A Voucher a Day Keeps the Doctor Away: Bounding the Effect of Housing Assistance on Recipients' Health*

Giuseppe Germinario†

May 28, 2022

Click here for the most recent version.

Abstract

Housing quality, both in terms of the physical structure and neighborhood characteristics, is positively associated with health outcomes. I examine the Housing Choice Voucher (HCV) program as a possible means of improving the health of recipients who would otherwise participate in place-based assistance programs such as public housing, or rent without a subsidy in the private market. Under three weak and plausible assumptions—voucher holders have weakly worse average potential health outcomes than non-recipients (non-positive monotone treatment selection); receiving a voucher does not result in strictly poorer health (monotone treatment response); on average, potential health outcomes are positively related to reported income (monotone instrumental variable)—I obtain nonparametric bounds on the average treatment effect of HCV on the health of individuals in recipient households using nationally-representative data from the 2018 Survey of Income and Program Participation (SIPP). My preferred estimates indicate that the causal effect is positive and statistically significant, with the likelihood of good or better self-reported health status increased by at least 4.8 percentage points and at most 21.3 percentage points. I also estimate that the probability of not having been hospitalized over the previous year is increased between 0.1 and 18.4 percentage points, though the effect is not statistically distinguishable from zero at conventional levels. Among Black members of voucher households, I estimate a larger lower bound for the effect on self-reported health of 11.7 percentage points, hinting that the effect on health is potentially larger for this subpopulation.

Key words and phrases: Health and poverty; Subsidized housing; Treatment effects; Bounds
JEL classification: R28, I14, C31

*I am incredibly grateful for the continued guidance and support from my advisor, Alfonso Flores-Lagunes. I have also greatly benefited from comments and feedback from Hugo Jales, Maria Zhu, Amy Ellen Schwartz, Gary Engelhardt, Jeffrey Kubik, Emily Wiemers, Alexander Rothenberg, Devashish Mitra, Vikesh Amin, Carlos Flores, Sam Saltmarsh, Joaquín Urrego, Christopher Rick, an anonymous reviewer, and seminar participants at Syracuse University and the U.S. Census Bureau. All errors are my own.

†Department of Economics, Syracuse University, Syracuse, NY 13210. Email: grgermin@syr.edu

1 Introduction

The Housing Choice Voucher program is the largest housing assistance program offered by the US federal government for low-to-moderate income renters on the private market. In 2018, the Department of Housing and Urban Development (HUD) reported that expenditures on the program totaled $21 billion and 2.5 million households received support via a housing voucher. Despite these figures, the supply of vouchers remains limited, and only around one quarter of eligible households receive benefits (Rosen, 2020). Given recent interest in expanding the voucher program—the American Rescue Plan of 2021 has so far led to an influx of $1.1 billion (or 70,000 vouchers) for Emergency Housing Vouchers targeting those most at risk of homelessness, and President Biden's proposed Build Back Better Framework included $15 billion in funding for additional vouchers—it is of critical interest to understand how recipients are impacted by receiving a voucher. This paper focuses on how receipt of a housing voucher affects the health status of individuals in the household relative to a counterfactual in which they receive place-based rental assistance (*e.g.*, residing in public housing, project-based Section 8, or Low Income Housing Tax Credit units), or rent in the unsubsidized private market.

This question is motivated by the strong link between housing and public health. Currently, the World Health Organization Housing and Health Guidelines (WHO) cite five key priorities in this area: inadequate living space (crowding), both low and high indoor temperatures, injury hazards in the home, and accessibility of housing for people with functional impairments. The first of these is associated with the spread of close-contact infectious diseases. The latter four focus on physical qualities of housing and how these can be harmful to the health of residents. For example, the report notes that cold indoor temperatures are correlated with high blood pressure and asthma symptoms, high indoor temperatures can be linked to dehydration and sleep disorders, and, to cite a more well-known example, exposure to lead carries a host of negative health consequences (anemia, kidney damage, brain damage), particularly among children. Finally, the guidelines also emphasize neighborhood characteristics such as air quality, crime rates, or noise levels. Persistent throughout is a WHO recommendation that governments address these issues, in part, by improving access to affordable housing.

My goal is to look specifically at one means of increasing affordability—vouchers—and assess its effectiveness as a public health intervention. In this paper, I consider that the causal path between receiving a housing voucher and experiencing better health operates through at least three potential mechanisms. First, receiving a voucher improves stability in the housing situation, which may help to reduce stress or anxiety. The second I call a "housing choice effect" that the voucher may allow households to relocate into housing that performs better with respect to the WHO priorities. The third is an income effect that arises from the fact that the voucher program is implemented as an economic transfer: a housing voucher limits the recipient household's rent to 30% of its monthly income, with the federal government subsidizing the remaining balance, up to a payment standard set by the local public housing authority. When this is a reduction in monthly expenditures on housing, the household's budget constraint for non-housing goods is relaxed, freeing up resources that can be devoted to improving health.

To my knowledge, this is the first study attempting to answer this question for adults in voucher households using data representative at the national level. Elsewhere in the literature, Katz, Kling, and Liebman (2001) do include adult health as an outcome in their study of the Moving to Opportunity

(MTO) experiment in Boston in 1998. The main idea of MTO was to randomly assign housing vouchers to households living in public housing or project-based Section 8 units in high-poverty neighborhoods, which would allow for identification of treatment effects. The place-based rental assistance counterfactuals for voucher holders are meant to emulate this setting. I am ultimately interested in a broader question of how vouchers compare when the counterfactual includes unsubsidized renting, and my analysis differs in two other key ways. First, there was one-sided non-compliance in the MTO treatment groups (in the sense that not all households assigned a voucher ended up using it), such that the estimate of the average treatment effect on the treated (ATT) in Katz et al. (2001) identifies a local average treatment effect (LATE) on health. In contrast, my strategy in this paper allows me to identify a population average treatment effect (ATE). In addition, the previous specificity to Boston leaves the distinction that I am identifying an effect for a broader population, even if the parameters identified were equivalent.

Outside of the random assignment as part of Moving to Opportunity, several aspects of the voucher program make it difficult to develop a compelling strategy for point identification of effects, particularly in public datasets. First, the primary structure of the program has not undergone any substantial changes since its inception in 1974 (Olsen, 2003; Zhang, 2021), leaving little room for natural experiments or IV/RDD approaches. Second, long waiting lists for vouchers induce an issue of negative selection. Previous empirical work has relied on administrative panel data and fixed effects strategies (Horn et al., 2014; Ellen et al., 2016; Schwartz et al., 2020) or on estimation of structural models (Zhang, 2021) to overcome these difficulties. In general, the first of these relies on within-person variation in voucher status and in the outcomes of interest. If this variation is random throughout the population, then the fixed effects strategies identify a population average treatment effect. If the variation is instead specific to certain subgroups, then a more local (and perhaps less policy relevant) effect is identified. A structural approach, on the other hand, relies on the validity of the parametric modeling choices, which could potentially be in question for applications.

My primary contribution in this paper is that I take a different approach which allows me to identify the causal effect of housing voucher receipt on individual health using only a single cross-section of the Survey of Income and Program Participation (SIPP). Specifically, I employ the partial identification method of Manski and Pepper (2000) to obtain nonparametric bounds on the average treatment effect in the population of low-income renter households who meet the HCV eligibility requirements. To do so, I make three key monotonicity assumptions: (1) there is negative monotone treatment selection (MTS), (2) total reported personal income serves as a valid monotone instrumental variable (MIV), and (3) there is monotone treatment response (MTR). The first assumption of MTS formalizes the above intuition that voucher recipients are negatively selected based on potential health outcomes, such that observed voucher recipients would, on average, have weakly worse potential health outcomes in any state of the world. The second assumption that total personal income is a valid MIV requires that, on average, potential health outcomes are non-decreasing in this reported income; this is supported, for example, by a consistent empirical finding of a health gradient with respect to income. The third assumption of MTR posits that an individual's potential health outcome under voucher receipt is no worse than without a voucher, and relies upon what I call the housing choice and income effects of voucher receipt above. I assess the plausibility of these assumptions in detail below. While they are relatively mild restrictions and are reasonable in this context, they do come at the price that I can only estimate bounds on the

ATE. I argue, however, that the price is worth it: by relaxing the stronger assumptions that I would need to obtain point identification (*e.g.*, an exclusion restriction for a traditional IV strategy or parametric assumptions), I can estimate a credible range of values for the average causal effect of vouchers on health.

I am able to statistically rule out a null effect, estimating that the probability of good or better self-reported health is increased by at least 4.8 percentage points (9.2%) when receiving a voucher. The same effect has an upper bound of 21.3 percentage points (59.5%). Additionally, the effect on the probability of zero nights spent in the hospital over the previous year is bounded below by a 0.1 percentage point (0.1%) increase and bounded above by an 18.4 percentage point (29.7%) increase, though the ATE cannot be statistically distinguished from zero at conventional levels. Together, these suggest that the health benefit of this form of housing assistance has a non-negligible magnitude. From analysis of demographic subgroups, I also find that Black members of voucher households are at least 11.7 percentage points (19.6%) and up to 26.9 percentage points (60.3%) more likely to self-report good health, and that the ATE is positive and statistically significant. If this suggests that the true effect is larger for Black voucher holders, then housing vouchers may also be a policy tool that can be leveraged to combat health disparities in the US.

With my primary outcome being a self-reported (and thus subjective) health status, it is important to be thoughtful with respect to the interpretation of the results. One way to understand self-reported health is as an imperfect—yet positively correlated—measure of true underlying physical health (Miilunpalo et al., 1997; Idler and Benyamini, 1997; Crossley and Kennedy, 2002; Wilson et al., 2007). This is the primary interpretation I take here, such that the effect on subjective health reflects a change in objective health. Alternatively, the self-reported status may more so capture an individual's perception of a given level of underlying health (Jürges, 2007; Lima-Costa et al., 2012). In this sense, the effects I bound represent changes to the "frameworks of evaluations" (Jylhä, 2009) after a housing voucher is received and recipients have the potential to improve the quality of their housing. While the former interpretation is more straightforward, the latter would still imply that my estimates demonstrate a meaningful improvement for voucher holders: a more positive perception of one's own health may indicate improved mental health.

The remainder of the paper proceeds as follows. Section 2 provides institutional details for the voucher program and reviews the relevant empirical literature on the impacts of housing vouchers. Section 3 describes my use of the SIPP and provides summary statistics. Section 4 provides a brief conceptual framework for my setting. Section 5 delves into the three main identifying assumptions and describes my procedure for estimation of and inference on the bounds, and the results are presented in Section 6. I conclude with Section 7.

2 Background

2.1 The Housing Choice Voucher Program

The use of housing vouchers in the United States originates from the Housing and Community Development Act of 1974, and was a policy response to concerns that public housing was causing "concentrations of poverty," where participants were located in low-income geographic areas (Massey and Kanaiaupuni, 1993). The legislation amended Section 8 of the Housing Act of 1937 to create a Section 8 Certificates (or Section 8 Existing) program. This awarded a Certificate to eligible households that was a subsidy for renting in the private market that would limit rent to 30% of income, under the conditions that properties met housing quality conditions and rent limitations established by the Department of Housing and Urban Development (HUD).

Later, the Housing and Community Development Act of 1987 created the Section 8 Voucher program, which was largely similar to the Certificate program but allowed recipients to spend more than 30% of income on rent payments. Finally, the Certificate and Voucher programs were merged in 1998 via the Quality Housing and Work Responsibility Act to create the modern Housing Choice Voucher (HCV) program.

The basic eligibility for a voucher under the HCV program is based on the household's total family income relative to the median income for the corresponding family size in the local county or metro area. The specific income thresholds are determined annually by HUD, and, in general, a recipient household's income may not exceed 50% of the median income for its size. Additionally, the Public Housing Authorities (PHAs) who administer the program locally are required to assign 75% of its available vouchers to households with income at or below 30% of the corresponding median.

Once an eligible household is selected (typically from a waiting list due to limited supply), the voucher works in the following way. First, a potential housing unit must be inspected by the PHA to ensure that it meets an "acceptable level of health and safety."[1] Next, the PHA sets a payment standard based on HUD's calculation of the local Fair Market Rent (FMR), which is typically equal to the 40th percentile of local rents, adjusted for the number of bedrooms required for the household size. The FMR represents the maximum permitted rent for voucher holders participating in the HCV program.

If the housing unit passes the PHA inspection and meets the FMR payment standard (and the property owner agrees to participate in the program),[2] then the voucher holder is required to pay 30% of household income towards the rent and utilities for the unit. The remaining difference between this amount and the payment standard is the subsidy amount the PHA pays directly to the landlord each month.[3]

Each party—the voucher holder, the landlord, and the PHA—retains obligations after the approved lease is signed. The voucher holder agrees to maintain the lease for at least one year, refrain from damaging the unit, limit the stays of guests who were not reported as household members when the application was submitted to the PHA, as well as follow other terms of the lease. Landlords are expected to provide "decent, safe, and sanitary" housing at the rental rate agreed upon in the lease and approved by the PHA. The PHA is responsible for ensuring the landlord receives the subsidy amount in a timely manner and for verifying annually that households remain income-eligible for the voucher and that the property continues to meet the minimum quality standard.

[1]Much of the information here comes directly from the HUD website.

[2]Federal law does not force landlords to accept tenants who have a voucher. However, some states and municipalities have instituted protections for tenants that prohibit discrimination based on source of income (SOI), which includes housing vouchers (Han, 2021).

[3]In some cases, a PHA may allow a recipient household to select a unit with rent exceeding the FMR, with the household having the responsibility of paying the excess. Overall, any excess payment cannot cause the proportion of income spent towards rent to rise above 40%.

2.2 Related Literature

A great deal of the empirical literature on housing vouchers has come from analysis of the Moving to Opportunity (MTO) experiment begun by the Department of Housing and Urban Development in 1994. From 1994 to 1998, HUD recruited 4600 low-income families with children and living in public housing across five major cities (Baltimore, Boston, Chicago, Los Angeles, and New York City) to participate in the experiment. The participants were then randomized into one of three treatment arms: (1) an experimental group offered a housing voucher that could only be used in a sufficiently low-poverty (a poverty rate below 10%) census tract as of 1990, (2) a comparison group offered a traditional housing voucher, or (3) a control group. Comparisons between groups (2) and (3) would then identify average treatment effects for the population included in the experiment.

Most highly relevant to the current paper is an early study of the MTO experiment in Boston by Katz, Kling, and Liebman (2001). The authors conducted follow-up surveys in 1997-1998 of 520 participant families. One question in the survey asked the heads of household for self-reported health status, with possible answers identical to the variable I use from the SIPP. Randomization allows for identification of both the effect of being offered a housing voucher (an intent-to-treat effect, or ITT) and of eventually receiving one (an effect of treatment on the treated, or ATT). They estimate the former as a 16.2 percentage point increase in the probability of good health, and the latter as a 26.4 percentage point increase. Although the ATT and ATE may differ due to one-sided non-compliance, this provides a useful benchmark for my analysis as a local average treatment effect (LATE): the experimental estimate lies near or slightly above all of my estimated bounds on the ATE, indicating that the magnitudes of my range of estimates are reasonable for the effect on a broader population.

Relatedly, Ludwig et al. (2011) conducted follow-up surveys with MTO participants from 2008 to 2010, collecting measures of body mass index (BMI) and glycated hemoglobin. They are thus able to identify the ITT of assignment to the voucher groups on the probability of obesity and diabetes. The estimates indicate that assignment to the traditional voucher group led to a statistically-significant (at the 10% level) 5.3 percentage point decrease in the likelihood of obesity and a 0.8 percentage point decline in the likelihood of diabetes (not significant at conventional levels). While, again, I am bounding a different parameter, the fact that effects on specific health conditions may be at the lower end of my identified sets for the ATE on overall self-reported health appears reasonable. Obesity and diabetes are only two components that an individual may consider when evaluating her own health, so it is possible that the effects on particular conditions accumulate into a total effect.

Apart from studies of MTO, other strands of the empirical vouchers literature seek to identify impacts using non-experimental data. This work has primarily focused on the effects on neighborhood or labor supply choices, the former of which can be informative for my MTR assumption. That is, the attributes of the neighborhood in which a voucher holder locates is a critical part of what I have called the housing choice effect of the voucher on health. Studies can generally be placed into one of two categories.

First are approaches which have relied on panel data or other fixed effects strategies to identify causal effects. Horn, Ellen, and Schwartz (2014) take one such approach by combining administrative HUD data with school district information to estimate the effect of voucher receipt on neighborhood choice through the lens of school quality. They find that voucher households locate near schools with 3 percentage point higher proficiency rates than for schools near households in public housing. On the other hand, schools near voucher households have slightly lower proficiency rates than schools near non-subsidized low-income households or Low Income Housing Tax Credit (LIHTC) recipients. They explain that this is likely due to the demographics of voucher holders, which is similar to my negative MTS assumption in that voucher holders have average characteristics potentially related to a lower demand for education quality. In addition to estimates that voucher holders locate in neighborhoods with greater amenities, Schwartz et al. (2020) provide evidence using data from New York City over the 2002-2012 period that voucher recipients also move into higher quality structures. Although not intended as causal estimates, they find that students in voucher households are 20% less likely to live in a building with a hazardous code violation. This indicates that, at least on average, the effect of voucher receipt through the direct effect on housing choice is positive. However, as I emphasize below, the MTR assumption is at the individual level, and so this evidence is merely suggestive.

The second approach common in the literature is to calibrate and estimate structural models. Leung, Sarpça, and Yilmaz (2012) set up a model with two household types (skilled/high-income and unskilled/low-income) and two neighborhoods. They compare equilibria reached under public housing versus a voucher program. Relative to the public housing equilibrium, their model predicts voucher holders end up with higher average lot sizes and locate in a higher quality neighborhood (measured by school quality) than public housing participants. Inside the voucher equilbrium, the voucher holders also reside on larger average lot sizes than the other low-income households without a voucher. Like the results cited in the previous paragraph, this suggests that the housing choice effect of the voucher on health is positive for at least some portion of the population.

3 Data

I utilize Wave 1 of the 2018 Survey of Income and Program Participation (SIPP) for my analysis. Administered by the U.S. Census Bureau, the SIPP is a nationally-representative survey begun in 1983 of households' income and government program participation in the United States. The survey is designed as a continuous series of panels that last between 2.5 and 4 years, where, beginning with the 2018 SIPP, a new panel begins each year so that there is overlap at each wave. Wave 1 of the 2018 SIPP therefore contains responses for the first panel in its first year. To focus on annual data, I use the month 12 response and the corresponding annual person weights in the analysis.[4]

A major advantage of using the SIPP is that, to the best of my knowledge, it is the only publicly-available dataset with information on both participation in the Housing Choice Voucher program and on individuals' health. My primary outcome of health status is measured by the respondent's answer to a question asking about her current health. Possible answers are on a scale including Poor, Fair, Good, Very Good, and Excellent health. I construct my main outcome of interest as an indicator variable equal to 1 if the self-reported health status is Good, Very Good, or Excellent, and equal to zero otherwise. My secondary outcome is the response to a question which asks how many nights the individual spent in the hospital over the previous year. I binarize this variable such that it is equal to 1 if zero nights were spent in the hospital, and equal to zero otherwise.

Information on voucher status comes from the following series of questions. First, the respondent is

[4]This follows the guidance in the Users' Guide to Selecting Weights published on the SIPP website.

asked if her living quarters are owned, rented, or "occupied without payment of rent." Those who report the latter two tenure statuses are then asked if their rent is lower because of participation in a federal, state, or local government housing program. Respondents who answer affirmatively to this question are then asked if their household has a housing voucher. Thus, the "treated" group in my sample consists of individuals in renter households possessing a housing voucher. The "control" or comparison group is made up of renters who (i) have their rent lowered by a government program which is not a housing voucher (*e.g.*, public housing, project-based Section 8, or LIHTC), or (ii) do not receive any government rental assistance. The former makes up 13% of my sample, and the latter 87%.

Unfortunately, the SIPP does not allow me to directly identify which non-voucher housing assistance is received by those receiving place-based subsidies in the comparison group.[5] To better understand the composition of this group, I perform the following back-of-the-envelope calculation. Using the figures available in the United States Federal Rental Assistance Fact Sheet from the Center on Budget and Policy Priorities for 2018,[6] 3 million households receive assistance due to a program other than HCV. This includes nearly 1 million households residing in public housing and 1.2 million in project-based Section 8 units. It is also relevant to consider households who may be living in units whose construction was subsidized by the Low Income Housing Tax Credit. Available data from HUD indicates that there were between 2.2 and 2.7 million LIHTC units in 2018. Only 80% of the households in these units have family incomes low enough to qualify for HCV.[7] Further, O'Regan and Horn (2013) provide evidence that 53% of these households also receive rental assistance, half of which are voucher recipients. Putting this together, I can conjecture that there are between 466,000 and 572,000 households residing in LIHTC units with no other assistance. Thus, I conclude that approximately 33% of this portion of my control group resides in public housing, 40% in project-based Section 8 housing, and 20% in LIHTC units.

Finally, I also use the individual's report of her total personal income. This variable is defined in the SIPP as including earnings from all jobs, total personal investment/property income, total personal means-tested transfer income, total personal social insurance payments, and total personal 'other' income (for example, survivor's benefits, government pensions, child support payments, or life insurance policy payouts). This income measure is reported monthly, so I sum over the 12 months of the SIPP to obtain my measure of total annual personal income.

I further restrict the sample by focusing on adults aged 18 to 65 and on those with a family income-to-poverty ratio below 1. The latter choice is a compromise. Public Housing Authorities establish eligibility for the voucher program based on thresholds relative to the local median family income. The geographic information available in the SIPP is limited to census region, so I use the federal poverty line adjusted for family size as a proxy for the main eligibility threshold of 50% of local median family income. With a national median family income in 2018 of $78,646 and a poverty guideline of $25,100 for a family of four, this is a decent proxy, if a bit overly restrictive. However, due to variation in median income over areas, I find it preferable to be on the restrictive side rather than including individuals in ineligible households. I am thus left with 3529 observations of individuals with non-missing data on health outcomes, voucher

[5] In a follow-up to a conversation with Census Bureau economists about this limitation, I was informed that specific questions about public housing and project-based Section 8 had been asked in the SIPP panel covering 2008-2013. The conjecture was that these were likely dropped due to concerns for respondent burden or disclosure risk.

[6] The fact sheet is available here.

[7] These figures are obtained from Tables 1 and 11 in the HUD document here.

status, and total annual personal income.

Table 1 describes key objects of interest in the sample. Overall, 76.4% of individuals report good or better health. In my sample, 7.7% of individuals are in a household having a housing voucher. Among these voucher holders, 57.1% report good or better health, versus 78.1% among the non-voucher holders. This unadjusted gap is indicative of the negative selection I describe in Section 5.

Table 1: Observed Probabilities of Good or Better Health and Distribution of Vouchers

t	$\mathbb{E}[Y \mid T = t]$	$\mathbf{Pr}[T = t]$	N_t
Voucher	0.571	0.077	279
	(0.033)	(0.006)	-
No Voucher	0.781	0.923	3250
	(0.008)	(0.006)	-
Total	0.764	1.000	3529
	(0.008)	-	-

Estimates use the annual SIPP weights for individual-level analysis. Standard errors in parentheses are based on 240 BRR replications with the SIPP replicate weights.

Table 2 provides descriptive statistics for my sample. Overall, the sample is 58.7% female, 22.8% Black, 45.2% white, and 22.2% Hispanic.[8] The average age is 35.34 with 12.36 years of schooling, 19.3% of the sample is married, and 36.5% reported having been employed for the entire previous month. The average total personal income is $4,853 and 26.3% are food insecure, while 9.9% report a plumbing problem in the home and 20.8% report a problem related to pests.

4 Conceptual Framework

Before proceeding to my key identifying assumptions, I first offer a simple framework to better understand the behavior of two main parties in this setting: the public housing authorities and the households.

Consider first a generic PHA which receives HCV funding from HUD and must decide how to allocate the corresponding housing vouchers. I remain agnostic on the preferences and objectives of the PHA in fulfilling this task. Rather, I take as given the empirical fact that housing authorities ration their limited supply of vouchers by creating a take-it-or-leave-it waiting list (Thakral, 2016).

For each household, let $U_i(1)$ denote potential utility obtained with a voucher, and let $U_i(0)$ denote potential utility from place-based rental assistance or renting without a subsidy. Households that join an HCV waiting list would incur a waitlist cost denoted θ_i. I suppose that this generates self-selection into voucher receipt à la a simple Roy model (Roy, 1951), based on the gains from the voucher relative to the waiting cost. Concretely, observed voucher status can be written

$$V_i = \begin{cases} 1\,, & \text{if } U_i(1) - U_i(0) > \theta_i \\ 0\,, & \text{if } U_i(1) - U_i(0) \leq \theta_i \end{cases}$$

Thus, voucher holders are those for whom the benefit of receiving a voucher exceeds the cost of entering

[8] My sample is also 5.8% Asian and 3.9% are categorized as "residual." These are an insufficient number of observations for the racial/ethnic subgroup analyses in Section 6, and so I exclude them from Table 2.

Table 2: Summary Statistics

	By Treatment Status			By Subpopulation				
	Overall	Voucher	No Voucher	Female	Male	Black	White	Hispanic
Female	0.587	0.695	0.578	·	·	0.649	0.557	0.608
	(0.492)	(0.461)	(0.494)	·	·	(0.477)	(0.497)	(0.489)
Black	0.228	0.456	0.209	0.252	0.193	·	·	·
	(0.419)	(0.499)	(0.406)	(0.434)	(0.395)	·	·	·
White	0.452	0.237	0.471	0.429	0.485	·	·	·
	(0.498)	(0.424)	(0.499)	(0.495)	(0.500)	·	·	·
Hispanic	0.222	0.228	0.222	0.230	0.211	·	·	·
	(0.415)	(0.421)	(0.416)	(0.421)	(0.408)	·	·	·
Age	35.34	41.25	34.85	35.32	35.39	36.81	35.08	35.35
	(13.93)	(14.66)	(13.76)	(13.71)	(14.26)	(13.87)	(14.57)	(12.52)
Total Personal Income	4852.50	5945.01	4761.26	5359.54	4132.07	5152.40	4670.93	5460.13
	(10,710.11)	(5856.08)	(11,014.83)	(8763.46)	(12,952.79)	(10,785.74)	(10,896.73)	(8879.13)
Education (years)	12.36	11.70	12.42	12.29	12.47	12.37	12.77	11.03
	(2.52)	(2.00)	(2.55)	(2.49)	(2.55)	(1.83)	(2.17)	(3.14)
Married	0.193	0.150	0.197	0.178	0.216	0.151	0.154	0.291
	(0.395)	(0.357)	(0.398)	(0.382)	(0.411)	(0.358)	(0.361)	(0.454)
Employed	0.365	0.221	0.377	0.358	0.374	0.327	0.368	0.419
	(0.481)	(0.415)	(0.485)	(0.480)	(0.484)	(0.469)	(0.482)	(0.494)
Food Security	0.737	0.611	0.748	0.719	0.762	0.712	0.726	0.751
	(0.440)	(0.489)	(0.434)	(0.449)	(0.426)	(0.453)	(0.446)	(0.433)
Plumbing Prob	0.099	0.081	0.100	0.097	0.101	0.094	0.104	0.089
	(0.298)	(0.273)	(0.300)	(0.296)	(0.302)	(0.292)	(0.306)	(0.285)
Pest Prob	0.208	0.208	0.208	0.221	0.191	0.227	0.206	0.208
	(0.406)	(0.407)	(0.406)	(0.415)	(0.393)	(0.419)	(0.404)	(0.406)
Observations	3529	279	3250	2094	1435	757	1591	834

Standard deviations in parentheses are computed using 240 BRR replications with the SIPP replicate weights. Sample means are weighted with the annual person weights.

the waiting list. Those observed with non-HCV place-based rental assistance or no government assistance are those with either a lower benefit from a voucher, a larger waiting cost, or both. For simplicity below, I assume that households have either a low or high waitlist cost type, such that $\theta_i \in \{\theta_L, \theta_H\}$. For the low-cost types with $\theta_i = \theta_L$, the Roy selection can be approximately described as saying those who would benefit from a voucher receive one, and those who would not benefit are observed in place-based assistance programs or do not receive any subsidies (*i.e.*, close to a scenario for an assistance program without a waiting list, such as SNAP).

Upon reaching the top of the waiting list, voucher holders obtain the utility level $U_i(1)$ by maximizing utility through choices of housing services h and non-housing consumption b. I follow a simplified version of the housing choice model of Geyer (2017) and suppose that households have Stone-Geary preferences over these consumption levels. Amenities A_j in neighborhood j also enter the utility function additively with a preference parameter γ_i. I make a further addition that utility depends on health $G_i(h, \tilde{b}, A_j)$, which is an increasing function of housing services, a subset $\tilde{b}$ of non-housing consumption that is beneficial to health, and neighborhood amenities. The function $G_i(\cdot)$ simply maps these choices into an indicator for good or better health. Thus, voucher holders have an objective function

$$u_{ij} = \alpha \ln(h - H) + (1 - \alpha) \ln(b - B_{ij}) + \gamma_i A_j + G_i(h, \tilde{b}, A_j) ,$$

where H is a minimum level of housing services and B_{ij} a minimum level of non-housing consumption, which may depend on neighborhood j amenities. This is subject to a budget constraint $w = p_j^h h + b$. With the voucher, households pay 30% of income on housing, meaning that $p_j^h h = 0.3w$, and the expenditure on non-housing consumption is equal to $0.7w$.

Immediately, it is evident that voucher receipt may generate an income effect on non-housing consumption (and thus indirectly on health)[9] by changing the household's budget constraint (Hoynes and Schanzenbach, 2009). The extent of this effect depends upon the relevant counterfactual for the household. For the public housing and project-based Section 8 counterfactuals, where housing expenditures are also equal to 30% of income, the available budget share for non-housing consumption is unchanged. There is subsequently no income effect for total expenditure relative to these alternative states of the world. For the LIHTC counterfactual, on the other hand, housing expenditures would instead be equal to 30% of the local adjusted median income, rather than the individual household's income. Since the income eligibility limit for HCV is 50% of median income, voucher holders pay at most $0.3 \times 0.5 = 15\%$ of the local median. Thus, the post-voucher budget constraint can allow for potentially significant increases in non-housing expenditures when the counterfactual is a LIHTC unit. Similar reasoning applies when the relevant counterfactual is unsubsidized renting. In any case, the basic setup of the housing choice model indicates a non-negative effect on the budget constraint.

Finally, receiving a voucher has a direct effect on the specific bundle of housing services, non-housing consumption, and neighborhood amenities.[10] While the Roy selection places a restriction that the ultimate utility level $U_i(1)$ is greater than the sum of the counterfactual utility $U_i(0)$ and waiting list costs θ_i, there are a number of allocations consistent with this outcome. Thus, my identifying assumptions in the next section are additional restrictions on the behavior or characteristics of low-income renters eligible for housing assistance—they are not directly implied by this framework.

5 Identification and Estimation

Let the health outcome Y_i be an indicator function for an individual being in good health. For the self-reported health status (hospitalization outcome), $Y_i = 1$ if good or better health is reported (zero nights

[9]The positive indirect effect on health can be conceived as suggesting that health is a normal good with respect to post-transfer income. Such a result would be consistent with the human capital model of demand for health (Grossman, 1972). To summarize the intuition, let $p \equiv \mathbf{Pr}[Y_i(v) = 1]$ denote the probability of good or better potential health outcomes, so that $\Delta p > 0$ represents the change in this probability when the health stock is increased. Let w_i and I_i denote an individual's income and total expenditure on health investments, respectively. Then, the optimality conditions for an individual choosing over a consumption good and health capital (Cropper, 1977; Grossman, 1999) yield an expression for the marginal benefit of increasing the health stock, written $\Delta p \cdot u(w_i - I_i)$, where $u(\cdot)$ is a utility function with $u'(\cdot) > 0$. Inspection of this quantity reveals that an increase in w_i raises the marginal benefit of increasing the health stock, such that increased income will prompt a greater demand for health.

[10]Note that, even if the total expenditure on consumption other than housing is unchanged (*e.g.*, remains at 70% of income for the public housing and project-based Section 8 counterfactuals), the choice of b can be influenced by the new level of amenities A_j. For example, if a new neighborhood is closer to work, then some spending on transportation can be reallocated to other uses.

spent in the hospital are reported), and $Y_i = 0$ otherwise. The treatment variable denoted V_i takes the value of 1 if the individual is in a voucher household, and the value of zero if she is not. In potential outcomes notation (Rubin, 2005), the potential health outcome for individual i under voucher receipt is denoted $Y_i(1)$, and the potential health outcome under place-based assistance or unsubsidized renting is denoted $Y_i(0)$. The observed health outcome (either the self-reported status or no-hospitalizations status) can thus be written $Y_i \equiv V_i \cdot Y_i(1) + (1 - V_i) \cdot Y_i(0)$. My object of interest is the population average treatment effect of voucher receipt on health status, written as[11]

$$ATE(1,0) = \mathbb{E}[Y_i(1)] - \mathbb{E}[Y_i(0)] \tag{1}$$

or, equivalently,

$$ATE(1,0) = \mathbf{Pr}[Y_i(1) = 1] - \mathbf{Pr}[Y_i(0) = 1] \ , \tag{2}$$

since the outcome is binary in nature. I will use the expectation form of Equation (1) for notational convenience.

Identification of the ATE is not straightforward due to the fundamental problem of causal inference: $Y_i(1)$ is observed only for those with $V_i = 1$, and $Y_i(0)$ is likewise only observed for those with $V_i = 0$. To see this explicitly, iterate expectations over voucher status for the mean potential outcome $\mathbb{E}[Y_i(1)]$:

$$\mathbb{E}[Y_i(1)] = \mathbb{E}[Y_i(1) \mid V_i = 1] \cdot \mathbf{Pr}[V_i = 1] + \mathbb{E}[Y_i(1) \mid V_i = 0] \cdot \mathbf{Pr}[V_i = 0] \tag{3}$$

Each of the above terms is identified by the data, apart from $\mathbb{E}[Y_i(1) \mid V_i = 0]$, which is a counterfactual object; a similar expansion of the expectation of $Y_i(0)$ shows the other key counterfactual object to be $\mathbb{E}[Y_i(0) \mid V_i = 1]$. To make progress, I must therefore consider assumptions on such counterfactuals.

One basic assumption from Manski (1989) is to impose that the support of the outcome variable is bounded. In the present paper with a binary outcome, this assumption is trivial: Y_i is either equal to 0 or to 1. Despite this triviality, the assumption of bounded support for the outcome allows for the conclusion that the counterfactual conditional expectation $\mathbb{E}[Y_i(1) \mid V_i = 0]$ lies between 0 and 1. Returning to the iterated expectations expansion in Equation (3), I can write "worst-case" bounds[12] on $\mathbb{E}[Y_i(1)]$ (Manski, 1989):

$$\begin{gathered} \mathbb{E}[Y_i \mid V_i = 1] \cdot \mathbf{Pr}[V_i = 1] \\ \leq \mathbb{E}[Y_i(1)] \leq \\ \mathbb{E}[Y_i \mid V_i = 1] \cdot \mathbf{Pr}[V_i = 1] + \mathbf{Pr}[V_i = 0] \ , \end{gathered} \tag{4}$$

where I have made use of the fact that $\mathbb{E}[Y_i(1) \mid V_i = 1]$ is identified by $\mathbb{E}[Y_i \mid V_i = 1]$ in the data. Similarly,

[11] It may also be of interest to consider the average treatment effect on the treated (ATT); that is, policymakers could be interested in the effect of voucher receipt specific to those actually observed with a voucher. The same assumptions used here to partially identify the ATE can also be used to bound the ATT. I return to this in Section 6.3 below.

[12] Some of the expressions for the bounds are greatly simplified due to the setting of a binary treatment variable and a binary outcome. For a more general treatment with multiple treatment levels and continuous outcomes, see the original Manski and Pepper (2000) discussion.

$\mathbb{E}[Y_i(0)]$ can be bounded with

$$\begin{aligned} &\mathbb{E}[Y_i \mid V_i = 0] \cdot \mathbf{Pr}[V_i = 0] \\ &\leq \mathbb{E}[Y_i(0)] \leq \\ &\mathbb{E}[Y_i \mid V_i = 0] \cdot \mathbf{Pr}[V_i = 0] + \mathbf{Pr}[V_i = 1] \end{aligned} \tag{5}$$

The worst-case lower bound on the ATE is then obtained by subtracting the upper bound on $\mathbb{E}[Y_i(0)]$ in (5) from the lower bound on $\mathbb{E}[Y_i(1)]$ in (4). The worst-case upper bound is equal to the upper bound on $\mathbb{E}[Y_i(1)]$ minus the lower bound on $\mathbb{E}[Y_i(0)]$. Note that, in applications, these worst-case bounds are not able to identify the sign of the ATE and are also generally wide—with a binary treatment and a binary outcome, the width of the identified set will always be equal to 1. Despite this, the worst-case bounds do give the possibility of ruling out moderate-to-large magnitudes of the treatment effect in one direction. Nevertheless, assumptions beyond bounded support of the outcome are necessary to draw more meaningful conclusions on the causal effect of housing vouchers on health.

5.1 Assumptions

I turn to three key assumptions—monotone treatment selection (MTS), a monotone instrumental variable (MIV), and monotone treatment response (MTR)—originally proposed by Manski and Pepper (2000) to aid in tightening the worst-case bounds on the ATE. Overall, these are fairly mild assumptions that I argue are plausible in the present context. They lead to nonparametric bounds on the effect of interest.

5.1.1 Monotone Treatment Selection

The first assumption captures the idea that households receiving vouchers (and thus the individuals within them) are negatively selected into the voucher program based on potential health outcomes. In other words, I assume that, on average, individuals observed in voucher households have weakly poorer potential health outcomes in any state of the world versus individuals observed in public housing, project-based Section 8, LIHTC, or unsubsidized units. This is stated formally in Assumption 5.1, which follows Manski and Pepper (2000).

Assumption 5.1 (Monotone Treatment Selection)**.** *Suppose that*

$$\mathbb{E}[Y_i(v) \mid V_i = 1] \leq \mathbb{E}[Y_i(v) \mid V_i = 0]$$

for each treatment status $v \in \mathcal{V} = \{0, 1\}$.

Notice that this inequality holds for potential outcomes under both treatment arms. The probability of good health *under voucher receipt* for those observed in a voucher household is no greater than the same probability for those observed in a non-voucher household. At the same time, those observed with a voucher also have a weakly lower probability of good health *under no voucher receipt* than those observed without a voucher. Thus, MTS is an assumption about the inherent characteristics of the individuals observed in voucher and non-voucher households rather than an assumption about the impact of vouchers.

In this paper, I argue that the limited supply of housing vouchers which leads to lengthy waiting lists induces self-selection by low-income renters into voucher receipt, as summarized in the Roy model of the previous section. Note here that the Roy selection is based on overall utility benefit, while MTS is an assumption specific to potential health outcomes. In principle, then, the self-selection at play is consistent with either negative or positive MTS. However, if characteristics associated with larger potential gains from receiving a voucher (*e.g.*, low educational attainment, food insecurity) are also correlated with poor health, then a negative MTS assumption would hold: on average, those observed with a voucher would have had weakly worse health outcomes, regardless of realized voucher status.

A primary concern with this is that the ordeal (Black et al., 2003) of the voucher waiting list is particularly costly for individuals who have these same characteristics associated with worse health. With many wait times spanning multiple years, high-cost types having the poorest potential health outcomes may never make it to the top of the waiting list. In less extreme cases, individuals with high waiting costs and relatively poor potential health outcomes may have difficulty verifying their eligibility when they reach the top of the list and subsequently lose out on the voucher. Each of these scenarios would leave the possibility of some positive selection into HCV based on health.

It may indeed be true that this type of positive selection is present among some subset of my population of interest. This, however, need not contradict non-positive MTS. The key point here is that MTS is a mean-level assumption that needs only to hold for the expected potential health outcomes conditional on voucher status. Thus, it is permissible that the sample is a mix of positively- and negatively-selected individuals, as long as the size of the positively-selected subset or the magnitude of the positive selection is sufficiently small, such that the overall direction of selection remains non-positive.

To make this more precise, I iterate expectations over the Roy model waiting costs $\theta_i \in \{\theta_L, \theta_H\}$ on the terms from Assumption 5.1 above. This yields

$$\begin{aligned}\mathbb{E}\left[Y_i(v) \mid V_i = 1\right] = {} & \mathbb{E}\left[Y_i(v) \mid V_i = 1, \theta_i = \theta_H\right] \cdot \mathbf{Pr}\left[\theta_i = \theta_H\right] \\ & + \mathbb{E}\left[Y_i(v) \mid V_i = 1, \theta_i = \theta_L\right] \cdot \mathbf{Pr}\left[\theta_i = \theta_L\right]\end{aligned} \tag{6}$$

and

$$\begin{aligned}\mathbb{E}\left[Y_i(v) \mid V_i = 0\right] = {} & \mathbb{E}\left[Y_i(v) \mid V_i = 0, \theta_i = \theta_H\right] \cdot \mathbf{Pr}\left[\theta_i = \theta_H\right] \\ & + \mathbb{E}\left[Y_i(v) \mid V_i = 0, \theta_i = \theta_L\right] \cdot \mathbf{Pr}\left[\theta_i = \theta_L\right]\end{aligned} \tag{7}$$

The sign of the difference between Equation (6) and Equation (7) is the sign for the selection present; for non-positive MTS this difference must be non-positive. Carrying out this subtraction and grouping terms by the waiting cost type probabilities, I obtain

$$\begin{aligned}& \mathbf{Pr}\left[\theta_i = \theta_H\right] \cdot \Big(\mathbb{E}\left[Y_i(v) \mid V_i = 1, \theta_i = \theta_H\right] - \mathbb{E}\left[Y_i(v) \mid V_i = 0, \theta_i = \theta_H\right]\Big) \\ & \quad + \mathbf{Pr}\left[\theta_i = \theta_L\right] \cdot \Big(\mathbb{E}\left[Y_i(v) \mid V_i = 1, \theta_i = \theta_L\right] - \mathbb{E}\left[Y_i(v) \mid V_i = 0, \theta_i = \theta_L\right]\Big).\end{aligned} \tag{8}$$

The overall selection is therefore a weighted average of the selection specific to the high-cost and low-cost waiting types. I assume that the bottom term in parentheses—the selection term for the subset with low waiting list costs—is weakly negative. This reflects that, when waiting costs are more negligible, individuals with poorer potential health outcomes (who have a stronger incentive to improve their housing and join a waiting list) are able to reach the top of the waiting lists and receive a housing voucher.

The selection term for those with a high waiting list cost, on the other hand, is the source of the potential concern for the MTS assumption. Here, one could argue that individuals who have a high waiting cost and yet still receive a voucher have better potential health outcomes on average, versus those receiving place-based assistance or renting without a subsidy. If the magnitude of either this selection term or the proportion of high-cost types is large enough, then this may dominate the negative selection among low-cost types.

To get a sense of what this means for my sample, consider the following exercise. First, while the waiting costs θ_L and θ_H are unobservable characteristics, suppose that I can proxy the high waiting cost type with an indicator for both fewer than 12 years of education and low food security. In my sample, this produces a rough estimate of 0.06 for $\mathbf{Pr}[\theta_i = \theta_H]$. Second, I assume the largest possible magnitude for the positive selection among high-cost types: every individual observed to have a voucher has $Y_i(v) = 1$ (good health, regardless of voucher status), and every individual observed without one has $Y_i(v) = 0$ (less than good health, regardless of voucher status). In this worst case for the MTS assumption, the high-cost selection term is equal to 1.

Substituting the above parameters into expression (8), I find that MTS would hold in the case that

$$0.06 \cdot 1 + 0.94 \cdot \Big(\mathbb{E}[Y_i(v) \mid V_i = 1, \theta_i = \theta_L] - \mathbb{E}[Y_i(v) \mid V_i = 0, \theta_i = \theta_L]\Big) \leq 0\,,$$

or

$$\Big(\mathbb{E}[Y_i(v) \mid V_i = 1, \theta_i = \theta_L] - \mathbb{E}[Y_i(v) \mid V_i = 0, \theta_i = \theta_L]\Big) \leq -0.06$$

This says that, even with the strongest possible positive selection in the subset of individuals with high waiting list costs, the MTS assumption requires only a mild degree of negative selection among those with low waiting list costs. While my proxy for waiting-cost type is clearly far from perfect, this exercise is suggestive that my assumption is robust to this concern about possible positive selection.

The fact that MTS is an assumption about inherent characteristics of individuals in the sample suggests an additional means of assessing its plausibility.[13] If those with poorer potential health outcomes are selected into the voucher program, then it should be the case that voucher recipients have other average characteristics which are also linked to worse health. Returning to Table 2, compare the average values in the SIPP sample for a set of baseline characteristics between voucher recipients and non-recipients. Individuals in households with a voucher are more likely to be female, Black, and are on average older than those in households without a voucher. They also have slightly less schooling (and are on the opposite sides of high school graduation), are less likely to have been employed for the entire previous

[13] This type of exercise has been used elsewhere in the empirical partial identification literature to assess similar monotonicity assumptions (Flores and Flores-Lagunes, 2010, 2013; Germinario et al., 2021b).

month, and are less likely to be food secure; each of these three are statistically significant at the 1% level.

The presence of negative selection is also hinted at by other findings in the empirical literature on housing vouchers. Zhang (2021) also uses SIPP data to study the voucher program and reports several "stylized facts" that voucher holders have lower average employment and earnings, a lower marriage rate, and a higher divorce rate. While these are only suggestive comparisons for assessing MTS—receiving a voucher may influence labor supply decisions, for example—the differences are consistent with poorer underlying health of voucher recipients. Elsewhere, Schwartz et al. (2020) argue in the context of studying student performance of children in voucher households that waiting lists may induce negative selection on educational outcomes, such that cross-sectional comparisons between those with and without vouchers would yield spurious negative relationships, as I find in my SIPP sample for health outcomes.

Under MTS, the worst-case bounds on the mean potential outcomes are tightened in the following way. Returning to the counterfactual conditional expectations $\mathbb{E}[Y_i(1) \mid V_i = 0]$ and $\mathbb{E}[Y_i(0) \mid V_i = 1]$ from expansions as in Equation (3), MTS implies

$$\mathbb{E}[Y_i(1) \mid V_i = 1] \leq \mathbb{E}[Y_i(1) \mid V_i = 0] \leq 1$$

and

$$0 \leq \mathbb{E}[Y_i(0) \mid V_i = 1] \leq \mathbb{E}[Y_i(0) \mid V_i = 0] \ .$$

The first inequality bounds the counterfactual conditional expectation for those observed without a voucher. The upper bound continues to reflect that, because the health outcome is binary, this mean cannot exceed one. The lower bound uses the content of MTS: health outcomes under a voucher would be weakly better on average for those observed in place-based assistance or renting without a subsidy. The second inequality is similar: the counterfactual conditional expectation for those observed with a voucher cannot be less than zero, and negative MTS entails that, on average, health outcomes in the absence of a voucher would be no better for observed voucher holders than for non-voucher holders. Plugging these two inequalities into the iterated expectations expansions allows for bounds on the mean potential outcomes $\mathbb{E}[Y_i(1)]$ and $\mathbb{E}[Y_i(0)]$ due to Manski and Pepper (2000):

$$\begin{gathered} \mathbb{E}[Y_i \mid V_i = 1] \\ \leq \mathbb{E}[Y_i(1)] \leq \\ \mathbb{E}[Y_i \mid V_i = 1] \cdot \mathbf{Pr}[V_i = 1] + \mathbf{Pr}[V_i = 0] \quad , \end{gathered} \tag{9}$$

and

$$\begin{gathered} \mathbb{E}[Y_i \mid V_i = 0] \cdot \mathbf{Pr}[V_i = 0] \\ \leq \mathbb{E}[Y_i(0)] \leq \\ \mathbb{E}[Y_i \mid V_i = 0] \quad , \end{gathered} \tag{10}$$

where I have again used the result that $\mathbb{E}[Y_i(1) \mid V_i = 1]$ and $\mathbb{E}[Y_i(0) \mid V_i = 0]$ are identified by the data.

The identified set for the ATE is obtained by subtracting the upper bound in (10) from the lower bound in (9) for its lower bound, and by taking the difference between the upper bound in (9) and the lower bound in (10) for the upper bound. In this setting with a binary treatment and outcome, the lower bound is equal to the observed difference in mean health status between voucher holders and those without a housing voucher. The upper bound remains identical to that of the worst-case bounds. Thus, the identifying power of MTS comes from narrowing the ATE bounds from below.

5.1.2 Monotone Instrumental Variable

The second assumption is that an individual's self-report of total personal income serves as a valid monotone instrumental variable (MIV), denoted Z_i. Despite the similarity in name, using total personal income as an MIV relies on a weaker assumption than if it were to be used as a traditional instrument. This comes with the advantage that it is plausible in a wider array of situations. Following Manski and Pepper (2000), I make the following assumption:

Assumption 5.2 (Monotone Instrumental Variable)**.** *For any two values z_1 and z_2 of total personal income Z_i such that $z_1 \geq z_2$, it is the case that*

$$\mathbb{E}\left[Y_i(v) \mid Z_i = z_1\right] \geq \mathbb{E}\left[Y_i(v) \mid Z_i = z_2\right]$$

for each treatment status $v \in \mathcal{V} = \{0, 1\}$.

It is worth expanding further upon how Assumption 5.2 differs from the more familiar IV assumptions. First are the objects identified when the assumptions hold. With a valid IV, this is a *local* average treatment effect (LATE) for a subpopulation of "compliers" whose treatment statuses are affected by the instrument (Imbens and Angrist, 1994). For example, if I were to possess a valid IV for housing voucher receipt (say, something that influences eligibility for households) and effects are heterogeneous, then I would point identify an average causal effect for those households impacted by the instrument. With a valid MIV, on the other hand, (partial) identification of the ATE for the full population of interest is obtained. For this reason, an MIV may allow conclusions of wider policy relevance in many cases.

Another important distinction lies in the content of the assumptions themselves. In the LATE framework of Imbens and Angrist (1994), a valid instrument relies heavily on an exclusion restriction, which would replace the weak inequality in the MIV assumption with a strict equality. This conditional mean independence would yield the intuition that an IV affects the outcome only through its effect on the treatment.

A valid MIV, on the other hand, is not required to satisfy anything beyond Assumption 5.2: on average, a higher self-reported income is associated with weakly better potential health outcomes, regardless of voucher status. Note that no restrictions are placed on the relationship between the MIV and the treatment variable. This means that total personal income is allowed to be endogenous with respect to voucher status, or to have a non-monotonic impact on voucher status.

Although the MIV assumption is considerably weaker than the IV assumptions, it is still a substantive statement about potential outcomes that cannot be tested in the data. It is therefore crucial to carefully assess the plausibility of a candidate MIV in the context of its application. Here, it is useful to note that, in economic terms, the assumed relationship between potential health outcomes and self-reported total personal income would be implied by the normality of health as a good for the low-income population I study in this chapter. This condition is reasonable in light of the Grossman (1972) model of health demand which is consistent with the framework of Section 4. Since an increase in total personal income increases full income (unearned income plus the market value of an individual's leisure endowment), the normality of health would imply that total personal income is a valid MIV. Note, however, that a restriction of health as a normal good is stronger than Assumption 5.2, and is therefore sufficient, but not necessary, for potential health outcomes to have a non-negative association with reported total personal income. Indeed, a large empirical literature finds a positive health gradient with respect to income, and supports the MIV assumption even if health may not be a normal good in the proper sense (Gundersen and Kreider, 2009; Kreider et al., 2012). For example, Deaton (2002) and Chetty et al. (2016) document an increasing relationship between life expectancy and income, particularly for those with incomes towards the bottom of the distribution. Case et al. (2002) additionally demonstrate that this gradient likely originates from childhood and accumulates over the life cycle. These results offer strong evidence in favor on the MIV assumption in this context.

The MIV bounds on the ATE are obtained in the following way. First, and without loss of generality,[14] discretize the values of (continuous) total personal income into L bins denoted $\mathcal{B}_1, \ldots, \mathcal{B}_L$. These are arranged such that $z \in \mathcal{B}_1$ and $z' \in \mathcal{B}_2$ implies that $z < z'$. Using these bins, iterate expecations on the mean potential outcomes to write

$$\mathbb{E}\left[Y_i(1)\right] = \sum_{\ell=1}^{L} \mathbb{E}\left[Y_i(1) \mid Z_i \in \mathcal{B}_\ell\right] \cdot \mathbf{Pr}\left[Z_i \in \mathcal{B}_\ell\right] \ , \tag{11}$$

with a similar expansion for $\mathbb{E}\left[Y_i(0)\right]$. The probability $\mathbf{Pr}\left[Z_i \in \mathcal{B}_\ell\right]$ is identified by the data, while I am able to partially identify the conditional expectations $\mathbb{E}\left[Y_i(1) \mid Z_i \in \mathcal{B}_\ell\right]$ using any of the bounds discussed above. In this chapter, I first focus on an MIV combined with the MTS assumption, and so I apply the MTS bounds within each bin. For ease of notation, I will write the MTS lower bound on $\mathbb{E}\left[Y_i(1)\right]$ (or $\mathbb{E}\left[Y_i(0)\right]$) in bin ℓ as $LB^{\ell}_{\text{mts},1}$ $\left(\text{or } LB^{\ell}_{\text{mts},0}\right)$, and likewise for the upper bound.

The resulting MIV+MTS bounds are able to further narrow the identified set. To see this, note that the MIV assumption implies that $\mathbb{E}\left[Y_i(1) \mid Z_i \in \mathcal{B}_\ell\right]$ weakly increases from bin 1 to bin L. At the same time, it also implies that the *lower bound* on this conditional mean must be non-decreasing over the bins, and that its *upper bound* must be non-increasing when moving from bin L to bin 1. Thus, the bounds on $\mathbb{E}\left[Y_i(1) \mid Z_i \in \mathcal{B}_\ell\right]$ can be written:

$$\sup_{\underline{\ell} \leq \ell} \left\{LB^{\underline{\ell}}_{\text{mts},1}\right\} \leq \mathbb{E}\left[Y_i(1) \mid Z_i \in \mathcal{B}_\ell\right] \leq \inf_{\bar{\ell} \geq \ell} \left\{UB^{\bar{\ell}}_{\text{mts},1}\right\} \ , \tag{12}$$

and likewise for $\mathbb{E}\left[Y_i(0) \mid Z_i \in \mathcal{B}_\ell\right]$.

Plugging these bounds into Equation (11) yields the MIV+MTS bounds on $\mathbb{E}\left[Y_i(1)\right]$ (Manski and

[14]Manski and Pepper (2000) derive the MIV bounds by iterating expectations over each point in the support of the MIV. My exposition is equivalent by setting L to the cardinality of the support set (if it is finite) or letting $L \to \infty$ (if the support is not finite). I use the terminology of bins because I believe it allows for clearer intuition in this case. Additionally, feasible estimation of the MIV bounds as described in the next subsection requires discretization of a continuous MIV.

Pepper, 2000):

$$
\begin{aligned}
&\sum_{\ell=1}^{L}\left(\sup_{\underline{\ell}\,\leq\,\ell} LB^{\underline{\ell}}_{\text{mts},1}\right)\cdot \mathbf{Pr}\left[Z_i \in \mathcal{B}_\ell\right] \\
&\qquad\qquad \leq \mathbb{E}\left[Y_i(1)\right] \leq \\
&\sum_{\ell=1}^{L}\left(\inf_{\bar{\ell}\,\geq\,\ell} UB^{\bar{\ell}}_{\text{mts},1}\right)\cdot \mathbf{Pr}\left[Z_i \in \mathcal{B}_\ell\right] .
\end{aligned}
\tag{13}
$$

Bounds on $\mathbb{E}[Y_i(0)]$ are written in a similar fashion. The MIV+MTS bounds are then constructed via the now-familiar order of subtraction between the two sets of bounds. Notice how the MIV can add identifying power relative the the MTS assumption alone. The lower bound on the ATE is constructed by taking a weighted average of the largest MTS lower bounds on $\mathbb{E}[Y_i(1)]$ across the L bins and subtracting a weighted average of the smallest MTS upper bounds on $\mathbb{E}[Y_i(0)]$ over the L bins. The upper bound on the ATE subtracts a weighted average of the greatest lower bounds on $\mathbb{E}[Y_i(0)]$ from a weighted average of the least upper bounds on $\mathbb{E}[Y_i(1)]$. In this way, the non-negative relationship between the MIV and the mean potential outcomes implies that the MIV+MTS lower bound is weakly greater than the MTS lower bound on the ATE and that the MIV+MTS upper bound is weakly less than the MTS upper bound.

5.1.3 Monotone Treatment Response

The final assumption is made directly on how an individual's health responds to receiving a housing voucher. The idea here is that potential health outcomes under voucher receipt are no worse than potential health outcomes without a voucher. Recalling that $Y_i(v) = 1$ denotes good or better health, I assume (Manski and Pepper, 2000):

Assumption 5.3 (Monotone Treatment Response)**.** *Suppose that*

$$Y_i(1) \geq Y_i(0)$$

for all individuals $i = 1, \ldots, N$.

Contrary to MTS and the MIV assumption, note that MTR is assumed to hold at the individual level, rather than at the mean. That is, Assumption 5.3 requires that no individual in a voucher household would have strictly better health in the counterfactual world where she is receiving place-based rental assistance or renting without a subsidy, and that no individual receiving place-based assistance or renting unsubsidized would have strictly worse health in the counterfactual where she is in a voucher household. This is the strongest assumption I make, and its plausibility is not obvious *a priori*.

The two mechanisms through which voucher receipt can impact health—indirectly by relaxing the budget constraint and directly by influencing the housing choice—are at the heart of the MTR assumption. The first follows from the discussion of non-housing expenditures in Section 4. There, households who would have counterfactually been in LIHTC or unsubsidized units are able to now allocate 70% of income towards the non-housing consumption b. If the new bundle also includes more goods from the subset $\tilde{b}$ which are conducive to health, then the voucher-induced income effect will lead the consumer to devote more resources to goods which promote better health (Chen, Flores, and Flores-Lagunes, 2018).[15] The second mechanism is precisely the Moving to Opportunity idea: the voucher may give recipients access to higher quality housing or higher quality neighborhoods. This in turn can promote better health when this quality represents improvements with respect to the WHO housing guidelines.

There are two primary concerns for this direct "housing choice effect" of voucher receipt. First, it might be the case that landlords only accept vouchers at properties with poor physical characteristics and/or properties located in low-amenity neighborhoods. Rosen (2020) describes this type of steering in her case study of HCV in Baltimore. If such lower-quality housing leads to worse health for some voucher holders (say, for example, an apartment building accepting vouchers is in a neighborhood with a high crime rate or with low air quality), then this could represent a violation of MTR.[16]

While the above could be true of the overall supply of housing relative to the supply available to voucher holders, it does not necessarily imply a failure of the assumption. Recall that the sample I use from the SIPP consists of individuals living in households with an income-to-poverty ratio less than 1 and for which the counterfactual housing of voucher holders consists of public housing, project-based Section 8, LIHTC units, or unsubsidized units offered to such low-income populations. A scenario in which the housing options available upon receiving a voucher are strictly worse than in the counterfactual without one seems unlikely in a population at these income levels. Indeed, the very motivation for tenant-based assistance itself (as well as the Moving to Opportunity experiment) was that vouchers offer an improved choice set relative to public housing. Additionally, the requirement that a potential unit pass an inspection by the PHA before a voucher can be used gives at least a nominal floor on the quality.

This question of the relative housing quality for voucher holders has also been well-studied in the empirical literature. Many results offer suggestive evidence that the housing choices available to voucher holders are no worse on average than for non-voucher holders. For example, Lens et al. (2011) use administrative HUD data to study census tracts across 91 large US cities over the period 1999-2001 and find that voucher households were 6.4 and 6.9 percentage points less likely to live in a high-crime neighborhood than, respectively, public housing tenants and residents of Low-Income Housing Tax Credit units (each statistically significant at the 1% level). If other neighborhood amenities are "bundled" with lower crime, then this could be indicative that voucher holders also experience other neighborhood characteristics associated with better health. As one example, Bondy et al. (2018) find a positive link between air pollution and crime in London over 2004-2005. It thus appears reasonable that MTR is not violated through the mechanism of voucher holders being restricted to low-quality neighborhoods. Additionally, the finding by Schwartz et al. (2020) that voucher holders in New York City are less likely to live in a building with hazardous code violations (where the counterfactual was housing in unsubsidized private units) potentially suggests that MTR can also be justified on the margin of physical housing characteristics.

[15] A potential issue here is that voucher receipt likely reduces labor supply (Jacob et al., 2015). This is because (1) the same income effect increases the demand for leisure, and (2) the 30% of income allocated to rent can be understood as a tax on earnings, generating also a substitution effect into more leisure. However, for my purposes, this is not a threat to MTR so long as the reduction in labor supply does not result in a post-transfer income below that which would have been earned in the absence of the voucher (leaving the recipient only able to afford bundles which harm health). I find such a scenario unlikely.

[16] Note that, at this point, I am considering "steering" as a landlord observes a voucher holder and perceives it as a signal for expected profitability from offering a lease. Given the usual meaning of steering based on race, I am cognizant of the fact that what I discuss here may be compounded for Black voucher holders. I estimate separate bounds by race below.

Estimates from the spatial equilibrium model of Leung et al. (2012) mentioned in Section 2 provide similar evidence. With two household types (skilled/high-income and unskilled/low-income) and two neighborhoods, their model predicts higher average lot sizes and higher quality neighborhoods (measured by school quality)[17] for voucher households relative to public housing participants. To the extent that more living space is conducive to better health outcomes (*e.g.*, less crowding in the home [Solari and Mare, 2012]), their result that voucher holders end up with higher average lot sizes versus public housing tenants would be consistent with the MTR assumption.

A second potential concern for the housing choice effect is related to the idea that the choice of housing for voucher recipients is a bundle of both physical structure quality (or housing services) and neighborhood amenities. For example, a voucher recipient with children may prefer to locate in a lower-quality structure that, at the same time, allows access to a high quality school district.[18] In this sense, some voucher holders may (perhaps indirectly) choose to trade off some health relative to their no-voucher counterfactual.

This simultaneous choice of structure quality and neighborhood characteristics is captured by the housing choice model adapted into the conceptual framework of Section 4. Geyer (2017) estimates the original model using administrative data from the Housing Authority of the City of Pittsburgh for voucher recipients in 2006. Relevant to the current paper, the estimation recovers three cross-elasticities for neighborhood amenities and the demand for housing services. For school quality (measured by average standardized test scores for eighth graders) and average public transit commute times, the elasticities are estimated at 0.11 and 0.03, respectively. These indicate that both amenities are complementary goods with respect to housing services: with the decrease in the relative price of housing services induced by the voucher, recipient households consume more of both the amenities and housing services.[19] This would plausibly imply that MTR is not violated through voucher holders trading off health by choosing a lower-quality structure in favor of neighborhood amenities.

Like with MTS, the bounds on the treatment effect of interest under the MTR assumption are tightened relative to the worst-case bounds. First note that the restriction of $Y_i(1) \geq Y_i(0)$ implies that $\mathbb{E}[Y_i(1) \mid V_i = v] \geq \mathbb{E}[Y_i(0) \mid V_i = v]$ for any v; that is, MTR also holds in expectation, conditional on either voucher status. This is again useful with respect to the unobservable conditional means $\mathbb{E}[Y_i(1) \mid V_i = 0]$ and $\mathbb{E}[Y_i(0) \mid V_i = 1]$. With MTR, I can conclude that:

$$\mathbb{E}[Y_i(0) \mid V_i = 0] \leq \mathbb{E}[Y_i(1) \mid V_i = 0] \leq 1$$

and

$$0 \leq \mathbb{E}[Y_i(0) \mid V_i = 1] \leq \mathbb{E}[Y_i(1) \mid V_i = 1]\ .$$

Using once more the iterated expecations expansion of Equation (3), the mean potential outcomes can be

[17] Recall from Section 1.2 that empirical evidence on this particular margin of school quality from Horn, Ellen, and Schwartz (2014) is somewhat mixed.

[18] I thank Amy Ellen Schwartz for suggesting further consideration of this issue.

[19] The third amenity considered is the proportion of neighborhood area designated as public parks. The estimated cross-elasticity is -0.001, suggesting that access to parks and housing services are slight substitutes. However, since the elasticity is small in magnitude and parks may be conducive to better health, I do not view this as a contradiction of my "housing choice effect" for MTR.

set identified as (Manski, 1997; Manski and Pepper, 2000):

$$
\begin{aligned}
\mathbb{E}\left[Y_i \mid V_i = 1\right] \cdot \mathbf{Pr}\left[V_i = 1\right] + \mathbb{E}\left[Y_i \mid V_i = 0\right] \cdot \mathbf{Pr}\left[V_i = 0\right] \\
\leq \mathbb{E}\left[Y_i(1)\right] \leq \\
\mathbb{E}\left[Y_i \mid V_i = 1\right] \cdot \mathbf{Pr}\left[V_i = 1\right] + \mathbf{Pr}\left[V_i = 0\right] \quad ,
\end{aligned}
\tag{14}
$$

and

$$
\begin{aligned}
\mathbb{E}\left[Y_i \mid V_i = 0\right] \cdot \mathbf{Pr}\left[V_i = 0\right] \\
\leq \mathbb{E}\left[Y_i(0)\right] \leq \\
\mathbb{E}\left[Y_i \mid V_i = 0\right] \cdot \mathbf{Pr}\left[V_i = 0\right] + \mathbb{E}\left[Y_i \mid V_i = 1\right] \cdot \mathbf{Pr}\left[V_i = 1\right] \quad .
\end{aligned}
\tag{15}
$$

The bounds on the ATE are computed via the same subtractions as in the previous subsection. In the binary treatment setting, it can be plainly seen that the lower bound on the ATE is equal to zero, as both the lower bound in (14) and the upper bound in (15) are equivalent to the observed mean of the outcome. This is consistent with the MTR assumption, which implies $\mathbb{E}\left[Y_i(1)\right] \geq \mathbb{E}\left[Y_i(0)\right]$, and thus that the average effect is non-negative. The MTR upper bound remains the same as both the worst-case and MTS upper bounds.

In many empirical applications (*e.g.*, Manski and Pepper, 2000; Gundersen and Kreider, 2009; de Haan, 2011; Germinario et al., 2021a,b), the bounds on the ATE are narrowed further by imposing MTS and MTR simultaneously to obtain MTS+MTR bounds. The combination, however, only adds identifying power when the two assumptions are in the same direction (*i.e.*, when positive MTS and positive MTR or negative MTS and negative MTR are assumed). In the present application, I assume (weakly) opposite directions for MTS and MTR, and so the MTS+MTR bounds can be no narrower than the tighter of the MTS and MTR bounds separately.

To be more precise, consider first the upper bound when MTS and MTR are imposed together. As noted above, the MTS and MTR upper bounds on the ATE are identical, and so this common value is also the MTS+MTR upper bound. For the lower bound, it was previously seen that the MTS lower bound is equal to the observed difference in mean health status between voucher recipients and non-recipients, while the MTR lower bound is equal to zero. The greatest lower bound when combining these is therefore the maximum of the MTS and MTR lower bounds. If the MTS lower bound is larger, then the MTS+MTR bounds are identical to the MTS bounds; if the MTR bound (zero) is larger, then the MTS+MTR bounds are equivalent to the MTR bounds. Since this is an empirical question, I still report results for MTS+MTR as a separate assumption below.

Finally, all three assumptions can be imposed simultaneously to identify MIV+MTS+MTR bounds on the ATE. The identification is identical to the discussion above surrounding inequalities (12) and (13), with the only difference being that the MTS+MTR bounds are applied within the MIV bins in this case.

The sharpest possible identified set under MIV+MTS+MTR is obtained with one final step. For this, observe that the MTR assumption also applies inside each bin of the MIV (McCarthy et al., 2015), or that $\mathbb{E}\left[Y_i(1) \mid Z_i \in \mathcal{B}_\ell\right] \geq \mathbb{E}\left[Y_i(0) \mid Z_i \in \mathcal{B}_\ell\right]$, for each ℓ. This yields a useful restriction that the MTS+MTR lower bound on $\mathbb{E}\left[Y_i(1)\right]$ in bin ℓ cannot be less than MTS+MTR upper bound on $\mathbb{E}\left[Y_i(0)\right]$ in that same bin.

5.2 Estimation and Inference Procedure

The worst-case, MTS, MTR, and MTS+MTR bounds all are estimated by replacing the expectations and probabilities with their sample counterparts.[20] I also compute 95% Imbens and Manski (2004) confidence intervals for the ATE under each assumption.

To operationalize the MIV+MTS+MTR bounds as in the inequality of (13), the supremum and infimum operators are replaced by maxima and minima, respectively. I also discretize total personal income into $L = 5$ bins, where the cutoffs to define the bins are based on the quintiles of its empirical distribution. This choice of the number of bins is meant to balance two sources of bias. Choosing too few bins does not utilize the identifying power of the MIV discussed at the end of Section 5.1 above. Using a larger number of bins can therefore reduce bias. At the same time, increasing the number of bins exponentially increases the number of objects that must be estimated (see Appendix A). Simulation evidence reported in Germinario, Flores, and Flores-Lagunes (2021) suggests 5 bins as a reasonable choice: compared to estimates using 2 bins, the lower bound using 5 bins was 124% more likely to capture the true lower bound and the upper bound was 29% more likely to capture the true upper bound.

Regardless of the choice for the number of bins, implementing the MIV+MTS+MTR estimators for the bounds in (13) introduces two complications. First, these are an instance of intersection bounds (due to the maximum and minimum operators), and it is well-documented that such bounds suffer from bias in finite samples (Manski and Pepper, 2000; Chernozhukov et al., 2013; Flores and Flores-Lagunes, 2013; Flores and Chen, 2018). This appears as an upward bias for the lower bound and a downward bias in the upper bound. Left uncorrected, this results in bounds which are narrower than the true identified set, which can suggest misleading conclusions on treatment effects. Compounding the problem, Hirano and Porter (2012) demonstrate that there are no locally asymptotically unbiased estimators when the parameters contain non-differentiable functionals of the data (such as a maximum or minimum). The second issue is that confidence intervals for the ATE with the desired coverage cannot be obtained with the standard bootstrap.

I address both issues by applying the Chernozhukov, Lee, and Rosen (2013) procedure (hereafter, CLR) to obtain bias-corrected estimates of the MIV+MTS and MIV+MTS+MTR bounds and valid confidence intervals for the ATE under those combined assumptions. This allows for estimates of the upper and lower bounds on the ATE which satisfy a half-median unbiasedness property. This means that the estimate of the lower bound is below the true lower bound with a probability of at least 0.5, and the estimate of the upper bound exceeds the true upper bound with a probability of 0.5 or above. In light of the Hirano and Porter (2012) result, this is a desirable property for the estimates. The simulation evidence in Germinario, Flores, and Flores-Lagunes (2021) also demonstrates that substantial bias can occur without the CLR correction. In estimates using 5 bins for the MIV, the CLR-corrected lower bound was below the true lower bound in 48% of the simulations; the CLR-corrected upper bound was above the true upper bound in 97% of the simulations. Without the correction, these were 8.4% and 72.4%, respectively. Further details on my implementation of the CLR procedure are contained in Appendix A.

[20]All estimation and inference described in this subsection is implemented in Stata using the package `mpclr` by Germinario et al. (2021).

6 Results

6.1 Main Results for Self-Reported Health

Estimation results for the effect of receiving a housing voucher on the self-reported health status of individuals within the household are summarized in Table 3. The first column contains the naïve OLS result from regressing health status on the indicator for voucher receipt; under negative MTS, this coefficient is a downward-biased (*i.e.*, more negative) estimate of the ATE. This indicates the large gap in health status between those with and without a housing voucher, even in my sample focused on those in households with an income-to-poverty ratio less than 1. That individuals in a household without a voucher are 20.9 percentage points more likely to report good or better health is likely due to the negative selection into the voucher program which I assume is present in results making use of MTS.

Table 3: Bounds on the ATE for Self-Reported Health: Total Personal Income as MIV, 5 Bins

	(1) OLS	(2) Worst-Case	(3) MTS	(4) MIV+MTS	(5) MTR	(6) MTS+MTR	(7) MIV+MTS+MTR
$ATE(1,0)$	-0.209^{***}	[-0.753 , 0.247]	[-0.210 , 0.247]	[-0.092 , 0.213]	[0.000 , 0.247]	[0.000 , 0.247]	[0.048 , 0.213]
	(0.033)	(-0.767 , 0.260)	(-0.266 , 0.261)	(-0.195 , 0.239)	(0.000 , 0.260)	(0.000 , 0.261)	(0.023 , 0.238)

[·] denotes estimated bounds in Columns 2, 3, 5, and 6, or half-median unbiased MIV bounds in Columns 4 and 7. (·) denotes the standard error in Column 1, or 95% CLR confidence intervals on the ATE in Columns 2-7. Bounds are estimated using 3529 observations. The variance-covariance matrix in the CLR first step is based on 240 BRR replications, where the replicate weights provided by SIPP are used.

Column 2 reports the worst-case bounds on the ATE, where I make use only of the fact that the outcome is binary. As is typical, these bounds are not particularly informative of the sign of the treatment effect. While the lower bound of a 75.3 percentage point decline in the probability of good or better health is not feasible relative to the baseline for voucher holders, the upper bound only allows for up to a 24.7 percentage point increase in the same probability.

It is interesting that this worst-case upper bound—where, again, I have not yet imposed any substantive assumptions—excludes the 26.4 percentage point ATT estimated by Katz et al. (2001) for MTO in Boston, and this point estimate also narrowly lies outside the 95% confidence interval I construct for the ATE. This may suggest three things. First, since the MTO experiment had only one-sided non-compliance (no members of the control group could receive the treatment, but not all those assigned to the voucher group accepted one), the estimated ATT is a LATE for those who complied with treatment assignment. The effect for these compliers may be larger than the effect for the population at large. Second, Katz et al. (2001) analyze only data from Boston, and it may be the case that vouchers have a greater effect on health in that city than at the national level. Finally, it is possible that the exclusion restriction used to identify the LATE—treatment assignment impacts health only through its effect on voucher uptake—was violated in the MTO experiment.[21] In this case, an estimate outside of the worst-case bounds could also reflect upward bias in the estimate of the LATE.

Turning to the more substantive assumptions, the estimated bounds from imposing non-positive MTS are reported in Column 3. These bounds account for negative selection into voucher receipt, and are

[21]This is similar to a point Finkelstein et al. (2012) and Chen, Flores, and Flores-Lagunes (2018) make in the context of the Oregon Health Insurance Experiment: learning about the housing voucher program through treatment assignment may increase awareness of other public assistance programs, which in turn can affect health, even if the voucher is not ultimately used.

considerably tightened from below. Here, I am able to rule out any negative health effect from receiving a voucher beyond 21.0 percentage points. In other words, any detrimental impact is *at worst* equal to the observed difference in the probability of good health across voucher statuses. The MTS upper bound continues to exclude any positive effect of housing voucher receipt beyond 24.7 percentage points, or a 76% percent increase using the baseline probability of good health observed for voucher holders.

In Column 4 are results combining the total personal income MIV with MTS. The MIV+MTS lower bound is increased relative to MTS alone, limiting any negative effect to 9.2 percentage points. The bounds are also tightened from above, with the upper bound excluding a positive effect larger than 21.3 percentage points.

The estimates of the MTR and combined MTS+MTR bounds in Columns 5 and 6 are identical, due to the impact of the MTR assumption. In both cases, MTR rules out any negative effect of the vouchers on health, while the upper bounds continue to be those of the worst-case upper bounds, excluding a positive effect on the likelihood of good or better health greater than 24.7 percentage points.

My preferred results using the total personal income MIV along with both MTS and MTR are presented in Column 7, where I have used 5 bins in the estimation of the bounds. These MIV+MTS+MTR bounds provide evidence that the causal effect of housing vouchers on the self-reported health of those in recipient households is positive and statistically significant. Under the relatively weak assumptions considered in this paper, I am able to conclude that receiving a housing voucher increases the likelihood of reporting good or better health by at least 4.8 percentage points, and by no more than 21.3 percentage points. With the observed proportion of those in voucher households reporting good or better health equal to 57.1, this lower bound implies that these individuals are at least 9.2% more likely to report good health than they would have been in the absence of the voucher. The upper bound suggests that voucher holders are at most 59.5% more likely to report good health than they would have been without the voucher.

With the tightened upper bound under MIV+MTS+MTR, the 26.4 percentage point LATE estimated using the Boston MTO data is further outside of my identified set for the population ATE. This reinforces the idea above that the targeted population in the experiment differs somewhat from the national-level population of renters with family income-to-poverty ratios less than 1 that I study in this chapter. It is comforting for my analysis, however, that the range of values I identify for the causal effect (ATE) of housing vouchers on self-reported health is of a magnitude similar to a treatment effect (ATT/LATE) identified under random assignment.

6.1.1 Subgroup Analyses

Next, I consider the same assumptions for demographic subgroups as for the overall sample. Though comparisons of partially-identified treatment effects are difficult across subgroups of sex or racial/ethnic identity because bounds often overlap, there are a few potentially interesting findings.

The two panels of Table 4 describe the main objects of interest for the subsamples by sex as self-reported in the SIPP. For the female subgroup, 75% report good or better health, while 78.5% do so in the male subgroup. The proportion of females in voucher households (9.1%) is nearly double that of the proportion of males in voucher households (5.7%).

In Table 5, I report the estimation results by sex. The top panel consists of estimates for the subsample of female household residents, and the bottom panel for male residents. Column 1 again reports the

Table 4: Observed Probabilities of Good or Better Health and Distribution of Vouchers, by Sex

t	$\mathbb{E}[Y \mid T=t]$	$\mathbf{Pr}[T=t]$	N_t
Female			
Voucher	0.545	0.091	197
	(0.037)	(0.007)	-
No Voucher	0.771	0.909	1897
	(0.010)	(0.007)	-
Total	0.750	1.000	2094
	(0.010)	-	-

t	$\mathbb{E}[Y \mid T=t]$	$\mathbf{Pr}[T=t]$	N_t
Male			
Voucher	0.629	0.057	82
	(0.059)	(0.008)	-
No Voucher	0.794	0.943	1353
	(0.011)	(0.008)	-
Total	0.785	1.000	1435
	(0.011)	-	-

Estimates use the annual SIPP weights for individual-level analysis. Standard errors in parentheses are based on 240 BRR replications with the SIPP replicate weights.

difference in means across voucher status. For both the female and male subgroups, the negative and statistically significant observed health gap remains, with females in voucher households 22.5 percentage points less likely to report good or better health, and males 16.5 percentage points less likely to report good health. This difference in the unadjusted voucher health gap may indicate that the degree of negative selection into the program is stronger among females. The worst-case bounds in Column 2 are quite similar for female and male residents, and these also do not much differ from the same bounds for the full sample. As usual, the worst-case bounds cannot identify the sign of the ATE. However, I am again able to exclude large negative effects beyond a 74.2 (77.0) percentage point decrease for female (male) household members, while also ruling out positive effects greater than 25.8 percentage points (89.9%) for females and greater than 23.0 percentage points (57.6%) among males.

Table 5: Bounds on the ATE for Self-Reported Health, by Sex: Total Personal Income as MIV, 5 Bins

	(1) OLS	(2) Worst-Case	(3) MTS	(4) MIV+MTS	(5) MTR	(6) MTS+MTR	(7) MIV+MTS+MTR
ATE for Female	-0.225^{***}	[-0.742 , 0.258]	[-0.225 , 0.258]	[-0.102 , 0.213]	[0.000 , 0.258]	[0.000 , 0.258]	[0.062 , 0.213]
Subsample	(0.039)	(-0.758 , 0.275)	(-0.291 , 0.275)	(-0.209 , 0.246)	(0.000 , 0.275)	(0.000 , 0.275)	(0.028 , 0.245)
ATE for Male	-0.165^{***}	[-0.770 , 0.230]	[-0.165 , 0.230]	[-0.109 , 0.210]	[0.000 , 0.230]	[0.000 , 0.230]	[0.013 , 0.210]
Subsample	(0.059)	(-0.789 , 0.249)	(-0.269 , 0.251)	(-0.258 , 0.253)	(0.000 , 0.249)	(0.000 , 0.251)	(0.000 , 0.252)

[·] denotes estimated bounds in Columns 2, 3, 5, and 6, or half-median unbiased MIV bounds in Columns 4 and 7. (·) denotes the standard error in Column 1, or 95% CLR confidence intervals on the ATE in Columns 2-7. Bounds are estimated using 2094 observations for the female subsample and 1435 observations for the male subsample. The variance-covariance matrix in the CLR first step is based on 240 BRR replications, where the replicate weights provided by SIPP are used.

MTS tightens each set of bounds in Column 3, as the assumed negative selection places the observed differences in means as the lower bound on the ATE. The addition of the total personal income MIV in Column 4 further increases the lower bound so that I can rule out a negative effect beyond 10.2 percentage points for females and beyond 10.9 percentage points for males. The MTR and MTS+MTR bounds in Columns 5 and 6 alternatively tighten the bounds through the implication of a non-negative treatment effect. The upper bounds in Columns 3, 5, and 6 remain identical to the worst-case upper bounds. Adding the total personal income MIV to MTS+MTR in Column 7 yields bounds which can identify the ATE as strictly positive for both female and male household members, although the CLR confidence intervals indicate that the effect can be statistically distinguished from zero at the 5% level only in the female subsample. For females, I estimate a lower bound that indicates receiving a housing voucher increases the probability of good or better health by at least 6.2 percentage points, which is somewhat larger in magnitude compared to the lower bound for the overall population studied in this chapter. Relative to the baseline likelihood of good health in the female subsample, this represents an increase of at least 12.8%. The upper bound for females allows for up to a 21.3 percentage point (59.5%) increase in the probability of reporting good health; this is equal to the upper bound for the full population, meaning there is a slightly narrower set of possible values for females than for the overall sample. For males, on the other hand, the estimated lower bound reflects an increase in the probability of good health by at least 1.3 percentage points (2.1%), but again a zero effect cannot be statistically ruled out. The greater lower bound for females does not necessarily imply a greater positive effect, as there is still a substantial amount of overlap between the two identified sets. However, it *can* rule out a range of smaller effects (*i.e.*, between 1.3 and 6.2 percentage points) included in the MIV+MTS+MTR bounds for males.

Next, the three panels of Table 6 report the conditional probabilities of good self-reported health and the treatment probabilities by racial/ethnic identity (Black, non-Hispanic; white, non-Hispanic; and Hispanic).[22] A few things are noteworthy. First is that a substantially higher proportion of Black individuals live in a voucher household compared to both white and Hispanic individuals, which is consistent with the overall makeup of voucher recipients.[23] Second is that the unadjusted gap in the probability of good health between voucher holders and non-voucher holders is substantially smaller among Black recipients than among white or Hispanic recipients. Combined with the results I describe below, I believe this may signal that the positive effect of vouchers on health is larger for Black recipients. In other words, receiving a voucher has a large enough effect within this subpopulation that the observed mean difference in health status has substantially shrunk.

Table 7 contains estimated bounds for the ATE of voucher receipt on self-reported health separately for Black, white, and Hispanic household members. First, Column 1 again shows that the observed mean difference in the likelihood of good or better health for those in voucher households versus non-voucher households remains negative within each subgroup, although here it is only statistically significant for white and Hispanic household members.

The worst-case bounds in Column 2 are broadly similar in the white and Hispanic subsamples, and these are in turn similar to the worst-case bounds for the overall sample. The bounds for Black residents are shifted rightward (*i.e.*, contain fewer negative possible ATEs) and rule out a health detriment of more than 67.2 percentage points or a health benefit of more than 32.8 percentage points.

Estimates in Column 3 are narrowed substantially for Black household members under MTS. By assuming weakly negative selection into the voucher program, I can rule out a negative effect on the probability of good health beyond 2.7 percentage points. These bounds are narrowed for the white and Hispanic subgroups as well, albeit to a lesser extent.

Notably, the MIV+MTS bounds in Column 4 are able to exclude zero for Black voucher holders. That is, even without my strongest assumption of MTR, I estimate that the probability of good health is increased by at least 10.2 percentage points (16.6%) for this subpopulation. I cannot, however, conclude

[22]Recall from Section 3 that I have an insufficient sample size to estimate bounds for the Asian subsample.

[23]In terms of the racial/ethnic breakdown of housing voucher recipients in my sample, 45.5% are Black, 23.4% are white, and 22.8% are Hispanic. Compared to the figures reported by the National Low Income Housing Coalition, my sample has a nearly identical proportion of Black voucher recipients (45%), a somewhat lower proportion of white recipients (35%), and a somewhat higher proportion of Hispanic recipients (16%).

Table 6: Observed Probabilities of Good or Better Health and Distribution of Vouchers, by Racial/Ethnic Identity

t	$\mathbb{E}[Y \mid T = t]$	$\mathbf{Pr}[T = t]$	N_t
Black			
Voucher	0.715	0.154	119
	(0.044)	(0.017)	-
No Voucher	0.742	0.846	638
	(0.018)	(0.017)	-
Total	0.738	1.000	757
	(0.017)	-	-

t	$\mathbb{E}[Y \mid T = t]$	$\mathbf{Pr}[T = t]$	N_t
White			
Voucher	0.434	0.040	71
	(0.058)	(0.006)	-
No Voucher	0.767	0.960	1520
	(0.013)	(0.008)	-
Total	0.753	1.000	1591
	(0.013)	-	-

t	$\mathbb{E}[Y \mid T = t]$	$\mathbf{Pr}[T = t]$	N_t
Hispanic			
Voucher	0.384	0.079	64
	(0.065)	(0.008)	-
No Voucher	0.809	0.921	770
	(0.015)	(0.008)	-
Total	0.775	1.000	834
	(0.016)	-	-

Estimates use the annual SIPP weights for individual-level analysis. Standard errors in parentheses are based on 240 BRR replications with the SIPP replicate weights.

Table 7: Bounds on the ATE for Self-Reported Health, by Race: Total Personal Income as MIV, 5 Bins

	(1) OLS	(2) Worst-Case	(3) MTS	(4) MIV+MTS	(5) MTR	(6) MTS+MTR	(7) MIV+MTS+MTR
ATE for Black	-0.027	[-0.672 , 0.328]	[-0.027 , 0.328]	[0.102 , 0.269]	[0.000 , 0.328]	[0.000 , 0.328]	[0.117 , 0.269]
Residents	(0.045)	(-0.700 , 0.357)	(-0.106 , 0.358)	(-0.005 , 0.329)	(0.000 , 0.358)	(0.000 , 0.348)	(0.019 , 0.329)
ATE for White	-0.333^{***}	[-0.759 , 0.241]	[-0.333 , 0.241]	[-0.234 , 0.213]	[0.000 , 0.241]	[0.000 , 0.241]	[0.026 , 0.213]
Residents	(0.060)	(-0.780 , 0.262)	(-0.436 , 0.263)	(-0.439 , 0.246)	(0.000 , 0.263)	(0.000 , 0.263)	(0.000 , 0.244)
ATE for Hispanic	-0.424^{***}	[-0.793 , 0.207]	[-0.425 , 0.207]	[-0.418 , 0.137]	[0.000 , 0.207]	[0.000 , 0.207]	[0.022 , 0.137]
Residents	(0.068)	(-0.819 , 0.232)	(-0.541 , 0.233)	(-0.560 , 0.192)	(0.000 , 0.233)	(0.000 , 0.233)	(0.000 , 0.194)

[·] denotes estimated bounds in Columns 2, 3, 5, and 6, or half-median unbiased MIV bounds in Columns 4 and 7. (·) denotes the standard error in Column 1, or 95% CLR confidence intervals on the ATE in Columns 2-7. Bounds are estimated using 757 observations for the Black subsample, 1591 observations for the white subsample, and 834 observations for the Hispanic subsample. The variance-covariance matrix in the CLR first step is based on 240 BRR replications, where the replicate weights provided by SIPP are used.

that the effect is statistically significant at the 5%, as the 95% CLR confidence interval includes zero. However, the lower bound on the 90% confidence interval (not reported in the table) is equal to 0.014, meaning that the ATE is statistically significant at the 10% level.

The full MIV+MTS+MTR set of assumptions reveals potentially interesting results in Column 7. First, the identified sets for each group are estimated to exclude zero, meaning that, under the assumptions, vouchers have a positive effect on health status across all races in the sample (though the effect is not statistically significant at the 5% level for the white or Hispanic subgroups). The second item of note is that the lower bound for Black household members at 11.7 percentage points is more than double the magnitude of the lower bound in the overall sample, and represents at least a 19.6% relative increase in the probability of good health for this subpopulation. As with the results by sex, the larger lower bound does not definitively mean that the ATE for Black voucher holders is larger than for white or Hispanic voucher holders. However, the extent of overlap between the identified sets is lower for these results by racial/ethnic identity. The bounds for Black and white household members share only the values between an 11.7 and a 21.3 percentage point increase; for Black and Hispanic household members, the bounds overlap only for increases of 11.7 to 13.7 percentage points.

6.2 Results for Hospitalization

I now turn to nights spent in the hospital during the previous year as a secondary, more objective measure of overall health. To maintain the convention that $Y_i(v) = 1$ represents good health, I binarize the hospitalization variable such that $Y_i = 1$ if the individual reported zero nights spent in the hospital, and $Y_i = 0$ if the individual was hospitalized for one or more nights. In this way, the MTS, MIV, and MTR assumptions are invoked in the same way as discussed in Section 5.1 for self-reported health, and I continue to identify bounds on the ATE of voucher receipt on the probability of good health.

Table 8 reports the estimates of the conditional expectations and probabilities which are the key building blocks of the bounds. The sample from the SIPP is identical to the one used above, and so the proportion of voucher holders remains 7.7%. Interesting here is that the difference in the observed probability of no hospitalizations is much smaller than the difference for self-reported good health: 80.4% of voucher holders report zero nights spent in the hospital versus 86.1% of non-voucher holders.

Table 8: Observed Probabilities of Zero Hospitalizations and Distribution of Vouchers

t	$\mathbb{E}[Y \mid T = t]$	$\mathbf{Pr}[T = t]$	N_t
Voucher	0.804	0.077	279
	(0.025)	(0.006)	-
No Voucher	0.861	0.923	3250
	(0.007)	(0.006)	-
Total	0.857	1.000	3529
	(0.007)	-	-

Estimates use the annual SIPP weights for individual-level analysis. Standard errors in parentheses are based on 240 BRR replications with the SIPP replicate weights.

The main estimation results for the full sample are in Table 9. Column 1 regresses the no-hospitalizations indicator on the voucher status variable. This shows that the unadjusted 5.7 percentage point gap in the probability of zero hospitalizations is statistically significant at the 5% level. As before, I argue that this is likely a result of negatively-selected recipients.

Table 9: Bounds on the ATE for Zero Hospitalizations: Total Personal Income as MIV, 5 Bins

	(1) OLS	(2) Worst-Case	(3) MTS	(4) MIV+MTS	(5) MTR	(6) MTS+MTR	(7) MIV+MTS+MTR
$ATE(1,0)$	-0.057^{**}	[-0.810 , 0.190]	[-0.057 , 0.190]	[-0.074 , 0.184]	[0.000 , 0.190]	[0.000 , 0.190]	[0.001 , 0.184]
	(0.025)	(-0.822 , 0.202)	(-0.100 , 0.203)	(-0.122 , 0.202)	(0.000 , 0.202)	(0.000 , 0.203)	(0.000 , 0.201)

[·] denotes estimated bounds in Columns 2, 3, 5, and 6, or half-median unbiased MIV bounds in Columns 4 and 7. (·) denotes the standard error in Column 1, or 95% CLR confidence intervals on the ATE in Columns 2-7. Bounds are estimated using 2094 observations for the female subsample and 1435 observations for the male subsample. The variance-covariance matrix in the CLR first step is based on 240 BRR replications, where the replicate weights provided by SIPP are used.

The worst case bounds on the ATE are presented in Column 2. Using no information other than the binary nature of the outcome, the upper bound indicates that I can exclude any positive effect larger than 19.0 percentage points. MTS is imposed in Column 3, which dramatically increases the lower bound: at worst, voucher receipt makes zero hospitalizations 5.7 percentage points (7.1%) less likely. Using total personal income for the MIV+MTS bounds in Column 4 results in slightly wider bounds than MTS alone, suggesting that the MIV may not have identifying power for the lower bound on the effect for hospitalizations in the full sample. The lower bound rules out a negative effect larger than 7.4 percentage points (9.2%) and a positive effect larger than 18.4 percentage points (29.7%).

Columns 5 and 6 respectively impose MTR and the combination of MTS+MTR, and the estimated bounds are identical. The MTR assumption restricts the effect of vouchers to be non-negative, and the upper bounds allow for up to a 19.0 percentage point increase in the probability of no hospitalizations. Finally, my preferred MIV+MTS+MTR bounds which combine all three assumptions are presented in Column 7. Here, the lower bound indicates that receiving a housing voucher causes the likelihood of no hospitalizations to increase by at least 0.1 percentage points (0.1%). However, the 95% (and unreported 90%) CLR confidence interval includes zero, such that I cannot statistically distinguish the ATE from zero at conventional levels. Despite this, the bounds remain informative in the sense that I can rule out large positive effects in excess of a 29.7% relative increase in the probability.

6.2.1 Subgroup Analyses

I next apply the monotonicity assumptions within the same demographic subgroups based on sex or racial/ethnic identity as above for the no-hospitalizations outcome. There are again some potentially informative findings.

The panels of Table 10 report the point estimates of the key objects for the bounds on the ATE. The size of the unadjusted gap in the probability of zero hospitalizations across voucher statuses is nearly equal for both female and male household members at around 5 percentage points. There is, however, a difference in the level of this probability by sex. For both those with and without vouchers, females in the sample have lower likelihoods of no hospitalizations versus males. Overall, males are 6.6 percentage points more likely to have avoided a night in the hospital over the previous year.

Table 10: Observed Probabilities of Zero Hospitalizations and Distribution of Vouchers, by Sex

t	$\mathbb{E}[Y \mid T = t]$	$\mathbf{Pr}[T = t]$	N_t
Female			
Voucher	0.783	0.091	197
	(0.032)	(0.007)	-
No Voucher	0.834	0.909	1897
	(0.009)	(0.007)	-
Total	0.829	1.000	2094
	(0.010)	-	-

t	$\mathbb{E}[Y \mid T = t]$	$\mathbf{Pr}[T = t]$	N_t
Male			
Voucher	0.851	0.057	82
	(0.041)	(0.008)	-
No Voucher	0.898	0.943	1353
	(0.009)	(0.008)	-
Total	0.895	1.000	1435
	(0.009)	-	-

Estimates use the annual SIPP weights for individual-level analysis. Standard errors in parentheses are based on 240 BRR replications with the SIPP replicate weights.

The estimates for the effect of voucher receipt are contained in Table 11. Column 1 reports the observed differences in the probability of no hospitalizations. The difference is very slightly larger for females at 5.1 percentage points versus 4.7 percentage points for males; however, neither unadjusted gap is statistically significant at the 5% level. Next, Column 2 reports estimates of the worst-case bounds on the ATE. For female household members, I can rule out an effect larger than 22.2 percentage points (39.6%) based on the upper bound; for males, positive effects beyond 14.5 percentage points (20.5%) can be excluded. Imposing MTS in Column 3 renders the observed 5.1 and 4.7 percentage point differences in the probability of zero hospitalizations as the lower bounds on the ATE for females and males, respectively.

Table 11: Bounds on the ATE for Zero Hospitalizations, by Sex: Total Personal Income as MIV, 5 Bins

	(1) OLS	(2) Worst-Case	(3) MTS	(4) MIV+MTS	(5) MTR	(6) MTS+MTR	(7) MIV+MTS+MTR
ATE for Female	−0.051	[-0.778 , 0.222]	[-0.051 , 0.222]	[-0.037 , 0.214]	[0.000 , 0.222]	[0.000 , 0.222]	[0.000 , 0.214]
Subsample	(0.032)	(-0.793 , 0.237)	(-0.107 , 0.238)	(-0.094 , 0.238)	(0.000 , 0.237)	(0.000 , 0.238)	(0.000 , 0.237)
ATE for Male	−0.047	[-0.855 , 0.145]	[-0.047 , 0.145]	[-0.057 , 0.130]	[0.000 , 0.145]	[0.000 , 0.145]	[0.007 , 0.130]
Subsample	(0.041)	(-0.872 , 0.162)	(-0.121 , 0.164)	(-0.140 , 0.154)	(0.000 , 0.162)	(0.000 , 0.164)	(0.000 , 0.153)

[·] denotes estimated bounds in Columns 2, 3, 5, and 6, or half-median unbiased MIV bounds in Columns 4 and 7. (·) denotes the standard error in Column 1, or 95% CLR confidence intervals on the ATE in Columns 2-7. Bounds are estimated using 2094 observations for the female subsample and 1435 observations for the male subsample. The variance-covariance matrix in the CLR first step is based on 240 BRR replications, where the replicate weights provided by SIPP are used.

The addition of the total personal income MIV in Column 4 yields the MIV+MTS bounds. For females, the lower bound is increased to -3.7 percentage points (-4.7%), ruling out any larger negative effects on the probability of avoiding a night in the hospital. The upper bound is slightly decreased relative to MTS alone, such that an effect larger than 21.4 percentage points can be excluded. For males, the lower bound decreases slightly and indicates that decreases in the likelihood of no hospitalizations of more than 5.7 percentage points (6.7%) can be ruled out. This again likely indicates that the MIV does not add identifying power for bounding the effect from below in the male subsample. The identified set is slightly tightened from above, such that positive effects of more than 13.0 percentage points (18.0%) can be excluded.

The MTR and MTS+MTR estimates in Columns 5 and 6 are identical, each yielding bounds which rule out negative effects and any positive effects greater than the worst-case and MTS upper bounds. Finally, Column 7 simultaneously imposes MTS and MTR along with the total personal income MIV. The lower bound for female household members cannot exclude a null effect, but still only allows for a positive effect of up to 21.4 percentage points. The bounds for males are narrower from both sides, where it can be concluded that the voucher increases the likelihood of no hospitalizations by no less than 0.7 percentage points (0.8%) and no more than 13.0 percentage points (18.0%). However, the 95% CLR confidence intervals include zero, so I am unable to conclude that the ATE is positive as well as statistically significant for either sex.

Next, I consider the effect of housing voucher receipt on the probability of non-hospitalization separately for Black, white, and Hispanic household members. The three panels of Table 12 report the sample proportions of those reporting zero nights in the hospital for each subgroup. Most interesting here is that Black voucher holders have a slightly higher probability of having avoided hospitalizations over the previous year compared to Black members of households without a voucher. For white and Hispanic household members, the usual pattern of non-voucher holders observed to have better health continues to be present. This may be further suggestive evidence that a positive health effect is particularly pronounced among Black recipients.

Table 12: Observed Probabilities of Zero Hospitalizations and Distribution of Vouchers, by Racial/Ethnic Identity

t	$\mathbb{E}[Y \mid T = t]$	**Pr** $[T = t]$	N_t
Black			
Voucher	0.878	0.154	119
	(0.031)	(0.017)	-
No Voucher	0.843	0.846	638
	(0.015)	(0.017)	-
Total	0.848	1.000	757
	(0.014)	-	-

t	$\mathbb{E}[Y \mid T = t]$	**Pr** $[T = t]$	N_t
White			
Voucher	0.706	0.040	71
	(0.059)	(0.006)	-
No Voucher	0.854	0.960	1520
	(0.010)	(0.008)	-
Total	0.848	1.000	1591
	(0.010)	-	-

t	$\mathbb{E}[Y \mid T = t]$	**Pr** $[T = t]$	N_t
Hispanic			
Voucher	0.781	0.079	64
	(0.046)	(0.008)	-
No Voucher	0.872	0.921	770
	(0.012)	(0.008)	-
Total	0.865	1.000	834
	(0.012)	-	-

Estimates use the annual SIPP weights for individual-level analysis. Standard errors in parentheses are based on 240 BRR replications with the SIPP replicate weights.

The estimated bounds on the ATE of voucher receipt on the probability of no hospitalizations are reported in Table 13. Column 1 contains the results from regressing the no-hospitalizations indicator on the voucher receipt variable. As noted above, Black voucher recipients are 3.5 percentage points more likely to have not spent any nights in the hospital, although this is not statistically significant at conventional levels. White and Hispanic voucher holders, on the other hand, are respectively 14.8 and 9.0 percentage points less likely to have avoided hospitalization; both unadjusted gaps are statistically significant at the 5% level.

The worst-case bounds in Column 2 for white and Hispanic household members are generally similar to the bounds for the overall sample. The upper bounds rule out positive effects greater than 16.8 percentage points (31.2%) for the white subpopulation, or greater than 18.0 percentage points (30.0%) for Hispanic household members. For Black household members, these bounds are shifted somewhat to the right, such that the upper bound allows for an ATE of up to 26.9 percentage points (44.2%).

Column 3 imposes the MTS assumption, raising the lower bound for each subgroup. Here, MTS alone is sufficient to bound the ATE away from zero for Black household members, with the lower bound indicating at least a 3.5 percentage point (4.2%) increase in the likelihood of zero hospitalizations. The 95% confidence interval, however, includes a null effect, so I cannot conclude the ATE is statistically significant at this level. For the other two subgroups, I am able to rule out negative effects beyond a 14.9 percentage point (17.4%) decline in the probability of non-hospitalization for white household members, and a 9.1 percentage point (10.4%) decrease for Hispanic household members.

Interestingly, adding the total personal income MIV in Column 4 yields MIV+MTS bounds that

Table 13: Bounds on the ATE for Zero Hospitalizations, by Race: Total Personal Income as MIV, 5 Bins

	(1) OLS	(2) Worst-Case	(3) MTS	(4) MIV+MTS	(5) MTR	(6) MTS+MTR	(7) MIV+MTS+MTR
ATE for Black	0.035	[-0.732 , 0.269]	[0.035 , 0.269]	[0.056 , 0.262]	[0.000 , 0.269]	[0.000 , 0.269]	[0.054 , 0.262]
Residents	(0.033)	(-0.762 , 0.298)	(-0.022 , 0.300)	(-0.024 , 0.310)	(0.000 , 0.299)	(0.000 , 0.300)	(0.000 , 0.310)
ATE for White	-0.148^{**}	[-0.832 , 0.168]	[-0.149 , 0.168]	[0.001 , 0.157]	[0.000 , 0.168]	[0.000 , 0.168]	[0.099 , 0.157]
Residents	(0.059)	(-0.849 , 0.185)	(-0.254 , 0.187)	(-0.087 , 0.182)	(0.000 , 0.185)	(0.000 , 0.187)	(0.082 , 0.181)
ATE for Hispanic	-0.090^{**}	[-0.820 , 0.180]	[-0.091 , 0.180]	[-0.138 , 0.180]	[0.000 , 0.180]	[0.000 , 0.180]	[0.000 , 0.180]
Residents	(0.046)	(-0.843 , 0.203)	(-0.172 , 0.204)	(-0.244 , 0.209)	(0.000 , 0.203)	(0.000 , 0.204)	(0.000 , 0.208)

[·] denotes estimated bounds in Columns 2, 3, 5, and 6, or half-median unbiased MIV bounds in Columns 4 and 7. (·) denotes the standard error in Column 1, or 95% CLR confidence intervals on the ATE in Columns 2-7. Bounds are estimated using 757 observations for the Black subsample, 1591 observations for the white subsample, and 834 observations for the Hispanic subsample. The variance-covariance matrix in the CLR first step is based on 240 BRR replications, where the replicate weights provided by SIPP are used.

indicate a positive effect for both Black and white household members. For Black residents of voucher households, the probability of zero hospitalizations is increased by at least 5.6 percentage points (6.8%), even without invoking the assumption on treatment response. For white individuals, the lower bound is small in magnitude, but does rule out a zero effect. However, the confidence intervals suggest that the ATE is not statistically significant under the combination of only MIV and MTS. For Hispanic voucher holders, the MIV+MTS bounds are strictly wider than the MTS bounds, suggesting a lack of identifying power from total personal income for this group.

Columns 5 and 6 report the identified sets under MTR and MTS+MTR. These are equal to one another, each ruling out a negative impact on the probability of zero hospitalizations. Lastly, Column 7 uses total personal income as an MIV to obtain the MIV+MTS+MTR bounds. For Black household members, these are nearly equivalent to the MIV+MTS bounds. The lower bound for white household increases substantially, meaning that the probability of no hospitalizations increases by at least 9.9 percentage points (16.3%) and up to 15.7 percentage points (29.3%), and the 95% CLR confidence interval indicates that the ATE is statistically significant at the 5% level. While the bounds for Hispanic household members cannot rule out a zero effect, I can conclude that the effect is no larger than 17.9 percentage points (28.6%).

Overall, the bounds for the effect of voucher receipt on the probability of zero nights spent in the hospital are less conclusive for its sign than the bounds estimated for the self-reported overall health outcome. It is relevant to note that hospitalization is somewhat of an extreme outcome, and so it is reasonable to expect that housing voucher receipt would have a smaller positive impact. The small magnitude of the MIV+MTS+MTR lower bounds would be consistent with this, as such smaller effects are not ruled out.

6.3 The Average Treatment Effect on the Treated

The focus of this chapter has been on partially identifying the population average treatment effect of vouchers on health outcomes, but other treatment effect parameters may be of additional interest. Conceptually, the ATE answers the thought experiment of how, on average, the health of any resident i in an eligible household would be impacted if she were to be assigned a housing voucher. While this typically would be the information sought by a policymaker, the nature of HCV again complicates the matter. As discussed in the context of the MTS assumption, voucher supply is severely limited and eligible non-voucher holders greatly outnumber households which actually possess one. The ATE thought experiment would thus require a substantial increase in the supply of vouchers if it were to be carried out. Given the failures of Build Back Better and its moderate expansion to HCV, policymakers could then conceivably be less interested in the ATE.[24]

An alternative question to answer may therefore be: what is the health effect of receiving a voucher *for those who actually receive a voucher*? This is, of course, the average treatment effect on the treated (ATT), which can be written

$$ATT(1,0) = \mathbb{E}[Y_i(1) \mid V_i = 1] - \mathbb{E}[Y_i(0) \mid V_i = 1] \tag{16}$$

The first term is identified in the data by $\mathbb{E}[Y_i \mid V_i = 1]$, while the second remains a counterfactual. The same assumptions used in Section 5 can be used to bound this counterfactual object. Thus, they can also provide partial identification of the ATT.[25]

Bounded support for the outcome implies that

$$0 \leq \mathbb{E}[Y_i(0) \mid V_i = 1] \leq 1 \tag{17}$$

such that worst-case bounds on the ATT are identified as

$$\begin{aligned} &\mathbb{E}[Y_i \mid V_i = 1] - 1 \\ &\leq ATT(1,0) \leq \\ &\mathbb{E}[Y_i \mid V_i = 1] \end{aligned} \tag{18}$$

As was the case for the ATE, the worst-case bounds have a width equal to 1 and so cannot strictly identify the sign of the ATT, but may potentially be informative on its magnitude. Imposing the remaining assumptions can again tighten the identified set.

Under MTS, it can be concluded that

$$0 \leq \mathbb{E}[Y_i(0) \mid V_i = 1] \leq \mathbb{E}[Y_i(0) \mid V_i = 0] \tag{19}$$

with the rightmost term identified with $\mathbb{E}[Y_i \mid V_i = 0]$ in the data. MTS bounds on the ATT are then obtained as

$$\begin{aligned} &\mathbb{E}[Y_i \mid V_i = 1] - \mathbb{E}[Y_i \mid V_i = 0] \\ &\leq ATT(1,0) \leq \\ &\mathbb{E}[Y_i \mid V_i = 1] \end{aligned} \tag{20}$$

Relative to the worst-case bounds, the lower bound in (20) is increased to the extent that the mean outcome among non-voucher holders is less than 1, while the upper bound remains the same.[26]

[24]Note also that an expansion large enough to provide every eligible household with a voucher would be likely to generate general equilibrium effects in violation of the stable unit treatment value assumption (SUTVA) that is implicitly held for the identification.

[25]To the best of my knowledge, using the Manski and Pepper (2000) framework to bound an ATT is unique to the present study.

[26]One may also notice that the MTS lower bound on the ATT is equal to the MTS lower bound on the ATE from (9) and (10).

From the MTR assumption, I am able to write

$$0 \leq \mathbb{E}[Y_i(0) \mid V_i = 1] \leq \mathbb{E}[Y_i(1) \mid V_i = 1] \tag{21}$$

where the observed mean among voucher holders identifies the rightmost term. Using this, the bounds on the ATT under MTR become

$$\begin{aligned} &0 \\ &\leq ATT(1,0) \leq \\ &\mathbb{E}[Y_i \mid V_i = 1] \end{aligned} \tag{22}$$

This effectively enforces that the ATT is non-negative, while the upper bound is identical to both the worst-case and MTS upper bounds. Therefore, the combination MTS+MTR bounds also share the same upper bound, while the lower bound is equal to the larger of the difference $\mathbb{E}[Y_i \mid V_i = 1] - \mathbb{E}[Y_i \mid V_i = 0]$ and zero.

Finally, the MIV+MTS and MIV+MTS+MTR identified sets are obtained as in (12) and (13), with the ATT bounds of this subsection replacing the ATE bounds under MTS and MTS+MTR in those inequalities. Estimation and inference is carried out via the same procedure as described in Section 5.2 above.

Table 14: Bounds on the ATT for Self-Reported Health: Total Personal Income as MIV, 5 Bins

	(1) OLS	(2) Worst-Case	(3) MTS	(4) MIV+MTS	(5) MTR	(6) MTS+MTR	(7) MIV+MTS+MTR
$ATT(1,0)$	-0.209^{***}	[-0.429 , 0.571]	[-0.210 , 0.571]	[-0.092 , 0.529]	[0.000 , 0.571]	[0.000 , 0.571]	[0.101 , 0.529]
	(0.033)	(-0.483 , 0.625)	(-0.265 , 0.625)	(-0.190 , 0.597)	(0.000 , 0.626)	(0.000 , 0.626)	(0.000 , 0.598)
	·	⟨-0.471 , 0.613⟩	⟨-0.253 , 0.614⟩	⟨-0.169 , 0.583⟩	⟨0.000 , 0.614⟩	⟨0.000 , 0.625⟩	⟨0.018 , 0.585⟩

[·] denotes estimated bounds in Columns 2, 3, 5, and 6, or half-median unbiased MIV bounds in Columns 4 and 7. (·) denotes the standard error in Column 1, or 95% CLR confidence intervals on the ATT in Columns 2-7. ⟨·⟩ denotes 90% confidence intervals in Columns 2-7. Bounds are estimated using 3529 observations. The variance-covariance matrix in the CLR first step is based on 240 BRR replications, where the replicate weights provided by SIPP are used.

Table 14 presents the results for the ATT of housing voucher receipt on self-reported health in the full sample. Column 1 regresses the indicator for good or better self-reported health on the voucher dummy, which would identify the ATT (and the ATE) under exogenous treatment selection. This suggests a large 20.9 percentage decline in the probability of reporting good or better health. As before, my identification argument holds that this is a biased estimate of the ATT due to negative selection into voucher status on potential health outcomes.

Turning to the estimated bounds, Column 2 reports the worst-case bounds on the ATT. These indicate that the probability of good or better self-reported health is decreased by no more than 42.9 percentage points for voucher recipients, and increased by no more than 57.1 percentage points. As usual, no information about the sign of the effect can be learned from the worst-case bounds, and the fact that the ones here are nearly centered at zero means only extremely large effects in either direction can be ruled out.

MTS is imposed in Column 3, which increases the lower bound to a 21.0 percentage point decrease in the probability of good SRH. This signifies that any negative effect would be at worst equal to the observed

This is specific to the all-binary treatment and outcome setting and need not be true in the general case.

difference in means across voucher status, while the upper bound is equivalent to the worst-case upper bound. Adding the total personal income MIV in Column 4 further increases the lower bound, which estimates that a negative effect can be no greater than a 9.2 percentage point decrease in the probability of good health.

The MTR and MTS+MTR bounds in Columns 5 and 6 are identical, with both sets ruling out a negative ATT, as well as any effect larger than a 57.1 percentage point increase. Finally, Column 7 uses the total personal income MIV to estimate the MIV+MTS+MTR bounds. From the lower bound, the probability of reporting good or better health among recipient households is increased by at least 10.1 percentage points, which represents a 21.5% relative increase. From the upper bound, I am able to reject anything beyond a 52.9 percentage point increase (1260%). While I cannot statistically reject a null effect based on the 95% confidence interval, the 90% confidence interval (the bottom row) does indicate that the ATT is statistically different from zero at that level.

Next, I report results for self-reported health broken down by sex and by racial/ethnic identity in Table 15. Panels A and B show the estimates for the female and male subsamples, respectively. Column 1 contains the differences in means, which would indicate a reasonably large and statistically significant negative impact on self-reported health for both female and male voucher recipients if there were no issues of endogenous selection.

The worst-case bounds on the ATT in Column 2 for the female subsample are quite similar to the bounds for the full sample, where they are almost centered at zero and can only exclude greater than a 45.5 percentage point decrease or a 54.6 percentage point increase in the probability of good self-reported health. For males in voucher households, the worst-case bounds are shifted somewhat to the right, where any negative effect is limited to a 37.1 percentage point decrease in the likelihood of good health, and allowing for up to a 62.9 percentage point increase.

For the MTS bounds in Column 3, the lower bounds are increased such that the probability of reporting good or better health is estimated to decrease by no more than 22.5 percentage points for female voucher recipients, and by no more than 16.5 percentage points for male recipients. The upper bounds remain the same as under only the bounded outcome support assumption. The MTS and MTS+MTR bounds in Columns 5 and 6 are identical for both the female and male subsamples, which restrict the ATT to be non-negative and no greater than a 54.6 and 62.9 percentage point increase in the probability of good health, respectively.

Column 4 combines the MTS assumption with the total personal income MIV. This increases the lower bound for female voucher holders to a 10.2 percentage point decline in the probability of good or better SRH, and simultaneously decreases the upper bound to exclude increases larger than 47.8 percentage points. For male recipients, negative effects beyond a 10.9 percentage point decrease can be ruled out, as can positive effects larger than 62.0 percentage points. Finally, I report the MIV+MTS+MTR bounds in Column 7. For females in voucher households, these estimate that the probability of reporting good or better health is increased by at least 12.6 percentage points (30.1%) and at most 47.8 percentage points (813%), and the ATT can be distinguished from 0 at the 5% level. For the male subsample, the ATT is estimated to be at least a 4.4 percentage point (7.5%) increase in the probability of good or better SRH, and no greater than a 62.0 percentage point increase (6989%). Zero is included in both the 95% and 90% confidence intervals, meaning that the effect cannot be distinguished from 0 at conventional levels.

Table 15: Bounds on the ATT for Self-Reported Health, by Sex and by Racial/Ethnic Identity: Total Personal Income as MIV, 5 Bins

	(1) OLS	(2) Worst-Case	(3) MTS	(4) MIV+MTS	(5) MTR	(6) MTS+MTR	(7) MIV+MTS+MTR
Panel A: Female							
$ATT(1,0)$	-0.225^{***}	[-0.455 , 0.546]	[-0.225 , 0.546]	[-0.102 , 0.478]	[0.000 , 0.546]	[0.000 , 0.546]	[0.126 , 0.478]
	(0.039)	(-0.516 , 0.607)	(-0.289 , 0.607)	(-0.204 , 0.558)	(0.000 , 0.608)	(0.000 , 0.608)	(0.007 , 0.562)
	·	⟨-0.503 , 0.593⟩	⟨-0.275 , 0.594⟩	⟨-0.183 , 0.541⟩	⟨0.000 , 0.595⟩	⟨0.000 , 0.595⟩	⟨0.030 , 0.546⟩
Panel B: Male							
$ATT(1,0)$	-0.165^{***}	[-0.371 , 0.629]	[-0.165 , 0.629]	[-0.109 , 0.620]	[0.000 , 0.629]	[0.000 , 0.629]	[0.044 , 0.620]
	(0.059)	(-0.468 , 0.727)	(-0.264 , 0.728)	(-0.252 , 0.769)	(0.000 , 0.729)	(0.000 , 0.729)	(0.000 , 0.770)
	·	⟨-0.447 , 0.706⟩	⟨-0.244 , 0.707⟩	⟨-0.224 , 0.739⟩	⟨0.000 , 0.709⟩	⟨0.000 , 0.709⟩	⟨0.000 , 0.741⟩
Panel C: Black							
$ATT(1,0)$	-0.027	[-0.285 , 0.715]	[-0.027 , 0.715]	[0.102 , 0.641]	[0.000 , 0.715]	[0.000 , 0.715]	[0.125 , 0.641]
	(0.045)	(-0.357 , 0.787)	(-0.103 , 0.787)	(0.002 , 0.765)	(0.000 , 0.787)	(0.000 , 0.787)	(0.000 , 0.767)
	·	⟨-0.341 , 0.771⟩	⟨-0.086 , 0.772⟩	⟨0.022 , 0.740⟩	⟨0.000 , 0.772⟩	⟨0.000 , 0.772⟩	⟨0.007 , 0.742⟩
Panel D: White							
$ATT(1,0)$	-0.333^{***}	[-0.567 , 0.434]	[-0.333 , 0.434]	[-0.234 , 0.408]	[0.000 , 0.434]	[0.000 , 0.434]	[0.043 , 0.408]
	(0.060)	(-0.663 , 0.530)	(-0.434 , 0.532)	(-0.436 , 0.527)	(0.000 , 0.536)	(0.000 , 0.536)	(0.000 , 0.529)
	·	⟨-0.642 , 0.510⟩	⟨-0.413 , 0.512⟩	⟨-0.410 , 0.505⟩	⟨0.000 , 0.516⟩	⟨0.000 , 0.516⟩	⟨0.000 , 0.507⟩
Panel E: Hispanic							
$ATT(1,0)$	-0.424^{***}	[-0.616 , 0.385]	[-0.425 , 0.385]	[-0.418 , 0.423]	[0.000 , 0.385]	[0.000 , 0.385]	[0.000 , 0.423]
	(0.068)	(-0.724 , 0.493)	(-0.538 , 0.494)	(-0.556 , 0.532)	(0.000 , 0.500)	(0.000 , 0.500)	(0.000 , 0.534)
	·	⟨-0.701 , 0.470⟩	⟨-0.515 , 0.472⟩	⟨-0.530 , 0.509⟩	⟨0.000 , 0.479⟩	⟨0.000 , 0.479⟩	⟨0.000 , 0.512⟩

[·] denotes estimated bounds in Columns 2, 3, 5, and 6, or half-median unbiased MIV bounds in Columns 4 and 7. (·) denotes the standard error in Column 1, or 95% CLR confidence intervals on the ATT in Columns 2-7. ⟨·⟩ denotes 90% confidence intervals in Columns 2-7. Bounds are estimated using 3529 observations. The variance-covariance matrix in the CLR first step is based on 240 BRR replications, where the replicate weights provided by SIPP are used.

Panels C, D, and E report the estimates by racial and ethnic identity. Column 1 contains the naïve OLS regressions, which again would indicate negative effects under an exogenous selection assumption. The worst-case bounds for Black voucher holders in Panel C, Column 2 exclude negative impacts on the probability of good or better health larger than a 28.5 percentage point decrease. The bounds for white recipients in Panel D, on the other hand, are closer to centered at zero, while those for Hispanic voucher recipients in Panel E are shifted leftward, ruling out any large increase in the probability of good health above 38.5 percentage points.

Some other results are noteworthy. The MIV+MTS bounds for Black voucher recipients in Column 4 suggest that—even without the stronger MTR assumption—the probability of reporting good or better health is increased by at least 10.2 percentage points (16.6%), and that this ATT statistically differs from 0 at the 5% level. For the MIV+MTS+MTR bounds in Column 7, the lower bound for Black voucher holders increases to estimate that the probability of good or better health increases by at least 12.5 percentage points (21.2%), with the 90% confidence interval able to rule out a null effect. For white recipients, the lower bound indicates that the likelihood of good or better SRH is increased by no less than 4.3 percentage points (11.0%), although the confidence intervals do not allow me to rule out a zero ATT at conventional levels. Finally, there appears to again be a lack of identifying power from the total personal income MIV for the Hispanic subsample, to the extent that the upper bound on the ATT is estimated to be larger than the observed proportion of Hispanic voucher holders reporting good or better health in the SIPP.

Complete results on the ATT for the no-hospitalizations outcome can be found in Appendix B. Considering the full sample estimates in Appendix Table B.1, imposing the MTS assumption limits any negative effect to a 5.7 percentage point drop in the probability of avoiding a hospital stay in the previous year. Adding MTR further increases the lower bound to rule out any negative impact. The upper bounds, on the other hand, are all either equal to the sample proportion of those experiencing zero hospitalizations, or exceed this proportion. They are therefore not informative about the magnitude of the ATT, as the sample proportion is the largest possible effect even without any assumptions. Further, the total personal income MIV does not appear to add identifying power, as the MIV+MTS and MIV+MTS+MTR are wider than their counterparts without the MIV.

Though the upper bounds also are not especially informative for the demographic subgroups in Appendix Table B.2, some of the lower bounds using the total personal income MIV do have additional identifying power for certain groups. In Panels A, C, and D containing the estimates for female, Black, and white voucher holders, respectively, the lower bounds under MIV+MTS+MTR indicate a positive ATT. For female residents of voucher households, the probability of avoiding hospitalization is increased by at least 2.2 percentage points, or a 2.9% relative increase. Among Black recipients, the effect is at least a 4.6 percentage point (5.5%) increase, and the lower bound for white recipients indicates at least a 13.6 percentage point (23.9%) rise in the probability of avoiding a stay in the hospital. The latter result is statistically different from zero at the 5% level, while conventional confidence intervals cannot reject a null ATT for the female and Black subpopulations.

7 Conclusion

Credible estimates of a causal relationship between housing assistance and recipients' health are difficult to obtain, particularly with publicly-available survey data. I appeal to mild and relatively plausible assumptions in order to overcome the selection issues present in voucher receipt and the lack of exogenous variation induced by policy changes and identify bounds on the ATE. The use of monotone treatment selection (MTS), a monotone instrumental variable (MIV), and monotone treatment response (MTR) to achieve partial identification in this context provides novel evidence on an understudied aspect of the Housing Choice Voucher program. As a secondary contribution, I also demonstrate further the usefulness of the Manski and Pepper (2000) method in analyzing components of the US social safety net, which has elsewhere been applied fruitfully in studying the Supplemental Nutrition Assistance Program (SNAP) (Gundersen and Kreider, 2009; Kreider et al., 2012).

My first assumption of MTS captures the idea that voucher holders are negatively selected into treatment with respect to their health. Put another way, average potential health outcomes are weakly poorer among observed voucher recipients in any state of the world. I argue that this primarily follows from long waiting lists inducing Roy selection into HCV by those with characteristics correlated with "true need" for housing assistance.

The second assumption that total personal income is a valid MIV requires that potential health outcomes are on average non-decreasing in its values. Sufficient, but not necessary, for this is the normality of health as a good in the population of interest, a result which follows from the human capital model of the demand for health. The validity of income as an MIV is also consistent with a large literature documenting a positive relationship between income and health.

Finally, my MTR assumption states that each individual's health under voucher receipt would be no worse than her health in the place-based assistance or no-assistance state. The argument here proceeds by separately considering the possible income effect on health via a relaxed budget constraint and a housing choice effect through the ability of voucher holders to adjust their dual choice of housing services and neighborhood amenities. When relevant, the income effect is positive as long as health is a normal good. The housing choice effect is more complicated, but is arguably non-negative given the likely behavior of both landlords and voucher holders.

Under all three assumptions, the estimated MIV+MTS+MTR bounds suggest that the receipt of a housing voucher improves the likelihood of self-reporting good or better health by between 4.8 and 21.3 percentage points, where the effect is statistically significant at the 95% level. Using the baseline proportion of voucher holders in good health, this represents at least a 9.2% increase and up to a 59.5% increase. In terms of the probability of zero hospitalizations over the past year, I also estimate relative increases of at least a 0.1% and at most 29.7%, but I cannot statistically reject a null effect. Looking at subpopulations, the estimates for Black household members show a rightward-shifted identified set for the effect on self-reported health, with a lower bound of 11.7 percentage points (a 19.6% relative effect) and an upper bound of 26.9 percentage points (a 60.3% relative effect). While this means I can rule out a range of smaller effects identified for the overall population, the nature of partial identification does not allow me to draw any definitive conclusions from comparing across groups (*i.e.*, I cannot use the larger lower bound to say that Black voucher recipients benefit more than the population generally). On the other hand, I *can* conclude that the health of lower-income Black renters is meaningfully improved through the housing voucher program.

Overall, since these are bounds on a population ATE for adults in low-income renter households eligible for a voucher and who would counterfactually receive place-based rental assistance or no assistance at all, the results should be of policy interest for discussions of the voucher program. If the results are also relevant for the broader question of housing stability's health impact (*i.e.*, not just specific to the Housing Choice Voucher Program), they may also shed light on the Summer 2021 Supreme Court decision to disallow the Centers for Disease Control from extending the pandemic moratorium on evictions amidst the surging Delta variant of COVID-19, and then the eventual Omicron wave. From my bounds, this introduction of instability had the ability to cause at least a modest negative health consequence, with the potential for substantial declines in the probability of good health if the true ATE lies closer to the upper bound in my estimates. Further, the large lower bound I estimate for low-income Black renters means that the consequences for this subpopulation have the potential to be even greater and exacerbate the unequal burdens the pandemic has already placed across racial and ethnic groups.

References

Black, D., J. Smith, M. Berger, and B. Noel (2003). Is the threat of reemployment services more effective than the services themselves? Evidence from random assignment in the UI system. *The American Economic Review 93*(4), 1313–1327.

Bondy, M., S. Roth, and L. Sager (2018). Crime is in the air: The contemporaneous relationship between air pollution and crime. IZA DP No. 11492.

Case, A., D. Lubotsky, and C. Paxson (2002). Economic status and health in childhood: The origins of the gradient. *American Economic Review 92*(5), 1308–1334.

Chen, X., C. Flores, and A. Flores-Lagunes (2018). Bounds on average treatment effects with an invalid instrument, with an application to the Oregon Health Insurance Experiment. Working Paper.

Chernozhukov, V., S. Lee, and A. Rosen (2013). Intersection bounds: Estimation and inference. *Econometrica 81*, 667–737.

Chetty, R., M. Stepner, S. Abraham, S. Lin, B. Scuderi, N. Turner, A. Bergeron, and D. Cutler (2016). The association between income and life expectancy in the United States, 2001-2014. *JAMA 315*(16), 1750–1766.

Cropper, M. (1977). Health, investment in health, and occupational choice. *Journal of Political Economy 85*(6), 1273–1294.

Crossley, T. and S. Kennedy (2002). The reliability of self-assessed health status. *Journal of Health Economics 21*(4), 643–658.

de Haan, M. (2011). The effect of parents' schooling on child's schooling: A nonparametric bounds analysis. *Journal of Labor Economics 29*(4), 859–892.

Deaton, A. (2002). Policy implications of the gradient of health and wealth. *Health Affairs 21*(2), 13–30.

Ellen, I., K. Horn, and A. Schwartz (2016). Why don't housing choice voucher recipients live near better schools? Insights from big data. *Journal of Policy Analysis and Management 35*(4), 884–905.

Finkelstein, A., S. Taubman, B. Wright, M. Bernstein, J. Gruber, J. Newhouse, H. Allen, K. Baicker, and O. H. S. Group (2012). The Oregon Health Insurance Experiment: Evidence from the first year. *The Quarterly Journal of Economics 127*(3), 1057–1106.

Flores, C. and X. Chen (2018). *Average treatment effect bounds with an instrumental variable: Theory and practice*. Singapore: Springer.

Flores, C. and A. Flores-Lagunes (2010). Nonparametric partial identification of causal net and mechanism average treatment effects. Mimeographed: Department of Economics, University of Miami.

Flores, C. and A. Flores-Lagunes (2013). Partial identification of local average treatment effects with an invalid instrument. *Journal of Business & Economic Statistics 31*, 534–545.

Galiani, S., A. Murphy, and J. Pantano (2015). Estimating neighborhood choice models: Lessons from a housing assistance experiment. *The American Economic Review 105*(11), 3385–3415.

Germinario, G., V. Amin, C. Flores, and A. Flores-Lagunes (2021a). The use of polygenic scores as monotone instrumental variables. Mimeographed: Department of Economics, Syracuse University.

Germinario, G., V. Amin, C. Flores, and A. Flores-Lagunes (2021b). What can we learn about the effect of mental health on labor market outcomes under weak assumptions? Evidence from the NLSY79. Working paper.

Germinario, G., C. Flores, and A. Flores-Lagunes (2021). mpclr: A command for bias-corrected estimation and valid inference for nonparametric bounds on treatment effects. Mimeographed: Department of Economics, Syracuse University.

Geyer, J. (2017). Housing demand and neighborhood choice with housing vouchers. *Journal of Urban Economics 99*, 48–61.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy 80*(2), 223–255.

Grossman, M. (1999). The human capital model of the demand for health. NBER Working Paper 7078.

Gundersen, C. and B. Kreider (2009). Bounding the effects of food insecurity on children's health outcomes. *Journal of Health Economics 28*(5), 971–983.

Han, J. (2021). *Three essays on public housing and social inequality*. Ph. D. thesis, Syracuse University.

Han, J. and A. Schwartz (2021). Are public housing projects good for kids after all? EdWorkingPaper No. 21-437.

Hirano, K. and J. Porter (2012). Impossibility results for nondifferentiable functionals. *Econometrica 80*(4), 1769–1790.

Horn, K., I. Ellen, and A. Schwartz (2014). Do housing choice voucher holders live near good schools? *Journal of Housing Economics 23*, 28–40.

Hoynes, H. and D. Schanzenbach (2009). Consumption responses to in-kind transfers: Evidence from the introduction of the Food Stamp Program. *American Economic Journal: Applied Economics 1*(4), 109–139.

Idler, E. and Y. Benyamini (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior 38*(1), 21–37.

Imbens, G. and J. Angrist (1994). Identification and estimation of local average treatment effects. *Econometrica 62*(2), 467–475.

Imbens, G. and C. Manski (2004). Confidence intervals for partially identified parameters. *Econometrica 72*(6), 1845–1857.

Jacob, B., M. Kapustin, and J. 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.

Jürges, H. (2007). True health vs response styles: Exploring cross-country differences in self-reported health. *Health Economics 16*(2), 163–178.

Jylhä, M. (2009). What is self-rated health and why does it predict mortality? Towards a unified conceptual model. *Social Science & Medicine 69*(3), 307–316.

Katz, L., J. Kling, and J. Liebman (2001). Moving to opportunity in Boston: Early results of a randomized mobility experiment. *The Quarterly Journal of Economics 116*(2), 607–654.

Kreider, B., J. Pepper, C. Gundersen, and D. Jolliffe (2012). Identifying the effects of SNAP (food stamps) on child health outcomes when participation is endogenous and misreported. *Journal of the American Statistical Association 107*(499), 958–975.

Lens, M., I. Ellen, and K. O'Regan (2011). Do vouchers help low-income households live in safer neighborhoods? Evidence on the housing choice voucher program. *Cityscape 13*(3), 135–159.

Leung, C., S. Sarpça, and K. Yilmaz (2012). Public housing units vs. housing vouchers: Accessibility, local public goods, and welfare. *Journal of Housing Economics 21*(4), 310–321.

Lima-Costa, M., A. Steptoe, C. Cesar, C. De Oliveira, F. Proietti, and M. Marmot (2012). The influence of socioeconomic status on the predictive power of self-rated health for 6-year mortality in English and Brazilian older adults: The ELSA and Bambui cohort studies. *Annals of Epidemiology 22*(9), 644–648.

Ludwig, J., L. Sanbonmatsu, L. Gennetian, E. Adam, G. Duncan, L. Katz, R. Kessler, J. Kling, S. Tessler Lindau, R. Whitaker, and T. McDade (2011). Neighborhoods, obesity, and diabetes—a randomized social experiment. *The New England Journal of Medicine 365*, 1509–1519.

Manski, C. (1989). Anatomy of the selection problem. *Journal of Human Resources 24*(3), 343–360.

Manski, C. (1997). Monotone treatment response. *Econometrica 65*, 1311–1334.

Manski, C. and J. Pepper (2000). Monotone instrumental variables: With an application to the returns to schooling. *Econometrica 68*, 997–1010.

Massey, D. and S. Kanaiaupuni (1993). Public housing and the concentration of poverty. *Social Science Quarterly 74*(1), 109–122.

McCarthy, I., D. Millimet, and M. Roy (2015). Bounding treatment effects: A command for the partial identification of the average treatment effect with endogenous and misreported treatment assignment. *The Stata Journal 15*(2), 411–436.

Miilunpalo, S., I. Vuori, P. Oja, M. Pasanen, and H. Urponen (1997). Self-rated health status as a health measure: The predictive value of self-reported health status on the use of physician services and on mortality in the working age population. *Journal of Clinical Epidemiology 50*(5), 517–528.

Olsen, E. (2003). Housing programs for low-income households. In R. Moffitt (Ed.), *Means-tested transfer programs in the United States*, Chapter 6, pp. 365–442. University of Chicago Press.

O'Regan, K. and K. Horn (2013). What can we learn about the Low-Income Housing Tax Credit program by looking at the tenants? *Housing Policy Debate 23*(3), 597–613.

Rosen, E. (2020). *The voucher promise: "Section 8" and the fate of an American neighborhood*. Princeton: Princeton University Press.

Roy, A. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers 3*(2), 135–146.

Rubin, D. (2005). Causal inference using potential outcomes. *Journal of the American Statistical Association 100*(469), 322–331.

Schwartz, A., K. Horn, I. Ellen, and S. Cordes (2020). Do housing vouchers improve academic performance? Evidence from New York City. *Journal of Policy Analysis and Management 39*(1), 131–158.

Solari, C. and R. Mare (2012). Housing crowding effects on children's wellbeing. *Social Science Research 41*, 464–476.

Spence, A. (1973). Job market signaling. *Quarterly Journal of Economics 87*(3), 355–374.

Thakral, N. (2016). The public housing allocation problem: Theory and evidence from Pittsburgh. Working paper.

Wilson, K., S. Elliott, J. Eyles, and S. Keller-Olaman (2007). Factors affecting change over time in self-reported health. *Canadian Journal of Public Health 98*(2), 154–158.

World Health Organization (2018). *WHO housing and health guidelines*. Geneva: World Health Organization.

Zhang, N. (2021). Housing vouchers, labor supply, and household formation: A structural approach. Job market paper.

A Technical Appendix for CLR (2013)

To illustrate the method, consider, for example, the MIV+MTS+MTR lower bound on the mean potential outcome $\mathbb{E}[Y_i(1)]$:

$$\sum_{\ell=1}^{L} \left(\max_{\underline{\ell} \leq \ell} LB^{\underline{\ell}}_{\text{mtsr},1} \right) \cdot \mathbf{Pr}[Z_i \in \mathcal{B}_\ell]$$

where the supremum is replaced with a maximum.

In the terminology of Chernozhukov, Lee, and Rosen (2013), I create a set of bounding functions out of the above expression. These allow me to use the properties of the maximum (or minimum) operator to rewrite the single summation as a set of multiple summations over which a single maximum (or minimum) is taken. The intuition is that the bounding functions capture all possibilities from taking the maxima at each piece of the original form of the lower bounds. Choosing, for example, $L = 3$ bins for the MIV would yield the following 4 $\left(= 2^{3-1}\right)$ bounding functions:

$$\max \Big\{ LB^1_{\text{mtsr},1} \,,\ (p_1 + p_2) LB^1_{\text{mtsr},1} + p_3 LB^3_{\text{mtsr},1} \,,\ p_1 LB^1_{\text{mtsr},1} + p_2 LB^2_{\text{mtsr},1} + p_3 LB^2_{\text{mtsr},1} \,,$$
$$p_1 LB^1_{\text{mtsr},1} + p_2 LB^2_{\text{mtsr},1} + p_3 LB^3_{\text{mtsr},1} \Big\} \,,$$

where $p_\ell \equiv \mathbf{Pr}[Z_i \in \mathcal{B}_\ell]$.

A similar set of bounding functions for the MIV+MTS+MTR upper bound on $\mathbb{E}[Y_i(0)]$ can be obtained, where properties of the minimum operator are used instead. Then the full set of bounding functions for the lower bound on the ATE is obtained by exhausting the number of possible subtractions of one term from the set of upper bounds on $\mathbb{E}[Y_i(0)]$ from one term of the set of lower bounds on $\mathbb{E}[Y_i(1)]$. This yields a total of $\left(2^{5-1}\right)^2 = 256$ bounding functions for each bound on the average treatment effect in my application where I use 5 bins of the MIV. For convenience, I borrow the Chernozhukov et al. (2013) notation and use $\theta^l(v)$ and $\theta^u(v)$, $v = 1, \ldots, 256$, to denote these bounding functions for the lower and upper ATE bounds, respectively.

The key aspect of the CLR method is that the procedures for obtaining the half-median unbiased estimates of the bounds and for the valid confidence intervals are performed on each bounding function $\theta^l(v)$ and $\theta^u(v)$ prior to the evaluation of the associated maximum or minimum. The technical condition required to do so is that there exist estimators of the $\theta^l(v)$ and $\theta^u(v)$ which are consistent and asymptotically normal. Given that the bounding functions in this paper are composed of sample means and sample proportions, this condition is satisfied. Therefore, I proceed with a *precision adjustment*, which gives bias-corrected estimated bounds and valid confidence intervals for the ATE.

The CLR precision-adjustment step first entails taking the product of a critical value denoted $\kappa(p)$ and the pointwise standard error of the bounding function estimator written $s(v)$. For upper bound bounding functions this product is added to the estimator $\hat{\theta}^u(v)$, and it is subtracted from $\hat{\theta}^l(v)$ for lower bound bounding functions. The choice of p in the critical value governs whether the adjustment yields the median-unbiased bounding function (which sets $p = 0.5$) or yields the desired bound for the confidence interval (see below). In this way, the CLR method has the advantage that the bias correction and inference is done within the same procedure.

The precision-corrected estimators for the respective lower and upper ATE bounds are given by

$$\hat{\theta}^l(p) = \max_v \left\{ \hat{\theta}^l(v) - \kappa^l(p) \cdot s^l(v) \right\}$$

and

$$\hat{\theta}^u(p) = \min_v \left\{ \hat{\theta}^u(v) + \kappa^u(p) \cdot s^u(v) \right\}$$

where $\hat{\theta}^l(v)$ and $\hat{\theta}^u(v)$ are the unadjusted estimators of the bounding functions and $s^l(v)$ and $s^u(v)$ are their associated standard errors, which are defined below.

The critical values $\kappa^l(p)$ and $\kappa^u(p)$ are computed according to the following. Let $\hat{\gamma}^l$ be a 256-dimensional column vector of all the unadjusted bounding function estimators for the lower bound, with $\hat{\gamma}^u$ defined likewise for the upper bound. The first step obtains, using 240 BRR replications,[27] a consistent estimate $\widehat{\mathbf{\Omega}}^l$ of the asymptotic variance-covariance matrix of $\sqrt{N}\left(\hat{\gamma}^l - \gamma^l\right)$.[28] With $\widehat{\boldsymbol{g}}^l(\boldsymbol{v})'$ the v^{th} row of $\hat{\mathbf{\Omega}}^{1/2,l}$, I can thus define $s^l(v) \equiv \frac{\|\widehat{\boldsymbol{g}}^l(\boldsymbol{v})\|}{\sqrt{N}}$.

Next, I simulate $R = 100{,}000$ draws from a $\mathcal{N}(\mathbf{0}, \boldsymbol{I})$ distribution, where $\boldsymbol{I}$ is the 256×256 identity matrix. The draws are labelled $\boldsymbol{Z}_r$, $r = 1, \ldots, R$, and are used to compute $Z_r^*(v) \equiv \widehat{\boldsymbol{g}}^l(\boldsymbol{v})'\boldsymbol{Z}_r / \|\widehat{\boldsymbol{g}}^l(\boldsymbol{v})\|$ for each r and v. In each replication, the maximum over the set of $Z_r^*(1), \ldots, Z_r^*(4^{L-1})$ is selected. From the resulting R values, I compute $\kappa^l(c)$, defined as the c^{th} quantile of the values, where $c \equiv 1 - (0.1/\log N)$. This value $\kappa^l(c)$ is used to construct the following set:

$$\widehat{V}^l = \left\{ v \in \mathcal{V}^l : \hat{\theta}^l(v) \geq \max_{\tilde{v} \in \mathcal{V}^l} \left[\hat{\theta}^l(\tilde{v}) - \kappa^l(c) \cdot s^l(\tilde{v}) \right] - 2\kappa^l(c) \cdot s^l(v) \right\}$$

where $\mathcal{V}^l$ is the indexing set for the lower bound bounding functions $\theta^l(v)$. Returning to the values $Z_r^*(v)$, I now take the maximum from each replication r, this time restricting the search only to $v \in \widehat{V}^l$. The CLR critical value $\kappa^l(p)$ is then the p^{th} quantile of the resulting R values, such that $\kappa^l(0.5)$ gives the half-median unbiased estimate of the lower bound $\hat{\theta}^l(0.5)$ on the ATE.

Obtaining the lower bound on a $(1 - \alpha) \cdot 100\%$ confidence interval requires one final adjustment to account for the width of the bias-corrected bounds. Borrowing notation from Chernozhukov et al. (2013),

[27]This is again a limitation of the SIPP data. Replicate weights are provided for obtaining correct variance estimates, but only for up to 240 replications. While the CLR procedure performs better with more replications, the simulation results in Germinario, Flores, and Flores-Lagunes (2021) show reasonable performance using 199 replications. Thus, the limited number of replications here should be adequate.

[28]I illustrate here only the estimate of the lower bound. The process for the upper bound is analogous.

define

$$
\begin{aligned}
\widehat{\Gamma} &\equiv \hat{\theta}^u(0.5) - \hat{\theta}^l(0.5) \\
\widehat{\Gamma}^+ &\equiv \max\left\{0, \widehat{\Gamma}\right\} \\
\rho &= \max\left\{\hat{\theta}^u(0.75) - \hat{\theta}^u(0.25)\ ,\ \hat{\theta}^l(0.25) - \hat{\theta}^l(0.75)\right\} \\
\tau &\equiv 1/(\rho \log N) \\
\hat{p} &\equiv 1 - \Phi\left(\tau\widehat{\Gamma}^+\right) \cdot \alpha
\end{aligned}
$$

where $\Phi(\cdot)$ is the standard normal CDF. The lower bound of a 95% confidence interval is based on $\hat{\theta}^l(\hat{p})$, which uses the critical value $\kappa^l(\hat{p})$, with $\alpha = 0.05$ in the expression for $\hat{p}$.

B Additional Tables

Table B.1: Bounds on the ATT for Zero Hospitalizations: Total Personal Income as MIV, 5 Bins

	(1) OLS	(2) Worst-Case	(3) MTS	(4) MIV+MTS	(5) MTR	(6) MTS+MTR	(7) MIV+MTS+MTR
$ATT(1,0)$	$-0.057^{\star\star}$	[-0.196 , 0.804]	[-0.057 , 0.804]	[-0.074 , 0.825]	[0.000 , 0.804]	[0.000 , 0.804]	[0.000 , 0.825]
	(0.025)	(-0.237 , 0.845)	(-0.098 , 0.845)	(-0.119 , 0.868)	(0.000 , 0.845)	(0.000 , 0.845)	(0.000 , 0.868)
	·	⟨-0.228 , 0.836⟩	⟨-0.089 , 0.836⟩	⟨-0.109 , 0.858⟩	⟨0.000 , 0.836⟩	⟨0.000 , 0.836⟩	⟨0.000 , 0.858⟩

[·] denotes estimated bounds in Columns 2, 3, 5, and 6, or half-median unbiased MIV bounds in Columns 4 and 7. (·) denotes the standard error in Column 1, or 95% CLR confidence intervals on the ATT in Columns 2-7. ⟨·⟩ denotes 90% confidence intervals in Columns 2-7. Bounds are estimated using 3529 observations. The variance-covariance matrix in the CLR first step is based on 240 BRR replications, where the replicate weights provided by SIPP are used.

Table B.2: Bounds on the ATT for Zero Hospitalizations, by Sex and by Racial/Ethnic Identity: Total Personal Income as MIV, 5 Bins

	(1) OLS	(2) Worst-Case	(3) MTS	(4) MIV+MTS	(5) MTR	(6) MTS+MTR	(7) MIV+MTS+MTR
Panel A: Female							
$ATT(1,0)$	-0.051	[-0.217 , 0.783]	[-0.051 , 0.783]	[-0.037 , 0.778]	[0.000 , 0.783]	[0.000 , 0.783]	[0.022 , 0.778]
	(0.032)	(-0.269 , 0.835)	(-0.104 , 0.835)	(-0.090 , 0.837)	(0.000 , 0.835)	(0.000 , 0.835)	(0.000 , 0.837)
	·	⟨-0.258 , 0.824⟩	⟨-0.092 , 0.824⟩	⟨-0.078 , 0.824⟩	⟨0.000 , 0.824⟩	⟨0.000 , 0.824⟩	⟨0.000 , 0.824⟩
Panel B: Male							
$ATT(1,0)$	-0.047	[-0.149 , 0.851]	[-0.047 , 0.851]	[-0.057 , 0.882]	[0.000 , 0.851]	[0.000 , 0.851]	[0.000 , 0.882]
	(0.041)	(-0.217 , 0.918)	(-0.114 , 0.918)	(-0.132 , 0.955)	(0.000 , 0.918)	(0.000 , 0.918)	(0.000 , 0.955)
	·	⟨-0.202 , 0.903⟩	⟨-0.100 , 0.904⟩	⟨-0.116 , 0.941⟩	⟨0.000 , 0.904⟩	⟨0.000 , 0.904⟩	⟨0.000 , 0.942⟩
Panel C: Black							
$ATT(1,0)$	0.035	[-0.122 , 0.878]	[0.035 , 0.878]	[0.056 , 0.886]	[0.000 , 0.878]	[0.000 , 0.878]	[0.046 , 0.886]
	(0.033)	(-0.173 , 0.929)	(-0.019 , 0.929)	(-0.018 , 0.958)	(0.000 , 0.929)	(0.000 , 0.929)	(0.000 , 0.958)
	·	⟨-0.162 , 0.918⟩	⟨-0.007 , 0.918⟩	⟨-0.001 , 0.947⟩	⟨0.000 , 0.918⟩	⟨0.000 , 0.918⟩	⟨0.000 , 0.947⟩
Panel D: White							
$ATT(1,0)$	$-0.148^{\star\star}$	[-0.295 , 0.706]	[-0.149 , 0.706]	[0.001 , 0.729]	[0.000 , 0.706]	[0.000 , 0.706]	[0.136 , 0.729]
	(0.059)	(-0.392 , 0.803)	(-0.247 , 0.804)	(-0.079 , 0.836)	(0.000 , 0.805)	(0.000 , 0.805)	(0.074 , 0.837)
	·	⟨-0.371 , 0.782⟩	⟨-0.226 , 0.783⟩	⟨-0.063 , 0.813⟩	⟨0.000 , 0.785⟩	⟨0.000 , 0.785⟩	⟨0.085 , 0.815⟩
Panel E: Hispanic							
$ATT(1,0)$	$-0.090^{\star\star}$	[-0.219 , 0.782]	[-0.091 , 0.782]	[-0.138 , 0.825]	[0.000 , 0.782]	[0.000 , 0.782]	[0.000 , 0.825]
	(0.046)	(-0.294 , 0.856)	(-0.166 , 0.856)	(-0.236 , 0.913)	(0.000 , 0.857)	(0.000 , 0.857)	(0.000 , 0.913)
	·	⟨-0.277 , 0.840⟩	⟨-0.150 , 0.840⟩	⟨-0.215 , 0.894⟩	⟨0.000 , 0.840⟩	⟨0.000 , 0.840⟩	⟨0.000 , 0.894⟩

[·] denotes estimated bounds in Columns 2, 3, 5, and 6, or half-median unbiased MIV bounds in Columns 4 and 7. (·) denotes the standard error in Column 1, or 95% CLR confidence intervals on the ATT in Columns 2-7. ⟨·⟩ denotes 90% confidence intervals in Columns 2-7. Bounds are estimated using 3529 observations. The variance-covariance matrix in the CLR first step is based on 240 BRR replications, where the replicate weights provided by SIPP are used.