The Role of People vs. Places in Individual Carbon Emissions

Eva Lyubich *

September 9, 2024

Abstract

There is substantial spatial heterogeneity in household carbon emissions. I leverage movers in two decades of administrative Decennial Census and American Community Survey data to estimate place effects – the amount by which carbon emissions change for the same household living in different places – for almost 1,000 cities and roughly 61,500 neighborhoods across the US. I estimate that place effects account for 14-23 percent of overall heterogeneity. A change in neighborhood-level place effects from one standard deviation above the mean to one below would reduce household carbon emissions from residential energy and commuting by about 40 percent. *JEL Codes:* H41, Q40, R20.

*U.S. Census Bureau. eva.lyubich@census.gov. Any opinions and conclusions expressed herein are those of the author(s) and do not represent the views of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this data product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release (Release authorization numbers: CBDRB-FY24-CES014-016, CBDRB-FY24-CES014-017, and CBDRB-FY24-CES005-017). I am grateful to Reed Walker, Pat Kline, Emmanuel Saez, and Joe Shapiro for their guidance and support throughout this work. I also thank the editor and three anonymous referees for their constructive feedback. This article has benefited tremendously from comments and suggestions from Alan Auerbach, Matilde Bombardini, Stephanie Bonds, Severin Borenstein, David Card, Lucas Davis, Karl Dunkle Werner, Ben Faber, Meredith Fowlie, Cecile Gaubert, Kendall Houghton, Hilary Hoynes, Jon Kolstad, Enrico Moretti, Andrés Rodríguez-Clare, Raffaele Saggio, Jim Sallee, Elif Tasar, Danny Yagan, Katherine Wagner, Randall Walsh, Chris Walters, and seminar participants at the University of California, Berkeley. I am grateful to the National Science Foundation Graduate Research Fellowship Program (DGE 1752814), the Berkeley Opportunity Lab and Smith Richardson Foundation, and Washington Center for Equitable Growth for financial support.

1 Introduction

Increased carbon and other greenhouse gas emissions since the onset of the industrial revolution have caused global average temperatures to rise by over 1°C (1.8°F) relative to preindustrial levels (NASA 2020). In 2015, the United States signed the Paris Accord, a global agreement aimed at mitigating the potential damages from climate change by limiting overall warming to below 2°C. In search of opportunities for decarbonization, researchers and policymakers have pointed to substantial spatial heterogeneity in household carbon emissions, suggesting that perhaps higher-emissions places could adopt features of lower-emissions places, such as density and high-quality public transportation infrastructure, in order to lower household carbon emissions (e.g. Jones and Kammen 2014; International Energy Agency 2021; Pomponi et al. 2021; Wagner 2021).

However, differences in mean carbon emissions across places reflect a combination of local amenities, household characteristics, and taste-based sorting. The relative contributions of these pieces is a central determinant of whether place-based interventions that change urban form would lead to significant reductions in carbon emissions. For instance, if places with large single-family homes and car-oriented transportation infrastructure are high-emissions because the people who live there dislike multi-family homes and public transit, then deregulating zoning or building new rail lines would have little impact on household emissions. Conversely, if the lack of denser housing and transit options is a constraint on household choices, rather than a reflection of their preferences, then interventions that change these local public amenities have the potential to decrease carbon emissions for many households at once.

In this paper, I decompose variation in household carbon emissions into a component driven by household characteristics and a component driven by place effects – i.e., the amount by which the same household's carbon emissions would differ from place to place due to differences in the underlying features of those places. To do so, I construct a longitudinal panel of residential and transportation energy use for over one million individuals from 20 years of restricted-access Decennial Census and American Community Survey (ACS) microdata. The longitudinal nature of these data allows me to link individual survey respondents over time and across places. I use a mover design, examining changes to household carbon emissions for over 250,000 movers across roughly 1,000 cities and 61,500 neighborhoods, to estimate place effects and their contribution to heterogeneity in carbon emissions.

I begin my analysis by documenting observational patterns of city and neighborhood-level variation in household carbon emissions in my sample. While previous work has shown evidence of substantial variation in household carbon emissions (e.g. Jones and Kammen 2014; Ummel 2014; Green and Knittel 2020), the level of geographic granularity in publicly available data has limited researchers to predicting neighborhood-level carbon emissions from national data projected onto local place and household characteristics. In contrast, detailed geographic identifiers in the administrative Census Bureau microdata make it possible to directly estimate neighborhood-level means. I estimate that on average, households living in cities with emissions one standard deviation above the mean emit 50 percent more than those living in cities one standard deviation below the mean. Similarly, households in neighborhoods one standard deviation above the mean emit, on average, just under twice as much as those in neighborhoods one standard deviation below the mean. Accounting for variation driven by observed household characteristics such as household size and income decreases the dispersion across place estimates by less than 10 percent.

The heterogeneity that remains after accounting for observable household characteristics reflects some combination of unobserved household characteristics and causal place effects. Unobserved household characteristics might include preferences for spending time in a private yard versus a public park, risk tolerances for biking versus driving, sensetivities to hot or cold temperatures, or simply environmental consciousness. Place effects could stem from a variety of local amenities and supply-side factors that determine patterns of household energy use. They could reflect aspects of urban form such as public transportation, bike and pedestrian infrastructure, highway networks, density, or zoning regulations. They could also be driven by natural amenities such as climate. Lastly, they could arise from supply-side factors that determine fuel shares and electricity emissions factors, both of which shift the amount of carbon emitted for a given level of energy use.[1] I show how place effects can be interpreted through the lens of a consumer energy demand model in which average energy demand, energy demand elasticities, energy prices, and average emissions factors vary across places.

My empirical strategy uses movers to estimate the contributions of place effects and household characteristics to heterogeneity in household carbon emissions. The mover design accounts for unobserved differences between households by comparing carbon emissions for the same household living in different places. Consider the following thought experiment. Imagine two households identical in every observable way - same household size, same income, same education levels, etc. One lives in a dense, urban neighborhood well-served by public transit, while the other lives in a car-dependent suburb with large homes. The urban household likely generates lower carbon emissions. But is that because the place itself determines emissions, or because the urban household has different unobservable characteristics, like stronger environmental preferences, that led it to select into the lower emissions neighborhood? If I observe the same household move between neighborhoods, any resulting change in that household's emissions can no longer be attributed to fixed unobserved household preferences or characteristics, and I can use those changes to infer differences in place effects. In order for estimates from the mover design to be unbiased, the central assumption is that mover destinations are uncorrelated with changes to unobserved determinants of household carbon emissions. A crucial advantage of undertaking this analysis with Census Bureau microdata is that I observe, and can control for, many time-varying household characteristics that could correlate with both potential emissions and destination choices and thereby confound estimates of place effects. In other words, the identifying strategy does not allow for households to move in response to a sudden shift in environmental preferences, but it does allow for households to move in response to changes in income, the number of children, and age-based or lifecycle preferences, because I observe these characteristics.

1. For instance, a household in an area with predominantly coal-fired power plants could emit significantly more carbon than a household with identical electricity consumption in an area powered mainly by natural gas or renewable sources. Similarly, many households in the Northeast use heating oil, in part due to legacy equipment choices, which produces substantially more carbon dioxide for the same amount of heating than natural gas.

For the first set of results, I use an event study to estimate how much carbon emissions change after households move, as a share of the mean difference between their origin and destination. Intuitively, if spatial heterogeneity is driven in part by underlying differences between places, when a household moves, I should see its emissions shift towards the mean of its new location. The larger the shift, the more important the role of place. I find that, on average, when households move to a new city, their carbon emissions change by about 85 percent of the mean difference between origin and destination cities. Sorting plays a larger role in neighborhood-level variation than in variation across cities, but the role of place remains meaningful; when households move to a new neighborhood, their carbon emissions change by 53-60 percent of the mean difference between their origin and destination neighborhoods.

I explore several dimensions of heterogeneity, finding that my estimates remain stable when restricting the sample to households without significant changes in observable characteristics, when splitting the sample by duration between observations, and when grouping movers based on the magnitude of origin-destination differences in observational mean emissions. The consistency of estimates across these analyses suggests limited systematic sorting of households to places, as such sorting would manifest in a non-linear relationship between observational means and household emissions. In the absence of systematic sorting, the event study estimates can be interpreted as causal estimates of the effect on emissions of any household moving between any pair of places. While the heterogeneity analysis lends support to this interpretation, the additional assumption of no systematic sorting is quite strong, particularly at the neighborhood level. Under the weaker baseline assumptions, the event study estimates yield unbiased predictions about how household carbon emissions change for any set of observed moves. This is valuable, as it makes it possible to calculate the carbon emissions externality of policies or regulatory restrictions that drive existing patterns of household migration, for instance, policies that restrict housing supply in on-average lower emissions cities.

In the second set of results, I estimate the full non-parametric distribution of household and place effects using a two-way fixed effects model, and then do a variance decomposition to estimate the share of overall heterogeneity explained by each component. This approach allows for unrestricted patterns of sorting, but this weaker assumption comes at the cost of limited mobility bias (Andrews et al. 2008): estimates of place effects are noisy because they can be derived from a small number of movers to and from each place. This imprecision creates an upward bias in the naive plug-in variance estimate relative to the true variance of place effects, even if estimates of place effects themselves are unbiased. I account for this upward bias using the heteroskedasticity-robust "leave-out" estimator proposed by Kline, Saggio, and Sølvsten (2020). I find low correlations between unobserved household and place effects, even at the neighborhood level. This suggests that sorting on unobserved household characteristics contributes to differences between places through "segregation" of households, but not in a way that is systematically correlated with unobserved neighborhood attributes. City effects explain 14-16 percent of overall heterogeneity, while neighborhood effects explain roughly 22-23 percent of overall heterogeneity. Climate, electricity emissions factors, and energy prices together account for over half of the CBSA variance component. At the neighborhood level, controlling for these factors decreases the place share by less than half, to about 15 percent. While this leaves the majority of overall heterogeneity to other factors, my estimates nevertheless imply the potential for considerable reductions to household carbon emissions from interventions that decrease place effects: I estimate that if a neighborhood went from having a place effect one standard deviation above the national mean to having a place effect one standard deviation below the national mean, household emissions for residents of that neighborhood would decrease by about 40 percent.

I characterize low- and high-emissions neighborhoods by presenting correlations between estimated tract effects and observable tract-level characteristics. I characterize local amenities using observable characteristics from within the Census Bureau microdata, public-use data on climate and electricity, and commercial data from Walk Score, a private company that generates estimates of the walk-ability, transit-ability, and bike-ability of every address in the US. These data lend unique insight into highly granular variation in neighborhood characteristics. I find that the correlations between amenities and neighborhood place effects for the most part mirror the relationships in the observational data. Low-emissions places have clean electricity and mild climates, and they have amenities characteristic of urban areas. Among the local amenities, density and average home size, proximity to principal cities, and the quality of local bike infrastructure appear to have the most explanatory power.

I conclude my analysis by examining how household carbon emissions would differ if some households were exposed to place effects neighborhoods more urban than the one they currently live in. If suburban and rural households lived in a place with the average place effect of the principal city closest to them – a scenario that captures in spirit how households' exposure to place effects might shift in response to regulations that limit urban sprawl and encourage upzoning and infill development – I estimate that their emissions from residential energy use and commuting would decrease by about 15 percent. To put this estimate into context, the Inflation Reduction Act, which was signed into law in August 2022 and is the largest Federal effort to address climate change to date, is projected to decrease economy-wide emissions in 2030 by an additional 15 percent relative to projected reductions from 2005 levels under business as usual.[2]

In many ways, the basic physical design and urban fabric of cities, suburbs, and towns creates the foundational patterns of transportation and residential energy use. Decades of housing, transportation, and land use policies shape these features. The wide distribution of place effects estimated in this paper implies that there may be potential for "place-based climate policies" – policies that aim to reduce household carbon emissions from residential and transportation energy by changing the underlying characteristics of the places people live in – to lead to meaningful reductions in carbon emissions. While this paper does not estimate the causal drivers of place effects, the correlational analysis presented here, and many observational and model based studies (e.g. Shammin et al. 2010; Timmons, Zirogiannis, and Lutz 2016; Ribeiro, Rybski, and Kropp 2019; Pomponi et al. 2021; Ko 2013) provide some hypotheses. Estimating causal relationships between specific amenities and place effects using credible exogenous variation remains

2. Three separate efforts to model IRA reductions have been commonly cited by advocates and lawmakers. The Rhodium Group estimates that under the IRA emissions will fall 31-44 percent from 2005 levels by 2030, with 24-34 percent reductions under business as usual (Larsen et al. 2022). For the same time frame, Energy Innovation estimates 37-41 percent reductions under the IRA and 24% under business as usual (Mahajan et al. 2022), and The REPEAT project estimates 42 percent reductions under IRA and 17 percent under business as usual (Jenkins et al. 2022).

an important direction for future research. One might have worried about whether the settings in which such studies can be done are too selected; what if the households that reside in places that make changes to local infrastructure or regulatory restrictions are different from households in places that do not make those changes? The core results of this paper suggest that variation in place effects drives a meaningful share of differences in emissions between places, mitigating some of these concerns about external validity.

This paper makes several contributions. A large body of work in labor and urban economics finds significant wage, employment, and productivity benefits from density and integrated land use and transit policies (Tsivanidis 2022; Allen and Arkolakis 2022; Duranton and Puga 2020). These studies suggest that spatial equilibria are inefficient due to agglomeration economies and other externalities, but have largely not considered the carbon emissions externality in their analysis. The theoretical justification for using place-based policies in cases where agglomeration economies and other local externalities exist is well-established. Federal intervention can correct inefficient market equilibria and improve welfare by supplementing local government provision of under-provided amenities, offering a "big push" towards an optimal equilibrium when several exist, fostering the growth of productive areas and agglomeration externalities, or insuring residents against place-based shocks (Glaeser and Gottlieb 2008; Kline 2010; Kline and Moretti 2014; Glaeser 2013; Austin, Glaeser, and Summers 2018). However, the empirical evidence on the efficacy of place-based policies is mixed. While some studies find that tax incentives targeting areas with lower employment can improve welfare (Busso, Gregory, and Kline 2013; Austin, Glaeser, and Summers 2018; Bilal 2023), others suggest that spatial policies promoting growth in less developed areas may have negligible or even negative aggregate effects on productivity and welfare (Kline and Moretti 2013; Gaubert 2018). Duranton and Venebles (2021) highlight the challenges in evaluating place-based policies, including in the context of urban transport, housing, and infrastructure. The welfare implications of place-based climate policies would crucially depend on their design, implementation details, costs, and household preferences for local amenities; a welfare analysis is beyond the scope of this paper. Carbon emissions are a canonical example of a global externality, but many of the theoretical justifications for place-based policy outlined above could be relevant, given the observed relationship between carbon emissions and factors of urban form such as density and urban transport. This, together with evidence from this paper that places play an important role in driving household carbon emissions, suggests that that further research is warranted.

Methodologically, I build on a large literature in labor examining wage inequality across firms, and a growing literature that uses mover designs to estimate place effects on other individual outcomes, e.g. nutritional choices (Allcott et al. 2019), health outcomes and health care utilization (Eid et al. 2008; Finkelstein, Gentzkow, and Williams 2016; 2021), intergenerational mobility (Chetty and Hendren 2018), and wages (De la Roca and Puga 2017; Card, Rothstein, and Yi 2024). This paper is the first to use a mover design to study household energy use and carbon emissions, yielding new insights into spatial heterogeneity in these outcomes. Previous work has highlighted the consequences of spatial heterogeneity in carbon emissions for allocative efficiency (Glaeser and Kahn 2010; Colas and Morehouse 2022) as well as for distributional impacts and the political economy of hypothetical climate policies (Cronin, Fullerton, and Sexton

2019; Sallee 2019; Green and Knittel 2020), but these papers did not examine the causal role of places in their findings. Several papers in the literature have generated estimates of heterogeneous energy demand parameters, but they have necessarily done so in spatially limited and sector-specific contexts (Auffhammer and Rubin 2024; Gillingham 2014; Nowak and Savage 2013; Spiller et al. 2014). The estimates generated in this paper on the relative roles of place effects versus household sorting could assist in resolving some of the challenges identified in the aforementioned literature around distributional impacts and political economy. For instance, if lock-in of urban form limits the share of household carbon emissions that could be targeted by pricing instruments in the short-to-medium term, estimates of the share of spatial heterogeneity driven by place effects could inform decisions around how much to redistribute carbon dividends when using geography as a tag. And as highlighted previously, they raise an important question as to whether place-based climate policies could serve as a welfare-improving complement to traditional instruments for addressing this global externality.

2 Data and Stylized Facts about Carbon Emissions in the US

75 percent of US greenhouse gas emissions are from burning fossil fuels. Of these, 20 percent are from residential energy use (including electricity), and another 20 percent are from light duty (i.e. passenger) vehicles (U.S. Energy Information Administration 2020b). The focus of this analysis is on carbon emissions from these two sectors; they are a meaningful portion of overall emissions, and they are the forms of emissions that are most directly related to locked-in characteristics of urban form. In the remainder of this section, I first discuss my data and the construction of relevant analysis variables, and then provide some descriptive evidence on heterogeneity in carbon emissions in my sample.

2.1 Data and Key Variables

I build a 20-year panel of individual and household-level data using the 2000 restricted access Decennial Census long form and the 2001-2019 American Community Survey (ACS). The 2000 Decennial Census long form consists of a stratified random sample covering one in six households in the US. After 2000, the ACS replaced the Decennial Census long form in order to gather detailed information on individuals and households more regularly. The ACS is a stratified random sample covering roughly 0.4 percent of households in 2001-2005, and roughly one percent of households in each year after 2005 (U.S. Department of Commerce 2014). I link individuals across surveys using Protected Identification Keys, which are unique person identifiers assigned by the US Census Bureau based on names, addresses, dates of birth, other household members, and social security numbers (when available).[3]

3. Neither the Decennial Census nor the ACS ask respondents for their social security numbers. Wagner and Layne (2014) use data with social security numbers to show that the error rate in assigning Protected Identification Keys without social security numbers is below one percent. See Bond et al. (2014) for detailed discussion of the assignment algorithm used by the US Census Bureau. Assignment success rates vary across demographic groups – in particular white and higher income individuals are more likely to be successfully assigned a Protected Identification Key – but for all demographic subgroups the success rate is greater than 85 percent. See Bond et al. (2014) for additional discussion of the variation in assignment rates across population subgroups.

For every individual in the panel, I observe measures of residential and transportation energy use, and a rich set of demographic, household, workplace, and home characteristics, including detailed geographic identifiers. I supplement the Decennial Census and ACS with several external sources of data in order to convert energy expenditures to energy services and emissions, and to characterize places.

2.1.1 Geographic Units of Analysis

Throughout the study, I analyze spatial heterogeneity at two levels of geographic granularity which roughly represent a city or labor market and a neighborhood.

My first geographic unit of analysis is a Core Based Statistical Area (CBSA). CBSAs are designated by the Office of Management and Budget and cover the population of metropolitan and micropolitan areas in the US. Each CBSA is a set of contiguous counties with strong commuting ties and at least one urban core area of at least 10,000 people. In addition to formally designated CBSAs, I define state-level residual CBSAs from unassigned rural areas. My second geographic unit of analysis is a census tract. Census tracts are county subdivisions that typically cover contiguous areas, have populations of 1,200-8,000 people (4,000 on average), and are delineated with boundaries that follow identifiable physical features. They are designed to be relatively stable, but are split or merged every ten years if populations exceed or fall below the 1,200-8,000 window.[4]

2.1.2 Carbon Emissions

My primary outcome is metric tons of carbon emissions from residential energy and passenger vehicle use, which together account for roughly one third of US greenhouse gas emissions. I implement the main analysis at the household level: carbon emissions are given by household residential emissions plus the sum of individual commuting emissions over all individuals in the household.

I estimate carbon emissions from residential energy use from household-reported expenditures on electricity, natural gas, and other home heating fuels in the last year, combined with external data on local annual retail prices and fuel emissions factors. For electricity, I calculate county-level average prices using data from the Energy Information Administration (2020a) Annual Electric Power Industry Report. This report contains sales, revenues, and total customers for every major utility in the US, by sector and state. It also delineates counties contained in each utility's service territory. I calculate county-level retail electricity prices using customer-weighted average prices (revenue divided by sales) across all utilities with service territories containing the county, and I compute household electricity consumption by dividing reported expenditures by my price estimates. I then assign households to one of 12 National Electric Reliability Council (NERC) regions using a tract-level crosswalk from the U.S. Department of Homeland Security (2021) Infrastructure Foundation-Level Database, and compute emissions using the average

4. Census geographic definitions vary over time to account for changes in administrative boundaries and populations. To ensure that I don't erroneously identify people who live in places where the designation changed as movers, I use the 2000-2010 census block concordance to assign 2010 geographic definitions to all years in the data, combining blocks in cases where they correspond to a single 2000 block.

annual emissions rates assigned to each region by the U.S. Environmental Protection Agency (2021a) Emissions & Generation Resource Integrated Database. For natural gas and other home heating fuels, I obtain average retail prices at the state level from the Energy Information Administration (2020b) State Energy Data System. If a household reports non-zero expenditures on "other home heating fuels," I impute the fuel used from its answer to the question "What was the primary fuel used for home heating?" Finally, I obtain fuel emissions factors from the U.S. Environmental Protection Agency (2018) Emission Factors for Greenhouse Gas Inventories.

I estimate carbon emissions from transportation energy from the sum of individually-reported commuting behavior within a household.[5] I estimate commute distance using the geodesic distance between home and place of work census blocks, and I estimated commute speed from estimated mileage and reported time-length of commute. I estimate gasoline usage using annual national average fuel economy from the U.S. Environmental Protection Agency and Energy (2020), accounting for the fact that in general fuel economy is roughly 30 percent higher on highways than in cities. Finally, I estimate the number of annual commutes using reported weeks worked last year and hours worked last week, and convert annual gallons of gasoline to carbon emissions using the motor gasoline emissions factor from the Energy Information Administration (2020b) State Energy Data System. Individuals who commute by rail, subway, streetcar, bus, bike, or walk, and individuals who work from home are assigned an emissions factor based on their mode of transit and data from the National Transit Database (see Appendix A.1 for more details). Altogether, this portion of the outcome captures variation in carbon emissions driven by commute lengths, number of commutes, and mode of transit. I examine the sensitivity of my results to using the Federal Highway Administration (2019) National Household Travel Survey (NHTS) to predict heterogeneous fuel economy and non-commute miles from household and geographic characteristics available in both the Census and NHTS. This is not my baseline approach, as it infers how much of variation in vehicle fleets and fuel economy observed in the NHTS is driven by individual preferences vs. place-based factors from cross-sectional variation.[6]

2.1.3 Individual and Household Characteristics

Throughout the analysis, I use demographic and household characteristics to control for variation driven by observable characteristics. Specifically, I control for age, education (completion of a bachelor's degree), sex, race and ethnicity, household income (from salaries and wages, interest, social security, supplemental security, public assistance, retirement, and self employment), household size, number of children, and homeowner status. I aggregate individual-level demographics to the household level by taking the mean across individuals within a household.

I also observe whether a household lives in a detached single-family home, the number of rooms in a home, and the number of vehicles in the household. These characteristics are intermediate outcomes, which directly affect carbon emissions. They also very likely reflect a

5. Commuting accounts for about 28 percent of all vehicle-miles travelled, and 39 percent of person-miles travelled on transit systems (U.S. Department of Transportation 2015), which means I underestimate carbon emissions from overall personal vehicle use for most people in my sample.

6. Place-based factors that contribute to variation in vehicle fleets could include social norms, perceptions of safety (e.g. if everyone around you is driving a big car it is safer for you to drive a big car; certain types of cars may be able to handle adverse weather better), road widths, ease of parking, etc.

combination of household preferences and place characteristics – many places impose restrictions on multi-unit homes and/or minimum lot sizes, and lack transportation options for households without a car, so household choices are likely to differ from place to place depending on these constraints. Therefore, I do not treat these variables as observable household characteristics when estimating place and household effects, but I do use them later to explore correlates of unobserved place and household heterogeneity.

It is not obvious whether homeownership should be considered an observable household characteristic or part of a place effect. Homeownership rates vary dramatically across CBSAs (Raetz 2021; Mateyka and Mazur 2021), and housing regulations can price people out of homeownership. In these cases, treating homeownership as an intermediate outcome seems appropriate. On the other hand, the choice to become a homeowner simultaneously impacts where people live and factors related to their carbon emissions. For instance, homeowners may want extra space for potential family expansion or space-intensive leisure activities, or may choose homeownership in order to be able to install solar panels and have a place to charge their electric vehicle. To be conservative and err on the side of finding a smaller role of place effects, I treat homeowner status as an observable household characteristic in my baseline analysis.

2.1.4 Place Characteristics and Amenities

I supplement Census micro-data with several external sources of data to characterize amenities at the block, tract, city and regional level. I focus on amenities that are directly relevant to energy consumption and carbon emissions in the residential and transportation sectors.

To capture variation in climate, I use data on annual heating degree days (HDDs) and cooling degree days (CDDs) at the CONUS Climate Division level (National Oceanic and Atmospheric Administration 2020). The National Oceanic and Atmospheric Administration (NOAA) divides the contiguous states into a total of 344 climate divisions based on regional differences in climates within states. Degree days represent the annual sum of the daily difference between that day's temperature and 65°F, and are meant to quantify the heating and cooling requirements of a place.

To account for neighborhood-level variation in transportation and leisure amenities, I use data from Walk Score, a private company that generates estimates of the walk-ability, transit-ability, and bike-ability of every address in the US.[7] Walk Score® rankings capture proximity to different commercial amenities such as grocery stores, as well as street characteristics such as block lengths and intersection widths. Bike Score™ indices capture characteristics that make biking more or less accessible, such as the existence of bike lanes, road connectivity, and hilliness. Transit Score® ratings capture proximity to different types of transit, and the frequency and connectivity of nearby options. For transit, I also observe the number of bus routes and rail routes within a half-mile. Other than route counts, each score is an index from 0-100. I assign over six million unique Walk Score points reflecting data from early 2020, one to every populated census block in the US, by matching census block centroids to the nearest Walk Score latitude-longitude coordinate.

7. Data can be viewed at www.walkscore.com, and was provided by Redfin Real Estate.

Finally, I estimate density at the tract-level using 2010 census block-level information on area and population. I define urban tracts as those that are characterized as urban by the US Census Bureau *and* surpass the density threshold set for urban centers by the EU-OECD definition of a functional urban area (Dijkstra, Poelman, and Veneri 2019).[8] I define suburban tracts as those contained within a CBSA but not designated as urban. Tracts outside of CBSAs are classified rural.

2.1.5 Sample Restrictions

I restrict the analysis to individuals who are at least 18 years of age, who are not identified as the householder's child or grandchild, and who are not missing any of the outcome variables or key explanatory or control variables described above. I also impose several additional restrictions related to energy variables. I exclude from the sample households for which residential energy costs are included in rent or gas costs are included in electricity bills, because I don't observe expenditures in those cases. I discuss this sub-sample of households and the potential impact of its exclusion in detail in Appendix A. I also exclude individuals in households in which residential energy use is top coded or whose commute time is top coded, as the top-coding will obfuscate changes in individual consumption for the highest demand individuals. Lastly, I exclude individuals if the sum of their household residential energy expenditures is zero, if they are in the bottom one percent of non-zero residential energy cost observations, or if they are in the top one percent of commute distance observations, as these outliers more likely reflect survey misreporting. My full sample consists of all individuals who meet these restrictions across the 48 continental states and the District of Columbia – almost 17 million people across over 12 million households (Table 1, column (1)). I use the full sample to estimate observational geographic and household heterogeneity.

I construct a panel sample by restricting the full sample to individuals for whom I have at least two observations in which they did not indicate that they had moved within the last year. This restriction on very recent migration ensures that I am estimating household carbon emissions at the correct location, as households report their residential energy expenditures over the past year.[9] The panel sample consists of 1,097,000 people across 916,000 households (Table 1, column (2)).

Finally, I impose two additional sample restrictions which are necessary for the implementation of my empirical strategy. First, because residential energy is determined at the household level, and place effects are identified from the variation in outcomes of movers between places, I restrict the panel sample to only households consisting of the same full sample individuals across observations. This ensures that changes in emissions across observations are not driven by changes in the composition of adults in the household. Importantly, this approach does retain households where children are born or grow up and move out between observations, as individuals under 18 are not part of the full sample. Second, I restrict CBSAs and tracts to the "leave-out connected set" – the network of CBSAs or tracts that remain connected to each other

8. This threshold is 1,500 people per square kilometer.

9. This restriction applies to all ACS years. The 2000 Decennial Census, asked whether respondents had moved within the last five years. Since this is significantly more restrictive, I don't drop these individuals.

by at least one mover after I drop all the observations of any given household (see Appendix B for an illustration). I do this after dropping tracts with fewer than 10 full sample household observations. The leave-out connected sets are constructed separately at the CBSA and tract level. This means it is possible for a household to be in the CBSA panel but not the tract panel. The leave-out restriction drops a negligible share of (residual) CBSAs and roughly 12 percent of (disproportionately rural) tracts, yielding approximately a 5 percent reduction in the number of households in the sample (Table 1, columns (3) and (4)).

CBSA movers are households in the CBSA panel that live in different CBSAs across observations (93,000 households, column (5)), and similarly, tract movers are households in the tract panel that live in different tracts (within or across CBSAs) across observations (248,000 households, column (6)). The CBSA panel, tract panel, CBSA movers, and tract movers make up my four primary analysis samples.

2.2 Sample Statistics

Table 1 shows sample statistics for the full sample, unrestricted panel sample, the two geographically restricted panel samples, and the two mover samples.

A comparison across the samples yields three main take-aways. First, households in the panel are on average more likely to be white, have higher income, and are more likely to be homeowners than households in the full sample (columns (1) and (2)). This reflects known heterogeneity in Protected Identification Key assignment rates within the Census Bureau (Bond et al. 2014). The panel sample is also seven percentage points less likely to live in an urban tract, nine percentage points more likely to live in a detached home, and two percentage points more likely to commute by car. The appreciable drop in transit score when going from the full sample to the panel is consistent with these differences in urbanity, and is likely in part driven by disproportionately dropping households in the densest areas when dropping households whose electricity and/or heating is included in rent, as discussed earlier and also in Appendix A. Second, further restricting the baseline panel to the CBSA and tract panels (columns (3) and (4)) does not meaningfully change the distribution of demographics, (intermediate) outcomes, or place characteristics. Finally, movers (columns (5) and (6)) tend to be younger, more college educated, and have higher income than both stayers and the full sample. Movers also are more likely than stayers to live in urban tracts, less likely than stayers to live in detached homes, and they have higher rates of electric heating and lower emissions from residential energy, making them more comparable to the full sample on all of these dimensions.

Table 1: **Sample Statistics**

		Panel Sample			Mover Sample	
	(1) Full	(2) All	(3) CBSA	(4) Tract	(5) CBSA	(6) Tract
A: Demographics						
College	0.25	0.25	0.25	0.25	0.35	0.31
Age	44	46	46	46	43	43
White	0.82	0.89	0.89	0.90	0.89	0.88
Female	0.48	0.48	0.48	0.48	0.45	0.47
Household Income	103,400	114,100	114,200	114,900	116,100	115,400
Household Kids	1.0	1.0	1.0	1.0	1.0	1.0
Household Size	2.8	2.9	2.9	2.9	2.8	2.9
Homeowner	0.75	0.85	0.85	0.85	0.72	0.73
B: Outcomes						
Tons CO_2	15.1	16.1	16.0	16.1	15.0	15.0
Tons CO_2 - Residential	11.9	12.8	12.8	12.8	11.8	11.9
Tons CO_2 - Commute	3.3	3.3	3.3	3.3	3.3	3.1
C: Intermediate Outcomes						
Detached Home	0.72	0.81	0.81	0.81	0.73	0.73
Use Electricity Only	0.30	0.24	0.24	0.24	0.30	0.28
Commute by Car	0.94	0.96	0.96	0.96	0.95	0.96
Commute Minutes	25.1	24.9	24.9	24.8	26.0	25.6
D: Place Characteristics						
Urban	0.29	0.22	0.23	0.22	0.20	0.25
Suburban	0.62	0.62	0.63	0.63	0.68	0.67
Rural	0.09	0.15	0.15	0.15	0.11	0.08
Walk Score	27.2	23.0	23.1	22.4	23.0	25.2
Bike Score	35.9	33.4	33.4	33.1	34.1	35.4
Transit Score	9.0	6.7	6.7	6.3	6.7	7.9
N Bus Routes	1.6	1.1	1.1	1.0	1.2	1.3
N Rail Routes	0.15	0.09	0.09	0.07	0.10	0.10
Cooling Degree Days	1,359	1,213	1,215	1,205	1,356	1,335
Heating Degree Days	4,376	4,824	4,815	4,851	4,494	4,518
N People	16,900,000	1,097,000	1,073,000	1,042,000	107,000	290,000
N Households	12,600,000	916,000	860,000	833,000	93,000	248,000
CBSAs	1,000	1,000	1,000	1,000	1,000	1,000
Tracts	71,500	70,000	70,000	61,500	54,500	61,500

Note: Column (1) shows statistics for the full sample. Column (2) shows statistics for the panel sample, with no restrictions that individuals be in the same household or live in a connected geography. Columns (3) and (4) show the panel samples restricted to individuals in a consistent household overtime and the CBSA and tract leave-one-out connected sets, respectively. Columns (5) and (6) show statistics for the CBSA and tract mover samples. All sample statistics are weighted using Census sample weights. Sample counts are unweighted and rounded according to Census Bureau disclosure rules.

Table 2: **Panel Statistics**

	Panel Sample		Mover Sample	
	CBSA	Tract	CBSA	Tract
A: Sample Characteristics				
First Observed in 2000	0.10	0.10	0.16	0.14
Years Between Observations	7.9	7.9	10.3	9.8
B: Demographic Characteristics				
Age First Observed	42.0	42.0	37.1	37.2
Share with Large Change in Income	0.28	0.28	0.45	0.41
Share with Change in N Kids	0.45	0.45	0.55	0.55
Change in N Kids	-0.12	-0.12	0.08	0.08
Share Rent to Own	0.11	0.11	0.26	0.26
C: Mover Place Changes				
Δ Walk Score			-7.0	-7.1
Δ Bike Score			-4.2	-4.0
Δ Transit Score			-2.3	-2.7
Δ N Bus Routes			-0.57	-0.56
Δ N Rail Routes			-0.04	-0.04
Δ Tract Share Detached Home			0.05	0.05
% Moves Urban-to-Urban			0.08	0.14
% Moves Urban-to-Suburban			0.17	0.14
% Moves Suburban-to-Suburban			0.44	0.46
%Δ Cooling Degree Days			214	136
%Δ Heating Degree Days			-331	-188
N People	1,073,000	1,042,000	107,000	290,000
N Households	860,000	833,000	93,000	248,000
CBSAs	1,000	1,000	1,000	1,000
Tracts	70,000	61,500	54,500	61,500

Note: Columns (1) and (2) shows panel statistics for the CBSA and tract panel samples. Columns (3) and (4) show statistics panel statistics as well as summary measures of mobility patterns for the CBSA and tract mover samples. All sample statistics are weighted using census sample weights. Sample counts are unweighted and rounded according to Census Bureau disclosure rules.

Overall, a little under 80 percent of household carbon emissions in my sample are from residential energy, and a little over 20 percent are from commuting.[10] Close to three quarters of the sample lives in a detached, single family home, a vast majority of the sample commutes

10. Household carbon emissions from residential and transportation energy are roughly evenly split (U.S. Energy Information Administration 2020b). Given that commuting makes up about 30% of transportation energy emissions, we would expect a slightly higher than 3:1 ratio of residential to commuting energy. My estimates appear to overstate residential energy use relative to commuting by a few percentage points at most.

by car, and on average households live within half a mile of only one bus route and only 0.1 rail routes.

Table 2 shows additional statistics for the panel samples. I observe the vast majority of households in the panel sample exactly twice, with on average 8-10 years in between observations. Movers tend to be younger than stayers the first time I observe them, and are much more likely to have had a child, experience a large change in real household income[11], or transition from renting to owning their home. Households tend to move to places with higher shares of detached single-family homes and worse non-car transportation amenities. The majority of moves are from urban to urban tracts, urban to suburban tracts, or suburban to suburban tracts. Finally, consistent with secular trends of mobility in the US, households generally move to warmer places. For additional comparisons of movers vs. stayers, estimates of the likelihood of moving given shocks to household income, number of children, or changes in homeownership, and the full set of transition probabilities across urban, suburban, and rural places, see Appendix Table G.3, Table G.4, and Table G.5 respectively.

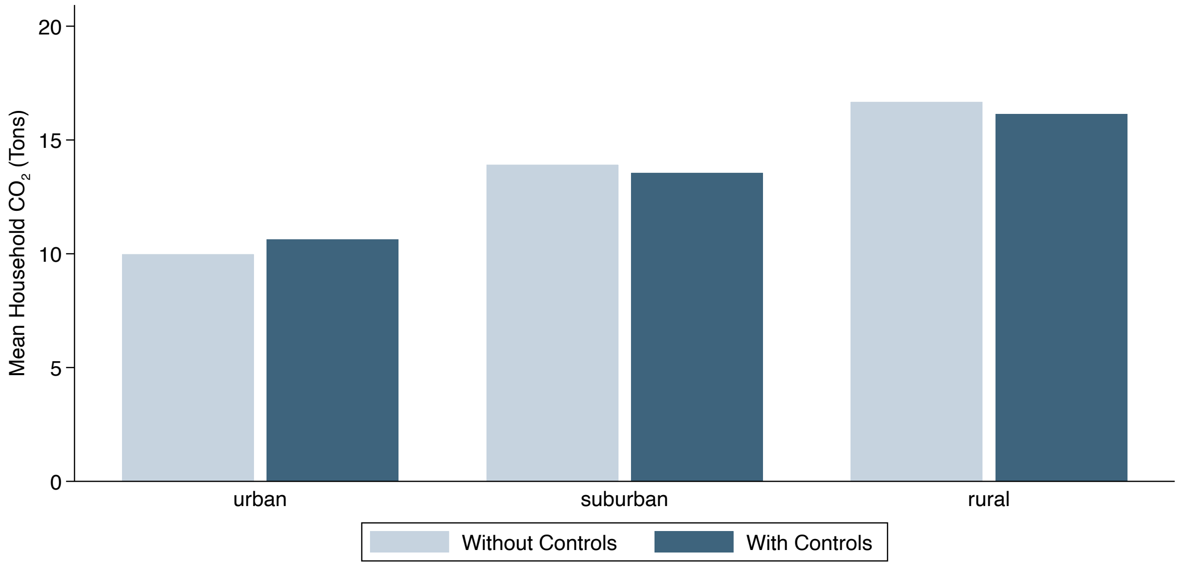
2.3 Observational Heterogeneity

Carbon emissions from residential energy and commuting vary immensely across individuals in the full sample (Figure G.1). This variation is strongly correlated with both geographic and household attributes. At the regional level, carbon emissions vary with climate and with the emissions intensity of local fuel sources. At the more local level, many have observed a relationship between emissions and local amenities that characterize urban form, such as local public transit, bike infrastructure, green space, and density.(Ou et al. 2013; Philips, Anable, and Chatterton 2022).

Figure 1 presents average household carbon emissions across urban, suburban, and rural neighborhoods. Households residing in suburban and rural areas have substantially higher emissions than those living in urban areas. While those differences could, in theory, result from sorting based on observable household characteristics such as college education, race and ethnicity, household income, or number of children, among others (Figure G.2), controlling for observable household characteristics only marginally reduces the differences between rural, suburban, and urban areas. There is some sorting of higher-emissions households to suburban and rural areas and lower-emissions households to urban areas, but even after accounting for these differences, households in suburban tracts still emit over 20 percent more per year than observationally similar households in urban tracts, while households in rural tracts emit about 50 percent more.

11. I define a large change in income as a greater than 0.5 (in absolute value) change in log income. This approximates a 50 percent increase/decrease in income, and corresponds to about the top quarter in absolute income change in my panel sample.

Figure 1: **Household Carbon Emissions in Urban, Suburban, and Rural Places**



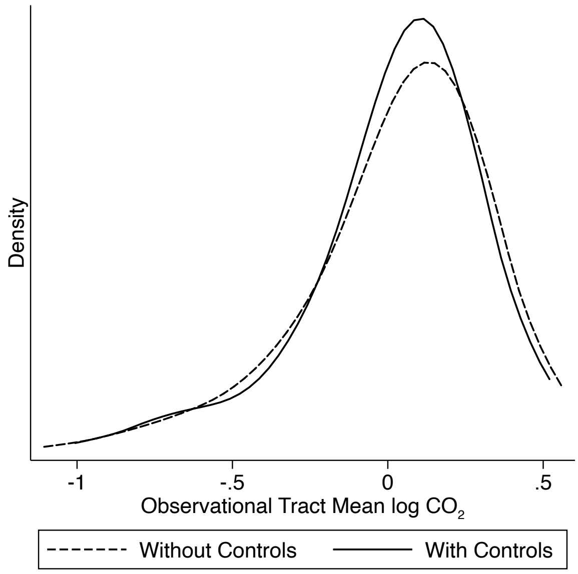
Note: This figure shows estimates of household carbon emissions by urbanity. Estimates are derived from regressions of log carbon emissions (in metric tons) on indicators for urban, suburban, and rural tracts, with year fixed effects. The specification with controls additionally includes age, gender, race, education, household size, number of children, and homeowner status. I define urban tracts according to Dijkstra, Poelman, and Veneri 2019. I define suburban tracts as non-urban tracts within a Core-Based Statistical Area (CBSA). I define rural tracts as those outside CBSAs. Mean household carbon emissions are precisely estimated and statistically different from one another. The unconditional regression has an R^2 of 0.11; the conditional regression has an R^2 of 0.29.

To examine spatial heterogeneity in more detail I estimate unconditional and conditional tract-level means, μ_j, using an ordinary least squares regression of log of household carbon emissions onto place fixed effects, year fixed effects τ_t, and in the conditional regression, individual and household observable characteristics X_{it}.

$$\ln CO_{2it} = \mu_{j(i,t)} + X_{it}\beta + \tau_t + \varepsilon_{it} \tag{1}$$

I use an Empirical Bayes "shrinkage estimator" to adjust the estimates for statistical noise (see Appendix C for details). In practice, the distributions and relevant moments are almost identical for the adjusted and unadjusted estimates. Figure 2 presents the adjusted conditional and unconditional distributions of $\hat{\mu_j}$. I estimate that households living in neighborhoods one standard deviation above the mean emit on average approximately 1.9 times more than those living in neighborhoods one standard deviation below the mean, or 1.8 times more after accounting for differences in observable characteristics. For the remainder of this paper, I refer to means conditional on observable characteristics as "observational means", following the terminology used by Abaluck et al. (2021).

Figure 2: **Heterogeneity in Tract-Level Carbon Emissions**



Note: This figure shows kernel density estimates, using a Gaussian kernel function, of tract-level observational mean emissions. The distributions are censored at the top and bottom 1% of observations in order to abide by Census Disclosure Avoidance rules. The dotted gray line labeled "Without Controls" corresponds to the distribution of log CO_2 conditional on year fixed effects only, and has a standard deviation of 0.33, while the solid line labeled "With Controls" conditions on observable household characteristics, and has a standard deviation of 0.30. Both distributions are de-meaned to match the model with controls. Observable characteristics include age, gender, race, ethnicity, education, home owner status, household income, household size, and number of children.

The remaining heterogeneity in observational means reflects some combination of place-based characteristics and unobserved household characteristics. To illustrate this, I rewrite Equation 1 as a two-way fixed effects model

$$\ln CO_{2it} = \alpha_i + \psi_j + X_{it}\beta + \tau_t + \varepsilon_{it} \tag{2}$$

where ψ_j represents the place-based characteristics and α_i represents the unobserved, fixed, household characteristics. Comparing Equation 1 and Equation 2 highlights the bias that can arise when inferring place effects from observational means, as $\mu_j = \psi_j + E[\alpha_i | i \in j]$. In other words, observational means reflect a combination of place effects and an average over the unobserved characteristics of the residents of a place.

The primary objective of this paper is to disentangle these components and quantify the extent to which heterogeneity in household carbon emissions is driven by unobserved household characteristics and preferences, and how much is driven by causal place effects, i.e. the amount by which the same household's carbon emissions would differ from place to place due to the underlying features of each place, holding household characteristics fixed. This motivates the mover design, which leverages this exact variation.

3 Model

In the previous section, I demonstrated how a two-way fixed effects model can be used to mitigate biases that arise when making inferences based on observational means. To provide further insight into the model, I now illustrate how it relates to a standard way of modeling consumer energy demand, and discuss the interpretation of place and household effects.

Consider a household, i, living in place, j, that consumes quantity Q of energy in the form of four categories of fuels, f. In the residential sector, it can consume electricity (e), natural gas (n), and other heating fuels (o). In the transportation sector, it can consume motor gasoline (m).[12]

Average demand a_j, price elasticities of demand ρ_j^f, and prices P_j^f are permitted to vary by place. Place-based differences in average energy demand and in price elasticities of demand could stem from a range of fixed and malleable characteristics of places. These characteristics include climate, local public goods and urban form (e.g. density, public transit, pedestrian and bike infrastructure, proximity to highways and availability of parking, and proximity to leisure and commercial amenities), and regulatory characteristics (e.g. zoning restrictions that change the size and density of homes, and building codes that change energy efficiency requirements or eliminate natural gas hook ups). All of these features could potentially shift energy demand, both on average and in slope.

In addition to place-based characteristics, demand also depends on observable fixed and time varying household characteristics (such as age, household size, and income) X_{it}, individual fixed unobserved determinants of demand (such as a person's intrinsic risk tolerance for biking on shared roads, aversion to public transit vs. traffic, or relative enjoyment of spending time in their own back yard vs. a public park) α_i, individual time-varying unobserved determinants of demand (perhaps, an increased willingness to bike after reading the latest Intergovernmental Panel on Climate Change report) ε_{it}, and national annual variation τ_t. Taken together, household demand for residential and transportation energy is given by:

$$\ln Q_{it} = a_j + \sum_{f \in \mathcal{F}} \rho_j^f \cdot \ln P_j^f + X_{it}\beta + \tau_t + \alpha_i + \varepsilon_{it} \tag{3}$$

If all households within a place used the same proportion of fuel types, it would be simple to express the above equation, with log carbon emissions as the outcome, in terms of a place-based average emissions factor $\bar{\phi}_j$, and to in turn rewrite that expression as the two-way fixed

12. Electric vehicles are a negligible share of driving in my sample time frame. If someone has an electric vehicle, I over-estimate their emissions, because the electricity they use to charge their vehicle is included in residential energy (if they charge at home) but I also assign them gasoline emissions. As electric vehicles (EVs) become a larger share of the market, observing household vehicle choices, and the place-based role of EV charging networks on these choices, will be critical for future research.

effects model defined in the previous section:

$$
\begin{aligned}
\ln CO_{2_{it}} &= \ln(\bar{\phi}_j Q_{it}) \\
&= \underbrace{\ln \bar{\phi}_j + a_j + \sum_{f \in \mathcal{F}} \rho_j^f \cdot \ln P_j^f}_{\psi_j} + X_{it}\beta + \tau_t + \alpha_i + \varepsilon_{it}
\end{aligned} \tag{4}
$$

In this simplified setting, place effects ψ_j capture the combination of place-based variation in five key factors: average energy demand, fuel price elasticities, fuel prices, electricity emissions factors, and fuel shares. In reality, fuel shares vary not only across places but also across households. For instance, some states and municipalities are attempting to ban natural gas connections for new construction (O'Brien 2023; Cornfield 2023), but even if natural gas is available as an option, some households may still opt to fully electrify their homes, while others may have a strong preference for cooking on gas stoves.

In Appendix D, I examine how the interpretation of the two-way fixed effects model is affected when fuel shares are allowed to vary across individuals. The more complex model includes a term that captures the interaction between fuel emissions factors and a household's fuel shares relative to the place-based average. Because electricity emissions factors vary across places, a household that disproportionately uses electricity will experience a larger drop in emissions when moving from a place with relatively high emissions electricity to a place with relatively clean electricity, compared to the average household. This lack of separability implies that there is some inherent mis-specification in the two-way fixed effects model in this setting.

The model makes two additional simplifications. First, place effects are assumed to be fixed, implying that any time variation, including changes in prices, is absorbed in the place effects, which reflect average differences between places over the sample time frame. Second, while the model allows price elasticities of demand to vary across places, it does not allow for them to systematically vary across households. For example, the model does not allow for different elasticities between high and low-income households. Allowing for heterogeneity in demand elasticities across household characteristics would introduce additional interaction terms in the error, as elasticities would be interacted with place-specific prices. Together, this interaction along with the interaction between household fuel shares and place-specific electricity emissions factors, motivate treating the errors as heteroskedastic.

4 Empirical Strategy

My empirical strategy uses movers to estimate place effects and their contribution to spatial heterogeneity in carbon emissions. The intuition behind the mover design is the following: Suppose high-emissions places are high-emissions because of causal place effects – for example, because there are no alternatives to commuting other than by car, or because zoning regulations impose constraints on minimum home sizes and density. If so, there would be households that live in high-emissions places for work, or to be near family or near other amenities they enjoy, which may otherwise make lower-emissions residential and transportation choices but are unable to. If those households move from an on-average high-emissions place to an on-average low-emissions place where lower-emissions alternatives become available to them, their carbon emissions should decrease. Conversely, if spatial heterogeneity is driven by strong preferences, then households currently living in detached single-family homes and commuting by car would continue to do so even given alternate options, and moving from on average high to low-emissions places should have little effect on household carbon emissions.

I use the mover design to estimate two versions of a heterogeneity decomposition of household carbon emissions. I begin with an event study, which, under strong assumptions, characterizes the share of differences *between* places attributable to place effects, inferred from movers' changes in emissions relative to origin-destination mean differences. Under weaker assumptions, the event study serves as additional descriptive evidence and model validation for the second decomposition, as well as unbiased prediction of how carbon emissions will change for the set of observed moves. I then estimate the non-parametric distribution of household and place effects and decompose *overall* heterogeneity into variance components. In the following subsections, I first discuss modeling and identifying assumptions, and then I describe each of these decompositions in more detail.

4.1 Main Assumptions

My empirical strategy at its foundation pairs a two way fixed effect model with a mover design. In order for estimates from this approach to be unbiased, three assumptions need to hold: (1) additive separability of place effects, or constant effects, (2) non-persistence of outcomes, and (3) exogenous mobility, or conditional orthogonality. I discuss each below.

Assumption 1: Additive separability of place effects, or constant effects.

A core modeling assumption of the two-way fixed effect design is that the outcome – log carbon emissions – is additively separable in household and place effects. The log specification is statistically appealing as it reflects the approximate log-normality of the household carbon emissions distribution (Figure G.1). It is also conceptually appealing, as it implies that place effects increase and decrease carbon emissions proportionally by the same amount for everyone, which aligns with many potential mechanisms through which place effects could arise. For example, it is natural to model climate as scaling residential heating or cooling needs up or down by the same factor for all households, regardless of their baseline energy consumption. If density drives place effects, it is reasonable to expect denser places to decrease the size of homes (and therefore residential energy requirements) or the length of commutes (and therefore transportation energy requirements) by the same factor for all households. Similarly, an increase in transportation alternatives to cars might decrease the share of trips taken by car for all households proportionally.

Nevertheless, the two-way fixed effects model imposes a substantial restriction: it does not allow for heterogeneous treatment effects or match effects. Heterogeneous place effects could arise in various scenarios. For instance, a new public transit option might significantly reduce emissions for relatively low emissions households while barely affecting high-emissions households' behavior. Paired with changes in density, this could even produce divergent effects if high emissions individuals find themselves idling in traffic more due to increased congestion. Alternatively, heterogeneous place effects might arise if low emissions households make the same consumption choices wherever they live, while high emissions households respond strongly to changes in amenities. Moreover, Section 3 reveals that there is an interaction between households' fuel shares and local emissions factors – a household that prefers to use electricity for heating will see a larger drop in emissions when moving to an area with cleaner electricity than a household that prefers natural gas heating, all else equal. This interaction introduces a degree of misspecification into the two-way fixed effect model, rendering it an approximation. The critical question is whether households systematically sort into locations based on this (or any other) interaction. As long as they do not, the mover design will continue to yield unbiased estimates of average place effects.

To rule out selection on heterogeneous effects, I follow Card, Heining, and Kline (2013) and test whether moves from a low-emissions place to a high-emissions place and moves from high-emissions place to a low-emissions place are associated with equal and opposite changes in household carbon emissions. To see why symmetry rules out selection on heterogeneous effects, consider differences in potential outcomes across an origin o and destination d, allowing now for there to be an interaction $\eta(\alpha_i, \psi_j)$ between person and place types:

$$E[CO_{2it}(d)] - E[CO_{2it}(o)] = (\psi_d - \psi_o) + \eta(\alpha_i, \psi_d) - \eta(\alpha_i, \psi_o)$$

As long as:

$$\eta(\alpha_h, \psi_d) - \eta(\alpha_h, \psi_o) \neq \eta(\alpha_l, \psi_d) - \eta(\alpha_l, \psi_o) \tag{5}$$

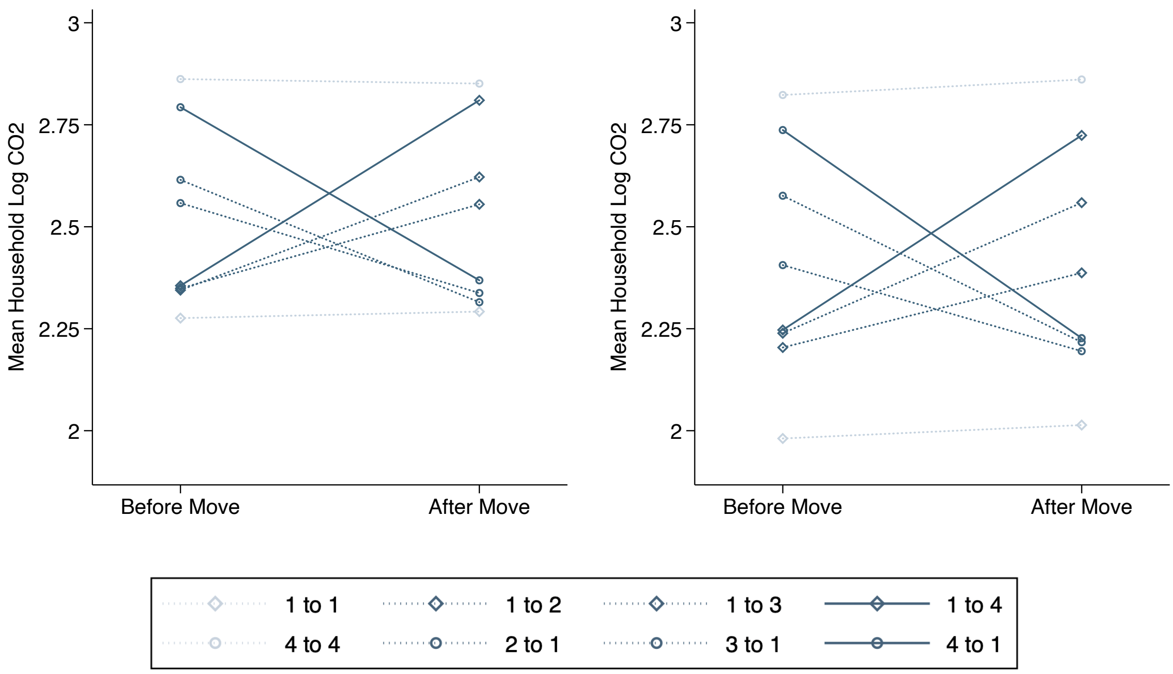
this type of interaction, paired with selection, would lead to asymmetries between changes in household carbon emissions for moves to higher on-average places vs. moves to lower on-average places. Note that condition (5) holds for a broad class of functions, including the simplest interaction, $\eta = \alpha_i \cdot \psi_j$, since $\alpha_i \cdot (\psi_d - \psi_o) \neq \alpha_j \cdot (\psi_d - \psi_o)\ \forall i \neq j, o \neq d$.[13] Returning to an earlier example, suppose place effects are due to a public transit option that only low-emissions households use. Suppose also, that households that wish to use public transit disproportionately move to places where it is available, whereas households who don't want to use public transit disproportionately move to places where it is not available; the subsequent decline in emissions for households moving from a high-emissions place to a low-emissions place will be larger than the increase in emissions for households moving in the opposite direction.

I group places into four quartiles based on observational averages of carbon emissions, and I estimate household carbon emissions for each origin-destination quartile pair, adjusting for annual trends and controlling for demographic and household characteristics. Figure 3 displays the results. For parsimony, the figure shows only moves from the lowest quartile emissions places to all four quartiles and vice versa, as well as moves within first-quartile places and moves within fourth-quartile places as bounds in gray.

13. More generally, Equation 5 implies that $\frac{\partial \eta}{\partial \alpha}[\psi_d - \psi_o] \neq 0$, and in turn, $\frac{\partial^2 \eta}{\partial \alpha \partial \psi} \neq 0$. In words, the symmetry check rules out any interaction in which the change between places grows or shrinks with household type.

Figure 3: **Changes in household CO_2 when moving across quartiles of Mean CO_2**

(a) CBSA (b) Tract



Note: This figure shows average household carbon emissions for movers across places classified into quartiles based on their observational mean carbon emissions in the full sample. Symmetric responses to moves in opposite directions provide support for the log-linear two-way fixed effect model of household carbon emissions because they are consistent with place effects entering log carbon emissions additively. This figure shows only the subset of moves to and from the lowest-emissions places (quartile 1), as well as moves within the highest-emissions places (quartile 4). Estimates are conditional on year fixed effects and the standard set of household characteristics used throughout this analysis. Estimates are weighted by Census sample weights.

First and foremost, this figure shows that moves across quartiles lead to equal and opposite changes in household carbon emissions. This symmetry suggests that the log-linear two-way fixed effect model of household carbon emissions is a good approximation for the role of place effects and alleviates concern about selection on heterogeneous treatment effects. Second, this figure provide evidence of selection at the tails, particularly for tract-level moves – households that move from the lowest quartile to a different place in the lowest quartile have lower emissions on average than households that move between the lowest quartile and any of the three higher quartiles. Similarly, households that move between places in the fourth quartile have higher emissions than those that move between the fourth and the first, though the difference at the top is less pronounced. Finally, note that households that move between tracts within the same quartile experience a small increase in emissions, consistent with the general trend of households moving to on-average higher emissions places.

Assumption 2: Nonpersistent outcomes.

As highlighted above, I identify relative place effects from pairwise comparisons of household carbon emissions between origin and destination,

$$E[CO_{2it}(d)|\alpha_i, X_{it}, \tau_t] - E[CO_{2it}(o)|\alpha_i, X_{it}, \tau_t] = \psi_d - \psi_o$$

This expression holds for any two households moving between o and d, which means that it cannot differ from household to household as a result of differences in their residential histories. In other words, the expected change for two households moving between the same origin and destination should be the same even if one of the households previously lived in Houston while the other previously lived in New York. Note, however, that the nonpersistent outcomes assumption does *not* rule out that the place someone was born and raised may have a persistent effect on their preferences and carbon emissions. Because I include household effects in the model, and only include individuals over the age of 18 in the sample, any persistent effect of place of birth and upbringing on carbon emissions will be captured by the household fixed effect.

Assumption 3: Exogenous mobility, or conditional orthogonality.

When the first two assumptions hold, the model serves as a reasonable approximation to the real world, and consequently random variation in exposure to place identifies place effects. Thus the final, identifying, assumption necessary for this empirical approach to be unbiased is that moves are conditionally exogenous; in other words, household destination choices are not related to changes in unobserved determinants of carbon emissions.

$$E[\epsilon_{it}|\alpha_i, \psi_{j(i,t)}, X_{it}, \tau_t] = 0 \tag{6}$$

It is important to emphasize here that the two-way fixed effects model allows for a broad set of sorting behaviors. First, it allows for unrestricted sorting of households on fixed or time-varying observable characteristics. A key advantage of the Census microdata is the ability to observe numerous time-varying household characteristics that could potentially bias the estimates if unobserved. Factors such as entering middle age, having children, experiencing a change in income, or becoming a homeowner are all associated with an increase in energy consumption generally (Figure G.2), and the latter three also significantly increase the probability of moving (Table G.4). However, this endogeneity does not bias the estimates, as I observe age, household size, number of children, household income, and home-owner status, and am therefore able to separate the effect of these characteristics on household emissions from their influence on a household's choice of new city or neighborhood.

Second, and crucially, the two-way fixed effect model allows for unrestricted sorting on fixed unobservable characteristics. In other words, if households have heterogeneous, but fixed, preferences for neighborhood amenities – for instance, if someone has a particular distaste for public transit, a strong preference for large homes, or a particular love for walking or biking – and their choice of what neighborhood to live in reflects those preferences, estimates of place effects are unbiased by this selection because household fixed effects capture these unobserved but fixed determinants of carbon emissions. The ability to account for, and allow unrestricted sorting on, these time invariant unobserved preferences is a critical benefit of the pairing of the two way fixed effect model and mover estimation strategy.

Note also that the fact that I might never observe a very high emissions household move to a very low emissions place is not a problem if you accept Assumption 1, that places do not have heterogeneous effects on different types of households. It is sufficient to have a connected network of pairwise moves between places in order to identify relative place effects, and household effects are identified relative to other households within the same place. Consequently, even if I observe only lower-emissions households moving from Pittsburgh to New York, and only higher emissions-households moving from Pittsburgh to Houston, a comparison of household emissions within Pittsburgh combined with the observed changes to household emissions at their destinations identify the relevant household and place effects. I have already provided some evidence that place effects do not appear to be heterogeneous in the symmetry check shown in Figure 3. I will provide additional evidence of this when I examine heterogeneity in the event study results, in Section 5.1. If this model assumption is violated and there is selection on heterogeneous treatment effects, then the place effects I estimate should be thought of as local average treatment effects for the population moving to those places.

Thus, the main threat to identification stems from the possibility that moves correspond to *changes* in unobserved preferences – either through an idiosyncratic shock or via preference "drift", i.e., a gradual evolution in preferences that is not captured by aging. Revisiting the example in the previous paragraph, this would reflect a scenario in which the two households in Pittsburgh were initially similar, but then one household becomes increasingly concerned about climate change and makes lifestyle choices to reduce its carbon emissions, subsequently relocating to New York to facilitate these choices, while the other household's preferences remain unchanged. In this example, the first household's carbon emissions would have changed to some extent even if it had stayed in Pittsburgh, and that portion of the change would be incorrectly attributed to the New York place effect.

A standard approach for ruling out endogenous moves is to test for parallel trends between movers and stayers prior to the move. Unfortunately, I observe the majority of my sample only twice, which makes this impossible. Instead, I use data from the Panel Study of Income Dynamics (PSID), over the same sample period, and assess whether movers in the PSID exhibit any changes to energy expenditures prior to their move. While I do not know where households move from or to, I find that energy expenditures increase following a move, consistent with life-cycle trends presented in Table 2 of people moving to places with larger homes and fewer non-car transportation amenities, and with the secular trend over my sample frame of people moving to places with higher cooling needs. Importantly, the data show no pre-trend in energy expenditures leading up to a move (Figure G.6).

In Section 5.1, I will also show evidence on preference "drift." If moves were endogenous to evolving preferences, the selection component, and consequently, the parameter estimate would likely increase with the duration since the move. Some evidence of drift emerges in the baseline analysis, but the magnitude of drift appears to be quite small, and becomes insignificant when looking at a restricted sample with no observable major life events.

4.2 Event Study Decomposition

The first decomposition I estimate is an event study, following the approach used by Finkelstein, Gentzkow, and Williams (2016). The event study begins with the two-way fixed effect model defined in Equation 2, but summarizes heterogeneity with a single parameter, rather than the full distribution of J place effects. For household i moving from origin o to destination d, the expected change in carbon emissions is given by:

$$E[\ln CO_{2it}(d) - \ln CO_{2it}(o) | \alpha_i, X_{it}, \tau_t] = \psi_d - \psi_o$$

I then express the change in place effects as a share of the differences between observational means:

$$\begin{aligned} \psi_d - \psi_o &= \frac{\psi_d - \psi_o}{\bar{y}_d - \bar{y}_o} \cdot (\bar{y}_d - \bar{y}_o) \\ &\equiv \theta_{o,d} \cdot (\bar{y}_d - \bar{y}_o) \end{aligned}$$

Incorporating this expression into the two-way fixed effect model yields the event study equation, which I use to estimate θ, the share of between-place differences attributable to place effects:

$$\begin{aligned} \ln CO_{2_it} &= \alpha_i + \psi_j + \tau_t + X_{it}\beta + \varepsilon_{it} \\ &= \alpha_i + \psi_o + \mathbb{1}[moved] \cdot (\psi_d - \psi_o) + \tau_t + X_{it}\beta + \varepsilon_{it} \\ &= \tilde{\alpha}_i + \mathbb{1}[moved] \cdot \theta \cdot (\bar{y}_d - \bar{y}_o) + \tau_t + X_{it}\beta + \varepsilon_{it} \end{aligned} \tag{7}$$

By characterizing the place share of heterogeneity with a single parameter, θ, the event study approach vastly reduces the dimensionality of the estimation problem. This efficiency comes at the cost of an additional assumption: In order for $\hat{\theta}$ to reflect an unbiased, causal parameter, it cannot be correlated with other parameters in the model. In other words, unlike the two-way fixed effect model, this approach does not permit systematic sorting of certain types of households, either based on observable characteristics or unobservable types, to certain types of places. This is because it infers place types from observational means; if there were sorting of, for example, high type households to high type places, $\bar{y}_d - \bar{y}_o$ would grow faster than $\psi_d - \psi_o$ as places grow more different, leading to biased estimates.

The stronger restriction on sorting is more plausible at the CBSA level, in which people are more likely to move for job opportunities or to be close to family, than at the neighborhood level, where choices are more likely to be driven by local amenities. In Section 5.1 I will provide evidence that estimates of the share parameter are not heterogeneous along several dimensions, suggesting little bias from this stronger assumption. However, even under weaker baseline assumptions, under which event study results can't be interpreted as causal, the results remain informative for two reasons. First, they serve as useful descriptive evidence and additional model intuition and validation for the KSS analysis to come. Second, they yield unbiased predictions about how household carbon emissions will change for any set of observed moves. This is particularly useful, as researchers have highlighted that restrictive zoning and high cost of living in productive areas drives people to higher emissions locations (Glaeser and Kahn 2010). The labor and urban literatures have identified that such moves lead to decreased welfare due to loss of agglomeration externalities. The event study specification makes it possible to additionally estimate the carbon emissions externality of these regulatory restrictions.

4.3 Variance Decomposition

The second decomposition is based on the full set of nonparametric fixed effect estimates from the model. Specifically, heterogeneity in household carbon emissions can be decomposed as below, lumping τ_t with X_{it} for brevity:

$$\begin{aligned} Var(y_{ij}) &= Var(\psi_j) + 2 \cdot Cov(\alpha_i, \psi_j) + Var(\alpha_i) \\ &\quad + Var(X_{it}\beta) + 2 \cdot Cov(\alpha_i, X_{it}\beta) + 2 \cdot Cov(\psi_i, X_{it}\beta) + Var(\varepsilon_{it}) \end{aligned} \tag{8}$$

This analysis focuses on the first three terms: the variance component of place effects, the variance component of unobserved person effects, and their covariance, which captures the spatial heterogeneity that results from systematic sorting on unobserved preferences. Each variance component describes the share of *overall* heterogeneity attributable to the relevant component.

In contrast to the event study decomposition, the two-way fixed effects decomposition allows unrestricted sorting of households across places; it imposes no limitations on the magnitude or sign of the covariance terms. This flexibility comes at an econometric cost. A well-documented challenge to estimating variance components in two-way fixed effect models is limited mobility bias (Andrews et al. 2008). Limited-mobility bias arises because place effects are estimated based on the outcomes of people who move between different locations; however, for any given place, there might only be a small number of people moving in or out. This can lead to imprecise estimates, which creates an upward bias in the naive plug-in variance estimate relative to the true variance of place effects, even if estimates of place effects themselves are unbiased. To address this, I estimate variance components using the heteroskedasticity-unbiased leave-out estimator proposed by Kline, Saggio, and Sølvsten (2020), henceforth KSS. The KSS estimator uses a leave-out estimate of standard errors to correct estimates of the variance components for sampling variability.

I implement the leave-out estimator at the household level, leaving out all observations corresponding to a household match. In the mover sample, the KSS estimator is robust to unrestricted heteroskedasticity and serial correlation within each match. Because it is not possible to leave out matches for stayers without dropping all their observations, if there is serial correlation in the error term, KSS estimates of the person variance component in the panel sample are an upper bound on the true value. See Appendix E for additional computational details, and KSS for a complete discussion of the leave-out estimator.

5 Results

This section presents the core results of my paper: estimates of the share of heterogeneity in household carbon emissions attributable to place effects. I begin by showing results from the event study specification, which – even if the stronger assumptions on selection are violated

– serve as additional descriptive evidence and can be used to predict how household carbon emissions will change for movers under existing patterns of mobility. I then present results from the variance decomposition of the unrestricted two-way fixed effect model. I conclude the section with a discussion on interpreting the two versions of the analysis.

5.1 Event Study Decomposition

This section presents estimates from the event study derived in Section 4.2

$$\ln CO_{2_{it}} = \tilde{\alpha}_i + \mathbb{1}[moved] \cdot \theta \cdot (\bar{y}_{d_{-i}} - \bar{y}_{o_{-i}}) + \tau_t + X_{it}\beta + \varepsilon_{it}$$

where $\bar{y}_{j_{-i}}$ are sample means estimated from the full sample, leaving out the household observation.[14]

Table 3 presents event study estimates of $\hat{\theta}$, which captures the share of differences between observational place means attributable to differences between place effects. Columns (1)-(3) show results from regressions examining CBSAs while columns (4)-(6) show results from regressions examining tracts. The top panel of the table shows results estimated from the panel sample, which consists of both movers and stayers; the bottom panel shows results using just the mover sample. Both samples use mover variation to identify place effects, but estimates may differ if the relationship between household characteristics and emissions varies systematically between movers and stayers. For instance, if having children induces divergent preference shocks – prompting movers to seek larger, potentially higher-emission homes while reinforcing stayers' preference to be within walking distance of friends and local amenities – including stayers in the analysis could underestimate the direct impact of children on movers' emissions, thereby overstating the place effect. Across specifications, the mover sample yields place share estimates that are at most four percentage points lower than the panel sample estimates. This implies that movers and stayers have only marginally different responses to changes in observables, and including stayers introduces a small upward bias in the place effect estimates.

Columns (1) and (4) present estimates from a model with just year fixed effects. These estimates show that when a household moves, its emissions change by 89-91 percent of origin-destination CBSA mean differences, and 76-78 percent of origin-destination tract mean differences. However, these estimates are confounded by the correlation between household characteristics and location choices. Columns (2) and (5) address this by controlling for observable household characteristics (age, household income, household size, number of children, and homeowner status), that likely influence both emissions and neighborhood choice. This adjustment reduces CBSA share estimates by up to five percentage points (to 85-86 percent) and tract share estimates by nearly 20 percentage points (to 57-60 percent). The larger impact on tract-level estimates suggests that, as you would expect, household sorting plays a more significant role in neighborhood-level emissions variation, while CBSA-level moves may be driven more by factors

14. To the extent that there is sampling variability in the distribution of observational means, my estimate of the relationship between origin-destination mean changes and individual changes in $\log CO_2$ may be biased. In practice, using a linear Empirical Bayes estimator to adjust means for sampling variability does not materially change the results.

like job opportunities or family considerations. In Table G.6, I re-estimate columns (2) and (5) controlling for average heating degree days, cooling degree days, and electricity emissions factors. Doing so decreases the CBSA-share to 68-70 percent, and the tract share to 51-54 percent. While these estimates lend insight into the mechanisms underlying place effects, I focus my main event study analysis on total place effects, which are more relevant for predicting emissions changes from observed migration patterns.

Table 3: **Share of Spatial Variation in Mean CO_2 Attributable to Place Effects**

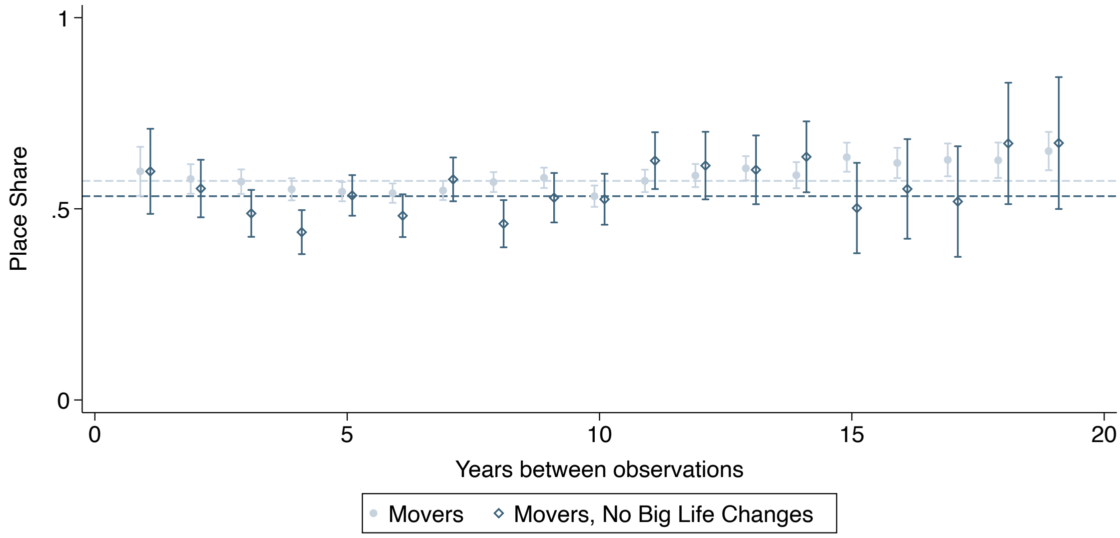
	CBSA			Tract		
	(1)	(2)	(3)	(4)	(5)	(6)
A: Panel Sample						
Place share of mean difs.	0.91***	0.86***	0.86***	0.78***	0.60***	0.55***
	(0.008)	(0.007)	(0.016)	(0.003)	(0.003)	(0.007)
N	1,764,000	1,764,000	633,000	1,710,000	1,710,000	613,000
R^2 (adj.)	0.73	0.75	0.77	0.74	0.76	0.77
B: Mover Sample						
Place share of mean difs.	0.89***	0.85***	0.84***	0.76***	0.57***	0.53***
	(0.009)	(0.009)	(0.020)	(0.004)	(0.004)	(0.009)
N	191,000	191,000	36,000	508,000	508,000	102,000
R^2 (adj.)	0.64	0.70	0.70	0.69	0.73	0.73
Household controls		X	X		X	X
No big life events			X			X

Note: This table reports event study estimates of the place share of spatial heterogeneity in household carbon emissions. The place share estimate ($\hat{\theta}$) represents the proportion of differences in average carbon emissions ($\bar{y}$) between a mover's origin and destination attributable to place effects. Panel A reports estimates from the panel sample, while panel B restricts the sample to movers only, allowing for systematic differences between movers and stayers. Columns (1) and 4 show estimates from CBSA and tract moves, respectively, with only year fixed effects. Column (2) and (5) add controls for the standard set of household characteristics. Columns (3) and (6) restrict the sample to households without changes in number of children, income changes exceeding 50 percent, or changes in homeownership. All estimates use Census sample weights.

Even after accounting for the rich set of observable characteristics available in the data, household destination choices may still be influenced by unobserved heterogeneous preference shocks. While I cannot rule this out entirely, I explore the potential bias from such selection by restricting the sample to households with minimal observable shocks, motivated by the premise that these households are less likely to have experienced large concurrent unobservable changes. Columns (3) and (6) present results from this "no big life events" subsample, which excludes households that experienced a change in the number of children, a greater than 0.5 log point change in income, or a change in homeownership status between observations. This subsample, comprising approximately 20 percent of the original observations, yields place share estimates at most five percentage points lower than the baseline estimates. The fact that households least likely to have experienced unobservable shocks yield qualitatively similar place share estimates to the baseline sample is reassuring and suggests limited bias from shocks to unobserved preference heterogeneity.

In addition to unobserved shocks, another potential source of bias arises from "preference drift" gradual changes in households' preferences over time that are not fully captured by changes in age or other observable characteristics (Card, Heining, and Kline 2013). If such drift exists and influences relocation decisions, place effect estimates would conflate true causal effects with selection, particularly for moves observed after long time intervals. This would manifest as place effect estimates that trend upward with the duration between observations. To explore this possibility, Figure 4 presents tract-level place share estimates by duration between observations, controlling for demographic and household characteristics. The light gray points depict estimates for the baseline mover sample, while the dark blue bars correspond to the restricted subsample of movers with no big life changes between observations.

Figure 4: **Place Share of Spatial Variation in Mean CO_2, over Time**



Note: This figure shows event study estimates of the share of spatial variation in mean carbon emissions that can be explained by place effects, by duration between mover observations. In other words, each coefficient is the estimate for place effects generated from the sub-sample of households that I observe x years apart. Coefficients plotted in light gray are estimated from the model using the full panel of stayers and movers. Coefficients plotted in the dark blue are estimated from the model using the sub-sample of stayers and movers with no changes in the number of children, less than 50 percent change in household income, and no change in homeownership status between observations. All estimates are weighted using Census sample weights.

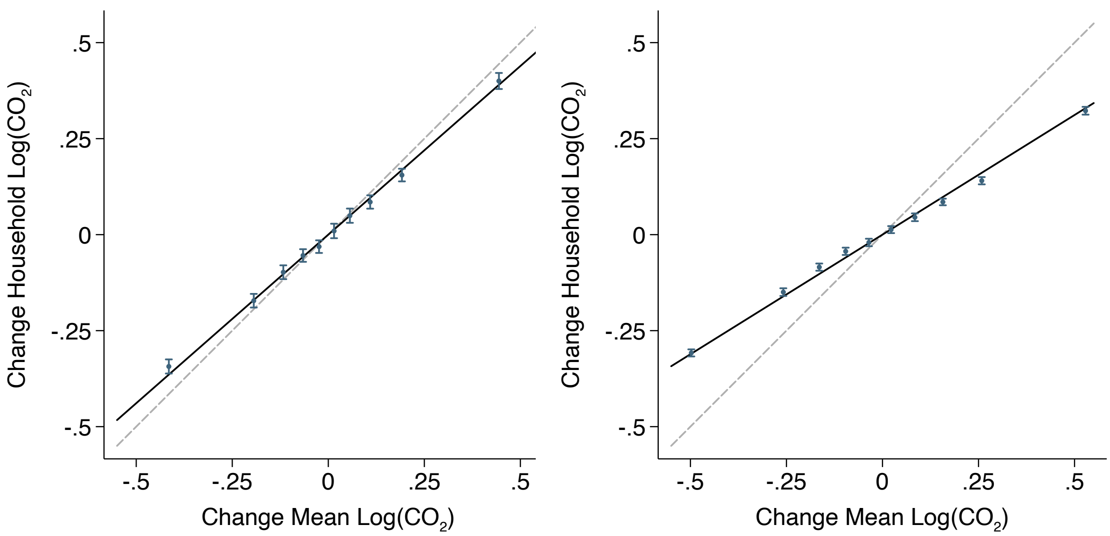
For the full mover sample, a slight upward trend emerges, with place share estimates for durations in the 15-19 year range exceeding the pooled estimate. This pattern dissipates when considering the restricted subsample of movers without observable big life changes; the pooled estimate for this group is not contained within the 95% confidence interval for four out of the 19 duration-specific estimates, but with no clear trend across durations. Figure G.7 presents analogous CBSA estimates, which exhibit a similar pattern, except the upward trend observed in the longer-duration coefficients become insignificant in the baseline mover sample as well.

While these results suggest some preference drift, potentially leading to upward-biased place share estimates, the magnitude appears small relative to observed changes in emissions between places, and this bias appears to be negligible in the sample restricted to households without big life events. One additional result that emerges from this analysis is that household carbon emissions appear to change instantaneously, suggesting place effects stem from attributes that directly impact emissions rather than gradual influences like peer effects or habit formation.[15]

As a final specification check, I explore heterogeneity across moves of different magnitudes in terms of place mean changes. A potential concern with the results presented thus far is that households choosing to relocate from low-emissions places to high-emissions places (or vice versa) may systematically differ from those moving between areas with similar emissions levels. If a household's choice of destination reveals newly observable information about unobserved preference shocks, the estimates would be biased, mistakenly attributing these shocks to place effects. To explore this possibility, Figure 5 depicts changes in mover households' carbon emissions across deciles of origin-destination mean differences, controlling for demographic and household characteristics. The x-axis represents conditional changes in mean emissions between origin and destination, while the y-axis shows to conditional changes in household emissions. The 45-degree line (in gray) represents a scenario in which place effects fully explain variation in carbon emissions across locations. The solid line's slope corresponds to the pooled estimate of the relationship between tract-level mean changes and household-level changes. Standard errors are estimated using bootstrap.

Figure 5: **Place Share of Spatial Variation in Mean CO_2 , by Move Type**

(a) CBSA (b) Tract



Note: This figure shows mover changes in household carbon emissions, by size of origin-destination differences in mean carbon emissions. All estimates are from a model that conditions on changes in observable household characteristics and year fixed effects. I split movers into ten deciles, according to the size of the (conditional) gap in mean carbon emissions across their origin and destination, and estimate standard errors using bootstrap. Point estimates are shown with 95% confidence intervals. The solid lines show the regression estimates from the pooled model, and the dotted gray line denotes 45°, i.e. the scenario in which moving to on average higher or lower emissions places leads to a one-for-one increase in own carbon emissions. All estimates are weighted using Census sample weights.

15. While I don't observe how long ago a household moved, the expected time since moving increases with the duration between observations.

I find that, for both moves across CBSAs and moves across neighborhoods, changes in household carbon emissions as a share of changes in origin-destination means are symmetric and linear across move types. In other words, the share of spatial heterogeneity attributable to place-based differences is consistent for moves from low- to high-emissions places (rightmost points), moves from high- to low- emissions places (leftmost points), and moves between places with similar average emissions (central points). The symmetry and stability of share estimates have several implications. First, this result extends the symmetry check presented in Figure 3, further validating the log-linear model specification and suggesting that the two-way fixed effect model reasonably approximates household carbon emissions. Recall that this implies place effect estimates are unbiased even if I never observe some household types moving to some place types. Second, this result indicates that the event study estimates presented in Table 3 are not driven by a subset of movers or mover destinations, alleviating concerns that estimates primarily reflect changes in the emissions of households who moved to a vastly different destination in response to a large unobserved preference shock. Lastly, this result provides another dimension along which heterogeneity in the estimated share parameter appears to be limited.

In summary, the event study estimates indicate that movers' carbon emissions change by over half of origin-destination differences in neighborhood means (and about 85 percent of origin-destination differences in CBSA means) when they move. Under strong assumptions on sorting, this can be interpreted as a causal share parameter, which would determine the change in carbon emissions resulting from any household moving between any pair of places. I estimate that heterogeneity in the share parameter is limited along several dimensions explored throughout this section – estimates are within a 5 percentage point range of each other across subsamples with vastly different changes to observed characteristics, across different observation time horizons, and across varying move types (between similar versus dissimilar places) – suggesting that perhaps bias from violations of this assumption may also be minimal.

Under weaker assumptions, the event study estimates serve as a useful specification check for the two-way fixed effect model, and perhaps more interestingly, they provide unbiased prediction of how household carbon emissions will change for any *observed* move. Recent research and media coverage have highlighted a pattern of increasing migration from more expensive but on average lower emissions places to less expensive but on average higher emissions places (Kolko 2021; Eisen 2019). For instance, in 2019, there was a net migration of about 45,00 people from California to Texas. While not all these moves were from San Francisco to Houston, using these two cities as an illustrative example: carbon emission in Houston are about 21% higher than they are in San Francisco (Jones and Kammen 2014). Combined with the event study findings, this implies that regulatory constraints that restrict housing, increase the cost of living, and drive this pattern of migration increase household carbon emissions by 12 percent or more, imposing a sizeable carbon externality.[16]

16. The 12 percent figure would correspond to treating San Francisco and Houston as neighborhoods. If we were to treat them more like CBSAs, the predicted increase in emissions would be closer to 19 percent.

5.2 Variance Decomposition

In this section, I weaken the restriction on sorting imposed by the event study, and present estimates from the variance decomposition shown in Equation 8

$$
\begin{aligned}
Var(y_{ij}) = Var(\psi_j) + 2 \cdot Cov(\alpha_i, \psi_j) + Var(\alpha_i) \\
+ Var(X_{it}\beta) + 2 \cdot Cov(\alpha_i, X_{it}\beta) + 2 \cdot Cov(\psi_i, X_{it}\beta) + Var(\varepsilon_{it})
\end{aligned}
$$

For each sample and specification, Table 4 presents the overall variance of the outcome, $\log(CO_2)$, the share of variance attributable to each of the unobserved heterogeneity components (place effects, ψ_j and household effects, ψ_i), the correlation between the unobserved heterogeneity components, and the bias-corrected standard deviation of place effects. The top panel presents estimates from the entire panel of movers and stayers, while the bottom panel presents estimates from the mover-only sample.

Columns (1) and (5) present the baseline analysis, which includes year fixed effects and the standard vector of household controls. I estimate that CBSA effects account for 14-16 percent of overall heterogeneity (column 1), and tract effects account for 22-23 percent of overall heterogeneity (column 5). Columns (2) and (6) introduce controls for mean heating degree days, cooling degree days, and log electricity emissions factors, excluding their impact on household emissions from estimated place effects. This specification not only provides insight into the mechanisms driving place effects, as in the event study analysis, but also allows for more precise counterfactual analysis within the KSS framework. When considering interventions that impact household carbon emissions through changes to local amenities and the built environment, isolating the place effect component not driven by climate or electricity generating sources may be more relevant, though the appropriate specification remains context-dependent. For instance, a city-wide initiative to install solar panels on parking lots, potentially coupled with EV charging stations, could decrease the average emissions intensity of electricity while promoting EV adoption. Moreover, while regional climate is, to first order, exogenous to local actions, there are many interventions cities and neighborhoods could enact to decrease heat island effects and in turn reduce local heating degree days (Druckenmiller 2023).

I find that controlling for climate and electric grid intensity together decreases the CBSA share of spatial heterogeneity by roughly 10 percentage points, or by more than half, to 4-7 percent of overall heterogeneity. At the neighborhood level, controlling for climate and electric grid intensity decreases the place share of heterogeneity by roughly 6-7 percentage points (column (6)), by less than half, leaving the remaining neighborhood attributes explaining approximately 15 percent. In columns (3) and (7) I additionally partial out variation driven by prices, using a price index constructed from interacting local lagged fuel shares with national retail prices. This does not further change the results at either the CBSA or tract level.

Table 4: **Unobserved Heterogeneity in CO2 – Variance Decomposition**

	CBSA				Tract		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel Sample							
Variance of log(CO_2)	0.31	0.31	0.31	0.31	0.31	0.31	0.31
Share attributable to places	0.16	0.07	0.08	0.17	0.23	0.15	0.15
Share attributable to hhs	0.50	0.50	0.50	0.30	0.36	0.36	0.36
Corr. of place and hh effects	0.01	0.03	0.03	0.02	0.02	0.08	0.11
SD of place effects	0.23	0.15	0.16	0.23	0.26	0.22	0.22
Mover Sample							
Variance of log(CO_2)	0.35	0.35	0.35		0.33	0.33	0.33
Share attributable to places	0.14	0.04	0.04		0.22	0.16	0.16
Share attributable to hhs	0.14	0.16	0.17		0.10	0.10	0.10
Corr. of place and hh effects	0.07	0.08	0.08		0.08	0.16	0.18
SD of place effects	0.22	0.12	0.12		0.27	0.23	0.23
Climate + Electricity CO_2		X	X			X	X
Price Index			X				X
Time-Varying FEs				X			

Note: This table reports results from the heteroskedasticity-robust KSS estimation of variance components. For each sample and specification, the table reports the overall outcome variance, the share of variance attributable to place effects (ψ_j), the share of variance attributable to household effects (α_i), the correlation between place effects and household effects (which captures how much households sort on unobserved characteristics), and the bias-corrected standard deviation of place effects. All specifications include year fixed effects and the standard set of household controls used throughout the paper. Columns (1) and (5) report the baseline variance decompositions at the CBSA and tract levels. Columns (2) and (6) add controls for local mean heating degree days, cooling degree days, and log electricity emissions factors. Columns (3) and (7) additional control for a price index, constructed from lagged fuel shares interacted with national retail prices. Finally, column (4) computes time-varying CBSA place effects using 5-year periods (2000-2004, 2005-2009, 2010-2014, and 2015-2019), using stayer observations across periods to identify time variation in place effects, while movers, as before, identify cross-sectional variation.

In Appendix Table G.7, I control separately for electricity emissions factors and climate, and I find that the majority of the decline in place share observed in Columns (2) and (6) comes from controlling for electricity emissions factors. It is well understood that climate has a robust effect on energy demand (e.g. Goldstein, Gounaridis, and Newell 2020; Levinson 2016), so why the disparity? Many places that are otherwise low-emissions are in regions that have made efforts to decarbonize their electricity (Murray and Maniloff 2015; Petek 2020) so there is a positive correlation between electricity emissions factors and high other emissions attributes. In contrast to this, I estimate a negative correlation between having above-average heating degree days and low other carbon emissions attributes; many colder cities are in the Northeast, a region which disproportionately contains older, denser cities and neighborhoods (Tomer et al. 2021). These correlation terms enter the place effect variance component, amplifying it when electricity emissions factors are included in the place effect, and diminishing it when heating and cooling degree days are included.

Across all three specifications discussed so far, estimates of the place share of heterogeneity are quite comparable between the panel and mover sample. This is not true of the estimates of the household components; a comparison between the panel sample and the mover sample reveals that the contribution of unobserved household characteristics to overall heterogeneity is highly sensitive to which sample the model is estimated on. In the panel sample, unobserved household heterogeneity accounts for 50 percent of overall heterogeneity when defining place at the CBSA level, and 36 percent when measuring place at the neighborhood level. Using the mover-only sample substantially decreases the unobserved household contribution across specifications, to 14-17 percent in the CBSA specification and 10 percent in the tract specification. The correlation estimates between unobserved place and household characteristics are also sensitive to sample choice, increasing across all specifications by up to eight percentage points when switching from the panel to the mover sample. This suggests some assortive matching of household types to place types (especially neighborhoods), but in the specification where the correlation coefficient is largest, it remains relatively low, at under 20 percent. The share of overall heterogeneity attributable to matching implied by the covariance term (which is not shown in Table 4 but can be computed from the correlation coefficient and variance components) is less than 5 percent across all specifications.

The panel and mover samples may yield different estimates due to inherent differences between stayers and movers, or KSS's inability to correct for serial correlation in stayers' error terms. To distinguish between these causes, I compare the KSS decomposition estimates with an uncorrected AKM decomposition (Abowd, Kramarz, and Margolis 1999), presented in Table G.8. If differences between stayers and movers drive the discrepancy, both KSS and AKM estimates should reflect this pattern. Conversely, if serial correlation in stayers' error terms is the main factor, AKM should show similar relative contributions of unobserved household heterogeneity across samples (both inflated due to limited mobility bias), with significant differences emerging only after the KSS correction. The AKM estimates reveal relatively stable place and household shares across panel and mover samples, with only minor reductions in the household component's relative size in some specifications. This contrasts with the more dramatic shifts observed in the KSS analysis, suggesting that KSS estimates of household variance components are likely more reliable in the mover sample. Panel sample estimates likely represent an upper bound, with upward bias primarily driven by serial correlation in stayer error terms. Appendix A.2 discusses potential sources of this serial correlation, many arising from the survey nature of the data, and their implications for result interpretation. Additionally, Table G.9 evaluates the sensitivity of results to alternative outcome definitions in the KSS decomposition.

Up until now, I have been estimating place effects pooling over the entire sample timeframe; however, place effects may evolve over time in ways that differ from national average trends in carbon emissions. Local or state governments particularly concerned about climate change may enact regulatory changes or make place-based investments aimed at reducing emissions for their residents. Changes to place effects could also arise from local or regional planning initiatives motivated by factors completely unrelated to decisionmakers' climate objectives. For instance, the Phoenix metropolitan area – one of the fastest growing metropolitan areas in the US – grew by nearly 1.6 million residents between 2000 and 2020. This period of growth has been accompanied by a mix of suburban expansion, urban development, the opening of a new light rail system, and several highway expansions (Maricopa Association of Governments 2020). To allow for such place-specific changes, I follow Lachowska et al. (2023) and estimate time-varying fixed effects ψ_{jt} at the CBSA level, using stayers to identify variation across time within place.[17] To maintain connectivity in my set of places, and because for the most part places evolve slowly, I define the time-varying place effects using 5-year intervals. Thus, there's a different time-varying place effect for each period 2000-2004, 2005-2009, 2010-2014, and 2015-2019. Results are shown in column (4) of Table 4 – allowing CBSA effects to evolve increases their variance share by only one percentage point relative to the baseline specification. This implies that changes in places over my sample period are either mostly captured by national secular trends, or are unrelated to household carbon emissions. In Appendix F, I provide some descriptive results on the nature of changing time-varying place effects in my sample. I also discuss some of the ways in which place effects may have evolved following the COVID-19 pandemic and the ensuing transition to remote work, which was accompanied by a steep decline in commuting and a shift towards larger homes to accommodate home offices (Van Nieuwerburgh 2023; D'Lima, Lopez, and Pradhan 2022).

As a final specification test, in Appendix Figure G.9, I show binned scatter plots similar to those presented in Section 5.1, but now with deciles of changes in estimated place effects, rather than observational means, on the x axis. I plot these against two sets of changes in household mean outcomes: changes for the full mover sample, and changes in the sample restricted to only households with no big life events. In a correctly specified model, changes in place effects should lead to one-for-one changes in household carbon emissions, though attenuation bias from noisily estimated place effects is expected to decrease the slope by some. This is roughly what I observe. Moreover, I find no discernible difference between the primary sample and the subsample of households experiencing no big life events, which provides additional reassurance that selection on heterogeneous preference shocks isn't a first order threat to identification in my analysis.

5.3 Interpreting Decomposition Results

How do the event study and KSS results inform one another? First, recall that event study estimates are unbiased only if heterogeneity in the share parameter is uncorrelated with observed and unobserved household characteristics. I have shown evidence that the event study estimates are fairly stable across several observable dimensions of heterogeneity in the data. I also showed in the KSS decomposition that the covariance between unobserved components of heterogeneity is relatively small – the largest correlation coefficient across the four baseline estimates corresponding to the specifications examined in the event study is .08. Together, this evidence suggest that bias from this assumption on selection should be minimal.

17. Because tracts by definition consist of many fewer observations than CBSAs, including a time varying component introduces either substantial noise, or a substantial geographic restriction to only the most populated tracts, so I do not estimate time-varying tract effects. Studying neighborhood-level changes is an important direction for future study.

Second, even when unbiased, the event study yields estimates of shares of mean differences *between* places attributable to place effects, while the KSS estimates yield a variance decomposition of *overall* variation, and this can lead to meaningful discrepancies in magnitudes. To help illustrate this, consider a simplified version of the two-way fixed effects model, where $y_{it} = \alpha_i + \psi_j + \varepsilon_{it}$. Song et al. (2019) show that unobserved heterogeneity can be decomposed further into a between-place component $Var_j(\bar{y}_j)$, which captures the variation in mean household carbon emissions across places, and a within-place component $Var_i(y_{it} - \bar{y}_j | i \in j)$, which captures the heterogeneity in carbon emissions of households living in the same place:

$$\begin{aligned} Var(y_{it}) &= Var_j(\bar{y}_j) + Var_i(y_{it} - \bar{y}_j | i \in j) \\ &= \underbrace{Var(\psi_j) + 2 \cdot Cov(\bar{\alpha}_j, \psi_j) + Var(\bar{\alpha}_j)}_{\text{Between}} + \underbrace{Var(\alpha_i - \bar{\alpha}_j) + Var(\varepsilon_{ij})}_{\text{Within}} \end{aligned} \tag{9}$$

Equation 9 highlights that heterogeneity between places reflects variation in place effects, sorting of certain types of households to certain types of places (the covariance term), and "segregation" of households – the extent to which households of different types segregate across places, whether or not this pattern reflects systematic sorting on place types.[18] In addition to the between-place heterogeneity, overall heterogeneity reflects heterogeneity in household carbon emissions within places, as well as heterogeneity that cannot be explained by the two-way fixed effects model.

For intuition, imagine two places, one ψ_{low} and one with ψ_{high}, with identical populations. If there is high variation in carbon emissions within populations and a small difference between ψ_{low} and ψ_{high}, the event study would yield a share coefficient of one (since populations are identical across places, all between differences are driven by place effects), but the KSS decomposition would yield a place variance component of close to zero (because of a large within component to the variance). In practice, this is very close to what happens at the CBSA level – the vast majority (85 percent) of differences between CBSAs can be attributable to variation in place effects and not household attributes, but there is much more variation in household carbon emissions within CBSAs than there is across, leading to a variance component of approximately 15 percent in the KSS estimation (more than half of which is attributable to climate and electric grid intensity). At the neighborhood level, household sorting contributes more to variation between places, dropping event study estimates of the place share to roughly 55-60 percent. Accounting for variation within neighborhoods in the denominator decreases the neighborhood share to 22-23 percent of overall heterogeneity, or about 15 percent when excluding climate and electric grid intensity from place effects.

6 The Characteristics of Low and High-Emissions Places

With estimates of place effects in hand, I proceed to characterizing high and low-emissions places. As highlighted in the conceptual model, place effects reflect a mix of differences in demand for energy services, energy prices, energy demand elasticities, fuel mixes, and emissions factors. The

18. $\bar{\alpha}_j \equiv E[\alpha_i | i \in j]$

urban and planning literature has identified many local amenities that could contribute to differences in average household energy demand and energy demand elasticities. In the residential energy sector, larger homes tend to use more energy, as do single-family homes; there is a strong relationship between carbon emissions and density, though it is potentially non-monotonic due to the effect of density on micro-climates; and parks, plants, and green surface coverage are all negatively correlated with energy use (see e.g. Ko 2013, for a review). In the transportation sector, car use is lower in places with better public transportation, less parking, and more direct road connectivity (e.g. Transportation Research Board 2009; Barrington-Leigh and Millard-Ball 2017), and mechanically, people drive fewer miles when they live closer to where they work, shop, and spend their leisure time. Many of these amenities are intertwined and simultaneously relate to residential and transportation emissions, underscoring the value of studying these sectors together. For instance, higher density not only implies smaller homes but also enables more efficient public transportation networks by reducing distances between transit stops and destinations.

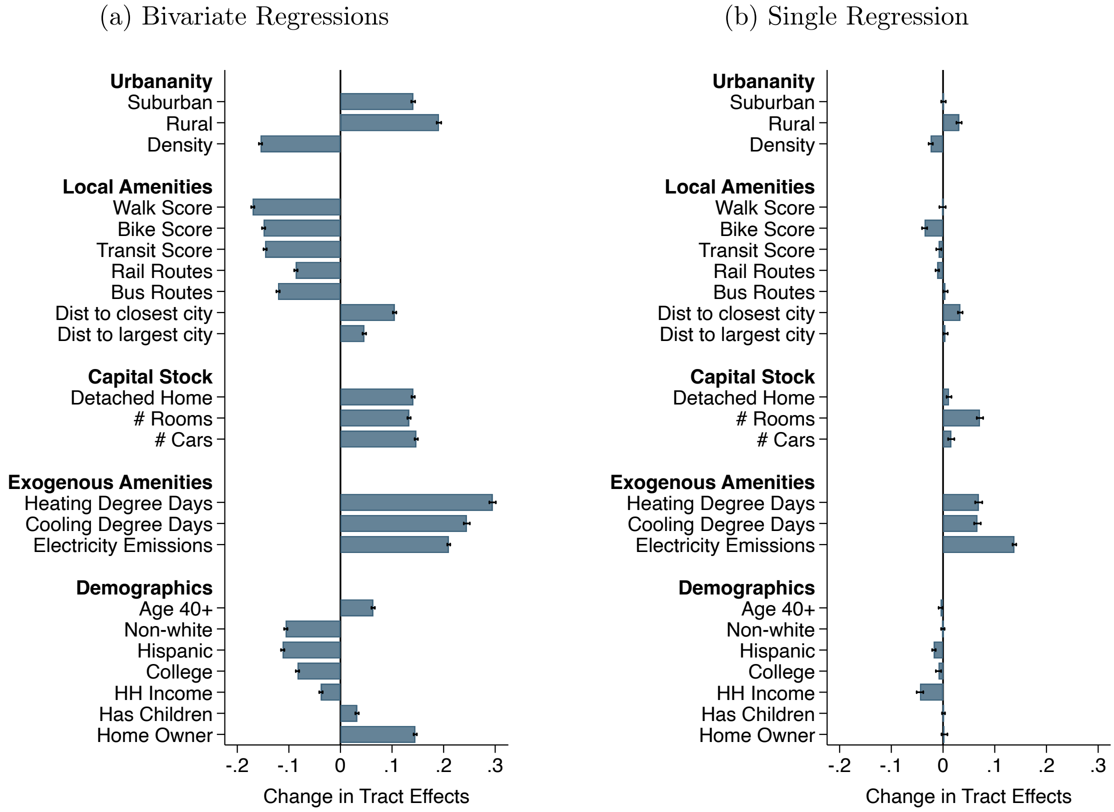
In Figure 6, I show the results from regressions of estimated tract effects onto a set of observational tract level characteristics. In the left panel, I show coefficients from bivariate regressions of tract effects onto each of the shown tract-level characteristics independently (except for suburban and rural, which are estimated in a single regression), while in the right panel, I show coefficients from a single regression onto all of the characteristics.[19] Observable characteristics are all measured at the tract level, and are normalized across tracts to have mean zero and a standard deviation of one (except for suburban and rural, which are retained as indicators). Thus, coefficients should be interpreted as showing approximately the average percent change in carbon emissions associated with a one standard deviation change in neighborhood-level observable characteristics.

Observable characteristics fall into five categories. Both "urbanity" and "local amenities" cover aspects of urban form that households effectively taken as given, but that vary neighborhood to neighborhood. These include whether a tract is urban, suburban, or rural, its density, measures of local public transportation amenities (walk scores, bike scores, transit scores, number of nearby rail routes and bus routes), and measures of sprawl (captured by geodesic distance between tract centroids and the centroid of the closest city and the largest city within the CBSA). "Capital stock" characteristics reflect household choices among the set of options available in a place, which are determined at least in part by local zoning regulations or other policies. These include the share of single-family detached homes, average house sizes, and the number of cars in a household. "Exogenous amenities" are predominantly shaped by factors beyond local control, and households also take these as given. These include annual heating degree days, annual cooling degree days, and electric grid intensity.[20] Finally, the "demographics" category captures tract level variation in household observable characteristics.

19. I used a LASSO regression with 10x crossfold validation to select characteristics, and it retained all of the variables.

20. As mentioned earlier, neither climate nor electricity emissions factors are impervious to local influence. Urban form can quite meaningfully impact local micro climates through urban heat island effects, but in this paper my measure of climate is at the NOAA climate division level, a geographically coarser definition. And municipal utilities can shape local electricity emissions factors through energy procurement choices, but currently make up a small share of overall electricity supply.

Figure 6: **Correlates of Tract Effects**



Note: This figure presents estimates from OLS regressions of estimated tract effects on a set of observable place-based and household characteristics. Panel (a) shows results from separate bivariate regressions, while panel (b) shows results from a single regression on all covariates. All amenity variables are tract level means, normalized to have mean zero and standard deviation one, except the rural and suburban indicators, which are retained as indicators. Regressions are weighted using ACS sample weights. Thus, estimated coefficients reflect the average change in place effect – and in turn the percent change in carbon emissions, given the log specification of the model – associated with a one standard deviation change in observable characteristics.

The results of the bivariate regressions (Panel A) are consistent with observational data and conclusions in the urban planning literature. Higher-emissions neighborhoods tend to be less urban and dense, with poorer walkability, bikability, and public transit options, and are typically situated farther from city centers within their CBSA. Higher-emissions neighborhoods feature a higher proportion of detached and larger homes, with households owning more cars on average. They experience less mild climates, and have higher electricity emissions factors. Finally, looking at the projection of tract effects onto demographics, I find that non-white, Hispanic, college-educated, and higher income households are more likely to live in low-emissions tracts, while older households, households with children, and homeowners are more likely to live in high-emissions tracts.

When including all of these characteristics in a single regression (Panel B), directionally the relationships remain the same, but many of the coefficient magnitudes decrease substantially. Electricity emissions factors, heating degree days, and cooling degree days emerge as three of the four strongest correlates of place effects. The second largest correlational coefficient is on the number of rooms in a house; a standard deviation change in tract mean number of rooms, is associated with a roughly seven percent change in tract level carbon emissions. The remainder of tract effects appear to load primarily onto measures of sprawl, the quality of local bike infrastructure, rurality, density, and household income, with standard deviation changes in these characteristics being associated with a 2-4 percent change in carbon emissions each. It is notable that measures of walkability, transit quality, share of detached homes, and number of cars appear to become less important after conditioning on density, sprawl, and home size. It is possible that this is because these characteristics are very co-linear in the observational data; as highlighted earlier, many of these relationships are interconnected.

Appendix Figure G.10 shows analogous projections of household fixed effects onto observable neighborhood-level characteristics. I find that correlations between household effects and observable characteristics are about an order of magnitude weaker than those between tract effects and observable characteristics in both the bivariate regressions and the full regression. This is consistent with minimal sorting on unobserved household characteristics that I estimate in KSS. The largest coefficients imply that households with high unobservable potential for carbon emissions sort to tracts where people are on average higher income and have more children, where it is warmer, where electricity is higher emissions, and where houses are bigger. Appendix Table G.10 presents additional results on the correlates between observable household characteristics and observable place characteristics.

These regressions do not elucidate causal relationships. As is the case with inference about place effects as a bundle, inference about the role of specific amenities from observational data alone is likely to be biased by household sorting. Nevertheless, the observational data suggests features of urban form whose effect on household carbon emissions should be studied further with credible exogenous variation in the amenity itself. Previously, one might worry about whether a setting for such a case study is selected. Can we learn about the effect of lifting zoning restrictions and allowing for smaller lot sizes if people who live in denser cities have significantly different preferences than those who don't? Are improvements to bike infrastructure only effective in cities where a lot of people already bike? The core results of my paper suggest that a meaningful share of variation between places is driven by variation in place effects, with evidence that place effects are constant across heterogeneous populations, mitigating some of these concerns about external validity.

7 Implications for Aggregate Carbon Emissions

The wide distribution of place effects suggests that there may be an opportunity to substantially reduce household carbon emissions from residential and transportation energy through what I refer to as place-based climate policies. Place-based climate policies – distinct from regional energy sector regulations or market-based mechanisms designed to encourage decarbonization – aim to reduce household carbon emissions by altering local characteristics that shape household energy choices.[21] They could be implemented at federal or local levels, and, building on insights

21. As transportation and residential sectors electrify and the power grid decarbonizes, the impact of place-based drivers of household emissions will diminish. However, given current challenges in clean electrification, particularly transmission capacity constraints and aging infrastructure (e.g. U.S. Department of Energy 2015), place-based policies that reduce energy demand and alleviate pressure on the electricity grid could be complementary to more traditional decarbonization instruments.

from labor and urban economics, could be designed to simultaneously reduce carbon emissions and address economic development goals, housing shortages, or local externalities like pollution and congestion.

Infrastructure and local public goods investments make up one category of potential place-based climate policies. The federal government has a history of investing into transportation infrastructure, with the Interstate Highway System being its most notable example (U.S. Department of Transportation 1977). More recently, the Bipartisan Infrastructure Law provided a significant increases to transit funds across the country; the Federal Transit Administration announced in early 2024 that it would be investing $9.9 billion to support local transit systems in urban areas across the US (U.S. Department of Transportation 2024). While adding and expanding public transit networks is perhaps the most obvious example of an intervention that might decrease place effects, many cities are implementing less costly initiatives that repurpose existing urban space for community use. Paris's Plan Velo has converted over 200 miles of roads into (often protected) bike lanes as of 2024, with further expansions planned through 2026 (Ville de Paris 2021). Barcelona's Superillas program aims to transform more than half of its car-dominated streets into mixed-use public spaces (Roberts 2019), while New York City's High Line project has repurposed an abandoned elevated railway into a popular pedestrian walkway and park. The COVID-19 pandemic accelerated this trend, sparking "slow" or "open" streets programs which temporarily restricting car traffic to benefit pedestrians and cyclists in many cities of the US, with some cities now working to make these changes permanent (New York City Department of Transportation 2023; Combs 2020). Expanding the network of electric vehicle chargers is another example of a lower cost place-based infrastructure investment that can reduce carbon emissions.

A second category of potential place-based policies includes regulatory changes intended to encourage – or remove barriers previously preventing – sustainable urban development. Zoning deregulation, land use reform, and transit-oriented development have all gained traction in the US in recent years. In 2018, Minneapolis became the first city in the US to enact a city-wide ban on exclusionary zoning, a common practice across the US that restricts land to be used for single-family homes only (Mervosh 2018). In 2021, the California State Assembly passed Bills 9 and 10, which streamlined the process of "up-zoning" residential land, allowing for the development of up to four units on land previously zoned for single-family homes only and facilitating higher-density construction near transit corridors. This was followed by AB 2097 in 2022, which eliminated most parking minimum requirements for new development (Fulton et al. 2023). At the federal level, President Biden's original infrastructure bill proposal in March 2021 included grants to cities that eliminated exclusionary zoning. While this portion of the bill did not get passed, and many states and municipalities are still debating but not implementing zoning reforms, these examples illustrate the relevance of such approaches in the current policy debate. At the same time, some states and municipalities are making regulatory efforts to make building codes more stringent by, for example, enacting energy efficiency minimum requirements (Levinson 2016), or mandating electrification through natural gas hook up bans (Payne 2020).

This paper does not identify a causal relationship between any specific amenity and place effects, or any specific intervention and place effects, but it shows evidence that low-emissions places are low-emissions not simply due to sorting of low-emissions people to those places, and that low emissions places tend to have amenities that are characteristic of more urban, less sprawling, neighborhoods. Using this as motivation, I examine how carbon emissions would change if the national distribution of place effects were more urban than it currently is. Specifically, to approximate a scenario in which the US limits suburban sprawl and increases the share of households within a region that live in an urban neighborhood, I estimate the effect on emissions if households currently living within a suburban or rural area lived instead in a place with the average place effect of the nearest principal city. A naive comparison of household emissions shows that, on average, households living in principal cities emit 20 percent less than households living in surrounding areas. After accounting for sorting of households between suburban and urban neighborhoods, my estimates suggest that if every tract had the place effect of the nearest principal city, the emissions of suburban and rural households would decrease by approximately 15 percent.

The majority of principal cities in the US would not be considered particularly urban on a global scale. To consider the potential effect of deeper urbanization, I examine how household carbon emissions would change if more people lived in places like Manhattan, which is uniquely dense, walkable, and transit oriented within the US context. Specifically, I examine how the emissions of household in the principal cities of the nine largest CBSAs after the New York Metropolitan Area in the US would change if those cities developed into a place with the average place effect of Manhattan. The observational gap is enormous: even using large cities as a comparison group, households in Manhattan emit about 73 percent less than observably comparable households in the other nine large cities. Accounting for unobserved fixed differences between households reduces the gap to 60 percent; people who choose to live in Manhattan are a selected sub-sample, but their emissions are nevertheless much lower as a result of living in Manhattan than they would be if they lived elsewhere.

These exercises lend insight into how development that shifts population shares across place types by "expanding" places with lower place effects – either by making their neighbors look more like them, or by allowing more people to live in such places without changing their fundamentals – could affect emissions in the future. My estimates yield only a first-order, partial equilibrium approximation to the effect of such interventions, as in practice there would be some re-sorting of populations in response to place-based changes, which would change the distribution of household types living in each place and thereby change aggregate carbon emissions. More importantly, my estimates don't lend insight into the specific interventions that would result in the largest changes to place effects – this is a critical direction for future research. My results highlight that the potential for urbanization-induced reductions to carbon emissions is overstated when inferred from observational means across places, because people choose where to live based in part on their potential carbon emissions. Nevertheless, because there is substantial variation in household carbon emissions and neighborhood effects explain 15-23 percent of this variation – depending on whether climate and electricity emissions are included in the place effects – my results also imply that even relatively small shifts in the distribution of place effects that households are exposed to could meaningfully decrease aggregate carbon emissions.

8 Discussion

This paper is the first to estimate the causal effects of places on household carbon emissions and decompose spatial heterogeneity in carbon emissions into a component driven by these place effects and a component driven by household characteristics and sorting. I find that up to 23 percent of overall heterogeneity in household carbon emissions from residential energy use and commuting across the US can be explained by neighborhood effects, or roughly 15 percent can be explained by neighborhood effects after accounting for variation driven by climate and electric grid intensity. Paired with high overall heterogeneity, these estimates imply that interventions that change the population weighted distribution of place effects across the US, either through direct changes to places, or through regulatory changes that make it possible for more people to live in low emissions places, could result in meaningful reductions in household carbon emissions.

There are several limitations of my empirical analysis that should be taken into consideration while interpreting my results. The first is that due to the survey nature of my data, carbon emissions are noisily measured. This leads to lower explanatory power of the model than is standard in papers using these methods with administrative data to estimate firm wage premia. The relatively low explanatory power of the model could also reflect model mis-specification, but with only two observations per household for the majority of my sample, the number of specification tests I can do is limited. Second, there is relatively little variation in urban form and transportation options across the US – 75 percent of residential land in the US is zoned for single family homes only (Badger and Bui 2019), 95 percent of commuters in my sample commute by car, and there is only one high speed rail line in the entire country, which operates at high speed over only roughly 50 miles of track. This is in stark contrast with other parts of the world, where many cities are denser and substantially less car oriented. Moreover, place effects are identified from movers, who differ from the general US population in meaningful ways. Thus, the external validity of my results relies upon estimates being stable to widening the distributions of place and household types.

Establishing that place matters – and how much it matters – for household carbon emissions takes a critical first step toward investigating the welfare impacts of specific place-based climate policies. The welfare effects of a given intervention depend on several parameters whose estimation is outside of the scope of this paper. First, they depend on the causal relationships between local amenities and place effects, and on household preferences for local amenities. While Tiebout (1956) posits that residential sorting allows for efficient provision of local public goods, his framework only applies to amenities without scale economies. Moreover, there is reason to believe that residential sorting is not efficient due to frictions and exclusionary policies (e.g. Rothstein 2017; Hausman and Stolper 2021; Christensen and Timmins 2023; Avenancio-León and Howard 2022). Estimating causal relationships between local public amenities and household carbon emissions, and quantifying whether emissions-relevant local public amenities are at an efficient level are important directions for future work. Second, the welfare impacts of place-based interventions would depend on the costs of implementing them relative to the cost of business as usual or other contender climate policies. Costs can vary dramatically for the same intervention in different settings (Goldwyn et al. 2022), making this estimation difficult, but incorporating cost estimates for a marginal value of public funds analysis (Hendren and Sprung-Keyser 2020) is another important avenue for future research. It is worth emphasizing that because built environment is sticky, infrastructure investments and regulatory choices made in the present day have the potential to increase or decrease the costs of carbon mitigation efforts in the future. Finally, welfare impacts would depend on other externalities or agglomeration benefits of the intervention. For example, the types of interventions highlighted in this paper could also impact local air pollution, congestion, traffic fatalities, and labor market productivity. These co-benefits and harms have been extensively studied in the environmental and urban economics literatures, and estimates could be incorporated into an aggregate welfare effect.

References

Abaluck, Jason, Mauricio Caceres Bravo, Peter Hull, and Amanda Starc. 2021. "Mortality Effects and Choice Across Private Health Insurance Plans." *Quarterly Journal of Economics* 136 (3): 1557–1610.

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica* 67 (2): 251–333.

Allcott, Hunt, Rebecca Diamond, Jean-Pierre Dubé, Jessie Handbury, Ilya Rahkovsky, and Molly Schnell. 2019. "Food Deserts and the Causes of Nutritional Inequality." *Quarterly Journal of Economics* 134 (4): 1793–1844.

Allen, Treb, and Costas Arkolakis. 2022. "The welfare effects of transportation infrastructure improvements." *The Review of Economic Studies* 89 (6): 2911–2957.

Andrews, Martyn, Laurence Gill, Thorsten Schank, and Richard Upward. 2008. "High wage workers and low wage firms: negative assortative matching or limited mobility bias?" *Journal of the Royal Statistical Society: Series A: Statistics in Society* 171 (3): 673–697.

Angrist, Joshua, Peter Hull, Parag Pathak, and Christopher Walters. 2017. "Leveraging Lotteries for School Value-Added: Testing and Estimation." *The Quarterly Journal of Economics* 132 (2): 871–919.

Auffhammer, Maximilian, and Edward Rubin. 2024. "Quantifying Heterogeneity in the Price Elasticity of Residential Natural Gas." *Journal of the Association of Environmental and Resource Economists* 11 (2): 319–357.

Austin, Benjamin, Edward Glaeser, and Lawrence Summers. 2018. "Saving the Heartland: Place-Based Policies in 21st Century America." *Brookings Papers on Economic Activity:* 151–255.

Avenancio-León, Carlos F, and Troup Howard. 2022. "The Assessment Gap: Racial Inequalities in Property Taxation." *The Quarterly Journal of Economics* 137 (3): 1383–1434.

Badger, Emily, and Quoctrung Bui. 2019. "Cities Start to Question an American Ideal: A House With a Yard on Every Lot." *New York Times.*

Bailey, Martha, Hilary Hoynes, Maya Rossin-Slater, and Reed Walker. 2024. "Is the Social Safety Net a Long-Term Investment? Large-scale Evidence from the Food Stamps Program." *The Review of Economic Studies* 91 (3): 1291–1330.

Barrington-Leigh, Chris, and Adam Millard-Ball. 2017. "More Connected Urban Roads Reduce US GHG Emissions." *Environmental Research Letters* 12 (4).

Bilal, Adrien. 2023. "The Geography of Unemployment." *The Quarterly Journal of Economics* 138 (3): 1507–1576.

Bond, Brittany, J. David Brown, Adela Luque, and Amy O'Hara. 2014. "The Nature of the Bias When Studying Only Linkable Person Records: Evidence from the American Community Survey." *CARRA Working Paper* 8:2–30.

Borenstein, Severin, and James B Bushnell. 2022. "Do two electricity pricing wrongs make a right? Cost recovery, externalities, and efficiency." *American Economic Journal: Economic Policy* 14 (4): 80–110.

Busso, Matias, Jesse Gregory, and Patrick Kline. 2013. "Assessing the Incidence and Efficiency of a Prominent Place Based Policy." *American Economic Review* 103 (2): 897–947.

Card, David, Ana Rute Cardoso, and Patrick Kline. 2016. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." *Quarterly Journal of Economics* 131 (2): 633–686.

Card, David, Joerg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics* 128:967–1015.

Card, David, Jesse Rothstein, and Moises Yi. 2024. "Location, Location, Location." *American Economic Journal: Applied Economics (forthcoming).*

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014a. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104 (9): 2593–2632.

———. 2014b. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review* 104 (9): 2633–79.

Chetty, Raj, and Nathaniel Hendren. 2018. "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *The Quarterly Journal of Economics* 133 (3): 1107–1162.

Christensen, Peter, and Christopher Timmins. 2023. "The damages and distortions from discrimination in the rental housing market." *The Quarterly Journal of Economics* 138 (4): 2505–2557.

Cicala, Steve. 2022. "Imperfect Markets versus Imperfect Regulation in US Electricity Generation." *American Economic Review* 112 (2): 409–41.

———. 2023. "JUE Insight: Powering Work from Home." *Journal of Urban Economics* 133.

Colas, Mark, and John M Morehouse. 2022. "The environmental cost of land-use restrictions." *Quantitative Economics* 13 (1): 179–223.

Combs, Tabitha. 2020. *Local Actions to Support Walking and Cycling During Social Distancing Dataset.*

Cornfield, Jerry. 2023. *Washington makes another run at heat pump rules,* November.

Cronin, Julie Ann, Don Fullerton, and Steven Sexton. 2019. "Vertical and Horizontal Redistribution from a Carbon Tax and Rebate." *Journal of the Association of Environmental and Resource Economists* 6 (S1): S169–S208.

De la Roca, Jorge, and Diego Puga. 2017. "Learning by Working in Big Cities." *Review of Economic Studies* 84 (1): 106–142.

Dijkstra, Lewis, Hugo Poelman, and Paolo Veneri. 2019. "The EU-OECD Definition of a Functional Urban Area." *OECD.*

D'Lima, Walter, Luis Arturo Lopez, and Archana Pradhan. 2022. "COVID-19 and housing market effects: Evidence from US shutdown orders." *Real Estate Economics* 50 (2): 303–339.

Druckenmiller, Hannah. 2023. "Urban Heat Islands 101." *Resources Magazine.*

Duranton, Gilles, and Diego Puga. 2020. "The Economics of Urban Density." *Journal of Economic Perspectives* 34 (3): 3–26.

Duranton, Gilles, and Anthony Venebles. 2021. "Place-based Policies: Principles and Developing Country Applications." *Handbook of Regional Science:* 1009–1030.

Eid, Jean, Henry G. Overman, Diego Puga, and Matthew Turner. 2008. "Fat City: Questioning the Relationship between Urban Sprawl and Obesity." *Journal of Urban Economics* 63 (2): 385–404.

Eisen, Ben. 2019. "Workers Are Fleeing Big Cities for Smaller Ones—and Taking Their Jobs With Them." *The Wall Street Journal.*

Energy Information Administration. 2020a. *Form EIA-861: Annual Electric Power Industry Report.*

———. 2020b. *State Energy Data System.*

Federal Highway Administration. 2019. *National Household Travel Survey.*

Federal Transit Administration. *NTD Data.*

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2021. "Place-Based Drivers of Mortality: Evidence from Migration." *American Economic Review* 111 (8): 2697–2735.

———. 2016. "Sources of Geographic Variation in Health Care: Evidence from Patient Migration." *Quarterly Journal of Economics* 131 (4): 1681–1726.

Fulton, Bill, David Garcia, Ben Metcalf, Carolina Reid, and Truman Braslaw. 2023. "New Pathways to Encourage Housing Production:A Review of CaliforniaâĂŹs Recent Housing Legislation." *Terner Center for Housing Innovation, UC Berkeley.*

Gaubert, Cecile. 2018. "Firm Sorting and Agglomeration." *American Economic Review* 108 (11): 3117–53.

Gillingham, Kenneth. 2014. "Identifying the Elasticity of Driving: Evidence from a Gasoline Price Shock in California." *Regional Science and Urban Economics* 47 (4): 13–24.

Glaeser, Edward L., and Matthew E. Kahn. 2010. "The Greenness of Cities: Carbon Dioxide Emissions and Urban Development." *Journal of Urban Economics* 67 (3): 404–418.

Glaeser, Edward. 2013. "Urban Public Finance." *Handbook of Public Economics* 5:195–256.

Glaeser, Edward, and Joshua Gottlieb. 2008. "The Economics of Place-Making Policies." *Brookings Papers on Economic Activity* 2008:155–239.

Goldstein, Benjamin, Dimitrios Gounaridis, and Joshua P. Newell. 2020. “The Carbon Footprint of Household Energy Use in the United State.” *Proceedings of the National Academy of Sciences* 117 (32): 19122–19130.

Goldwyn, Eric, Alon Levy, Elif Ensari, Marco Chitti, Abdirashid Dahir, Rachel Daniell, Yinan Yao, and Anan Maalouf. 2022. *The Transit Costs Project.*

Green, Tomas W, and Christopher R Knittel. 2020. “Distributed Effects of Climate Policy: a Machine Learning Approach.” *MIT Center for Energy and Environmental Policy Research Working Paper.*

Gupta, Arpit, Vrinda Mittal, Jonas Peeters, and Stijn Van Nieuwerburgh. 2022. “Flattening the Curve: Pandemic-Induced Revaluation of Urban Real Estate.” *Journal of Financial Economics* 146 (2): 594–636.

Hausman, Catherine, and Samuel Stolper. 2021. “Inequality, information failures, and air pollution.” *Journal of Environmental Economics and Management* 110:102552.

Hendren, Nathaniel, and Ben Sprung-Keyser. 2020. “A Unified Welfare Analysis of Government Policies.” *Quarterly Journal of Economics* 135 (3): 1209–1318.

International Energy Agency. 2021. *Net Zero by 2050: A Roadmap for the Global Energy Sector.* Technical report.

Jenkins, Jesse D., E.N. Mayfield, J Farbes, R Jones, N Patankar, Q Xu, and G Schivley. 2022. *Preliminary Report: The Climate and Energy Impacts of the Inflation Reduction Act of 2022.* Technical report. REPEAT Project.

Jones, Christopher, and Daniel M. Kammen. 2014. “Spatial Distribution of U.S. Household Carbon Footprints Reveals Suburbanization Undermines Greenhouse Gas Benefits of Urban Population Density.” *Environmental Science and Technology* 48 (2): 895–902.

Kline, Patrick. 2010. “Place Based Policies, Heterogeneity, and Agglomeration.” *American Economic Review* 100 (2): 383–87.

Kline, Patrick, and Enrico Moretti. 2013. “Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority.” *Quarterly Journal of Economics* 129 (1): 275–331.

———. 2014. “People, Places and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs.” *Annual Review of Economics* 6:629–662.

Kline, Patrick, Evan Rose, and Chris Walters. 2021. “Systemic Discrimination Among Large U.S. Employers.” *Quarterly Journal of Economics* 137 (4): 1963–2036.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten. 2020. “Leave Out Estimation of Variance Components.” *Econometrica* 88 (5): 1859–1898.

Ko, Yekang. 2013. “Urban Form and Residential Energy Use: A Review of Design Principles and Research Findings.” *Journal of Planning Literature* 28 (4): 327–351.

Kolko, Jed. 2021. "The Most Urban Counties in the U.S. Are Shrinking." *The New York Times.*

Lachowska, Marta, Alexandre Mas, Raffaele Saggio, and Stephen A Woodbury. 2023. "Do Firm Effects Drift? Evidence from Washington Administrative Data." *Journal of Econometrics* 233 (2): 375–395.

Larsen, John, Ben King, Hannah Kolus, Naveen Dasari, Galen Hiltbrand, and Whitney Herndon. 2022. *A Turning Point for US Climate Progress: Assessing the Climate and Clean Energy Provisions in the Inflation Reduction Act.* Technical report. Rhodium Group.

Levinson, Arik. 2016. "How Much Energy Do Building Energy Codes Save? Evidence from California Houses." *American Economic Review* 106 (10): 2867–2894.

Mahajan, Megan, Olivia Ashmoore, Jeffrey Rissman, Robbie Orvis, and Anand Gopal. 2022. *Modeling the Inflation Reduction Act using the Energy Policy Simulator.* Technical report. Energy Innovation.

Maricopa Association of Governments. 2020. *2040 Regional Transportation Plan (RTP) Update Executive Summary.* Technical report.

Mateyka, Peter, and Christopher Mazur. 2021. *Homeownership in the United States: 2005 to 2019.*

Mervosh, Sarah. 2018. "Minneapolis, Tackling Housing Crisis and Inequality, Votes to End Single-Family Zoning." *The New York Times.*

Murray, Brian C, and Peter T Maniloff. 2015. "Why Have Greenhouse Emissions in RGGI States Declined? An Econometric Attribution to Economic, Energy Market, and Policy Factors." *Energy Economics* 51:581–589.

NASA. 2020. *Vital Signs.*

National Oceanic and Atmospheric Administration. 2020. *US Climate Division.*

New York City Department of Transportation. 2023. *NYC DOT Taking New Steps to Expand Bike Infrastructure and Encourage Safe Operation of E-bikes as Overall Bike Ridership Reaches All-time High.* Press Release.

Nowak, William, and Ian Savage. 2013. "The Cross Elasticity between Gasoline Prices and Transit Use: Evidence from Chicago." *Transport Policy* 55 (10): 885–910.

O'Brien, Brendan. 2023. "New York State bans natural gas in some new construction." *Reuters.*

Ou, Jinpei, Xiaoping Liu, Xia Li, and Yimin Chen. 2013. "Quantifying the Relationship Between Urban Forms and Carbon Emissions Using Panel Data Analysis." *Landscape Ecology* 28:1889–1907.

Payne, Heather. 2020. "The natural gas paradox: Shutting down a system designed to operate forever." *Maryland Law Review* 80 (3): 693–762.

Petek, Gabriel. 2020. "Assessing California's Climate Policies –Electricity Generation." *Legislative Analysts Office Report.*

Philips, Ian, Jillian Anable, and Tim Chatterton. 2022. "E-bikes and their Capability to Reduce Car CO2 Emissions." *Transport Policy* 116:11–23.

Pomponi, Franceso, Ruth Saint, Jay H. Arehart, Niaz Gharavi, and Bernardino D'Amico. 2021. "Decoupling Density from Tallness in Analysing the Life Cycle Greenhouse Gas Emissions of Cities." *Npj Urban Sustainability* 1 (1): 33.

Raetz, Hayley. 2021. "Housing Characteristics of Small and Mid-Sized Cities." *The Stoop, NYU Furman Center Blog.*

Resume Builder. 2023. *90% of Companies Will Return to Office By the End of 2024.*

Ribeiro, Haroldo V., Diego Rybski, and Jurgen P. Kropp. 2019. "Effects of Changing Population or Density on Urban Carbon Dioxide Emissions." *Nature Communications* 10 (1): 3204.

Roberts, David. 2019. "Cars Dominate Cities Today. Barcelona has Set Out to Change That." *Vox.*

Rothstein, Richard. 2017. *The Color of Law.* Liveright Publishing Corporation.

Sallee, James. 2019. "Pigou Creates Losers: On the Implausibility of Achieving Pareto Improvements from Efficiency-Enhancing Policies." *NBER Working Paper,* no. 25831.

Shammin, Md, Robert Herendeen, Michelle Hanson, and Eric Wilson. 2010. "A Multivariate Analysis of the Energy Intensity of Sprawl versus Compact Living in the U.S. for 2003." *Ecological Economics* 69 (12): 2363–2373.

Song, Jae, David J. Price, Faith Guvenen, Nicholas Bloom, and Till Von Wachter. 2019. "Firming up Inequality." *The Quarterly Journal of Economics* 134 (1): 1–50.

Spiller, Elisheba, Heather Stephens, Christopher Timmins, and Allison Smith. 2014. "The Effect of Gasoline Taxes and Public Transit Investments on Driving Patterns." *Environmental and Resource Economics* 59 (4): 633–657.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 65 (5): 416–424.

Timmons, David, Nikolaos Zirogiannis, and Manuel Lutz. 2016. "Location Matters: Population Density and Carbon Emissions from Residential Building Energy Use in the United States." *Energy Research & Social Science* 22:137–146.

Tomer, Adie, Joseph Kane, Jenny Schuetz, and Caroline George. 2021. "We Can't Beat the Climate Crisis without Rethinking Land Use." *Brookings Institution.*

Transportation Research Board. 2009. *Driving and the Built Environment: The Effects of Compat Development on Motorized Travel, Energy Use, and CO_2 Emissions.* Technical report. National Research Council of the National Academies.

Tsivanidis, Nick. 2022. "Evaluating the Impact of Urban Transit Infrastructure:Evidence from Bogota's TransMilenio." *Working Paper.*

Ummel, Kevin. 2014. "Who pollutes? A household-level database of America's greenhouse gas footprint." *Center for Global Development Working Paper,* no. 381.

U.S. Department of Commerce, Bureau of the Census. 2014. *American Community Survey Design and Methodology.* Technical report.

U.S. Department of Energy. 2015. *Quadrennial Energy Review: Energy Transmission, Storage, and Distribution Infrastructure.* Technical report. U.S. Department of Energy.

U.S. Department of Homeland Security. 2021. *Homeland Infrastructure Foundation-Level Data.*

U.S. Department of Transportation. 1977. *America's Highways: 1776-1976: A History of the Federal-Aid Program.* U.S. Government Printing Office.

———. 2024. "Biden-Harris Administration Sends Another $9.9 Billion from the Bipartisan Infrastructure Law to Transit Agencies Across the Country."

U.S. Department of Transportation, Federal Highway Administration. 2015. *2015 Status of the Nation's Highways, Bridges, and Transit: Conditions & Performance.* Technical report.

U.S. Energy Information Administration. 2020a. "U.S. Energy-Related Carbon Dioxide Emissions, 2019."

———. 2020b. *U.S. energy-related carbon dioxide emissions fell in 2019, mainly in electric generation.* Technical report.

U.S. Environmental Protection Agency. 2018. *Emission Factors for Greenhouse Gas Inventories.* Technical report.

———. 2021a. *Emissions & Generation Resource Integrated Database.*

———. 2021b. *U.S. Transportation Sector Greenhouse Gas Emissions 1990-2019.* Technical report.

U.S. Environmental Protection Agency, Office of Energy Efficiency, and Renewable Energy. 2020.

Van Nieuwerburgh, Stijn. 2023. "The Remote Work Revolution: Impact on Real Estate Values and the Urban Environment: 2023 AREUEA Presidential Address." *Real Estate Economics* 51 (1): 7–48.

Ville de Paris. 2021. *Un nouveau Plan Vélo pour une ville 100% cyclable.*

Wagner, Deborah, and Mary Layne. 2014. "The Person Identification Validation System (PVS):Applying the Center for Administrative Records." *CARRA Working Paper,* no. 1.

Wagner, Gernot. 2021. "Sometimes a High Rise Is Better Than a Park." *Bloomberg.*

Xu, Dafeng. 2020. "Free Wheel, Free Will! The Effects of Bikeshare Systems on Urban Commuting Patterns in the U.S." *Journal of Policy Analysis and Management* 39 (3): 664–685.

A Data Appendix

A.1 Additional Details on Variable Construction

- **Missing and imputed variables:** I follow Chetty and Hendren (2018) and Bailey et al. (2024) in treating all imputed variables as missing, unless otherwise described. Dollar values are inflated to 2019 values using the Consumer Price Index (CPI). Throughout the analysis I use demographic and household characteristics to control for selection on time-varying observables, I use work characteristics to construct commuting variables, and I use home characteristics in the second half of the paper to characterize places and study associations between built environment and place effects.

- **Flags**: Residential energy expenditure is flagged as "allocated" (to zero) for many households in the 2014 ACS if they checked a box indicating that they did not use natural gas or fuel use and then left the expenditure question blank. Because of this, I make an exception to my rule of dropping flagged variables and allow for residential energy to be allocated to zero based on the checkbox question.

- **Work characteristics:** For each individual I retain information on employment status, place of work, weeks worked last year, and hours worked last week. I allow place of work tracts or more detailed geographies to be missing, but I drop observations if county of work is missing (unless the individual is unemployed or works from home, in which case I impute their place of work as their home). I also allow current employment status to be missing if weeks worked last year and hours worked last week are not missing and not imputed. In 2008-2018, the weeks worked variable is binned; I follow Chetty and Hendren (2018) and assign the midpoint to all individuals in the bin. Since these variables are an input into my measure of commuting energy use, I use the midpoint from the bin for all years to keep the variable definition consistent.

- **Demographic characteristics** I control for age using bins: 18-24, 25-29, 30-34, 35-39, 40-49, 50-64, and 65+. I control flexibly for number of children in the household using categorical variables for 0, 1, 2, or 3+ kids. As highlighted in Card, Cardoso, and Kline (2016), the normalization choice for categorical variables does not affect the estimated size of the place variance component or the variance component of the sum of fixed and observable household effects, but it does affect the relative sizes of the place and unobserved household effects, as well as the estimated covariances. Throughout my analysis, I choose the age bin 40-49, no college degree, male, white, and non-Hispanic as the omitted categories. Other than "white", these are the categories with the highest within-group variance in outcomes. Thus this normalization should err towards finding a larger unobservable person component relative to place component.

- **Observations dropped due to missing utility data:** I exclude households from the sample if their residential energy costs are included in rent, or if their gas costs are included in their electricity bill, because I do not observe expenditures in those cases. As shown in Table G.1, households in this sample have significantly lower income, are less likely to own their home, live in a detached single family home, or commute by car, and are more likely to live in an urban neighborhood. For the subsample who would have been in the panel if not for this restriction, I estimate these households are more likely to have an increase in income, an increase in the number of children in the household, and go from renting to owning than the baseline sample. They are also twice as likely to move from an urban neighborhood to another urban neighborhood when moving. To the extent that you might worry that 1) households who live in the most urban neighborhood are most selected and would respond differently to changes in place than a more average household or 2) large changes in observable characteristics signal that estimates may be biased by accompanying large changes in unobservable characteristics, dropping these households from the analysis is likely to decrease any upward bias in place share estimates that could arise from violations of either the constant effects assumption or the exogenous mobility assumption.

- **Vehicle fuel economy:** I assume individuals that commute by car or taxi do so in a vehicle with annual national average fuel economy, using data from the Federal Highway Administration (2019). For motorcycles, I scale miles per gallon (mpg) by two (U.S. Department of Transportation 2015). This is a minor point as motorcycles account for only roughly 0.6 percent of vehicle miles driven (U.S. Environmental Protection Agency and Energy 2020). I also account for the fact that in general fuel economy is roughly 30 percent higher when driving on highways than in cities by adjusting mpg up by 19 percent relative to the national average for drivers whose average commuting speed exceeds 55 miles per hour (mph), and down by nine percent relative to the national average for drivers whose average commuting speed is below 40 mph (U.S. Environmental Protection Agency 2021b). As a robustness check, I also use data from the National Household Travel Survey (NHTS) to estimate heterogeneous fuel economy values. I discuss the construction of these estimates in Appendix A.2.

- **Carpooling**: I divide carbon emissions by the number of carpoolers for individuals who report carpooling.

- **Emissions from public transportation:** I assign emissions factors to commutes by public transportation using estimates derived from the National Transit Database (NTD, Federal Transit Administration). Specifically, for each transit agency and year, I use reported data on fuel consumption and passenger miles traveled (PMT) by mode in order to estimate carbon emissions per PMT for six possible modes of commuting: subway or elevated rail, commuter rail, streetcar, bus, ferry, or taxi.[22] As with residential emissions, I assign fuel emissions factors using data from the U.S. Environmental Protection Agency 2018 and I assign electricity emissions factors using average values at the North American Electric Reliability Corporation (NERC) region level. After estimating mode-specific emissions for all reporting transit agencies, I estimate passenger mile weighted mode emis-

22. NTD modes of commute are more granular than ACS modes of commute. I group heavy rail, monorail/automated guideway, light rail, and aerial tramway into the subway & elevated rail group. I treat commuter rail and hybrid railroad as railroad. I group streetcar, cable car, inclined plane, and trolleybus into the streetcar mode. Bus, bus rapid transit, and commuter bus all get categorized as bus. And lastly, I categorize demand response, demand response taxis, and vanpools as taxi.

sions factors in each urban area, and assign individuals an emissions factor based on the urban area they live in. For individuals who live outside of an urban area with a reporting agency, I use their mode's national annual average emissions per passenger miles travelled to estimate emissions.

- **Emissions from walking or biking**: I assign zero emissions to commutes by walking or biking. This underestimates emissions from biking as electric bikes (i.e. e-bikes) grow in popularity. Unfortunately, I cannot distinguish in the ACS the kind of bikes commuters use, and in the majority of my sample traditional bicycles dominated the market. An important question for future research is how e-bike subsidies and local bike-share programs change commute mode choice and emissions.[23]

- **Commuting distance**: I estimate commute mileage using the GPS distance between reported home and place of work census blocks. To account for the fact that geodesic distances don't capture the indirect nature of roads, I re-scale my mileage estimates to match the national average commuting distance, by mode, reported in the NHTS. For individuals who only report their county of work but not their census block of work, I impute miles traveled using reported commute time and average commute speeds for people with similar residence-job geographic pairs. I use a similar imputation for individuals for whom the travel speeds implied by dividing estimated miles by commute time are infeasible – over 150 mph in a train,[24] or over 80 mph on average in other modes.

- **Number of annual commutes**: I estimate commuting days per week using reported hours worked last week and assuming people work eight hours a day up to five days a week, assuming people worked five days if they worked 40-50 hours a week, 6 days if they worked 50-60 hours in a week, and 7 days if they worked more than that. I assume everyone commutes twice a day, and that commuting behavior is consistent across all the weeks worked last year.

- **Identifying children:** I designate a household member a child and drop them from the analysis sample if they are under the age of 18, or if they are identified as a child via the Census' relationship to householder code.

A.2 Measurement Error in Household Carbon Emissions

There are several sources of measurement error in household carbon emissions from residential and transportation energy use. While an advantage of the ACS is that it makes observable many household characteristics that are unobservable in standard administrative datasets on energy use, making it possible to control for changes to household characteristics that are correlated with both changes to energy demand and move propensity and destinations and decrease potential bias from unobserved preference shocks, a disadvantage is that the survey nature of the data means that the outcome variables are constructed from a combination of survey responses (whose quality depends on household reporting) and local external data. This could introduce

23. Xu (2020) finds that bike commuting is more common in cities with bike share programs.
24. This is the fastest speed a train ever goes in the US, along a small segment of the Northeast Corridor.

bias in either estimates of household and place effects, estimates of the variance components, or both. Note that if errors are random but serially correlated within a household, both a naive variance decomposition and a KSS variance decomposition on a sample consisting of both stayers and movers will overstate the share of heterogeneity attributable to households; however, when I restrict to the mover only sample, the KSS correction accounts for serial correlation in the error term and gives unbiased estimates of variance components. Below, I discuss the various possible sources of measurement error, as well as potential biases that arise in my estimates as a result. In cases of greater concern, I discuss the construction of alternate variables used for robustness checks in the paper.

Household reporting of residential energy expenditures

Households may not accurately report energy expenditures. Inaccurate reporting could arise, for example, due to inattention to bills, or due to bias driven by the seasonality of energy expenditures – that is, if households use their last monthly bill to proxy for annual expenditures.

If household inattention is fixed it will be absorbed by the household effect. If inattention leads high types to overstate their expenditures, and low types to understate their expenditures, this would lead to an upward bias in the household component of heterogeneity, and vice versa. It is also reasonable to think that inattention may be random but serially correlated within household.

With fixed or random inattention, estimates of place effects themselves are unbiased. However, if moves are correlated with changes in attention, this could lead to bias in estimates of place effects. For example, if households move after positive income shocks, and higher income households pay less attention to their energy bills, *and* this inattention leads to systematic under- or over-estimation of expenditures, estimates of place effects with more inattentive residents would be biased.

Seasonality is unlikely to bias my estimates because surveys are sent out randomly, and therefore the season households were surveyed shouldn't be correlated with other components of the model.

Electricity prices

In the baseline specifications, I estimate electricity prices from total utility revenues divided by total utility customers, by county (using data from EIA Form 861). This introduces three sources of measurement error in electricity prices.

First, in counties served by more than one utility, I cannot match customers to the actual utility they are served by. If customers in an area can select their residential energy provider, this could lead to bias in the household component of heterogeneity. For example, if higher type customers are selecting into lower average price utilities, I will underestimate the household component of heterogeneity. Similarly, if there are several utilities serving different neighborhoods within the same county, this could lead to bias in the place component of heterogeneity. In particular, I will over-estimate consumption in neighborhoods served by more expensive utilities, and under-estimate consumption in neighborhoods served by cheaper utilities. If more expensive utilities generally serve lower consumption neighborhoods, this will lead me to underestimate the place component of heterogeneity.

Second, residential customers generally face a two-part tariff consisting of a fixed charge and a marginal volumetric charge, in which the marginal price can either be increasing or decreasing in consumption depending on the utility. Because I use average prices, calculated from utility residential revenues and quantities sold, I overestimate the average volumetric price and in turn underestimate consumption for everyone (more so for households in high fixed charge service territories). Moreover, for some utilities, marginal prices are either increasing or decreasing in consumption. When prices are increasing in consumption, I under-estimate prices faced by high-demand customers and over-estimate prices faced by low-use customers. This means I over-estimate quantities consumed by high-demand customers and under-estimate quantities consumed by low-demand customers, leading to an upward bias in my estimates of the household variance component. Conversely, if prices are decreasing in consumption, I underestimate the household variance component.

Borenstein and Bushnell (2022) estimate that in the US, roughly 37 percent of customers face increasing block pricing, and roughly 21 percent face decreasing block pricing, though in all cases the rate structure is fairly narrow. They also estimate that across territories, utilities that utilize increasing-block pricing generally serve lower demand customers on average. Thus, my estimates likely somewhat over-estimate variation across households within utility territories, and underestimate variation across territories. Overall, unobserved rate structures should lead me to estimate a lower bound on place-based heterogeneity and estimate an upper bound on preference-based heterogeneity.

Finally, residential rates can vary within utilities, and I don't observe which rate a household has selected. This leads to the same biases as not being able to observe which utility a customer chooses, discussed above. Additionally, I do not observe if a household has solar, and in many states solar customers face different price schedules with significant subsidies for selling generated power back to the grid. This lowers their average price per kilowatt hour (kwh), causing me to underestimate quantity consumed and in turn CO_2 from electricity purchased from the grid by these customers.

Alternate electricity price estimates for robustness checks

To test the sensitivity of my results to the issues described above, I construct several alternate estimates of residential electricity prices.

First, I account for fixed charges, closely following Borenstein and Bushnell (2022) in my approach. I supplement the EIA 861 data with annual data from the Utility Rate Database (URDB), which contains utility-level data on rate schedules. I collect fixed charges from the set of utilities in the URDB that report detailed retail pricing information, using the median fixed charge in the standard tariff for each utility-state pair in cases in which utilities reported multiple rates.

The URDB is not perfectly populated, and is much sparser in the earlier years[25]. In cases where I observe a fixed charge for some but not all years of a utility-state pair, I impute values for missing years using values from the closest available year. If I observe two different fixed charges with missing years in between, I impute the value for those missing years using the mean of the observed values.

I then estimate the variable price component for each utility-state pair by combining my fixed charge estimates with annual total revenue, generation, and customer data from the EIA 861. I subtract estimated total fixed revenue (fixed charge times number of customers) from total revenue reported in EIA 861, and then divide variable revenue by total sales to get a variable price per kwh of electricity. Consistent with the fact that fixed charges are generally a low share of the two-part tariff (I estimate that across my sample fixed charges make up roughly 9 percent of total revenues), the distributions of average and variable prices appear similar. I proceed as in the baseline estimation, constructing a county-level average variable price as the customer-weighted mean variable price of all utilities serving a given county. In the microdata, for counties without a variable price estimate, I continue to use my average price estimate.

Second, I account for the fact that sometimes utility tariffs follow a tiered pricing schedule, in which marginal prices either increase or decrease with the quantity of electricity consumed. URDB also contains some information on price schedules with tiered pricing, but these data are even more complex and sparse than the fixed charge data. I have no way of knowing which customers choose a rate with tiered pricing, or even what share of customers are on each schedule. To bound the issues that could arise from tiered pricing, I gather information on the mean price difference between the top and bottom price for each utility-state-year. I do this separately for tariffs with increasing block rates vs. decreasing block rates. As with the baseline and variable price estimation, I estimate a county-level average price difference for increasing and decreasing block prices. I then estimate top and bottom county-level prices as the variable price in that county plus/minus half the price difference. I estimate the price step as being at the median county level quantity consumed, as estimated using average variable price.

I then explore three bounding scenarios. In the first, I assume that every customer who lives in a county where an increasing block price schedule is available chooses the increasing block price schedule. In counties without any increasing block price schedules, customers are assigned the average variable price. In the second, I make an analogous assumption but with decreasing block prices. Finally, I consider a selected scenario, in which customers with below median electricity costs for their county select into an increasing block pricing schedule, while customers with above median electricity costs for their county select into a decreasing block pricing schedule. Note, this selected scenario also yields some insight into the bias that would arise from customers selecting across utilities based on price. While none of these perfectly capture the actual price schedules faced by all customers in the data, they should provide some bounds on the bias incurred by not accurately observing marginal prices.

Using estimated variable prices or assigning tiered pricing schedules to households does not meaningfully impact variance component estimates (Table G.9).

25. Coverage of EIA 861 utilities goes from 16% in 2000 to 79% in 2019.

Electricity carbon emissions factors

I estimate the carbon emissions intensity of electricity using average emissions factors at the NERC region. This does not capture the fact that electricity is generated from different fuels throughout the course of the day (e.g., solar peaks in the afternoon) and across seasons (e.g., there is less solar in the winter). The error in household carbon emissions that results from this is likely serially correlated within household, and can be accounted for in the mover-only KSS specification. However, if consumption profiles are also correlated with these patterns, my estimates of household carbon emissions will be biased. For example, if low-type users consume more electricity when marginal emissions are higher, then I would tend to under-estimate their carbon emissions and over-estimate the household component of heterogeneity.

Alternate electricity emissions estimates for robustness checks

In the baseline specifications, I estimate household electricity emissions using average emissions factors computed from aggregate production and fuel use at the NERC region level. Conceptually, I believe that this is the right emissions factor to use because a change in the place effect simultaneously affects all residents of the place, leading to non-marginal changes in electricity consumption. However, I also construct a measure of household electricity emissions using marginal emissions factors. Note that this doesn't address the issue of emissions varying across hours and seasons and that variation possibly being correlated with usage patterns, because I cannot distinguish differences in marginal emissions across households within a place.

I follow Borenstein and Bushnell (2022) and estimate marginal emissions for each of nine regions – the eight reliability regions of the NERC, with the Western Interconnection (WECC) region split into California and non-California sub-parts – by regressing hourly carbon emissions on hourly load using the following specification

$$CO_{2it} = \beta Load_{it} + \alpha_{mn} + \gamma_i LoadInterconnect_{-it} + \epsilon_t$$

where α_{mn} represents month of sample by hour of day fixed effects and γ_i represents the marginal effect of load from other parts of the interconnect onto carbon emissions in a given region. Marginal emissions from electricity load in a region are then given by $\beta_i + \sum_{j \neq i} \gamma_j$. In practice, allowing for the impact of other regions' load on marginal emissions does not make a big difference.

I construct hourly carbon emissions from power plants in each region using data from the Environmental Protection Agency (EPA) Continuous Emissions Monitoring System (CEMS). I extend estimates of hourly load from Cicala (2022) through 2019 using data from the Federal Energy Regulatory Commission's Form-714 Survey. In a few region-year pairs where supplementary data were required but unavailable, I interpolate region marginal emissions from the nearest available year.

Using marginal emissions estimates to construct carbon emissions from electricity increases the share of covariance attributable to places by about seven percentage points relative to baseline estimates (Table G.9).

Natural gas and other residential heating fuel prices and emissions

Many of the same price measurement errors arise with natural gas as with electricity, but generally individuals have less choice over their utility, fixed charges are larger, and there is less prevalence of block pricing. Unlike electricity, fuel emissions factors for other fuels are the same regardless of where a household lives. However, in the case of natural gas a significant source of emissions is upstream methane leaks, which I don't consider in this analysis.

Assignment of heating fuel

I estimate carbon emissions from fuel use by assigning all expenditures on "other home heating fuels" to the fuel reported as the primary fuel. If a household has non-zero other fuel expenditures, but it doesn't list a primary fuel, I impute its primary fuel based on the most commonly used primary fuel among other survey respondents in their state and year (out of residual oil, propane, and wood). If in reality households use more than one heating fuel, or use a heating fuel other than the one I imputed for them, there will be error in my measurement of carbon emissions, both as a result of dividing expenditures by the wrong fuel price, and as a result of assigning the wrong carbon emissions factor. I will overestimate household carbon emissions if reported or imputed fuel prices are lower than actual average fuel prices faced by the household, or if reported or imputed fuel types have higher emissions factors than the fuels actually used.

If I tend to overestimate carbon emissions from heating fuels for otherwise high-type households and underestimate carbon emissions from heating fuels for otherwise low-type households, then my household variance component will be biased upward, and vice versa. Moreover, if moves are correlated with shocks to unobserved fuel components, this could lead to bias in my estimates of place effects. For example, if a household uses the same heating fuel everywhere they live but doesn't report this fuel, if they move to a place where their neighbors use an on average higher emissions heating fuel, I would overestimate the place effect. In practice, the share of households reporting non-zero energy expenditures on heating other than electricity or natural gas is small, and my estimates are not meaningfully affected when I exclude other heating from the calculation.

Commuting distances:

Because I estimate commute miles from geodesic distances between coordinates, I will underestimate speed and miles traveled for individuals who have less direct commutes. If place-based constraints (e.g., the result of living in a gated community or a neighborhood with many winding roads and cul-de-sacs) shape the directness of a commute, and if these types of neighborhoods tend to be farther from employment centers and have longer commutes to begin with, then I will underestimate the place component of spatial heterogeneity.

Additionally, I impute miles for the people for whom I don't observe census block of work using average mph for home and place of work county pairs. This will lead me to overstate commute distances for people with slower than average commutes, and understate commute distance for people with faster than average commutes. If faster than average commutes are also longer than average, then I will underestimate the household component of spatial heterogeneity. The "Commute from hrs" rows in Table G.9 show that my estimates are not sensitive to using a simpler measure of commute distance, calculated from simply dividing reported commute time by the average national commute speed, 32 mph (Federal Highway Administration 2019), suggesting that errors in commute speeds are unlikely to bias my estimates.

Total commuting miles:

I use weeks worked last year to estimate total commuting from typical commuting behavior last week. This assumes that hours worked are stable, that people work at the same place all year, and that information about commutes reported for last week is representative of commutes generally. Any deviations along these dimensions introduces measurement error into my outcome. While such errors are more likely to arise for lower income households with less job stability, it is unlikely that it results in a systematic over- or under-estimate of commute miles on average.

Indeed, the results in the "Commute from hrs, fixed num." row in Table G.9 are qualitatively similar to the baseline estimates.

Commuting emissions:

I assume everyone drives a vehicle with the annual national average fuel economy, using data from the NHTS. This is a significant oversimplification – and my inability to observe fuel economy is a significant limitation of my data – as it ignores patterns of heterogeneity in fuel economy both across commute lengths and across regions. If people with longer commutes drive more fuel efficient vehicles, I will overstate heterogeneity. On the other hand, if people who want to conserve on gas both buy more fuel efficient vehicles and choose to have shorter commutes, I will understate heterogeneity. The bias in my estimates of relative shares is more ambiguous. As with my broader analysis, there is a question of whether regional patterns in fuel economy are driven by individual preferences or place-based differences. If regional variation in fuel economy is driven by individual preferences, I will understate the relative importance of the person component in spatial variation. On the other hand, if they are driven by local norms or place characteristics such as, for instance, the availability of parking and width of roads, I will understate the relative importance of the place effect.

Additionally, if households change their mode of transit over the year, or if they use multiple modes of transit in a single commute, I do not capture this variation. For example, if households report taking public transit as their primary mode, but in reality they drive part of the distance of their commute, I will under-estimate their carbon emissions and overstate overall heterogeneity. On the other hand, if they walk or bike part of the distance of their commute, I will overestimate these household's carbon emissions and understate overall heterogeneity.

Allowing for heterogeneous vehicle fuel economy

In the baseline specification, I assign a national average fuel economy to all households. To explore the sensitivity of my results to this assumption, I construct three estimates of fuel economy using data from the NHTS, allowing for heterogeneity across geographic characteristics (CBSA, state, urbanity) only, individual and household characteristics (age, race, household size, household income, gender, number of vehicles, and commute mode of transit interacted with commute length) only, and both sets of characteristics. For each specification, I use a penalized Lasso regression to predict individual-level vehicle fuel economy based on the included set of characteristics, and then I use these estimates of mpg to estimate emissions from commuting.

Results are presented in the three "MPG" rows of Table G.9, and are not qualitatively different from baseline estimates.

Non-commuting transportation emissions:

I don't observe transportation other than commuting. In particular, I don't observe local travel for errands or leisure, nor do I observe airplane travel. Thus, I (weakly) underestimate carbon emissions magnitudes. If commuting is a rank-preserving share of total transportation emissions, my results will be qualitatively correct but off in magnitudes. However, if for example places with long commutes have lower other transportation emissions (because everybody spends leisure time in their back yard) whereas places with short commutes have higher other transportation emissions (because people go away for the weekend), then my estimates cannot be used to infer anything about heterogeneity in overall transportation emissions.

Estimating total vehicle miles traveled

In the baseline specification, miles commuted serve as a proxy for total vehicle miles. To explore the sensitivity of my results to this assumption, I also construct an estimate of total miles traveled by using a penalized Lasso regression to predict total miles from the set of both household and geographic variables described above in the NHTS data. Table G.9 shows that this does not meaningfully affect the results.

B The Leave-One-Out Connected Set

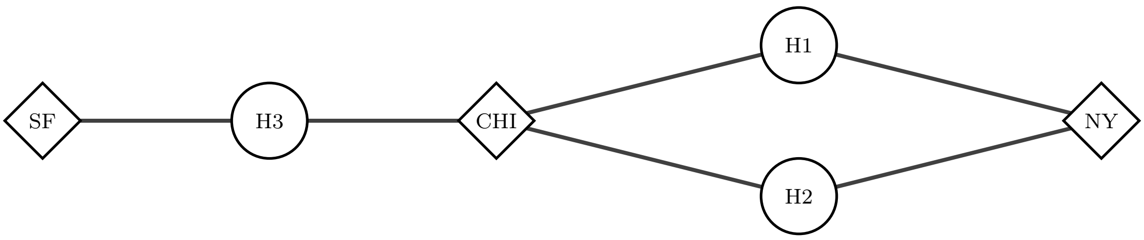
Consider the following data:

Individual & Household Geographic Locations

Year	Household	Place
1	1	NY
2	1	CHI
1	2	CHI
2	2	NY
1	3	SF
2	3	CHI

Household 1 moves from NY to Chicago, household 2 moves from Chicago to NY, and household 3 moves from San Francisco to Chicago. This data can be visualized as a network, where each place is a node, each household is a node, and edges connect households to each place they've lived in.

Household + Place Network



In this figure, San Francisco, Chicago, and New York are all connected by movers – this is a connected set. The leave-out connected set is the set of places that remains connected after dropping any household from the data. In this example, San Francisco is *not* in the leave-out connected set, because it is only connected to the rest of the network through H3.

C Empirical Bayes Adjustment

When discussing distributions of either observational means or place effects, I account for the fact that these parameters are noisily estimated by using linear Empirical Bayes, i.e. a shrinkage estimator. Many papers in the public and labor literatures have used this approach to predict for example teacher value add or neighborhood effects in other contexts (Chetty, Friedman, and Rockoff 2014a; 2014b; Angrist et al. 2017; Chetty and Hendren 2018; Finkelstein, Gentzkow, and Williams 2021; Abaluck et al. 2021). Although the linear approximation only corresponds to the true Empirical Bayes posterior when errors are normal and homoskedastic, Kline, Rose, and Walters (2021) show that even when errors are heteroskedastic, the linear shrinkage estimator doesn't do much worse than non-parametric Empirical Bayes. The shrinkage estimates are given by:

$$\hat{y}_j^{EB} = \lambda_j \hat{y}_j + (1 - \lambda_j) \frac{1}{J} \sum_j \hat{y}_j \tag{10}$$

where y represents the neighborhood-level parameter of interest, and the weights $\lambda_j = \frac{\hat{\sigma}_j^2}{s_j^2 + \hat{\sigma}_j^2}$ capture the signal-to-noise ratio of each estimate and down-weight noisy estimates to the grand mean.

D Model

Household i, living in place j, consumes quantity Q of energy in the form of four types of fuels: electricity (e), natural gas (n), other heating fuels (o), and motor gasoline (m). Each of these fuels has an emissions factor $\phi_{(jt)}$; these factors vary over time and place for electricity but are fixed along both of these dimensions for the other three fuel types. Household carbon emissions are therefore given by the following expression, with it subscripts temporarily supressed for easier legibility:

$$CO_2 = \phi^e_{jt} \cdot Q^e + \phi^n \cdot Q^n + \phi^o \cdot Q^o + \phi^m \cdot Q^m$$

Note that it is possible to re-express the above in terms of fuel shares, where for each fuel

$$s^f = \frac{Q^f}{\sum_f Q^f}$$

And therefore

$$CO_2 = \left(\sum_f s^f \cdot \phi^f_{(jt)} \right) \cdot Q$$

where, as before, Q represents total energy consumption across the four fuels.

Returning to Equation 3

$$\ln Q_{it} = a_j + \sum_{f \in \mathcal{F}} \rho^f_j \cdot \ln P^f_j + X_{it}\beta + \tau_t + \alpha_i + \varepsilon_{it}$$

it follows that

$$\ln CO_{2_{it}} = \ln \left(\sum_f s^f_{it} \cdot \phi^f_{(jt)} \right) + a_j + \sum_{f \in \mathcal{F}} \rho^f_j \cdot \ln P^f_j + X_{it}\beta + \tau_t + \alpha_i + \varepsilon_{it}$$

I add and subtract log of the average emissions factor, $\bar{\phi}_j$, which I used in the simplified exposition of the model in Equation 4, and rearrange terms to get the following expression:

$$\ln CO_{2_{it}} = \ln \bar{\phi}_j + a_j + \sum_{f \in \mathcal{F}} \rho^f_j \cdot \ln P^f_j + X_{it}\beta + \tau_t + \alpha_i + \varepsilon_{it} + \ln \left(\frac{\sum_f s^f_{it} \cdot \phi^f_{(jt)}}{\bar{\phi}_j} \right)$$

Observe that if not for the last term, this expression would be equivalent to Equation 4, but when household fuel shares vary, there is an interaction between household fuel shares relative to the average in the place where it lives, and place specific electricity emissions intensities. A household that disproportionately uses electricity wherever it lives will have a larger drop in emissions when moving from a place with relatively dirty electricity to a place with relatively clean electricity than the average household will. This variability gets absorbed by the error term in my regressions, and motivates the use of heteroskedastic errors.

E Computational Appendix

This section closely follows the description provided in KSS, as I replicate their method. I proceed in two steps, regressing $\log(CO_2)$ on observable characteristics and year fixed effects, and residualizing so that I am left with

$$\tilde{y}_{ij} = \alpha_i + \psi_j + \varepsilon_{it}$$

The share of overall variance attributable to place effects can then be captured by the variance component of place effects,

$$Var(\psi_j) \equiv \sigma^2_\psi = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\psi_{j(i,t)} - \bar{\psi})^2$$

and the covariance component between place effects and person effects

$$Cov(\alpha_i, \psi_j) \equiv \sigma^2_{\alpha,\psi} = \frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}(\psi_{j(i,t)} - \bar{\psi}) \cdot \alpha_i$$

KSS provides an estimate for the standard error $\psi_i^2 = Var(\varepsilon_i)$ based on a leave out estimate of σ_i^2:

$$\hat{\sigma}_i^2 = y_i(y_i - x_i'\hat{\beta}_{-i}) = y_i\frac{(y_i - x_i'\hat{\beta})}{1 - P_{ii}}$$

where $P_{ii} = x_i'(x_i x_i')^{-1} x_i$ is the observation leverage.

To reduce the computational burden of the KSS estimator, I use the Johnson-Lindenstrauss approximation (JLA) algorithm introduced by KSS to estimate the statistical leverages of each match, i.e. the amount by which estimates change when leaving out the match. KSS show that using JLA introduces an approximation error of roughly 10^{-4} relative to estimating statistical leverages directly. See KSS for a complete discussion of the leave-out estimator and JLA algorithm.

F The Evolution of Place Effects

Time-varying fixed effects:

I provide descriptive evidence on the changing nature of place effects from 2000-2019 in Appendix Figure G.8. Pooling all time-varying fixed effect estimates together and grouping pooled values into four quartiles, the vast majority of CBSAs either do not change rank or become lower emissions from the first period (2000-2004) to the last period (2015-2019), consistent with large declines in emissions from electricity production as a result of a dramatic decline in coal and increase in renewables (U.S. Energy Information Administration 2020a). In contrast, defining quartiles within year, the distribution of whether CBSAs become relatively lower or higher emissions than their counterparts between those two periods is roughly symmetric, but with over half of CBSAs not changing relative rank.

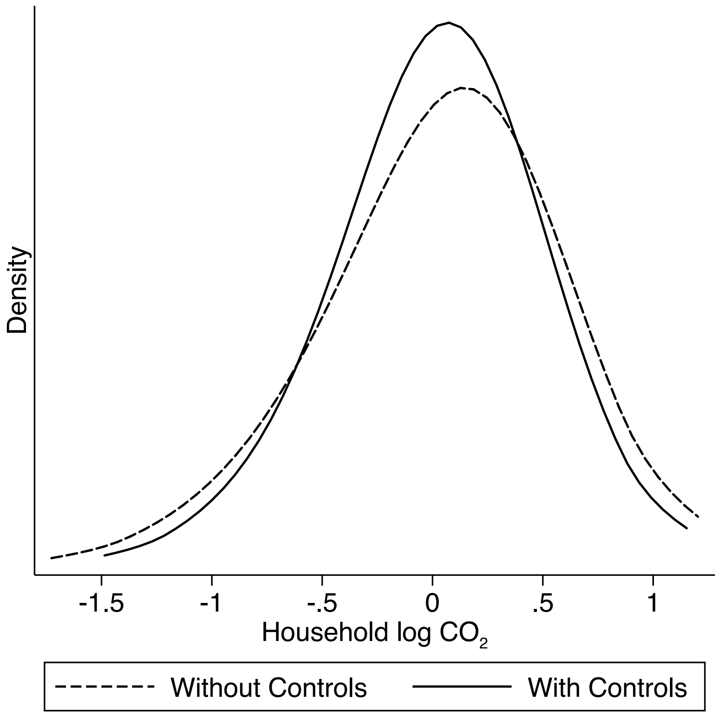
The impact of the COVID-19 pandemic:

There was likely a discontinuous change in CBSA effects following the COVID-19 pandemic and the shift to remote work, which was accompanied by a steep decline in commuting and a shift towards larger homes to accommodate home offices (Van Nieuwerburgh 2023; D'Lima, Lopez, and Pradhan 2022). Cicala (2023) finds that during the acute parts of the pandemic (Q2-Q4 of 2020), residential energy consumption increased by about eight percent, while the use of transportation fuel consumption declined by about 16 percent. The resulting increase in residential energy is likely to widen the gap in place effects between suburban and urban tracts, though the net impact on emissions should be modulated by the decrease in commercial energy. In contrast, the reduction in commuting is likely to decrease the gap between suburban and urban tracts. It is a limitation of my data that I only observe commuting miles, but in my sample time frame, using the NHTS to predict overall transportation from commuting does not substantively change the results. In the COVID-era this data limitation becomes prohibitive as commuting and overall transportation miles become completely disentangled. Finally, it is worth noting that while initially it seemed like there might be a permanent structural shift to remote work and a decline of cities (e.g. Gupta et al. 2022), as of 2024 it appears that many employers are requiring workers to return to the office (e.g. Resume Builder 2023), calling into question whether the pandemic will have had a long-term impact on cities. The net effect of all these countervailing forces, and the extent to which they result in a permanent, structural shift in place effects, is an empirical question which this paper does not have enough data to address at this time, but is an important avenue for future research.

G Additional Figures and Tables

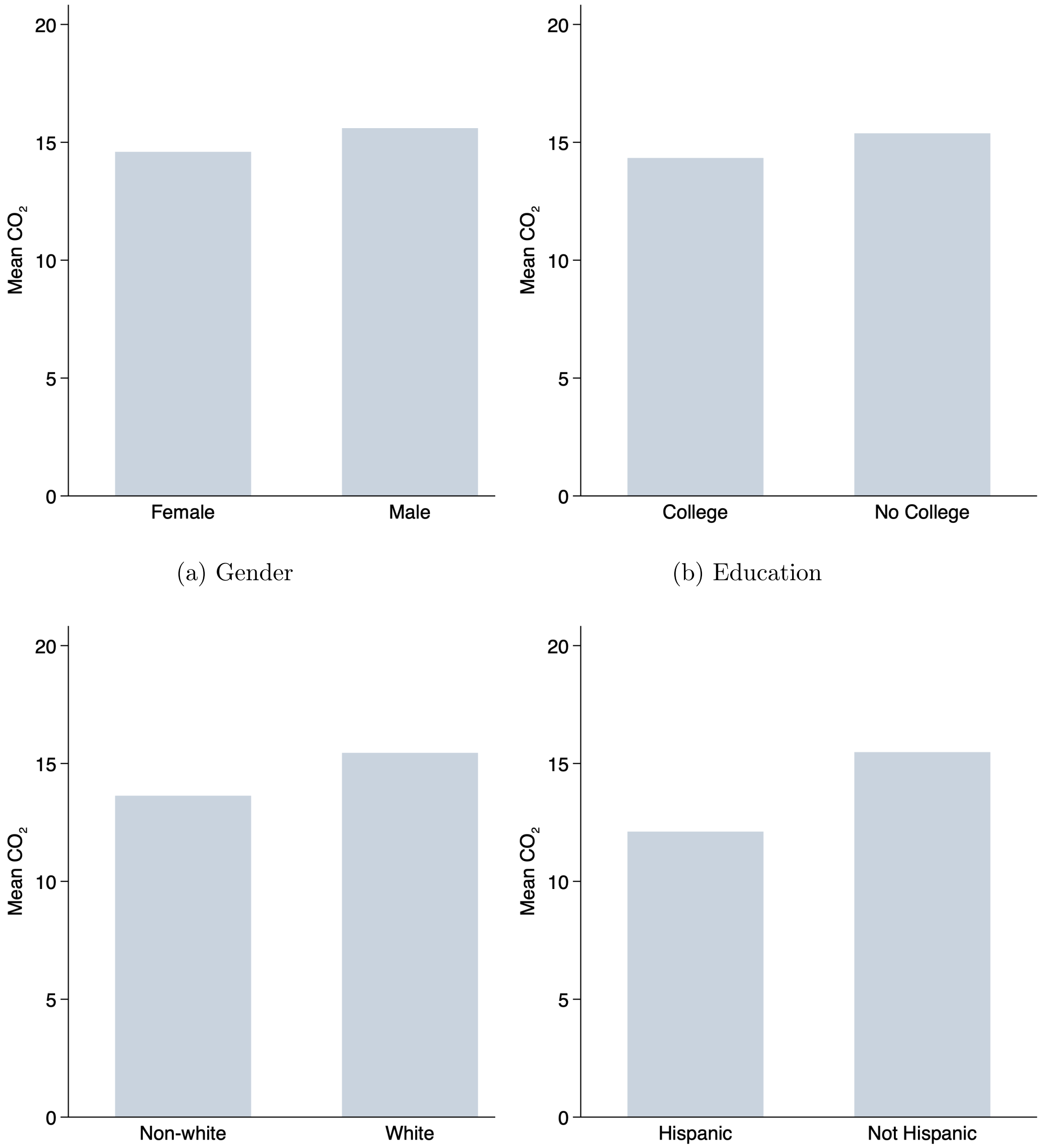
G.1 Additional Figures

Figure G.1: **Heterogeneity in Household Carbon Emissions**



Note: This figure shows Kernel Density estimates, using a Gaussian kernel function, of de-meaned household carbon emissions. Household carbon emissions are censored at the top and bottom 1% of observations in order to abide by Census Disclosure Avoidance rules. The dotted gray line labeled "Without Controls" corresponds to the distribution of log CO_2 conditional on year FEs only, and has a standard deviation of 0.59, while the solid line labeled "With Controls" conditions on observable household characteristics, and has a standard deviation of 0.52. Observable characteristics include age, gender, race, ethnicity, education, home owner status, household income, household size, and number of children.

Figure G.2: CO_2 Profiles by Demographic Characteristics (1/4)



(a) Gender

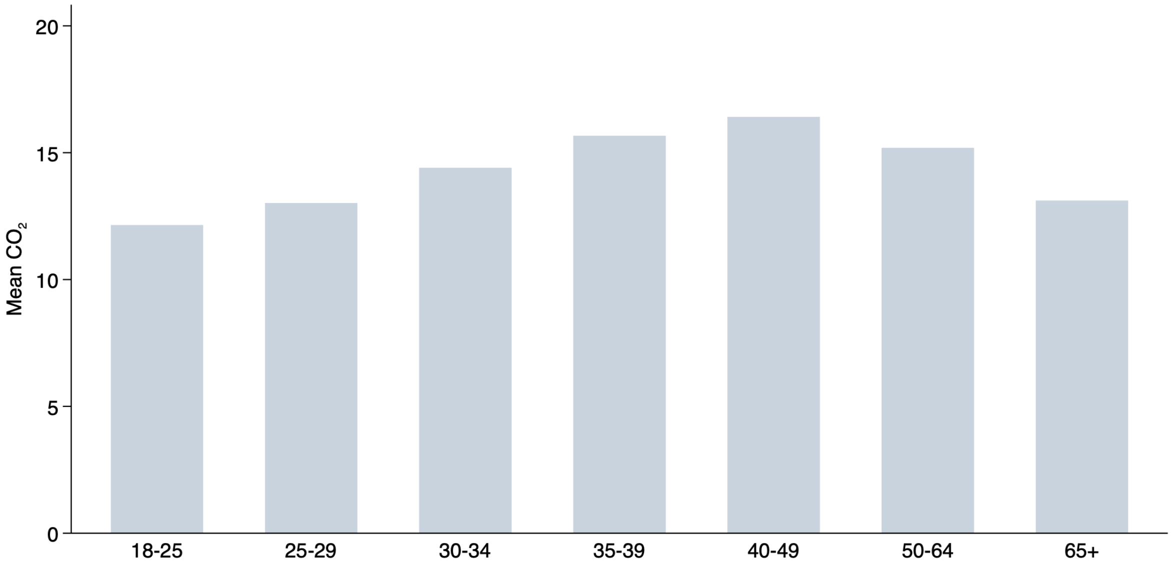
(b) Education

(c) Race

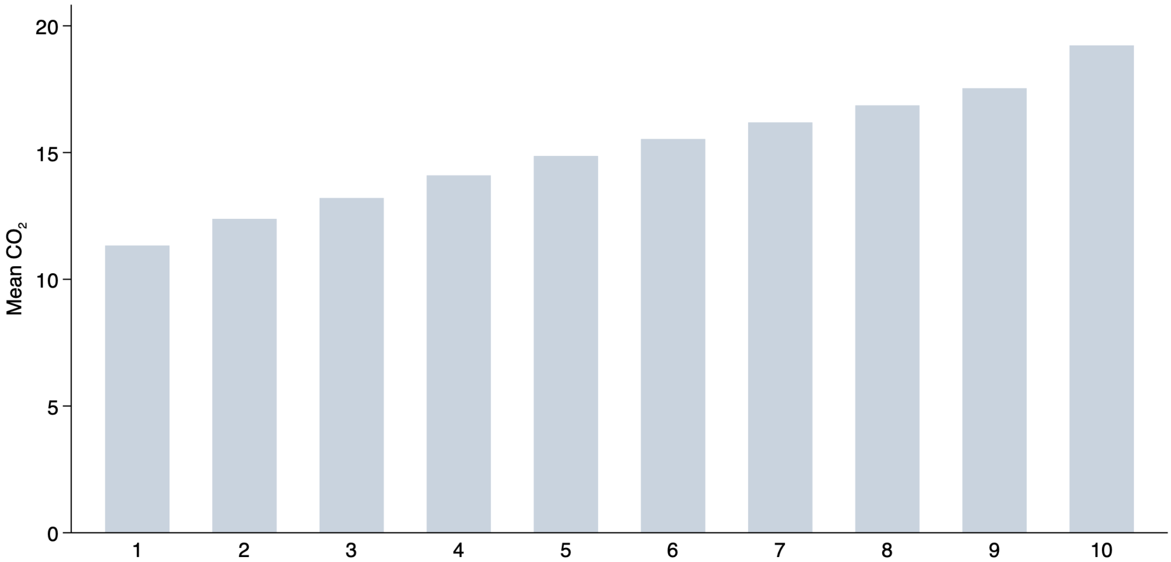
(d) Ethnicity

Note: This figure shows variation in household carbon emissions by household member demographics. Panel (a) shows that households with more women (age 18+) have slightly lower emissions (consistent with women having fewer and shorter commutes). Panel (b) shows that college educated households have slightly lower emissions. Panel (c) and (d) show large differences by race and ethnicity – white households and non-Hispanic households have higher emissions on average than non-white and Hispanic households. All estimates reflect the full sample, pooled 2000-2019, weighted by Census sample weights.

Figure G.3: CO_2 Profiles by Demographic Characteristics (2/4)



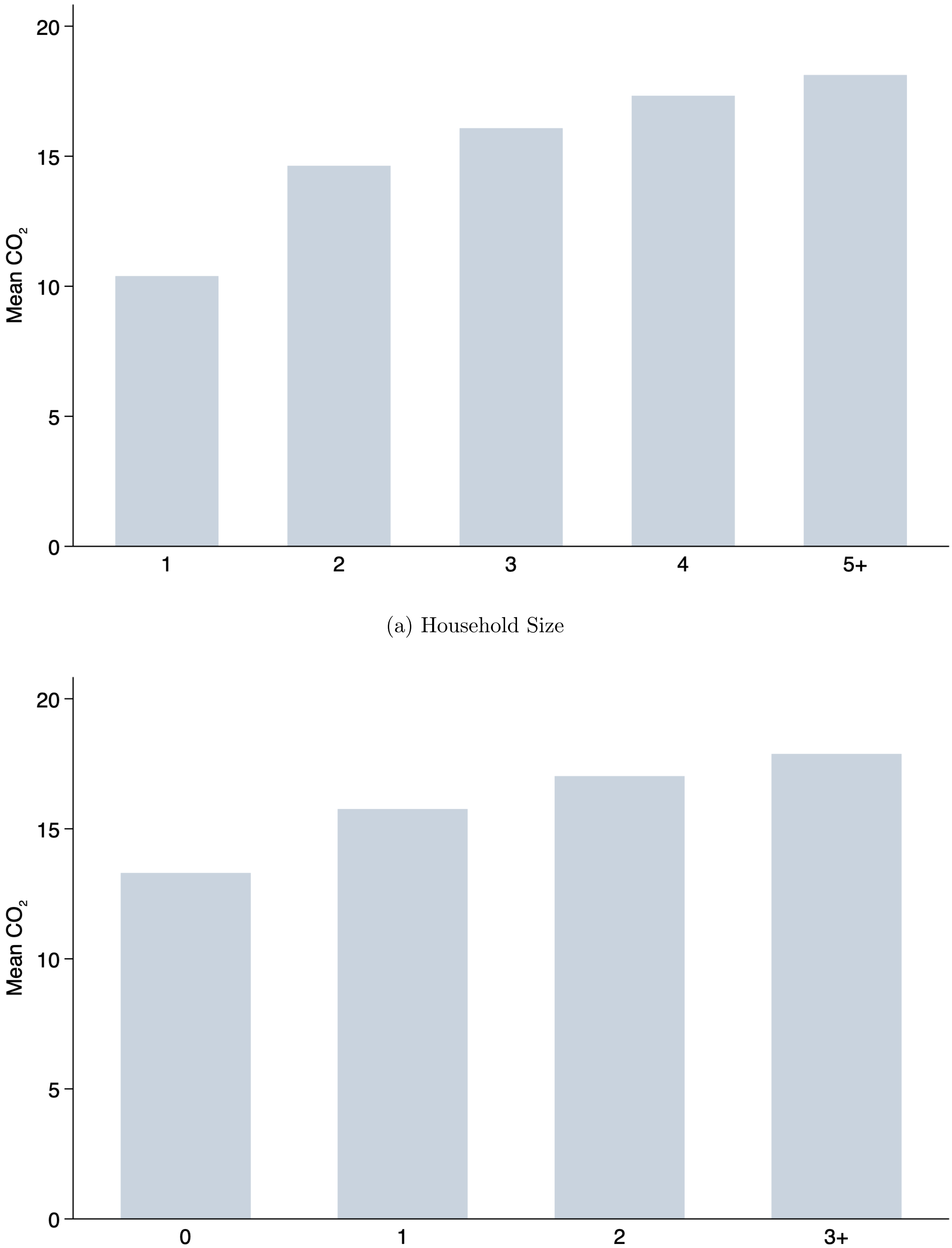
(a) Age



(b) Income Decile

Note: This figure shows variation in household carbon emissions by household member age and household income deciles. Panel (a) shows a non-linear relationship between the adult age of household members and mean carbon emissions which increases through people's 40s and then decreases again (likely reflecting a combination of higher incomes and children still being in the home). Panel (b) shows an increasing relationship between household income decile and carbon emissions. All estimates reflect the full sample, pooled 200-2019, weighted by Census sample weights. Household income is CPI-adjusted.

Figure G.4: CO_2 Profiles by Demographic Characteristics (3/4)

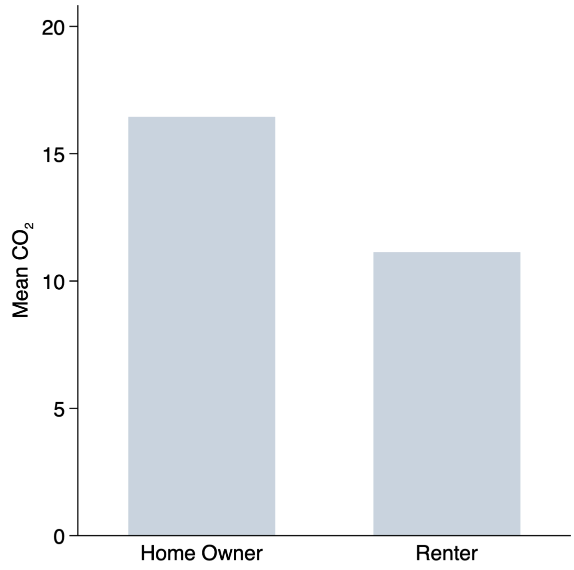


(a) Household Size

(b) Number of Kids

Note: This figure shows variation in household carbon emissions by household size (a) and number of children (b). Carbon emissions increase with household size and with the number of children, but less than proportionally, and the increase is fairly small going from 4 to 5+ people, or 2 to 3+ kids. All estimates reflect the full sample, pooled 200-2019, weighted by Census sample weights.

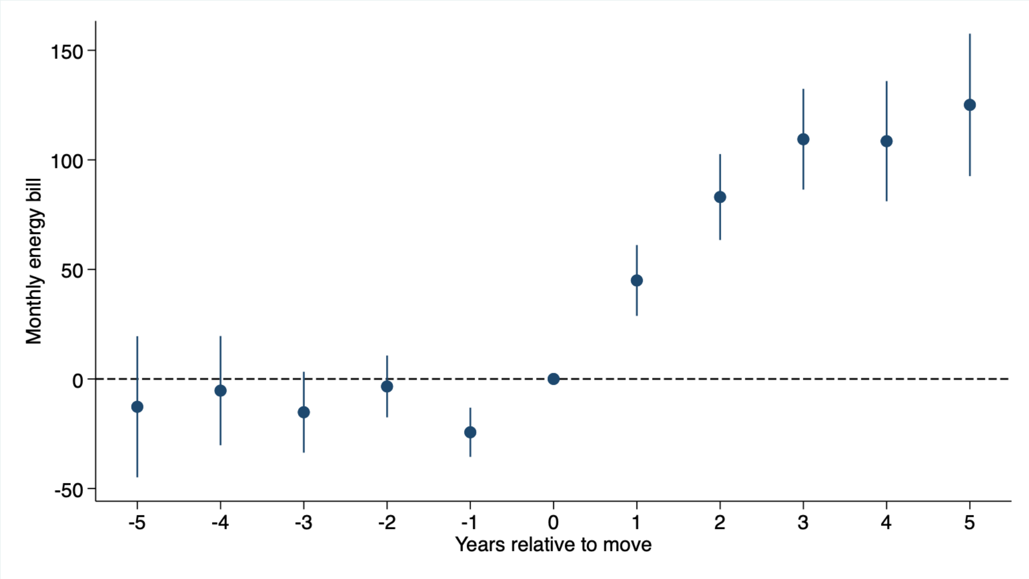
Figure G.5: CO_2 Profiles by Demographic Characteristics (4/4)



(a) Home Ownership

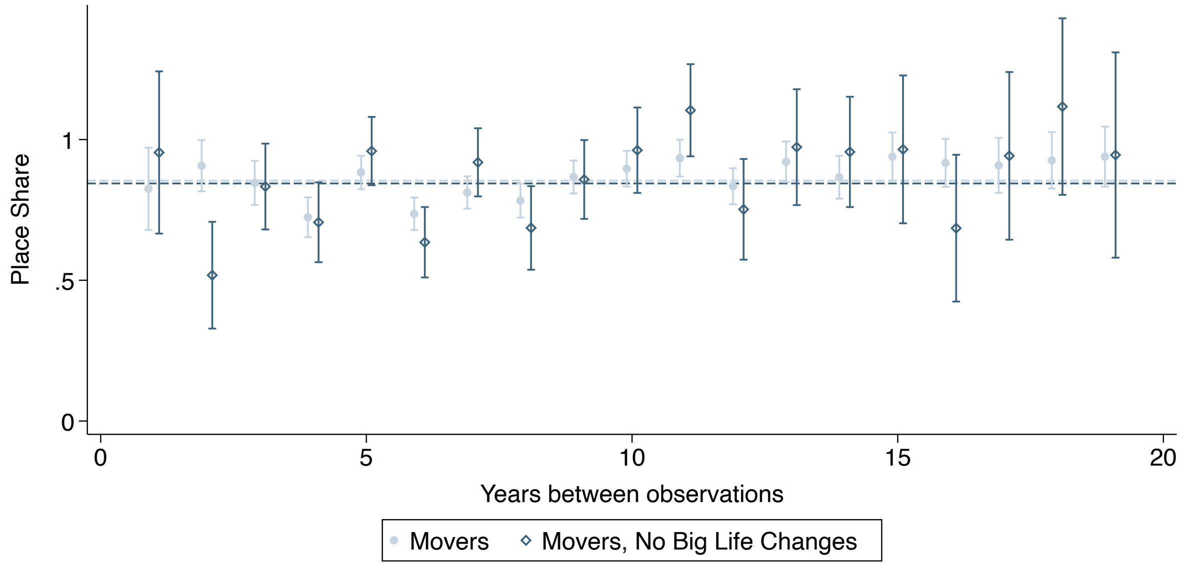
Note: This figure shows variation in household carbon emissions by homeowner status, highlighting that renters have lower emissions on average than homeowners. All estimates reflect the full sample, pooled 200-2019, weighted by Census sample weights.

Figure G.6: Energy Expenditures in Mover Households in the PSID



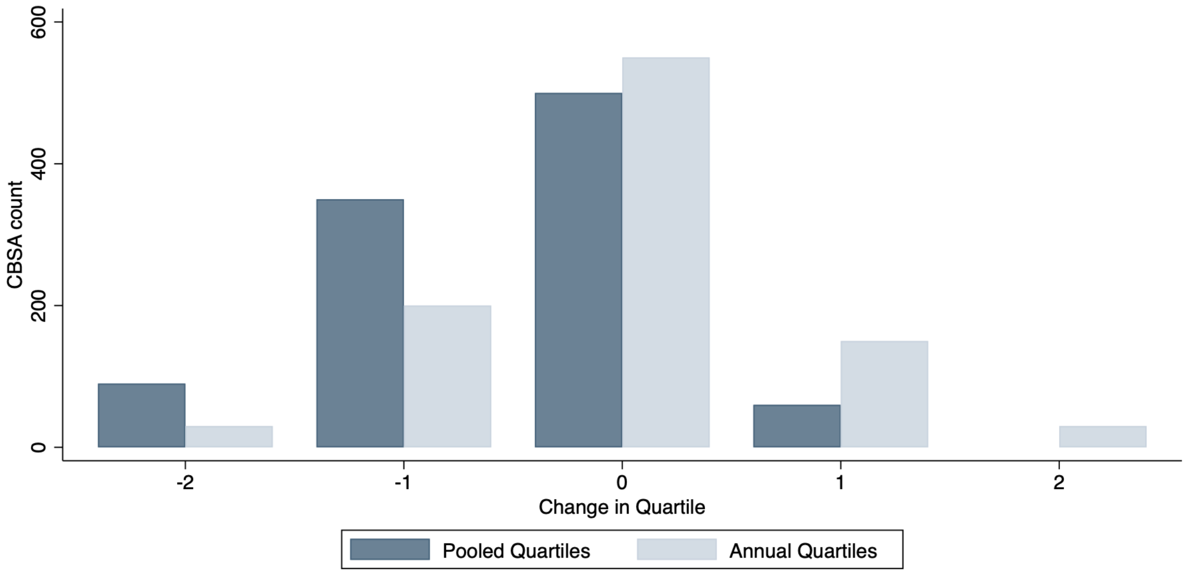
Note: I examine whether there are pre-trends in energy consumption for movers using data from the PSID, given data limitations in my baseline data. In particular, I test whether there are significant changes to monthly energy bills in the years prior to a move, after controlling for household characteristics such as income and household size. If anything, I find a slightly countervailing pre-trend for movers, with energy bills decreasing in the year before a move, and then increasing in the several years after (consistent with a secular trend of households moving to higher emissions places).

Figure G.7: **Event study by duration – CBSA**



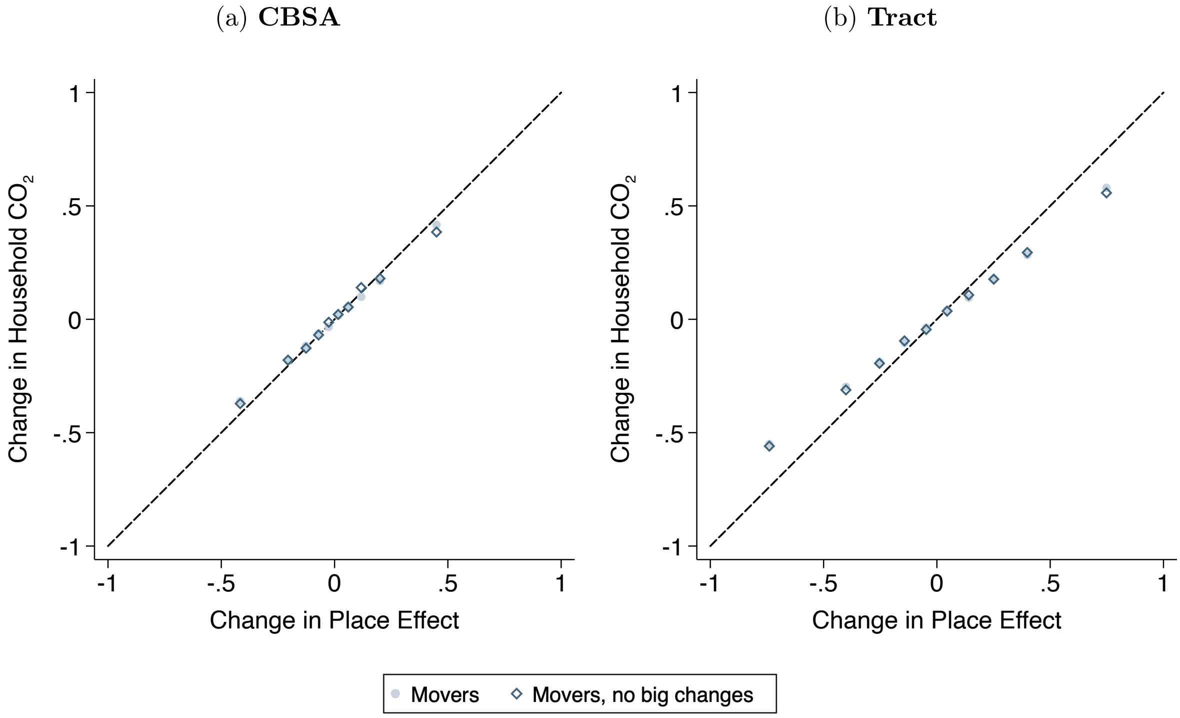
Note: This figure shows event study estimates of the share of spatial variation in mean carbon emissions that can be explained by place effects, by duration between mover observations. In other words, each coefficient is the estimate for place effects generated from the sub-sample of households that I observe X years apart. Coefficients plotted in light gray circles are estimated from the model using the entire sample of movers. Coefficients plotted in the dark blue diamonds are estimated from the model using the sub-sample of movers with no change in the number of children, a less than 0.5 log point change in household income, and no change in home-ownership status between observations. All estimates are weighted using Census sample weights.

Figure G.8: **Changes in Time-Varying CBSA Effect Ranks**



Note: This figure shows the distribution of rank changes in time varying CBSA effects from the 2000-2004 period to the 2015-2019 period. The dark blue bars show changes in pooled quartiles, while the light blue bars show changes in within-year quartiles.

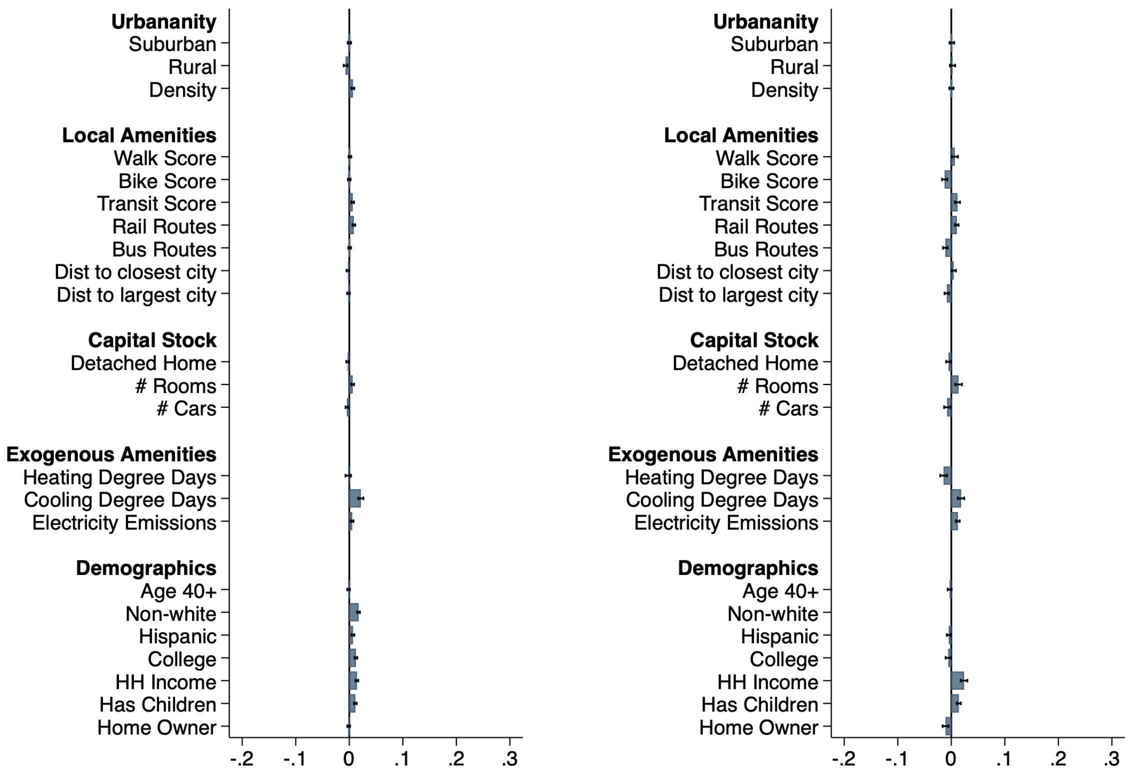
Figure G.9: **Place Effects vs. Household Carbon Emissions**



Note: This figure shows event study estimates of the share of spatial variation in mean carbon emissions that can be explained by place effects, by size of origin-destination differences in the KSS estimates of place effects. The two sets of points compare the full sample of movers (solid light grey circle) to the sub-sample of movers with no change in the number of children, a less than 0.5 log point change in household income, and no change in home-ownership status between observations (empty dark blue diamond). The dotted black line shows the 45°line. All estimates are weighted using Census sample weights.

Figure G.10: **Correlates of Unobserved Household Heterogeneity**

(a) Bivariate Regressions (b) Single Regression



Note: This figure presents estimates from OLS regressions of estimated tract effects on a set of observable place-based and household characteristics. Panel (a) shows results from separate bivariate regressions, while panel (b) shows results from a single regression on all covariates. All amenity variables are tract level means, normalized to have mean zero and standard deviation one, except the rural and suburban indicators, which are retained as indicators. Regressions are weighted using ACS sample weights.

G.2 Additional Tables

Table G.1: **Summary Statistics for Sample Dropped Due to Missing Energy Info**

	Full				Panel	Mover	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	All	Elec in Rent	Gas in Rent	Gas in Elec.	All	CBSA	Tract
A: Demographics							
College	0.31	0.21	0.28	0.33	0.33	0.41	0.36
Age	42	40	39	43	44	40	41
White	0.72	0.67	0.68	0.74	0.83	0.83	0.81
Female	0.48	0.46	0.48	0.48	0.48	0.45	0.48
Household Income	85,010	51,800	65,310	98,860	100,800	102,700	100,100
Household Kids	0.9	0.7	0.6	1.0	0.9	0.9	0.9
Household Size	2.6	2.2	2.1	2.8	2.7	2.7	2.6
Homeowner	0.47	0.10	0.16	0.67	0.65	0.53	0.55
B: Outcomes							
Tons CO_2 - Commute	2.5	1.7	1.8	3.0	2.6	2.7	2.5
C: Intermediate Outcomes							
Detached Home	0.44	0.15	0.08	0.65	0.60	0.54	0.52
Use Electricity Only	0.04	0.24	0	0	0.09	0.14	0.12
Commute by Car	0.83	0.70	0.70	0.89	0.88	0.88	0.88
Commute Minutes	25.4	23.9	25.4	25.7	25.0	25.0	25.5
D: Place Characteristics							
Urban	0.34	0.44	0.47	0.28	0.26	0.25	0.28
Suburban	0.18	0.15	0.19	0.18	0.15	0.12	0.15
Rural	0.48	0.41	0.34	0.54	0.59	0.64	0.57
Walk Score	46.1	53.8	59.3	40.2	38.3	36.2	39.7
Bike Score	47.5	52.3	55.6	43.8	43.2	43.5	44.5
Transit Score	18.8	23.7	26.6	15.2	14.4	14.0	15.9
N Bus Routes	4.7	6.9	7.9	3.3	3.5	3.5	3.8
N Rail Routes	0.93	1.24	1.84	0.54	0.63	0.55	0.65
Cooling Degree Days	947	1,110	924	919	932	1,062	983
Heating Degree Days	5,028	4,754	5,236	4,993	5,272	5,021	5,151
N People	1,810,000	389,000	722,000	980,000	165,000	24,500	68,000
N Households	1,410,000	322,000	593,000	721,000	272,000	44,500	121,000
CBSAs	1,000	1,000	1,000	950	950	950	950
Tracts	68,500	53,000	56,500	55,000	49,000	26,000	43,000

Note: This table shows summary statistics for households dropped from the analysis as a result of having their electricity bills included in rent, their natural gas bills included in rent, or their natural gas bills included in their electricity bills. Column (1) shows statistics for the entire set of households who would've been in the full sample but got dropped for any one of those three reasons. Columns (2)-(4) show summary statistics broken out by group. Column (5) shows summary statistics for the sub-sample of column (1) who would've been in the panel sample if not for these unobserved bills, and columns (6) and (7) show households who would have been in the mover sample. All statistics are weighted by ACS household weights.

Table G.2: **Panel Statistics for Sample Dropped Due to Missing Energy Info**

		Movers	
	(1)	(2)	(3)
	Panel	CBSA	Tract
A: Sample Characteristics			
First Observed in 2000	0.08	0.13	0.12
Years Between Observations	8.6	10.4	9.9
B: Demographic Characteristics			
Age First Observed	40.6	34.8	35.9
Share with Large Change in Income	0.43	0.62	0.56
Share with Change in N Kids	0.44	0.51	0.51
Change in N Kids	0.03	0.30	0.23
Share Rent to Own	0.22	0.39	0.38
C: Mover Place Changes			
Δ Walk Score		-7.8	-7.7
Δ Bike Score		-5.1	-4.9
Δ Transit Score		-2.7	-3.6
Δ N Bus Routes		-0.96	-1.22
Δ N Rail Routes		-0.06	-0.15
Δ Tract Share Detached Home		0.07	0.08
% Moves Urban-to-Urban		0.19	0.28
% Moves Urban-to-Suburban		0.22	0.21
% Moves Suburban-to-Suburban		0.34	0.32
%Δ Cooling Degree Days		146	122
%Δ Heating Degree Days		-265	-159
N People	165,000	24,500	68,000
N Households	142,000	22,000	60,000
CBSAs	950	950	950
Tracts	49,000	26,000	43,000

Note: This table shows panel statistics for households dropped from the main analysis as a result of having their electricity bills included in rent, their natural gas bills included in rent, or their natural gas bills included in their electricity bills. The Column (1) shows statistics for households who would have been in the panel if not for unobserved billing information, while Columns (2)-(3) show statistics for households who would have been in the mover sample. All statistics are weighted by ACS household weights.

Table G.3: **Mean CO_2 – Movers vs. Stayers**

	CBSA Panel				Tract Panel			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mover	-0.07***	-0.05***			-0.11***	-0.08***		
	(0.001)	(0.001)			(0.001)	(0.001)		
Mover x Orig.			-0.12***	-0.05***			-0.08***	-0.04***
			(0.002)	(0.002)			(0.001)	(0.001)
Mover x Dest.			-0.04***	-0.04***			-0.03***	-0.03***
			(0.001)	(0.001)			(0.001)	(0.001)
Cons.	2.64***	2.58***	2.64***	2.59***	2.67***	2.61***	2.65***	2.61***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)	(0.000)
Controls	No	Yes	No	Yes	No	Yes	No	Yes

Note: This table compares household carbon emissions for movers and stayers. Columns (1)-(2) and (5)-(6) compare movers overall to stayers overall, with and without controls. Movers have lower carbon emissions than stayers, with a slightly less pronounced difference after controlling for differences in income and other demographic characteristics. Columns (3)-(4) and (7)-(8) present within-comparisons of stayers and movers within a given place. The "Mover x Orig." coefficient compares movers with stayers at their origin, while the "Mover x Dest." coefficient compares movers with stayers at their destination. Movers have lower emissions than stayers at both their origin and their destination. The origin difference looks more pronounced in the specifications without controls, but is effectively the same as the destination difference after controlling for observable household characteristics. All estimates are weighted by ACS household sample weights.

Table G.4: **Probability of Moving**

	(1) Moved CBSA	(2) Moved Tract
Decrease in Kids	0.02***	0.05***
	(0.001)	(0.001)
Increase in Kids	0.06***	0.18***
	(0.001)	(0.001)
Large Decrease in Income	0.06***	0.13***
	(0.001)	(0.001)
Large Increase in Income	0.09***	0.16***
	(0.001)	(0.001)
Rent $\rightarrow$ Own	0.16***	0.48***
	(0.001)	(0.002)
Constant	0.06***	0.19***
	(0.000)	(0.001)
R^2 (adj.)	0.05	0.17

Note: This table shows that households with a change in the number of children at home, a larger than 0.5 log point change in income, or who go from renting to owning are much more likely to move than stay. This is most pronounced for households who go from renting to owning their home, and is also more pronounced for positive changes in children or income than negative changes. All estimates are weighted by ACS household sample weights.

Table G.5: **Mover Origin and Destination Types**

(a) CBSA Movers

	To Rural	To Suburban	To Urban	Total Share
From Rural	0.03	0.09	0.01	0.13
From Suburban	0.07	0.44	0.09	0.60
From Urban	0.01	0.17	0.08	0.26
Total Share	0.11	0.50	0.18	1.00

(b) Tract Movers

	To Rural	To Suburban	To Urban	Total Share
From Rural	0.03	0.06	0.01	0.10
From Suburban	0.05	0.46	0.09	0.60
From Urban	0.01	0.17	0.14	0.32
Total Share	0.09	0.69	0.24	1.00

Note: This table shows shares of origin-destination tract types for CBSA movers (panel (a)) and tract movers (panel (b)). The most common type of move, for both CBSA and tract movers, is from a suburban tract to another suburban tract. Moves between urban and rural tracts are exceedingly uncommon. All estimates are weighted by ACS household sample weights.

Table G.6: **Event Study – with Climate and Electricity Emissions Controls**

	CBSA		Tract	
	(1)	(2)	(3)	(4)
A: Panel Sample				
Place share of mean difs.	0.86***	0.70***	0.60***	0.54***
	(0.007)	(0.012)	(0.003)	(0.004)
N	1,764,000	1,764,000	1,710,000	1,710,000
R^2 (adj.)	0.75	0.75	0.76	0.77
B: Mover Sample				
Place share of mean difs.	0.85***	0.68***	0.57***	0.51***
	(0.009)	(0.014)	(0.004)	(0.004)
N	191,000	191,000	508,000	508,000
R^2 (adj.)	0.70	0.71	0.73	0.73
Household controls	X	X	X	X
Climate & electricity controls		X		X

Note: This table reports event study estimates of the place share of spatial heterogeneity in household carbon emissions. The place share estimate ($\hat{\theta}$) represents the proportion of differences in average carbon emissions ($\bar{y}$) between a mover's origin and destination attributable to place effects. Panel A reports estimates from the panel sample, while panel B restricts the sample to movers only, allowing for systematic differences between movers and stayers. Columns (1) and (3) replicate the baseline analysis presented in Table 3. Columns (2) and (4) add controls for mean heating degree days, mean cooling degree days, and mean electricity emissions factors. All estimates use Census sample weights.

Table G.7: **Place-Based Heterogeneity in CO_2 – Climate vs Electricity Emissions**

	CBSA				Tract			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel Sample								
Variance of log(CO_2)	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31
Share attributable to places	0.163	0.146	0.080	0.074	0.228	0.217	0.151	0.152
Share attributable to hhs	0.496	0.496	0.494	0.494	0.363	0.362	0.362	0.361
Corr. of place and hh effects	0.013	0.036	-0.003	0.027	0.016	0.070	0.024	0.081
SD of place effects	0.23	0.21	0.16	0.15	0.26	0.26	0.22	0.22
Mover Sample								
Variance of log(CO_2)	0.35	0.35	0.35	0.35	0.33	0.33	0.33	0.33
Share attributable to places	0.140	0.125	0.046	0.039	0.218	0.220	0.145	0.155
Share attributable to hhs	0.136	0.129	0.160	0.156	0.099	0.098	0.102	0.101
Corr. of place and hh effects	0.073	0.102	0.048	0.084	0.084	0.160	0.084	0.161
SD of place effects	0.22	0.21	0.13	0.12	0.27	0.27	0.22	0.23
Climate		X		X		X		X
Electricity CO_2			X	X			X	X

Note: This table reports KSS estimates of variance components, deliniating between the contribution of local climate conditions vs the contribution of local electricity emissions factors. All specifications include demographic and household controls as well as time fixed effects. To ease comparison, Columns (1) and (5) replicate baseline estimates shown in columns (1) and (5) of Table 4, while columns (4) and (8) of this table replicate estimates accounting for the contribution of both climate and electricity emissions factors simultaneously (i.e. columns (2) and (6) of Table 4). Columns (2) and (6) of this table show estimates accounting for just the role of climate in the variance components, while Columns (3) and (7) show estimates accounting for just the role of electricity carbon emissions. All estimates are weighted by ACS household sampling weights.

Table G.8: **Place-Based Heterogeneity in CO2 – No Bias Correction**

	CBSA				Tract		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel Sample							
Variance of log(CO_2)	0.31	0.31	0.31	0.31	0.31	0.31	0.31
Share attributable to places	0.175	0.086	0.090	0.188	0.523	0.449	0.453
Share attributable to hhs	0.558	0.554	0.554	0.557	0.748	0.747	0.746
Corr. of place and hh effects	-0.026	-0.028	-0.027	-0.026	-0.417	-0.429	-0.427
SD of place effects	0.23	0.16	0.17	0.24	0.40	0.37	0.37
Mover Sample							
Variance of log(CO_2)	0.35	0.35	0.35		0.33	0.33	0.33
Share attributable to places	0.153	0.051	0.053		0.461	0.399	0.403
Share attributable to hhs	0.505	0.502	0.502		0.582	0.581	0.581
Corr. of place and hh effects	0.008	-0.008	-0.005		-0.291	-0.298	-0.294
SD of place effects	0.23	0.13	0.14		0.39	0.36	0.37
Climate + Electricity CO_2		X	X			X	X
Price Index			X				X
Time-Varying FEs				X			

Note: This table reports results from the biased AKM estimation of variance components. All specifications include demographic and household controls as well as time fixed effects. Columns (1) and (5) report the baseline variance decompositions at the CBSA and tract levels. Columns (2) and (5) add controls for local mean heating degree days, cooling degree days, and electricity emissions factors (all in logs). Columns (3) and (6) additional control for a price index, constructed from lagged fuel shares interacted with national retail prices. Finally, column (4) computes time-varying CBSA place effects using 5-year windows (2000-2004, 2005-2009, 2010-2014, and 2015-2019), using stayer observations across time windows to identify time variation in place effects, while movers, as before, identify cross-sectional variation. All estimates are weighted by ACS household sampling weights.

Table G.9: **Place-Based Heterogeneity in CO_2 – Alternate Outcome Definitions**

	CBSA				Tract			
	V(y)	S(ψ_j)	S(α_i)	corr.	V(y)	S(ψ_j)	S(α_i)	corr.
Panel Sample								
Baseline	0.31	0.163	0.496	0.013	0.31	0.228	0.363	0.016
Electricity Emissions Estimates								
Marginal Emissions	0.34	0.236	0.470	-0.016	0.337	0.295	0.343	-0.011
Variable Prices	0.33	0.155	0.498	0.018	0.321	0.218	0.363	0.026
Decreasing Block Prices	0.35	0.150	0.503	0.016	0.341	0.216	0.373	0.011
Increasing Block Prices	0.31	0.186	0.480	0.013	0.303	0.246	0.345	0.033
Selected Block Prices	0.32	0.159	0.500	0.014	0.311	0.220	0.366	0.024
Transportation Emissions Estimates								
Commute from hrs	0.29	0.160	0.487	0.015	0.287	0.206	0.367	0.024
Commute from hrs, fixed num.	0.29	0.168	0.487	0.009	0.286	0.213	0.366	0.020
MPG from NHTS (dem. only)	0.32	0.169	0.495	0.008	0.313	0.235	0.365	0.006
MPG from NHTS (geo. only)	0.32	0.173	0.493	0.011	0.313	0.240	0.363	0.007
MPG from NHTS	0.32	0.174	0.490	0.011	0.312	0.238	0.360	0.010
Total Transportation from NHTS	0.20	0.177	0.468	0.037	0.200	0.218	0.376	0.015
Mover Sample								
Baseline	0.35	0.140	0.136	0.073	0.333	0.218	0.099	0.084
Electricity Emissions Estimates								
Marginal Emissions	0.39	0.219	0.136	0.004	0.378	0.281	0.094	0.025
Variable Prices	0.36	0.133	0.135	0.077	0.349	0.209	0.100	0.088
Decreasing Block Prices	0.39	0.131	0.136	0.069	0.372	0.207	0.105	0.062
Increasing Block Prices	0.34	0.161	0.137	0.067	0.329	0.240	0.097	0.091
Selected Block Prices	0.35	0.138	0.134	0.073	0.336	0.211	0.097	0.086
Transportation Emissions Estimates								
Commute from hrs	0.32	0.142	0.132	0.070	0.310	0.202	0.101	0.089
Commute from hrs, fixed num.	0.32	0.149	0.132	0.063	0.309	0.210	0.099	0.075
MPG from NHTS (dem. only)	0.36	0.145	0.138	0.070	0.341	0.221	0.102	0.077
MPG from NHTS (geo. only)	0.36	0.150	0.138	0.076	0.341	0.227	0.100	0.083
MPG from NHTS	0.35	0.150	0.138	0.075	0.341	0.225	0.097	0.089
Total Transportation from NHTS	0.21	0.169	0.171	0.0772	0.207	0.214	0.118	0.0862

Note: This table reports KSS estimates of variance components, using a variety of different outcome definitions to test robustness of the baseline estimates. Estimates are reported for both the full panel sample (top half of table) and the mover only sample (bottom half of table). Each outcome definition is a row in the table, with baseline estimates replicated in the first row of each sample to ease comparability. Outcome variants are grouped into two categories: one which impacts residential carbon emission estimates, and one which estimates transportation emissions estimates. Overall variance of the outcome, the share attributable to place effects, the share attributable to person effects, and the correlation between place and household effects are reported at the CBSA level in columns (1)-(4), respectively, and at the tract level in columns (5)-(9). All estimates are weighted by ACS household sampling weights.

Table G.10: **Place Correlates w/ Observable Characteristics**

	College	Age > 40	Non-white	HH Income	Has Kids	Homeowner
Density	-0.002*	0.004***	0.006***	0.033***	0.017***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Suburban	0.014***	0.005***	-0.025***	0.015***	-0.010***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Rural	0.001	0.009***	-0.025***	-0.028***	-0.019***	-0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Walk Score	0.002*	0.006***	-0.022***	0.025***	0.007***	-0.021***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
Bike Score	0.025***	-0.011***	-0.004***	0.027***	-0.006***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Transit Score	0.006***	0.004***	0.027***	0.030***	-0.007***	0.006***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
CDD	-0.015***	0.000	-0.063***	-0.044***	0.003	0.025***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
HDD	-0.030***	-0.003***	-0.103***	-0.070***	-0.012***	0.033***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.001)
Elec. CO_2	-0.018***	-0.020***	0.008***	-0.062***	-0.010***	-0.004***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)
N Rail Routes	0.007***	0.002***	-0.012***	0.036***	0.002	0.010***
	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)
N Bus Routes	0.016***	-0.003***	-0.012***	0.016***	-0.013***	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Detached Home	-0.043***	0.009***	-0.000	-0.085***	0.013***	0.047***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N Rooms	0.121***	0.055***	-0.004***	0.309***	0.032***	0.071***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
N Vehicles	-0.036***	0.017***	-0.055***	0.076***	0.043***	0.041***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Dist. Closest City	-0.013***	-0.005***	-0.011***	-0.041***	-0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Dist. Largest City	0.005***	0.012***	0.010***	0.064***	0.007***	0.007***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.25***	0.63***	0.14***	11.29***	0.53***	0.78***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)
Adj. R^2	0.451	0.465	0.326	0.643	0.155	0.767

Note: This table reports correlation coefficients between tract-level mean observable household characteristics and a detailed vector of observable place characteristics. All estimates are weighted by ACS household sampling weights.

Table G.11: **10 most populous CBSAs (2020)**

Rank	CBSA
1	New York-Newark, NY-NJ-CT-PA
2	Los Angeles-Long Beach, CA
3	Chicago-Naperville, IL-IN-WI
4	Dallas-Fort Worth, TX-OK
5	Houston-The Woodlands, TX
6	Washington-Baltimore-Arlington, DC-MD-VA-WV-PA
7	Philadelphia-Reading-Camden, PA-NJ-DE-MD
8	Miami-Port St. Lucie-Fort Lauderdale, FL
9	Atlanta-Athens Clarke County-Sandy Springs, GA-AL
10	Boston-Worcester-Providence, MA-RI-NH-CT

This table presents the ten most populous CBSAs as of 2020, which are used in the analysis evaluating how overall emissions would change under different distributions of place effects. Source: https://en.wikipedia.org/wiki/Combined_statistical_area