2021-01

A rising tide drowns unstable boats: how inequality creates homelessness

A Rising Tide Drowns Unstable Boats: How Inequality Creates Homelessness

Thomas Hugh Byrne, Benjamin F. Henwood, & Anthony W. Orlando

CONDITIONALLY ACCEPTED AT *THE ANNALS OF THE AMERICAN ACADEMY OF POLITICAL AND SOCIAL SCIENCE*

AUGUST 2020

Abstract

This article examines whether income inequality is a driver of homelessness at the community level. We theorize that inequality is likely to affect homelessness through both an "income channel"—by leading to the crowding out of low-income households from the rental market—and a "price channel" through which inequality leads to rising home prices. We construct a dataset combining information on inequality, homelessness, rent burden, and housing prices from 239 communities for the period from 2007 to 2018 to assess the income inequality-homelessness relationship. Our results suggest that income inequality is a significant driver of homelessness at the local level, and that the "income channel" is the more likely mechanism through which inequality creates homelessness. We argue that broader policy efforts to reduce income inequality are likely to have the collateral effect of reducing homelessness, and discuss the need for national and local policies to help low-income households afford housing.

Corresponding Author: Thomas Byrne, Boston University School of Social Work, 264 Bay State Rd., Boston, MA 02215; Phone: 617-358-0783; email: tbyrne@bu.edu

Benjamin Henwood: bhenwood@usc.edu

Anthony W. Orlando: anthony.w.orlando@gmail.com

Keywords: Homelessness; Income inequality; Housing; Housing cost; Cities;

Author biographies

Thomas Byrne is an Assistant Professor at the Boston University School of Social Work and Investigator at the U.S. Department of Veterans Affairs Center for Healthcare Outcomes and Implementation Research and National Center on Homelessness among Veterans.

Benjamin Henwood is an Associate Professor and the Director of the Center for Homelessness, Housing and Health Equity Research at the University of Southern California Suzanne Dworak-Peck School of Social Work. He is the lead author of the American Academy of Social Work and Social Welfare's grand challenge to end homelessness.

Anthony W. Orlando is Assistant Professor of Finance, Real Estate, & Law and Donor's Scholar of Analytics at California State Polytechnic University, Pomona. He is a Visiting Scholar at the Federal Reserve Bank of Atlanta and Faculty Affiliate of the USC Bedrosian Center on Governance and the Public Enterprise.

Acknowledgements

The authors wish to acknowledge Leah Boustan for providing the data and code that were essential in constructing the measures of income inequality used in this study.

Introduction

Does income inequality have an impact on homelessness? Answering this question is of obvious importance for improving our understanding of the structural drivers of homelessness and, consequently, for developing policy responses to reduce the scope of homelessness. To be clear, we are not the first to pose this question or to seek an answer to it. However, this study provides what we believe is the most rigorously empirical response to the question to date. Put simply: our answer is that "yes," increasing local income inequality appears to be an important driver of increasing homelessness in a given housing market. More pointedly, our analysis suggests that in a community of 740,000 people (the average size of the communities in the dataset we use to examine the income inequality-homelessness relationship, or roughly the population Seattle), an increase in local income inequality that tracked the average nationwide increase in income inequality in the United States over the period from 2007-2018 would translate into roughly 200 additional people experiencing homelessness in that community on a given night. While the precise impact of income inequality on homelessness appears to vary as a function of whether we consider persons who are homeless as individuals or persons in families, or those who are sheltered as opposed to unsheltered, our results point clearly to a key role for income inequality in generating homelessness.

Before delving too deeply into the answer to our central question, however, it is necessary to place this question in the appropriate conceptual and empirical context. Thus, the remainder of this paper proceeds first by examining the rise of contemporary homelessness in the United States and related research seeking to identify the structural determinants of homelessness. We then similarly discuss the rise of income inequality in the United States and its origins. Next, we present an explication of the theoretical mechanisms by which income

inequality and homelessness are likely to be linked. We then describe the data and analytic approach we use to assess the extent to which inequality is associated with homelessness in the United States, and a full description of the results of that analysis. We conclude with a brief discussion of the implications of our findings for policy responses to homelessness.

The Emergence and Search for Structural Determinants of Contemporary Homelessness in the United States

Homelessness in the United States is a long-standing problem. Since the Great Depression, there have been transient workers who lacked a fixed geographic home or were socially disconnected (Hopper 1991). "Modern day" homelessness, however, is a more recent phenomenon (Lee, Tyler, and Wright 2010). The archetypical portrayal of contemporary homelessness as individuals bedding down on sidewalks, in doorways, and other public places first emerged in the 1980s, and is widely believed to have increased during that decade. Today, the scope of the problem remains substantial by any measure, and homelessness continues to garner public attention. The U.S. Department of Housing and Urban Development (HUD) estimates that, on a given night in 2019, 567,715 people in the United States were experiencing homelessness, with one-third in unsheltered situations and two-thirds in sheltered situations (U.S. Department of Housing and Urban Development 2020).

The emergence and persistence of contemporary homelessness has spurred social scientists to increasingly look at societal or structural factors to understand its rise, persistence, and variation across communities and over time (Shinn and Khadduri 2020).[1] For example, demographic changes in the U.S. population after Word War II resulted in many, mostly African American, young adults being excluded from the labor maket during times of economic

recession in the 1970s and 1980s. This coincided with an epochal shift in federal housing policy away from direct production of low-income housing units toward less generous tenant subsidies targeted toward fewer households (Von Hoffman 2012). Even as poverty and homelessness grew, funds for low-income residents were diverted to other budget priorities, and the growth of federal housing assistance slowed (Wolch, Dear, and Akita 1988). That deceleration has continued to this day (Rice 2016) to the point that 75% of low-income, cost-burdened renters do not receive federal assistance (Fischer and Sard 2017).

Poverty is also a frequently hypothesized structural determinant of homelessness, but the fact that approximately 12% of Americans live in poverty (Semega et al. 2019) yet less than 1% of Americans experience homelessness (U.S. Department of Housing and Urban Development 2020) suggests something else, or more, is needed to explain homelessness. An extensive body of research suggests that housing affordability may be a key factor in explaining homelessness, with studies consistently finding a positive relationship between rent levels and homelessness (Hanratty 2017; Byrne et al. 2013; Lee, Price-Spratlen, and Kanan 2003; Glynn and Fox 2019).

While contemporary homelessness is believed to have emerged and grown over the past few decades, reliable data on the trajectory of the size of the homeless population over the entirety of this time period is not available. Consistent annual estimates of the size of the homeless population have only been available since 2007, when HUD began requiring communities to conduct annual "point in time" (PIT) counts of the number of persons experiencing homelessness on a given night at the end of January. This more recent time series points to a different aggregate trend and one that runs in the opposite direction of the conventional wisdom: the total number of people experiencing homelessness has actually decreased over the period from 2007 to 2019. However, as we also detail below, this reduction

in homelessness has not been uniform across all communities, with some seeing substantial rises in the number of people experiencing homelessness over this time period. Indeed, the rise of homelessness in some large communities has meant that a decade-long downward trend in homelessness in the United States between 2007 and 2016 has reversed, and there have been upticks in the number of people experiencing homelessness in the United States each year between 2016 and 2019 (U.S. Department of Housing and Urban Development 2020).

Why has homelessness increased significantly in some cities while the overall national rate has gone down? Figure 1 shows one potential explanation: differences in income inequality. At a local level, homelessness rates are positively correlated with inequality in every year for which data are available, from 2007 to 2018. (**INSERT FIGURE 1 Between Community Relationship of Income Inequality and Total Rate of Homelessness, 2007-2018 HERE)** This evidence has led a few scholars to suggest income inequality as a potential structural driver of homelessness in the United States (Orlando 2013; Shinn 2010; O'Flaherty 1996), which is one of the world's richest countries yet also one of the most unequal with higher rates of homelessness than its peer countries. Yet, empirical research on whether and how income inequality impacts homelessness remains limited. We seek to address this gap in the present study.

How Inequality Creates Homelessness: A Theoretical Model

The rise of inequality in the United States has attracted significant attention in recent years, both academically and politically, in part because it has a simple punchline: Since the late 1970s, incomes have grown faster for high-income households than for low- to middle-income (LMI) households. As a result, high-income households now earn a larger share of the country's annual output than they did only a few decades earlier.

Economists have tracked this evolution with several different measures. The classic Gini coefficient (Gini 1912), which we use in this analysis, has the benefit of being easy to approximate with only a rough breakdown of incomes across the distribution. The Gini coefficient estimates the degree to which the share of income earned by each percentile of the distribution differs from the share of the population in that percentile. It thus ranges from zero to one, with zero representing perfect equality (all households have the same income) and one representing perfect inequality (one household has all the income).[2] Figure 2 shows this measure alongside other measures of inequality, namely the portion of national income earned by the richest 10% and 1% of the distribution, for comparison. (**INSERT FIGURE 2 Comparison of Inequality Measures, 1947-2014 HERE).** In all three cases, inequality declines slowly from 1950 to 1970. By 1980, all three measures are rising, and though there are a few cyclical pauses, they continue rising until the end of the time series. So sharp is this rise that the data pass their 1950 peak by 1990—and they keep rising thereafter. By 2014, the top 10% is earning nearly half of national income, and the top 1% has doubled their share, from one-tenth to two-tenths of every dollar produced. These measures likely understate the full extent of inequality, as they only capture annual income, not the full accumulation over time in the form of wealth. Access to saved funds, such as intergenerational transfers, is an important determinant of the ability to purchase a home (Lee, Farrell, and Link 2004), though they likely play a smaller role in monthly rent payments for low-income households. For these households living paycheck-to-paycheck, it is income that matters most—and moreover, it is in these income statistics that we find the most available data for our analysis.

Alongside this rise in inequality has been an equally troubling rise in housing costs, both for homeowners and for renters. This trend has accelerated since the end of the Great Recession,

with housing prices and rents growing much faster than incomes, especially on the coasts where major urban areas have restricted housing supply, making it unable to keep up with the influx of demand. In the wake of the financial and foreclosure crises, many borrowers found themselves shut out of the mortgage market, "diverting" these homeowners into the rental market (Myers et al. 2016). As a result, nearly half of renters are now "cost-burdened," meaning they spend at least 30% of their income on housing. In all, over 37 million U.S. households are cost-burdened, of which 18 million spend more than *half* of their income on housing (Joint Center for Housing Studies of Harvard University 2019). While this outcome may seem to be yet another effect of rising inequality, researchers have not established a convincing theoretical or empirical connection between the two. The similarities are too glaring to ignore. Just as the economy has grown rapidly for some and sluggishly for others, so too have the housing markets in many cities seen rapid price growth for homeowners at the same time as increasingly unbearable cost burdens for others, which may be key to understanding the uneven distribution of homelessness across the U.S.

In order to understand this phenomenon, it is important to distinguish between two different types of inequality: between-city inequality and within-city inequality. As Diamond (2016), Ganong and Shaog (2017), and Hsieh and Moretti (2019) have shown, a significant contributor to the national rise in inequality has been the divergence between high-wage cities with high housing prices and high-skilled workers, on the one hand, and low-wage cities with low housing prices and low-skilled workers, on the other hand. For a local phenomenon such as homelessness, it is inequality *within* each city, not *between* cities, that matters. That is, homelessness in Los Angeles is likely driven more by the dynamics of the housing market in Los

Angeles than by the differences between Los Angeles and Kansas City. Thus, it is possible for homelessness to decline nationally and yet increase in Los Angeles if *local* inequality is increasing in Los Angeles but not in many other cities. (National inequality can still increase if between-city inequality is growing.) To better understand these dynamics, we consider how economic growth increases housing demand and propose two main mechanisms through which local inequality and local homelessness may be causally connected**:** an "income channel" and a "price channel."

Income channel

First, we consider the "income channel." Let us begin with the theoretical construct that economic growth increases the demand for housing and thus the price of housing. If housing supply is inelastic, as a large body of evidence indicates for many U.S. cities (Saiz 2010; Glaeser and Gyourko 2018), new housing construction does not respond to this price increase, and not enough new housing (even of the higher-quality housing that will filter down to low-income households) is produced. The result is increased housing cost burden for all households. If income growth is equally distributed, then housing cost burden is equally affected across the income distribution—that is, incomes rise proportionally across the income distribution to meet this increase in housing prices. Thus, the lowest-income households on the bottom tier of the housing ladder experience an income increase proportional to those at the top; if there is an increase in homelessness, it is because of the increase in prices, not differences in income growth. However, when income growth is skewed toward the highest earning households, housing burden increases disproportionately for the bottom and middle of the income distribution. Thus, households towards the bottom of the income distribution will be relatively poorer than their fellow residents, who continue to drive up housing costs for the city as a whole.

Even if the price growth is not abnormally high, low-income residents lose the competition to bidders with higher income growth. Some households will remain housed by devoting a larger share of their income towards housing, but others will be unable to afford housing at all, resulting in increased homelessness. This is what we refer to as the "income channel" through which inequality produces homelessness. In order for this channel to be true, we would expect income inequality to increase both homelessness and the share of low-income households who are rent-burdened.[3]

Price channel

We also consider the "price channel" as a mechanism through which income inequality may increase homelessness. Again, we return to the construct that economic growth increases the demand for housing and thus the price of such housing. However, cities with higher inequality attract faster housing price appreciation because they have more high-income, high-skilled, high-productivity workers, who increasingly command rents in the global economy. As a result, in the context of income inequality, low-income households will experience higher cost burdens not just because their income lags behind, but also because higher levels of income inequality will accelerate housing price growth. Importantly, this cost acceleration can take the form not just of an increase in the purchase price of the houses but also of all the investments that go into the houses, including energy prices and renovation costs. This is the price channel mechanism by which increasing income inequality might increase homelessness. Even if income growth is equally distributed, these cities are so desirable that prices exceed the threshold of affordability for many low-income households.

Despite the highly plausible theoretical link between income inequality and homelessness, empirical evidence of such a relationship is somewhat limited. O'Flaherty (1996)

presents some empirical support for his theory of the relationship between inequality and homelessness, although it is based on data from the 1970s and 1980s in only a handful of cities. Toro and colleagues (2007) and Shinn (2010) use cross-national comparisons between the United States and a number of European nations to show that countries with higher income inequality tend to have higher rates of homelessness. Finally, in the study most closely related to the present one, Wood et al. (2015) use panel data from 328 localities in Australia to examine the structural drivers of homelessness, finding that within-community increases in the Gini coefficient are positively associated with the rate of homelessness.

Study Question

We build on existing studies by leveraging data on income inequality and homelessness from a large set of U.S. communities over the period from 2007-2018. In doing so, our analysis offers the most robust analysis to date of how income inequality is related to homelessness in the United States. Specifically, we seek to address the following questions: Does income inequality have an impact on homelessness at the local level? And, on a more exploratory basis, does the relationship between income and inequality that may be impacting homelessness operate through the income channel, price channel, or both?

Methods

Data

The present study relies primarily on publicly-available data from the HUD and the U.S. Census Bureau, which we use to construct a novel dataset that allows us to capture changes in income inequality, homelessness, renter cost burden and home values in a sample of 239 communities over a 12-year period from 2007 to 2018. More precisely, these "communities" are comprised of single counties or aggregations of multiple counties known as Continuums of Care

(CoCs), which are geographic units at which HUD administers federal homeless assistance dollars and at which enumerations of the homeless population are conducted. To enable the construction of annual measures of income inequality, our sample includes the sub-set of CoCs comprised fully of counties above a certain population threshold (>65,000 people). For the sake of simplicity we use the terms "community" and "CoC" interchangeably throughout the remainder of the manuscript. Below, we briefly describe the steps we used to construct our dataset, with additional details provided in the Appendix.

First, to measure income inequality, we mirror the approach used by Boustan and colleagues (2013) by using aggregate data and microdata from the Census' American Community Survey (ACS) to construct county-level Gini coefficients on an annual basis for each year from 2007 to 2018. We also obtain measures of renter cost burden, home values and additional county-level variables from the ACS and other sources.

We then merge these county-level data with annual point-in-time (PIT) counts of homelessness and annual data on the number of permanent supportive housing (PSH) units in a community from HUD. The inclusion of the number of PSH units in our analysis is important to account for recent expansions in PSH and the documented association between such expansions and reductions in homelessness (Evans et al. 2019).

We merge our county-level income inequality and other measures with the HUD data to construct a CoC-level dataset using a modified version of the process described by Byrne and colleagues (2013). Because the ACS county-level measures we use in our analysis are only available on an annual basis for counties with a population of 65,000 or greater, our final sample is comprised of 239 CoCs (out of 398 active CoCs in 2018) for which we have up to 12 years of data. These 239 communities account for 77% of the total number of persons experiencing

homelessness nationwide on a given night in 2018. We conduct supplemental analysis using data from all CoCs for a more limited number of years. Results from that analysis are similar to the main results and are detailed in the Appendix.

Analysis

Our analytic strategy involves estimating a series of regression models that explore the relationship between income inequality and total homelessness (measured as a rate per 10,000 members of the general population), the proportion of rent-burdened low-income renters (defined as households with income <$20,000 paying >30% of income on rent); and median value of owner-occupied housing units (hereafter referred to as median home value), at the community-level. The models using the latter two outcome measures allow us to test the "income channel" and "price channel"[4] mechanisms by which inequality might drive homelessness, respectively. In addition, to further explore the relationship between income inequality and homelessness, we estimate models in which we use homeless individuals, persons in families, sheltered homeless persons, unsheltered homeless persons, and chronically homeless persons as outcomes.

In all of these models, the Gini coefficient is the primary independent variable of interest and we also control for the following set of time-varying control variables where appropriate: median household income, median rent for a two-bedroom apartment, median home value, poverty rate, real GDP, and the number of PSH beds per 10,000 people. We briefly describe our modeling strategy below, with full details provided in the Appendix.

We estimate two types of regression models. First, we estimate a series of ordinary least squares (OLS) regression models. In addition to the time-varying control variables described above, these models also include community and year fixed effects. The inclusion of these fixed effects is intended to control for, respectively, any time-invariant differences between

communities and any secular time trends that may confound the relationship between income inequality and homelessness.

Second, we use an instrumental variable (IV) approach to develop a more robust estimate of the causal impact of income inequality on our outcome. Specifically, a limitation of our OLS models is that they do not account for unobserved time-varying community-level characteristics that are possible confounders of the relationship between income inequality and our outcomes of interest and cannot rule out reverse causality. The IV approach helps address these potential sources of bias and reverse causality by identifying a variable (the instrument) that isolates the portion of variation in our predictor (the Gini coefficient in this case) that is exogenously determined, thus allowing for a more robust assessment of its causal impact on the outcome of interest. In constructing our instrument, we follow the approach used by Boustan and colleagues (2013). Briefly, this approach uses national changes in the income distribution as an instrument for variation in a community's actual income distribution, with the assumption that national changes in income distribution are a source of exogenous variation in a community's actual income distribution. We use this instrument to estimate the impact of income inequality on our outcomes using a two-stage least squares approach (2SLS). Our IV models also include community and year fixed effects.

Results

Table 1 (**INSERT Table 1 Summary Statistics, 2007-2018 HERE)** presents summary statistics of the variables used in the analysis, averaged across the entire study period and in the first year (2007) and last year (2018) of the study period. The average Gini coefficient, pooled across all years and communities in the sample was 0.438, and it increased, on average, within communities by about 0.004 (or just under 1 percent) over the entire study period. This

relatively small average increase in the Gini coefficient over the study period masks a period of sharper increase over the early years of the study period and then a slight decrease over the later years. It also masks substantial heterogeneity across communities: roughly a third of communities experienced a one point or more increase in the Gini coefficient, and 18% experienced a decrease in the Gini coefficient of one point or more. Similarly, whereas the total rate of homelessness decreased, on average, by 19% over the study period, this masks substantial variation, with the rate of homelessness increasing, on average, in the latter years of the study period, and nearly 30% of communities in the sample experiencing an increase in the rate of homelessness between 2007 and 2018. It is this heterogeneity that allows us to identify the impact of income inequality on rates of homelessness.

Tables 2 and 3 (**INSERT Table 2- OLS models of relationship between income inequality and homelessness AND Table 3- IV models of relationship between income inequality and homelessness HERE**) present the results of the OLS and IV models for the homelessness outcomes, respectively. The OLS estimates find that income inequality has a positive and significant relationship with the total rate of homelessness per 10,000 people, as well as with the rates of homelessness among individual, people in families, and the rates of sheltered and chronic homelessness. The OLS model finds that a one-point increase in the Gini coefficient is associated with an increase in the total rate of homelessness of about 1.1 persons per 10,000 members of the general population.[5] In a community with a population of 740,000 (roughly the average population of communities in our sample, or equivalently, approximately the 2018 population of Seattle), this would translate into an increase of about 81 people experiencing homelessness. The coefficient estimates for the models using the rate of individual and family homelessness as outcomes suggest that this overall relationship appears to be driven

more heavily by the relationship between income inequality and the rate of homelessness among individuals, as opposed to people in families. Similarly, the parameter estimate for the Gini coefficient in the model in which the rate of sheltered homelessness is the outcome suggests that the positive association between income inequality and the total rate of homelessness is driven largely by an increase in the rate of sheltered homelessness. The relatively weaker relationship between income inequality and chronic homelessness makes sense, as we would expect chronic homelessness to be less sensitive to the income and price channel mechanisms we consider as drivers of the relationship between inequality and homelessness.

The IV estimates shown in Table 3 represent our preferred specification and are consistent with the OLS estimates: the parameter estimates for our Gini coefficient measure are positive and significant in the models for the total rate of homelessness, the rates of homelessness among individuals and persons and families, and the rates of sheltered and chronic homelessness. Moreover, as with the OLS models, the parameter estimates for the Gini coefficient are relatively stronger in both the models for individuals and for sheltered homelessness. However, the magnitude of the coefficients in the IV models are substantially larger than in the OLS models. For example, in the model for the total rate of homelessness, the IV estimates indicate that a one-point increase in the Gini coefficient generates an increase in the total rate of homelessness of 7.6 persons per 10,000 members of the general population. Returning to our example of a hypothetical community with a population of 740,000, this means that a one-point increase in the Gini coefficient would result in an increase of about 562 people experiencing homelessness on a given night. Another way of contextualizing this estimate is by applying the average change in the Gini coefficient observed across all communities in our sample over the study period, which was an increase of 0.37 points. In a community with a

population of 740,000, this would translate into an additional 208 people experiencing homelessness on a given night.[6]

To examine whether our price channel and income channel mechanisms might be driving the relationship between income inequality and homelessness, Table 4 **(INSERT Table 4- Models of relationships between income inequality, proportion rent burdened low-income renters and home values HERE)** presents the results of the models in which we use the proportion of rent-burdened low-income renters and median home values as our outcomes. For the proportion of rent-burdened renters, the OLS model finds a positive, but not statistically significant, relationship between income inequality and the share of rent-burdened low-income renters. However, the IV model finds a positive and significant effect of income inequality on the proportion of rent-burdened low-income renters, with a one-point increase in the Gini coefficient leading to a 0.86 percentage point increase in the proportion of rent-burdened low-income renters. For the median home value models, the OLS specifications find a significant negative association between the Gini coefficient and home values, while the IV estimates do not find a significant association between the Gini coefficient and home values. Thus, in examining the mechanism by which income inequality generates homelessness, we find more support for the income channel than for the price channel.

Discussion

We conclude by returning to our original questions of whether local income inequality generates homelessness, and, if so, what is the mechanism through which it does. Our results are clear in showing a link between local income inequality and homelessness, and our IV estimates suggest that this relationship is causal in nature. Notably, our findings suggest that the impact of income inequality on homelessness is driven by changes in sheltered, but not unsheltered

homelessness. That we did not find evidence of a link between inequality and unsheltered homelessness could be due to measurement error as there are longstanding methodological challenges in enumerating the unsheltered population the degree of which varies across communities (Glynn and Fox 2019). Alternatively, because the rate of sheltered homelessness depends on shelter capacity, our findings may also suggest that income inequality is linked to whether a community responds to homelessness through an expansion of emergency shelter.

Our findings that income inequality has a positive impact on renter cost burden but not home values also suggests that the income channel is the more likely causal mechanism through which income inequality impacts homelessness. As our findings are not sufficiently robust to fully untangle the relationship between income inequality, renter cost burden, home values and homelessness, they offer only tentative evidence in favor of the income channel. We hope our findings offer a point of departure for future research to fully explore these mechanisms.

Of course, our investigation of the impact of income inequality on homelessness is not just an academic exercise. Our findings have implications for policy and programmatic efforts to address homelessness. We envision such efforts operating at three different levels.

First, our results suggest that broader policy efforts to reduce income inequality in their own right would have the collateral benefit of reducing homelessness. Thus, our study provides yet another brick in the wall of the argument that we should pursue aggressive policy responses to reverse decades-long trends in income inequality in the United States.

Second, our results suggest that policy and programmatic efforts intended specifically to address homelessness are needed more sorely in places where income inequality has been increasing more quickly. Our tentative evidence in favor of the income channel points to specific types of policy responses that may be warranted. In particular, while many housing policy advocates focus on slowing the growth of housing prices, this strategy alone is unlikely to be sufficient in preventing homelessness. It is important to include policies that increase the ability of low-income households to afford housing—e.g., increasing the minimum wage, public benefit levels, and the supply of Section 8 Housing Choice Vouchers. Because this problem is so specific to these high-inequality localities, it especially makes sense for these localities to adopt their own solutions rather than a one-size-fits all federal policy. For example, Los Angeles has effectively created its own section 8 program through its Department of Public Health for people experiencing homelessness and is currently providing pre-paid debit cards to low-income families who have been effective by the COVID-19 as a direct cash subsidy.

Finally, these findings point to the importance of *local*, as distinct from *national*, inequality. While most of the research literature and the public debate have focused on the macro level of analysis, this study reveals that the city level is what matters for homelessness. It is the ability of high-income residents to out-compete low-income residents that leads to the exclusion of the latter group from the housing market. The significance of the income channel, rather than the price channel, confirms this finding. Homelessness is less associated with the divergence between high-priced and low-priced cities than it is with the divergence between residents of the same city. Hopefully, this conclusion inspires future research to explore *why* some cities have higher levels and growth rates of inequality and what targeted policies can reverse these local trajectories, rather than relying on federal policies that are poorly attuned to the needs of

different local housing markets. If policymakers focus only on slowing the growth of prices in cities like New York and Los Angeles, it is likely to be insufficient to solve the homelessness crisis. They must also consider the ways in which some neighborhoods are given more access to opportunity than others, whether through segregation or locally financed education or the spatial mismatch of jobs and housing. Homelessness is more than a problem of dollars and cents. It is a symptom of a deeper challenge, fundamental to the very structure of our society: How do we share this Earth with our fellow human beings, with all that each location has to offer and all that each person deserves to experience?

---

[1] Importantly, greater exposure to the local homeless population seems to change the public's view, leading them to put more weight on structural causes (Lee, Farrell, and Link 2004).

[2] Some economists have recently proposed alternative measures of income inequality that are more easily interpretable. For example, Thomas Piketty and Emmanuel Saez (2003) have popularized the use of "top income shares," simply tracking the share of income earned by the households at the top of the income distribution, where "top" is defined as 10%, 1%, 0.1%, or even 0.01% depending on the analysis. Unfortunately, this measure is difficult to replicate, as it requires a detailed breakdown of the distribution, typically found in confidential income tax filings.

[3] This dynamic is similar to the "superstar city" effect identified by Gyourko, Mayer, and Sinai (2013), who articulate a similar concept as follows: "If the growth in latent housing demand for a particular location exceeds the growth in local housing supply due to supply inelasticity, housing rents must rise to clear the market, with lower income households crowded out by higher income households." It is the relative difference between residents, not the high demand in the city per se, that drives them out of the market. O'Flaherty (1996) offers a related explanation. In his model, the supply of low-quality housing only occurs through "filtering," the gradual deterioration of higher-quality housing originally built to meet the demand from higher-earning, middle-class households. In this model, the highest quality housing of the type for which demand is concentrated among the highest earning households remains profitable to maintain and does not filter down. Thus, the supply of low-quality housing for low-income households is not affected by the supply of luxury housing. Rising income inequality and a resulting shrinking middle class thus ultimately leads to a decrease in the supply of such low-quality housing due to fewer higher-quality units filtering down to low-income households. In turn, competition for this limited supply is more intense, driving up the price of these units and pushing more households who cannot afford these prices off the lowest rung of the housing ladder into homelessness. Our theory encompasses this possibility without relying on such stringent assumptions about filtering.

[4] We use home values instead of rents because they are a better measure of the value of the land. Our price channel posits that inequality makes high-wage cities more desirable and therefore land more valuable, which naturally translates into higher prices. Rents depend on prices and

landlord profit margins, which vary between cities and neighborhoods based on factors that may be unrelated to inequality: competition, concentration, ownership, etc. Thus, prices are the preferred metric in urban economics to capture the desirability of the location.

[5] Although the Gini coefficient is typically measured on a decimal scale (i.e. ranging from 0 to 1), we transform this measure so that it ranges from 0 to 100 in all models to facilitate interpretation of its coefficient.

[6] In the Appendix, we also estimate a one-way fixed-effects model that only controls for CoC-specific differences, u_i, and we find the same positive and significant results consistent with our preferred specification, thus confirming the robustness of our findings independent of the exact specification.

References

Boustan, Leah, Fernando Ferreira, Hernan Winkler, and Eric M. Zolt. 2013. "The Effect of Rising Income Inequality on Taxation and Public Expenditures: Evidence from U.S. Municipalities and School Districts, 1970-2000." *Review of Economics and Statistics* 95 (4): 1291–1302.

Byrne, T., E.A. Munley, J.D. Fargo, A.E. Montgomery, and D.P. Culhane. 2013. "New Perspectives on Community-Level Determinants of Homelessness." *Journal of Urban Affairs* 35 (5).

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review*. American Economic Association.

Evans, William N., Sarah Kroeger, Caroline Palmer, and Emily Pohl. 2019. "Housing and Urban Development–Veterans Affairs Supportive Housing Vouchers and Veterans' Homelessness, 2007–2017." *American Journal of Public Health* 109 (10): 1440–45.

Fischer, Will, and Barbara Sard. 2017. "Federal Housing Spending Is Poorly Matched to Need." 2017. https://www.cbpp.org/research/housing/chart-book-federal-housing-spending-is-poorly-matched-to-need.

Ganong, Peter, and Daniel Shoag. 2017. "Why Has Regional Income Convergence in the U.S.

Declined?" *Journal of Urban Economics* 102 (November): 76–90.

Gini, C. 1912. "Variabilità e Mutabilità." *Vamu*.

Glaeser, Edward, and Joseph Gyourko. 2018. "The Economic Implications of Housing Supply." In *Journal of Economic Perspectives*, 32:3–29. American Economic Association.

Glynn, Chris, and Emily B. Fox. 2019. "Dynamics of Homelessness in Urban America." *Annals of Applied Statistics* 13 (1): 573–605.

Gyourko, Joseph, Christopher Mayer, and Todd Sinai. 2013. "Superstar Cities." *American Economic Journal: Economic Policy* 5 (4): 167–99.

Hanratty, Maria. 2017. "Do Local Economic Conditions Affect Homelessness? Impact of Area Housing Market Factors, Unemployment, and Poverty on Community Homeless Rates." *Housing Policy Debate* 27 (4): 640–55.

Hoffman, Alexander Von. 2012. "History Lessons for Today's Housing Policy: The Political Processes of Making Low-Income Housing Policy." Working Paper W12-5. Cambridge, MA.

Hopper, Kim. 1991. "Homelessness Old and New: The Matter of Definition." *Housing Policy Debate* 2 (3): 755–813.

Hsieh, Chang Tai, and Enrico Moretti. 2019. "Housing Constraints and Spatial Misallocation." *American Economic Journal: Macroeconomics* 11 (2): 1–39.

Joint Center for Housing Studies of Harvard University. 2019. "The State of the Nation's Housing, 2019." Cambridge, MA. https://www.jchs.harvard.edu/sites/default/files/Harvard_JCHS_State_of_the_Nations_Housing_2019.pdf.

Lee, Barrett A., Chad R. Farrell, Bruce G. Link. 2004. "Revisiting the Contact Hypothesis: The

Case of Public Exposure to Homelessness."*American Sociological Review* 69 (1): 40–63.

Lee, Barrett A., Townsand Price-Spratlen, and James W. Kanan. 2003. "Determinants of Homelessness in Metropolitan Areas." *Journal of Urban Affairs* 25 (3): 335–56.

Lee, Barrett A., Kimberly A. Tyler, and James D. Wright. 2010. "The New Homelessness Revisited." *Annual Review of Sociology* 36 (1): 501–21.

Myers, D, G Painter, H Lee, and J Park. 2016. "Diverted Homeowners, the Rental Crisis and Foregone Household Formation." *Research Institute for Housing America. Https*.

O'Flaherty, Brendan. 1996. *Making Room : The Economics of Homelessness*. Harvard University Press.

Orlando, AW. 2013. *Letter to the One Percent*. Lulu Press. h

Piketty, T., and E. Saez. 2003. "Income Inequality in the United States, 1913-1998." *The Quarterly Journal of Economics* 118 (1): 1–41.

Rice, Douglas. 2016. "Cuts in Federal Assistance Have Exacerbated Families' Struggles to Afford Housing." 2016. https://www.cbpp.org/research/housing/chart-book-cuts-in-federal-assistance-have-exacerbated-families-struggles-to-afford.

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply." *Quarterly Journal of Economics* 125 (3): 1253–96.

Semega, Jessica, Melissa Kollar, John Creamer, and Abinash Mohanty. 2019. "Income and Poverty in the United States:2018." Washington, D.C.

Shinn. 2010. "Homelessness, Poverty and Social Exclusion in the United States." *European Journal of Homelessness* 4 (19–44).

Shinn, Marybeth, and Jill Khadduri. 2020. *In the Midst of Plenty*. *In the Midst of Plenty*. Hoboken, NJ: John Wiley & Sons.

Toro, Paul A., Carolyn J. Tompsett, Sylvie Lombardo, Pierre Philippot, Hilde Nachtergael, Benoit Galand, Natascha Schlienz, et al. 2007. “Homelessness in Europe and the United States: A Comparison of Prevalence and Public Opinion.” *Journal of Social Issues* 63 (3): 505–24.

U.S. Department of Housing and Urban Development. 2020. “The 2019 Annual Homeless Assessment Report To Congress: Part 1-PIT Estimates of Homelessness in the U.S.” Washington, D.C.

Wolch, Jennifer R., Michael Dear, and Andrea Akita. 1988. “Explaining Homelessness.” *Journal of the American Planning Association* 54 (4): 443–53.

Wood, Gavin, Deb Batterham, Melek Cigdem, and Shelley Mallett. 2015. “The Structural Drivers of Homelessness in Australia, 2001-11.” Melbourne, Australia.

Table 1 Summary Statistics, 2007-2018

	All years		**2007**	**2018**	**% change, 2007-2018**
	Mean	SD	Mean	Mean	
Gini coefficient	.4382	.278	.4298	.4335	0.86
% rent burdened low-income renters	90.04	4.41	89.11	89.68	0.64
Rate of homelessness per 10,000 residents					
Total	24.98	20.92	28.4	22.98	-19.08
Individuals	16.23	14.03	18.79	15.47	-17.67
People in families	8.75	10.56	9.61	7.51	-21.85
Sheltered	16.49	16.46	17.27	15.05	-12.85
Unsheltered	8.5	12.39	11.13	7.93	-28.75
Chronically homeless	4.24	5.16	5.49	3.76	-31.51
Median home value, $	302411.5	170618.4	331556.6	362894.3	9.45
Median household income, $	60406.72	15270.42	56749.91	70240.79	23.77
Number of PSH beds per 10,000 residents	12.63	12.3	8.73	15.95	82.7
Median rent, 2 bedroom unit	11.75	3.62	9.98	13.82	38.48
Poverty rate	14.11	4.69	12.25	12.61	2.94
Real GDP (2012 dollars), in $ billions	109.46	145.18	105.16	123.14	17.1
Total population	742851	1036722	711963	772222	3.95

Table 2- OLS models of relationship between income inequality and homelessness

	Rate of homelessness per 10,000 population					
	Total homeless	Homeless individuals	Homeless people in families	Sheltered homeless	Unsheltered homeless	Chronic homeless
Gini coefficient	1.119***	0.649***	0.470***	0.969***	0.150	0.242**
	(0.221)	(0.165)	(0.114)	(0.139)	(0.149)	(0.086)
Log(Median household income)	28.678***	17.459***	11.219*	22.959***	5.719	2.370
	(7.144)	(4.346)	(5.340)	(4.334)	(5.599)	(2.821)
PSH beds per 10,000 population	-0.225**	-0.257***	0.032	0.017	-0.243**	-0.154***
	(0.079)	(0.053)	(0.052)	(0.039)	(0.067)	(0.039)
Median rent, 2 bedroom apartment (in $100)	0.603*	0.498**	0.105	0.440**	0.162	0.124
	(0.216)	(0.158)	(0.099)	(0.147)	(0.157)	(0.072)
Log(Median home value)	-2.873	-2.881	0.007	-3.855**	0.981	-1.084
	(2.487)	(1.859)	(0.992)	(1.310)	(2.076)	(0.929)
Poverty rate	-0.304	-0.279*	-0.025	-0.083	-0.221	-0.151
	(0.164)	(0.135)	(0.087)	(0.090)	(0.142)	(0.077)
Real GDP,2012 dollars ($ billions)	0.128***	0.085***	0.043***	0.017**	0.111***	0.033***
	(0.010)	(0.007)	(0.006)	(0.006)	(0.008)	(0.004)
Constant	-356.624***	-193.444***	-163.180**	-242.878***	-113.746	-26.302
	(82.095)	(51.597)	(59.867)	(43.905)	(64.510)	(32.963)
Observations	2,857	2,857	2,857	2,857	2,857	2,857

Note: *p<.05; **p<.01;***p<0.001

All models include CoC and year fixed effects

Robust standard errors clustered at CoC level shown in parentheses

Table 3- IV models of relationship between income inequality and homelessness

	Rate of homelessness per 10,000 population					
	Total homeless	Homeless individuals	Homeless people in families	Sheltered homeless	Unsheltered homeless	Chronic homeless
Gini coefficient	7.568***	4.263***	3.305***	7.237***	0.331	1.316**
	(1.457)	(1.088)	(0.777)	(1.154)	(0.713)	(0.385)
Log(Median household income)	90.696***	52.213***	38.483***	83.236***	7.460	12.694*
	(17.582)	(13.670)	(9.856)	(12.460)	(9.687)	(5.847)
PSH beds per 10,000 population	-0.202*	-0.245***	0.042	0.040	-0.242**	-0.150***
	(0.078)	(0.046)	(0.056)	(0.047)	(0.066)	(0.036)
Median rent, 2 bedroom apartment (in $100)	1.146***	0.803***	0.344**	0.969***	0.178	0.214**
	(0.258)	(0.176)	(0.121)	(0.194)	(0.156)	(0.073)
Log(Median home value)	-0.746	-1.688	0.943	-1.787	1.041	-0.730
	(2.785)	(2.021)	(1.161)	(1.873)	(2.088)	(0.912)
Poverty rate	-1.410***	-0.899***	-0.511**	-1.158***	-0.252	-0.335***
	(0.268)	(0.156)	(0.157)	(0.218)	(0.149)	(0.067)
Real GDP,2012 dollars ($ billions)	0.131***	0.087***	0.044***	0.020*	0.111***	0.033***
	(0.012)	(0.008)	(0.007)	(0.009)	(0.008)	(0.004)
Constant	-1,344.005***	-746.753***	-597.252***	-1,202.545***	-141.459	-190.679*
	(259.321)	(204.259)	(140.832)	(186.859)	(138.478)	(80.925)
Observations	2,857	2,857	2,857	2,857	2,857	2,857

Note: *p<.05; **p<.01;***p<0.001

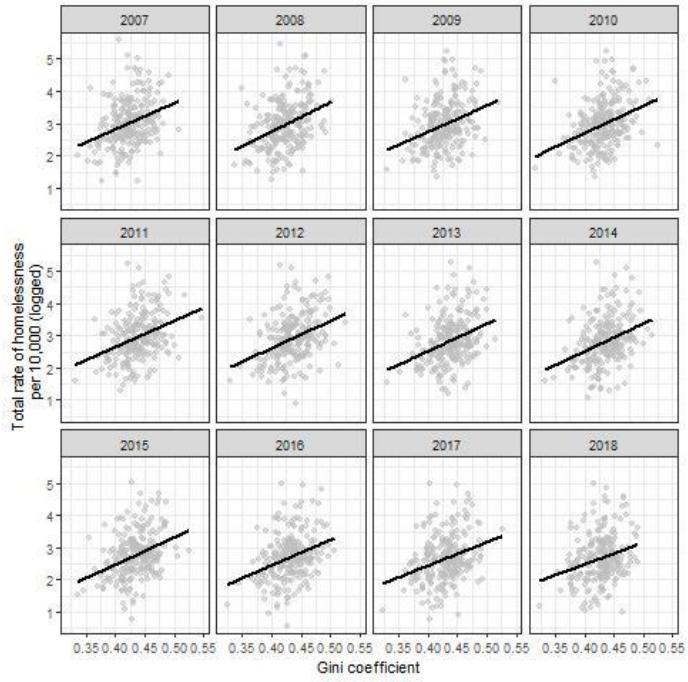
All models include CoC and year fixed effects
Robust standard errors clustered at CoC level shown in parentheses

Table 4- Models of relationships between income inequality, proportion rent burdened low-income renters and home values

	% rent burdened low-income renters		Log(Median home value)	
	OLS Estimates	IV Estimates	OLS Estimates	IV Estimates
Gini coefficient	0.089	0.861*	-0.004*	-0.006
	(0.059)	(0.358)	(0.002)	(0.018)
Log(Median household income)	-2.671	4.749	0.801***	0.789***
	(1.601)	(3.722)	(0.064)	(0.176)
PSH beds per 10,000 population	-0.007	-0.004	-0.0003	-0.0003
	(0.010)	(0.009)	(0.001)	(0.001)
Median rent, 2 bedroom apartment (in $100)	-0.121*	-0.056	-0.006**	-0.006*
	(0.049)	(0.054)	(0.002)	(0.002)
Log(Median home value)	2.104***	2.358***		
	(0.425)	(0.477)		
Poverty rate	0.171***	0.039	-0.014***	-0.014***
	(0.042)	(0.075)	(0.002)	(0.004)
Real GDP,2012 dollars ($ billions)	0.007***	0.008***	0.001***	0.001***
	(0.002)	(0.002)	(0.0001)	(0.0001)
Constant	84.231***	-33.904	4.116***	4.311
	(17.938)	(57.401)	(0.717)	(2.683)
Observations	2,857	2,857	2,857	2,857

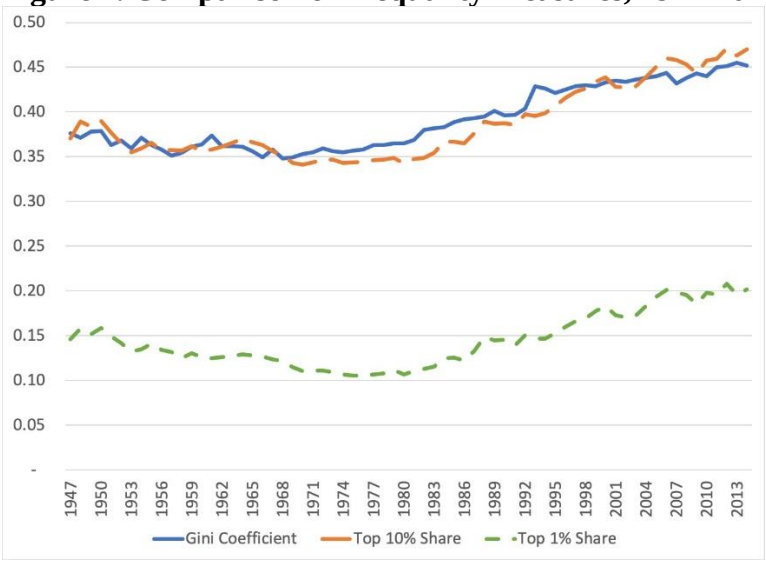
*p<.05; **p<.01;***p<0.001
All models include CoC and year fixed effects
Robust standard errors clustered at CoC level shown in parentheses

Figure 1–Between Community Relationship of Income Inequality and Total Rate of Homelessness, 2007-2018



Note: Each dot represents one Continuum of Care (CoC) and each line represents bivariate regression of the total rate of homelessness in each CoC on the Gini coefficient. The total rate of homelessness shown in natural log transformation to more clearly visualize relationship. Because CoCs have irregular geographic boundaries, we construct CoC-level measures of the Gini coefficients using a modified version of the process described by Byrne and colleagues (2013) to match CoCs to counties and take population weighted averages of county-level Gini coefficient measures where necessary.

Figure 2. Comparison of Inequality Measures, 1947-2014



*Notes: Gini coefficient calculated by U.S. Census Bureau, retrieved from Federal Reserve Bank of St. Louis. Top 10% share and top 1% share calculated by Piketty, Saez, and Zucman (2018)***.**

APPENDIX

Method Used to Construct Gini Coefficients

To measure income inequality, we mirror the approach used by Boustan and colleagues (2013) to calculate county-level Gini coefficients on an annual basis for each year from 2007 to 2018. This approach entails combining information obtained from the Census Bureau's American Community Survey (ACS) 1-Year estimates Detailed Tables on the number on the number of households in each of 16 different income categories[1] with Census 1-year ACS microdata, which was obtained from IPUMS USA.[2] Specifically, we use the ACS microdata to estimate the annual state-specific median income for persons in each of these 16 income categories.[3] We then construct the income distribution for each county for each year by assigning each household in the 16 income categories its state-specific median income. This, in turn, allows us to calculate annual county-level Gini coefficients. Importantly, because we do not observe the full income distribution and thus must use the assumption that all households in a particular income category have the estimated state-level median income for all households in that income category, our income distribution is only an approximation of the true income distribution, and thus our Gini coefficient is likewise only approximate. The Census Bureau does report county-level Gini coefficients as part of the 1-Year ACS estimates Detailed tables that are presumably on more fine-grained information about the true income distribution. Nonetheless, we chose to use our constructed Gini coefficient measure for all analyses both because our measure is highly correlated with the Census Bureau's constructed measure (r= 0.83) and because our instrumental variable strategy (which is also adopted from Boustan and colleagues, 2013) relies on this constructed Gini coefficient.

Process Used to Match Counties to Department of Housing and Urban Development (HUD) Continuums of Care (CoCs) and Construct CoC-level Measures

As we note in the main text, we merge our county-level income inequality and other measures with the HUD data to construct a CoC-level dataset. Specifically, in addition to our county-level income inequality measure described above, we obtain several additional annual county-level measures from the ACS 1-Year estimate Detailed Tables: the proportion of low-income renter households (defined as households with income <$20,000) with rent burden (defined as paying

[1] These income categories are: <$10,000, $10,000; $10,000 to $14,999; $15,000 to $19,999; $20,000 to $24,999; $25,000 to $29,999; $30,000 to $34,999; $35,000 to $39,999; $40,000 to $44,999; $45,000 to $49,999; $50,000 to $59,999; $60,000 to $74,999; $75,000 to $99,999; $100,000 to $124,999; $125,000 to $149,999; $150,000 to $199,999; and $200,000 or more

[2] Steven Ruggles, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas and Matthew Sobek. IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN: IPUMS, 2020. https://doi.org/10.18128/D010.V10.0

[3] Publicly available ACS microdata do not include county-level geographic information, and we rely on state-level median income estimates in constructing county-level income distributions. This will bias the Gini coefficients in all counties towards the state-level mean Gini coefficient. In some cases (where the true income distribution in a county is more unequal than the state's income distribution) this will downwardly bias the Gini coefficient for a given county, and in some cases (where the true income distribution in a county is more equal than the state's income distribution) this will bias the Gini coefficient in an upwards direction. However, the degree of this bias is likely small, as we note that our constructed Gini coefficient is strongly correlated with a county-level Gini coefficient measure made publicly available by the Census Bureau.

>30% of income on rent); median value of owner-occupied housing units (hereafter referred to as median home value); median household income; median rent for a two-bedroom apartment; poverty rate; and total population. We supplement these variables with a measure of county-level Gross Domestic Product (in 2012 dollars) obtained from the U.S. Bureau of Economic Analysis.

We then merge these county-level data with the HUD data on homelessness. The merger of HUD data on homelessness with county-level inequality and other measures is complicated by two factors. First, PIT counts are conducted and reported at the Continuum of Care (CoC) level, which is the geographic unit at which federal homeless assistance grants are awarded. While some CoCs have boundaries that are coterminous with a single county, CoC boundaries do not always align neatly with county boundaries. Thus, we convert all of our county-level measures into CoC-level measures using a modified version of the process described by Byrne and colleagues (2013) to match CoCs to counties. Second, as an added complication, 1-Year ACS Detailed Table estimates are only available for counties with a population of 65,000 or greater. Thus, to ensure that we could accurately create annual CoC-level Gini coefficients from our county-level data, we initially restrict our sample to the 253 CoCs that were active in 2018 and that were made up entirely of one (or more) of the 798 counties that were large enough to be included in the 1-Year ACS estimates for the entire duration of the study period. Based on the HUD PIT count data, these 253 CoCs account for 77% of the total number of persons experiencing homelessness nationwide on a given night in 2018.

To construct CoC-level measures from county-level data, we first matched counties to CoCs based on the boundaries of the 398 CoCs that were active in 2018. To do so, we used geospatial mapping software to match the centroids of Census tracts (i.e. points representing the geographic centers of Census tracts) to CoCs. Based on the counties in which tracts were located, we then created a county-to-CoC crosswalk, with a county matching to a CoC if the centroid of any tract in that county was located within the boundaries of a CoC. This resulted in several types of CoC-to-county relationships including: 1) one CoC fully encompassing a single county (i.e. a one-to-one match); 2) one CoC fully encompassing multiple counties (i.e. a one-to-many match); 3) One county fully encompassing multiple CoCs (i.e. a many-to-one-match); and 4) matches in which part of a county is located in multiple CoCs and at least one of those CoCs includes parts of multiple counties (i.e. a many-to-many match). Construction of CoC-level measures from county-level data thus varied as a function of CoC-county relationship type. In instances when a single CoC mapped onto a single county, this process was straightforward with no manipulation of county-level data required. In instances when a single CoC mapped onto multiple counties (i.e. a one-to-many match) we create CoC-level measures by either calculating the sum (e.g. in the case of total population or population in poverty to calculate poverty rate) or the population weighted mean (e.g. in the case of median household income or median rent), as appropriate, of the county-level measures. Cases where multiple CoCs were located in a single county (i.e. a many-to-one match) did not require any transformation of county-level measures, but does require the aggregation of PIT homeless counts and PSH bed inventory in all CoCs in a given county. This results in the collapse of multiple CoCs into a single new "pseudo-CoC." In the case of "many to many" matches (i.e. the fourth type of relationship described above) we used two different approaches. In some cases, we used the share of a county's population (calculated based on which Census tracts in a given county were located in each CoC and using tract-level

population from the 2013-2017 ACS 5-year estimates) to assign a county to the single CoC in which the majority of its population was located. In other cases, this entailed manually combining multiple CoCs into a single pseudo-CoC. This occurred when the vast majority of a county's population was housed within one or two CoCs (as was the case with the Atlanta and Fulton County CoCs), with a very small share in one or more other CoCs. A total of 253 of the 398 CoCs that were active in 2018 were comprised entirely of counties that were large enough to be included in the ACS 1-Year Estimate Detailed tables for the entire period from 2007-2018. After merging CoCs that had either many-to-one CoC-to-county relationships or many-to-many relationships, the effective number of communities was reduced to 239. A total of 236 of these CoCs have complete data for the entire 12 year study period. Three CoCs have incomplete data for one or more years either because PIT data were not included in HUD's data for a given year, or because these CoCs did not exist for the entire time series. Thus, we do not have a completely balanced panel, but it is only slightly unbalanced.

It is important to note that, while this panel of 239 communities captures the majority of the homeless population in the United States, by definition, it excludes CoCs comprised of smaller communities and for which it is not possible to construct annual measures of income inequality. Thus, we conduct supplemental analysis using data from all CoCs for 2009, 2014, and 2018. Results from that analysis are similar to the main results and are provided below in the Appendix.

Modeling Approach

We use our data to estimate several models exploring the relationship between income inequality and homelessness, all of which take the following general form:

$$y_{it} = \beta_1 Gini_{it} + \alpha X_{it} + u_i + \lambda_t + \epsilon_{it} \,,$$

where *i* indexes a community of interest and *t* indexes the year, and thus y_{it} and $Gini_{it}$ represent the outcome of interest and Gini coefficient in CoC *i* in year *t*; X_{it} is a vector of time-varying control variables including median household income, median rent for a two-bedroom apartment, median home value, poverty rate, real GDP, and the number of PSH beds per 10,000 people; u_i and λ_t represent CoC and year fixed-effects, respectively; and ϵ_{it} is a random error term. This two-way fixed-effects model identifies the relationship between a city that experiences higher deviations from its typical inequality level, compared to the deviations in other cities, and that city's deviations in the outcome variable, compared to deviations in other cities, where deviations are all defined relative to each city's average level over time. In this modeling framework, β_1 is the coefficient of primary interest and represents the relationship between how changes in income inequality, relative to a given city's average level, differ across cities and how changes in the outcome, y_{it}, relative to a given city's average level, differ across cities (Kropko & Kubinec, 2020). If it is positive and significant, as we hypothesize, it indicates that larger increases in inequality result in larger increases in the outcome variable, without the confounding influence either of CoC-specific differences (because each CoC is compared to its own average over time) or national trends over time (because each CoC is compared to other CoCs). Critics have argued that such two-way fixed effects models are not substantively meaningful or statistically valid (Kropko & Kubinec, 2020). We thus also estimate one-way fixed effects

models which have the more straightforward interpretation of representing the relationship between changes in the Gini coefficient and changes in rates of homelessness over time *within* the same CoC. The results of these models are presented later in the Appendix.

We first estimate a series of linear regression models using this specification and with the following set of primary outcome variables: the total rate of homelessness per 10,000 members of the population, the proportion of rent-burdened low-income renters, and median home value (logged). The first two outcomes models allow us to test the "income channel" mechanism by which inequality might drive homelessness, while the third outcome allows us to test the "price channel" mechanism. To further explore the relationship between income inequality and homelessness, we estimate several additional models in which we use homeless individuals, persons in families, sheltered homeless persons, unsheltered homeless persons, and chronically homeless persons (all measured as a rate per 10,000 members of the population) as outcomes. In following prior research (Evans et al., 2019; Hanratty, 2017) in all models, we weight all observations by the contemporaneous CoC population and use robust standard errors, clustered at the CoC-level.[4] Finally, although the Gini coefficient is typically measured on a decimal scale (i.e. ranging from 0 to 1), we transform this measure so that it ranges from 0 to 100 in all models to facilitate interpretation of this coefficient.

While these two-way fixed effects models controls for time-invariant differences between communities, time-varying factors that affect all communities, as well as a set of time-varying community level characteristics, there may be unobserved time-varying community-level characteristics that are confounders of the relationship between income inequality and our outcomes of interest. For instance, income inequality may contribute to policy changes at the local level, such as the introduction of right-to-counsel ordinances in housing court or the passage of locally funded initiatives to fund homeless services (such as Measure HHH in Los Angeles), that would downwardly bias the relationship between income inequality and homelessness. Likewise, income inequality may be confounded with unobserved drivers in housing costs that could bias the relationship between inequality and homelessness in an upward direction. These models also cannot rule out reverse causality, which may also be an issue: homelessness could be a driver of increases in income inequality because more people experiencing homelessness (assuming people experiencing homelessness are universally people with extremely low-incomes) means a downward skewing of the income distribution, assuming the number of higher earners stays fixed or increases. Increased homelessness could also cause reductions in inequality if, for example, a highly visible homeless problem causes outmigration of higher income earners, thus flattening the income distribution.

Thus, to address these potential sources of bias, and to develop a more robust estimate of the causal effect of income inequality and our outcomes of interest, we also employ an instrumental variable (IV) approach. This approach addresses the above noted concerns about bias and reverse causality by identifying a variable (the instrument) that is both correlated with the primary predictor of interest but uncorrelated with the error term in the model estimating the relationship between this predictor and the outcome. Put differently, an IV approach isolates the

[4] We use the R clubSandwich package (Pustejovsky, 2020) to calculate these standard errors and p-values. Because the outcome is a standardized rate (i.e. number homeless per 10,000 people), in calculating these standard errors, we assume that the error variances are inversely proportional to population size.

portion of variation in our predictor that is exogenously determined, thus allowing for a more robust assessment of its causal impact on the outcome of interest.

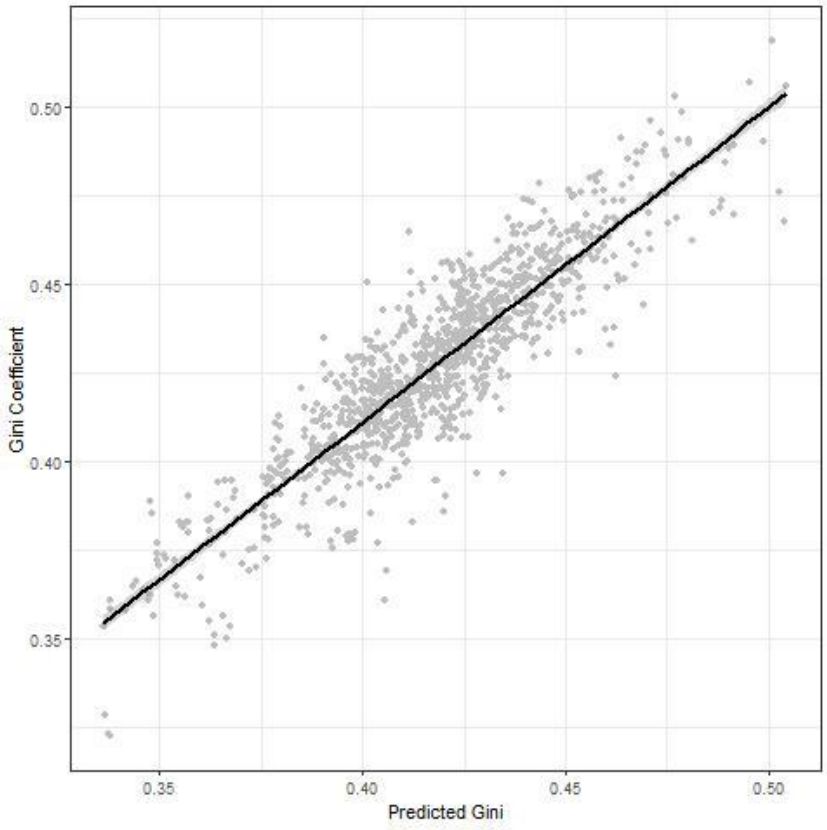
In constructing our instrument, we follow the approach used by Boustan and colleagues (2013). Specifically, Boustan and colleagues (2013) approach involved constructing predicted income distributions for each community in their sample for each time period, based on a community's initial income distribution and national patterns of income growth. They then calculate the Gini coefficient based on this predicted income distribution and used this measure as an instrument for the actual Gini coefficient. In the context of the present study, constructing this instrument entailed first using data from the 2006 ACS 1-year estimates to tabulate the number of households in each of the above described 16 income categories in each county. Doing so provided us with an estimate of the income distribution in each county for an initial year (2006) that preceded our observation period. Using ACS microdata for the years 2007-2018, we then allowed the income level of households in these 2006 income categories in each county to change over time in accordance with the actual *national* level change in income experienced by households in that income category.[5] This resulted in a predicted income distribution for each county and each year. We then used the Gini coefficient calculated based on this predicted income distribution as our instrument for the actual observed Gini coefficient, which we describe above. In short, this instrument allows us to identify the portion of the variation in local changes in inequality that are driven by national trends. As we detail below, a key assumption is that these national trends, by themselves, are not expected to have any impact on the rate of homelessness at the local level.

This instrument will provide an asymptotically unbiased estimate of the relationship between income inequality and our outcomes when two conditions are satisfied. First, the instrument must be correlated with the actual, observed Gini coefficient (this is also known as the relevance condition) and second, the instrument only affects our outcomes of interest through the observed Gini coefficient (this is also known as the exclusion restriction). To assess the first condition, Figure A-1 plots the relationship between our instrument (the Gini coefficent based on the predicted income distribution) and our actual observed Gini coefficient, that we constructed using the process described above. There is a clear and strong positive relationship between these two variables (r = 0.83). We conducted a more formal test of this assumption via a first-stage model in which we regressed the observed Gini coefficient on our instrument and in which we also controlled for our full set of time-varying covariates, and that also included CoC and year fixed effects. The coefficient for the instrument in this model was 1.96 ($p < .01$) and the first-stage F-statistic was 11.05 ($p < .01$), which is above the conventional threshold for a strong instrument. The second condition is satisfied by construction. Because the instrument is based on an income distribution that is frozen in the in the base year (2006), it cannot be affected by community characteristics in later years (e.g. it cannot be affected by households moving into/out of an area that might be correlated with the rate of homelessness and also result in a change in the income distribution). Moreover, changes in the national income distribution should, by themselves, not be expected to have any impact on the *local* rate of homelessness, which is the assumption we rely on to satisfy the exclusion criterion. We use this instrument to estimate the impact of income

[5] Boustan and colleagues convert the end points of their income bins into corresponding percentiles of the income distribution, and calculate changes based on corresponding changes in these percentiles. We leave the income bins in their raw values when constructing our instrument.

inequality on our outcomes using a two-stage least squares approach (2SLS). All analysis were conducted using the R environment for statistical computing (R Core Development Team, 2015) version 3.6.2; 2SLS models were estimated using the ivreg function from the AER package (Kleiber & Zeileis, 2020) and cluster robust standard errors were calculated via the clubSandwich package (Pustejovsky, 2020).

Figure A-1- Relationship between Observed and Predicted Gini Coefficients



IV Estimates of Impact of Income Inequality on Homelessness Using One-Way Fixed Effects Model

Table A1 presents the results of the IV models estimating the relationship between income inequality and homelessness. This model is identical to the specification shown in Table 3 of the main analysis, except it is a one-way fixed effect model, in which we include a CoC-level fixed-effect, but not a year fixed-effect term. The models thus represent the relationship between changes in the Gini coefficient and changes in rates of homelessness over time *within* the same CoC. The results are consistent with the main analysis, although the magnitude of the coefficients are smaller.

Table 4- IV models of relationship between income inequality and homelessness

	Rate of homelessness per 10,000 population					
	Total homeless	Homeless individuals	Homeless people in families	Sheltered homeless	Unsheltered homeless	Chronic homeless
Gini coefficient	1.636**	0.934**	0.701*	1.518***	0.117	0.218
	(0.511)	(0.288)	(0.343)	(0.318)	(0.336)	(0.162)
Log(Median household income)	-31.558***	-18.255***	-13.303***	-11.545***	-20.012***	-8.422***
	(4.270)	(2.771)	(2.605)	(2.693)	(3.162)	(1.183)
PSH beds per 10,000 population	-0.377***	-0.347***	-0.030	-0.076	-0.301***	-0.182***
	(0.099)	(0.063)	(0.056)	(0.046)	(0.071)	(0.040)
Median rent, 2 bedroom apartment (in $100)	0.802***	0.637***	0.165	0.515***	0.287	0.193**
	(0.207)	(0.148)	(0.097)	(0.108)	(0.152)	(0.065)
Log(Median home value)	0.026	-0.570	0.596	-1.717	1.743	-0.068
	(2.173)	(1.610)	(0.951)	(1.167)	(1.969)	(0.704)
Poverty rate	-1.422***	-0.991***	-0.431**	-0.794***	-0.628***	-0.374***
	(0.237)	(0.151)	(0.127)	(0.130)	(0.169)	(0.091)
Real GDP,2012 dollars ($ billions)	0.123***	0.081***	0.042***	0.015**	0.108***	0.029***

	(0.009)	(0.006)	(0.006)	(0.005)	(0.007)	(0.003)
Constant	255.664***	163.887***	91.778**	91.215**	164.450***	83.639***
	(45.168)	(27.311)	(30.257)	(27.528)	(34.368)	(13.952)
Observations	2,857	2,857	2,857	2,857	2,857	2,857

Note: *p<.05; **p<.01;***p<0.001

Results of Analysis Based on Full Set of CoCs Active in 2018 and PIT Count Data from 2009, 2014 and 2018

We conducted supplemental analysis based on the full set of 398 CoCs that were active in 2018. As noted above, 1-Year ACS estimate Summary File tables are only available for counties above a certain population threshold. This means that we cannot use all CoCs in our main analysis. However, data needed for our analysis are available for all counties in the United States from the 5-Year ACS estimates, which pool data from across rolling 5-year periods. The first five year period for which these data are available is 2005-2009, and then they are updated annually (i.e. based on data from 2006-2010, 2007-2011, etc). The most recent available data are from the 2014-2018 5-year estimates. Because data are pooled across 5-year periods, it would be possible to use each of these 5-year datasets in sequence to create full panel data for all communities for the period from 2009-2018, with the end year of each set of 5-year estimates standing in as a proxy for each year from 2009 to 2018. However, because the 5-year estimates pool data from across 5-year periods, doing so would not accurately capture year-to-year changes in income inequality. We thus constructed a panel dataset in which we sought to use data from sequential non-overlapping 5-year periods. This was not fully possible as data from only two non-overlapping 5-year periods are currently available. We thus, constructed a dataset in which we paired 2005-2009 ACS 5-year estimates with 2009 PIT homeless count data; 2010-2014 ACS 5-year estimates with 2014 PIT count data and 2014-2018 ACS 5-year estimates with 2018 PIT count data. We used the same methods described above to create a CoC-level dataset based on county-level measures. Doing so again, resulted in the merger of several CoCs such that the original 398 CoCs that were active in 2018 became a panel dataset of 369 CoCs for which we had a maximum of 3 time periods for each community; this panel was only slightly unbalanced with 3 CoCs having data for less than 3 years.

We then re-estimated our full set of main models on this dataset. The results of this analysis is shown below in Tables A2, A3, A4 and A5, which correspond to Tables 2, 3, 4 and 5 in the main analysis.

Table A2- Model of relationship between income inequality and proportion rent burdened low-income renters

	OLS Estimates	IV Estimates
Gini coefficient	0.381***	0.623*
	(0.097)	(0.247)
Log(Median household income)	-0.696	1.105
	(2.326)	(2.734)
PSH beds per 10,000 population	-0.049***	-0.050***
	(0.011)	(0.012)
Median rent, 2 bedroom apartment (in $100)	-0.078	-0.056
	(0.052)	(0.055)
Log(Median home value)	4.924***	5.035***
	(0.556)	(0.599)
Poverty rate	0.199**	0.164*
	(0.067)	(0.075)
Real GDP,2012 dollars ($ billions)	-0.008***	-0.007***
	(0.002)	(0.002)
Constant	21.317	-10.861
Gini coefficient	(27.124)	(40.158)
Observations	1,094	1,094

*p<.05; **p<.01;***p<0.001

All models include CoC and year fixed effects

Robust standard errors clustered at CoC level shown in parenthesis

Table A3- OLS models of relationship between income inequality and homelessness

	Rate of homelessness per 10,000 population					
	Total homeless	Homeless individuals	Homeless people in families	Sheltered homeless	Unsheltered homeless	Chronic homeless
Gini coefficient	2.877***	2.073***	0.805*	2.356***	0.521	0.828***
	(0.491)	(0.308)	(0.310)	(0.273)	(0.381)	(0.151)
Log(Median household income)	32.028	36.809*	-4.781	20.876*	11.152	17.051
	(21.032)	(17.370)	(8.898)	(8.169)	(18.975)	(10.350)
PSH beds per 10,000 population	-0.216	-0.252	0.035	-0.079	-0.137	-0.188*
	(0.152)	(0.132)	(0.051)	(0.039)	(0.140)	(0.086)
Median rent, 2 bedroom apartment (in $100)	0.723*	0.473*	0.251	0.787***	-0.064	0.062
	(0.271)	(0.197)	(0.139)	(0.169)	(0.198)	(0.114)
Log(Median home value)	1.561	-1.635	3.195	1.968	-0.407	-3.185
	(5.186)	(4.369)	(2.119)	(1.606)	(4.831)	(2.563)
Poverty rate	-1.157**	-0.837**	-0.320	-0.409	-0.748*	-0.589***
	(0.371)	(0.291)	(0.220)	(0.232)	(0.301)	(0.142)
Real GDP,2012 dollars ($ billions)	0.116***	0.080***	0.036***	-0.001	0.117***	0.032***
	(0.009)	(0.006)	(0.004)	(0.005)	(0.007)	(0.003)
Constant	-515.937**	-476.209**	-39.728	-347.807***	-168.130	-179.391*
	(185.240)	(152.679)	(81.059)	(86.203)	(161.064)	(87.023)
Observations	1,094	1,094	1,094	1,094	1,094	1,094

Note: *p<.05; **p<.01;***p<0.001

All models include CoC and year fixed effects

Robust standard errors clustered at CoC level shown in parenthesis

Table A4- IV models of relationship between income inequality and homelessness

	Rate of homelessness per 10,000 population					
	Total homeless	Homeless individuals	Homeless people in families	Sheltered homeless	Unsheltered homeless	Chronic homeless
Gini coefficient	5.776	2.551	3.225**	8.004***	-2.228	-0.378
	(3.584)	(3.223)	(0.930)	(0.947)	(3.498)	(2.069)
Log(Median household income)	53.568**	40.365**	13.203	62.841***	-9.274	8.091
	(19.543)	(13.829)	(11.586)	(11.879)	(16.528)	(7.681)
PSH beds per 10,000 population	-0.230	-0.254*	0.024	-0.107*	-0.123	-0.182*
	(0.138)	(0.119)	(0.056)	(0.046)	(0.129)	(0.079)
Median rent, 2 bedroom apartment (in $100)	0.986*	0.516	0.470*	1.299***	-0.313	-0.048
	(0.426)	(0.352)	(0.172)	(0.203)	(0.374)	(0.227)
Log(Median home value)	2.884	-1.416	4.301*	4.547*	-1.663	-3.736
	(6.546)	(5.783)	(2.120)	(1.972)	(6.303)	(3.567)
Poverty rate	-1.575**	-0.906*	-0.670**	-1.224***	-0.352	-0.415
	(0.521)	(0.451)	(0.224)	(0.292)	(0.475)	(0.272)
Real GDP,2012 dollars ($ billions)	0.124***	0.081***	0.043***	0.014*	0.109***	0.029***
	(0.011)	(0.008)	(0.005)	(0.006)	(0.008)	(0.005)
Constant	-900.767*	-539.742	-361.025*	-1,097.554***	196.787	-19.305
	(398.714)	(343.472)	(142.330)	(157.089)	(373.350)	(214.143)
Observations	1,094	1,094	1,094	1,094	1,094	1,094

Note: *p<.05; **p<.01;***p<0.001

All models include CoC and year fixed effects

Robust standard errors clustered at CoC level shown in parenthesis

Table A5- Model of relationship between income inequality and home values

	Log(Median home value)	
	OLS Estimates	IV Estimates
Gini coefficient	-0.012*	0.065**
	(0.005)	(0.020)
Log(Median household income)	1.580***	2.209***
	(0.136)	(0.176)
PSH beds per 10,000 population	-0.001	-0.001
	(0.001)	(0.001)
Median rent, 2 bedroom apartment (in $100)	-0.010***	-0.004
	(0.002)	(0.003)
Poverty rate	-0.031***	-0.043***
	(0.004)	(0.005)
Real GDP,2012 dollars ($ billions)	-0.0002**	0.00000
	(0.0001)	(0.0001)
Constant	-2.794	-13.145***
	(1.586)	(2.622)
Observations	1,094	1,094

Note: *p<.05; **p<.01;***p<0.001

All models include CoC and year fixed effects
Robust standard errors clustered at CoC level shown in parenthesis