Urban Sprawl and Residential Carbon Emissions: Evidence from Indonesia and the Philippines*

Yi Jiang
Asian Development Bank

Taylor Lathrop
Syracuse University

Alexander D. Rothenberg†
Syracuse University

Yao Wang
The Ohio State University

December 2025

Abstract

This paper studies how neighborhood density affects residential carbon emissions, using detailed data from Indonesia and the Philippines. To address simultaneity, we instrument density with soil characteristics, and to address sorting, we control for community averages of observed characteristics. Unlike cities in developed countries, we find that density is positively correlated with residential energy use. After controlling for sorting, we find a precise null relationship between density and residential carbon emissions. Our results suggest that policies to control urban sprawl may not be successful in reducing residential carbon emissions in developing country cities.

JEL Classifications: R11, O13, O18, Q41, L94, Q54

Keywords: urban sprawl, carbon emissions

*Acknowledgement: Katrina Miradora and Marjorie Remolador provided excellent research assistance. We are grateful to seminar participants at the Asian Development Bank (ADB), AREUEA, Fordham University, Syracuse University, and the UEA for helpful comments and suggestions. The views expressed in this paper are those of the authors and do not necessarily reflect the views and policies of ADB, its Board of Governors or the governments they represent. ADB does not guarantee the accuracy of the data included in this paper and accepts no responsibility for any consequence of their use. All errors remain our own.

†Corresponding author: 426 Eggers Hall, Syracuse, NY 13244-1020. Email: adrothen@maxwell.syr.edu.

> But as India and China get richer, their people will face a choice that could dramatically affect all our lives. Will they follow America and move towards car-based exurbs or stick with denser urban settings that are far more environmentally friendly? ... Driving and urbanization patterns in these countries may well be the most important environmental issues of the twenty-first century.
>
> — Glaeser (2011)

1 Introduction

Over the next 30 years, much of the world's growth in energy demand is expected to come from developing countries. According to the EIA (2019), energy consumption in non-OECD countries is expected to increase by nearly 70 percent from 2020-2050, in contrast to an expected increase of 15 percent in OECD countries over the same period. Some of the rise in energy demand will be fueled by poverty reduction. As the world's poor grow increasingly affluent, they will invest in durable household appliances and motor vehicles, and these extensive margin purchases may sustain increased energy consumption for decades (Wolfram et al., 2012; Gertler et al., 2016).

The same rising incomes that trigger rising energy demand in the developing world are also increasing urban sprawl. Globally on average, cities are expanding spatially twice as fast as their population growth rates (Angel et al., 2011). Driven by motorization, the rapid sprawl in developing country cities raises concerns that low-density development may further exacerbate residential carbon emissions.

As cities sprawl, they can enable higher income commuters to escape from the massive levels of density and crowding experienced in inner cities (Henderson and Turner, 2020). In developed countries, when more affluent households sort into the suburbs, they often purchase larger homes that need to be cooled in hot weather, and they need to commute farther to their jobs, often using private transport modes instead of public transit alternatives. As a consequence, the newly constructed, large homes in the suburbs could be particularly harmful for the environment. If rapid sprawl is associated with rising carbon intensity, policies to regulate the spatial extent of cities could be justified as ways to mitigate negative environmental externalities. Such policies—which are often second-best to carbon pricing—could include growth controls, binding limits on new construction, or open space dedications (Cunningham, 2007; Glaeser and Ward, 2009; Brueckner and Sridhar, 2012).

At the same time, income sorting patterns vary greatly across cities throughout the world (Brueckner et al., 1999). Many central cities in developing countries are home to higher income households, while slums and lower income housing lie on the outskirts of those cities (Deffebach et al., 2025). The relative locations of the rich and the poor within cities are driven by the income elasticity of housing demand, the income elasticity of commuting expenditures, and differences in amenities, and as these factors differ across cities, they may alter the relationship between density and carbon emissions.

In this paper, we investigate how urban form shapes residential carbon emissions, using the experiences of sprawling cities in Indonesia and the Philippines as a case study. Indonesia and the Philippines are both lower-middle income countries that have experienced rapid industrialization. Because both countries have electricity grids that rely heavily on coal, they have some of the most emissions-intensive electricity sectors in the world (IEA, 2022a). Residential energy demand and transportation also comprise large shares of total energy use in both countries. To quantify household residential carbon emissions, we use detailed cross-sectional data from various waves of both Indonesia's National Socioeconomic Survey (*Survei Sosial Ekonomi Nasional*, or *Susenas*) and the Family Income and Expenditure Survey (FIES) in the Philippines.

Our primary objective is to estimate how within cities, a community's population density affects residential carbon emissions. We want to understand how an average household's emissions would change if that household moved from a dense neighborhood in the center of the city to a low density suburb of the same city. To estimate the causal place effect of density on residential carbon emissions, we confront two fundamental identification challenges. First, omitted, place-specific variables may drive correlations between both density and carbon emissions, creating simultaneity problems. For example, favourable natural amenities may facilitate the construction of power lines and also attract more people.

Prior research has confronted the simultaneity problem by instrumenting for population density using geologic instruments (e.g Black et al., 2002; Rosenthal and Strange, 2008; Combes et al., 2010). We follow a similar approach here, instrumenting density within urban areas using soil characteristics and depth to bedrock. Soil characteristics were determined millions of years ago, and favorable soils lead to the formation of denser settlements because they facilitated greater agricultural productivity historically. We document that these initial advantages within cities persist even today, despite the fact that most workers in cities do not depend on agricultural production for income. We measure soil characteristics and depth to bedrock using data from SoilGrids, a global dataset providing high-resolution measures of many soil properties (Hengl et al., 2017).

Because we work with a large vector of candidate instruments, many of which may be weak on their own, we use post-double-selection lasso techniques to select instruments, following Belloni et al. (2012). This approach delivers the efficiency gains from optimal instruments while reducing problems associated with many instruments. A key concern with our IV strategy is that even within urban areas, soil characteristics could affect residential carbon emissions through channels other than density. We undertake several robustness checks to provide evidence in favor of the exclusion restriction.

After addressing simultaneity, there is still a second identification problem, which is that estimates of the relationship between density and carbon emissions could be confounded with sorting. If higher income individuals sort into denser neighborhoods, as we show that they do in cities in Indonesia and the Philippines, this can bias estimates of the place effects of density. To tackle sorting, we combine instruments for density with controls for community averages of observed individual characteristics. We include a rich vector controls for observed population and demographic characteristics, computed from census microdata in both countries. Altonji and Mansfield (2018) show that under certain assumptions, these controls for sorting on observables will also control for sorting on unobservables. We follow Civelli et al. (2023) who combine their approach—which obtains partial identification of overall group effects—with instruments, to point-identify the place effects of density in a way that is unconfounded by sorting or simultaneity.

We find that in Indonesia and the Philippines, less dense areas within cities tend to contain households with lower total consumption expenditures and lower levels of educational attainment. Consequently, when we estimate the relationship between density and different measures of residential energy use, including consumption quantities of electricity, liquified petroleum gas (LPG), gas for motor vehicles, and kerosene, we find significant positive associations. These correlations also typically survive simultaneity bias in instrumental variables specifications. However, after we introduce controls for sorting, the positive relationship between density and energy consumption tends to disappear. Our point estimates decline, and we generally cannot detect significant effects of density on residential carbon emissions.

We find similar patterns using data on household asset ownership: households in denser neighborhoods of cities in both countries tend to own refrigerators, air conditioners, and LPG cylinders, but once we control for sorting and address simultaneity, the effects of density on asset ownership disappear. We investigate exactly how the sorting controls we use impact our estimates, and we find that controlling for economic sorting, as opposed to ethnic or religious sorting, seems to eliminate significant effects of density on carbon emissions.

Using carbon intensity weights, carefully calculated to reflect differences in electricity grids across countries and cities, we aggregate our coefficient estimates into a single elasticity of carbon emissions with respect to density. We cannot detect a significant relationship between urban density and residential carbon emissions in either Indonesia or the Philippines. Specifically, we find that increasing density by 1 percent leads to a reduction of carbon emissions by -0.005 percent in Indonesian cities and an increase in emissions by 0.004 percent in Philippines cities. Both null results are precise and statistically indistinguishable from zero. Using different waves of survey and census data, we also find that both precise null density elasticity estimates are remarkably stable over time.

Because of concerns surrounding the exclusion restriction, we conduct a number of identification checks, including dropping agricultural households and controlling for historical measures of infrastructure. In both cases, our point estimates remain largely unchanged. We also investigate the effect of soil characterstics in rural areas where they should not predict density, and we find little evidence for any reduced form effects of the soil characteristics on residential carbon emissions outcomes (Altonji et al., 2005; van Kippersluis and Rietveld, 2018). We take the weight of this evidence as favoring the exclusion restriction.

From 2000-2010, total carbon emissions increased by 63 percent in Indonesia, and from 2010-2018, emissions increased by 71 percent in the Philippines (Friedlingstein et al., 2024). Given our lack of significant density effects, we finally attempt to explain drivers of these large increases in emissions. To do so, we use a growth decomposition, which combines our estimated parameters with census data on the average characteristics of households in different communities. We find that apart from a few cities, such as Jakarta, population growth is generally not responsible for large increases in residential carbon emissions. Instead, changes in the relationship between household or neighborhood characteristics and energy demand play a more prominent role in Indonesia, while changes in neighborhood composition play a larger role in cities in the Philippines. Overall, changes in neighborhood density explain only 1.1-1.8 percent of the changes in overall emissions for large cities in Indonesia and the Philippines.

Our work focuses on cities in Indonesia and the Philippines, two large, middle-income countries that together are home to nearly 400 million people, a population larger than that of the United States (see Table 1). Both countries face significant development challenges; despite rapidly urbanizing, 20 percent of urban residents in Indonesia and 37 percnet in the Philippines live in slums. While rising incomes and population growth are driving a substantial increase in energy demand, energy supply in both countries remains heavily dependent on fossil fuels. Examining these two diverse economies, our work is representative of middle-income nations that contain 75 percent of the world's population and 62 percent of the world's poor (Bank, 2024).

We also study the relationship between density and residential carbon emissions in these countries over time. At the start of our observation period (2000), Indonesia's real GDP per capita ($6,190, in PPP terms) was similar to Angola's in 2000 and is roughly equivalent to Ghana's today. Although our study is limited to two countries, the consistently estimated null elasticity of residential emissions with respect to density, which is robust across distinct urban forms, income levels, and time periods, suggests that our findings are likely generalizable to a wide array of lower and upper-middle-income countries navigating different stages of urban and economic development.

In seminal work, Glaeser and Kahn (2010) and Zheng et al. (2011) study the relationship between city location choices and residential carbon emissions, but they tend not to focus on the impact of neighborhood choice within cities. de Thé et al. (2021) use data from French cities to study how urban forms affect car usage and emissions. They find a bell-shaped relationship between city size and car emissions, so that smaller and larger cities tend to be greener than medium-sized cities. Borck and Schrauth (2021) use panel data from German cities and find that air quality decreases with population density, while Carozzi and Roth (2023) use U.S. data to study the effect of urban density on air pollution.[1] Our work complements this literature by extending the focus to a rapidly sprawling developing country setting, and to emphasizing the role of income sorting in mediating the findings.

In a related paper, Lyubich (2025) studies spatial heterogeneity in carbon emissions across census tracts in the U.S. and finds a large role for place effects. Her work uses a movers design to separate place effects from cross-sectional sorting, while we work with multiple waves of cross-sectional data, an instrumental variables design, and a control function approach to address sorting. Our paper identifies a component of overall place effects—namely the effect of neighborhood density—and while we find that this particular place effect is small, other components of place may be more important drivers of spatial heterogeneity.

The rest of this paper is organized as follows. Section 2 presents background information on economic development, urbanization, and residential carbon emissions in cities in Indonesia and the Philippines. Section 3 describes the different datasets we analyze. Section 4 describes our empirical strategy, and Section 5 presents our results. Section 6 presents the results from a decomposition exercise to understand what factors are responsible for driving the growth in residential carbon emissions. Section 7 concludes.

2 Background: Growth, Urbanization, and Residential Emissions

After World War II, the Philippines was one of the wealthiest countries in East and Southeast Asia. In 1960, its real GDP per capita of $1,124 (in constant 2015 USD) was similar to South Korea's (Lucas Jr, 1993; Boquet, 2017). At that time, Indonesia had a much lower standard of living, with a real per capita GDP of $598, just over half that of the Philippines. However, from 1960-2010, Indonesia's GDP per capita grew

[1]Older work includes Brownstone and Golob (2009) who study the joint relationship between residential density and fuel usage in cities in California.

by an average of 6.9 percent per year, while the Philippines grew by only 2.2 percent per year. Hence, Indonesia is now roughly 12.5 percent wealthier than the Philippines in GDP per capita terms.

Indonesia's economic growth and structural transformation was accompanied by rapid urbanization. Today, 151 million people, or roughly 56 percent of Indonesia's population, live in urban areas (Roberts et al., 2019). In 1975, commensurate with its relatively higher level of development, the Philippines was more urbanized than Indonesia, and 67 percent of Filipinos lived in urban areas. By 2020, 95.3 million people—or 85 percent of the population—lived in cities in the Phillipines (Santillan and Heipke, 2024).

As populations moved to urban areas in both countries, the share of built up surfaces also grew substantially. In 1975, only 0.28 percent of the Philippines' land area consisted of built up surfaces, but this share increased to 0.80 percent in 2020, more than doubling (Santillan and Heipke, 2024). In 1990, 0.46 percent of Indonesia's land area consisted of built-up surfaces, but by 2014, this figure had nearly doubled to 0.75 percent (Civelli et al., 2023). The faster rates of built up area expansion, relative to urban population growth, are suggestive of high rates of urban sprawl in both countries (Angel et al., 2005).

Although different rates of sprawl can be explained by variation in demographic, geographic, and economic characteristics (Burchfield et al., 2006), national policies in both countries also played an important role. The Philippines lacks a comprehensive national land use law, and this facilitates the conversion of agricultural land on the urban fringe into residential land. While the Philippines removed the majority of its gasoline subsidies in the late 1990s, Indonesia has maintained its fuel subsidies, favoring motorization. Policymakers in both countries have also seldom enacted land use regulations to curb sprawl, such as open space dedications, limits on new construction, or environmental regulations (Rukmana, 2015).

Sprawl and Income Sorting. In many U.S. cities, higher income households sort into suburban areas, while lower income households remain in the center. However, the relationship between income and distance to the city center is very different in other countries (e.g. Brueckner et al., 1999; Glaeser et al., 2008; Gaigné et al., 2022; Deffebach et al., 2025). Figure 2 plots the relationship between a community's distance to the central business district (CBD) and several community-level variables. Panels A-C focus on Indonesia, using 2010 Census data and *Susenas* data, while Panels D-F focus on the Philippines, using data from the 2020 Census. In Indonesia, our definition of a community is a *desa*, while in the Philippines, it is a *barangay*. Both communities constitute the smallest level of local government in both countries, and within cities appropriately represent neighborhoods.

This figure uses data from our sample of urban areas (described in more detail below). Estimated local polynomial regression lines are plotted in blue, with 95% confidence bands in gray. Panel A shows that in Indonesian cities, population density declines rapidly as distance to the CBD increases.[2] Panel B shows that households with higher total expenditures tend to locate closer to city centers, although this relationship is not very strong. Another indicator of affluence is secondary school completion, and Panel C shows that more highly educated households also cluster in city centers.

In the Philippines, the sorting patterns are somewhat different. Panel D shows that density peaks at around 10 km away from the CBD, perhaps suggestive of higher levels of commercial real estate in city centers. High school completion rates are higher in the Philippines than Indonesia (72.6 percent vs 57.5 percent in Indonesia), and variation across distance to the CBD is somewhat muted (Panel E). However, Panel F shows that head of household college completion rates decline substantially with distance to the

[2] Panel A reproduces Figure 4 from Civelli et al. (2023).

CBD, suggesting that wealthier households tend to cluster in the center of Filipino cities.

Based on Figure 2, we might expect households living in denser areas in both Indonesian and Filipino cities to produce more residential emissions. Our empirical approach, described in more detail below, aims to identify the place effects of neighborhood density on residential carbon emissions, disentangling them from the effects of income and taste-based sorting.

Emissions and Electrification. Indonesia is now the world's twelth-largest consumer of energy (IEA, 2022b). In 2022, Indonesia was the largest source of carbon emissions in the ASEAN region at 651.7 metric tons of CO_2 equivalents (IEA, 2022a).[3] Although the Philippines produces fewer total emissions, as the country has grew wealthier, CO_2 emissions per capita increased by 40 percent from 2000-2022. Together, residential energy and transportation constitute large shares of total energy use in both countries (50.8 percent in Indonesia and 63.3 percent in the Philippines).

As the economies of Indonesia and the Philippines expanded, so too did the reach of their electricity networks, with nearly complete coverage by 2023, according to World Bank Data. However, this electrification comes with environmental costs, as both countries also have some of the most emissions-intensive electricity sectors in the world. According to CTR (2020), in 2020, Indonesia's grid produced 804 grams of CO_2 per kWh, while the Philippines produced 691 grams of CO_2/kWh. This compares to 556 g CO_2/kWh in China and 383 g CO_2/kWh in the United States. Only 12 percent of Indonesia's electricity came from renewable sources, while 23 percent came from renewables in the Philippines.

Residential Energy Use. Indonesia's tremendous economic growth was accompanied by rising rates of vehicle ownership and growing gasoline expenditures. By 2015, over 121 million vehicles were registered in Indonesia (roughly 1 vehicle for every 2 people).[4] Motorization is somewhat lower in the Philippines; according to the Dept. of Transportation, only 14.6 million vehicles were registered there in 2024 (1 vehicle for every 8 people). Note that U.S. households in 2020 owned roughly 1 vehicle for every 1.2 people, so as development continues, rates of vehicle ownership could grow further.

Cooking fuel is another important emissions source. Although kerosene has historically been important for Indonesian households, in 2007, the nation began a large-scale conversion program designed to help households transition to liquid petroleum gas (LPG) (Dwi Cahyani et al., 2020). From 2000-2016, the number of Indonesians using LPG increased by 4 fold and reached 100 million urban residents (IEA, 2017).[5] By 2021, almost 85 percent of households used LPG as their primary fuel (BPS, 2022).

In the Philippines, the fuel mix is somewhat different. According to the 2018 Family Income and Expenditure Survey (FIES), urban households' consumption of LPG constituted 20 percent of their total energy demand, electricity constituted 63 percent, and kerosene constituted 1 percent. Firewood and charcoal constituted 15 percent, but this share has declined substantially over time as LPG penetration has increased.

[3]The Association of Southeast Asian Nations (ASEAN) includes Brunei Darussalam, Burma, Cambodia, Indonesia, Laos, Malaysia, the Philippines, Singapore, Thailand, and Vietnam.

[4]Appendix Figure A.1 plots the number of registered vehicles in Indonesia over time, using data from various waves of Transportation Statistics (*Statistik Transportasi*) published by the Central Bureau of Statistics (*Badan Pusat Statistik*, or BPS).

[5]Per-capita usage of LPG also increased from 4.7 kg in 2007 to 24.4 kg in 2015 (Thoday et al., 2018).

3 Data

Delineating Urban Areas. We use nighttime lights satellite imagery to demarcate the spatial extent of urban areas in Indonesia and the Philippines, following Jiang (2021). The U.S. Air Force's Defense Meteorological Satellite Program (DMSP) collected data from Operational Linescan System (OLS) sensors to measure the intensity of Earth-based light. Raw DMSP-OLS data were aggregated to an annual panel at a 30 arc second resolution from 1992-2013. In 2013, the DMSP-OLS data were replaced by a different set of sensors, known as the Visible and Infrared Imager/Radiometer Suite (VIIRS). VIIRS data are available at a higher resolution, with each pixel equal to about 0.22 km^2 at the equator.

We first deblur and merge these data, following Abrahams et al. (2018). Then, we delineate human settlements based on a luminosity threshold equal to zero, which means that pixels with positive luminosity are all considered human settlements. We then aggregate clusters of luminous pixels, allowing for small gaps between them (1 km for DMSP-OLS data and 0.5 km for VIIRS data), into a single polygon. This process generates 12,522 polygons for Indonesia and 3,940 polygons for the Philippines in 1992. Most of them are quite small and discrete, so we cross-reference them with place names from the Global Rural-Urban Mapping Project (GRUMP) that describe urban settlements with populations exceeding 100,000 in 2000. After matching, we retain 93 polygons for Indonesia and 42 for the Philippines that contain these urban settlements. These polygons constitute our set of urban areas.[6]

Figure 1 presents a map of the locations of these cities and their defined urban spatial extents. Panel A shows that in Indonesia, more than half of the 93 cities identified are located on Java and Bali, roughly a quarter are on the island of Sumatra, and the remainder are located in the remaining Outer Islands. The largest metropolitan area is Greater Jakarta (*Jabodetabekpunjur*), the economic and political center of Indonesia, which is a megacity with over 26 million people in 2016. Three other cities have more than 2 million inhabitants—these are Bandung, Surabaya, and Medan. Panel B shows the locations of the 42 cities in our sample from the Philippines. The Greater Manila Area is far and away the largest city, with a population of over 24 million in 2016. Three other cities have populations above 1 million (Cebu, Angeles, and Davao).[7]

Measures of Residential Carbon Intensity. We construct measures of residential carbon emissions from the consumption modules of various waves of Indonesia's National Socioeconomic Survey (*Survei Sosial Ekonomi Nasional*, or *Susenas*). The consumption data in the *Susenas* were matched on household identifiers and to individual-level data from the *Susenas* core survey.[8] Our primary analysis uses three waves of *Susenas* data (2010, 2011, and 2012), which we pool together into a single 2010 epoch.[9]

We derive similar measures for the Philippines using the Family Income and Expenditure Survey (FIES). The Philippine Statistics Authority (PSA) fields this survey every 3 years to provide information on citizens' socio-economic status and expenditure patterns. For Filipino cities, we predominantly work

[6] A single polygon may correspond to multiple GRUMP urban settlements if those settlements were close to each other and already appeared connected in the 1992 NTL data. The polygon is named after the largest settlement.

[7] Appendix Table A.1 summarizes the cities in our sample from Indonesia, while Appendix Table A.2 summarizes the cities in our sample from the Philippines.

[8] Note that in household-level regressions, we sometimes use individual-level data for the head of household to construct covariates like age, sex, education level and employment status.

[9] Surbakti and Statistik (1995) provides an overview for the design of the *Susenas* survey. Historically, consumption modules were only included every 3 years, but beginning in 2005, consumption modules were included every year.

with data from the 2018 FIES. A sample of 170,917 households was interviewed in the FIES 2018 round for this survey, a substantially larger sample size than previous surveys (PSA, 2018).

The key variables used to measure residential carbon emissions, which are available in both countries across all waves of data, include consumption quantities and total expenditures on the following items: (1) electricity; (2) kerosene; (3) liquefied petroleum gas (LPG); and (4) vehicle gas. Our primary specifications use total quantities of consumption of different energy sources as dependent variables. While total expenditures are very often reported, quantities are sometimes missing in the Indonesia data (less than 3 percent of observations).[10] To account for this, we impute missing quantities by first estimating unit values (i.e. expenditures divided by quantities) for each item separately at the city-survey year level. We then impute missing quantities by dividing reported expenditures by the average unit values we recover. In our results below, we explore the sensitivity of our findings to the use of imputed quantity data.[11]

Table 2 presents summary statistics for monthly energy consumption variables across households in our urban sample, in the 2010 epoch for Indonesia and in 2018 for the Philippines. Filipino households in 2018 consumed nearly twice as much electricity as Indonesian households in 2010, but they also consumed relatively lower amounts of LPG, vehicle gas, and kerosene. The final row shows estimates of residential carbon emissions that are implied by these expenditure quantities. To obtain these estimates, we multiply each category's monthly fuel consumption by carbon emissions factors (reported in Appendix Table A.3). Different factors are used for households that source their electricity from different types of power plants.[12]

Both surveys asked questions about monthly energy consumption. Ignoring seasonal differences in energy consumption, we find that urban households produced an average of 0.924 tons of CO_2 per year in Indonesia and 0.972 tons of CO_2 in the Philippines. According to Goldstein et al. (2020), the average household in the U.S. produces 2.83 tons of CO_2 per year from residential emissions sources. Based on this, in 2010, Indonesian households produced about 33 percent of residential emissions of U.S. households, while Philippines households in 2018 produced about 34 percent of the emissions of U.S. households.

Community-Level Demographic Characteristics. We construct community-level demographic characteristics using Census data. For Indonesia, we merge *desa* / *kelurahan*-level aggregates from the 2010 Census to our pooled *Susenas* sample. For the Philippines, we work with *barangay*-level aggregates from the 2020 Census. Both community definitions approximate neighborhoods in cities.[13] These census data

[10]For all items, we interpret any observations with zero expenditures and missing quantities as zeroes. For kerosene, LPG, and vehicle gas consumption, we interpret observations that are missing both quantities and expenditures as zeros. We also drop households that have zero consumption across all categories of energy expenditures and quantities.

[11]For the 2010 epoch in Indonesia, we imputed: (1) 2.4 percent of missing electricity quantities; (2) 1.1 percent of missing LPG quantities; (3) 2.7 percent of missing vehicle gas quantities; and (4) essentially 0 percent (2/119485) of missing kerosene quantities.

[12]We use government data from both countries on the total capacity of different types of power plants and their locations to calculate electricity emissions factors. In Indonesia, such data is available at the province level; for the 2010 vintage, we use data from BPS (2012), and for the 2000 vintage, we use data from 2003 in BPS (2008). Table A.3 provides emissions factors for different electricity generation sources, and we use provincial capacity weights to calculate a weighted average emissions factor for electricity emissions in each city. In the Philippines, for both waves of data, we use data from the 2015-2017 National Grid Emission Factor issued by the Philippine Department of Energy. This data is available only for the two major Philippine power grids: the Luzon-Visayas Grid and the Mindanao Grid.

[13]The *desa* / *kelurahan* is the 4th level administrative unit in Indonesia, below the province, district, and subdistrict. The

allow us to construct multiple measures, including population density at the community level, the share of community members with different levels of educational attainment, the share belonging to different ethnic groups, and the share that is married or migrated from another district. As we describe below, community averages of individual-level characteristics, which we calculate with these data, are crucial for our empirical strategy.

Geospatial Data on Administrative Boundaries and Topography. Our analysis relies on administrative boundary shapefiles that identify community borders. These datasets are created by Indonesia's national statistical agency, *Badan Pusat Statistik* (BPS), and the Philippine Statistics Authority (PSA). We use these boundaries in combination with data from the Harmonized World Soil Database (HWSD) to construct basic topographic characteristics (e.g., ruggedness, slope, and elevation).

Soil Characteristics and Depth to Bedrock. We use data from SoilGrids to measure the characteristics of soils in urban communities of Indonesia and the Philippines. Hengl et al. (2017) train machine learning algorithms to predict soil attributes on hand-collected soil profiles from roughly 150,000 sites. This produces a global dataset—at a 250-meter resolution—of several soil properties, including: (1) bulk density; (2) water content; (3) sand content; (4) clay content; (5) texture classification; and (6) soil taxonomy information. Absolute depth to bedrock (in cm), which is also predicted using the global compilation of soil ground observations, is also available from SoilGrids.[14]

Although SoilGrids also provides measures of organic carbon content and soil pH, we did not use these measures as instruments because they can be directly manipulated by human activity. We also only use soil characteristics measured at a depth of 60 cm or more, as these reflect variation in the subsoils and parent material of soils which were determined millions of years ago.

4 Empirical Strategy

The goal of this paper is to estimate how population density affects residential carbon emissions within cities. Our thought experiment is to imagine that an average household—with fixed demographic and employment characteristics—moves from a low density area in the suburbs to a higher density area in the center of the same city. How would that household's residential emissions change?

This section explains our approach for addressing the two key identification challenges that confound estimates of the relationship between a community's population density and household residential carbon emissions: (1) sorting of households with different income levels or different tastes for carbon intensity; and (2) place-specific unobservables that may simultaneously determine density and residential energy use. Altonji and Mansfield (2018) show that in a certain class of sorting models, adding controls for sorting on observables can help to bound the variance of overall community-level treatment effects when households sort into communities. Following Civelli et al. (2023), we combine these sorting controls with instruments to point identify the effect of density on residential carbon emissions. We first sketch the sorting model, and we provide key features of the sorting controls and instrumental variables procedure below.[15]

barangay—historically referred to as the *barrio*—is the smallest administrative division in the Philippines.

[14]SoilGrids data are publicly available and can be processed using Google Earth Engine and ISRIC Data Hub.

[15]Many details on this empirical strategy—used for a different research question—can also be found in Civelli et al. (2023).

4.1 Sorting into Communities

Let i index households and let $v \in \{1, ..., V\} \equiv \mathcal{S}$ index communities that comprise different cities. The set $\mathcal{S}$ includes both suburban and central neighborhoods that constitute all cities throughout a country.[16] Each community contains a $(K \times 1)$ vector of amenities. This vector includes both exogenous amenities, such as natural features, but also endogenous amenities, like density, school quality, or congestion, which are determined through the sorting equilibrium. Let P_v denote housing prices in community v, which are also endogenously determined.

We assume that household i's indirect utility from living in community v is given by:

$$U_i(v) = \mathbf{W}_i \mathbf{A}_v - P_v + \varepsilon_{iv} \, , \tag{1}$$

where $\mathbf{W}_i$ is a $(1 \times K)$ vector measuring household i's willingness to pay for different amenities, and ε_{iv} is an idiosyncratic preference term. Following Altonji and Mansfield (2018), we partition the willingness to pay coefficients, $\mathbf{W}_i$, into three additively separable terms:

$$\mathbf{W}_i = \mathbf{X}_i \mathbf{\Theta} + \mathbf{X}_i^U \mathbf{\Theta}^U + \mathbf{Q}_i \mathbf{\Theta}^Q \, . \tag{2}$$

This partition includes the following components: (1) $\mathbf{X}_i$, a vector of household-level observables affecting tastes for amenities and also impact residential carbon emissions; (2) $\mathbf{X}_i^U$, a vector of household-level unobservables that shape tastes for amenities and emissions; and (3) $\mathbf{Q}_i$, a vector of variables (both observed and unobserved) that may influence preferences over amenities—and therefore impact sorting—but have no impact on residential emissions. In equation (2), the vectors $\mathbf{\Theta}$, $\mathbf{\Theta}^U$, and $\mathbf{\Theta}^Q$ measure each component's respective willingness to pay coefficients.

Note that the partition in equation 2 defines $\mathbf{X}_i$, $\mathbf{X}_i^U$, and $\mathbf{Q}_i$ so that they represent the complete set of household-level factors that determine sorting and residential emissions. This formulation is fairly general and allows for arbitrary correlations between household characteristics—both observable and unobservable—and tastes for amenities. The major caveat is that indirect utility function, equation (1), is additively separable.

When making location decisions, we assume that agents form expectations about the levels of P_v and $\mathbf{A}_v$ that will prevail in each community before moving. When agents move, they act atomistically, ignoring their impact on prices and endogenous amenities that emerge in equilibrium. Using their full information set, including expectations about prices of communities in different cities, expectations of endogenous amenities, their full set of preference weights $\mathbf{W}_i$, and realizations of ε_{iv} for all v, households choose a location to maximize expected indirect utility. Let $v(i)$ denote household i's optimal community choice.

Given this framework, Altonji and Mansfield (2018) prove that under a relatively weak set of additional assumptions, the community-level expectation of household-level unobservables that influence residential carbon emissions, denoted by $\mathbf{X}_v^U \equiv \mathbb{E}[\,\mathbf{X}_i^U \,|\, v(i) = v\,]$, is linearly dependent on community-level average observables, $\mathbf{X}_v \equiv \mathbb{E}[\,\mathbf{X}_i \,|\, v(i) = v\,]$. The sorting process creates two vector-valued mappings: (1) a mapping between community-level average observables and amenities, denoted by $\mathbf{X}_v =$

[16] We estimate the effects of community-level population density on residential carbon emissions separately in both countries, so we do not allow households to migrate across countries in the model.

$\mathbf{f}(\mathbf{A}_v)$; and (2) a mapping between community average unobservables and amenities, denoted by $\mathbf{X}_v^U = \mathbf{f}^U(\mathbf{A}_v)$. The authors provide conditions under which the first mapping, $\mathbf{f}$, is invertible, so we can write: $\mathbf{X}_v^U = \mathbf{f}^U\left(\mathbf{f}^{-1}(\mathbf{X}_v)\right)$. Under an additional assumption, the authors show that $\mathbf{f}^U\left(\mathbf{f}^{-1}(\cdot)\right)$ is actually linear.[17]

Fortunately, we can directly test for the invertibility of $\mathbf{f}$. In particular, a necessary condition for invertibility is that the dimension of $\mathbf{A}^{\mathbf{X}}$, the subset of amenities that affect sorting on observables, is less than the number of elements in $\mathbf{X}_v$. This would occur if $\mathbb{V}(\mathbf{X}_v)$ is rank deficient.

To test this condition, we use a vector of 38 variables constructed from Indonesian census data to measure $\mathbf{X}_v$, and we use a different vector of 38 variables constructed from Philippines census data to measure $\mathbf{X}_v$. These variables include the community's average age, years of schooling, household size, the percentage of the community that is female, the percent who self-identify with different religions or ethnicities, the share of different types of employment status and marital status.[18]

Appendix Table A.4 uses a test proposed by Kleibergen and Paap (2006) to formally test the rank of $\mathbb{V}(\mathbf{X}_v)$. In the Indonesian urban sample (column 2), we cannot reject the null hypothesis that the rank of the variance-covariance matrix of $\mathbf{X}_v$ is 14 against the alternative that it is 15 or greater (p-value = 0.237). In urban areas of the Philippines (column 4), we cannot reject the null hypothesis that the rank of the variance-covariance matrix of $\mathbf{X}_v$ is 17 against the alternative that it is 18 or higher (p-value = 0.228).[19] The results from Appendix Tables A.4 suggest that $\mathbf{X}_v$ is rank deficient in both countries. This implies that $\mathbf{f}$ will be invertible, so that $\mathbf{X}_v$ can be used as a linear control function for sorting on unobservables.

4.2 Production of Residential Carbon Emissions

After choosing locations, we assume that an energy consumption outcome for household i living in community v, denoted by y_{vi}, is produced according to the following log-linear, additively separable function:

$$\log y_{vi} = \mathbf{X}_i\beta + x_i^U + \theta \log \text{density}_v + \mathbf{C}_v\mathbf{\Gamma} + c_v^U + \eta_{vi} + \xi_{vi}\,. \tag{3}$$

In our application, $\log y_{vi}$ denotes the log of the total quantity of electricity used per month, or logs of the total quantity of different types of fuel used per month (e.g. vehicle gasoline, kerosene, or LPG). This equation is composed of three sets of terms: (1) a household-level component; (2) a community-level component; and (3) an idiosyncratic component.

The household component, $\mathbf{X}_i\beta + x_i^U$, measures the contribution of observed and unobserved household characteristics to energy consumption. The community-level component, $\theta \log \text{density}_v + \mathbf{C}_v\mathbf{\Gamma} + c_v^U$, contains three terms. The first measures log population density at the community level. Here, density is defined as the population of community v in the survey year divided by the area of that community (in km^2). The key object of interest, θ, measures the elasticity of energy consumption with respect to density. The second component, $\mathbf{C}_v$, measures how other community-level characteristics affect energy

[17] See Altonji and Mansfield (2018) and Civelli et al. (2023) for more details on the specific assumptions required.

[18] We discuss the exact variables used to construct $\mathbf{X}_v$ for different countries below in Section 5.

[19] Appendix Table A.5 also reports a principal components analysis of $\mathbf{X}_v$ in Indonesia and the Philippines. In our urban *Susenas* sample for Indonesia (column 2), 27 factors explain 95 percent of the total variation in $\mathbf{X}_v$ and 31 factors explain 99 percent of the total variation in $\mathbf{X}_v$. In the urban FIES sample from the Philippines (column 4), 28 factors explain 95 percent of the total variation in $\mathbf{X}_v$, while 33 factors explain 99 percent of the total variation in $\mathbf{X}_v$.

consumption. We include urban-area fixed effects in $\mathbf{C}_v$, so that θ is pinned down by within city variation. Importantly, urban-area fixed effects also allow energy prices to vary by city. Finally, the third term, c_v^U, is a scalar summarizing how unobserved neighborhood characteristics contribute to y_{vi}.

The idiosyncratic component, $\eta_{vi} + \xi_{vi}$, also contains two terms. The first, η_{vi}, captures unobserved variation in community-level energy demand.[20] The second term, ξ_{vi}, captures other factors influencing energy demand that are determined after the household arrives in community v. We assume that these are unpredictable given the other observed and unobserved community characteristics. For example, unanticipated local labor market shocks could affect households' demands for residential energy, or shocks to infrastructure may impact the prevailing levels of energy consumption in certain areas.

We can expand the group-level observables–excluding log density–in equation (3) by writing $\mathbf{C}_v = [\mathbf{X}_v, \mathbf{C}_{2v}]$. Here, $\mathbf{X}_v$ is a row vector containing our sorting controls, namely community v's averages of household-level observables. The term $\mathbf{C}_{2v}$ includes city-specific intercepts, as well as pre-determined community-level characteristics, such as elevation, ruggedness, and distance to the coast or rivers, which may affect the provision of infrastructure or impact the prices of different sources of sources. This notation allows us to we rewrite equation (3) as follows:

$$\log y_{vi} = \mathbf{X}_i \beta + x_i^U + \theta \log \text{density}_v + \mathbf{X}_v \mathbf{\Gamma}_1 + \mathbf{C}_{2v} \mathbf{\Gamma}_2 + c_v^U + \eta_{vi} + \xi_{vi} \,. \tag{4}$$

where $\mathbf{\Gamma} = [\mathbf{\Gamma}_1, \mathbf{\Gamma}_2]$ collects the coefficients of the community-level variables on energy demand. In Section 4.1 above, we explain why adding $\mathbf{X}_v$ controls for sorting on both observables and unobservables in the production of energy demand. Typical control function procedures would use a non-linear or semi-parametric functional form, but the spanning assumption (assumption A5 in Altonji and Mansfield, 2018) implies that these controls need only enter linearly.

4.3 Identifying θ with Soil Characteristics Instruments

Altonji and Mansfield (2018) study how sorting controls, $\mathbf{X}_v$, can help identify group-level treatment effects in situations where individuals or households sort into groups—here neighborhoods within cities. Although these controls for sorting successfully eliminate sorting bias, they also absorb peer effects that could depend on neighborhood-averages. As a result, including them in a linear model to estimate neighborhood effects may result in overcontrolling, so identification is only partial. Civelli et al. (2023) extend their approach by using instruments to point-identify a component of the overall group-level treatment effect in a way that is unconfounded by sorting.

We follow Civelli et al. (2023) and use a vector of deep soil characteristics as instruments for population density within urban aras. We use data from SoilGrids to capture various attributes of different soils in urban areas of Indonesia and the Philippines. These include: (1) bulk density; (2) water content; (3) sand content; (4) clay content; (5) texture; and (6) soil taxonomy information. For robustness, we also include depth to bedrock as a plausibly exogenous shifter of neighborhood density.

Historically, fertile and stable soils lead to denser settlements in certain areas. Different soil types also impacted colonial investments, as shown by Dell and Olken (2019) in Indonesia. We allow soil types to affect neighborhood density differently in Indonesia and the Philippines, and we show that within cities

[20] Note that both observed and unobserved community-level variables may be correlated with η_{vi}.

today, they have a strong first stage relationship with population density. Soil characteristics are similar to geologic instruments for density like depth to bedrock that have also seen use in prior work (e.g Hoxby, 2000; Black et al., 2002; Rosenthal and Strange, 2008; Combes et al., 2010; Ahlfeldt et al., 2024).

Although we have a strong first stage relationship, there are several concerns with our identification strategy. One issue is reverse causality: soil characteristics measured in cities today may be influenced by human activity. As we discuss in Section 3, we only use deep soil attributes—measured at a depth of 60 cm or more—which helps to ensure that they are unaffected by humans. We also drop certain measures, like organic carbon content or acidity (pH), that are easily modified by human behavior.

A larger concern is the exclusion restriction; deep soil characteristics within cities need to only affect energy consumption through their impact on density. Although soil minerology and soils' parent materials were determined millions of years ago, fertile soils may still drive local wealth within cities, especially if agricultural employment is common in urban areas. We try to carefully examine this treat and several others in our analysis below.

Interpreting θ. We think that θ is a useful object of interest for evaluating how policies that shape the density of a community may impact its residential carbon emissions. For example, policymakers in the U.S. and other developed countries use many tools to directly regulate the level of density that is permitted in certain locations. Such policies include: (1) minimum-lot size zoning; (2) binding limits on new construction; (3) open space dedications; (4) growth controls; (5) environmental regulations; (6) septic system regulations; (7) subdivision requirements; and (8) historic preservation (Glaeser and Ward, 2009). While these policies limit the amount of density permitted in certain locations, they do not directly determine where certain people are allowed to live. Our estimates measure how density impacts energy demand conditional on the spatial distribution of the population, which is more relevant for evaluating the effects of such policies.

5 Results

First Stage. We begin by using post-double-selection lasso techniques, following Belloni et al. (2012), to predict log density of urban communities in Indonesia and the Philippines using soil characteristics. Table 3 reports parameter estimates from the following regression equation:

$$\text{log density}_{iv} = \alpha_c + \mathbf{z}_v'\beta + \mathbf{X}_i\mathbf{\Theta}_1 + \mathbf{C}_{2v}\mathbf{\Theta}_2 + \varepsilon_{iv} \,, \tag{5}$$

where i indexes households, v indexes communities and c indexes urban areas. The dependent variable, log density$_{iv}$ is the log of the density of the community in which household i resides. The term α_c denotes a city-specific intercept, $\mathbf{z}_v$ denotes the lasso selected vector of soil characteristics, $\mathbf{X}_i$ is a vector of household-level controls, $\mathbf{C}_{2v}$ denotes a vector of additional exogenous community-level characteristics (including ruggedness, elevation, and distance to the nearest coast and river), and ε_{iv} is an error term.

Out of 67 candidate soil characteristics instruments for density, all measured at a depth of 60 cm or more, the post-double-selection lasso procedure selected only 5 instruments for Indonesian cities and 3 instruments for Filipino cities. Table 3 reports their first-stage coefficients. The overall F-statistics for Indonesian cities (77.2 from column 2) and Filipino cities (96.0 from column 4) are large, and the first stages explain 58-67 perent of the variation in density within cities.

Columns 1 and 2 show that within cities in Indonesia and conditional on $\mathbf{C}_{2v}$ controls, population density in 2010 was positively related to the bulk density of soils' parent material. Columns 3 and 4 show similar results for cities in the Philippines, although with a slightly greater depth of bulk density selected (200 cm vs. 60 cm selected in Indonesia). These results are reasonable, given that more compact soils are more favorable for construction.

We also find that in Indonesian cities, sand content of soils' parent material is negatively related to density, while in Philippines cities, water content of soils' parent material is negatively related to density. Although sandy soils may be favourable for construction, they are also difficult for growing crops and likely reduced historical agricultural productivity. Watery soils can also be unstable, making it harder to build residential or commercial floorspace.[21]

In addition, four soil types were also significant predictors of population density. Haplustolls (from the order Mollisols) are grassland soils used for growing grains and feed crops. Chromusterts (Vertisols) are fertile soils that are often used for growing grains. Haplustalfs (Alfisols) are also naturally productive soils with high fertility (USDA, 2015). Each of these soil types are prominent in areas that were initially favorable for rice production and influenced historical settlement patterns, and as expected, these soil characteristics have positive effects on modern density within urban areas. On the other hand, Tropodults (Ultisols) are predominant in tropical forests and may have disuaded historical settlements (USDA, 1975), and the density coefficient for this soil type is negative. In summary, the selected soil characteristics within cities are either associated with attracting historical development through favourable agricultural production or with the ease of constructing buildings.[22]

In columns 2 and 4, we add controls for sorting on observables and unobservables, $\mathbf{X}_v$, to equation (5). The coefficients on the selected soil characteristics retain their signs and statistical significance. Overall, the results from Table 3 suggest that conditional on city fixed effects, the selected soil characteristics have strong first stage relationships with community-level population density, the key dependent variable in our analysis.[23]

Baseline Results. To estimate the impact of density on residential energy consumption expenditures, we run linear instrumental variable regressions of the following form:

$$\log\left(1+y_{vi}\right)=\alpha_c+\mathbf{X}_i\beta+\theta\log\text{density}_v+\mathbf{X}_v\mathbf{\Gamma}_1+\mathbf{C}_{2v}\mathbf{\Gamma}_2+\varepsilon_{vi}\,, \tag{6}$$

where α_c denotes a city fixed effect, $\mathbf{X}_i$ is a vector of household-level observables, $\mathbf{X}_v$ is a vector of community averages of the characteristics of people who live in community v (described above), $\mathbf{C}_{2v}$ are controls for community characteristics that are not mechanically related to sorting, and ε_{vi} is an error term. We estimate separate effects for urban households in both Indonesia and the Philippines. To illustrate our empirical strategy, Table 4 shows results for a single outcome variable, namely the log total

[21]Elevation, ruggedness, and distance to the coast and rivers are also correlated with density as one might expect. However, we include them in $\mathbf{C}_{2v}$ but not as instruments, because they could directly impact residential carbon emissions. For example, elevation may directly affect energy demand by reducing the need to use AC for cooling in higher altitudes.

[22]Note that Civelli et al. (2023) selected a slightly different set of soil characteristics to predict density within Indonesian cities. The differences between these two papers owe to different definitions of urban areas, and to the fact that this paper uses an expanded sample of urban Susenas communities because we pool data from 2010, 2011, and 2012.

[23]First-stage regression results at the community-level can be found in Appendix Table A.6.

quantity of electricity consumed by household i last month. Panel A shows results for households in Indonesian cities, while Panel B focuses on households in urban areas of the Philippines. For inference, we report robust standard errors, clustered at the sub-district level in Panel A and at the municipality level in Panel B, in parentheses.[24]

We begin by reporting estimates of θ from specifications that omit controls for sorting, setting $\mathbf{\Gamma}_1 = 0$ as a baseline. In column 1, our OLS specification finds that increasing density by 1 percent increases electricity consumption by 0.08 percent in Indonesian cities. Although highly significant, this is a moderate effect size, equivalent to roughly 0.1 percent of a standard deviation in electricity consumption. The OLS coefficient for cities in the Philippines is also positive (0.06, from Panel B).

Column 2 reports the relationship between density and electricity consumption estimated from IV-Lasso specifications, where we instrument density with soil characteristics. Overall, the estimate of θ grows larger, increasing to 0.099 in Indonesia and 0.181 in the Philippines. In column 2, the Kleibergen and Paap (2006) Wald Rank F-Stat, a generalization of the first-stage F-statistic for multiple instrumental variables, is large at over 92 in Indonesia and 27 in the Philippines. The Kleibergen-Paap rank LM test also strongly rejects the null of underidentification for the endogenous density variable in both countries.

Appendix Table A.7 reports coefficients on household and community-level controls for Indonesian households (columns 1 and 2 of Table 4), while Appendix Table A.8 does the same for Filipino households (Table 4, Panel B). We see that relative to no schooling (the omitted category in Indonesia) and relative to no secondary education (the omitted category in the Philippines), higher levels of educational attainment are associated with greater quantities of electricity use. As expected, household size also increases electricity use, and electricity consumption has an inverse U-shaped relationship with the age of the household head.

In Table 4, column 3, we report results of the full model, where we include household-level covariates, $\mathbf{X}_i$, community-specific characteristics, $\mathbf{C}_{2v}$, and averages of different individual-level variables at the community level, $\mathbf{X}_v$, to control for sorting. The addition of sorting controls reduces the elasticity of electricity consumption with respect to density to 0.027 in Indonesia and 0.068 in the Philippines. In both cases, the effect becomes statistically insigificant. Although the Kleibergen-Paap Wald Rank F-Stat falls in both countries, the Kleibergen-Paap LM tests still reject the underidentification. Moreover, the Sargan-Hansen J-test statistic for overidentifying restrictions is also small in both countries, and we cannot reject the null that the soil characteristics instruments are correctly excluded from the estimation equation at the 1 percent significance level. Overall, the results in Table 4 suggests that the IV models are well specified, even after introducing controls for sorting.

In Appendix Table A.9, we report the estimated coefficients on Indonesian sorting controls (Table 4, Panel A), and in Appendix Table A.10, we report estimated coefficients for sorting controls for Philippines cities. Some ethnic group shares are negatively correlated with electricity use, but these are typically weak correlations. However, average years of schooling in the community seems to be strongly predictive of increased electricity use, suggesting that educational sorting and income sorting may be responsible for the naive OLS findings. Overall, the results from Table 4 suggest that while greater density is associated with higher electricity consumption, this effect is not causal and is not robust to controls for

[24]Indonesian cities are collections of urban *desa* / *kelurahan*, which span multiple sub-districts (*kecamatan*), the third-level of administrative division. Philippines cities are collections of urban *barangay*, which span possibly multiple municipalities, also the third-level of administrative division.

sorting on observables and unobservables.

Full Results. Table 5 shows parameter estimates from equation (6) for all of the residential carbon intensity outcomes we study. Note that this approach allows for flexibility in energy demand across different sources, instead of estimating a single carbon emissions demand equation, as in Lyubich (2025). Columns 1-3 replicate the results on electricity from Table 4. Columns 4-6 report estimates of the relationship between density and LPG consumption. We obtain similar results to electricity: in both countries, density is positively associated with LPG consumption, and the effects increase after instrumenting for density. However, once we control for sorting, the relationships with density attenuate. Column 6 suggests that a 1 percent increase in density reduces LPG consumption by 0.05 percent in Indonesia and increases LPG consumption by 0.01 percent in the Philippines, but neither estimate is statistically significant.

Columns 7-9 report estimates of the relationship between density and vehicle gas consumption. Despite what one might expect from a monocentric city model, density and vehicle gas consumption are positively correlated in Indonesian cities in the OLS and IV specifications (column 6-7, Panel A). We also find a positive, though insignificant, correlation between density and vehicle gas consumption for urban households in the Philippines in the IV specification. However, after we control for sorting, in both countries, we cannot reject that the coefficient on density is statistically indistinguishable from zero. These results are quite different from findings in the U.S., where lower density neighborhoods are associated with increased miles traveled and fuel consumption (e.g. Brownstone and Golob, 2009).

Columns 10-12 show the results for kerosene, an important source of fuel for Indonesian households but less so for the Philippines. Here, we do find a significant relationship between kerosene consumption and density in Indonesia, but the effect is positive. The estimate in column 12 suggests that a 1 percent increase in density is associated with a 0.1 percent increase in kerosene consumption. However, the coefficient is quite small in the Philippines and is statistically insignificant.

Density and Asset Ownership. Given that we find an insignificant relationship between many aspects of residential carbon emissions and population density, we would also expect this finding to hold for measures of asset ownership. The 2010-2012 waves of the *Susenas* ask Indonesian households about ownership of several assets, including: (1) refrigerators; (2) motorcycles; (3) cars; (4) air conditioning; (5) gas cylinders; and (6) water heaters. Similar, but not all, asset ownership outcomes are available in the Philippines' FIES. These variables are summarized in Appendix Table A.11.

Table 6 presents IV estimates of the relationship between community-level population density and household asset ownership, using linear probability models based on the same specification as equation (6). These estimates suggest that refrigerator and AC ownership are positively associated with density in both Indonesia and the Phillipines, but those effects are not robust to controls for sorting.

Although one might expect that density would be negatively related to car and motorcycle ownership, we do not find statistically significant relationships between density and vehicle ownership outcomes in either Indonesia or the Philippines. If anything, car ownership is positively associated with density in Philippines cities, but neither vehicle outcome is significantly related to density after controlling for sorting. We also do not find significant relationships between density and ownership of LPG cylinders or hot water heaters, but these variables are only measured in Indonesia.

In summary, just as in the case of residential carbon emissions measures, many of the relationships between asset ownership and density reflect income sorting. After accouting for income sorting, we find no role for neighborhood population density in shaping household-level asset ownership.

Intensive vs. Extensive Margin Effects. Our main expenditure results rely on outcome variables that combine extensive and intensive margin effects. Given that we do not find significant relationships between density and asset ownership after controlling for sorting, we should expect to see similar null results on the extensive margin of expenditure. Appendix Table A.12 uses a linear probability model specification and reports estimates of the effects of density on indicators for whether or not the household reported any expenditures on electricity, vehicle gas, natural gas, or kerosene. The density effects in IV results with controls for sorting are not statistically significant at conventional levels, and these null results hold for all outcomes and for both countries.

Appendix Table A.13 drops the $\log(1 + y_{iv})$ transformation in our main specification and examines the effect of density on $\log y_{iv}$. Movement in this outcome reflects intensive margin changes in residential energy expenditures for households who report positive expenditure values. For Indonesian cities, we report null results on electricity, vehicle gas consumption, and kerosene in IV specifications with sorting controls. We do find that density reduces LPG consumption, though the coefficient only marginally significant. For the Philippines, we find insigificant results for electricity, kerosene, and vehicle gas consumption. However, we also find positive, but small and significant effects of density on LPG consumption.

Probing the Effect of Sorting. In Figure 3, we explore the role that different types of sorting controls have on our IV estimates of the relationship between density and different sources of residential carbon emissions. Each point reports the coefficient on log population density from equation (6) where the dependent variable is listed in the panel header. The bars depict 95 percent confidence intervals.

The first set of estimates ("None") replicate our estimates from Table 5 without controls for sorting, which tend to be positive and significant. In the second row of estimates ("Non-Econ"), we add non-economic sorting controls, including ethnic composition shares and shares of different members of the community in different religious groups. The positive estimates of the effects of density on sources of residential carbon emissions do not change much in terms of magnitude or significance.

However, in the third set of rows ("Econ"), we replace the non-economic sorting controls with variables that capture sorting on economic factors, including education shares, household size, and employment characteristics. These controls dramatically reduce many of our coefficients, often making them insignificant. This suggests that economic sorting is playing a substantial role in mediating the observed relationship between density and residential carbon emissions. The fourth set of rows ("All") reproduce our preferred estimates from Table 5 with the full set of $\mathbf{X}_v$ controls.

The Density Elasticity of Residential Emissions. To more precisely estimate how density impacts overall residential carbon emissions, we aggregate the separate effects of density on different categories of residential energy-related expenses. Let CE_{iv} denote the total carbon emissions of household i in community v:

$$\text{CE}_{iv} = \sum_k \omega_v(k) y_{iv}(k)\,, \tag{7}$$

where $y_{iv}(k)$ is the total annual quantity of energy consumed by household i from source k, and $w_v(k)$

is a carbon intensity weight for source k. We allow the carbon intensity weights to differ by community because electricity generation sources differ across grids, and different types of power plants can vary in their carbon intensity.[25]

Estimating equation (6) for different outcome variables, indexed by $k = 1, ..., K$, provides different estimates of $\theta_k = \partial \log(1 + y_{iv}) / \log \text{density}_v$, which is the elasticity of quantities of energy category k with respect to density. It is easy to show that we can write the elasticity of total household carbon emissions with respect to density as:

$$\mathcal{E}_{\text{CE}_{iv},\text{density}} = \sum_{k=1}^{K} \left(\frac{w_v(k) y_{iv}(k)}{\text{CE}_{iv}} \right) \theta_k \quad . \tag{8}$$

This is just a sum across energy sources of different density elasticities, where each elasticity is weighted by that source's share in total household carbon emissions.

Table 7 provides estimates of $\mathcal{E}_{\text{CE}_{iv},\text{density}}$, using national average carbon intensity weights, separately for urban households in Indonesia and the Philippines. We estimate all equations simultaneously using a seemingly unrelated regressions system to calculate standard errors. Panel A shows that a 1 percent increase in density is associated with a 0.007 percent increase in carbon emissions in Indonesia and a nearly 0 percent increase in carbon emissions in Philippines households. Column 2 shows that these effects increase to roughly 0.01 in both countries after instrumenting for density. However, in column 3, we add controls for sorting, and the elasticity falls to -0.005 in Indonesia and 0.004 in the Philippines. In both cases, the density elasticity of residential emissions becomes statistically insigificant.

Figure 4 plots estimates of $\mathcal{E}_{\text{CE}_{iv},\text{density}}$ for different cities, where we replace the national carbon intensity weights with city-specific weights to calculate equation (8). Because the carbon intensity weights for LPG, kerosene, and vehicle gas are fixed internationally (see Appendix Table A.3), our city-specific weights allow electricity consumption to contribute differently to emissions depending on where the city is located in the electricity grid and the carbon intensity of electricity production nearby. Despite allowing for more heterogeneity, nearly every city's density elasticity of residential carbon emissions is statistically indistinguishable from zero after we include sorting controls.

5.1 Probing Internal Validity and Robustness

Agriculture Households. A major concern with the results presented so far is the exclusion restriction: for soil characteristics to be a valid instrument, they must only affect residential carbon emissions through their impact on population density. One potential violation of the exclusion restriction is that if favorable soils were important for agricultural productivity in cities today, they could affect residential carbon emissions by affecting income or wealth. This is clearly an issue, since approximately 19 percent of households in Indonesian cities and 3 percent of urban households in the Philippines report employment in agriculture.

In Appendix Table A.14, we examine the sensitivity of our estimates to dropping agricultural households and to dropping communities with large shares of employment in agriculture. Each cell in this

[25] See our discussion of carbon intensity weights in Section 3 for more details. Appendix Table A.3 shows the various emissions factors used to calculate $\omega_v(k)$.

table reports results from a different regression, focusing on Indonesian cities.[26] Different panels focus on different emissions outcomes, and different rows denote estimates with and without sorting controls. Column 1 reports our baseline IV-Lasso estimates (from columns 2 and 3, Table 5). In column 2, we drop all households that report any employment in agriculture, and our results are mostly unchanged. In columns 3-6, we drop different communities based on the share of households that are employed in agriculture. While the magnitudes of density's effects vary slightly, qualitatively our results remain unchanged. We continue to find that unconditionally, greater energy use takes place in denser communities, but these results do not survive the inclusion of sorting controls.

Controlling for Historical Infrastructure. Although our results are robust to dropping agricultural households, the soil characteristics IVs may still be correlated with omitted community characteristics that affected historical sorting patterns and explain today's carbon emissions outcomes. We already control for certain aspects of geography (e.g. ruggedness, elevation, distance to the coast and rivers), but other omitted community characteristics may have affected the development of historical infrastructure and be correlated with modern density. This infrastructure may have induced people of different types to sort into different neighborhoods, determining persistent income sorting that lingers today.

In Appendix Table A.16, we show that our main estimates of the elasticity of residential carbon emissions with respect to density in Indonesia (reported in Table 7, Panel A), are robust to controlling for different types of historical infrastructure. To measure these controls, we use *Podes* data from 1983 to measure the number of education facilities, health facilities, religious facilities, irrigation facilities, electricity coverage, as well as the number of agricultural and social organizations. We also include controls for distance to major historical roads.[27] In all cases, our main estimate of the density elasticity of residential carbon emissions is nearly unchanged. While we lack data from the 1980s for the Philippines, Appendix Table A.17 shows how the 2018 density elasticity results (reported in Table 7, Panel B) are robust to controlling for measures of infrastructure from the 2000 census, including education, health, and religious facilities as well as distance to major roads.

Soil Characteristics in Rural Areas. If the exclusion restriction is satisfied, then in a subsample with no first stage relationship, there should also be no reduced-form relationship (Altonji et al., 2005; van Kippersluis and Rietveld, 2018). To provide more evidence in favor of the exclusion restriction, we performed a placebo exercise, estimating the effect of soil characteristics on residential energy consumption in rural communities where those IVs do not predict population density.

To implement this exercise, we identified rural communities in Indonesia and the Philippines with a population density of less than 300 inhabitants per square km.[28] Appendix Table A.18 shows that in these communities, soil characteristics IVs are not predictive of density. In the last row of each panel, we show that the first stage relationship between the selected soil characteristics and density is quite weak, with Kleibergen-Paap Wald F-stats of 2.4 in Indonesia and 0.06-0.1 in the Philippines. The table also reports p-values for tests of the joint significance of the soil characteristics in predicting each dependent variable. Of the four dependent variables we study, the soil characteristics IVs are only significant in predicting LPG consumption in the Philippines (p-value $= 0.01$), while the other p-values are above

[26]A similar table for households in cities in the Philippines can be found in Appendix Table A.15.
[27]For more details on these variables, see the Online Appendix to Civelli et al. (2023).
[28]This definition of rural areas is consistent with the definition used by the UN Statistical Commission (UN, 2020).

conventional significance levels. We take this evidence as weighing in favor of the exclusion restriction.

Depth-to-Bedrock Instruments. So far, our IVs have focused solely on soil characteristics. However, depth to bedrock has also been used as an IV for density in prior work, and it may be even more likely to be excludable, given that bedrock depth may not be as impactful for agricultural productivity as other soil characteristics. In Appendix Table A.19, we add depth to bedrock to the instrument choice set, together with other soil characteristics, before implementing the post-double-selection IV lasso estimator. The first-stage F-statistics are similar to our preferred estimates reported in Table 5. Moreover, the estimates of the effect of density on residential energy use are also qualitatively similar, showing that our estimates are robust to expanding the instrument set.

Unit Values and Exact Quantity Measurements. As discussed in Section 3, for the Indonesian data, we sometimes impute missing quantity measures based on estimates of unit values. To explore the sensitivity of our results to this choice, Appendix Table A.20 re-estimates our main specification but drops observations with predicted quantities. Although point estimates differ slightly, the results are qualitatively similar to our baseline estimates, suggesting that our imputation procedure is not responsible for the results we obtain.

5.2 Mechanisms and Heterogeneity

Floor Area and Commuting Distance. One reason why many policymakers are concerned about urban sprawl is the tendency for homes to get larger as they move away from city centers, because larger homes require more energy consumption. However, the relationship between distance to the city center and home size may be different in cities with different income sorting patterns. In Table 8, we examine how population density affects a household's floor area, using self-reported data from the *Susenas* 2010-2012. In Indonesia, we find that conditional on household size, increasing density reduces floor area as expected, with a 1 percent increase in density reducing floor area by 0.76 percent in the IV specification (Panel A, column 2). When we add the sorting controls, the density coefficient loses some significance, but the density coefficient remains similar and negative. In the Philippines, the OLS specification also suggests that increasing density reduces floor area, and the IV estimates, while no longer significant, are still negative and similarly sized. These results suggest that households in denser urban neighborhoods do tend to use smaller amounts of residential floorspace, but despite this, there are no differences in residential energy consumption.

Another surprising finding from Table 5 is that density is unrelated to vehicle gas consumption after controlling for sorting and simultaneity. If anything, the coeffients in Table 8, columns 4-6 also suggest that commuting distances *increase* in denser neighborhoods, at least for households in Indonesian cities. These patterns reinforce the idea that monocentric city models may lack explanatory power for cities in lower-middle income countries.[29]

Heterogeneity by City Size. As countries grow richer, one concern is that they may converge towards more U.S. style income sorting, with richer households who own cars and move into bigger homes in the suburbs (Glaeser, 2011). On the other hand, if neighborhood amenities are persistent, richer households

[29]Unfortunately, we lack similar commuting distance variables in the Philippines FIES data.

may bid up housing prices in the center and take advantage of those amenities. Some of these central amenities may even become more desireable through endogenous sorting of higher income households (Brueckner et al., 1999).

In Appendix Table A.21, we examine the extent to which the effect of density on residential carbon emissions differs in large, affluent cities with more than 1 million residents (e.g. Jakarta and Manila) vs. other smaller cities. We generally find that the effect of density is not significantly different in larger cities relative to other smaller cities, suggesting that our results are invariant to size and potentially income differences across cities. The one exception to this finding is that the positive effects of density on kerosene consumption seem to come from smaller cities in Indonesia.

Time Stability of the Density Elasticity. Finally, we document that the elasticity of residential carbon emissions with respect to density is remarkably stable over time, despite large changes in urban populations and rapid migration experienced in Indonesia and the Philippines. To do so, we estimate the density elasticity using two older waves of data: a 2000 epoch for Indonesia and a 2010 epoch for the Philippines. In Indonesia, we combine data on residential energy expenditures from the 1996 *Susenas* with neighborhood sorting controls calculated from the 2000 Census. In the Philippines, we pool data from the 2006, 2009, and 2012 FIES together with sorting controls from the 2010 Census.

Appendix Table A.22 shows the results of using these earlier waves of cross-sectional data to estimate $\mathcal{E}_{\mathrm{CE}_{iv},\mathrm{density}}$ as defined in equation (8). We use national average carbon intensity weights and report separate for urban households in Indonesia and the Philippines. As before, to calculate standard errors, we estimate the individual energy expenditure equations simultaneously using a seemingly unrelated regressions system. Panel A shows that in the OLS and IV-Lasso specifications without sorting controls, a 1 percent increase in density is associated with a 0.02 percent increase in carbon emissions in Indonesia, with a similar effect size in the Philippines. However, in column 3, when we add controls for sorting, the effect sizes fall and become statistically insignificant. These precise zero effect sizes in models with sorting controls are very similar to what we found in Table 7 using more recent data waves.

6 Decomposing Overall Changes in Residential Emissions

So far, we have shown that energy demand in Indonesian and Philippines cities does not seem to be related to changes in community-level population density. Our estimates of the density elasticity are small, and the null density elasticity appears to be stable over time. If urban sprawl—as proxied by changes in population density—does not seem to play a role in driving increased residential carbon emissions, what is driving growth instead? In this section, we develop a novel Oaxaca-Blinder type of decomposition to explain growth in residential carbon emissions across cities. We calculate the components of this decomposition by combining multiple waves of detailed census data with our parameter estimates.

Decomposition. Let $c = 1, ..., C$ index cities in our sample, and let $v = 1, ..., N_v$ index communities in city c. Let $s \in \mathcal{S}$ also index types of households.[30]

[30] For this analysis, we group households into 3 size groups (1-2 person households, 3-4 person households, and households with more than 4 members) and 4 household-head education groupings (no schooling, some schooling, only completed high school, greater than high school completion). Therefore, there are 12 household types in our analysis (i.e. $|\mathcal{S}|$ = 12).

Appendix B explains how we combine our parameter estimates averages from census data, taken for each household type s in each neighborhood v, to predict city c's total residential carbon emissions at time t. Our estimate, denoted $\widehat{\text{CE}}_{ct}$, can be expressed as follows:

$$\widehat{\text{CE}}_{ct} = \sum_{v=1}^{N_c} \sum_{i \in s} N_{sv,t} \cdot \widehat{\text{CE}}_{sv,t}$$

where $N_{sv,t}$ measures the number of type s households in community v at time t and $\widehat{\text{CE}}_{sv,t}$ is the predicted average residential carbon emissions for a type s household in community v at time t. Our predictions combine census averages with estimates of the parameters in equation (6). Because our primary regression equations are log linear, we follow Duan (1983) to predict emissions levels.

The growth in predicted total residential carbon emissions from census wave t to $t+1$ in city c is given by:

$$\Delta\widehat{\text{CE}}_c \equiv \widehat{\text{CE}}_{c,t+1} - \widehat{\text{CE}}_{c,t} = \sum_{v=1}^{N_c} \sum_{s \in \mathcal{S}} \left(N_{sv,t+1} \cdot \widehat{\text{CE}}_{sv,t+1} - N_{sv,t} \cdot \widehat{\text{CE}}_{sv,t} \right)$$

Appendix B shows that $\Delta\widehat{\text{CE}}_c$ can be decomposed into three terms, as follows:

$$\Delta\widehat{\text{CE}}_c = \sum_{v=1}^{N_c} \sum_{s \in \mathcal{S}} \underbrace{(N_{sv,t+1} - N_{sv,t}) \cdot \widehat{\text{CE}}_{sv,t}}_{(\mathbf{A})} + \underbrace{\Delta\widehat{\text{CE}}_{sv}\,(\Delta\mathbf{x}) \cdot N_{sv,t+1}}_{(\mathbf{B})} + \underbrace{\Delta\widehat{\text{CE}}_{sv}\,(\Delta\beta) \cdot N_{sv,t+1}}_{(\mathbf{C})} \tag{9}$$

In this expression, term A reflects *population growth* of type s households in community v, multiplied by the predicted levels of carbon emissions in time t for those household types in those communities. Term B reflects growth in emissions due to *changes in the mapping* between $\mathbf{x}_{v,s}$ characteristics and emissions over time. Term C reflects growth in emissions due to *changes in characteristics* over time. Because we are fixing household types in this analysis, one source of these changes will be changes in community-level population density. The last two terms of the decomposition in equation (9) are similar to a Oaxaca-Blinder decomposition, but for time periods as opposed to groups. They also need adjustments in how they are calculated because of the log linear model we use.

Results. According to Global Carbon Budget Data (Friedlingstein et al., 2024), from 2000 to 2010, Indonesia produced an additional 175.0 million metric tons (MMT) of CO_2. This represents a 63 percent increase in emissions from a baseline of 276.6 MMT in 2000. Over a similar time period, from 2010 to 2018, the Philippines produced an additional 58.9 MMT of CO_2, which was a 71 percent increase from emissions in 2010 (83.0 MMT).

Before discussing the decomposition results, we first present descriptive evidence on the growth of residential carbon emissions in major cities in our sample, based on household survey and census data we use. In Figure 5, we plot the overall changes in emissions for the 10 largest cities in Indonesia (from 2000-2010, Panel A) and for some of the largest cities in the Philippines (from 2010-2018, Panel B).[31] The figure shows both overall changes in emissions as well as changes in emissions by energy source. In most of Indonesia's largest cities, overall emissions increased despite large reductions in emissions

[31] We show the 10 largest cities with coverage in the FIES in our results for the Philippines.

from kerosene over this period. The decline in kerosene emissions was offset by even larger increases in emissions from vehicle gas and electricity. In cities in the Philippines, emissions also grew, but these were driven largely by increases in electricity consumption.

In Indonesia, the 10 largest cities shown in Figure 5, Panel A, produced an additional 2.38 million metric tons of CO_2 per month from electricity, vehicle gas, LPG, and kerosene consumption. If we assume these emissions trends—which are based on monthly data—are constant throughout the year, this would account for an additional 28.6 MMT of CO_2 per year, or about 16 percent of Indonesia's total emissions growth from 2000-2010. In the Philippines, the cities in Figure 5, Panel B, produced an additional 0.71 million metric tons of CO_2 per month—or 8.6 million metric tons per year—from 2010-2018. This accounts for 14.6 percent of the total emissions growth in the Philippines over this period.

Figure 6 shows the results of decomposing this overall emissions growth by city into the different components described in equation (9). Panel A shows that in Indonesia, apart from Jakarta, overall emissions growth is typically not explained by changes in population. On average across the top 10 cities, 66.4 percent of the total absolute changes in emissions come from changes due to the mapping between community and household characteristics and emissions (term C). A further 18.6 percent comes from changes in characteristics (term B), while the remaining 15 percent comes from population growth (term A).

Figure 6, panel B shows somewhat different results for the Philippines. For a typical city in this panel, population growth only explains 18.1 percent of total emissions growth, and only 21.6 percent is explained by changes due to the mapping between community and household characteristics and emissions. Instead, changing characteristics explains 60.3 percent of the growth of emissions.

A key message of the paper is that density only modestly contributes to changes in emissions. Appendix Figure A.2 shows the results of further separating the effects of changes in characteristics (term B in equation 9) into a component owing to changes in density and a component due to changes in other factors.[32] On average, across the largest Indonesian cities, changing density explains only 1.1 percent of the changes in overall emissions by city. For the Philippines, although changing charateristics explains a larger share of total emissions growth, density changes still only explain 1.8 percent of the changes in overall emissions by city.

7 Conclusion

This paper presents causal estimates of the effect of neighborhood population density on residential carbon emissions, using data from cities in Indonesia and the Philippines. Estimating the effect of a place-specific feature like population density is plagued by two key identification problems: (1) simultaneity, in which omitted variables may drive correlations between both density and residential carbon emissions; and (2) sorting, where individuals with specific tastes for different emissions-based lifestyles may sort systematically into low or high density places. We confront the first identification challenge by instrumenting for density within cities using characteristics of the soils prevalent in different neighborhoods. We address sorting using a control function approach, following Altonji and Mansfield (2018).

We find that unlike many wealthier cities in the U.S. and Europe, higher income individuals tend to

[32]The details of this additional decomposition can be found in Appendix B.

sort closer to the city center, and they also settle in denser neighborhoods. Consequently, neighborhoods with higher levels of density tend to produce greater carbon emissions. Our IV specifications suggest that this positive correlation survives simultaneity bias. However, after we control for sorting, the effect of density attenuates and becomes statistically insignificant. Our preferred estimates suggest that increasing density by 1 percent reduces carbon emissions by -0.005 percent in Indonesian cities and increases emissions by 0.004 percent in Philippines cities, but these effects are statistically indistinguishable from zero.

We also find that this precise null density elasticity of residential carbon emissions is remarkably stable over time. Our results are robust to many threats to the exclusion restriction, and we find similar patterns using data on asset ownership. We also find surprisingly similar estimates in cities of different sizes. When probing the effect of sorting, economic sorting controls seem to dampen effect sizes instead of social or ethnic sorting controls.

If urban sprawl leads to undesireable externalities, it may be prudent to use growth controls as a policy to reduce these outcomes (Brueckner and Largey, 2008). Given that we do not find much scope for density-based place effects in Indonesia and the Philippines, such growth controls may not be warranted on these grounds, and it would be far better to use carbon pricing to better align marginal social costs with average private costs. However, other work has found a larger role for place effects in explaining spatial heterogeneity in residential carbon emissions (Lyubich, 2025). While population density is correlated with many important aspects of sprawl (Ewing and Cervero, 2001; Brownstone and Golob, 2009), other aspects of place may be more important. Further work should try to causally identify the effect of these other place-specic effects, including access to public transit, walkability, and neighborhood composition, to determine whether other policies may successfully reduce residential carbon emissions.

References

ABRAHAMS, A., C. ORAM, AND N. LOZANO-GRACIA (2018): "Deblurring DMSP nighttime lights: A new method using Gaussian filters and frequencies of illumination," *Remote Sensing of Environment*, 210, 242–258.

AHLFELDT, G. M., N. BAUM-SNOW, AND R. JEDWAB (2024): "The Skyscraper Revolution: Global Economic Development and Land Savings," Working Paper.

ALTONJI, J. G., T. E. ELDER, AND C. R. TABER (2005): "An evaluation of instrumental variable strategies for estimating the effects of catholic schooling," *Journal of Human resources*, 40, 791–821.

ALTONJI, J. G. AND R. K. MANSFIELD (2018): "Estimating group effects using averages of observables to control for sorting on unobservables: School and neighborhood effects," *American Economic Review*, 108, 2902–46.

ANGEL, S., J. PARENT, D. L. CIVCO, A. BLEI, AND D. POTERE (2011): "The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050," *Progress in Planning*, 75, 53–107.

ANGEL, S., S. SHEPPARD, D. L. CIVCO, R. BUCKLEY, A. CHABAEVA, L. GITLIN, A. KRALEY, J. PARENT, AND M. PERLIN (2005): *The dynamics of global urban expansion*, Citeseer.

BADAN PUSAT STATISTIK (BPS) (2008): *Statistik Listrik PLN 2003-2007*, Jakarta: BPS.

——— (2012): *Statistik Listrik 2007-2011*, Jakarta: BPS.

——— (2022): *Statistical Yearbook of Indonesia*, Jakarta: BPS.

BANK, W. (2024): "World Development Report 2024: the Middle-Income Trap," .

BELLONI, A., D. CHEN, V. CHERNOZHUKOV, AND C. HANSEN (2012): "Sparse Models and Methods for Optimal Instruments With an Application to Eminent Domain," *Econometrica*, 80, 2369–2429.

BERTANI, R. AND I. THAIN (2002): "Geothermal power generating plant CO2 emission survey," *IGA news*, 49, 1–3.

BLACK, D., K. DANIEL, AND S. SANDERS (2002): "The impact of economic conditions on participation in disability programs: Evidence from the coal boom and bust," *American Economic Review*, 92, 27–50.

BOQUET, Y. (2017): *The Philippine Archipelago*, Springer.

BORCK, R. AND P. SCHRAUTH (2021): "Population density and urban air quality," *Regional Science and Urban Economics*, 86, 103596.

BROWNSTONE, D. AND T. F. GOLOB (2009): "The impact of residential density on vehicle usage and energy consumption," *Journal of urban Economics*, 65, 91–98.

BRUECKNER, J. K. AND A. G. LARGEY (2008): "Social interaction and urban sprawl," *Journal of Urban Economics*, 64, 18–34.

BRUECKNER, J. K. AND K. S. SRIDHAR (2012): "Measuring welfare gains from relaxation of land-use restrictions: The case of India's building-height limits," *Regional Science and Urban Economics*, 42, 1061–1067.

BRUECKNER, J. K., J.-F. THISSE, AND Y. ZENOU (1999): "Why is Central Paris rich and Downtown Detroit poor?: An Amenity-based Theory," *European Economic Review*, 43, 91–107.

BURCHFIELD, M., H. G. OVERMAN, D. PUGA, AND M. A. TURNER (2006): "Causes of sprawl: A portrait from space," *The Quarterly Journal of Economics*, 121, 587–633.

CAROZZI, F. AND S. ROTH (2023): "Dirty density: Air quality and the density of American cities," *Journal of Environmental Economics and Management*, 118, 102767.

CIVELLI, A., A. GADUH, A. D. ROTHENBERG, AND Y. WANG (2023): "Urban Sprawl and Social Capital: Evidence from Indonesian Cities," *The Economic Journal*, 133, 2110–2146.

COMBES, P.-P., G. DURANTON, L. GOBILLON, AND S. ROUX (2010): "Estimating Agglomeration Economies with History, Geology, and Worker Effects," in *Agglomeration Economics*, ed. by E. L. Glaeser, National Bureau of Economic Research, 15–65.

CTR (2020): "Climate Transparency Report 2022," *URL: https://www. climate-transparency. org/g20-climate-performance/the-climatetransparency-report-2020# 1531904804037-423d5c88-a7a7*.

CUNNINGHAM, C. R. (2007): "Growth controls, real options, and land development," *The Review of Economics and Statistics*, 89, 343–358.

DE THÉ, C. B., B. CARANTINO, AND M. LAFOURCADE (2021): "The carbon 'carprint'of urbanization: New evidence from French cities," *Regional Science and Urban Economics*, 89, 103693.

DEFFEBACH, P., D. L. LAGAKOS, Y. MIYAUCHI, AND E. YAMADA (2025): "The Spatial Distribution of Income in Cities: New Global Evidence and Theory," Working Paper.

DELL, M. AND B. A. OLKEN (2019): "The development effects of the extractive colonial economy: the Dutch cultivation system in Java," *The Review of Economic Studies*, 87, 164–203.

DUAN, N. (1983): "Smearing estimate: a nonparametric retransformation method," *Journal of the American Statistical Association*, 78, 605–610.

DWI CAHYANI, A., N. D. NACHROWI, D. HARTONO, AND D. WIDYAWATI (2020): "Modern residential energy inequalities in Indonesia: spatial and income analyses," *Energy Sources, Part B: Economics, Planning, and Policy*, 15, 329–350.

EWING, R. AND R. CERVERO (2001): "Travel and the built environment: a synthesis," *Transportation research record*, 1780, 87–114.

FRIEDLINGSTEIN, P., M. O'SULLIVAN, M. W. JONES, R. M. ANDREW, J. HAUCK, P. LANDSCHÜTZER, C. LE QUÉRÉ, H. LI, I. T. LUIJKX, A. OLSEN, ET AL. (2024): "Global Carbon Budget 2024," *Earth System Science Data Discussions*, 2024, 1–133.

GAIGNÉ, C., H. R. KOSTER, F. MOIZEAU, AND J.-F. THISSE (2022): "Who lives where in the city? Amenities, commuting and income sorting," *Journal of Urban Economics*, 128, 103394.

GERTLER, P. J., O. SHELEF, C. D. WOLFRAM, AND A. FUCHS (2016): "The Demand for Energy-using Assets Among the World's Rising Middle Classes," *American Economic Review*, 106, 1366–1401.

GLAESER, E. (2011): *Triumph of the City: How Urban Spaces Make Us Human*, Pan Macmillan.

GLAESER, E. L. AND M. E. KAHN (2010): "The greenness of cities: Carbon dioxide emissions and urban development," *Journal of urban economics*, 67, 404–418.

GLAESER, E. L., M. E. KAHN, AND J. RAPPAPORT (2008): "Why do the poor live in cities? The role of public transportation," *Journal of urban Economics*, 63, 1–24.

GLAESER, E. L. AND B. A. WARD (2009): "The causes and consequences of land use regulation: Evidence from Greater Boston," *Journal of Urban Economics*, 65, 265–278.

GOLDSTEIN, B., D. GOUNARIDIS, AND J. P. NEWELL (2020): "The carbon footprint of household energy use in the United States," *Proceedings of the National Academy of Sciences*, 117, 19122–19130.

HENDERSON, J. V. AND M. A. TURNER (2020): "Urbanization in the Developing World: Too Early or Too Slow?" *Journal of Economic Perspectives*, 34, 150–73.

HENGL, T., J. M. DE JESUS, G. B. HEUVELINK, M. R. GONZALEZ, M. KILIBARDA, A. BLAGOTIĆ, W. SHANGGUAN, M. N. WRIGHT, X. GENG, B. BAUER-MARSCHALLINGER, ET AL. (2017): "SoilGrids250m: Global gridded soil information based on machine learning," *PLoS one*, 12.

HOXBY, C. M. (2000): "Does competition among public schools benefit students and taxpayers?" *American Economic Review*, 90, 1209–1238.

International Energy Agency (IEA) (2017): *World Energy Outlook*, Paris: IEA.

——— (2022a): *An Energy Sector Roadmap to Net Zero Emissions in Indonesia*, Paris: IEA.

——— (2022b): *World Energy Balances*, Paris: IEA.

Jiang, Y. (2021): "Asian cities: spatial dynamics and driving forces," *The Annals of Regional Science*, 66, 609–654.

Kleibergen, F. and R. Paap (2006): "Generalized reduced rank tests using the singular value decomposition," *Journal of Econometrics*, 133, 97 – 126.

Lucas Jr, R. E. (1993): "Making a miracle," *Econometrica*, 251–272.

Lyubich, E. (2025): "The Role of People versus Places in Individual Carbon Emissions," *American Economic Review*, 115, 1439–84.

Philippine Statistics Authority (PSA) (2018): *Family Income and Expenditure Survey: Volume I: National and Regional Estimates*, Quezon City, Philippines.

Roberts, M., F. G. Sander, and S. Tiwari (2019): "Time to Act: Realizing Indonesia's Urban Potential," Tech. rep., World Bank.

Rosenthal, S. S. and W. C. Strange (2008): "The Attenuation of Human Capital Spillovers," *Journal of Urban Economics*, 64, 373 – 389.

Rukmana, D. (2015): "The Change and Transformation of Indonesian Spatial Planning after Suharto's New Order Regime: The Case of the Jakarta Metropolitan Area," *International Planning Studies*, 20, 350–370.

Santillan, J. R. and C. Heipke (2024): "Assessing patterns and trends in urbanization and land use efficiency across the Philippines: a comprehensive analysis using global Earth observation data and SDG 11.3. 1 indicators," *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 92, 569–592.

Sari, D., M. Rizki, B. Nathania, M. Ahmad, P. G. Gan, and N. Noor (2021): "Indonesia Zero Emissions Application (EMISI): Methodology for Calculating Individual Emissions from Food, Clothing, Electricity Consumption, and Solid Waste," Technical note.

Surbakti, P. and I. B. P. Statistik (1995): "Indonesia's National Socio-Economic Survey: a continual data source for analysis on welfare development," *(No Title)*.

Thoday, K., P. Benjamin, M. Gan, and E. Puzzolo (2018): "The Mega Conversion Program from kerosene to LPG in Indonesia: Lessons learned and recommendations for future clean cooking energy expansion," *Energy for Sustainable Development*, 46, 71–81.

UN (2020): "A Recommendation on the Method to Delineate Cities, Urban and Rural Areas for International Statistical Comparisons," Background paper, United Nations Statistical Commission.

USDA (1975): *Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys*, 436, US Government Print.

——— (2015): *Illustrated Guide to Soil Taxonomy*, National Resources Conservation Service Lincoln, Nebraska.

U.S. Energy Information Administration (EIA) (2019): *International Energy Outlook 2019*, Washington, D.C.: EIA.

——— (2023): "Carbon Dioxide Emissions Coefficients," https://www.eia.gov/environment/emissions/co2_vol_mass.php, accessed: 2023-11-27.

van Kippersluis, H. and C. A. Rietveld (2018): "Beyond plausibly exogenous," *The Econometrics Journal*, 21, 316–331.

Wolfram, C., O. Shelef, and P. Gertler (2012): "How Will Energy Demand Develop in the Developing World?" *Journal of Economic Perspectives*, 26, 119–38.

Zheng, S., R. Wang, E. L. Glaeser, and M. E. Kahn (2011): "The greenness of China: household carbon dioxide emissions and urban development," *Journal of Economic Geography*, 11, 761–792.

Table 1: Comparing Indonesia and the Philippines

	Indonesia		Philippines	
	Value	**Rank**	**Value**	**Rank**
Population (million)	284.4	4 (of 195)	114.1	12 (of 195)
GDP Per Capita (PPP)	16,448	101 (of 186)	11,794	117 (of 186)
Gini	34.9	90 (of 171)	39.3	56 (of 171)
Poverty Rate	19.9	44 (of 216)	16.9	47 (of 216)
Urbanization Rate	55.3	128 (of 221)	46.9	154 (of 221)
Urban Primacy	7	149 (of 153)	27	92 (of 153)
Share Living in Slums	20.2	75 (of 172)	37.3	56 (of 172)

Notes: This table provides summary statistics for broad measures of population, output, and urbanization for Indonesia and the Philippines. Population measures come from the United Nations (2023-2025). GDP per capita is purchasing-power parity adjusted (PPP), from the World Bank in 2023. The Gini coefficient is from the World Bank (2021-2024). The poverty rate is the poverty headcount ratio, measured in 2021 PPP, of the percent of people living on less than $4 per day, from The World Bank. The urbanization rate measures the share of the total population living in urban areas in 2018, also from the World Bank. Urban primacy measures the share of the urban population living in the largest city in 2018, from the World Bank. The share living in slums measures the percent of the urban population living in slums in 2018, from the World Bank.

Table 2: Summary Statistics on Monthly Energy Consumption

	Indonesia (2010)		Philippines (2018)	
Variable	**Mean (Std. Dev)**	*N*	**Mean (Std. Dev)**	*N*
Electricity (kWh)	77.844 (53.858)	40,920	138.687 (136.884)	47,635
LPG (kg)	5.027 (4.069)	40,920	5.621 (3.883)	47,635
Vehicle Gas (L)	13.149 (14.819)	40,920	9.692 (23.912)	47,635
Kerosene (L)	1.219 (4.531)	40,920	0.137 (0.899)	47,635
Carbon emissions (tons)	0.077 (0.053)	40,920	0.081 (0.091)	47,635

Notes: This table provides summary statistics on our main energy consumption outcomes for urban households in the Indonesia (2010) sample and for urban households in the Philippines (2018) sample. The Indonesian data are pooled from three *Susenas* rounds in 2010, 2011, and 2012, while the Philippines data come from the 2018 FIES. The final row summarizes estimates of the household residential carbon emissions implied by these expenditure quantities. Carbon emissions estimates are obtained by multiplying each category's fuel consumption by carbon emissions factors (reported in Appendix Table A.3).

Table 3: First Stage: Density and Soil Characteristics

	Indonesia		Philippines	
	(1)	(2)	(3)	(4)
Soil bulk density at 60cm depth	0.033***	0.015***		
	(0.002)	(0.002)		
Sand content at 60 cm depth (% (kg / kg))	-0.049***	-0.020***		
	(0.006)	(0.004)		
Great Group: Haplustolls (Mollisols)	0.676***	0.123*		
	(0.106)	(0.068)		
Great Group: Tropudults (Ultisols)	-1.119***	-0.848***		
	(0.196)	(0.173)		
Great Group: Chromusterts (Vertisols)	0.745***	0.473***		
	(0.107)	(0.092)		
Soil bulk density at 200cm depth			0.020***	0.012***
			(0.003)	(0.003)
Soil water content at 200cm depth			-0.037***	-0.017**
			(0.009)	(0.007)
Great Group: Haplustalfs (Alfisols)			0.377***	0.164*
			(0.096)	(0.085)
N	40,207	40,207	41,249	41,249
N Clusters	1,329	1,329	212	212
Adj. R^2	0.517	0.755	0.586	0.740
Adj. R^2 (Within)	0.342	0.666	0.335	0.582
Regression F-Stat	78.4	77.2	28.9	96.0
City FE	Yes	Yes	Yes	Yes
$\mathbf{X}_i$ Controls	Yes	Yes	Yes	Yes
$\mathbf{C}_{2v}$ Controls	Yes	Yes	Yes	Yes
$\mathbf{X}_v$ Controls	No	Yes	No	Yes

Notes: This table reports estimates of equation (5), the household-level first stage relationship between log population density (the dependent variable) and different soil characteristics variables. We use post-double-selection lasso regressions, following Belloni et al. (2012), to select instruments in these regressions from a set of 67 soil characteristics. Columns 1 and 2 are limited to the sample of urban communities in Indonesia covered by the 2010 *Susenas* epoch, while columns 3 and 4 are limited to urban communities in the Philippines covered by the 2018 FIES. All regressions include controls for household characteristics, city fixed effects, elevation, ruggedness, distance to the nearest point on the coast, and distance to the nearest river. In columns 2 and 4, we add village-level controls for sorting on observables and unobservables, denoted by $\mathbf{X}_v$. Robust standard errors, clustered at the subdistrict-level in columns 1-2 and clustered at the municipality level in columns 3-4, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 4: The Effect of Density on Electricity Use

	OLS	IV-Lasso	
Panel A: Indonesia	**(1)**	**(2)**	**(3)**
Log Density (2010)	0.082***	0.099***	-0.002
	(0.009)	(0.019)	(0.040)
N	40,207	40,207	40,207
N Clusters	1,329	1,329	1,329
Adjusted R^2	0.221	0.217	0.232
Adjusted R^2 (within)	0.129	0.125	0.142
Kleibergen-Paap Wald Rank F Stat		92.079	26.289
Under Id. Test (KP Rank LM Stat)		181.466	84.511
p-Value		0.000	0.000
Sargan-Hansen Test (Overidentification)		11.889	3.905
p-Value		0.018	0.419
Panel B: Philippines	**(1)**	**(2)**	**(3)**
Log Density (2018)	0.057***	0.181***	0.068
	(0.010)	(0.033)	(0.042)
N	41,249	41,249	41,249
N Clusters	212	212	212
Adjusted R^2	0.283	0.283	0.296
Adjusted R^2 (within)	0.193	0.193	0.208
Kleibergen-Paap Wald Rank F Stat		27.066	14.546
Under Id. Test (KP Rank LM Stat)		31.607	20.518
p-Value		0.000	0.000
Sargan-Hansen Test (Overidentification)		1.641	3.280
p-Value		0.440	0.194
City FE	Yes	Yes	Yes
$\mathbf{X}_i$ Controls	Yes	Yes	Yes
$\mathbf{C}_{2v}$ Controls	Yes	Yes	Yes
$\mathbf{X}_v$ Controls	No	No	Yes

Notes: Each cell reports the coefficient on log population density from equation (6) where the dependent variable is the log quantity of electricity (kWh) consumed last month. Panel A reports results for households in Indonesian cities, while Panel B reports results for households in Philippines cities. Column 1 reports OLS estimates, while Columns 2 and 3 apply a post-double-selection IV-lasso estimator, following Belloni et al. (2012). All regressions include city fixed effects and controls for $\mathbf{X}_i$ and $\mathbf{C}_{2v}$. Column 3 additionally includes $\mathbf{X}_v$ controls for sorting. The specific variables we include in $\mathbf{X}_i$, $\mathbf{C}_{2v}$ and $\mathbf{X}_v$, as well as their coefficients, are reported in Appendix Tables A.7 and A.9 for Indonesia (and Appendix Tables A.8 and A.10 for the Philippines). Robust standard errors, clustered at the subdistrict-level in Panel A (and the municipality level in Panel B), are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 5: Density and Residential Energy Use

	Electricity			LPG			Vehicle Gas			Kerosene		
	OLS	**IV-Lasso**		**OLS**	**IV-Lasso**		**OLS**	**IV-Lasso**		**OLS**	**IV-Lasso**	
Panel A: Indonesia	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Log Density (2010)	0.082***	0.099***	-0.002	0.086***	0.070***	-0.052	0.002	0.027	-0.050	0.037***	0.056***	0.100**
	(0.009)	(0.019)	(0.040)	(0.009)	(0.019)	(0.041)	(0.012)	(0.023)	(0.049)	(0.008)	(0.018)	(0.041)
N	40,207	40,207	40,207	40,207	40,207	40,207	40,207	40,207	40,207	40,207	40,207	40,207
N Clusters	1,329	1,329	1,329	1,329	1,329	1,329	1,329	1,329	1,329	1,329	1,329	1,329
Kleibergen-Paap Wald Rank F Stat		92.079	26.289		92.079	26.289		92.079	26.289		92.079	26.289
Panel B: Philippines	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Log Density (2018)	0.057***	0.181***	0.068	0.014	0.091***	0.007	-0.106***	0.025	0.047	-0.000	-0.002	-0.001
	(0.010)	(0.033)	(0.042)	(0.009)	(0.029)	(0.037)	(0.013)	(0.033)	(0.055)	(0.000)	(0.002)	(0.003)
N	41,249	41,249	41,249	43,687	43,687	43,687	42,887	42,887	42,887	45,312	45,312	45,312
N Clusters	212	212	212	212	212	212	212	212	212	212	212	212
Kleibergen-Paap Wald Rank F Stat		27.066	14.546		27.163	14.117		26.841	13.915		27.747	15.646
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\mathbf{X}_i$ Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\mathbf{C}_{2v}$ Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\mathbf{X}_v$ Controls	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Each cell reports the coefficient on log population density from equation (6) where the dependent variable is listed in the column header. Panel A reports results for households in Indonesian cities, while Panel B reports results for households in Philippines cities. Columns 1, 4, 7, and 10 report OLS estimates, while the other columns apply a post-double-selection IV-Lasso estimator, following Belloni et al. (2012). All regressions include city fixed effects and controls for $\mathbf{X}_i$ and $\mathbf{C}_{2v}$. Columns 3, 6, 9, and 12 additionally include $\mathbf{X}_v$ sorting controls. Robust standard errors, clustered at the subdistrict-level in Panel A (and the municipality level in Panel B), are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 6: Density and Asset Ownership: IV-Lasso Estimates

	Indonesia		**Philippines**	
	(1)	**(2)**	**(3)**	**(4)**
Refrigerator (0 1)	0.051***	-0.017	0.065***	0.016
	(0.009)	(0.019)	(0.016)	(0.021)
AC (0 1)	0.021***	0.021	0.049***	0.010
	(0.005)	(0.015)	(0.011)	(0.014)
Motorcycle (0 1)	0.001	-0.010	-0.006	0.030
	(0.008)	(0.019)	(0.013)	(0.021)
Car (0 1)	0.008	0.002	0.018**	-0.008
	(0.006)	(0.018)	(0.008)	(0.012)
LPG Cylinder (0 1)	0.029***	-0.007		
	(0.009)	(0.027)		
Water Heater (0 1)	-0.001	0.005		
	(0.007)	(0.021)		
N	40,207	40,207	46,754	46,754
N Clusters	1,329	1,329	212	212
City FE	Yes	Yes	Yes	Yes
$\mathbf{X}_i$ Controls	Yes	Yes	Yes	Yes
$\mathbf{C}_{2v}$ Controls	Yes	Yes	Yes	Yes
$\mathbf{X}_v$ Controls	No	Yes	No	Yes

Notes: Each cell reports the coefficient on log population density from equation (6) where the dependent variable is listed in the row header. All columns report results from post-double-selection IV-Lasso estimators, following Belloni et al. (2012). Columns 1-2 report results for households in Indonesian cities, while columns 3-4 show results for households in Philippines cities. All regressions include city fixed effects and controls for $\mathbf{X}_i$ and $\mathbf{C}_{2v}$. Columns 2 and 4 additionally include $\mathbf{X}_v$ sorting controls. Robust standard errors, clustered at the subdistrict-level in Columns 1-2 and the municipality level in Columns 3-4, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 7: The Average Density Elasticity of Residential Carbon Emissions

	OLS	IV-Lasso	
Panel A: Indonesia	**(1)**	**(2)**	**(3)**
Log Density (2010)	0.007***	0.008***	-0.005
	(0.001)	(0.002)	(0.004)
N	40,207	40,207	40,207
Panel B: Philippines	**(1)**	**(2)**	**(3)**
Log Density (2018)	0.000	0.012***	0.004
	(0.001)	(0.002)	(0.003)
N	43,284	43,284	43,284
City FE	Yes	Yes	Yes
$\mathbf{X}_i$ Controls	Yes	Yes	Yes
$\mathbf{C}_{2v}$ Controls	Yes	Yes	Yes
$\mathbf{X}_v$ Controls	No	No	Yes

Notes: This table reports estimates of the average density elasticity of residential carbon emissions for urban households in Indonesia (Panel A) and the Philippines (Panel B). The object estimated is described in equation (8). To obtain this, we estimate the density elasticity for our four carbon intensity outcomes simultaneously, using a SUR system and national carbon intensity weights. All regressions include city fixed effects and controls for $\mathbf{X}_i$ and $\mathbf{C}_{2v}$. Columns 2 and 3 report results from post-double-selection IV-lasso estimators, following Belloni et al. (2012), and column 3 additionally includes $\mathbf{X}_v$ controls. Robust standard errors, clustered at the subdistrict-level in Panel A and the municipality level in Panel B, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table 8: Mechanisms of Density: Floor Area and Commuting Distance

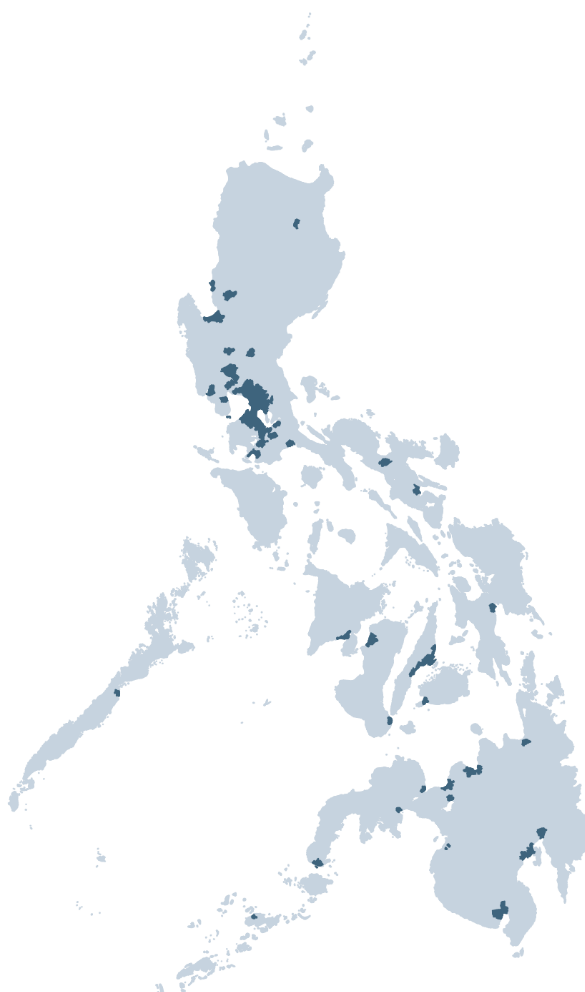
	log(Floor Area Per Capita)			log(Commuting Distance)		
	OLS	IV-Lasso		OLS	IV-Lasso	
Panel A: Indonesia	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Log Density (2010)	-0.071***	-0.058***	-0.020	0.008	0.004	0.006
	(0.006)	(0.013)	(0.028)	(0.014)	(0.032)	(0.056)
N	40,207	40,207	40,207	40,056	40,056	40,056
N Clusters	1,329	1,329	1,329	1,325	1,325	1,325
Kleibergen-Paap Wald Rank F Stat		91.999	26.265		92.006	26.201
Panel B: Philippines	**(1)**	**(2)**	**(3)**			
Log Density (2018)	-0.080***	-0.027	-0.057			
	(0.011)	(0.029)	(0.039)			
N	46,754	46,754	46,754			
N Clusters	212	212	212			
Kleibergen-Paap Wald Rank F Stat		26.864	14.807			
City FE	Yes	Yes	Yes	Yes	Yes	Yes
$\mathbf{X}_i$ Controls	Yes	Yes	Yes	Yes	Yes	Yes
$\mathbf{C}_{2v}$ Controls	Yes	Yes	Yes	Yes	Yes	Yes
$\mathbf{X}_v$ Controls	No	No	Yes	No	No	Yes

Notes: Each cell reports the coefficient on log population density from equation (6) where the dependent variable is listed in the column header. Panel A reports results for households in Indonesian cities, while Panel B reports results for households in Philippines cities. Columns 1 and 4 report OLS estimates, while the other columns apply a post-double-selection IV-Lasso estimator, following Belloni et al. (2012). All regressions include city fixed effects and controls for $\mathbf{X}_i$ and $\mathbf{C}_{2v}$. Columns 3 and 6 additionally include $\mathbf{X}_v$ sorting controls. Robust standard errors, clustered at the subdistrict-level in Panel A (and the municipality level in Panel B), are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Figure 1: Urban Areas in Indonesia and the Philippines

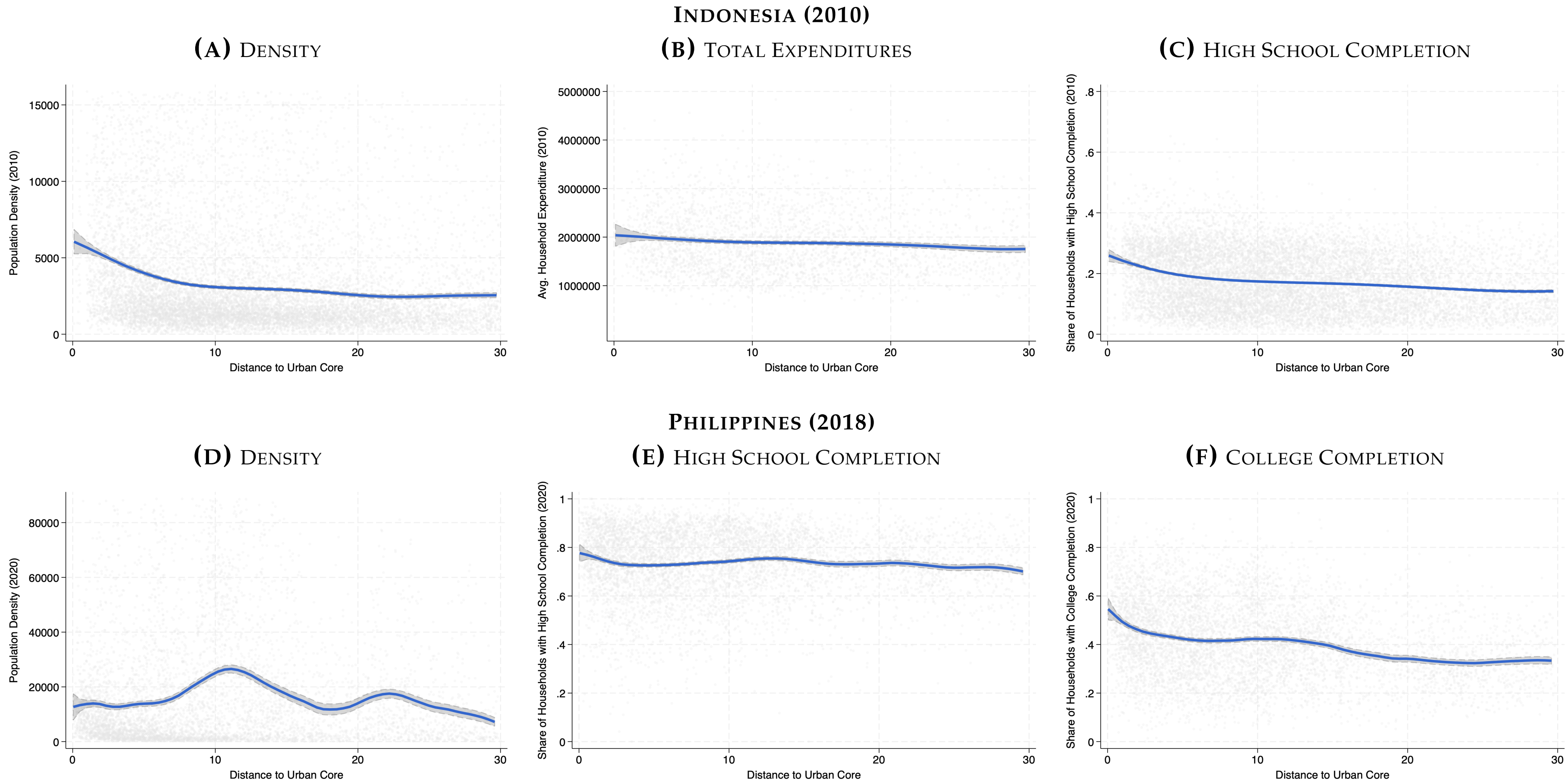
(a) Indonesia

(b) Philippines



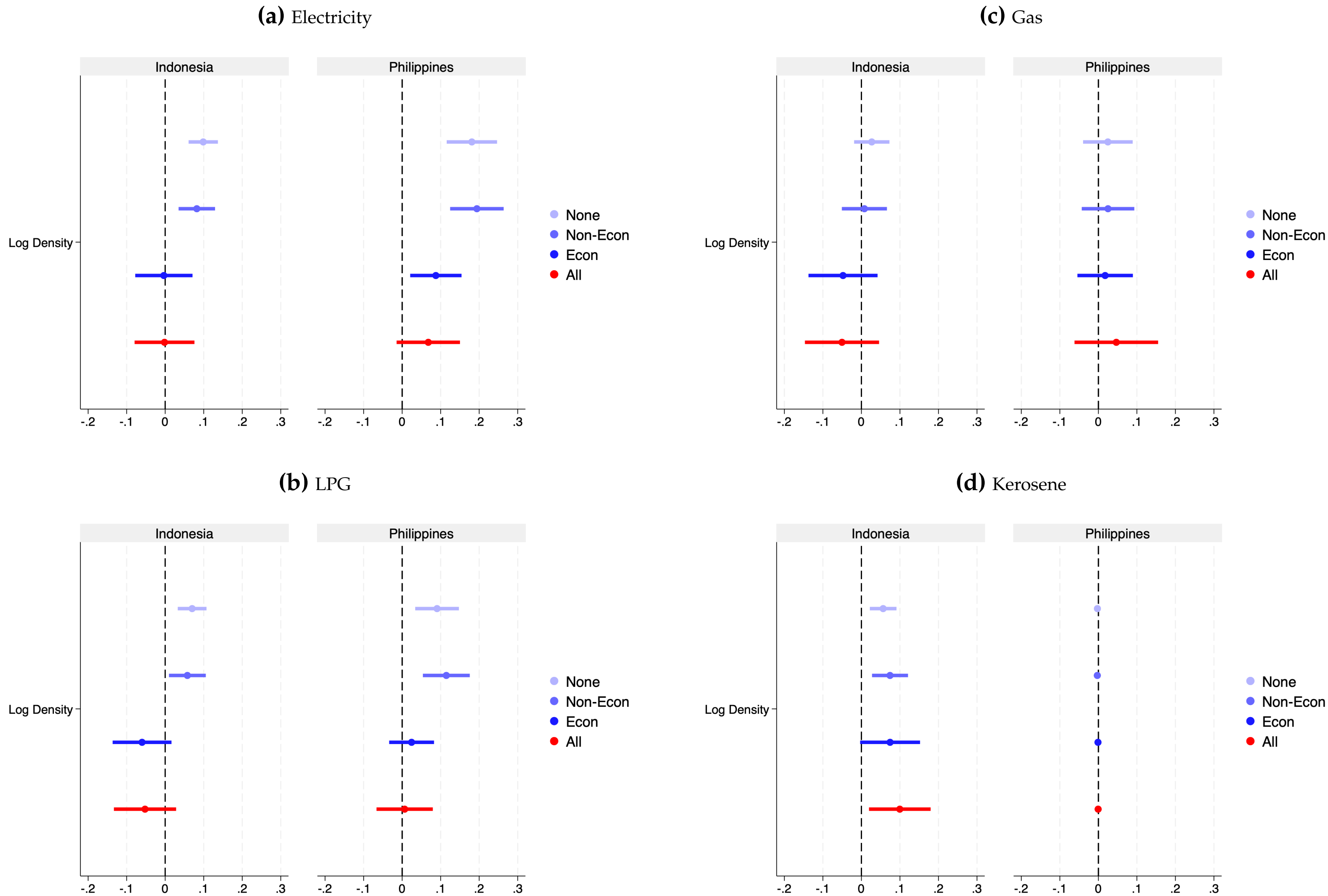
Notes: This figure presents a map of 91 urban areas in Indonesia and 42 urban areas in the Philippines, following Jiang (2021) and using nighttime lights satellite imagery to delineate cities.

Figure 2: Population Density, Income, Skill, and Distance to the CBD



Notes: These figures plot local polynomial regressions of census community aggregates against distance to the CBD, pooled across cities. Regression lines are plotted in blue, along with 95% confidence bands in gray. Panels A-C focus on *desa*-level data from Indonesia, using the 2010 Census and *Susenas* data, while Panels D-F focus on *barangay*-level data from the Philippines, using data from the 2020 Census and *FIES* data. Local polynomial regressions use an Epanechnikov kernel, rule-of-thumb bandwidth and a local cubic function, and are estimated on data from our sample of metropolitan areas.

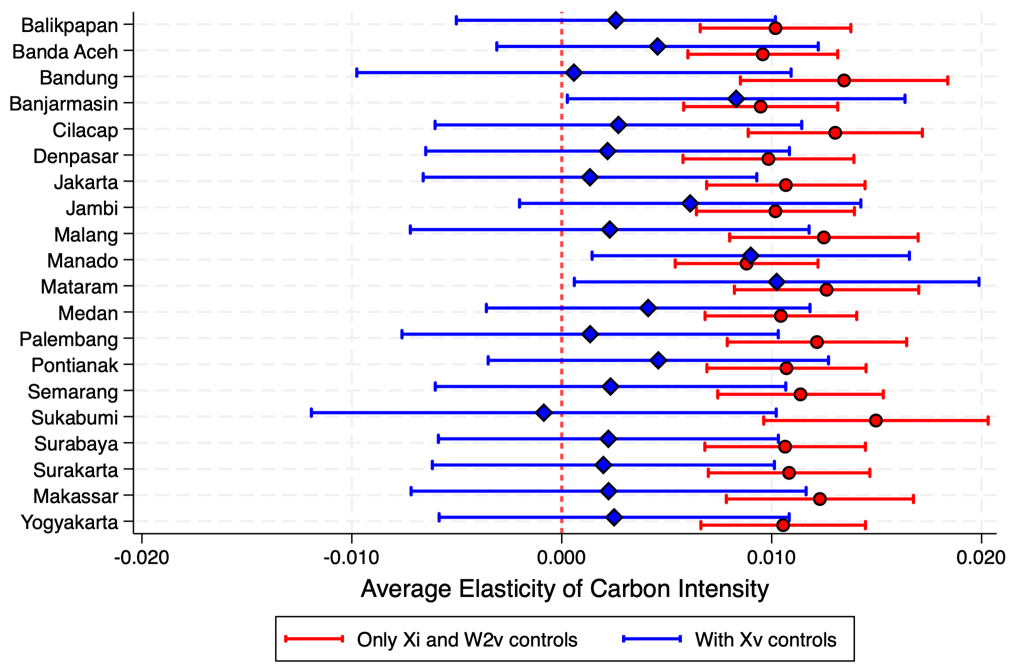
Figure 3: Density and Residential Energy Consumption: Probing the Effect of Sorting



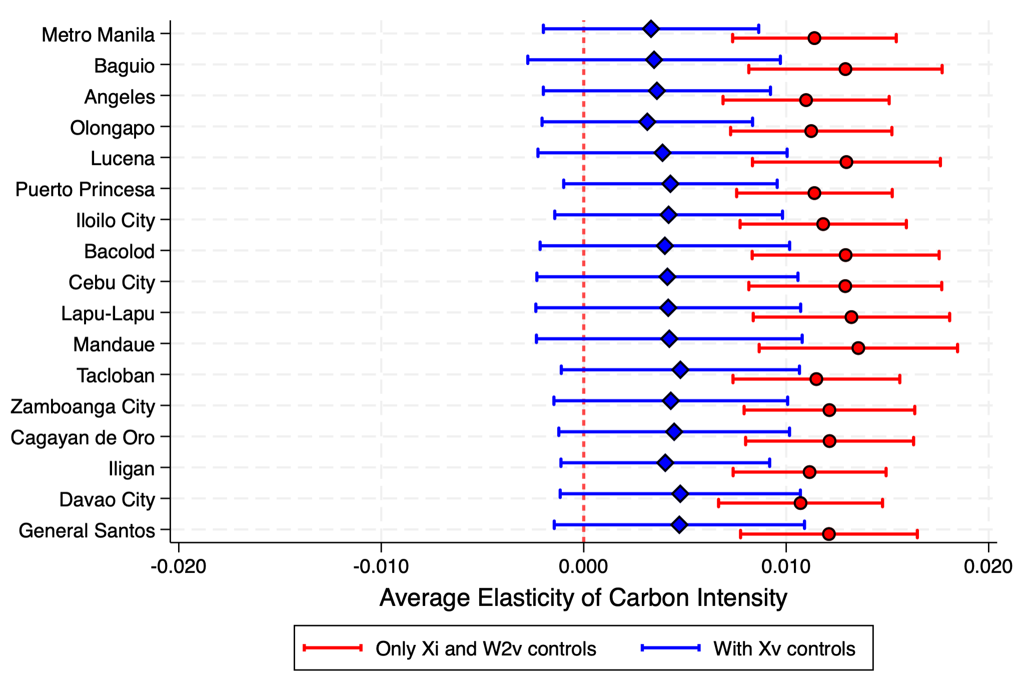
Notes: Each point reports the coefficient on log population density in 2010 from equation (6) where the dependent variable is listed in the panel header. The lines report 95% confidence intervals. All estimates are results from a post-double-selection IV-lasso estimator, following Belloni et al. (2012). The left sub-panels results for households in Indonesian cities, while the right sub-panels show results for households in Philippines cities. All regressions include city fixed effects and controls for $\mathbf{X}_i$ and $\mathbf{C}_{2v}$. The "non-econ" bars include non-economic $\mathbf{X}_v$ controls (e.g. ethnic composition shares and shares of different members of the community in different religious groups). The "econ" bars include economic $\mathbf{X}_v$ controls (e.g. education shares, household size, and employment characteristics). The "All" bars include the full set of $\mathbf{X}_v$ controls.

Figure 4: The Density Elasticity of Residential Carbon Emissions, by City

(a) Indonesia



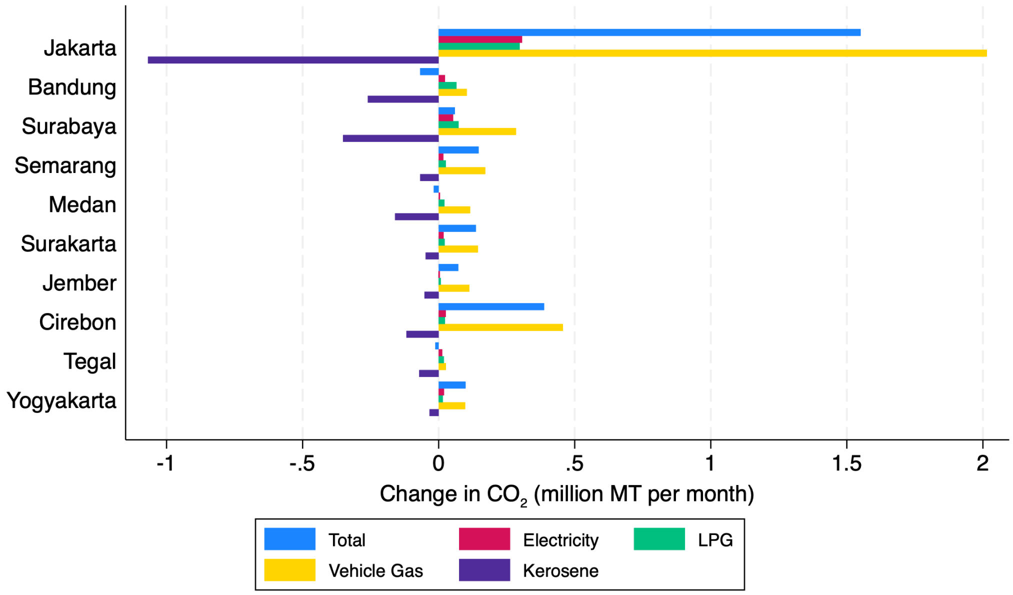
(b) Philippines



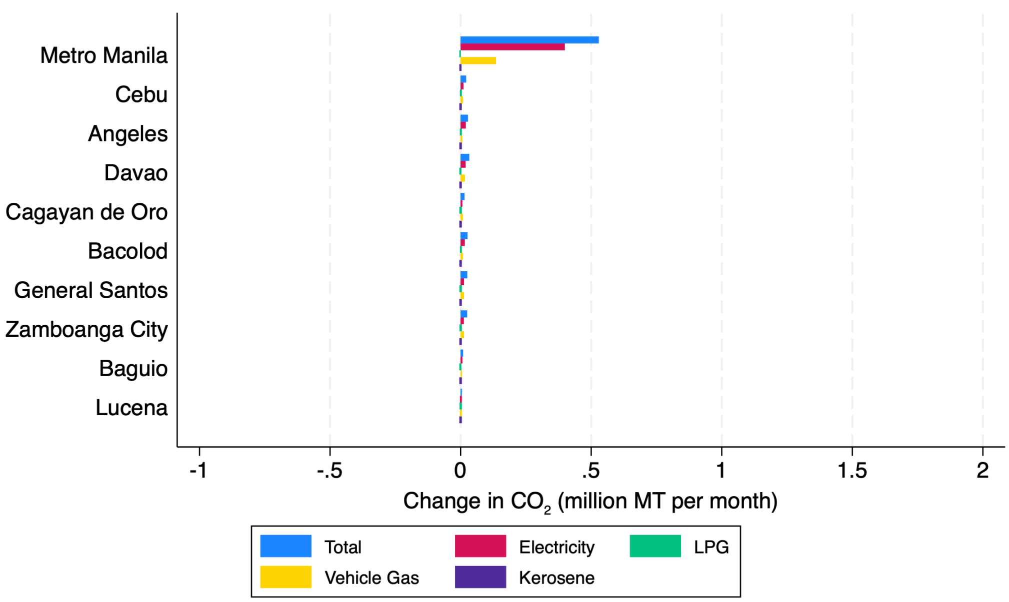
Notes: This figure reports estimates of the average density elasticity of residential carbon emissions for urban households in Indonesia (Panel A) and the Philippines (Panel B). The object estimated is described in equation (8), but with city-specific carbon intensity weights. To obtain these results, we estimate the four carbon intensity outcomes simultaneously, using a SUR system and city-specific carbon intensity weights. The point estimates (in dots) and 95 confidence intervals (lines) come from post-double-selection IV-lasso estimators, following Belloni et al. (2012). The red lines report IV-Lasso estimates without $\mathbf{X}_v$ controls, while the blue lines include $\mathbf{X}_v$ controls. The 95 percent confidence intervals account for clustering at the subdistrict-level in Panel A and the municipality level in Panel B.

Figure 5: Changes in Emissions by City and Emissions Type

(A) Indonesia



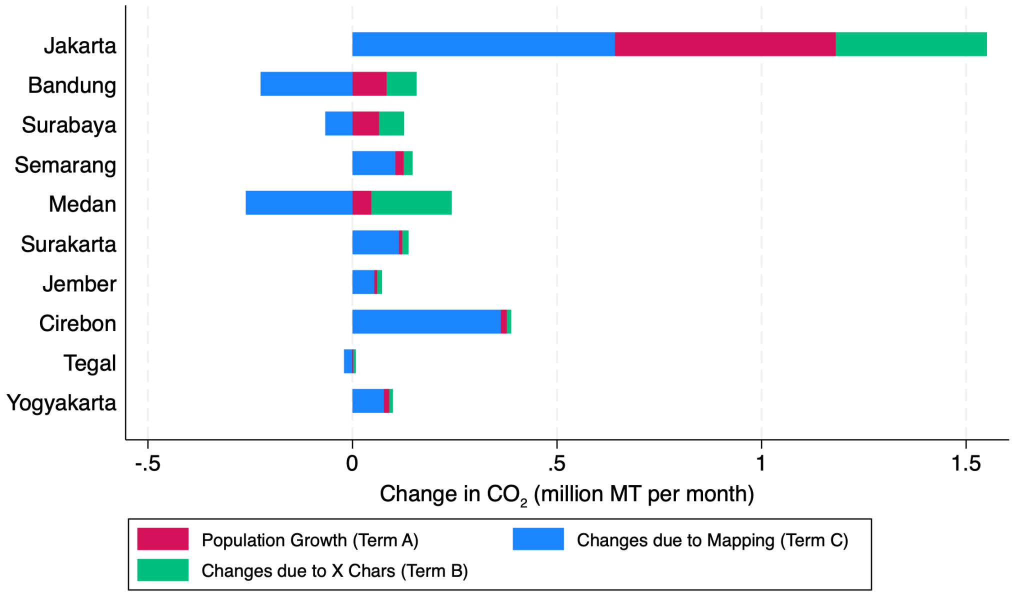
(B) Philippines



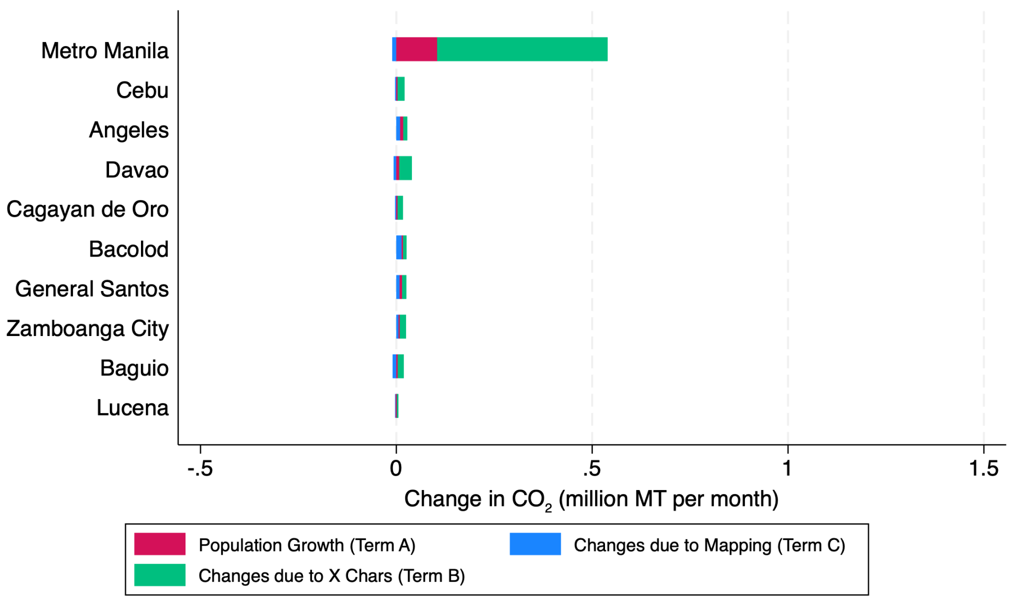
Notes: This figure presents estimates of changes in overall residential carbon emissions by city, $\Delta\widehat{\text{CE}}_c$, and growth in emissions by different sources.

Figure 6: Overall Emissions Growth Decomposition: Results for Top 10 Cities

(A) INDONESIA



(B) PHILIPPINES



Notes: This figure presents our overall decomposition of changes in predicted emissions by city, using equation (9). Appendix B explains how we calculate each of these terms in detail.

Online Appendix

Lathrop, T., Rothenberg, A., and Wang, Y. (2025): "Urban Sprawl and Residential Carbon Emissions: Evidence from Indonesian Cities"

Table of Contents

List of Tables

List of Figures

A Additional Tables and Figures

Table A.1: List of Cities in Indonesia

	2016 Population		2016 Population
Jakarta	26,369,334	Lumajang	393,315
Bandung	6,982,046	Palu	392,189
Surabaya	6,529,726	Bengkulu	385,937
Medan	4,586,536	Cianjur	383,911
Semarang	2,611,846	Gorontalo	376,157
Cirebon	2,190,797	Salatiga	367,692
Malang	2,024,502	Pematangsiantar	360,012
Surakarta	1,948,117	Binjai	357,775
Palembang	1,894,083	Kuningan	351,951
Yogyakarta	1,870,661	Tulungagung	316,244
Denpasar	1,858,784	Singkawang	315,124
Tegal	1,840,919	Jepara	291,567
Ujung Pandang	1,800,338	Padangsidempuan	285,540
Jember	1,755,505	Tebingtinggi	271,924
Kudus	1,367,919	Jaya Pura	267,991
Padang	1,245,331	Kupang	267,959
Mataram	1,210,029	Duri	262,885
Karawang	1,202,381	Tanjungpinang	261,386
Bandar Lampung	1,193,621	Batang	254,634
Banjarmasin	1,097,384	Pemalang	246,337
Tasikmalaya	1,053,857	Banyuwangi	244,678
Pontianak	1,004,170	Batu	212,171
Jombang	995,394	Singaraja	205,661
Pekan Baru	946,627	Ponorogo	201,991
Batam	945,092	Metro	194,940
Pekalongan	840,266	Blitar	188,452
Samarinda	825,220	Sibolga	177,241
Purwokerto	809,891	Tarakan	169,193
Bojonegoro	733,778	Subang	159,798
Indramayu	723,226	Lhokseumawe	159,756
Magelang	695,771	Kisaran	154,074
Cikampek	688,233	Pacet	153,814
Jambi	624,013	Tanjung Balai	149,399
Cilegon	603,137	Bontang	146,554
Manado	593,758	Bukit Tinggi	134,541
Purwakarta	578,127	Pangkalpinang	121,012
Mojokerto	532,297	Rantauprapat	117,003
Garut	515,288	Parepare	108,366
Balikpapan	493,790	Brebes	104,199
Sukabumi	482,345	Watampone	74,085
Pasuruan	467,265	Palopo	71,195
Karangampel	447,372	Sorong	57,472
Pamekasan	446,891	Bitung	43,828
Dumai	446,388		
Banda Aceh	422,231		
Kebumen	412,782		
Cilacap	411,080		
Kediri	411,024		
Probolinggo	404,936		
Madiun	393,678		

Notes: This table provides a list of 93 urban areas in Indonesia, identified following the procedure in Jiang (2021) that uses nighttime lights satellite imagery to delineate cities. Population totals are based on 2016 Landscan Data.

Table A.2: List of Cities in the Philippines

	2016 Population
Metro Manila	24,253,738
Cebu	2,801,102
Angeles	1,569,482
Davao	1,545,521
Cagayan de Oro	777,611
Bacolod	667,295
General Santos	639,191
Zamboanga	585,040
Iloilo	537,811
Batangas	486,125
Baguio	451,965
Dagupan	383,634
Naga	373,422
Cotabato	361,203
Lipa	338,079
Iligan	331,163
Olongapo	326,828
Tarlac	312,772
Lucena	290,920
Tacloban	258,498
San Pablo	218,385
Butuan	182,180
Dumaguete	170,545
San Pedro	159,905
Marawi	147,534
Tuguegarao	129,505
Urdaneta	123,855
Silay	119,167
San Fernando	112,496
Ormoc	92,578
Calapan	89,885
Roxas	76,138
Ozamis	74,127
Surigao	47,994
Toledo	43,836
Kidapawan	43,413
San Carlos, Negros Occidental	38,182
Calbayog	33,272
Cadiz	33,242
Kabankalan	22,650
San Carlos, Pangasinan	22,451
Bago	12,548

Notes: This table provides a list of 42 urban areas in the Philippines, identified following the procedure in Jiang (2021) that uses nighttime lights satellite imagery to delineate cities. Population totals are based on 2016 Landscan Data.

Table A.3: Carbon Dioxide Emissions Factors

Emissions Source	**Factor**	**Unit**	**Source**
Propane (LPG)	5.750	kilograms of CO_2 per gallon	EIA (2023)
Motor Gasoline	8.780	kilograms of CO_2 per gallon	EIA (2023)
Kerosene	9.880	kilograms of CO_2 per gallon	EIA (2023)
Electric Power Source	**Factor**	**Unit**	**Source**
Diesel Power	0.264	U.S. tons of CO_2 per MW hour (Fuel = Fuel Oil)	Sari et al. (2021)
Gas Machine Power	0.200	U.S. tons of CO_2 per MW hour (Fuel = Natural Gas)	Sari et al. (2021)
Gas Turbine Power	0.200	U.S. tons of CO_2 per MW hour (Fuel = Natural Gas)	Sari et al. (2021)
Geothermal Power	0.134	U.S. tons of CO_2 per MW hour	Bertani and Thain (2002)
Hydro Power	0.000		
Micro Hydro Power	0.000		
Steam Gas Power	0.404	U.S. tons of CO_2 per MW hour (Fuel = Coal)	Sari et al. (2021)
Steam Power	0.404	U.S. tons of CO_2 per MW hour (Fuel = Coal)	Sari et al. (2021)
Wind Power	0.000		

Notes: This table reports carbon emissions factors by source that we use to estimate household residential carbon emissions. Different factors are used for households that obtain their electricity from different types of power plants. We use government data from both countries on the locations of different types of power plants, and we match households to their nearest power plants to obtain the appropriate emissions factor for their electricity consumption. We assume that hydro, micro-hydro, and wind power produce zero tons of CO_2 per MW hour.

Table A.4: Kleibergen and Paap (2006) Cluster-Robust Tests of the Rank of the $\mathbf{X}_v$ Covariance Matrix

# Factors		**Indonesia**		**Philippines**	
		Full *Susenas* (p-value)	**Urban *Susenas* (p-value)**	**Full FIES (p-value)**	**Urban FIES (p-value)**
H_0	H_A	**(1)**	**(2)**	**(3)**	**(4)**
0	1+	0.000	0.000	0.000	0.000
1	2+	0.000	0.000	0.000	0.000
2	3+	0.000	0.000	0.000	0.000
3	4+	0.000	0.000	0.000	0.000
4	5+	0.000	0.000	0.000	0.000
5	6+	0.000	0.000	0.000	0.000
6	7+	0.000	0.000	0.000	0.000
7	8+	0.000	0.000	0.000	0.000
8	9+	0.000	0.000	0.000	0.000
9	10+	0.000	0.000	0.000	0.000
10	11+	0.000	0.000	0.000	0.000
11	12+	0.000	0.004	0.000	0.000
12	13+	0.000	0.025	0.000	0.000
13	14+	0.000	0.039	0.000	0.000
14	15+	0.005	0.237	0.000	0.000
15	16+	0.004	0.436	0.000	0.000
16	17+	0.036	0.452	0.028	0.028
17	18+	0.325	0.748	0.228	0.228
18	19+	0.412	0.855	0.542	0.542
19	20+	0.793	0.939	0.870	0.870
20	21+	0.733	0.956	0.967	0.967
21	22+	0.939	0.930	0.895	0.895

Notes: Each element of this table reports a p-value from a test of the null hypothesis that the rank of the covariance matrix of $\mathbf{X}_v$ is equal to the value associated with the row label, against the alternative that the rank exceeds this value. The test statistic is described in Kleibergen and Paap (2006), and the p-values we report are robust and account for clustering at the sub-district level in Indonesia and clustering at the muncipality level in the Philippines. Column 1 performs these tests on the full *Susenas* sample in Indonesia, while column 2 restricts the sample to urban areas in Indonesia. Column 3 performs these tests on the full FIES sample in the Philippines, while column 4 restricts the sample to urban areas in the Philippines.

Table A.5: Principal Components Analysis of $\mathbf{X}_v$

	Indonesia (*Susenas*)		Philippines (*FIES*)	
	Full (1)	**Urban (2)**	**Full (3)**	**Urban (4)**
# of Variables in $\mathbf{X}_v$	38	38	38	38
# of factors needed to explain:				
... 75% of total $\mathbf{X}_v$ variation	16	16	19	17
... 90% of total $\mathbf{X}_v$ variation	23	23	26	24
... 95% of total $\mathbf{X}_v$ variation	27	27	29	28
... 99% of total $\mathbf{X}_v$ variation	31	31	35	33
... 100% of total $\mathbf{X}_v$ variation	36	36	38	38

Notes: This table reports a principal components analysis of the 38 $\mathbf{X}_v$ variables for the Indonesia analysis and the 38 variables in the Philippines analysis, both for the full *Susenas* / FIES samples (columns 1 and 3) and for our urban *Susenas* / FIES samples described in Section 3 (columns 2 and 4). The first row lists the number of variables in $\mathbf{X}_v$. The next set of rows report the number of factors needed to explain 75%, 90%, 95%, 99% and 100% of the total variation in $\mathbf{X}_v$.

Table A.6: First Stage, Community Level

	Indonesia		Philippines	
	(1)	(2)	(3)	(4)
Soil bulk density at 60cm depth	0.027***	0.014***		
	(0.002)	(0.002)		
Sand content at 60 cm depth (% (kg / kg))	-0.042***	-0.020***		
	(0.005)	(0.004)		
Great Group: Haplustolls (Mollisols)	0.488***	0.123*		
	(0.103)	(0.069)		
Great Group: Tropudults (Ultisols)	-1.026***	-0.799***		
	(0.174)	(0.167)		
Great Group: Chromusterts (Vertisols)	0.678***	0.430***		
	(0.105)	(0.085)		
Sand content % 200cm depth median			0.046***	0.042***
			(0.013)	(0.010)
Soil water content % at 33kPa 200cm depth median			-0.047***	-0.032***
			(0.009)	(0.008)
Great Group: Haplustalfs (Alfisols)			0.372***	0.212**
			(0.103)	(0.085)
N	3,502	3,502	1,844	1,844
N Clusters	1,328	1,328	209	209
Adj. R^2	0.548	0.734	0.614	0.719
Adj. R^2 (Within)	0.405	0.650	0.442	0.594
Regression F-Stat	83.4	76.6	39.3	70.4
City FE	Yes	Yes	Yes	Yes
$\mathbf{C}_{2v}$ Controls	Yes	Yes	Yes	Yes
$\mathbf{X}_v$ Controls	No	Yes	No	Yes

Notes: This table reports estimates of a community-level version of equation (5), the relationship between log population density (the dependent variable) and different soil characteristics variables. We use post-double-selection lasso regressions, following Belloni et al. (2012), to select instruments in these regressions from a set of 67 soil characteristics. Columns 1 and 2 are limited to the sample of urban communities in Indonesia covered by the 2010 *Susenas* epoch, while columns 3 and 4 are limited to urban communities in the Philippines covered by the 2018 FIES. All regressions include controls for city fixed effects, elevation, ruggedness, distance to the nearest point on the coast, and distance to the nearest river. In columns 2 and 4, we add village-level controls for sorting on observables and unobservables, denoted by $\mathbf{X}_v$. Robust standard errors, clustered at the subdistrict-level in columns 1-2 and clustered at the municipality level in columns 3-4, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels. A version of this table at the household level can be found in Table 3.

Table A.7: Density and Electricity Use: Indiv. and Exog. Community Controls, Indonesia

	OLS	IV-Lasso
Panel A: Indonesia	**(1)**	**(2)**
Log Density (2010)	0.082***	0.099***
	(0.009)	(0.019)
Age	0.033***	0.033***
	(0.002)	(0.002)
Age^2	-0.000***	-0.000***
	(0.000)	(0.000)
Female (0 1)	0.018	0.021
	(0.018)	(0.018)
not completed high school	0.279***	0.272***
	(0.018)	(0.018)
completed high school	0.549***	0.531***
	(0.021)	(0.023)
college and higher education	0.787***	0.768***
	(0.025)	(0.028)
Married (0 1)	0.334***	0.346***
	(0.034)	(0.034)
Divorced (0 1)	0.163***	0.170***
	(0.041)	(0.041)
Widowed (0 1)	0.267***	0.273***
	(0.036)	(0.036)
Non-employed (0 1)	0.052***	0.044***
	(0.013)	(0.013)
Household Size	0.081***	0.081***
	(0.003)	(0.003)
Elevation	-0.000**	-0.000
	(0.000)	(0.000)
Ruggedness	0.043	0.093*
	(0.052)	(0.053)
Distance to the coast	-0.003	-0.004
	(0.014)	(0.015)
Distance to nearest river	-0.002	-0.000
	(0.008)	(0.008)
N	40,207	40,207
N Clusters	1,329	1,329
Adjusted R^2	0.221	0.217
Adjusted R^2 (within)	0.129	0.125
Kleibergen-Paap Wald Rank F Stat		92.079
Under Id. Test (KP Rank LM Stat)		181.466
p-Value		0.000
Sargan-Hansen Test (Overidentification)		11.889
p-Value		0.018
City FE	Yes	Yes

Notes: This table reports the the main density coefficient, the coefficients on the individual-level controls ($\mathbf{X}_i$), and the coefficients on the exogenous community controls ($\mathbf{C}_{2v}$) from the regressions reported in Table 4, Panel A, Columns 1-2. Robust standard errors, clustered at the subdistrict-level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.8: Density and Electricity Use: Indiv. and Exog. Community Controls, Philippines

	OLS	IV-Lasso
Panel B: Philippines	**(1)**	**(2)**
Log Density (2018)	0.057***	0.181***
	(0.010)	(0.033)
Household size	0.081***	0.079***
	(0.002)	(0.002)
Age of HH head	0.029***	0.030***
	(0.002)	(0.002)
Age of HH head (sq)	-0.000***	-0.000***
	(0.000)	(0.000)
HH head is female (1 0)	0.081***	0.075***
	(0.012)	(0.012)
HH head has secondary educ (1 0)	0.349***	0.341***
	(0.019)	(0.018)
HH head has tertiary educ (1 0)	0.766***	0.763***
	(0.040)	(0.038)
HH head is married (1 0)	0.249***	0.255***
	(0.024)	(0.024)
HH head is widowed (1 0)	0.064**	0.067***
	(0.025)	(0.025)
HH head is divorced/separated (1 0)	0.012	0.009
	(0.029)	(0.028)
HH head is self employed (1 0)	0.091***	0.096***
	(0.016)	(0.016)
HH head is an employer (1 0)	0.417***	0.429***
	(0.039)	(0.038)
HH head is not employed (1 0)	0.187***	0.192***
	(0.017)	(0.016)
HH head is employed in agriculture (1 0)	-0.229***	-0.163***
	(0.055)	(0.046)
HH head is employed in services (1 0)	0.101***	0.093***
	(0.014)	(0.015)
Elevation (m)	-0.000	-0.000
	(0.000)	(0.000)
Slope (deg)	-0.013	-0.001
	(0.018)	(0.023)
Ruggedness (m)	0.000	0.001**
	(0.000)	(0.000)
Distance to coast (km)	-0.006***	-0.001
	(0.002)	(0.002)
Distance to river (km)	-0.002	-0.001
	(0.003)	(0.003)
N	41,249	41,249
N Clusters	212	212
Adjusted R^2	0.283	0.283
Adjusted R^2 (within)	0.193	0.193
Kleibergen-Paap Wald Rank F Stat		27.066
Under Id. Test (KP Rank LM Stat)		31.607
p-Value		0.000
Sargan-Hansen Test (Overidentification)		1.641
p-Value		0.440
City FE	Yes	Yes

Notes: This table reports the the main density coefficient, the coefficients on the individual-level controls ($\mathbf{X}_i$), and the coefficients on the exogenous community controls ($\mathbf{C}_{2v}$) from the regressions reported in Table 4, Panel B, Columns 1-2. Robust standard errors, clustered at the municipality-level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.9: Density and Electricity Use: Sorting Controls, Indonesia

	OLS	IV-Lasso
Panel A: Indonesia	**(1)**	**(2)**
Log Density (2010)	0.011	-0.002
	(0.012)	(0.040)
Percent Jawa	-0.019	-0.057
	(0.100)	(0.095)
Percent Sunda	-0.105	-0.163**
	(0.084)	(0.079)
Percent Batak	-0.167	-0.246
	(0.326)	(0.318)
Percent Ethnicities from Nusa Tenggara	-1.928**	-1.884**
	(0.756)	(0.747)
Percent Madura	-0.171	-0.201
	(0.166)	(0.165)
Percent Betawi	0.149	0.100
	(0.123)	(0.118)
Percent Aceh	-1.100	-1.111
	(0.880)	(0.878)
Percent Minangkabau	-0.271	-0.402
	(0.544)	(0.562)
Percent Bugis	-0.976**	-1.013**
	(0.406)	(0.393)
Percent Malay	0.142	0.041
	(0.284)	(0.276)
Percent Ethnicities from South Sumatra	0.114	0.003
	(0.516)	(0.523)
Percent Ethnicities from Banten	-0.052	-0.128
	(0.117)	(0.109)
Percent Banjar	-0.025	-0.077
	(0.639)	(0.635)
Percent Dayak	-0.978	-1.098
	(0.885)	(0.877)
Percent Chinese	-0.738	-0.638
	(0.624)	(0.617)
Percent Ethnicities from Central Sulawesi	-1.118	-1.068
	(1.510)	(1.546)
Percent Ethnicities from Papua	-3.716	-4.221
	(6.168)	(5.983)
Percent Makassar	-0.385	-0.405
	(0.370)	(0.357)
Avg. Age	-0.008	-0.008
	(0.009)	(0.009)
Percent Female	-0.706	-0.659
	(0.885)	(0.873)
Avg. Household Size	-0.000	-0.000*
	(0.000)	(0.000)
Percent Single	-1.000	-0.850
	(1.406)	(1.425)
Percent Married	-0.444	-0.293
	(1.316)	(1.362)
Percent Divorced	-0.572	0.355
	(2.811)	(2.753)
Percent Religion: Islam	10.802**	9.611**
	(4.819)	(4.759)
Percent Religion: Christian	10.551**	9.397**
	(4.846)	(4.792)
Percent Religion: Catholic	11.225**	9.961**
	(4.848)	(4.781)
Percent Religion: Hindu	11.017**	9.794**
	(4.832)	(4.775)
Percent Religion: Buddhist	11.626**	10.291**
	(4.940)	(4.871)
Percent Religion: Confucian	9.284*	7.524
	(5.585)	(5.463)
Percent Unemployed	0.293	0.318
	(0.200)	(0.196)
Percent Self-employed	0.398	0.416
	(0.252)	(0.257)
Percent Employer	0.053	0.059
	(0.307)	(0.302)
Avg. Years of Schooling	0.040***	0.043***
	(0.013)	(0.015)
Percent Ever Migrants	0.119	0.137
	(0.103)	(0.120)
Percent Recent Migrants	0.233	0.186
	(0.201)	(0.234)
Percent Working in Agriculture	-0.301***	-0.321***
	(0.081)	(0.109)
N	40,207	40,207
N Clusters	1,329	1,329
Adjusted R^2	0.232	0.232
Adjusted R^2 (within)	0.142	0.142
Kleibergen-Paap Wald Rank F Stat		26.289
Under Id. Test (KP Rank LM Stat)		84.511
p-Value		0.000
Sargan-Hansen Test (Overidentification)		3.905
p-Value		0.419
City FE	Yes	Yes

Notes: This table displays the coefficients on log density and the community-level sorting controls ($\mathbf{X}_v$) from the regression reported in Table 4, Panel A, Column 3. Robust standard errors, clustered at the subdistrict-level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.10: Density and Electricity Use: Sorting Controls, Philippines

	OLS	IV-Lasso
Panel B: Philippines	**(1)**	**(2)**
Log Density (2018)	0.031***	0.068
	(0.011)	(0.042)
Average family size in the barangay	0.001***	0.001***
	(0.000)	(0.000)
Percent females	0.108	-0.326
	(0.773)	(0.812)
Percent 7-15 years old	-1.191	-1.122
	(0.836)	(0.830)
Percent 16-24 years old	-2.816***	-2.680***
	(0.735)	(0.745)
Percent 25-54 years old	-1.191	-0.738
	(0.745)	(0.847)
Percent 55-64 years old	-2.080**	-1.603*
	(1.036)	(0.951)
Percent 65 years old and up	-1.532	-0.759
	(0.937)	(1.473)
Percent highest grade completed secondary	0.412	0.181
	(0.425)	(0.448)
Percent highest grade completed tertiary	1.833***	1.527***
	(0.304)	(0.500)
Percent single	-91.211	-41.662
	(115.275)	(109.422)
Percent married	-0.905	-0.286
	(0.601)	(1.183)
Percent widowed	-2.365	-3.949
	(1.784)	(2.425)
Percent common-law	-0.243	-0.856
	(0.665)	(0.940)
Percent overseas workers	0.676	1.118
	(0.823)	(0.925)
Percent Roman Catholics	0.408	0.450*
	(0.256)	(0.250)
Percent Muslims	1.075***	1.092***
	(0.368)	(0.330)
Percent Iglesia ni Cristo	0.133	0.074
	(0.393)	(0.368)
Percent other religious affiliations	-0.557	-0.765
	(0.567)	(0.534)
Percent Aglipay	0.184	0.433
	(0.535)	(0.546)
Percent Seventh Day Adventists	1.155**	1.251**
	(0.558)	(0.573)
Percent Tagalog	0.662***	0.611***
	(0.144)	(0.149)
Percent Bisaya	0.362**	0.302*
	(0.155)	(0.163)
Percent Cebuano	0.582**	0.521**
	(0.228)	(0.236)
Percent Ilocano	-0.006	-0.097
	(0.177)	(0.194)
Percent Ilonggo	0.542	0.458
	(0.385)	(0.390)
Percent Bikol	1.127***	1.127**
	(0.433)	(0.443)
Percent Waray	1.719***	1.622***
	(0.526)	(0.527)
Percent Kapampangan	0.851***	0.820***
	(0.201)	(0.211)
Percent Boholano	-0.660	-0.789
	(1.020)	(0.958)
Percent Pangasinan	-0.368	-0.648
	(0.444)	(0.463)
Percent Maguindanao	-0.028	-0.105
	(0.187)	(0.181)
Percent Maranao	-0.321	-0.533
	(0.380)	(0.349)
Percent Tausug	-1.630***	-1.710***
	(0.293)	(0.282)
Percent Capizeno	-6.339	-6.653
	(6.368)	(6.353)
Percent Masbatenon	3.853	4.297*
	(2.589)	(2.602)
Percent Karay-a	76.675	76.446
	(59.334)	(57.366)
Percent Manobo	15.529	14.473
	(11.384)	(10.920)
Percent Subanen	167.223	159.520
	(134.352)	(134.143)
N	41,249	41,249
N Clusters	212	212
Adjusted R^2	0.296	0.296
Adjusted R^2 (within)	0.209	0.208
Kleibergen-Paap Wald Rank F Stat		14.546
Under Id. Test (KP Rank LM Stat)		20.518
p-Value		0.000
Sargan-Hansen Test (Overidentification)		3.280
p-Value		0.194
City FE	Yes	Yes

Notes: This table displays the coefficients on log density and the community-level sorting controls ($\mathbf{X}_v$) from the regression reported in Table 4, Panel B, Column 3. Robust standard errors, clustered at the municipality-level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.11: Summary Statistics on Asset Ownership

	Indonesia (2010)		Philippines (2018)	
Asset	**Mean (Std. Dev)**	**N**	**Mean (Std. Dev)**	**N**
Fridge	0.35 (0.48)	40,920	0.59 (0.49)	47,742
Air Conditioner	0.02 (0.14)	20,705	0.22 (0.41)	47,742
Motorcycle	0.59 (0.49)	40,920	0.24 (0.43)	47,742
Car	0.05 (0.22)	20,705	0.12 (0.32)	47,742
Gas Cylinder (<15kg)	0.09 (0.29)	20,705		
Water Heater	0.02 (0.14)	20,705		

Notes: This table provides summary statistics for different asset ownership outcomes for households in the Indonesia (2010) sample and for households in the Philippines (2018) sample. The Indonesian data are pooled from three *Susenas* rounds in 2010, 2011, and 2012, while the Philippines data come from the 2018 FIES.

Table A.12: Density and Residential Energy Consumption: Extensive Margin

Panel A: Indonesia	**Any Electricity? (1)**	**Any LPG? (2)**	**Any Vehicle Gas? (3)**	**Any Kerosene? (4)**
Log Density (2010)	-0.001	-0.018	-0.012	0.011
	(0.003)	(0.019)	(0.014)	(0.018)
N	40,207	40,207	40,207	40,207
N Clusters	1,329	1,329	1,329	1,329
Kleibergen-Paap Wald Rank F Stat	26.289	26.289	26.289	26.289
Panel B: Philippines	**(1)**	**(2)**	**(3)**	**(4)**
Log Density (2018)	0.002	-0.034*	0.013	-0.002
	(0.005)	(0.019)	(0.024)	(0.011)
N	46,754	46,754	46,754	46,754
N Clusters	212	212	212	212
Kleibergen-Paap Wald Rank F Stat	14.785	14.785	14.785	14.785
City FE	Yes	Yes	Yes	Yes
$\mathbf{X}_i$ Controls	Yes	Yes	Yes	Yes
$\mathbf{C}_{2v}$ Controls	Yes	Yes	Yes	Yes
$\mathbf{X}_v$ Controls	Yes	Yes	Yes	Yes

Notes: This table replicates columns 3, 6, 9, and 12 of Table 5 but replaces the dependent variable to be an indicator for whether the household consumed any quantity of residential energy. The dependent variable is listed in the column header. Panel A reports results for households in Indonesian cities, while Panel B reports results for households in Philippines cities. All columns apply a post-double-selection IV-Lasso estimator, following Belloni et al. (2012), and they include city fixed effects, controls for $\mathbf{X}_i$, $\mathbf{C}_{2v}$, and $\mathbf{X}_v$. Robust standard errors, clustered at the subdistrict-level in Panel A (and the municipality level in Panel B), are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.13: Density and Residential Energy Use, Intensive Margin

Panel A: Indonesia	**Electricity (1)**	**LPG (2)**	**Vehicle Gas (3)**	**Kerosene (4)**
Log Density (2010)	0.011	-0.045*	-0.025	-0.017
	(0.040)	(0.026)	(0.030)	(0.087)
N	40,081	29,289	25,121	6,166
N Clusters	1,322	1,299	1,323	873
Kleibergen-Paap Wald Rank F Stat	26.083	19.274	29.846	25.799
Panel B: Philippines	**(1)**	**(2)**	**(3)**	**(4)**
Log Density (2018)	0.040	0.050**	-0.195	0.119
	(0.035)	(0.021)	(0.148)	(0.075)
N	46,163	39,593	19,846	2,378
N Clusters	212	211	212	154
Kleibergen-Paap Wald Rank F Stat	14.527	17.528	3.108	17.064
City FE	Yes	Yes	Yes	Yes
$\mathbf{X}_i$ Controls	Yes	Yes	Yes	Yes
$\mathbf{C}_{2v}$ Controls	Yes	Yes	Yes	Yes
$\mathbf{X}_v$ Controls	Yes	Yes	Yes	Yes

Notes: This table replicates columns 3, 6, 9, and 12 of Table 5 but replaces the dependent variable to be $\log y_{iv}$ instead of $\log(1 + y_{iv})$. The dependent variable is listed in the column header. Panel A reports results for households in Indonesian cities, while Panel B reports results for households in Philippines cities. All columns apply a post-double-selection IV-Lasso estimator, following Belloni et al. (2012), and they include city fixed effects, controls for $\mathbf{X}_i$, $\mathbf{C}_{2v}$, and $\mathbf{X}_v$. Robust standard errors, clustered at the subdistrict-level in Panel A (and the municipality level in Panel B), are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.14: Density and Residential Energy Use: Dropping Agricultural HHs, Indonesia

	Baseline IV-Lasso	Dropping Households Employed in Agriculture	Dropping Communities with Agricultural Household Share >80%	>60%	>40%	>20%
Panel A: Electricity	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
1. Only Xi and W2v Controls	0.099***	0.090***	0.097***	0.096***	0.101***	0.106***
	(0.019)	(0.020)	(0.019)	(0.020)	(0.019)	(0.021)
2: Adding Xv controls	-0.002	0.006	-0.005	-0.006	0.017	0.020
	(0.040)	(0.039)	(0.039)	(0.039)	(0.038)	(0.042)
N	40,207	32,441	40,125	39,872	38,965	35,870
N Clusters	1,329	1,328	1,324	1,318	1,294	1,201
Kleibergen-Paap Wald Rank *F* Stat	26.289	27.661	26.390	26.208	27.187	23.345
Panel B: LPG	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
1. Only Xi and W2v Controls	0.070***	0.044**	0.069***	0.068***	0.076***	0.064***
	(0.019)	(0.018)	(0.019)	(0.019)	(0.020)	(0.020)
2: Adding Xv controls	-0.052	-0.059	-0.051	-0.057	-0.038	-0.053
	(0.041)	(0.039)	(0.041)	(0.041)	(0.040)	(0.040)
N	40,207	32,441	40,125	39,872	38,965	35,870
N Clusters	1,329	1,328	1,324	1,318	1,294	1,201
Kleibergen-Paap Wald Rank *F* Stat	26.289	27.661	26.390	26.208	27.187	23.345
Panel C: Vehicle Gas Consumption	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
1. Only Xi and W2v Controls	0.027	-0.003	0.027	0.029	0.034	0.054**
	(0.023)	(0.025)	(0.023)	(0.024)	(0.025)	(0.025)
2: Adding Xv controls	-0.050	-0.024	-0.048	-0.049	-0.020	0.029
	(0.049)	(0.050)	(0.049)	(0.049)	(0.049)	(0.052)
N	40,207	32,441	40,125	39,872	38,965	35,870
N Clusters	1,329	1,328	1,324	1,318	1,294	1,201
Kleibergen-Paap Wald Rank *F* Stat	26.289	27.661	26.390	26.208	27.187	23.345
Panel D: Kerosene Consumption	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
1. Only Xi and W2v Controls	0.056***	0.063***	0.056***	0.056***	0.059***	0.067***
	(0.018)	(0.019)	(0.018)	(0.018)	(0.018)	(0.020)
2: Adding Xv controls	0.100**	0.108***	0.097**	0.097**	0.103**	0.108**
	(0.041)	(0.042)	(0.041)	(0.041)	(0.041)	(0.045)
N	40,207	32,441	40,125	39,872	38,965	35,870
N Clusters	1,329	1,328	1,324	1,318	1,294	1,201
Kleibergen-Paap Wald Rank *F* Stat	26.289	27.661	26.390	26.208	27.187	23.345

Notes: Each cell reports the coefficient on log population density from equation (6) where the dependent variable is listed in the panel header. All coefficients are obtained from a post-double-selection IV-Lasso estimator, following Belloni et al. (2012). The first row in each panel includes city fixed effects and controls for $\mathbf{X}_i$ and $\mathbf{C}_{2v}$. The second row in each panel additionally includes $\mathbf{X}_v$ controls. Robust standard errors, clustered at the subdistrict-level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.15: Density and Residential Energy Use: Dropping Agricultural HHs, Philippines

	Electricity		**LPG**		**Vehicle Gas**		**Kerosene**	
	Baseline IV-Lasso (1)	**Dropping Households Employed in Agriculture (2)**	**Baseline IV-Lasso (3)**	**Dropping Households Employed in Agriculture (4)**	**Baseline IV-Lasso (5)**	**Dropping Households Employed in Agriculture (6)**	**Baseline IV-Lasso (7)**	**Dropping Households Employed in Agriculture (8)**
1. Only Xi and W2v Controls	0.181***	0.237***	0.091***	0.098***	0.025	0.030	-0.002	0.000
	(0.033)	(0.060)	(0.029)	(0.029)	(0.033)	(0.036)	(0.002)	(0.002)
2. Adding Xv controls	0.068	0.234	0.007	0.024	0.047	0.058	-0.001	0.001
	(0.042)	(0.145)	(0.037)	(0.037)	(0.055)	(0.058)	(0.003)	(0.005)
N	41,249	39,734	43,687	42,186	42,887	41,390	45,312	43,812
N Clusters	212	212	212	212	212	212	212	212
Kleibergen-Paap Wald Rank F Stat	14.546	5.871	14.117	14.074	13.915	13.859	15.646	6.409

Notes: Each cell reports the coefficient on log population density from equation (6) where the dependent variable is listed in the panel header. All coefficients are obtained from a post-double-selection IV-Lasso estimator, following Belloni et al. (2012). The first row in each panel includes city fixed effects and controls for $\mathbf{X}_i$ and $\mathbf{C}_{2v}$. The second row in each panel additionally includes $\mathbf{X}_v$ controls. Robust standard errors, clustered at the municipality-level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.16: Density and Residential Energy Use: Controlling for Historical Infrastructure, Indonesia

	Baseline		**Infrastructure Controls**							
	OLS	**IV-Lasso**	**IV-Lasso**	**IV-Lasso**	**IV-Lasso**	**IV-Lasso**	**IV-Lasso**	**IV-Lasso**	**IV-Lasso**	**IV-Lasso**
Panel A: Only $\mathbf{X}_i$ controls	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
1. Only $\mathbf{X}_i$ and $\mathbf{W}_{2v}$ Controls	0.008***	0.011***	0.013***	0.013***	0.012***	0.012***	0.012***	0.012***	0.011***	0.012***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)
Panel B: Adding $\mathbf{X}_v$ controls	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**
2. Adding $\mathbf{X}_v$ Controls	0.001	0.003	0.004	0.004	0.003	0.004	0.003	0.004	0.003	0.003
	(0.001)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Education Facilities (1983)	.	.	Yes	.	.	.	.	.	.	Yes
Historical Medical Facilities (1983)	.	.	.	Yes	.	.	.	.	.	Yes
Places of Worship(1983)	.	.	.	.	Yes	.	.	.	.	Yes
Irrigation and PLN (1983)	.	.	.	.	.	Yes	.	.	.	Yes
Agricultural Organizations (1983)	.	.	.	.	.	.	Yes	.	.	Yes
Social Activities (1983)	.	.	.	.	.	.	.	Yes	.	Yes
Distance to Major Road	.	.	.	.	.	.	.	.	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	151,520	151,520	151,520	151,520	151,520	151,520	151,520	151,520	151,260	151,260

Notes: Each cell reports estimates of the average density elasticity of residential carbon emissions for urban households in Indonesia (Panel A). The object estimated is described in equation (8). Column 1 and 2 reproduce the OLS estimates and IV-Lasso estimates from Table 7, Panel A. Notice that the estimates are slightly different due to imperfect merging with the historical infrastructure data. In Panel A, we only control for $\mathbf{X}_i$ and $\mathbf{C}_{2v}$, setting $\mathbf{\Gamma}_1 = 0$. Panel B reports the full, unrestricted model. All regressions are limited to the sample of villages within urban areas in Indonesia and include city-fixed effects. From Column 3-9, we separately includes different groups of historical infrastructure variables, as indicated from the bottom panel. In Column 10, we include all historical infrastructure controls. Robust standard errors, clustered at the subdistrict-level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% level.

Table A.17: Density and Residential Energy Use: Controlling for Historical Infrastructure, Philippines

	Baseline		Infrastructure Controls				
	OLS	**IV-Lasso**	**IV-Lasso**	**IV-Lasso**	**IV-Lasso**	**IV-Lasso**	**IV-Lasso**
Panel A: Only X_i controls	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
1. Only $\mathbf{X}_i$ and $\mathbf{W}_{2v}$ Controls	0.000	0.012***	0.011***	0.011***	0.012***	0.012***	0.011***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Panel B: Adding X_v controls	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
2. Adding $\mathbf{X}_v$ Controls	-0.001*	0.004	0.003	0.003	0.003	0.003	0.003
	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Education Facilities (2000)	.	.	Yes	.	.	.	Yes
Medical Facilities (2000)	.	.	.	Yes	.	.	Yes
Places of Worship(2000)	.	.	.	.	Yes	.	Yes
Access to Major Road (2000)	.	.	.	.	.	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	173,135	173,135	169,873	170,820	170,892	170,112	168,755

Notes: Each cell reports estimates of the average density elasticity of residential carbon emissions for urban households in Indonesia (Panel A). The object estimated is described in equation (8). Column 1 and 2 reproduce the OLS estimates and IV-Lasso estimates from Table 7 (Panel B). Notice that the estimates are slightly different due to imperfect merging with the historical infrastructure data. In Panel A, we only control for $\mathbf{X}_i$ and $\mathbf{C}_{2v}$, setting $\mathbf{\Gamma}_1 = 0$. Panel B reports the full, unrestricted model. All regressions are limited to the sample of villages within urban areas in Indonesia and include city-fixed effects. From Column 3-6, we separately includes different groups of historical infrastructure variables, as indicated from the bottom panel. In Column 7, we include all historical infrastructure controls. Robust standard errors, clustered at the municipality-level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% level.

Table A.18: The Effects of Soil Characteristics on Energy Consumption in Rural Areas

	Electricity	LPG	Vehicle Gas	Kerosene	CO_2
Panel A: Indonesia	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Soil bulk density at 60 cm depth (kg / m3)	-0.001	-0.001	0.001	-0.000	
	(0.001)	(0.001)	(0.001)	(0.001)	
Sand content at 60 cm depth (% (kg / kg))	-0.005	-0.000	0.010**	0.003	
	(0.004)	(0.003)	(0.004)	(0.003)	
Great Group: Haplustolls (Mollisols)	-0.171*	-0.006	-0.021	-0.100	
	(0.096)	(0.052)	(0.089)	(0.079)	
Great Group: Tropudults (Ultisols)	0.006	-0.011	0.040	-0.023	
	(0.029)	(0.023)	(0.030)	(0.026)	
N	30,435	30,435	30,435	30,435	
N Clusters	1,894	1,894	1,894	1,894	
Adj. R^2	0.289	0.341	0.341	0.451	
Adj. R^2 (Within)	0.056	0.066	0.162	0.019	
$H_o : \beta = 0$ (p-value)	0.131	0.801	0.093	0.545	
Kleibergen-Paap Wald Rank F Stat	2.438	2.438	2.438	2.438	
Density Cutoff	300	300	300	300	
Fixed Effects	Kabu	Kabu	Kabu	Kabu	
Panel B: Philippines	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Soil bulk density at 200cm depth	0.001	-0.001	-0.001	-0.000*	
	(0.002)	(0.001)	(0.001)	(0.000)	
Soil water content at 200cm depth	0.009	-0.002	0.003	-0.000	
	(0.006)	(0.002)	(0.003)	(0.001)	
Great Group: Haplustalfs (Alfisols)	0.164	-0.062	-0.089	-0.062	
	(0.177)	(0.038)	(0.267)	(0.040)	
N	37,579	36,718	36,959	34,634	
N Clusters	82	82	82	82	
Adj. R^2	0.332	0.426	0.208	0.235	
Adj. R^2 (Within)	0.227	0.146	0.152	0.040	
$H_o : \beta = 0$ (p-value)	0.345	0.006	0.784	0.139	
Kleibergen-Paap Wald Rank F Stat	0.094	0.085	0.106	0.068	
Density Cutoff	300	300	300	300	
Fixed Effects	Province	Province	Province	Province	

Notes: This table reports reduced-form regression coefficients of the dependent variable (listed in the column headers) on our selected soil characteristics in rural areas with a population density of less than 300 inhabitants per km^2. Panel A focuses on rural areas in Indonesia, while Panel B focuses on rural areas in the Philippines. All columns include city fixed effects and controls for $\mathbf{X}_i$, $\mathbf{C}_{2v}$, and $\mathbf{X}_v$. Robust standard errors, clustered at the subdistrict-level in Panel A and the municipality level in Panel B, are reported in parentheses. The "Ho : $\beta = 0$ (p-value)" row reports the p-value of an F-test for the null hypothesis that the coefficients on the soil characteristics variables are all equal to zero. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.19: Density and Residential Energy Use, Using Depth to Bedrock + Selected Soil IVs

	Electricity			LPG			Vehicle Gas			Kerosene		
	OLS	IV-Lasso		OLS	IV-Lasso		OLS	IV-Lasso		OLS	IV-Lasso	
Panel A: Indonesia	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Log Density (2010)	0.082***	0.099***	0.017	0.086***	0.086***	-0.003	0.002	0.041*	-0.011	0.037***	0.029*	0.035
	(0.009)	(0.018)	(0.035)	(0.009)	(0.018)	(0.037)	(0.012)	(0.022)	(0.043)	(0.008)	(0.016)	(0.036)
N	40,207	40,180	40,180	40,207	40,180	40,180	40,207	40,180	40,180	40,207	40,180	40,180
N Clusters	1,329	1,329	1,329	1,329	1,329	1,329	1,329	1,329	1,329	1,329	1,329	1,329
Kleibergen-Paap Wald Rank F Stat		84.072	25.886		84.072	25.886		84.072	25.886		84.072	25.886
Panel B: Philippines	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Log Density (2018)	0.057***	0.167***	0.055	0.014	0.087***	0.022	-0.106***	-0.015	0.062	-0.000	-0.003**	-0.005*
	(0.010)	(0.030)	(0.042)	(0.009)	(0.023)	(0.036)	(0.013)	(0.029)	(0.045)	(0.000)	(0.002)	(0.003)
N	41,249	41,249	41,249	43,687	43,687	43,687	42,887	42,887	42,887	45,312	45,312	45,312
N Clusters	212	212	212	212	212	212	212	212	212	212	212	212
Kleibergen-Paap Wald Rank F Stat		22.905	12.083		23.193	11.658		23.119	11.462		23.135	8.682
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\mathbf{X}_i$ Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\mathbf{C}_{2v}$ Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\mathbf{X}_v$ Controls	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Each cell reports the coefficient on log population density from equation (6) where the dependent variable is listed in the column header. Panel A reports results for households in Indonesian cities, while Panel B reports results for households in Philippines cities. Columns 1, 4, 7, and 10 report OLS estimates, while the other columns apply a post-double-selection IV-Lasso estimator, following Belloni et al. (2012). All regressions include city fixed effects and controls for $\mathbf{X}_i$ and $\mathbf{C}_{2v}$. Columns 3, 6, 9, and 12 additionally include $\mathbf{X}_v$ sorting controls. Robust standard errors, clustered at the subdistrict-level in Panel A (and the municipality level in Panel B), are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.20: Density and Residential Energy Use: Only Reported Quantities, Indonesia

	Electricity			**LPG**			**Vehicle Gas**			**Kerosene**		
	OLS	**IV-Lasso**		**OLS**	**IV-Lasso**		**OLS**	**IV-Lasso**		**OLS**	**IV-Lasso**	
Panel A: Indonesia	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
Log Density (2010)	0.081***	0.113***	0.006	-0.003	-0.024	-0.097***	-0.016	0.021	-0.026	0.031***	0.021*	0.011
	(0.009)	(0.023)	(0.054)	(0.009)	(0.016)	(0.032)	(0.010)	(0.019)	(0.038)	(0.006)	(0.012)	(0.027)
N	40,612	40,612	40,612	53,845	53,845	53,845	51,073	51,073	51,073	55,045	55,045	55,045
N Clusters	1,329	1,329	1,329	1,405	1,405	1,405	1,405	1,405	1,405	1,405	1,405	1,405
Kleibergen-Paap Wald Rank F Stat		124.229	32.902		81.240	25.042		81.634	25.099		70.164	22.682
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\mathbf{X}_i$ Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\mathbf{C}_{2v}$ Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\mathbf{X}_v$ Controls	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes

Notes: Each cell reports the coefficient on log population density from equation (6) where the dependent variable is listed in the column header. Columns 1, 4, 7, and 10 report OLS estimates, while the other columns apply a post-double-selection IV-Lasso estimator, following Belloni et al. (2012). All regressions include city fixed effects and controls for $\mathbf{X}_i$ and $\mathbf{C}_{2v}$. Columns 3, 6, 9, and 12 additionally include $\mathbf{X}_v$ sorting controls. For the sample, unlike our main Indonesia results (reported in Table 5, Panel A), we drop observations without reported quantities in the consumption expenditures data. Robust standard errors, clustered at the subdistrict-level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

Table A.21: Heterongeneity Analysis: Big City (> 1 million) VS Other Cities

	Electricity				**LPG**				**Vehicle Gas**				**Kerosene**			
	Big Cities		**Other**		**Big Cities**		**Other**		**Big Cities**		**Other**		**Big Cities**		**Other**	
	IV	**IV control $\mathbf{X}_v$**	**IV**	**IV control $\mathbf{X}_v$**	**IV**	**IV control $\mathbf{X}_v$**	**IV**	**IV control $\mathbf{X}_v$**	**IV**	**IV control $\mathbf{X}_v$**	**IV**	**IV control $\mathbf{X}_v$**	**IV**	**IV control $\mathbf{X}_v$**	**IV**	**IV control $\mathbf{X}_v$**
Panel A: Indonesia	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**	**(16)**
Log Density (2010)	0.100***	0.033	0.059**	-0.007	0.046**	-0.037	0.044	-0.011	0.083**	-0.014	-0.031	-0.089	0.008	-0.007	0.136***	0.133**
	(0.024)	(0.046)	(0.030)	(0.042)	(0.021)	(0.042)	(0.031)	(0.049)	(0.033)	(0.061)	(0.032)	(0.058)	(0.016)	(0.024)	(0.035)	(0.056)
N	27,620	27,620	12,587	12,587	27,620	27,620	12,587	12,587	27,620	27,620	12,587	12,587	27,620	27,620	12,587	12,587
N Clusters	874	874	471	471	874	874	471	471	874	874	471	471	874	874	471	471
Kleibergen-Paap Wald Rank F Stat	57.559	19.848	64.885	30.631	57.559	19.848	64.885	30.631	57.559	19.848	64.885	30.631	57.559	19.848	64.885	30.631
Panel B: Philippines	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**	**(13)**	**(14)**	**(15)**	**(16)**
Log Density (2018)	0.072***	0.046	0.122	0.089	0.003	-0.015	0.029	0.011	-0.147***	-0.112**	-0.023	-0.108**	-0.000	0.000	0.001	-0.001
	(0.021)	(0.034)	(0.100)	(0.056)	(0.024)	(0.039)	(0.045)	(0.033)	(0.026)	(0.049)	(0.072)	(0.047)	(0.001)	(0.003)	(0.005)	(0.002)
N	25,565	25,565	15,684	15,684	26,303	26,303	17,384	17,384	25,880	25,880	17,007	17,007	27,411	27,411	17,901	17,901
N Clusters	176	176	36	36	176	176	36	36	176	176	36	36	176	176	36	36
Kleibergen-Paap Wald Rank F Stat	30.970	18.920	19.508	37.087	30.909	19.825	19.513	35.924	37.817	21.672	20.207	37.892	38.968	22.318	19.099	38.898
$\mathbf{X}_v$ Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes

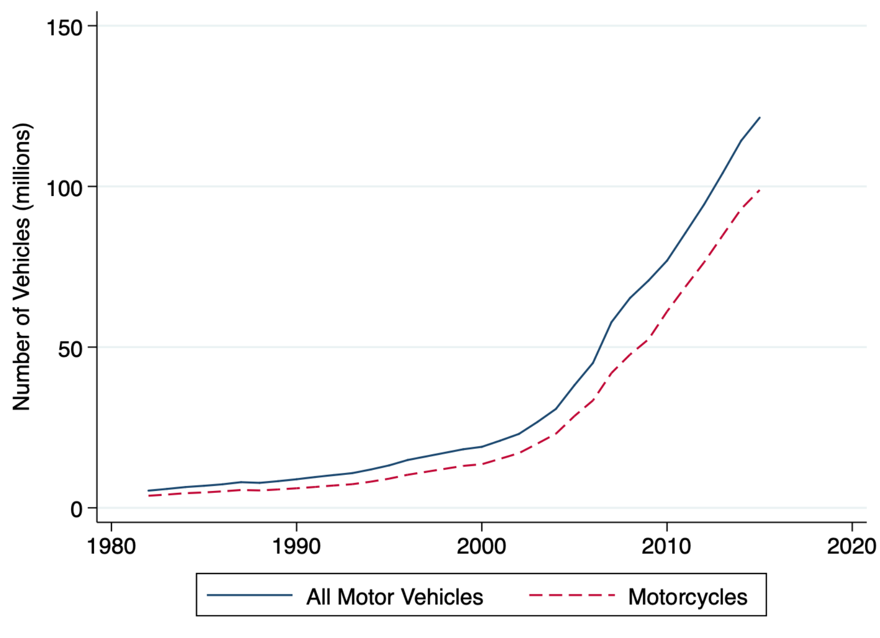
Notes: Each cell reports estimates of the effect of density on different types of residential energy consumption, with the type of energy listed in the column header. All coefficients are obtained from a post-double-selection IV-Lasso estimator, following Belloni et al. (2012). The odd columns only control for city fixed effects, $\mathbf{X}_i$, and $\mathbf{C}_{2v}$, setting $\mathbf{\Gamma}_1 = 0$. The even columns add the $\mathbf{X}_v$ sorting controls. Columns 1-2, 5-6, 9-10, and 13-14, restrict the sample to urban areas with populations greater than 1 million, while the other columns restrict the sample to smaller urban areas. Robust standard errors, clustered at the municipality-level, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% level.

Table A.22: The Average Density Elasticity of Residential Carbon Emissions: Previous Waves

	OLS	IV-Lasso	
Panel A: Indonesia	**(1)**	**(2)**	**(3)**
Log Density (2000)	0.017***	0.020***	-0.002
	(0.002)	(0.005)	(0.009)
N	16,847	16,833	16,833
Panel B: Philippines	**(1)**	**(2)**	**(3)**
Log Density (2010)	0.003**	0.014***	0.005
	(0.001)	(0.003)	(0.004)
N	11,376	11,376	11,376
City FE	Yes	Yes	Yes
$\mathbf{X}_i$ Controls	Yes	Yes	Yes
$\mathbf{C}_{2v}$ Controls	Yes	Yes	Yes
$\mathbf{X}_v$ Controls	No	No	Yes

Notes: This table reports estimates of the average density elasticity of residential carbon emissions for urban households in Indonesia (Panel A) and the Philippines (Panel B). The object estimated is described in equation (8). To obtain this, we estimate the density elasticity for our four carbon intensity outcomes simultaneously, using a SUR system and national carbon intensity weights. All regressions include city fixed effects and controls for $\mathbf{X}_i$ and $\mathbf{C}_{2v}$. Columns 2 and 3 report results from post-double-selection IV-lasso estimators, following Belloni et al. (2012), and column 3 additionally includes $\mathbf{X}_v$ controls. Robust standard errors, clustered at the subdistrict-level in Panel A and the municipality level in Panel B, are reported in parentheses. */**/*** denotes significant at the 10% / 5% / 1% levels.

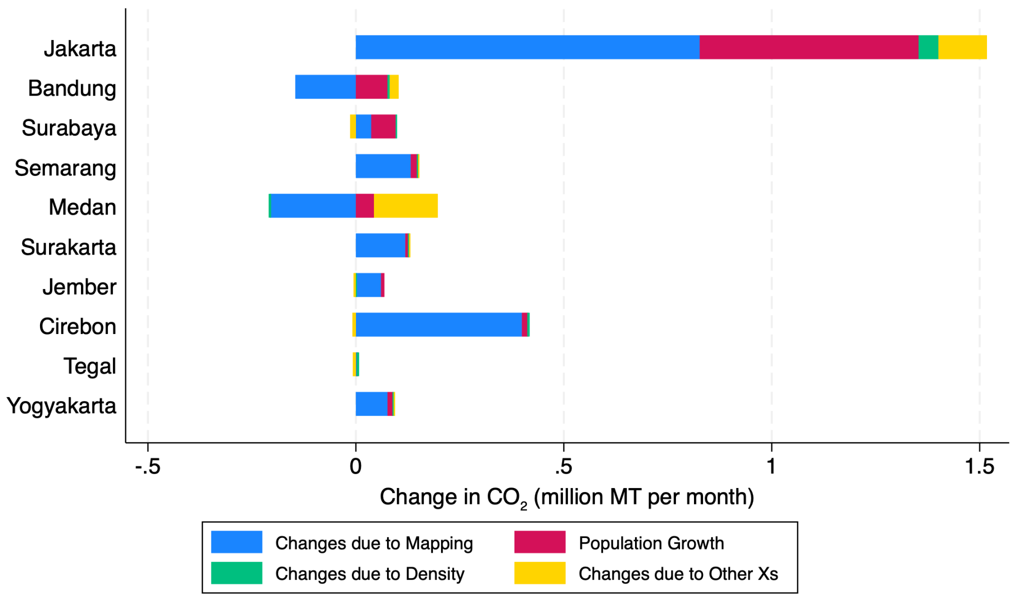
Figure A.1: Number of Registered Vehicles in Indonesia



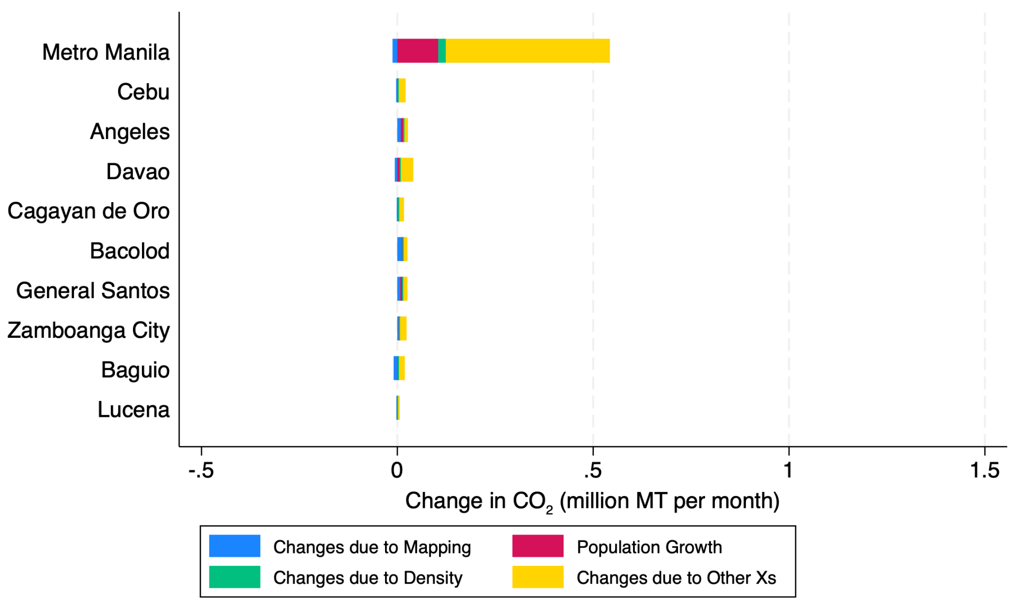
Notes: This figure uses data from various waves of Indonesia's Transportation Statistics (*Statistik Transportasi*), published by BPS, to plot the number of registered vehicles in Indonesia over time. The data compiled by BPS are based on vehicle registrations that come from Indonesia's State Police (*Kepolisian Republik Indonesia*).

Figure A.2: Overall Emissions Growth Decomposition: Results for Top 10 Cities + Density

(A) Indonesia



(B) Philippines



Notes: This figure presents our overall decomposition of changes in predicted emissions by city. This decomposition further separates term B in equation (9) into density and non-density characteristic effects. To do so, we follow equation (B.2) in Appendix B.

B Growth Decomposition

Predicting Total Residential Emissions. Let $c = 1, ..., C$ index cities in our sample, and let $v = 1, ..., N_v$ index communities in city c. Let $s \in \mathcal{S}$ also index types of households. In our analysis, we group households into 3 size groups (1-2 person households, 3-4 person households, and households with more than 4 members) and 4 household-head education groupings (no schooling, some schooling, only completed high school, greater than high school completion). Therefore, there are 12 household types in our analysis (i.e. $|\mathcal{S}|$ = 12).

We can express city c's estimated total residential carbon emissions at time t as follows:

$$\widehat{\text{CE}}_{ct} = \sum_{v=1}^{N_c} \sum_{i \in s} N_{sv,t} \cdot \widehat{\text{CE}}_{sv,t}$$

where $N_{sv,t}$ measures the number of type s households in community v at time t and $\widehat{\text{CE}}_{sv,t}$ is the predicted average residential carbon emissions for a type s household in community v at time t. From equation (7), we can write:

$$\widehat{\text{CE}}_{sv,t} = \sum_{k} w_v(k) \cdot \widehat{y}_{sv,t}(k)$$

where k indexes residential energy outcomes, $w_v(k)$ are carbon intensity weights, and $\widehat{y}_{sv,t}(k)$ are the predicted *levels* of consumption for household type s in community v at time t.

Our main regression model, shown in equation (6), is a log linear regression. Abusing notation, let $\mathbf{x}_{sv,t}$ be a vector of variables in that regression (including city-specific intercepts, individual controls, log density, and community-level controls), and let β be a vector collecting their respective parameters (i.e. $\beta = [\alpha_c, \beta', \theta, \mathbf{\Gamma}_1', \mathbf{\Gamma}_2']'$. We can rewrite equation (6) as:

$$\log\left(1 + y_{sv,t}(k)\right) = \mathbf{x}_{sv,t}'\beta_{k,t} + \varepsilon_{sv,t}(k)$$

where k indexes the type of energy consumption, and t indicates the time period for the cross-sectional regression analysis. We can write:

$$\begin{aligned} \log\left(1 + y_{sv,t}(k)\right) &= \mathbf{x}_{sv,t}'\beta_k + \varepsilon_{sv,t}(k) \\ \implies \exp\left\{\log\left(y_{sv,t}(k) + 1\right)\right\} &= \exp\left\{\mathbf{x}_{sv,t}'\beta_k\right\}\exp\left\{\varepsilon_{sv,t}(k)\right\} \\ \implies y_{sv,t}(k) &= \exp\left\{\mathbf{x}_{sv,t}'\beta_{k,t}\right\}\exp\left\{\varepsilon_{sv,t}(k)\right\} - 1 \end{aligned}$$

So we can use our regression results, which provide estimates of β_k, to calculate a "smearing estimate" of the predicted levels of y (Duan, 1983):

$$\begin{aligned} \widehat{y}_{sv,t}(k) &= \exp\left\{\mathbf{x}_{sv,t}'\widehat{\beta}_{k,t}\right\}\left(\frac{1}{N}\sum_{i=1}^{N}\exp\left\{\mathbf{e}_{sv,t}(k)\right\}\right) - 1 \\ &= \exp\left\{\mathbf{x}_{sv,t}'\widehat{\beta}_{k,t}\right\}\widehat{\mathbf{s}}_t - 1 \end{aligned}$$

where $\mathbf{e}_{sv,t}(k) = \log\left(y_{sv,t}(k) + 1\right) - \mathbf{x}_{sv,t}'\widehat{\beta}_{k,t}$ is the residual of the log linear regression for outcome k for a type s household in community v, and $\widehat{\mathbf{s}}_t$ is defined as the average of the exponentiated residuals:

$$\widehat{\mathbf{s}}_t \equiv \frac{1}{N}\sum_{i=1}^{N}\exp\left\{\mathbf{e}_{sv,t}(k)\right\}$$

Therefore, we can express our estimate of average residential carbon emissions for a type s household in community v at time t is given by:

$$\begin{aligned}
\widehat{\mathrm{CE}}_{sv,t} &= \sum_k w_v(k) \cdot \widehat{y}_{sv,t}(k) \\
&= \sum_k w_v(k) \cdot \left[\exp\left\{ \mathbf{x}'_{sv,t} \widehat{\beta}_{k,t} \right\} \left(\frac{1}{N} \sum_{i=1}^{N} \exp\left\{ \mathbf{e}_{sv,t}(k) \right\} \right) - 1 \right] \\
&= \sum_k w_v(k) \cdot \left[\exp\left\{ \mathbf{x}'_{sv,t} \widehat{\beta}_{k,t} \right\} \widehat{\mathbf{s}}_t - 1 \right]
\end{aligned}$$

We calculate this expression for each community v, household type s, and census wave t. To do so, we combine our parameter estimates with community-specific, household-type averages of the characteristics included in $\mathbf{x}_{sv,t}$. These average characteristic variables are calculated directly from census data, and we combine them with the same geographic and community-level variables that we include in our main regressions.

Growth Decomposition. The growth in predicted total residential carbon emissions from t to $t+1$ in city c is given by:

$$\Delta\widehat{\mathrm{CE}}_c \equiv \widehat{\mathrm{CE}}_{c,t+1} - \widehat{\mathrm{CE}}_{c,t} = \sum_{v=1}^{N_c} \sum_{s \in \mathcal{S}} \left(N_{sv,t+1} \cdot \widehat{\mathrm{CE}}_{sv,t+1} - N_{sv,t} \cdot \widehat{\mathrm{CE}}_{sv,t} \right)$$

Adding and subtracting $N_{sv,t+1} \cdot \widehat{\mathrm{CE}}_{sv,t}$ and rearranging, it is straightforward to show that ΔCE_c can be written as the sum of two terms:

$$\Delta\widehat{\mathrm{CE}}_c = \underbrace{\sum_{v=1}^{N_c} \sum_{s \in \mathcal{S}} \left(N_{sv,t+1} - N_{sv,t} \right) \cdot \widehat{\mathrm{CE}}_{sv,t}}_{\textbf{(I)}} + \underbrace{\sum_{v=1}^{N_c} \sum_{s \in \mathcal{S}} \left(\widehat{\mathrm{CE}}_{sv,t+1} - \widehat{\mathrm{CE}}_{sv,t} \right) \cdot N_{sv,t+1}}_{\textbf{(II)}} \tag{B.1}$$

This equation decomposes the overall growth in predicted carbon emissions for city c into two terms. Term I reflects growth in the population of type s households in community v, multiplied by the predicted levels of carbon emissions in time t for those types of households in those communities. Changes in predicted emissions growth in term II owe to population growth or changes in the composition of households in community v. Term II instead reflects growth in average carbon emissions for type s households.

We can make more progress on Term II by noting:

$$\begin{aligned}
\left(\widehat{\mathrm{CE}}_{sv,t+1} - \widehat{\mathrm{CE}}_{sv,t} \right) &= \sum_k w_{v,t}(k) \cdot \widehat{y}_{sv,t+1}(k) - \sum_k w_v(k) \cdot \widehat{y}_{sv,t}(k) \\
&= \sum_k w_v(k) \left[\widehat{y}_{sv,t+1}(k) - \widehat{y}_{sv,t}(k) \right] \\
&= \sum_k w_v(k) \left[\left(\exp\left\{ \mathbf{x}'_{sv,t+1} \widehat{\beta}_{k,t+1} \right\} \widehat{\mathbf{s}}_{t+1} - 1 \right) - \left(\exp\left\{ \mathbf{x}'_{sv,t} \widehat{\beta}_{k,t} \right\} \widehat{\mathbf{s}}_t - 1 \right) \right] \\
&= \sum_k w_v(k) \left[\exp\left\{ \mathbf{x}'_{sv,t+1} \widehat{\beta}_{k,t+1} \right\} \widehat{\mathbf{s}}_{t+1} - \exp\left\{ \mathbf{x}'_{sv,t} \widehat{\beta}_{k,t} \right\} \widehat{\mathbf{s}}_t \right]
\end{aligned}$$

Focusing on the $\widehat{y}_{sv,t+1}(k) - \widehat{y}_{sv,t}(k)$ terms, and adding and subtracting $\exp\left\{ \mathbf{x}'_{sv,t+1} \widehat{\beta}_{k,t} \right\} \widehat{\mathbf{s}}_t$, we obtain:

$$\begin{aligned}
\Delta\widehat{y}_{sv}(k) &\equiv \widehat{y}_{sv,t+1}(k) - \widehat{y}_{sv,t}(k) \\
&= \exp\left\{ \mathbf{x}'_{sv,t+1} \widehat{\beta}_{k,t+1} \right\} \widehat{\mathbf{s}}_{t+1} - \exp\left\{ \mathbf{x}'_{sv,t} \widehat{\beta}_{k,t} \right\} \widehat{\mathbf{s}}_t \\
&= \left(\exp\left\{ \mathbf{x}'_{sv,t+1} \widehat{\beta}_{k,t+1} \right\} \widehat{\mathbf{s}}_{t+1} - \exp\left\{ \mathbf{x}'_{sv,t+1} \widehat{\beta}_{k,t} \right\} \widehat{\mathbf{s}}_t \right) + \left(\exp\left\{ \mathbf{x}'_{sv,t+1} \widehat{\beta}_{k,t} \right\} \widehat{\mathbf{s}}_t - \exp\left\{ \mathbf{x}'_{sv,t} \widehat{\beta}_{k,t} \right\} \widehat{\mathbf{s}}_t \right)
\end{aligned}$$

$$= \underbrace{\left(\exp\left\{\mathbf{x}'_{sv,t+1}\widehat{\beta}_{k,t+1}\right\}\widehat{\mathbf{s}}_{t+1} - \exp\left\{\mathbf{x}'_{sv,t+1}\widehat{\beta}_{k,t}\right\}\widehat{\mathbf{s}}_{t}\right)}_{(*)} + \underbrace{\left(\exp\left\{\mathbf{x}'_{sv,t+1}\widehat{\beta}_{k,t}\right\} - \exp\left\{\mathbf{x}'_{sv,t}\widehat{\beta}_{k,t}\right\}\right)\widehat{\mathbf{s}}_{t}}_{(**)}$$

Using the $(*)$ and $(**)$ terms, we define the following:

$$\Delta\widehat{\mathrm{CE}}_{sv}\left(\Delta\mathbf{x}\right) \equiv \sum_{k} w_v(k)\left(\exp\left\{\mathbf{x}'_{sv,t+1}\widehat{\beta}_{k,t}\right\} - \exp\left\{\mathbf{x}'_{sv,t}\widehat{\beta}_{k,t}\right\}\right)\widehat{\mathbf{s}}_{t}$$
$$\Delta\widehat{\mathrm{CE}}_{sv}\left(\Delta\beta\right) \equiv \sum_{k} w_v(k)\left(\exp\left\{\mathbf{x}'_{sv,t+1}\widehat{\beta}_{k,t+1}\right\}\widehat{\mathbf{s}}_{t+1} - \exp\left\{\mathbf{x}'_{sv,t+1}\widehat{\beta}_{k,t}\right\}\widehat{\mathbf{s}}_{t}\right)$$

Here, $\Delta\widehat{\mathrm{CE}}_{sv}\left(\Delta\mathbf{x}\right)$ denotes the change in predicted residential carbon emissions for a type s household in community v that owe to changes in the $\mathbf{x}$ characteristics. The expression $\Delta\widehat{\mathrm{CE}}_{sv}\left(\Delta\beta\right)$ reflects changes in emissions that owe to changes in the mapping between characteristics and outcomes. Using these definitions, it is straightforward to show that:

$$\begin{aligned}
\Delta\widehat{\mathrm{CE}}_{c} &= \sum_{v=1}^{N_c}\sum_{s\in\mathcal{S}}\left(N_{sv,t+1} - N_{sv,t}\right)\cdot\widehat{\mathrm{CE}}_{sv,t} + \sum_{v=1}^{N_c}\sum_{s\in\mathcal{S}}\left(\widehat{\mathrm{CE}}_{sv,t+1} - \widehat{\mathrm{CE}}_{sv,t}\right)\cdot N_{sv,t+1} \\
&= \underbrace{\sum_{v=1}^{N_c}\sum_{s\in\mathcal{S}}\left(N_{sv,t+1} - N_{sv,t}\right)\cdot\widehat{\mathrm{CE}}_{sv,t}}_{(\mathbf{A})} + \underbrace{\sum_{v=1}^{N_c}\sum_{s\in\mathcal{S}}\Delta\widehat{\mathrm{CE}}_{sv}\left(\Delta\mathbf{x}\right)\cdot N_{sv,t+1}}_{(\mathbf{B})} + \underbrace{\sum_{v=1}^{N_c}\sum_{s\in\mathcal{S}}\Delta\widehat{\mathrm{CE}}_{sv}\left(\Delta\beta\right)\cdot N_{sv,t+1}}_{(\mathbf{C})}
\end{aligned}$$

This expression shows that we can decompose the growth in carbon emissions per household into two additional terms. Term B reflects growth in emissions due to changes in the $\mathbf{x}_{v,s}$ characteristics over time. Because we are fixing household types in this analysis, one source of these changes will be changes in community-level population density. Term C reflects growth in emissions due to changes in the mapping between $\mathbf{x}_{v,s}$ characteristics and emissions over time.

Using a similar add and subtract trick, we can also split term B into two separate terms:

$$\Delta\widehat{\mathrm{CE}}_{sv}\left(\Delta\text{density}\right) \equiv \sum_{k} w_v(k)\left(\exp\left\{\mathbf{x}'_{sv,t}\widehat{\beta}_{k,t}\right\} - \exp\left\{\mathbf{x}_{sv,t}(\text{density}_{t+1})'\widehat{\beta}_{k,t}\right\}\right)\widehat{\mathbf{s}}_{t}$$
$$\Delta\widehat{\mathrm{CE}}_{sv}\left(\Delta\text{other}\right) \equiv \sum_{k} w_v(k)\left(\exp\left\{\mathbf{x}_{sv,t}(\text{density}_{t+1})'\widehat{\beta}_{k,t+1}\right\} - \exp\left\{\mathbf{x}'_{sv,t+1}\widehat{\beta}_{k,t}\right\}\right)\widehat{\mathbf{s}}_{t}$$

where $\mathbf{x}_{sv,t}(\text{density}_{t+1})$ denotes the vector of $\mathbf{x}_{sv,t}$'s using density_{t+1} in place of density_{t}. It is easy to show that we can write:

$$\underbrace{\sum_{v=1}^{N_c}\sum_{s\in\mathcal{S}}\Delta\widehat{\mathrm{CE}}_{sv}\left(\Delta\mathbf{x}\right)\cdot N_{sv,t+1}}_{(\mathbf{B})} = \sum_{v=1}^{N_c}\sum_{s\in\mathcal{S}}\left[\Delta\widehat{\mathrm{CE}}_{sv}\left(\Delta\text{density}\right) + \Delta\widehat{\mathrm{CE}}_{sv}\left(\Delta\text{other}\right)\right]\cdot N_{sv,t+1} \tag{B.2}$$