from 2019-2021

VARIABLES	(1) $\Delta ShelterPopulation_{Treated}$	(2) $\Delta ShelterPopulation_{Treated}$
3K Treatment	**-72.28***	**-67.70**
	(46.51)	(82.10)
$\Delta ShelterPopulation_{Untreated}$	1.66**	1.63*
	(0.89)	(1.05)
Month Variable		-0.41
		(5.99)
Constant	-16.85	-13.97
	(28.05)	(50.90)
Observations	27	27
R-squared	0.34	0.34

Standard errors in parentheses
**** p<0.01, *** p<0.05, ** p<0.10, * p<0.15

Table 9 contains the last flow regression table that is a bit more explanatory at 34% for R^2 than some of the previous flow tables. The 3K treatment is significant at the 15% level. Once we add a time trend line though to this regression, all of the coefficients become less significant.

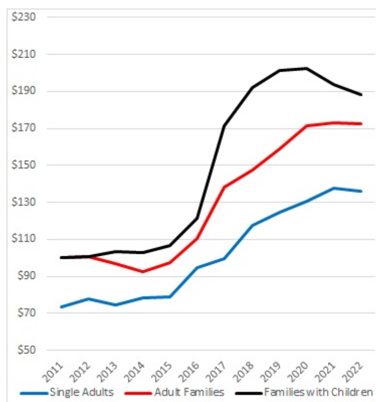
5.3 Overview of Results

My hypothesis did not prove to be correct in every case. At least 100 regressions were run of various forms, and although in many cases the treatment effect aligned with my hypothesis as both negative and significant, there were also many cases in which the treatment effect was close to 0 and insignificant. I was only able to find two cases where the treatment effect was positive and significant by manipulating the variables. This leads me to believe there is probably a small reduction in the homeless population due to the expansion of early childhood education.

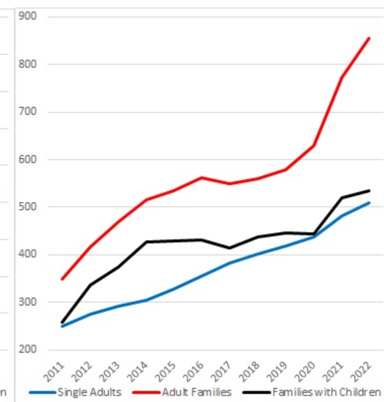
6 Cost-Benefit Analysis

6.1 Cost of New York City Homeless Shelters

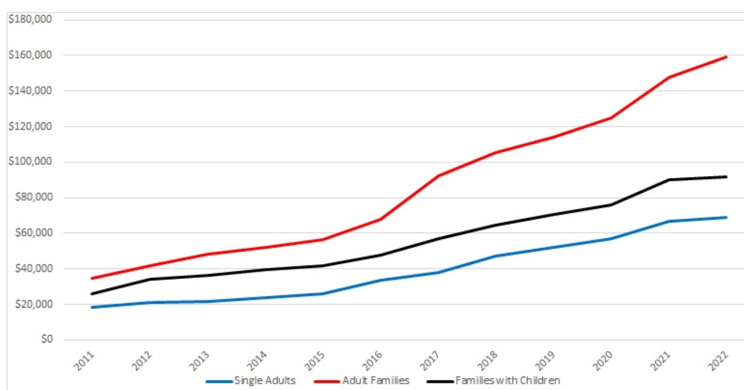
The total cost of the homeless shelter system levied to the taxpayer is composed of two metrics - the average cost per night multiplied by the current shelter population. These two measures can be further decomposed into even more fundamental statistics. The nightly unit shelter cost is determined by facility costs and administrative efficiency. Shelter population is determined through the length of shelter stay and the number of new entrants to the system.



(a) Unit Shelter Cost per Night



(b) Length of Stay



(c) Shelter Entry Costs

Figure 6: Shelter Category Statistics

New York City annual *Mayor's Management Report* (MMR) (NYC Mayor's Office of Operations n.d.) is used for homeless shelter expense data provided in this this subsection. All years in fig. 6 and 7 are New York City fiscal years. A fiscal year runs from July 1, FY-1 to June 30, FY. So fiscal year 2011 is July 1, 2010 - June 30, 2011. In 2010, the cost per night was approximately $70 for a single adult and $100 for a family. By 2020, these costs had doubled. Fig. 6a shows that the growth wasn't constant, but ballooned in the four years between 2016 and 2019 to an annual growth rate of 18.5% (compound annual growth rate will be used for all growth rates unless otherwise noted). For the eleven years between 2011 and 2022, the annual growth rate in unit shelter costs per night is 5.8%. To give context, this far outpaces inflation and median rent growth in New York City, but is about half the growth rate of the S&P 500. The inflation rate hovered around 2% during the entire period with the exception of the last two years in which it spiked to 4.5% and 7% respectively as we emerged from COVID-19 (World Bank 2021). The average growth rate of the S&P 500 over this period was approximately 11% (S&P Dow Jones Indices LLC 2022).

Mayor Bill de Blasio (2014-2021) was very critical of the education and homelessness policies of the previous two administrations (Rudy Giuliani 1994-2001 and Michael Bloomberg 2001-2013). Prior to his mayoral run, as New York City's Public Advocate, he opposed a housing voucher cut and created a bad landlord watchlist. Almost immediately after he was elected, he implemented a program for homeless to obtain identification cards. With the homeless population continuing to rise through his tenure, his policies quickly began to face criticism even among his supporters, for instance when he spent millions in mental health programs to aid the homeless in 2015 (Grynbaum and Stewart 2015). He doubled down on his promises and 2016 saw shelter city inspections ramp up as "Shelter Repair Scorecards" were used to publicly report on every city shelter. de Blasio voiced his "determin[ation] to give every family and individual in a homeless shelter decent living conditions" (NYC Office of the Mayor 2016).

Money was certainly being spent at a higher rate than before on the homeless. de Blasio's 2017 MMR (NYC Mayor's Office of Operations n.d.), attributed the per night shelter cost increase shown in fig. 6a to "security costs; enhancements, including mental health services and program activities in adult shelters, cost of living adjustments for contract providers, utilization of commercial hotels to accommodate increases in entrants, and a reduction in the use of "cluster" apartment-type shelters, as well as the opening of new shelters, which generally have higher rent and fixed costs than do older, established facilities."

O'Flaherty 2019a notes that most of the shelter expansion was complete by 2015. So while some of the increase in per capita shelter cost was definitely due to facility upgrades and higher operating costs, administrative efficiency almost certainly plays a role. Even NYC Office of the Mayor 2016 admits that many shelters had relatively few violations prior to de Blasio's big push to ensure "decent shelter conditions". Before major shelter repairs were undertaken, family shelters averaged about 0.5 violations per living area, roughly the same as all New York City buildings.

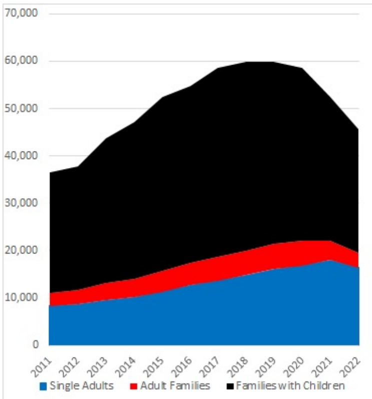
Over the last decade, fig. 6b shows the length of stay of those entering the shelter system increased from eight to 17 months for both single adults and families with children and from 12 to 29 months for adult families. There are many plausible explanations for why individuals and families opt to increase their shelter stay. Some explanations give rise to moral hazard. As shelters improve, they may outpace the quality of residences that could be rented by the poor. Another explanation supposes an increase in housing subsidies upon exiting shelters from an administration that is attempting to improve homeless conditions. Individuals and families may be willing to wait their turn in the shelter for these more lucrative subsidies instead of moving on without them. A third explanation may be the decline of unskilled labor positions. A final reason may be that rising rents kept shelter residents from exiting sooner. This is not likely, however, since from 2011 to 2019, median New York City rents were stable and only grew at inflationary rates (Engebreth 2020) before exploding the last two years following the COVID-19 pandemic, reaching annual growth rates of 30% (Zumper Inc. 2022).

One area in which the city government should be commended is the declining recidivism rates among all three categories over the last decade. In 2013 when the MMR began reporting this statistic, the percentage of shelter residents who returned within a year of exit was 22%, 15%, and 10% for single adults, adult families, and families with children respectively. By 2022, all categories dropped to 10%, 2%, and 4%. While I do not have the data regarding re-entries greater than a year after shelter exit, this is a promising sign that individuals that exit are able to live in a more stable situation.

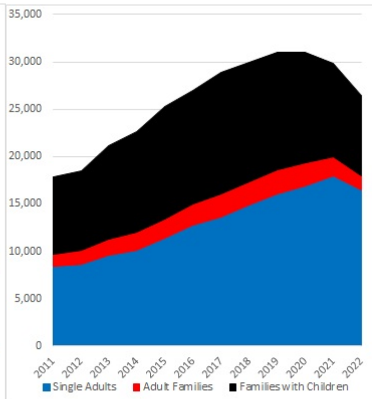
Combining the unit shelter cost per night and the average length of stay gives shelter entry cost for an individual or family entering the shelter system (see fig. 6c). In 2011, a typical family with children that entered the shelter system cost the taxpayer $25,831 for their time in a city shelter. In 2022, this same family costs $92,014 to host in a city shelter. An even more drastic increase is money spent per adult family, rising from $34,942 in 2011 to $159,049 in 2022. The lowest-increasing category, single adults, had their entry cost increase almost four-fold from $18,395 in 2011 to $69,137 by 2022. The staggering growth rates of these costs have been relatively constant at 13% per year. To put it another way, when a case investigator at the intake office makes a determination that a family with children has no other housing options, she obligates approximately $100,000 for that single family's entry. A decade ago that same case worker only committed $25,000 for a positive homelessness determination for a family with children.

Fig. 7 displays cumulative totals for New York City shelter services. Fig. 7a charts individuals while fig. 7b and fig. 7c organize data by category unit (number of single adults, adult families, and families with children). Remember that adult families and families with children average 2.1 and 3.1 individuals per unit respectively.

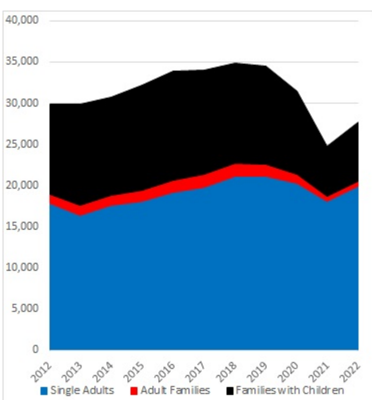
Fig. 7a and 7b display the homeless shelter population from 2011 to 2022. Overtime the average daily number of individuals in a New York City homeless shelter rose from 36,460 in 2011 to its peak of 59,953 in 2018 and has since decreased to 45,564 in 2022.



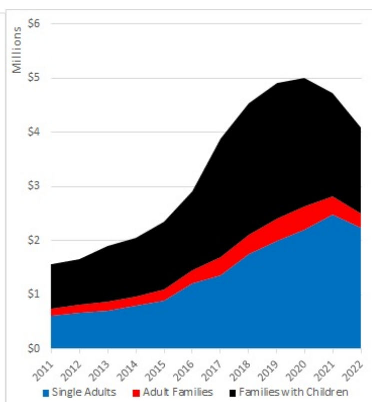
(a) Average Daily Shelter Population by Individual



(b) Average Daily Unit Shelter Population by Category



(c) New Shelter Entrants per Year



(d) Total Shelter Cost per Night

Figure 7: Shelter Category Totals

Fig. 7c displays the flow of individuals and families into the homeless system each year. Intakes rose moderately until 2018. For the next three years, intakes dropped precipitously, coinciding with the COVID-19 pandemic. Overall, the shelter system is admitting less individuals in 2022 than they did in 2011. The timing of the decrease in shelter population, shelter intakes, and shelter costs per night is heavily correlated with the timing of the COVID-19 pandemic. Intakes had actually begun a moderate decrease prior to the pandemic but it appears to have been attenuated. The initial, moderate decrease is likely due to burgeoning length of stay and per night costs. O'Flaherty 2019a would say the "front door is harder to open" as fewer people are admitted into the shelter system. Interestingly, because the length of stay has continued to rise, the average daily population in the shelters has increased substantially (see fig. 7b). The amount exiting the shelter spigot has slowed, so, while intakes decline, the shelter population rises. Intake levels are already increasing again, up 11% from 2021.

Finally, the total city-wide shelter cost per night is depicted in fig. 7d by multiplying the nightly unit cost in fig. 6a and average daily shelter population in fig. 7b. Both of the two composite values have increased, so shelter costs have greatly increased as well. Overall, the compound annual growth rate of shelter costs from 2011 to 2022 is 9.1%. Costs for single adults have increased at a higher rate of 12.4% per year, while family costs have risen by 6.3% annually. Each night the city spends $4.1 million operating its homeless shelters for a yearly cost of $1.5 billion. This is up from the $1.5 million per night and $572 million per year the city spent on homeless shelter operation in 2011.

Many of the cost growth rates associated with New York City's homeless shelter system are triple or quadruple inflation rates. This is unsustainable and indicative of poor policy design and/or implementation. If the shelter system is broken, then a short-lived cash infusion is probably needed, followed by sustainable growth in maintaining the quality of service. Incentives in this arena are very complicated because the higher quality or easier subsidies are to obtain, the greater the chance services will be abused. A balance must be struck between providing shelter for those in need while also incentivizing that same group to provide for themselves. Instead of as many shelter improvements, there may be benefits to shifting some of this budget towards increased housing subsidies (or making more subsidies available) for those low-income earners who have the potential to become homeless. This not only decreases the number showing up at the intake center but also has the effect of lowering the demand for shelter by decreasing shelter quality (or at least not increasing as much). This suppresses both lengths of stay and cost per person per night. What is needed is a dignified, temperate, and clean place to sleep. The desire for privacy, a larger area than just a bed, and the pride of having a home should motivate many to seek a permanent solution.

It is important to note that any policy that is successful at decreasing the number of individuals that may potentially become homeless or decreasing applicants to the homeless system (prevention programs), will not necessarily decrease the number of homeless entrants themselves. Similar to yearly fluctuations in applicants to a college not affecting the number of offers or enrollees, decreasing the number of applicants doesn't necessarily decrease the intakes and ultimately the costs. If DHS case managers operate on a quota system similar to university admissions, then unless the applicants decrease below that quota, entrants will not decrease. Similarly, if there are objective metrics for accepting or denying shelter entry, a given policy that reduces applicants still might not have the desired effect of decreasing entrants. For example, a housing subsidy increase as part of a homeless prevention program may disproportionally target the applicants that were already getting denied shelter entry before the new policy went into effect. The applicant pool decreases, but those individuals weren't going to be allowed to enter the shelter anyway. Only a policy that affects individuals that would be accepted into the shelter system is cost-effective. The loose or stiff "front door" concept greatly depends on the administration and political party in charge. In this case, it ultimately comes down to a single individual: the New York City mayor. In this study, Mayor de Blasio was in office from January 1, 2014 - December 31, 2021 covering the implementation of both 4K and 3K so the "front door" should be relatively stable throughout this time. Further examination needs to be completed with how closely the number of applicants at the intake center correlates with the number of accepted entrants. This data is included in my FOIL request and should be made available to me in early 2023. Its analysis will be included in an updated version of this paper.

6.2 Cost of Early Education Expansion

Within four months of New York City's fiscal year close on June 30$^{\text{th}}$, the Comptroller publishes the Annual Comprehensive Financial Report (ACFR) (New York City Comptroller 2022). It lists the budget and expenditures of city programs. Previous to 2014, the New York City Department of Education (DOE) provided an enrollment count of Pre-K in its schools. Since Mayor de Blasio took office in 2014, the total Pre-K enrollment includes the additional public early education center children age 4 (and now age 3) in the total Pre-K count. In 2014 the ACFR also began a separate budgeting line for UPK and in 2016 began tracking UPK expenditures separately. Special Education Pre-K Contract Payments have been included in the ACFR for at least 20 years and actually decreased since expansion of Pre-K in 2014. Because these payments were previously ongoing, these were not included in the additional UPK program costs instituted in 2014.

Table 10 shows UPK expenditures, UPK enrollment, and per-student cost of the program. The expenditure tab comes from the respective year's ACFR. The ACFR also contains UPK enrollment data, but only total enrollment. The DOE provides a yearly demographic snapshot (NYC Department of Education 2022b) that includes enrollment data for each grade at every public school in the city on October 31$^{\text{st}}$ each year. Total enrollment data is slightly different (less than 1%) between the reports, but this merging of the two sources allows the reader to see the introduction of 3K into the fold over the last five years. Included in appendix B table 14, are UPK costs dating back to 2014.

Table 10: Last Five Years Universal Pre-Kindergarten Costs

Fiscal Year	2018	2019	2020	2021	2022
Expenditures					
Personal Services	$490,203,709	$567,631,665	$666,931,144	$641,100,166	$747,447,490
Other Than Personal Services	$409,786,548	$438,502,405	$405,607,801	$443,547,230	$820,647,727
Total UPK Expenditures	$899,990,257	$1,006,134,070	$1,072,538,945	$1,084,647,396	$1,568,095,217
Enrollment					
Total UPK Enrollment	71,528	73,904	87,480	75,981	92,055
4K Enrollment	70,704	70,590	69,894	60,501	57,315
3K Enrollment	824	3,314	17,586	15,480	34,740
UPK Full Day Enrollment	67,881	67,886	67,589	58,469	56,045
Percentage in Full Day UPK	95%	92%	77%	77%	61%
Per Student Cost	**$12,582**	**$13,614**	**$12,260**	**$14,275**	**$17,034**

The ACFR breaks down expenditures between "personal" and "other than personal". Personal services include salaries and fringe benefits of City employees. Other than personal services include expenses other than salaries such as supplies, equipment, utilities, and contractual services (NYC Mayor's Office of Management and Budget n.d.). One can see the total UPK expenses have increased over the past five years from $900M in 2018 to $1.57B in 2022. Expenditures and budgets are not broken down between 3K and 4K as everything is lumped into UPK or Pre-K expenses.

Total enrollment has increased every year since 2014 except for 2021. This is most likely due to the effects of the COVID-19 pandemic when schools closed, workplaces closed, and people moved out of the city to less expensive locations as remote work became more prevalent. 4K was already universally available to all when the pandemic hit, so there has been a drop from 2020-2022. 3K was just being implemented in 2020 when the pandemic hit, so its enrollment continues to grow through the pandemic with the first 824 enrollees in 2018 and currently up to 34,740 enrollees during the 2021-2022 school year.

The per-student cost of the program has remained stable for six of the past seven years the ACFR has reported expenditures. Between 2016 and 2021, the cost per student fluctuated between $11,707 and $14,275. In 2022, the city made a large financial blunder by substantially overestimating the number of enrollees and the cost per student rose to $17,034. The city budgeted, hired, and had

[1] Depicts average cost of childcare for children under the age of 2 in New York City in 2022. As age of child rises, costs decrease. Median rent for a 1 bedroom apartment in New York city is $2,570 in 2021 and median monthly New York City income is $5,333 (U.S. Census Bureau). (TOOTRiS 2022)

classrooms setup for many more students than actually enrolled. While 3K students have continued to grow since the program was implemented, the 4K enrollment has not begun to increase back to its pre-pandemic levels. The Pre-K students (both 3K and 4K) that attended the full day declined immensely from 95% in 2018 to only 61% in 2022. This is most likely due outcomes referenced in the previous paragraph, of declining total employment that is just returning to pre-pandemic levels in 2023 (U.S. Bureau of Labor Statistics 2022) and more parents working remotely and choosing to keep their young children home.

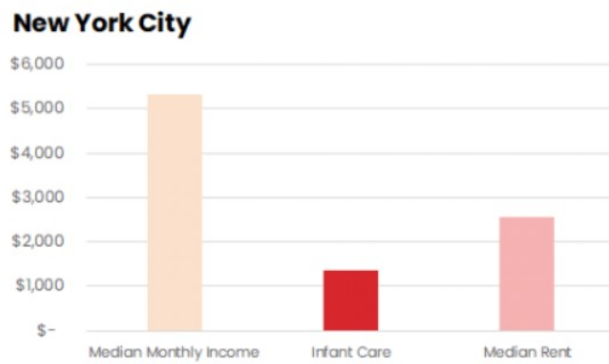


Figure 8: Major Expenses with Young Child [1]

Mayor de Blasio's 4K expansion in 2014 to city-wide by 2016 was largely seen as his signature policy achievement as mayor. At its height between 2016 and 2019, 4K enrolled over 70,000 students, approximately 90% of the number that would attend kindergarten the following year. The program saved parents on average $1,000 per month in childcare and allowed some parents to go back to work (TOOTRiS 2022). The 3K program, introduced in 2019, had a delayed rollout due to the pandemic. In his last year, 2021, there was a major push from de Blasio to expand the 3K program city-wide to spur lagging employment. This obligated hiring of many teachers, purchasing equipment, and readying facilities for the children. Many of these seats are located at early education centers which are paid as contractors of the city. With the mis-estimation of Pre-K enrollments and tax revenues lower than expected, some of these contracting centers were paid late or not at all. Mayor Eric Adams is facing mounting criticism and protests for his handling of the situation. There is fear from 3K supporters that he may scale back or even cut the 3K program completely (Fitzsimmons 2022).

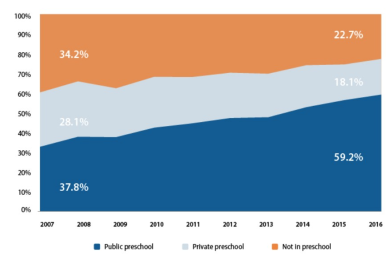
Another unique cost to early education expansion is the effect it has on private childcare centers. As I alluded to above, as a child matures from an infant to a toddler and then to preschool age, childcare costs decrease. Many states have laws that require a lower student-to-teacher (or caregiver) ratio, the younger the age. This increases costs as the average age of attendees at a daycare center decreases. If many three and four-year-olds are pulled out of a childcare center, this

[2](Malik 2018)

could raise the costs for parents of younger children that attend, pricing them out or even forcing a center to close if the price outpaces the decreased demand at a higher cost. Hurley 2019 states that New York seems to have a balanced approach to this problem by incorporating childcare centers in its four-pronged approach to UPK. In New York, it's not just the public schools that offer Pre-K, but millions of dollars are being obligated to private centers that contract for the city. Brown 2018 does not seem to be as optimistic. Her calculations show that the cost difference between a four-year-old and a two-year-old is about 15% and usually it is the poorest daycares that are hurt the most. If the current events regarding payment troubles New York City is having with its childcare contractors are any indication, it appears these unintended consequences of creating a benefit for 3 and 4-year-olds may harm infants, toddlers, their families, and the lower-income childcare industry.

6.3 Cost Effectiveness of Pre-K to Homeless Reduction

Figure 9: Private Market Crowd Out of Childcare[2]



After the misestimation of UPK enrollment in 2022, the cost per student was $17,034. Assuming enrollment levels will return to pre-pandemic levels and the city estimates a reasonable Pre-K enrollment, the cost per student should fall to roughly $15,000 per student per year. In 2022, the average shelter entry cost of a family with children was $92,014. This amount has risen consistently and will probably pass $100,000 per family with children entry in 2023 or 2024. Using $100,000 for shelter entry cost and $15,000 cost of an enrollee in Pre-K, UPK is more cost-effective than homeless shelters only if one out of every 6.6 Pre-K students prevents his or her family from entering a shelter. Intuitively, one out of 6.6 Pre-K students is not in danger of becoming homeless so UPK is not a cost-effective policy for reducing the homeless family population.

An even clearer way to come to this conclusion would be to recall fig. 7d, which shows that families with children cost the city $1.6M per night to shelter. Over one year, this totals $584M. The 2022 yearly UPK costs of $1.56B is roughly triple the cost to shelter all families with children. If free Pre-K were offered to a very targeted low-income demographic in which homelessness was a high probability (1 in 6.6), then it would be cost-efficient. UPK is not a targeted program, and it is not built with a homeless reduction effect in mind. UPK may reduce homelessness, but it does so at a much greater monetary expense than traditional sheltering can provide.

UPK is designed to give students better educational and employment opportunities in the future. This should increase future societal output and reduce inequality. It also has a likely secondary effect of increasing employment and societal output in the present as it allows parents who were unemployed because they were involved in childcare the opportunity of employment and creates jobs for Pre-K teachers.

7 Conclusion

Based on my results with the publically available aggregate data from DHS, I believe there is evidence for a slight treatment effect for homeless intake reduction in families with young children due to the expansion of childcare in the form of universal early education. Based on what looks like minimal effectiveness and high costs compared to sheltering an individual or family, childcare policies must be very targeted if they aim to break even from a cost-benefit perspective in terms of homeless reduction.

While individualized DHS and DOE data does exist and has been used in various papers, with the public aggregate data available, I cannot provide a more definitive conclusion regarding how much, if any, early childhood education expansion or childcare affects the number of people experiencing homelessness or entering homelessness.

7.1 Recommendations for further research

I intend to continue editing this paper based on recommendations I receive and to make the statistics more sound with individualized applicant and intake data for the New York City shelter system with data from DHS. I am curious if the micro-level data matches the macro data presented here and if it allows one to draw stronger correlations between UPK policies and homeless families with children reductions.

I intended to conduct interviews with a limited number of homeless and gather their feedback and perspective. Because this was part of a research project, it required me to complete the Human Subject training and a protocol review from Columbia University's Institutional Review Board (IRB). The protocol was approved only three days before the paper's class due date, but the interviews will be conducted the week of May 14, 2023. I believe it is important to take the time to engage in these conversations with groups of people I study and not become academically detached. The script stems from interviews with homeless and shelter employees I conducted in Florida in 2016, and incorporates a framework similar to the qualitative study undertaken by Mabhala, Yohannes, and Griffith 2017. The script is listed in Appendix C. Appendix D is a table built for quick entry of respondents' answers into the paper.

References

Brown, Jessica H (2018). "Does public pre-k have unintended consequences on the child care market for infants and toddlers?" In: *Princeton University Industrial Relations Section Working Paper* 626.

Cassidy, Mike (2020a). "A closer look: Proximity boosts homeless student performance in New York City". In: URL: https://miketcassidy.com/img/jmp_mc.pdf.

— (2020b). "Short Moves and Long Stays: Homeless Family Responses to Exogenous Shelter Assignments in New York City". In: URL: https://miketcassidy.com/img/jmp_fam_mc.pdf.

Coalition for the Homeless (n.d.). *Why Are So Many People Homeless?* Accessed: 2022-10-3. URL: https://www.coalitionforthehomeless.org/why-are-so-many-people-homeless/.

Engebreth, Ben (2020). *New York Residential Rent and Rental Statistics.* Accessed: 2022-11-28. URL: https://www.deptofnumbers.com/rent/new-york/.

Fitzsimmons, Emma G. (Sept. 22, 2022). *3-K for All? Adams Retreats From Expanding N.Y.C. Preschool Program.* URL: https://www.nytimes.com/2022/09/22/nyregion/prekindergarten-adams-nyc-3k.html.

Frazier, Ian (2013). "Hidden city". In: *The New Yorker*, pp. 39–49.

Grynbaum, Michael M. and Nikita Stewart (Aug. 5, 2015). "De Blasio, Facing Criticism, Is Taking On Homelessness With $22 Million Initiative". In: *The New York Times.* URL: https://www.nytimes.com/2015/08/06/nyregion/de-blasio-facing-criticism-is-taking-on-homelessness-with-22-million-initiative.html (visited on 08/06/2015).

Hanratty, Maria (2017). "Do local economic conditions affect homelessness? Impact of area housing market factors, unemployment, and poverty on community homeless rates". In: *Housing Policy Debate* 27.4, pp. 640–655. URL: https://www.tandfonline.com/doi/full/10.1080/10511482.2017.1282885.

Hayes, Robert (1987). "Litigating on Behalf of Shelter for the Poor". In: *Harv. Cr-cLL Rev.* 22, p. 79. URL: https://heinonline.org/HOL/LandingPage?handle=hein.journals/hcrcl22&div=12&id=&page=.

Hurley, Kendra (Dec. 13, 2019). *How Universal Pre-K Drives Up Families' Infant-Care Costs.* URL: https://www.bloomberg.com/news/articles/2019-12-13/how-universal-pre-k-affects-infant-daycare-costs.

iPropertyManagement (Apr. 2022). *Averge Rent by Year.* Accessed: 2022-11-3. URL: https://ipropertymanagement.com/research/average-rent-by-year.

Jones, Marian Moser (2015). "Creating a science of homelessness during the Reagan era". In: *The Milbank Quarterly* 93.1, pp. 139–178.

Mabhala, Mzwandile A, Asmait Yohannes, and Mariska Griffith (2017). "Social conditions of becoming homelessness: qualitative analysis of life stories of homeless peoples". In: *International journal for equity in health* 16.1, pp. 1–16.

Malik, Rasheed (Sept. 26, 2018). *The Effects of Universal Preschool in Washington, D.C.* URL: https://www.americanprogress.org/article/effects-universal-preschool-washington-d-c/.

National Academies of Sciences (July 2018). "Permanent Supportive Housing: Evaluating the Evidence for Improving Health Outcomes Among People Experiencing Chronic Homelessness". In: Accessed: 2022-10-3, Appendix B, The History of Homelessness in the United States. URL: https://www.ncbi.nlm.nih.gov/books/NBK519584/.

National Alliance to End Homelessness (Jan. 2020). *Housing.* Accessed: 2022-10-05. URL: https://endhomelessness.org/homelessness-in-america/what-causes-homelessness/housing/.

New York City Comptroller (July 2022). *Annual Comprehensive Financial Report (ACFR).* Accessed: 2022-11-25. URL: https://comptroller.nyc.gov/reports/annual-comprehensive-financial-reports/.

NYC Department of Education (2022a). *3-K.* Accessed: 2022-10-10. URL: https://www.schools.nyc.gov/enrollment/enroll-grade-by-grade/3k.

— (June 2022b). *Demographic Snapshot.* Accessed: 2022-10-20. URL: https://data.cityofnewyork.us/Education/2017-18-2021-22-Demographic-Snapshot/c7ru-d68s.

NYC Department of Homeless Services (Nov. 2022). *Daily Report.* Accessed: 2022-11-4. URL: https://www.nyc.gov/site/dhs/about/stats-and-reports.page.

NYC Mayor's Office of Management and Budget (n.d.). *Frequently Asked Questions.* Accessed: 2022-11-25. URL: https://www.nyc.gov/site/omb/faq/frequently-asked-questions.page.

NYC Mayor's Office of Operations (n.d.). *Mayor's Management Report*. Accessed: 2022-11-2. URL: https://www.nyc.gov/site/operations/performance/mmr.page.

NYC Office of the Mayor (Feb. 1, 2016). *Mayor Bill de Blasio Announces City Aggressively Inspecting and Fixing Homeless Shelter Conditions*. URL: https://www.nyc.gov/office-of-the-mayor/news/120-16/mayor-bill-de-blasio-city-aggressively-inspecting-fixing-homeless-shelter-conditions (visited on 02/01/2016).

O'Flaherty, Brendan (2019a). "Homelessness research: A guide for economists (and friends)". In: *Journal of Housing Economics* 44, pp. 1–25. ISSN: 1051-1377. DOI: https://doi.org/10.1016/j.jhe.2019.01.003. URL: https://www.sciencedirect.com/science/article/pii/S1051137718302109.

— (Apr. 2019b). *The Dean's Table with Fredrick Harris: Economics Professor Dan O'Flaherty*. Podcast. Columbia University. URL: https://deanstable.com/economics-professor-dan-oflaherty-dt5/.

Piña, Gabriel and Maureen Pirog (2019). "The impact of homeless prevention on residential instability: Evidence from the Homelessness Prevention and Rapid Re-Housing Program". In: *Housing Policy Debate* 29.4, pp. 501–521.

S&P Dow Jones Indices LLC (2022). *S&P 500*. FRED, Federal Reserve Bank of St. Louis. URL: https://fred.stlouisfed.org/series/SP500#0.

TOOTRiS (2022). *COST OF CHILD CARE IN NEW YORK: A BREAKDOWN (2022 UPDATE)*. URL: https://tootris.com/edu/blog/parents/the-cost-of-child-care-in-new-york-a-breakdown-for-2021/.

U.S. Bureau of Labor Statistics (Nov. 2022). *All Employees, Total Nonfarm*. FRED, Federal Reserve Bank of St. Louis. URL: https://fred.stlouisfed.org/series/PAYEMS.

U.S. Department of Housing and Urban Development (2022a). *Homeless Management Information System*. Accessed: 2022-10-05. URL: https://www.hudexchange.info/programs/hmis/.

— (Feb. 2022b). *PIT and HIC data since 2007*. Accessed: 2022-10-05. URL: https://www.hudexchange.info/resource/3031/pit-and-hic-data-since-2007/.

World Bank (2021). *Inflation, consumer prices for the United States [FPCPITOTLZGUSA]*. FRED, Federal Reserve Bank of St. Louis. URL: https://fred.stlouisfed.org/series/FPCPITOTLZGUSA.

Zumper Inc. (Nov. 28, 2022). *New York, NY Rent Prices.* URL: https://www.zumper.com/rent-research/new-york-ny.

A 4K Yearly Regression Results

A.1 Comparison Between Homeless Children in Families from 0-5 and 6-13

VARIABLES	(1) Avg Shelter Population	(2) Avg Shelter Population	(3) Total Individuals Sheltered
4K Treatment	Continuous	Binary	Binary
	3,650.61****	3,225.09****	4,374.92****
	(430.66)	(340.67)	(669.17)
Age Group 0-5	1,000.17***	1,031.85***	2,796.33****
	(242.65)	(233.70)	(459.05)
4K & 0-5	**-443.54**	**-396.41**	**-1,603.83***
	(522.31)	(369.51)	(725.81)
Year			
2012	1,513.78***	1,513.78***	1,139.00*
	(302.30)	(286.22)	(562.21)
2013	2,355.28****	2,355.28****	1,837.50***
	(302.30)	(286.22)	(562.21)
2014	1,604.64****	455.33	-46.50
	(264.31)	(286.22)	(562.21)
Constant	7,442.86****	7,427.02****	14,783.83****
	(245.79)	(233.70)	(459.05)
Observations	10	10	10
R-squared	0.98	0.99	0.97

Standard errors in parentheses
**** p<0.01, *** p<0.05, ** p<0.10, * p<0.15

Table 11: Comparison Between Sheltered Homeless Children in NYC Ages 0-5 and 6-13 from 2011-2015

The regression results in table 11 compares the groups of homeless children ages 0-5 and 6-13. 0-5 becomes the treated group and 6-13 the untreated group. This decision is made because the 4K expansion would affect 4-year-olds in the 0-5 group directly, while only affecting the 6-13 group indirectly if they had 4 year old siblings. The first regression contains a continuous treatment variable as I capture the fact that while 4K program was implemented in 2014, it wasn't fully implemented. The remaining regressions are simplified, using a binary variable for treatment. This change does not have a large effect on the results. The treatment becomes more significant with this change. This is most likely due because 4K was implemented in poorer neighborhoods initially and higher income neighborhoods later. So even as the treatment rolled out, the low income group being studied - those who have 4 year old children with the chance of becoming homeless - were treated earlier rather than later in the process. The change in the third regression is the response variable changes from "average daily shelter population" to "total individuals serviced".

The age group variable is binary if the child is a member of that age group or not. The year variables are also binary indicators. Years 2011 and 2015 are not included as they are the base cases of untreated (2011) and treated (2015). The treatment begins in 2014. You can see the number of observations is quite small with five year observations across two groups.

A.2 Comparison Between Homeless Adults in Families with and without Children

VARIABLES	(1) Avg Shelter Population
4K Treatment	27.09
	(214.09)
Families w/ Children	-2,943.86***
	(1,287.20)
Families w/ Children & 4K	**-947.37******
	(269.27)
Shelter Exits	1.22****
	(0.18)
Year	
2013	-40.87
	(175.21)
2014	-8.89
	(146.73)
Observations	16
R-squared	1.00

Standard errors in parentheses
**** p<0.01, *** p<0.05, ** p<0.10, * p<0.15

Table 12: Comparison Between Sheltered Homeless Families with and without Children in NYC Ages 21-29 and 30-44 from 2012-2015

This regression in table 12 compares groups of adults with indirect 4K effects and the same age range with no 4K effects. Both of these groups are the most likely age ranges to have children 4 years old. Of the two adult groups, one is a member of adult families and the other is part of families with children. This latter group has the potential to be affected by the 4K expansion policy, the former does not. From the previous regressions, shelter exists are added. Shelter exits become the most significant variable and the only explanatory variable that is not binary, so it has the most explanatory power as its coefficient is multiplied by a generally large number.

A.3 Comparison Between Various Age Groups in Families with and without Children

VARIABLES	(1) Avg Shelter Population	(2) Avg Shelter Population
4K Treatment	8.92	-1,407.52**
	(174.64)	(728.29)
Families w/ Children		-39,098.95**
		(19,620.90)
Families w/ Children & 4K	**-1,045.96******	**-9,217.68*****
	(192.69)	(4,104.52)
Shelter Exits	1.35****	6.60***
	(0.02)	(2.64)
Year		
2013	122.51	-1,260.34**
	(125.13)	(703.20)
2014	176.68	602.23***
	(122.90)	(240.97)
Age Groups		
6-13	-741.26****	-740.31****
	(150.51)	(136.70)
21-29 w/ Children	-4,188.03****	-4,187.08****
	(150.51)	(136.70)
30-44 w/ Children	-3,823.07****	-3,822.13****
	(150.51)	(136.70)
21-29 w/o Children	-234.10*	-3,929.54**
	(142.17)	(1,858.96)
30-44 w/o Children	-144.45	-3,839.90**
	(142.17)	(1,858.96)
Observations	24	24
R-squared	1.00	1.00

Standard errors in parentheses
**** p<0.01, *** p<0.05, ** p<0.10, * p<0.15

Table 13: Comparison Between Sheltered Homeless Families in NYC from 2012-2015

These regressions in table 13 are similar to the previous regression, but with additional age groups, including the 0-5 and 6-13 which are directly or indirectly affected by treatment. All four age groups in the group of "families with children" are considered to be in the treated group as they all have potential to be treated. The two age groups of the "adult families" comprise the untreated group. The second regression has the additional "families with children indicator variable". Shelter exits remain the most significant part of these regressions. 2012 and 2015 are base years for no 4K treatment and 4K treatment and thus are not included. All age groups in "families with children" are considered treated in this case. These age groups were specifically chosen as other age groups would most likely not have a 4-year-old sibling (older high school children) or be parents of a 4-year-old child (exceeding the age of 44). Thus the age range of subjects in this regression is 0-13 and 21-44.

B Universal Pre-Kindergarten Costs

Fiscal Year	2014	2015	2016	2017	2018	2019	2020	2021	2022
Expenditures									
Personal Services			$439,140,425	$432,501,091	$490,203,709	$567,631,665	$666,931,144	$641,100,166	$747,447,490
Other Than Personal Services			$423,099,429	$416,883,869	$409,786,548	$438,502,405	$405,607,801	$443,547,230	$820,647,727
Total UPK Expenditures			$862,239,854	$849,384,960	$899,990,257	$1,006,134,070	$1,072,538,945	$1,084,647,396	$1,568,095,217
Enrollment									
4K Enrollment	55,734	66,403	71,847	72,553	70,704	70,590	69,894	60,501	57,315
3K Enrollment	0	0	0	0	824	3,314	17,586	15,480	34,740
Total UPK Enrollment	55,734	66,403	71,847	72,553	71,528	73,904	87,480	75,981	92,055
UPK Full Day Enrollment					67,881	67,886	67,589	58,469	56,045
% in Full Day UPK					95%	92%	77%	77%	61%
Per Student Cost			**$12,001**	**$11,707**	**$12,582**	**$13,614**	**$12,260**	**$14,275**	**$17,034**

Table 14: Universal Pre-Kindergarten Costs

C Homeless Interview Script

Recruitment

Introduction: “Hi, my name is Chris. I’m a Columbia University graduate student researching homelessness in New York City. Would you be willing to talk with me for about 10 minutes and share some of your experiences?”

If “no”: “Thank you for your time, and have a good day.”

If “yes”: “I’d like to ask you some questions, some of which may bring up painful memories. If you feel uncomfortable, you can stop me at any time. I will not be offended if you decide to stop the interview. You are welcome to see the questions I am going to ask. Would you like to see them?”

Interview

”Let’s begin with some basics:”

- Alias:
- Gender:
- Age:
- Race/Ethnicity:
- Living Condition (Street, Shelter, Apartment):
- Location:
- Education:
- Primary Caregiver:
- Family History:
- Physical / Mental Health Problems:
- Homelessness
  - Onset Event:
  - Date:
  - Episodes:
  - What brought you to NYC:
  - Why unsheltered:
- Employment:

- Yearly Income:
- Spending Habits:
- Donations - do you ever give money to others on the street:
- Recommendations for policies to decrease homelessness:

Thank you for your time and for talking with me today. Your input helps provide better research in this area.

D Homeless Interview Table

Table 15: Homeless Interview Notes

Alias	Biographical Information	Education	Primary Caregiver	Family History	Medical History	Homelessness History	Employment History	Current Income	Spending Habits	Donations	Homeless Research
Ruddle	Age: 60 Gender: M Location: Flat for homeless people	Highest Grade: 9 School Type: public Performance: poor Behavior: bad	Both parents	Dad died and house sold when he was 12 Disruptive family life	'If I need see a doctor I just to the pharmacy'	Onset Event: both parents died when 15 years-old Date: NYC: Sheltered	In and out of fast food jobs	$15,000	Food, clothes, alcohol	Yes, $1 to friends, will split a pizza	Cash Transfers
John	Age: 52 Gender: M Location: 110th & Broadway	Highest Grade: 10 School Type: public Performance: poor Behavior: in trouble with the principal	Both parents	Normal working class family	No health problems	Onset Event: Divorce & Alcoholism Date: NYC: Unsheltered	Electrical factory worker for 10 years. Fired 3 years ago.	$12,000	Food, clothes, debt	Yes, $5 to Jason	Access to housing help/vouchers
NAME	Age: Gender: Location:	Highest Grade: School Type: Performance: Behavior:	CAREGIVER	FAMILY	MEDICAL	Onset Event: Date: NYC: Unsheltered	EMPLOYMENT	INCOME	SPENDING	DONATIONS	REC