Do Evictions Cause Income Changes? An Instrumental Variables Approach

Grace Mok

Professor Christopher Timmins, Faculty Advisor

Professor Michelle Connolly, Faculty Advisor

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University.

Duke University

Durham, North Carolina

2019

Table of Contents

Acknowledgements

They say it takes a village to raise a thesis:

To Professor Christopher Timmins and Professor Michelle Connolly who have seen me through this labor: thank you for the faith and kindness you generously poured into me and this project.

To John Killeen of DataWorks and all of the staff, advocates and members at the Community Empowerment Fund: thank you for lending your experiences and wisdom in the conception of this project.

To my family of friends— Albert Hsiung, Claire Wang, Jay Zussman, David Wohlever Sanchez, Ben Welton, Amulya Vadapalli, Mitchell Ochse, Khuong Do and the rest of the thesis class gang—thank you for having the patience to believe in and support me as I bumbled and bloomed into adulthood this year.

Though my thesis is still a pretty ugly baby, I'm proud. Maybe she can grow up to be great and helpful one day.

Abstract

Evictions are an important aspect of the affordable housing crisis facing low-income American renters. However, there has been little research quantifying the causal impact of evictions, which poses challenges for academics interested in understanding inequality and policy-makers interested in reducing it. Merging two datasets both new to the literature, I address this gap in the causal literature by using an instrumental variables strategy to examine the impact of evictions on household income over time in Durham, North Carolina. Exploiting gentrification-related evictions as an instrument, I find a 2.5% decrease in household income after eviction. This is a small, but significant decrease in income given that median household income for households at time of eviction is about $15,000.

JEL Codes: I32, R29

Keywords: urban economics, eviction, housing, gentrification, North Carolina

Introduction

Low-income American renters face an affordable housing crisis. Though an important aspect of this crisis, the causes and effects of evictions have been little studied. In this paper, I explore how evictions may cause changes in household income using data from Durham, North Carolina.

Over the past half-century, the cost-burdened share of renters—those who spend more than the recommended 30 percent on housing—has doubled from 23.8 percent in the 1960s to 47.5 percent in 2016 (Joint Center for Housing Studies of Harvard University, 2018). Driving this trend are three other patterns. First, increases in median renter income has severely lagged behind skyrocketing median rent payments. Adjusting for inflation between 1960 and 2016, the median rent payment grew 61 percent, whereas the median renter income grew only 5 percent (Joint Center for Housing Studies of Harvard University, 2018). Second, federal rental assistance has failed to keep up with the growth in low-income renters. Between 1987 and 2015, the share of very low-income renters assisted by the federal government decreased from 29 percent to 25 percent (Joint Center for Housing Studies of Harvard University, 2018). Finally, availability of low-cost rentals has not kept pace with the number of low-income renters (Joint Center for Housing Studies of Harvard University, 2018). According to a 2018 study, only 35 rental units are affordable and available for every 100 extremely low-income renters, defined as those who earn no more than 30% of area median income (National Low Income Housing Coalition, 2018).

Another important component of the affordable housing crisis is evictions—involuntary, landlord-initiated displacements of renters. Using court data between 2000 and 2016, the Eviction Lab estimates that there is an average of almost one million evictions annually, with annual eviction rates ranging from between 2.34 and 3.11. For perspective, consider that in 2010—after the height of the Great Recession—there were slightly over one million foreclosures in total (Eviction Lab). Lundberg & Donnelly (2016) estimate that 1 in 7 urban American children experience an eviction for nonpayment of rent or mortgage by age 15.[1]

Not only are evictions widespread, existing research also indicates that evictions may play a role in health, income, neighborhood and employment disparities. Mothers who have experienced an eviction within the previous year are more likely to experience depression and material hardship afterwards (Desmond & Kimbro, 2015). After eviction, families are also more likely to move to neighborhoods with greater poverty and crime (Desmond & Schollenberger, 2015). Furthermore, evictions can manifest on tenants' credit reports, which impacts their ability to pursue future housing and future employment opportunities (Desmond et. al., 2013). Nevertheless, no confident conclusions can be drawn from the current literature due to an emphasis on correlational results, reliance on self-report data and small datasets. These limitations make it difficult for researchers to understand relationships between forced displacement and other social outcomes and for policymakers to accurately assess and address the affordable housing crisis.

In this paper, I extend this nascent literature on the impact of evictions by using an instrumental variables approach to estimate the causal impact of eviction on household income in Durham, North Carolina. First, I estimate the "exogeneity" of an eviction by relating it to the change in eviction rate in the block group. An "exogenous" eviction occurs outside of the control

---

[1] These figures are likely *under*estimates, since they are based only on court record. Hence, they exclude informal evictions—coercive displacements initiated by landlords outside of the court system such as removing a door or paying tenants to move out earlier. Though illegal in most jurisdictions, such informal evictions are estimated to make up almost three-fourths of all evictions (Desmond & Kimbro, 2015). Similarly, it is likely that *more* than 1 in 7 American urban children experience eviction, because Lundberg & Donnelly (2016)'s estimate is based on families' self reports of whether or not they had been evicted—which are likely downward biased for two reasons. First, some tenants who have indeed experienced eviction may not label their experience as "eviction." According to Desmond (2012), some tenants whose names were on court records of their eviction still did not name the experience as an "eviction" when spoken with in person. Second, there is little incentive to report having experienced an eviction when one has not. Eviction histories make it difficult to rent again and being evicted due to nonpayment of rent may evoke feelings of shame or trauma.

of the renter. For instance, one might be evicted because their apartment complex changed owners or due to a sudden economic downturn in the area. If "exogenous" evictions occur, they are likely to occur as a shock to the area. Then, I regress "exogeneous" evictions onto households' estimated incomes over time.

The first way I contribute to the eviction literature is by using an empirical strategy that allows me to identify the causal impact of evictions on income change. Rather than relying on self-reported indices for material hardship that can only proxy for income (e.g. "Did you borrow money from friends or family to help pay bills?"), I use third-party estimates of household income. This data source allows me to have more precise numerical estimates of how income is impacted. Moreover, because of an instrumental variables approach, my estimates will be *causal* and not merely correlational, which will lend to better understanding of the role that evictions play in urban inequality.

The second way I contribute is by introducing two novel datasets to the literature: household information purchased from the marketing company InfoUSA and community-level data from Durham, North Carolina on evictions. By merging these two datasets, I address a key limitation in the eviction literature—data. If combining national, household-level data with local, high-resolution information on evictions is successful, future researchers can leverage partnerships with local non-profits and governments to better understand evictions.

The following is a roadmap for the remainder of my paper. In Section 2, I review the existing literature. In Section 3, I make the theoretical case for my approach by describing the housing market and eviction processes in Durham, North Carolina. In Section 4, I review my empirical specification. In Section 5, I describe my data sources, including their strengths and weaknesses. In Section 6, I discuss my results and place them into the context of other literature on eviction. In Section 7, I conclude.

Literature Review

In this section, I provide an overview of the literature's two major limitations: reliance on self-report data and correlational empirical strategies like multivariate regression and propensity score matching. These two limitations likely bias the results of the five main papers of the eviction literature, summarized in Table 1.

Table 1: Overview of Eviction Literature

Study	Data Source	Empirical Strategy	Outcome Variables
Desmond, Gershenson & Kiviat (2015)	Milwaukee Area Renters Study (MARS)	Multivariate regression	Tenant mobility Housing quality
Desmond & Shollenberger (2015)	Milwaukee Area Renters Study (MARS)	Multivariate regression	Neighborhood poverty Neighborhood crime
Desmond & Kimbro (2015)	Fragile Families and Child Wellbeing Survey (FFCWS)	Propensity score matching	Material hardship Maternal depression Maternal health Child health Parenting stress
Desmond & Gershenson (2016)	Milwaukee Area Renters Study (MARS)	Propensity score matching	Employment security
Humphries, Mader, Tannenbaum & van Dijk (2018)	Data on eviction court, credit bureau and payday loans	Quasi-experimental	Credit score Total bal. collections Open auto loan or lease Any mortgage Any Payday Inquiry x 100 Any Payday Account x 100 Cumulative Zipcode moves Poverty rate (x 100) Any Eviction Case Eviction at Dif. Address
Collinson & Reed (2018)	Data on public assistance (food stamps, Medicaid, and cash	Quasi-experimental	Homelessness Health outcomes Earnings

	assistance)		Receipt of public assistance

The two most common sources of data thus far in the literature are self-report surveys: the Milwaukee Area Renters Study (MARS) and the Fragile Families and Child Wellbeing Survey (FFCWS). Both survey limited, though informative populations. MARS is a one-time cross-sectional survey of 1,086 renting households, selected through multistage stratified probability sampling (Desmond, Gershenson & Kiviat, 2015) while FFCWS is a panel study that follows a birth cohort of new unmarried parents and their children (Reichman et. al., 2001; Desmond & Kimbro, 2015). Due to stigma, MARS respondents may have minimized any troubles they may have faced regarding housing and employment. Their assessments may also not also accurately depict neighborhood quality, due to vagueness in question wording. Similarly, FFCWS responses regarding sensitive issues such as material hardship, maternal depression, health and parental stress may not be reliable or accurate assessments of outcome variables either. It is unclear what direction this self-report data would bias results, because the stigma in self-report data would bias results more conservatively, but the specific at-risk populations surveyed may bias results towards significance.

Correlational empirical strategies dominate the eviction literature, as overviewed by Humphries, Mader, Tannenbaum and van Dijk (2018) (Desmond, Gershenson & Kiviat, 2015; Desmond & Shollenberger, 2015; Desmond & Kimbro, 2015; Desmond & Gershenson, 2016). Using either multivariate regression or propensity score matching, these studies estimate the effect of eviction—generally adverse—on a variety of outcome variables after controlling for variables observable in the data. However, these correlational results cannot easily be assigned a straightforward causal interpretation, because they are inherently subject to omitted variable bias.

One example of a variable that the majority of the literature fails to control for is mere contact with eviction court—even if that individual is not later evicted. Humphries et. al., (2018) find that the impact of eviction on financial strain is significantly minimized after controlling for contact with eviction court. This result suggests that correlational studies that do not (or lack the data to) control for contact with eviction court likely overestimate the negative impacts of eviction on outcome variables. Being summoned to eviction court is just one of many variables that, if omitted, could bias the results of a correlational approach.[2]

To my knowledge, the only existing papers that avoid the eviction literature's main limitations of self-reported data and correlational empirical strategy are Humphries et. al., (2018) and Collinson & Reed (2018). First, both papers use official eviction court data linked to observed, not self-reported, data from agencies such as credit bureaus and public assistance agencies. Thus, the papers avoid the response bias to which self-report surveys are subject. Second, the papers do not opt for a correlational empirical strategy. Instead, they exploit the random assignment of cases to judges in eviction court to identify the causal impact of evictions. In contrast to previous literature, both papers find that the causal effects of eviction are small. Humphries et. al., (2018) find the impact of eviction small relative to the financial troubles tenants face shortly before entering eviction court. Collinson and Reed (2018) find that evictions lower earnings slightly, but do not substantially worsen employment outcomes or increase receipt of public assistance.

In the same vein as these recent papers, my paper uses official eviction court data linked to

[2] In Appendix C, I present the OLS benchmark results that also control for whether a household had experienced summons before. I originally intended to include this indicator variable for summons in my main regressions. I wanted to avoid omitted variable bias, reasoning that summoned households might have lower incomes regardless of whether they were actually evicted. By controlling for summons, I thought I would be able to isolate the impact of eviction from the impact of mere contact with eviction court system. However, when I included summons as a control, the coefficients were positive and significant. Because I could not identify any theoretical or qualitative reason why summoned households would have higher incomes than their non-summoned counterparts, I opted to omit summons as a control variable in my main regressions and only include those results in Appendix C for intellectual honesty.

observed—not self-reported—data on income and uses a non-correlational empirical strategy to estimate the causal impact of eviction on household income in Durham, North Carolina.

Theoretical Framework

In this section, I make the theoretical case for both stages of my instrumental variables analysis. First, I draw upon sociological literature to provide possible explanations for why evictions may cause household income to decrease. Second, I make the distinction between evictions that are *endogenous*—related to some characteristic of the individual, such as previous incarceration and mental health problems—and those that are *exogenous*—related to situations outside of the control of the individual, such as judge/courtroom leniency and gentrification. When evaluating the causal impact of evictions, it is critical to distinguish between these two types of evictions—necessitating an empirical strategy such as instrumental variables.

The Impact of Evictions on Income

There are several reasons why evictions may have negative impacts on income. First, because eviction court records are publicly available, tenants' credit scores may decrease, which could impact their employment opportunities. Humphries et. al., (2018) demonstrate small, though statistically significant impacts of eviction on credit scores. Second, simply moving homes is a stressful event, which could impact performance at work. Third, there is some evidence that evicted tenants move to poorer and higher-crime neighborhoods, which could also be farther from high-paying and stable employment opportunities (Desmond & Shollenberger, 2015).

Endogenous Reasons for Eviction

Endogenous reasons for eviction are those related to some characteristic of the individual or circumstance that could also explain changes in income. These endogenous reasons for eviction pose challenges for the identification of the causal impact of eviction on income, because changes in income may be misattributed to eviction, as opposed to some other factor that may itself have been a cause of eviction. Some simple examples of endogenous reasons are job loss and health problems in the absence of health insurance, which could simultaneously cause the eviction and the change in income.

Previous literature suggests two other examples of endogenous reasons for eviction: paternal incarceration and maternal depression. Using a logit model, Geller & Franklin (2014) found that mothers with recently incarcerated male partners were 50% more likely to face housing insecurity than their counterparts without incarcerated partners, controlling for socioeconomic status, race and other characteristics. I classify paternal incarceration as an *endogenous* reason for eviction because it is possible that families that are more likely to have paternal incarceration are also more likely to become evicted. In other words, paternal incarceration is not necessarily an event that occurs independently of the family. Similarly, using a logit model, Curtis et. al. (2014) found that mothers who experienced postpartum depression were more likely to experience or be at risk of homelessness, which included being evicted. Maternal depression is an *endogenous* reason for eviction because it could also cause mothers to be less capable of working longer hours or more demanding jobs, independently of being evicted.

Exogenous Reasons for Eviction

In contrast, exogenous reasons for eviction are those related to situations outside of the control of the individual. Being able to isolate these exogenous reasons for eviction is essential in identifying the causal impact of eviction on income.

Two examples of exogenous reasons for eviction are landlord discretion and judge/courtroom stringency. According to ethnographic work conducted in Milwaukee, landlords have considerable control over deciding whether or not to file for eviction (Desmond, 2016). Depending on their relationship with the tenant or other demographic factors, landlords may choose to file for eviction or give extra chances to tenants (Desmond, 2016). Similarly unrelated to the individual is judge stringency—their tendency to grant eviction orders. To explore the impact of eviction on economic status, Humphries et. al. (2018) and Collinson and Reed (2018) exploit this judge stringency in recent working papers using the random assignment of judges.

The exogenous reason for eviction that my paper will exploit is gentrification, defined as "the buying and renovation of houses and stores in deteriorated urban neighborhoods by upper- or middle-income families or individuals, raising property values but often displacing low-income families and small businesses" (Dictionary). The Herald Sun, a newspaper local to the Durham-Raleigh region, reported that between 2010 and 2018, six out of Durham County's sixty census tracts had increases in median household income levels of more than 40 percent (Vaughan & Eanes, 2018). During the same eight-year span, the county's median household income levels grew only 8 percent, indicating rapid influxes of wealthier residents into certain parts of Durham (Vaughan & Eanes, 2018).

As more and more affluent individuals move into an area, landlords may be incentivized to raise rents. Landlords are able to shorten the lengths of leases, which would allow them to increase the rents more frequently and give a shorter notice for tenants to leave (North Carolina, 2016; The City of Durham Department of Neighborhood Improvement Services, 2012). Based on local news reporting, although current tenants are unable to afford the increased rents, they are also unable to find a new place to stay and move out under the short notice (Vaughan, 2018). As a result, these tenants receive eviction notices due to exogenous increases in rent requirements. Essentially, gentrification could cause many people living in a given geographic area to be evicted without being related to any individual characteristics.

To take advantage of the exogeneity of gentrification, I use the summons rate at a given block group—the smallest geographic unit for which there is census information—and year to measure "how exogeneous" the eviction of a household in the same block group-year is (Rossiter). The more evictions at a given block group-year, the more likely that the eviction of a household in that block group and time was due to gentrification, an exogenous reason. While exogenous and endogenous reasons for eviction can certainly co-occur, my instrumental variables strategy will be able to isolate the degree to which an individual's eviction was exogenous.

In order for an instrument to be valid, it must satisfy both the relevance and excludability conditions. I demonstrate empirically how this instrument satisfies the relevance condition in the empirical specification section. As for the excludability condition, it is not empirically testable with an overidentification test because I do not have more instruments than endogenous variables. As a result, to justify the excludability of this instrument, I can only provide the aforementioned theoretical reasons why eviction rate may be a good proxy for gentrification, an exogenous reason for eviction.

Empirical Specification

In order to provide a benchmark for my instrumental variables strategy, I estimate the OLS regression—the correlational empirical strategy typical of most of the eviction literature. Let $E_{i,j}$ be an indicator equal to 1 if household i is evicted in block group j; Y_i be the household income and X be a set of controls:

$$Y_i = \beta_0 + \beta_1 E_{i,j} + \beta_2 X + u_{i,t} \quad (1)$$

The set of controls X includes an indicator variable for the presence of children, which

Desmond, An, Winkler & Ferriss (2013) found to be associated with more evictions. In order to control for changes to income due to increases in work experience, it also includes household head age. Similarly, in order to control for annual changes in the regional economy, I include year indicators. Finally, X also includes household fixed effects, which sweep out time-invariant household characteristics such as race.

The instrumental variables strategy of this paper improves upon this basic OLS regression. In order to identify the causal impact of eviction on income, this paper exploits gentrification-related evictions, proxied through $Z_{j,t}$ —the number of evictions in the block group and at the time of a household's eviction. Essentially, the more evictions in a block group-year, the more likely an eviction in that block group-year is exogenous. In order to implement the instrumental variables empirical strategy, I use two stage least squares (2SLS) with first and second stage equations:

$$Logit(E)_{i,j} = \gamma_0 + \gamma_1 Z_{j,t} + \gamma_2 X + \epsilon_i \qquad (2)$$

$$Y_i = \beta_0 + \beta_1 E'_{i,j} + \beta_2 X + u_{i,t} \qquad (3)$$

Equation (2) purges endogeneity from evictions through the use of the instrument $Z_{j,t}$. Equation (3) then uses the "purged" version of eviction to estimate the causal impact of eviction on household income.

In order to test the relevance condition, which is whether the instrument $Z_{j,t}$ strongly predicts $E'_{i,j}$, I simply run the first stage and check whether the coefficient on the instrument is significant. Table 2 presents the results from the first stage, where the coefficient on eviction rate in block group-year is indeed significant at the .01 level.

Table 2: First Stage Least Squares

VARIABLES	(1) First Stage Least Squares
Eviction Rate in Block Group-Year of Eviction	10.07***

	(0.754)
2013	-0.355***
	(0.0433)
2014	-0.140***
	(0.0470)
2015	-0.232***
	(0.0504)
2016	-0.383***
	(0.0533)
2017	-0.576***
	(0.0573)
Presence of Children	-0.0378
	(0.0752)
Age: < 25	0.0902
	(0.224)
Age: 25-29	0.176
	(0.217)
Age: 30-34	0.0738
	(0.214)
Age: 35-39	0.0295
	(0.213)
Age: 40-44	-0.0199
	(0.213)
Age: 45-49	-0.00756
	(0.212)
Age: 50-54	-0.0574
	(0.212)
Age: 55-59	-0.120
	(0.211)
Age: 60-64	-0.127
	(0.212)
Age: 65+ (inferred)	-0.273
	(0.296)
Age: 65-69 (reported)	-0.166
	(0.214)
Age: 70-74 (reported)	0.0238
	(0.198)
Age: 75+ (reported) (omitted due to collinearity)	-
Observations	28,427

The outcome variable is an indicator variable that is 1 if evicted.
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Data

In this section, I describe the characteristics of and the merging strategy for two datasets new to the eviction literature: consumer data from marketing company InfoUSA and official eviction court data collected from by the non-profit DataWorks. I provide detail on how I generated specific variables and provide summary statistics on Durham.

InfoUSA Consumer Data

InfoUSA is a marketing and sales company that aims to "help businesses acquire, manage and retain customers" (InfoUSA). The company does so through its consumer database, which tracks 120 million households and 292 million individuals. The database, currently available from 2006 to 2017, is built and maintained using 29 billion records from 100 sources, such as census statistics, billing statements, telephone directory listings and mail order buyers/magazine subscribers. The unique identifiers assigned to each individual and to each household allow me to follow them over time. Particular variables of interest to this paper are gender, ethnicity, age, address, renter/owner status and estimated household income. The scope, length of time and variables captured by this dataset introduce a degree of resolution thus far unseen in the literature.

However, the dataset is limited in a few ways. First, it is unclear who is missing from the database and whether missing records are at random, making it difficult to determine possible bias. Second, several variables from InfoUSA used in this paper are estimated, which introduces some possible measurement error and uncertainty. For instance, household income is estimated by "one of three regression formulas using a combination of the Infogroup Consumer Database and census 2010 variables" (InfoUSA). This estimation may introduce inherent collinearity between household income and the geographic data I use to estimate the "exogeneity" of the eviction. Another variable used by this paper is household head age, which is estimated by InfoUSA in 5 year-wide bins. This bin size may lack the resolution necessary to absorb household income changes within the five years that this study tracks households.

Official Eviction Court Records from DataWorks

DataWorks is an "independent data intermediary dedicated to democratizing the use of quantitative information" based in Durham, North Carolina (Dataworks NC). Each month, DataWorks acquires civil process records from the Durham County Sheriff's Department, which contain records for both those who have only been summoned to eviction court and those who have actually been evicted after being summoned to eviction court.[3] These records, along with the residential address and dates received associated with them, are currently available from 2012 to 2018, inclusive.

Like other datasets based on court records, the advantages and disadvantages of this dataset from DataWorks are go hand-in-hand. On one hand, official eviction court records avoid issues of self-report data. On the other hand, they therefore do not capture informal evictions, which are estimated to make up almost three-fourths of all evictions despite being technically illegal (Desmond & Kimbro, 2015). Nevertheless, this dataset of eviction court data in Durham is the only one available to me to work with and merge with the InfoUSA dataset on the address level.[4]

---

[3] As the first of two court processes for formal eviction, summary ejectments summon tenants to a court proceeding which formally decides whether the tenant should be evicted or not. Summary ejectments represent a landlord's formal filing for eviction ("Summary Ejectments per Square Mile"). If the judge decides to evict the tenant, then the eviction process moves to the second stage and the tenant is served a "writ of possession," which orders the tenant and their belongings out of the property ("Summary Ejectments per Square Mile"). Far fewer writs of possession are served than summary ejectments. Using data on eviction filings and eviction rates from the Eviction Lab, about one-third of those who receive eviction filings are later formally evicted (Eviction Lab, 2018).

[4] Although the Eviction Lab, a Princeton University team of researchers, has created a national database of evictions on the address level, this data is only available to other researchers if they also share their own data for merging purposes. Because contract agreements with InfoUSA prevent me from sharing the data with the Eviction Lab, DataWorks is the best alternative. Moreover, the datasets should be identical, since they are both compiled using official court records.

Merging

I considered all households residing at the same address and year of a household that was evicted to also be evicted. I had to make this decision because the only way I could uniquely identify households from the eviction court data was through a combination of address and year. Therefore, if *multiple* households at the same year were found in the income data from InfoUSA, I could not distinguish which household was actually evicted—even if only one eviction was reported for that address. These false positives should conservatively bias any negative impacts of eviction on income because many households that were not actually evicted will be treated as "evicted" in the data. Appendix A contains further technical details on merging.

To reduce the number of false positives, I used income and ownership status data from InfoUSA to only apply "eviction" to households with incomes of $100,000 or less and households that were considered to be renters. Appendix B documents how the benchmark OLS results change depending on the aggressiveness of the trimming specification and discusses in further detail how I arrived at the cutoff of $100,000 for income and to only include renters. For any forthcoming regressions, I use the liberal trimming.

In order to prepare the dataset for the xt commands in Stata, I removed 9,824 household-year duplicates from the InfoUSA Dataset. Just 18 household-years comprised these 9,824 household-year duplicates—the remaining 823,455 household-year observations had no duplicates. This divergent distribution of duplicates was surprising to me and possibly indicates some validity issues with the InfoUSA dataset.

Variables of Interest

Post-eviction indicator

I used a household's *earliest* instance of eviction to mark which household-year observations were considered post- eviction.[5] For example, if a household were evicted in both 2013 and 2015, any observations of that household after 2013, not just 2015, would be considered post- eviction. It is not clear what direction this method of marking households as post- eviction would bias my results. On one hand, it may conservatively bias my results because the impacts of eviction tend to fade-out over time (Desmond & Kimbro, 2015; Humphries et. Al., 2018). By considering households as still being impacted by the effect of evictions even up to five years after the event, I am conservatively biasing my results. See Table 3, which displays how many household-year observations are considered "post- eviction." On the other hand, this method of marking households as post-eviction would attribute the impacts of any evictions after the first to the first. The number of evictions that households experience is shown in Table 4.

Table 3: Household-Year Observations Considered "Post-Eviction"[6]

Years after Earliest Eviction Observed	1	2	3	4	5
Number of Families	3,607	1,209	814	470	200

Table 4: Number of Evictions Experienced by Households[7]

Number of Evictions Experienced	1	2	3	4	5	6
Number of Families	23,180	7,541	2,816	1,163	487	263

Eviction Rate in Block Group-Year of Time of Eviction

---

[5] Some evictions/summons in the court records were marked as "returned"— meaning that the order returns unexecuted to the court. Summons orders can be returned if the plaintiff proves their case, the defendant admits the allegations of the complaint or the defendant fails to appear the day of court. Eviction orders can be returned for several reasons: after filing for eviction, the landlord may allow the tenant and/or their property to remain on the premises or the landlord may refuse to comply with a sheriff's request for advance transportation and storage cost for the tenant's property (North Carolina General Statutes).
For the 53 summons marked returned, I treated those observations as if they never had contact with eviction court. For the 877 evictions marked returned, I considered those cases as summons only, not evictions.
[6] This table includes false positives.
[7] This table includes false positives. It is impossible to separate them out because the only way I can identify unique households is through merging with InfoUSA.

Using renter populations from the American Community Survey and official court data from DataWorks, I calculate eviction rates per block group-year. Because I could not find the 2012 renter populations, I substitute the 2013 renter populations for those. I drop 26 block group-years with 0 renter population so as not to accidentally create eviction rates of infinity. After using the eviction rate in the first stage of instrumental variables, I attach the predicted probability of eviction *E'* to all household-years after their first eviction to run the reduced form regression.

Descriptive Statistics, Durham County (2012-2017)

Although the number of summons in Durham County has been decreasing from 2012 to 2018, the number of evictions has actually been increasing, indicating that a higher proportion of those summoned to eviction court have been evicted over time. Furthermore, although the number of summons has been decreasing, it is still the highest per capita rate among North Carolina's ten largest counties (Eviction Lab).

Figure X: Court Summons and Evictions in Durham (2012-2018)

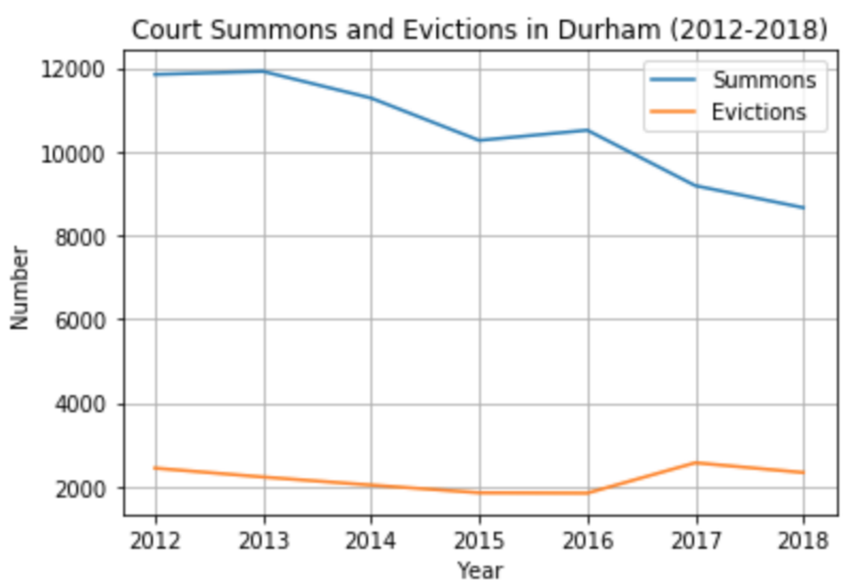

Table 5: Descriptive Statistics on Durham County (2012-2017)

Variables	**Full Durham Household-Year Sample** n = 823,473	**Household-Years at Year of Summons**[8] n = 86,175 False positives = 27,150	**Households-Years at Year of Eviction** n = 55,375 False positives = 45,117
Household Income[9] $5,000 means $5,000 or less, $500,000 means $500,000 or more.	Min: 5,122 Max: 546,850 Median: 45,556 Mean: 65.083 Std: 62,734	Min: 5,122 Max: 99,393 Median: 15,910 Mean: 20,358 Std: 15,363	Min: 5,122 Max: 99,369 Median: 15,366 Mean: 19,642 Std: 14,799
Owner/Renter Status[10]	Min: 0 Max: 9 Median: 7 Mean: 6.12 Std: 2.87	Min: 0 Max: 3 Median: 2 Mean: 1.72 Std: .79	Min: 0 Max: 3 Median: 2 Mean: 1.69 Std: .77
Race[11]	White: 61.8% Black: 21.9% Latinx: 5%	White: 62.8% Black: 21.0% Latinx: 5.2%	White: 62.4% Black: 21.4% Latinx: 5.3%
Age	< 25: 7.1% 25-29: 10.3% 30-34: 11.8% 35-39: 10.0% 40-44: 9.7% 45-49: 9.5%	< 25: 16.9% 25-29: 20.4% 30-34: 16.0% 35-39: 10.2% 40-44: 8.0% 45-49: 7.0%	< 25: 17.7% 25-29: 20.0% 30-34: 15.6% 35-39: 10.3% 40-44: 8.3% 45-49: 7.2%

Results

Table 6 presents the effect of eviction on household income estimated through two

[8] Summons here includes households that were summoned, but not as evicted, as well as those who were eventually evicted.
[9] Because the household incomes have been adjusted for inflation, incomes originally at the lower and upper bounds of $5,000 and $500,000 are slightly different now.
[10] These values are estimated from InfoUSA. The numbers mean as follows: 9 = Homeowner (reported), 8-7 = most likely Homeowner, 6-4 = unknown, 3-1 = most likely Renter, 0 = Renter (reported).
[11] InfoUSA estimates race through name-ethnicity matching. Because that process is proprietary, it is not clear how accurate it is or how they define particular races like Hispanic/Latinx.

methods—both with household fixed effects. The benchmark OLS regression typical of the majority of the eviction literature is presented in column 1, without controls, and column 2, with controls. The instrumental variables (IV) approach, with controls, is presented in column 3.

Table 6: The effect of eviction on household income (levels)

VARIABLES	(1) Benchmark OLS, without Controls	(2) Benchmark OLS, with Controls	(3) Instrumental Variables
Eviction	1,914***	-4,429***	-3,437.51***
	(394.6)	(383.1)	(677.3)
Presence of Children		3,688***	3,686***
		(195.6)	(195.6)
2012 (omitted due to collinearity)		-	-
2013		-1,736***	-1,716***
		(122.4)	(122.4)
2014		-15,414***	-15,363***
		(126.0)	(125.9)
2015		-14,875***	-14,801***
		(131.3)	(131.2)
2016		-14,321***	-14,225***
		(138.5)	(138.3)
2017		-14,595***	-14,482***
		(144.5)	(144.3)
Age: < 25		14,672***	14,638***
		(582.3)	(582.3)
Age: 25-29		21,568***	21,593***
		(532.2)	(532.2)
Age: 30-34		24,613***	24,664***
		(508.5)	(508.6)
Age: 35-39		28,224***	28,276***
		(497.2)	(497.2)
Age: 40-44		36,370***	36,418***
		(486.7)	(486.7)
Age: 45-49		34,347***	34,395***
		(475.9)	(476.0)
Age: 50-54		34,696***	34,738***
		(465.9)	(466.0)
Age: 55-59		36,377***	36,414***
		(455.1)	(455.2)
Age: 60-64		34,062***	34,095***
		(441.9)	(441.9)

Age: 65+ (inferred)		25,489***	25,511***
		(676.0)	(676.1)
Age: 65-69 (reported)		21,117***	21,139***
		(427.2)	(427.2)
Age: 70-74 (reported)		11,683***	11,696***
		(375.0)	(375.0)
Age: 75+ (reported) (omitted due to collinearity)		-	-
Constant	65,057***	48,432***	48,303***
	(32.32)	(434.8)	(434.8)
Observations	823,473	823,473	823,473
R-squared	0.000	0.076	0.076
Number of Households	289,983	289,983	289,983
Household FE	NO	YES	YES

The outcome variable is household income, as estimated by InfoUSA and inflated to 2019 dollars.
Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

I discuss the implications of these models in turn.

The comparison of the benchmark OLS with and without controls (in Columns (1) and (2), respectively) demonstrates the necessity of controls.[12] Without controls, the benchmark OLS has a positive and significant coefficient on the impact of eviction. This result seems to suggest that those who have been evicted will have household incomes of about $2,000 higher than those who have not—despite the financial and emotional stress associated with a forced move. However, once controls are included in Column (2) for the presence of children, year and household head age, the coefficient on impact of eviction on income becomes negative and significant, as expected. Furthermore, the R-squared increases from 0.000 to 0.076. Although 0.076 is low for an R-squared, it seems reasonable given the number of false positives in my dataset, the estimated variables from InfoUSA and the noisy nature of household poverty.

[12] See Appendix C and footnote 2 for more details on the summons variable.

The IV estimate in column (3) attenuates the OLS estimate in column (2) towards 0—about -$3,500 as opposed to about -$4,500. This difference of $1,000 is about one fourth of the OLS estimate. Because the IV estimate purges the OLS estimate of some of the endogeneity associated with eviction, we may consider the IV estimate as the lower bound and the OLS estimate as the upper bound of the impact of eviction on income. This IV estimate is this paper's major contribution to the eviction literature, suggesting that future researchers may consider utilizing this instrumental variables strategy to pin down the true causal impact of eviction.

Nonetheless, regardless of whether we consider the lower or upper bound, the impact of eviction is large— especially given that data limitations forced both regressions include a false positive ratio of about 4:1—theoretically biasing the results towards 0. With either estimate, given that the median income of households at the year of summons/evictions is about $15,000 (see Table 5 from the data section), eviction's impact on income is substantial. Table 7 presents the impact of eviction on log household income, showing a 2.5% to 5.7% decrease in income. Again, the IV regression results are less negative than the OLS results, providing a lower bound on the impact of eviction that is still significant and substantial.

Table 7: The effect of eviction on household income (%)

VARIABLES	(1) Benchmark OLS, without Controls	(2) Benchmark OLS, with Controls	(3) Instrumental Variables
Eviction	-0.00434	-0.0569***	-0.0251**
	(0.00566)	(0.00552)	(0.00976)
Presence of Children		0.0438***	0.0437***
		(0.00282)	(0.00282)
2012 (omitted due to collinearity)		-	-
2013		-0.00705***	-0.00676***
		(0.00176)	(0.00176)
2014		-0.187***	-0.186***
		(0.00182)	(0.00181)

2015		-0.177***	-0.176***
		(0.00189)	(0.00189)
2016		-0.139***	-0.138***
		(0.00200)	(0.00199)
2017		-0.142***	-0.140***
		(0.00208)	(0.00208)
Age: < 25		0.0984***	0.0978***
		(0.00839)	(0.00839)
Age: 25-29		0.330***	0.330***
		(0.00767)	(0.00767)
Age: 30-34		0.405***	0.406***
		(0.00733)	(0.00733)
Age: 35-39		0.449***	0.450***
		(0.00716)	(0.00717)
Age: 40-44		0.563***	0.563***
		(0.00701)	(0.00701)
Age: 45-49		0.522***	0.523***
		(0.00686)	(0.00686)
Age: 50-54		0.551***	0.551***
		(0.00671)	(0.00671)
Age: 55-59		0.558***	0.558***
		(0.00656)	(0.00656)
Age: 60-64		0.524***	0.524***
		(0.00637)	(0.00637)
Age: 65+ (inferred)		0.407***	0.407***
		(0.00974)	(0.00974)
Age: 65-69 (reported)		0.323***	0.323***
		(0.00616)	(0.00616)
Age: 70-74 (reported)		0.181***	0.181***
		(0.00540)	(0.00540)
Age: 75+ (reported) (omitted due to collinearity)		-	-
Constant	10.63***	10.33***	10.33***
	(0.000464)	(0.00627)	(0.00627)
Observations	823,473	823,473	823,473
R-squared	0.000	0.067	0.067
Number of Households	289,983	289,983	289,983
Household FE	NO	YES	YES

The outcome variable is log household income, as estimated by InfoUSA and inflated to 2019 dollars.
Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

This decrease in income is somewhat larger than what Collinson and Reed (2018) find using New York City eviction court data, which is that cumulative earnings drop by about $3,000 after eviction. However, after introducing controls such as quarterly earnings and employment history up to and including the quarter of filing, Collinson and Reed (2018) find that these OLS estimates reduce to about -$1,760. If this paper had access to those controls, my OLS estimates would perhaps also become smaller and more similar to Collinson and Reed (2018). Finally, this comparison of the OLS results between this paper and Collinson and Reed (2018) suggests that there may be significant differences in labor markets between cities, leading to different income changes after eviction.

The implications of these results for policy depend on the cost-benefit ratios of eviction prevention policies in comparison to the cost-benefit ratios of other welfare policies. The IV estimates presented in this paper, Humpries et. al. (2018) and Collinson and Reed (2018) suggest significant, though small detriments due to eviction. While eviction prevention policies can certainly play a role in mitigating or avoiding these consequences, they must be considered in a cast of poverty prevention policies. Furthermore, if gentrification and lack of affordable housing options are the main exogenous causes of eviction (as I discuss in the theoretical framework section), then overemphasizing the importance of eviction prevention would only be *re*active to poverty, not *pro*active. Future research on evictions may consider investigating the impact of eviction prevention policies and comparing their worth vis-à-vis to other poverty prevention policies.

Conclusion

In this paper, I contribute to the burgeoning literature on evictions, which are an important aspect of the affordable housing crisis facing low-income American renters. Similar to Humphries et. al. (2018) and Collinson and Reed (2018)—other recent empirical papers on the impact of eviction— I successfully use an instrumental variables strategy to estimate the impact of eviction on household income. Specifically, I exploit gentrification-related evictions by using eviction rates in block group-years as a proxy. Using this instrumental variables strategy, I find a 2.5% decrease in household income, which is a small, but significant blow to vulnerable households. My paper contributes not only a concrete number on the impact of evictions, but also furthers the ability to make causal inference.

Works Referenced

"2010 Census Tallies of Census Tracts, Block Groups & Blocks." *United States Census Bureau.* Web. <https://www.census.gov/geo/maps-data/data/tallies/tractblock.html>.

Collinson, Robert, and Davin Reed. *The Effect of Evictions on Low-Income Households*., 2018. Print.

Curtis, Marah A., et al. "Maternal Depression as a Risk Factor for Family Homelessness." *American Journal of Public Health* 104.9 (2014): 1664-70. Web.

"Dataworks NC." Web. <https://dataworks-nc.org/>.

Desmond, Matthew. *Evicted: Poverty and Profit in the American City*. Crown, 2016. Print.

---. "Eviction and the Reproduction of Urban Poverty." *American Journal of Sociology* 118.1 (2012): 88-133. Web.

Desmond, Matthew, et al. "Evicting Children." *Social Forces* 92.1 (2013)Web.

Desmond, Matthew, and Carl Gershenson. "Housing and Employment Insecurity among the Working Poor." *Social Problems* 63.1 (2016): 46-67. Web.

Desmond, Matthew, Carl Gershenson, and Barbara Kiviat. "Forced Relocation and Residential Instability among Urban Renters." *Social Service Review* 89.2 (2015): 227-62. Web.

Desmond, Matthew, and Rachel Tolbert Kimbro. "Eviction's Fallout: Housing, Hardship and Health." *Social Forces* 94.1 (2015): 295-324. Web.

Desmond, Matthew, and Tracey Shollenberger. "Forced Displacement from Rental Housing: Prevalence and Neighborhood Consequences." *Demography* 52.5 (2015): 1751-72. Web.

Dictionary. "Gentrification." Web. <https://www.dictionary.com/browse/gentrification>.

Eviction Lab. "Eviction Lab." Sept 18, 2018. Web. <https://evictionlab.org/>.

Geller, Amanda, and Allyson Walker Franklin. "Paternal Incarceration and the Housing Security of Urban Mothers." *Marriage and Family* 76.2 (2014): 411-27. Web.

Green Building Registry. “US Address-Scourgify.” April 12, 2019. Web.

<https://github.com/GreenBuildingRegistry/usaddress-scourgify>.

Humphries, John Eric, et al. *Does Eviction Cause Poverty? Quasi-Experimental Evidence from Cook County, IL*., 2018. Print.

InfoUSA. "About Us: InfoUSA." Web. <https://www.infousa.com/about-us/about-infousa/>.

---. *Consumer Data Overview*., 2018. Print.

Joint Center for Housing Studies of Harvard University. *The State of the Nation's Housing 2018*., 2018. Print.

Lundberg, Ian, and Louis Donnelly. *The Prevalence of Housing Eviction in Urban Families*., 2016. Print.

National Low Income Housing Coalition. *The Gap: A Shortage of Affordable Homes*., 2018. Print.

North Carolina. *North Carolina General Statutes*., 2016. Print.

Reichman, Nancy E., et al. *Fragile Families: Sample and Design*. 23 Vol. , 2001. Web.

Rossiter, Katy. *What are Census Blocks?*., 2011. Print.

"Summary Ejectments per Square Mile." *Durham Neighborhood Compass.* Web. <https://compass.durhamnc.gov/#SUMEJECT/blockgroup/>.

The City of Durham Department of Neighborhood Improvement Services. *Durham, North Carolina Tenant Training Program Manual*., 2012. Print.

The Eviction Lab. "National Estimates: Eviction in America." May 11, 2018. Web. <https://evictionlab.org/national-estimates/>.

Vaughan, Dawn B. "Gentrification in Durham: 'It's all about the Dollar Bill.' A Renter's Story." *Herald Sun,* May 23, 2018, Web. <https://www.heraldsun.com/news/local/counties/durham-county/article211591644.html>.

Vaughan, Dawn B., and Zachery Eanes. "'Story of My Street': Gentrification and Neighborhood Change in Durham." *Herald Sun,* April 24, 2018, Web. <https://www.heraldsun.com/news/local/counties/durham-county/article204694739.html>.

Appendix A

The first matching technique I pursued was simply standardizing the addresses and then matching the strings. I used a Python library found on Github called usaddress-scourgify to "remove unacceptable special characters, extra spaces, predictable abnormal character sub-strings and phrases" as well as "abbreviate directional indicators and street types according to the abbreviation mappings found in address_constants" (Green Building Registry). Then I merged on the address strings.

There were two main reasons why address strings from the eviction court data were not able to be matched with addresses from InfoUSA. First, typos in the eviction court data addresses prevented a complete merge (e.g. 1 BUCKEAD CT v. 1 BUCKHEAD CT). A priori, these typos should be randomly distributed among the addresses. Second, some addresses present in the eviction court data were not present in the InfoUSA data, though I could manually identify some of them via Google Maps.

Appendix B

Table 8: Benchmark OLS Regressions Under Different Trimming Specifications

VARIABLES	(1) No Trimming	(2) Conservative Trimming	(3) Liberal Trimming
Eviction	-5,108***	-5,452***	-4,429***
	(332.8)	(347.6)	(383.1)
Presence of Children	3,689***	3,689***	3,688***
	(195.6)	(195.6)	(195.6)
2012 (omitted due to collinearity)			-
2013	-1,763***	-1,753***	-1,736***
	(122.5)	(122.4)	(122.4)
2014	-15,463***	-15,453***	-15,414***
	(126.1)	(126.0)	(126.0)
2015	-14,942***	-14,929***	-14,875***

	(131.5)	(131.4)	(131.3)
2016	-14,408***	-14,396***	-14,321***
	(138.7)	(138.6)	(138.5)
2017	-14,702***	-14,689***	-14,595***
	(144.9)	(144.8)	(144.5)
Age: < 25	14,701***	14,709***	14,672***
	(582.2)	(582.2)	(582.3)
Age: 25-29	21,577***	21,575***	21,568***
	(532.1)	(532.1)	(532.2)
Age: 30-34	24,607***	24,604***	24,613***
	(508.4)	(508.4)	(508.5)
Age: 35-39	28,214***	28,209***	28,224***
	(497.1)	(497.1)	(497.2)
Age: 40-44	36,358***	36,356***	36,370***
	(486.6)	(486.6)	(486.7)
Age: 45-49	34,333***	34,331***	34,347***
	(475.9)	(475.9)	(475.9)
Age: 50-54	34,678***	34,672***	34,696***
	(465.9)	(465.9)	(465.9)
Age: 55-59	36,366***	36,358***	36,377***
	(455.1)	(455.1)	(455.1)
Age: 60-64	34,055***	34,045***	34,062***
	(441.8)	(441.8)	(441.9)
Age: 65+ (inferred)	25,493***	25,486***	25,489***
	(675.9)	(675.9)	(676.0)
Age: 65-69 (reported)	21,113***	21,108***	21,117***
	(427.1)	(427.1)	(427.2)
Age: 70-74 (reported)	11,684***	11,678***	11,683***
	(375.0)	(375.0)	(375.0)
Age: 75+ (reported) (omitted due to collinearity)	-	-	-
Constant	48,530***	48,521***	48,432***
	(434.9)	(434.8)	(434.8)
Observations	823,473	823,473	823,473
R-squared	0.076	0.076	0.076
Number of Households	289,983	289,983	289,983

The outcome variable is household income, as estimated by InfoUSA and inflated to 2019 dollars.
Standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

I chose the more liberal trimming specification for two reasons. First, the coefficient on eviction became less negative and closer to other estimates found in the literature. Second, from a theoretical perspective, aggressively excluding owners and rich people made intuitive sense given that my instrumental variables is predicated on the impact of gentrification. Owners and rich people are less likely to be impacted by gentrification and eviction.

Appendix C

Table 9: Including Summons as a Control for Benchmark OLS Regressions

VARIABLES	(1) Benchmark OLS, without Controls	(2) Benchmark OLS, with Controls
Eviction	-1,368***	-5,572***
	(456.2)	(440.1)
post_first_s_ad_yr	6,521***	2,318***
	(455.4)	(439.3)
Presence of Children		3,686***
		(195.6)
2012 (omitted due to collinearity)		-
2013		-1,727***
		(122.4)
2014		-15,393***
		(126.0)
2015		-14,844***
		(131.4)
2016		-14,282***
		(138.7)
2017		-14,545***
		(144.8)
Age: < 25		14,661***
		(582.3)
Age: 25-29		21,578***
		(532.2)
Age: 30-34		24,635***
		(508.5)
Age: 35-39		28,248***
		(497.2)
Age: 40-44		36,393***
		(486.7)
Age: 45-49		34,371***
		(476.0)
Age: 50-54		34,714***

		(465.9)
Age: 55-59		36,395***
		(455.1)
Age: 60-64		34,077***
		(441.8)
Age: 65+ (inferred)		25,498***
		(676.0)
Age: 65-69 (reported)		21,124***
		(427.2)
Age: 70-74 (reported)		11,688***
		(375.0)
Age: 75+ (reported)		-
(omitted due to collinearity)		
Constant	65,013***	48,376***
	(32.46)	(434.9)
Observations	823,473	823,473
R-squared	0.000	0.076
Number of Households	289,983	289,983

The outcome variable is household income, as estimated by InfoUSA and inflated to 2019 dollars. Standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1