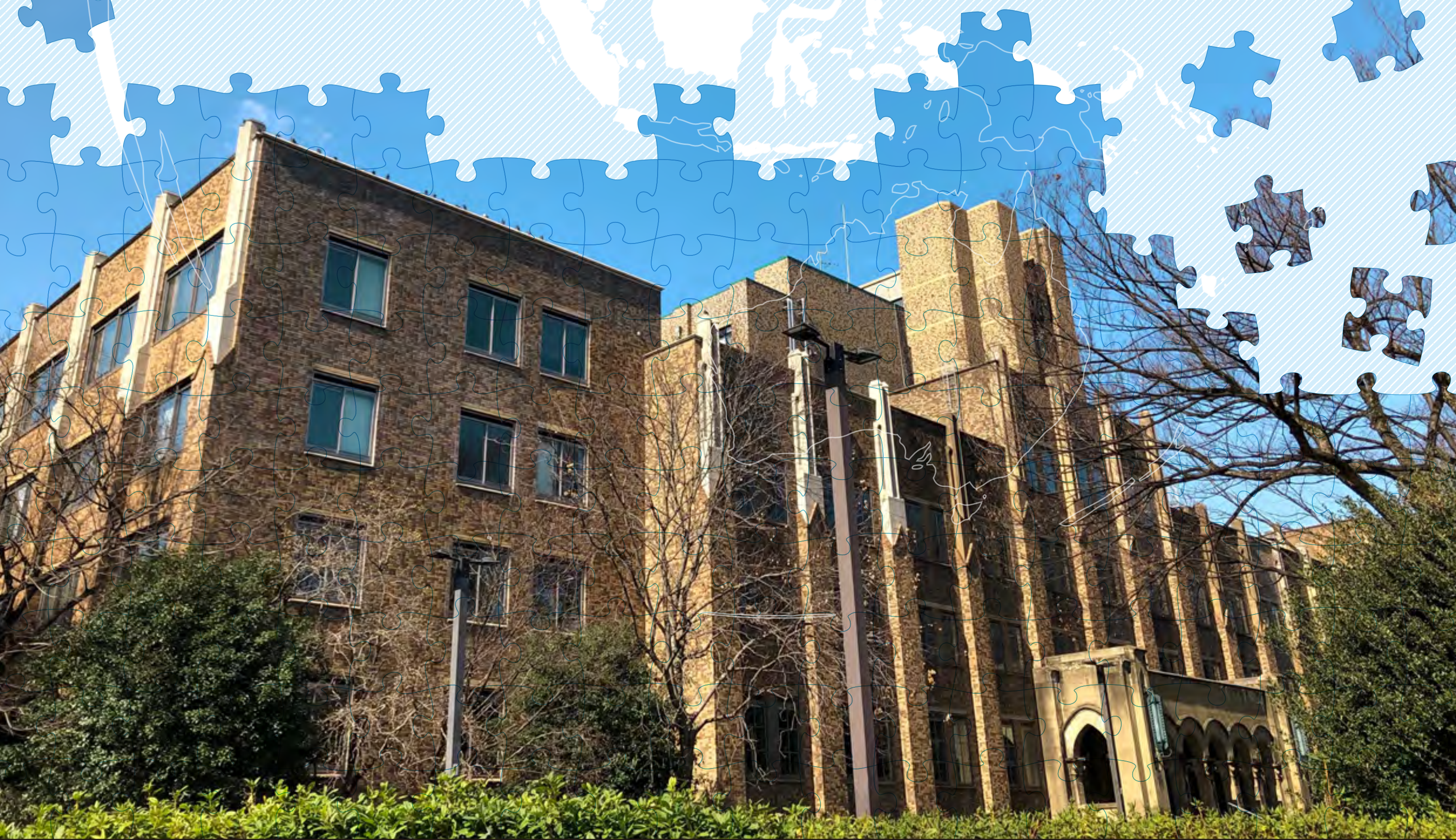
CSRDA
Center for Social Research and Data Archives, Institute of Social Science, The University of Tokyo
SUSTAINABLE DEVELOPMENT GOALS
CSRDA supports the Sustainable Development Goals

CSRDA Discussion Paper

Heat as a Measure of Living Standards: Exploring Social Interaction Benefits in Squatter Communities



No. 10
Date August.2025
SDGs
1 NO POVERTY
3 GOOD HEALTH AND WELL-BEING
10 REDUCED INEQUALITIES

Name
Mariko Nakagawa, Kotaro Iizuka

Heat as a Measure of Living Standards: Exploring Social Interaction Benefits in Squatter Communities*

Mariko Nakagawa† Kotaro Iizuka‡

August 1, 2025

Abstract

Collecting reliable socio-economic data from marginalized populations is difficult when cognitive or addiction-related issues hinder survey responses. This study presents a novel, non-survey approach for assessing current living standards among squatters in contexts where survey responses are unreliable, using winter roof temperatures captured by drone-mounted thermal cameras. Unlike roof footprint, which reflects accumulated wealth, roof temperature indicates present living conditions. Validation thermal experiments confirm a positive association between indoor and roof temperatures. Applied to squatter communities, the method reveals that those in larger clusters have higher roof

*The authors are grateful to Kenzo Asahi, Sam Asher, Donald Davis, Chishio Furukawa, Marco Gonzalez-Navarro, Clément Gorin, Vernon Henderson, Tatsuhito Kono, Melanie Morten, Kentaro Nakajima, Manabu Nose, Masayuki Sawada, Takaaki Takahashi, Kensuke Teshima, Joaquin Urrego, David Weil, Junichi Yamasaki for their thoughtful comments and suggestions. We would also like to thank seminar participants at the workshops at Tokyo, Tohoku, Osaka, Hitotsubashi, Sogang, ADB Conference, 12th European Meeting of UEA, and 17th North American Meeting of UEA. We are grateful to Nairu Kimura for sharing his experience staying in the study area. All remaining errors are our own.

†Institute of Economic Research, Hitotsubashi University and Graduate School of Information Sciences, Tohoku University. E-mail: nakagawa.m@r.hit-u.ac.jp.

‡Center for Spatial Information Science, University of Tokyo. E-mail: kiizuka@csis.u-tokyo.ac.jp.

temperatures, suggesting social interaction benefits. This approach offers a new tool for studying hard-to-reach populations.

1 Introduction

Obtaining socioeconomic information about vulnerable populations is crucial, as they are often difficult to reach and therefore excluded from social welfare services, despite facing severe hardships in their daily lives. Ensuring adequate support for these marginalized groups is important not only from a humanitarian perspective but also for strengthening the broader social security system. Homeless individuals, for example, are one such vulnerable group that has drawn public attention. However, collecting reliable quantitative socioeconomic data—such as their living standards or income—from this population is known to be challenging. This is largely due to the high prevalence of alcoholism, illegal drug use, and intellectual disabilities among the homeless in many countries. In Western Europe and the United States, studies consistently report elevated rates of mental illness, substance abuse, and cognitive impairments among homeless populations (Philippot et al., 2007; Early, 2005; Oakes and Davies, 2008; Van Straaten et al., 2014).

The squatter communities in suburban Tokyo, the focus of this study, are no exception to such challenges of surveying vulnerable populations.[1] These residents—classified in Japan as "settled homeless persons" who build informal cage-like houses on public land[2]—commonly suffer from mental health issues such as hallucinations, delusions, and cognitive impairments (Ito et al., 2014; Morikawa et al., 2011; Nishio et al., 2017; Okamura et al., 2014, 2015). Alcoholism is also widespread (Morikawa et al., 2011; Takano et al., 1999).

[1] In addition to the challenges of obtaining reliable quantitative information from settled homeless individuals, similar difficulties arise when collecting data in informal settlements or slums. Gathering economic data in these areas is particularly challenging due to safety risks for fieldworkers (Marx et al., 2013), as well as the high costs, time demands, and logistical complexities involved in conducting in-person surveys (Huang et al., 2021).

[2] The population we study represents a mix of urban squatters (in a developed country context) and homeless individuals. Hereafter, we use the terms "(settled) homeless persons" and "(urban) squatters" interchangeably, except where doing so might cause confusion for the reader.

In addition, studies have found the prevalence of intellectual disabilities in this population (Nishio et al., 2017; Okamura et al., 2015). For instance, Morikawa et al. (2011) reported difficulties interviewing individuals who were intoxicated or hallucinating. Salize et al. (2002) further highlights a selection bias: interviews to vulnerable populations often only reach those who use shelters or attend free meal services.

To overcome these challenges, we introduce a novel, non-survey-based method for measuring living standards,[3] utilizing unmanned aerial system (usually called "drone"). Drawing on the unique characteristics of informal housing among settled homeless individuals, we use winter roof temperatures—captured by drone-mounted infrared cameras—as a proxy for the household current living standards. The idea is straightforward: homes with more heat-generating appliances, often associated with higher living standard, emit more heat through their roofs.[4] This approach is particularly feasible in our study area. As described in Section 2, many homeless individuals in our study area possess heaters, generators, and gas appliances (Murata, 2015). Their homes are lightweight and well-ventilated to prevent carbon monoxide poisoning, allowing indoor heat to escape through the roof. Using a drone, we collected high-resolution visual and thermal imagery, identified informal houses from daytime images, and measured each roof's nighttime tempera-

[3]We do not deny the value of research based on interviews. However, the main reason we adopt a non-survey-based approach is that numerical data from surveys can be unreliable in contexts where respondents' answers are inconsistent or inaccurate, making them unsuitable for formal quantitative analysis. For this reason, we choose not to validate our roof thermal data—collected under uniform and objective conditions by researchers—against potentially unreliable survey-based figures. Instead, we prefer to validate our thermal data through a thermal experiment, as described in Section 3.

[4]While not focused on homelessness or informal settlements, the Santamouris et al.'s (2014) winter survey of low-income households in Athens found that higher thermal and electricity spending was linked to higher indoor temperatures. It also showed that households with higher annual incomes consumed more energy for heating, resulting in warmer homes—reflecting their greater financial capacity to afford energy. Similar results were observed in a nationwide Japanese survey: lower-income households tended to have colder living rooms due to limited ability to pay for heating (Umishio et al., 2020). Studies like Huebner et al. (2016) also show that owning and using more appliances significantly increases electricity consumption. Taken together, these findings suggest that higher income allows households to purchase and use more appliances, driving up energy use and indoor temperatures.

ture using the overlaid thermal data.

To validate this approach, we conducted a thermal experiment by temporarily constructing a replica of a typical house in a private drone field. We recorded both indoor and nearby outdoor temperatures, and measured the roof temperature using the same drone-based method. The results confirmed that roof temperature closely reflects indoor temperature variation and that heat radiation is localized, supporting the reliability of this data at the individual house level.

We also showed that roof temperature captures a different economic dimension than roof footprint—a proxy for the total amount of wealth proposed by Huang et al. (2021). In our dataset, roof temperature (adjusted for outside temperature) and roof footprint have almost no correlation, indicating they reflect different aspects of economic status. While footprint represents accumulated wealth in the past, roof temperature better indicates current living standards and earnings. This distinction becomes even clearer when analyzing oldcomer versus newcomer houses identified from past aerial photos. Therefore, our method is especially useful for assessing present economic activity rather than accumulated wealth.

Finally, we applied this method to examine whether settled homeless individuals (squatters) gain benefits from social interactions within their communities. Homeless individuals in the area typically earn small, unstable incomes through informal or illicit work, relying on social networks for job-related information from nearby homeless persons, and mutual support among them (See Section 2). Geographic proximity facilitates these interactions, making social networks crucial for accessing resources and improving daily life stability.

The study uses hierarchical clustering based on house proximity to group homeless persons into clusters, assuming social interactions occur within these clusters. Cluster size is used as a proxy for the size of social networks. Results show that larger clusters—indicating larger social networks—are associated with higher present living standards, as measured by roof temperature. To address potential endogeneity, the study employs an instrumental variable: a dummy indicating whether a house is located on the inner curve of a meandering river—areas that typically have flatter terrain and are more suitable for construction. Even with the instrumental variable estimation, the positive association between cluster size and roof temperature remains significant, reinforcing the role of social interaction networks in improving material welfare.[5] In addition, we compare roof footprint and temperature in this application to confirm that these two proxies reflect distinct dimensions of economic status.

To our knowledge, this is the first study to assess individual living standards using drone-collected daytime imagery and nighttime thermal data, offering a new tool for economic research on hard-to-reach, vulnerable populations—a critical step toward addressing broader societal challenges. Our method is best suited for studying populations who remain in fixed locations, live in dwellings lacking sufficient thermal insulation, and are difficult to survey reliably due to factors such as substance abuse, addiction, alcoholism, or cognitive and mental health challenges. One immediate application is investigating unhoused individuals who semi-permanently reside in tents on the street or in homeless encampments.[6]

This paper also contributes to the growing body of literature that leverages geospatial alternative data for economic analysis. Since the pioneering use of nighttime light intensity as a proxy for GDP (Chen and Nordhaus, 2011; Henderson et al., 2012), spatial imagery data have become increasingly common in economics research. Notable applications of nightlight data in the social sciences include analyses of regional inequality within countries (Lee, 2018), panel of city-level economic activity (Storeygard, 2016), regional favoritism based on the nightlight intensity of political leaders' birthplaces (Hodler and Raschky, 2014), and the association between ethnic political centralization and regional development using nightlight data from ethnic

[5] The study cannot use individual fixed effects to control for unobservable characteristics (e.g., ability, intellect, physical condition). However, a bias from these unobservables likely skews results downward, as individuals with lower abilities may cluster together for support (negative sorting). Therefore, the true positive effect of cluster size on current living standards may be even stronger than estimated (See Section 4.6).

[6] One important consideration when applying our thermal method to tents on the street or in homeless encampments is the need to assess heat radiation between neighboring tents. This can be done through thermal experiments, as described in Section 3.

homelands (Michalopoulos and Papaioannou, 2013).

High-resolution aerial imagery also has been increasingly applied to economic analysis in various ways. For example, Henderson et al. (2021) examine the dynamics of slum growth and (in)formal land use. Marx et al. (2019) use the brightness of corrugated iron roofs in high-resolution satellite images as a proxy for housing quality to examine how ethnic patronage affects rental prices and housing investment. Go et al. (2022) use convolutional neural networks to detect vehicles and study the economic impact of airport openings. Gendron-Carrier et al. (2022) investigate whether subway openings influence air pollution levels, using remotely sensed measure of particulates. In another example, Baragwanath et al. (2021) use daytime satellite imagery as well as nightlight intensity to detect urban market locations.

The noncontact sensing method used in this study—thermal imaging captured by drone-mounted infrared cameras—has only recently been introduced in the natural sciences. Applying this technique to the social sciences expands the possibilities for accessing types of information that have long been difficult for researchers to obtain. Using thermal data to measure current living standards, along with the adoption of interdisciplinary data collection methods, may usher in a new era of social science research with broader and more innovative perspectives.

The remainder of this paper is organized as follows. Section 2 introduces the study area and explains the advantage of our choice of drone-based data collection. Section 3 demonstrates the data validity by presenting the characteristics of the thermal data collected in the replication experiment. It also compares roof temperature and footprint, highlighting how each captures different economic aspects. Section 4 explores whether settled homeless persons benefit from social interaction within close geographic proximity, using roof temperature as a proxy for current living standards. Finally, Section 5 concludes the paper.

2 Study Area and Drone Data Collection

The area we focus on in this study is a riverbank in a suburban part of Tokyo, where clusters of settled homeless communities have formed. Although these informal settlements are the subject of our research, there is no straightforward way to identify their exact residential locations or assess their living standards. In this section, we discuss the challenges involved in obtaining such information and explain how we addressed or navigated these obstacles. Specifically, we highlight why this area is well suited for our data collection method, which combines high-resolution daytime aerial photographs with thermal imagery captured by an infrared camera mounted on a drone.

Sections 2.1–2.3 begin with an overview of the background of the homeless persons in the study area, followed by an explanation of why our data collection approach is appropriate for these circumstances. Section 2.4 provides a brief summary of the data gathered through drone flights.

2.1 Aerial Imagery and House Locations

The homeless persons in our study area, often categorized as settled homeless persons, build cage-like houses in which they illegally occupy public land. Some unique features of the exteriors of these houses and sociodemographic characteristics enable us to approximate the house locations.

Unlike nomadic young homeless persons, whose temporary locations can be difficult to detect, some older homeless persons, mainly single males without families, settle in stations, city parks, or riverbanks (DHWKC, 2019). In 2018, 49.7% of homeless men in the ward where our study area is located were settled in cage-like houses that they had built on the river banks in our study area (DHWKC, 2019). Many of the current homeless persons in Japan served as construction workers or were engaged in the manufacturing sector during the era of high economic growth, and such workers lost their jobs after the industrial structural shift at the end of Japan's high economic growth period after WWII (DHWKC, 2019; Takano et al., 1999). These ex-construction workers may well have the ability to build simple-structure houses if they find an affordable location. The riverbank is evidently such a site to build houses. Moreover, most of the homeless persons in our study area are single old men (DHWKC, 2019).[7] It is therefore reasonable to assume that each house is occupied by a single resident, allowing us to use the number of houses as a proxy for the number of homeless individuals. This implies that our unit of observation is at the individual (person) level.

To obtain the location information of the cage-like houses, we captured aerial images with coordinate and height information via drone and detected house-like objects appearing in the high-resolution aerial pictures. The typical exterior characteristics for the houses built by homeless persons in the study area enable us to detect what objects in the images are houses (as displayed in Online Appendix Figure A1). We detected house objects by selecting polygons that satisfy typical features of the cage-like houses in our study area, as will be more detailed in Section 2.4.

2.2 Thermal Imagery and Living Standards

We describe how homeless persons earn money and what they own. As for their earnings, one main income source for homeless persons in this area is collecting aluminum cans or used metal found in garbage disposal locations in non-homeless residential areas.[8] Selling collected cans to scrap firms illegally is one way that homeless persons can earn money.[9][10]

Turning to what they have in houses, several homeless persons in our

[7]According to the DHWKC (2019), the average age of homeless persons in the ward where our study area lies was 62.3 years: 47.5% were in their sixties and 24.9% were older than 70 years in 2018.

[8]According to the interview with the municipality division of the medial welfare services in our study area, approximately 15,000 garbage collection stations are evenly distributed near this area.

[9]In most administrative divisions in Japan, local ordinances forbid the collection or transportation of resources disposed in garbage collection areas, except by business firms entrusted by the local government.

[10]It is rare for homeless persons in Japan to beg for money or food from passersby (Okamoto, 2007). In the neighborhoods of our study area, approximately 70% earn money independently. On average, they earn approximately $ 350–400 US per month. Among homeless earners, 82% obtain income by collecting and selling cans to illegal scrap firms and 11.5% work as daily construction workers (DHWKC, 2019). Homeless persons in this area rarely beg for money or food and earn money on their own (DHWKC, 2019; Okamoto, 2007).

study area have electronic generators, electromechanical devices, and gas bottles, as observed in Murata (2015). Because using these products generates heat, the roof temperature of an individual who owns multiple electromechanical devices should be higher than that of a person who does not have such heat-producing devices. Thus, our data collection method is based on an idea that roof surface temperature sensed from an infrared camera can be a proxy for their present living standards. We use nighttime roof temperature in winter as the measurement of the individual living standards in each house.

Despite this idea in combination with infrared camera and drones, one may cast doubt on the validity of using surface temperature of a roof as a proxy for the living standard of its resident. Specifically, the following two questions about the validity may arise: (i) Does the nighttime roof temperature in winter express the level of resident living standards? (ii) If so, does the surface temperature of a house necessarily capture the inside temperature?

For the first question, we can safely assume that the roof temperature at night during the winter accurately represents the living standard of a homeless resident for the following reasons. The choice of data collection timing at winter evenings is the key to answer this question. As winter is when homeless persons suffer the most from the cold (DHWKC, 2019), they would want to avoid being cold if possible, and indeed, more well-off homeless persons can increase their warmth using heat-generation equipment. By contrast, poor homeless persons may not be able to stay warm, as they lack sufficient heat-generation products.[11][12]

How appropriate the data collection in nighttime is, is related to whether residents are staying in their houses. Usually, homeless persons earn money by collecting used cans and selling them to illegal scrap firms, as already mentioned. In the morning, each Japanese household disposes of used cans

[11] In summer, people prefer to cool down their house, so it would be unclear whether the houses of more well-off homeless persons show higher or lower temperatures; this ambiguity does not occur in winter.

[12] Data collection in winter is suitable also in a sense that not a small number of houses are under trees in the study area. In winter, deciduous trees shed their leaves, allowing high-resolution daytime aerial images to capture houses located beneath the trees through the gaps in the branches, making it possible to identify their locations.

at garbage sites. Homeless persons collect these cans in the morning, which means there is scant possibility they are at home during the daytime. It may not be appropriate to measure roof temperature during the daytime, and we conjecture that the homeless are only consistently in residence in the evening. Thus, data collection at night is preferable, which is why we collected the thermal data at night. Such flexibility in data collection timing is one of the strengths of data collection by drones. Unlike nightlight intensity from satellites whose data collection timing is ready-made and inflexible, drone data collection is order-made and more flexible.

We discuss the second question about external vs internal heat. Balaras and Argiriou (2002) note that thermal infrared images taken by thermography cameras are useful for noncontact detection of how energy leaks from a building's envelope. As an object's temperature correlates with the infrared radiation it emits, infrared inspection of building envelopes can detect heat losses from walls or roofs. We can measure inside temperature that leaks out through walls or roofs when they are thin. Indeed, the walls and roofs of homeless persons' houses tend to be very thin, because houses are simply constructed of frameworks covered by plastic sheets such that the houses are well ventilated to avoid carbon monoxide (CO) intoxication (Murata, 2015).

Moreover, as discussed in Section 3.1, we conducted a thermal experiment in which we built a facsimile of a typical house from the study area. We measured the indoor temperature and sensed the roof temperature using a thermal infrared camera attached to a drone. The obtained temperature data suggest that a higher roof temperature sensed by drone is associated with a higher indoor temperature. Hence, the roof temperature recorded by drone is reasonably a good proxy for the inside temperature.

2.3 Why Is Survey-Based Data Collection Difficult?

As in many other countries, there is a strong tendency for homeless persons in Japan to have mental disorders or cognitive impairment, such as hallucinations and delusions (Ito et al., 2014; Morikawa et al., 2011; Nishio et al., 2017; Okamura et al., 2014, 2015). Likewise, alcoholism is prevalent among homeless persons, and the morbidity rate of alcohol psychosis or alcohol-dependence syndrome is much higher than the Japanese average (Morikawa et al., 2011; Takano et al., 1999). In addition to the mental illness and alcoholism, there is a problematic prevalence of low IQ, cognitive disorders, and intellectual impairment among homeless persons (Nishio et al., 2017; Okamura et al., 2015). These issues make it particularly difficult to collect reliable survey data from homeless individuals, especially when it comes to obtaining accurate quantitative information. In a study of mental illness among homeless persons in Tokyo conducted by Morikawa et al. (2011), the authors were unable to interview individuals who were drunk or experiencing hallucinations.

Another limitation in interviewing homeless persons is noted by Salize et al. (2002). Even when homeless persons do not suffer from intellectual disability, mental illness, or substance abuse, the interviewers can only easily reach accessible groups of homeless persons such as users of shelters or people who come for free meals; this challenge can induce potential selection bias problems.

These problems would also emerge in our study area if we were to conduct a survey based on questionnaires or interviews of homeless persons. DHWKC (2019) notes that homeless persons in our study area also suffer from alcohol or gambling addictions. Thus, instead of adopting an interview- or questionnaire-based approach, we collected roof temperature data using an infrared thermal camera loaded on a drone to elicit information about living standards. This approach was based on the assumption that roof temperature in the winter can be used as a proxy for the present living standard of residents in houses. The advantage of using a drone loaded with a thermal sensor is that it can collect thermal data by sensing various objects from the sky under nearly uniform conditions.[13] Such uniformity in data collection

[13] Akiyama et al. (2019) report the possibility of detecting vacant houses using drones. The presence of life activity producing waste heat was observed, and vacant houses showed no heat loss from house windows and walls. The authors emphasize that this observation is more valid during winter and at night. During winter, heat from solar radiation resolves faster during the day. Specular reflectance is also avoided at night, allowing the researchers to sense only the heat transfer from human activity.

can circumvent sample selection problems and unreliable information from homeless persons with intellectual and mental disorders.

2.4 Data Collected by Drone

A flight campaign was conducted using a Matrice 210 drone with a thermal Zenmuse-XT2 camera (DJI, Shenzhen, China).[14] The altitude of the drone was 140 m above the ground, and aerial images were continuously taken with the camera facing a nadir along the river edges. The aerial data collection obtained thermal information and location data of houses of homeless persons on one side of the river we study. Flight permission was obtained following aviation laws, and extreme care was taken with the altitude and geographic position of the flights due to the presence of the nearby airport.[15]

The study area is known for homeless persons illegally occupying the land in cage-like, rectangular-shaped houses that are usually covered with blue, green, or gray plastic sheets.[16] Since these house features are most common in the target area, we first detected house locations by viewing drone-collected image data from the daytime on a day in February, 2019 (hereafter, Date F). We extracted house-like objects that have the following features: (i) a rectangular or box shape, (ii) covered with blue, green, or gray plastic sheets, and (iii) sufficient size for a person to reside (e.g., sides longer than approximately 1.5 m and height greater than approximately 1 m).[17][18]

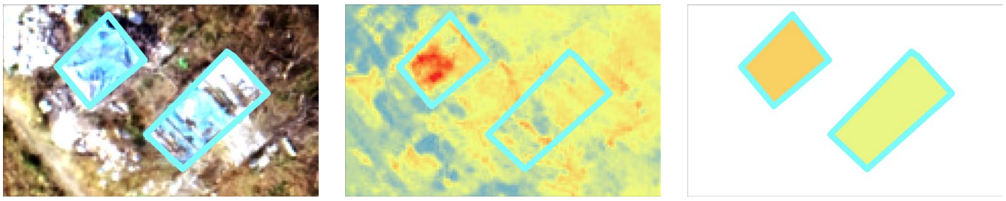
[14]This data collection protocol was approved by the Institutional Review Board (IRB) of Center for Spatial Information Science, University of Tokyo (No. CSIS-RER-2021001).

[15]A limitation that arose in the data collection is that we were unable to collect data from houses underneath bridges to prevent possible accidents caused by crashing into the bridge. In Section 4.8 in an application to the homeless persons' social interaction analysis, we conduct a robustness check by dropping observations near bridges.

[16]Online Appendix Figure A1 shows typical houses of homeless individuals in this area.

[17]The possibility remains that some of the detected houses were vacant. To overcome this issue, we only considered those house polygons that show a variation in roof temperature for the final sample of houses. As will be explained later, we consider roof temperature a sign of human activity, so house polygons without temperature variations, which have a high possibility of having been abandoned, should be excluded from the sample.

[18]We calculate the house height, based on a 3D model reconstructed using Structure-from-Motion (SfM) techniques. We carefully examine the 3D structure and extract a representative height value from the roof top of the house. The determination of the ground level is considered by choosing a representative value from the surrounding terrain,



(a) Sample of Houses (Orthophoto) (b) Sample of Thermal Images (c) Sample of Average Roof Temperatures

Figure 1: House Polygons & Roof Temperatures

Figure 1a: Polygons marked with light blue lines are detected houses on Date F. Figure 1b: Reddish (bluish) area indicates higher (lower) temperatures. Temperature is measured in °C. Data was collected on Date F. Figure 1c: Each polygon is colored by the average temperature within each polygon area. Data was collected on Date F.

We detected 63 house-like objects that satisfy these conditions. Figure 1a displays a sample of the detected house polygons highlighted by light blue lines.

Regarding the geographic locations of detected houses in the study area, each house is assigned a unique ID whose numbering starts from the house located in the southeast end (house 1) and ends at the northwest end (House 63), since houses are aligned in a narrow space along the meandering river.[19] The ID number is based on the number of houses up from house 1 along the river, and the distance from house 1 to 63 along the river is about 7.1 km.

To collect the thermal temperature data from each house,[20] we conducted three drone flights in the winter of 2019: once on Date F and twice in March. The first and second March flights are referred to as Date M1 and Date M2, respectively. Our drone flights were conducted at 18:48–21:13 on Date F,

deliberately excluding areas such as grasslands or sloped terrain. Consequently, we use data from flat areas or bare ground in the vicinity to establish a reliable ground level. Our method of height calculation derives the difference between the house's roof top value and the determined ground level.

[19] Online Appendix Figure A9a shows the entire study area. Each point in Figure A9a represents the centroid of a house, and the blue band running from the southeast to the northwest is the river.

[20] The thermal temperature of an object can be calculated from the emitted electromagnetic energy based on the Stefan–Boltzmann law. The Zenmuse-XT2 (DJI, Shenzhen, China) thermal infrared camera observes the spectral range of 7.5–13.5μm. The system senses thermal infrared waves and converts electronic signals into absolute temperature information.

19:15–22:07 on Date M1, and at 19:10–21:24 on Date M2. Our nighttime collection of temperature data is appropriate because of the disturbance of solar radiation during daytime. Heating material during the day causes limitations in observing the direct source of heat waste, and the specular reflectance of solar radiation causes observation errors, which is avoided in nighttime data collection.[21]

We match the house-location data and roof-temperature data by assuming that, during our study period (February and March), there was no relocation of houses detected in the daytime imagery on Date F.[22] Figure 1b displays a sample of the thermal images matched with house polygons.[23] After separating out the part covered by edges and leaves to avoid capturing the temperature of vegetation covering rooftops, we calculate the average of temperature over the house polygon for each house and assign this average roof temperature to each house location. Figure 1c illustrates the roof temperature averaged over each polygon, corresponding to the house polygons in Figures 1a and 1b. Hereafter, we call this average temperature over each house polygon the roof temperature for each house.

Notably, even though we have 63 house-location observations from the daytime flight, some of these 63 houses lack nighttime temperature values for two reasons, the first of which is technical limitations. Some houses have no temperature values due to the daytime visual images having higher resolution levels than the nighttime thermal images. For observations with roofs covered by tiny leaves or thin edges, we know from the higher-resolution daytime aerial images that houses exist there even when their roofs are covered

[21]According to Balaras and Argiriou (2002), infrared measurements should be performed at night or during cloudy days without high wind speeds. Drone flights using Zenmuse-XT2 cannot be performed on windy or rainy days, and our roof temperature data (infrared thermal image data) were collected by a drone with a thermal sensor at night, meaning that the temperature data were collected under ideal conditions.

[22]It is unlikely that house relocation would occur in such a short time period (approximately one month).

[23]The collected aerial images are in radiometric JPEG (R-JPEG) format. The temperature information is extracted from the R-JPEG data using the FLIR Tools software ver. 6.4 (FLIR, Wilsonville, U.S.). The "box tool" is used to cover the area of the house polygons, and the average and maximum temperature values for each house polygon are extracted.

by vegetation. In the nighttime thermal images characterized by lower resolution, we cannot distinguish roof parts covered by tiny leaves or edges from those not covered by such vegetation. Essentially, mixing of roof parts covered by vegetation and those uncovered occurs in the thermal images. This issue prevented us from calculating the average roof temperature over the polygon by removing the part covered by vegetation. Hence, houses 30, 48, 49, and 53–55 have missing roof temperature values for all dates.[24] Finally, we obtained 46 observations for Date F, 57 observations for the other two dates, and thus 160 records in total.

One thing to note here is that the roof temperature variable shows values lower than the outside temperature, and they sometimes take negative values. For instance, roof temperature is on average 0.219°C (Online Appendix Table A2). This seemingly unnatural tendency of roof temperature is in fact natural, because the temperature of an artificial building measured from the outside is often lower than that in the vegetation area or the outside temperature (Michell and Biggs, 1979).[25] To further show the validity of thermal data we collected in Section 3.1, we present the summary statistics of the temperature data collected in a thermal experiment that we replicated the actual situation.

3 Discussion on Thermal Data as a New Measurement

So far, we have highlighted the potential of thermal data as a novel indicator of current living standards. In this section, we further assess the measure by: (i) validating the thermal data through an experimental analysis, and

[24]Another reason for missing nighttime roof temperature values is simply human error. The nighttime flights were less visible than the daytime flights, so the former have more difficulty distinguishing objects than the latter. Houses 1, 12–17, 47, and 50–52 lack roof temperature data for Date F for this reason, whereas there are no missing roof temperature values for Date M1 and M2.

[25]The fact that the artificial materials show lower temperatures than natural objects does not hinder analyses based on temperature data if investigators are conducting a comparison between artificial objects (houses of homeless persons), not between artificial and natural objects, as we do in the application analysis in Section 4.

(ii) comparing it with an alternative measure for economic status—namely, roof footprint.

3.1 Data Validity: Thermal Experiment

This section addresses three concerns regarding the use of thermal data and supports its validity using results from our thermal experiment. The experiment was designed to confirm: (i) whether indoor temperatures are high enough for habitation, (ii) whether roof surface temperatures reflect changes in indoor temperature, and (iii) whether heat radiation is confined to a narrow spatial area—ensuring the data can be interpreted at the observation unit level. To test these points, we constructed a replica of a typical house from the study area and analyzed the thermal data it produced.

3.1.1 Thermal Experiment

An experiment replicating thermal data collection was conducted on the evening of December 9, 2022, at a private drone field with grass-covered ground, mimicking riverbank conditions. A temporary, simplified replica of a house typical of the study area was built.[26] The frame was covered with blue plastic sheets, duckboards for flooring, and cardboard for interior walls and floor.[27] Two portable heaters and mobile batteries were placed inside as heat sources.

Using the same method as in the study area, we conducted drone flights capturing aerial and thermal images every 3 minutes from 18:12 to 19:24, with a pause at 18:51 for a battery change. The same drone and infrared camera as used in the main study were flown at the altitude of 140 m. Roof temperature was defined as the average within the replica house polygon in the thermal image. Five portable thermometers, logging every 10 seconds,

[26] Due to time constraints, we used a ready-made tent frame—adjusted to match the average house size observed in the study area—and modified it accordingly.

[27] The design of the replica was based on reference materials from field visits to homeless communities conducted by Nairu Kimura, who generously shared his observations from a visit to a riverbank dwelling. We also drew on photographs and descriptions of the study area from Murata (2015).

Table 1: Comparison Between Experimental & Actual Thermal Data

	(1)	(2)	(3)	(4)
Experimental/Actual:	Experimental		Actual	
Site:	Drone field		Riverbank	
Date:	December 2022	Date F, 2019	Date M1, 2019	Date M2, 2019
Recorded temperatures (Mean, deg. C)				
Roof (Raw)	-5.294	-3.778	-0.281	3.946
Outside	3.078	6.000	9.026	11.457
Roof (Adjusted by outside temperature)	-8.372	-9.779	-9.307	-7.512

Outside temperature in the experiment is the temperature recorded by a thermometer set 3.0 m apart from the house wall. "Roof temperature (Adjusted by outside temperature)" is defined as "Roof (Raw) temperature" minus "Outside temperature".

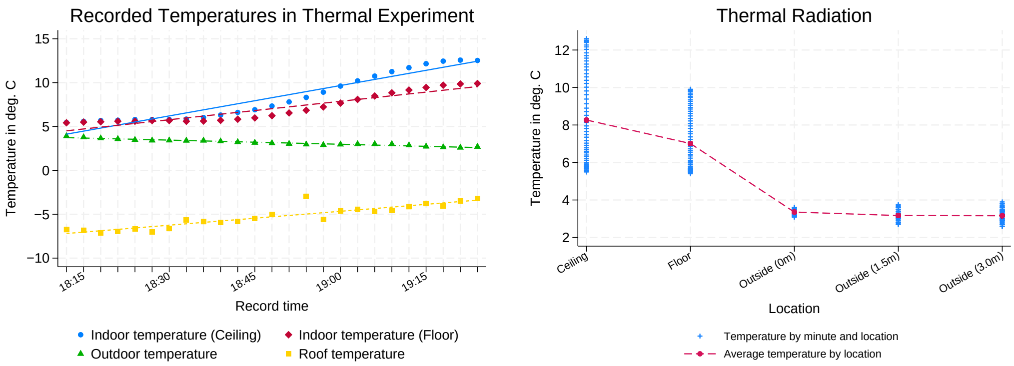
recorded inside and outside temperatures. Two were placed inside (near the ceiling and floor), and three outside at distances of 0, 1.5, and 3.0 m. These readings were averaged to 3-minute intervals. For more details, see Online Appendix Section A.

3.1.2 Surface vs Internal Temperature

Based on the temperature data collected in the experiment, we investigate whether the inside temperature is sufficiently high even when the roof surface temperature recorded by the drone is low. In addition, we investigate if the roof surface thermal information can capture the internal temperature variations.

To begin, we confirm the similarity between the actual and experimental temperatures. Table 1 compares the roof temperatures collected in the experiment and those on the riverbank. Column (1) in Table 1 corresponds to the recorded temperatures in the experiment, and columns (2)–(4) correspond to the actual data by date. "Roof (Raw)" displays the roof temperature recorded by drone, "Outside" displays the outside temperature at the same record time as the roof temperature, and "Roof (Adjusted by the outside temperature)" is defined as the "Roof (Raw)" minus "Outside" temperatures. The values in Table 1 are averages of observations for each date.

The raw roof temperatures in the experiment (column (1)) and the actual data (columns (2)–(4)) do not show comparable values. If adjusted by



(a) Recorded Temperatures (b) Thermal Radiation

Figure 2: Recorded Temperatures in Thermal Experiment & Heat Radiation

Figure 2a shows time series of recorded temperatures in the thermal experiment on December 9, 2022. Inside temperatures (ceiling and floor) and outside temperature were recorded every 10 sec by thermometers and converted to data with 3-min intervals by taking averages by 3 min. Outside temperature was recorded at a point 3.0 m apart from the house exterior wall. Roof temperature was recorded every 3 min except 18:51 due to a drone battery exchange. Straight lines are linear fits for each temperature plot. In Figure 2b, inside (ceiling and floor) and outside temperatures were recorded by thermometers in the thermal experiment. Temperature data were recorded every 10 sec by thermometers and converted to data with 1-min intervals by taking averages by 1 min.

the outside temperatures, however, they are comparable (-8.372°C for the experimental data; while -9.779°C, -9.307°C, and -7.512°C for the actual data), suggesting the reliability of our thermal experiment.

Based on the validity of our experiment from the discussion above, we next make a comparison between the indoor temperatures and roof temperature recorded by drone flights. Figure 2a shows time series plots of the recorded temperatures. From an instant observation of the plots in Figure 2a, indoor temperatures are much higher than the roof temperatures. Comparing the indoor temperatures (those on the ceiling depicted by blue plots and those on the floor by red plots) and the roof temperature recorded by the drone (those marked with yellow plots), the drone temperatures are much lower than the indoor temperatures. This means that roof temperature data based on the drone thermal imagery we collected in the winter of 2019 was not unnaturally low.

Another important observation from Figure 2a is the apparent positive relationship between the indoor temperatures and the roof temperature. It can be safely asserted that a higher roof temperature is accompanied by a higher indoor temperature because (i) the only heat source was the heaters set in the replica house, (ii) the outside temperature plotted by green triangle marks exhibits a slight decline as the experiment time passes, (iii) the indoor temperature plotted by blue or red marks rises through time, and (iv) the roof temperature plotted by yellow squares shows an increasing behavior. This finding supports the assumption that the higher the indoor temperature is, the higher the roof temperature indirectly sensed from outside.

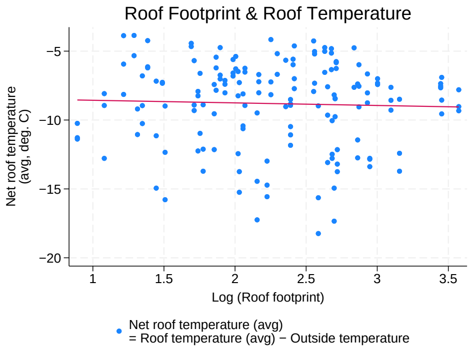
3.1.3 Heat Radiation

Turning to the issue in regard with whether heat radiation from a house is limited to a narrow geographic area, we utilize the temperature data recorded by thermometers. For this exercise, temperature values recorded by thermometers are aggregated at a minute level.

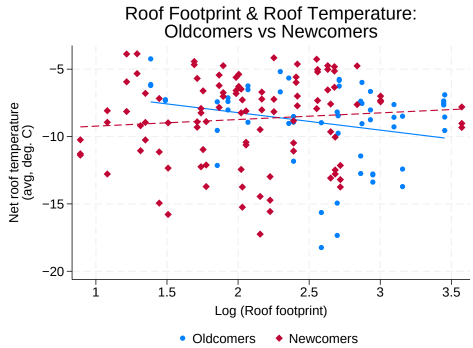
Figure 2b plots the temperatures by minute. Blue cross marks depict the temperature by minute at each location where the thermometers were set (the ceiling and floor inside the imitation house; and 0, 1.5, and 3.0 m apart from the house exterior wall). Red circles are the mean temperatures at each location. From a quick look at the figure, high indoor temperature immediately drops at outside locations, suggesting that heat radiation across houses seems to rarely occur. Hence, at least from our experiment, it is reasonable to assume that heat radiation is limited to a narrow spatial range and that a roof temperature is not affected by the temperatures of neighboring houses through the channel of heat radiation.

3.2 Roof Footprint vs Temperature

While this paper introduces roof temperature as a new measure of current living standards, previous studies have proposed alternative economic indicators that capture wealth. For example, Huang et al. (2021) use roof footprint as a proxy for the total amount of household wealth, and roof material as an



(a) Roof Footprint vs Roof Temperature: Whole Sample



(b) Roof Footprint vs Roof Temperature: Old- & Newcomers

Figure 3: Roof Footprint vs Temperature: By Oldcomers & Newcomers

Figure 3a: Blue points are the scatter plot of the "Net roof temperature" (= "Roof temperature (avg)" minus "Outside temperature on the record time") and the log of roof footprint (m^2). Red line is a linear fit. Figure 3b: Blue points (red diamonds) are the scatter plot of the oldcomers (newcomers). Blue solid line (red dashed line) is a liner fit of the oldcomers (newcomers).

indicator of the quality of wealth. In this section, we compare roof temperature and roof footprint in our sample to examine whether these two variables capture distinct dimensions of a household's economic condition.

Figure 3a compares the log of house floor area (roof footprint) with net roof temperature, defined as the roof temperature minus the outside temperature at the time each infrared image was captured. The two variables show almost no correlation (Corr$= -0.038$). As interpreted in Huang et al. (2021), roof footprint reflects accumulated wealth still in use today. In contrast, roof temperature serves as a proxy for current living standards, reflecting present-day economic activity such as earnings. The lack of correlation between the two is therefore expected, as they capture different dimensions of economic status. This supports the relevance of our thermal data approach for research focused on current economic conditions.[28]

To further explore how these two variables capture different aspects of

[28]Unlike roof temperature, roof footprint is generally fixed once built or extended and cannot reflect declines in economic activity, such as reduced earnings. In contrast, roof temperature captures declines in economic conditions over time.

economic status, we separately plot them for oldcomer and newcomer houses. We identify these groups as follows: using aerial photographs from the summer of 2009–2010 purchased from the Geographic Information Authority of Japan, we delineate house polygons using the same criteria applied to the 2019 drone imagery. Houses that appear in both the past (2009–2010) and present (2019) images in the same location are classified as oldcomers, while those that appear only in the 2019 imagery are classified as newcomers.[29]

Figure 3b plots the relationship between roof footprint and temperature separately for oldcomer and newcomer houses, revealing slightly different patterns. For newcomer houses, there is a weak positive correlation (Corr=0.091), whereas oldcomer houses show a slight negative correlation (Corr=-0.245). We interpret this divergence as follows: newcomer houses represent more recently established wealth, meaning their roof footprints (a proxy for wealth at the time of house construction or extension) are likely more aligned with their current economic activity (reflected by roof temperature). In contrast, oldcomer houses may reflect a larger gap between past and present economic conditions, leading to a weaker or negative correlation. This suggests that, for newer households, wealth and current living standards tend to be more closely linked.

Looking at the oldcomer plots, which show a negative correlation, we interpret this in contrast to the newcomers. Despite having engaged in greater economic activity in the past—as reflected by their larger roof footprints—many oldcomers now experience reduced economic activity. This decline is likely due to aging and physical limitations, which make it harder to earn income through activities like collecting cans or scrap metal. As a result, their lower roof temperatures relative to their larger footprints reflect a drop in current living standards, leading to the observed negative correlation. The comparison between oldcomer and newcomer houses is revisited in Section 4.7 during the application analysis, which focuses on the benefits of social interaction among nearby homeless individuals.

[29]For more details on the classification method, see Online Appendix B. Online Appendix Figure A8b shows the spatial distribution of oldcomer and newcomer houses.

4 Application: Benefits of Social Interaction

This section applies the new measure of current living standards—based on roof temperature—to analyze the benefits of social interaction among homeless individuals. Our study area reveals clusters of homeless communities, and we explore whether individuals in larger clusters gain advantages from increased social interaction with nearby homeless neighbors, using the house location and roof temperature data introduced earlier.

4.1 Sources of Benefits from Social Interaction

A key factor in the positive impact stemming from forming residential clusters is social interaction among homeless persons who live close together. Two types of benefits regarding material well-being may arise for homeless persons living in a cluster from the perspective of social interaction: (i) gains from information diffusion and (ii) mutual aid and risk-sharing ties.[30]

Information sharing is important among homeless persons (Wolch et al., 1993; Yamakita, 2006). In our study area, information diffusion via social interaction is considered an essential benefit to homeless persons. The benefit from information diffusion for homeless persons is related to daily income and expenditures. As explained in Section 2, homeless persons in the study area earn money by illegally selling cans or scrap metals they collected to scrap firms. To do so, individuals must have knowledge of which scrap firms conduct informal transactions. Without this information, individuals can neither sell collected cans nor exchange them for money because only a few scrap firms conduct illicit transactions with homeless persons. Such informal knowledge is only available through interaction with homeless friends; thus, information diffusion via interaction with other homeless persons is vital. On the expenditure side, knowing which supermarkets sell inexpensive food at discounted prices or food shops let homeless persons have waste food is

[30] Mutual friendships within clusters may also contribute to mental well-being, as regular contact with other homeless individuals could enhance emotional satisfaction. However, since roof footprint reflects material rather than mental well-being, our data collection does not capture psychological aspects, which are beyond the scope of this paper.

helpful. Since food scarcity is a major obstacle for the homeless (DHWKC, 2019), shared knowledge about access to inexpensive food providers alleviates anxieties about food deficiency.

In terms of mutual aid and risk-sharing ties, one of the most important benefits of living in a cluster is protection from violence from passersby or juvenile delinquents. Harassment and violence from non-homeless people are ranked as the greatest anxieties for homeless people (DHWKC, 2019; Sugita et al., 2010). Living near other homeless persons may keep them from being badly injured. In addition, in daily life, sharing food when a person cannot earn enough money is one example that highlights the necessity of mutual aid among homeless neighbors. Because income is obtained by collecting and selling cans and metal illicitly, the amount of money earned will change each day. An unreliable income may be relieved by mutual aid among homeless persons who live close together. Likewise, individuals can obtain hand-me-down electromechanical devices or useful goods from other homeless persons.

These social interaction factors may improve the quality of life for the homeless, so we expect that larger social networks (i.e., more homeless persons living close together) may raise their current living standards.

4.2 Cluster Analysis

We utilize the house-location information to group individuals into clusters based on the between-house distances. Doing so enables us to obtain information about geographic proximity among houses. We implicitly assume that geographic proximity can be used as a proxy for social interactions,[31] and we use cluster size (the number of houses in each cluster) as a proxy for social network size.

We first detect the clusters of homeless persons' houses in our sample, using hierarchical cluster analysis. We adopt single-linkage clustering because

[31] Bayer et al. (2008), Hellerstein et al. (2011), and Schmutte (2015) use geographic proximity as a proxy for social interaction. This paper follows this strand of identifying social interactions, and we assume that homeless persons in the same cluster interact with each other rather than with those in different clusters. Moreover, Section 4.3 takes into account geographic distance between any two persons within the same cluster.

of its chaining property, which means that a cluster created in single-linkage clustering has the shape of a chain. As the houses of homeless persons are located in a narrow area along the river, they can be viewed as aligned in this manner. Thus, single-linkage clustering, which generates chain-like clusters, should be appropriate. In single-linkage clustering, the between-cluster distance (linkage function between two clusters, C_1 and C_2) is defined as

$$d_{\text{single}}(C_1, C_2) = \min_{i \in C_1, j \in C_2} d_{ij}, \tag{1}$$

where the distance between houses i and j is denoted as d_{ij}. In single-linkage clustering, the linkage score, or between-cluster distance, in (1) is the distance between the closest pair of houses where one belongs to cluster C_1 and the other belongs to cluster C_2, which has no common elements (houses) with C_1. Starting from singleton clusters, we repeat the assignment to merge pairs of clusters with the shortest between-cluster distance (1), and the process ends when all houses are included in one identical cluster.

The distance between houses, d_{ij}, expresses how far homeless person i must travel to reach another homeless person j. The river is meandering (Online Appendix Figure A9a), so it is inadequate to calculate the direct distance between two houses because homeless persons do not walk over the river. Then, for $i < j$, where i and j indicate house IDs from 1 to 63, we define d_{ij} as $d_{ij} = \sum_{k=i}^{j-1} d_{k,k+1}$, which is obtained by summing the direct length between adjacent house centroids over house i through house j on the line connecting house centroids from house 1 to house 63.

Running single-linkage clustering induced by the dissimilarity matrix whose (i, j)-th entry is d_{ij} and determining the optimal number of clusters based on the Calinski-Harabasz index (CH-index) proposed by Calínski and Harabasz (1974), we optimally grouped 63 houses into 4 clusters. The houses are grouped as follows: cluster 1 = [1–17], cluster 2 = [18–46], cluster 3 = [47–52], and cluster 4 = [53–63], where the numbers in the brackets indicate the house IDs.[32]

[32]For more details about the procedures, see Online Appendix Section C. The detected clusters are marked in Online Appendix Figure A9a.

4.3 Measures of Social Interaction

4.3.1 Cluster Size: Social Interaction Size

We consider variables that measure the size of social interaction (i.e., cluster size) of homeless persons. As mentioned in Section 4.1, homeless persons may improve their living standards by helping each other, so the larger the number of homeless persons residing in a neighborhood, the more the person's life will improve. The number of houses in the same cluster or the cluster size, captures this aspect. As an immediate measure of cluster size, we adopt $N(i)_{-i} \equiv N(i) - 1$, where $N(i)$ is the number of houses in cluster $C(i)$, which house i belongs to, except house i.

Although $N(i)_{-i}$ is a direct variable, it does not consider the distance between houses within a cluster, which may matter because people will interact more frequently with those located closer than those who are farther away. Thus, we modify $N(i)_{-i}$ to take into this aspect by discounting the number of houses by distance to other houses in the same cluster. The simplest modification of $N(i)_{-i}$ is the distance-discounted number of houses in $C(i)$ except house i:[33]

$$DN_{-i}(i) \equiv \sum_{j \in C(i) \setminus \{i\}} e^{-d_{ij}}. \tag{2}$$

4.3.2 Cluster Affluence: Quality-Adjusted Social Interaction Size

We consider a variable expressing the aggregate living standards of each cluster, not just the cluster size itself. The idea is that a homeless person may be better off if homeless persons in their neighborhood live relatively well. We calculate the sum of roof temperatures in cluster $C(i)$, exclusive of house i's contribution. Assuming that roof temperatures in the winter express how well-off homeless persons are, the sum of roof temperatures of houses in a cluster may represent the overall current living standards in that

[33] d_{ij} in (2) is measured in kilometers. Then, for example, in cluster 1, the magnitude of $e^{-d_{ij}}$ for house 1, the most peripheral house in the cluster, ranges from $e^{-d_{1,2}} \approx 0.797$ to $e^{-d_{1,17}} \approx 0.438$. Additionally, for House 9, the most central location in the same cluster, it ranges from $e^{-d_{9,8}} \approx 0.974$ to $e^{-d_{9,1}} \approx 0.596$.

cluster. Such a cluster-affluence variable can be interpreted as the quality-adjusted social interaction size.

When calculating the index, note that the sum should be corrected due to the missing values for some houses, as explained in Section 2.4. Therefore, we propose a modified version of the sum of house roof temperatures in cluster $C(i)$ that replaces missing temperature values with the average of the observed roof temperatures in cluster $C(i)$. The variable is computed as follows:

$$ST_{-i}(i) \equiv \sum_{j \in C'(i) \setminus \{i\}} T_j + \sum_{j \in C(i) \setminus C'(i)} \bar{T}_{C'(i)},$$

where T_j is the roof temperature of house j, $C'(i)$ is the subset of houses in cluster $C(i)$ whose roof temperatures were observed, and $N'(i)$ is the number of houses in the set of houses in $C'(i)$. $\bar{T}_{C'(i)} \equiv \frac{1}{N'(i)} \sum_{j \in C'(i)} T_j$, which is the average roof temperature of houses whose temperatures were observed. This index is computed by assuming that the missing roof temperatures are, on average, equal to non-missing roof temperatures in the same cluster. As in the case of $N(i)_{-i}$ and $DN(i)_{-i}$, the distance-discounted version of $ST(i)_{-i}$ can also be considered:

$$DST_{-i}(i) \equiv \sum_{j \in C'(i) \setminus \{i\}} e^{-d_{ij}} T_j + \sum_{j \in C(i) \setminus C'(i)} e^{-d_{ij}} \bar{T}_{C'(i)}.$$

4.4 Empirical Model

Our hypothesis is that homeless persons residing in larger clusters are better off, which is translated into the hypothesis that the roof temperature of a house located in a larger cluster is higher. Note here that $N(i)_{-i}$ and $DN(i)_{-i}$ are time-invariant variables (i.e., they take the same values across all three dates, as they are based on only the house locations). Then, the estimation model is forced to be a random effects (RE) model:

$$\begin{aligned}\text{RoofTemperature}_{ictd} = {} & \beta_0 + \beta_1 \text{SocialInteraction}_{ic} + \beta_2 \text{OutsideTemperature}_{itd} \\ & + \beta_3 X_i + \theta_d + \nu_i + \epsilon_{ictd}, \end{aligned} \tag{3}$$

where $\text{RoofTemperature}_{ictd}$ is the roof temperature of house i in cluster c at time t (recording time of the roof temperature of house i) on date d. θ_d is the date fixed effects, ν_i is house random effects assumed to satisfy $E(\nu_i) = 0$ and is uncorrelated with other explanatory variables of house i,[34] and ϵ_{ictd} is a stochastic disturbance. $\text{SocialInteraction}_{ic}$ is either $N(i)_{-i}$, $DN(i)_{-i}$, $ST(i)_{-i}$, or $DST(i)_{-i}$.

$\text{OutsideTemperature}_{itd}$ represents the outside temperature near house i at time t on date d. While roof temperature was recorded at minute-level precision, outside temperature data was only available every 10 minutes from the two automated meteorological data acquisition system (AMeDAS) stations nearest to and sandwiching our study area.[35] Since exact-time, on-site measurements were not feasible, we used the closest available AMeDAS time to match each house observation. We then computed a distance-weighted average of the outside temperatures from the two AMeDAS stations based on each house's proximity.[36]

X_i is a set of time-invariant variables of house i other than $\text{SocialInteraction}_{ic}$ to control for natural and geographic aspects and the local amenity.[37] We include a forest-coverage variable, a grass-coverage variable, the distance to the river shore, and the distance to the nearest freely available toilet/tap in the riverside area.[38] A sufficiently large number of toilets and taps are located in the riverside area, so homeless persons can have frequent and daily access.[39] Hence, we expect that the access to toilets and taps improves the living standards of homeless persons in this area.

[34] This assumption of the RE model is not supported in some of the specifications, and we think that ν_i is correlated with $\text{SocialInteraction}_{ic}$. This endogeneity problem is discussed in Section 4.6.

[35] The AMeDAS is a collection of automatic weather stations operated by the Japan Meteorological Agency for automatic observation of precipitation, wind direction/speed, temperature, and sunshine duration to support real-time monitoring of weather conditions.

[36] For more detail, see Online Appendix Section D.1.

[37] Data sources are listed in Online Appendix Table A1.

[38] Online Appendix Figure A2 is a picture of a toilet and a tap in the riverside area. Tap water is available at the site where the toilet is located. In Japan, tap water is good for drinking, so better access to taps means not only better access to water for washing but also for drinking.

[39] Online Appendix Figure A9b is a map of a toilet/tap location layer superimposed on the house-location layer.

For the forest- and grass-coverage variables, we create a 5-m buffer from the house edges and calculate the percentage of land covered by either trees or grass.[40] The variable distance to the river shore is the shortest distance between a house edge and the river shore. These three variables capture geographic and natural characteristics that may directly affect the roof temperature; land areas covered by grass or trees may be warmer than barren areas in the winter. Distance to the river shore may affect roof temperature because, due to the specific heat of water, the water temperature may be higher than the land temperature on winter evenings. Distance to the nearest toilet/tap is considered to affect the current living standards of homeless persons via a channel of improved access to the amenity. We expect that better access to the sanitary amenity improves a homeless persons' living standards, which leads to a higher roof temperature.

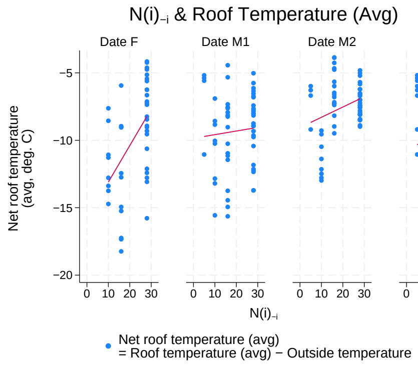
4.5 Estimation: Baseline Results

Because we aim to test whether there is a positive effect on the living standards of homeless persons from residing in a larger cluster, we expect a positive sign for β_1 in (3). Under the assumption that geographic proximity (in our case, belonging to the same cluster) is a proxy for social interactions, $\beta_1 > 0$ can be interpreted as the existence of benefits from social interactions within a cluster.

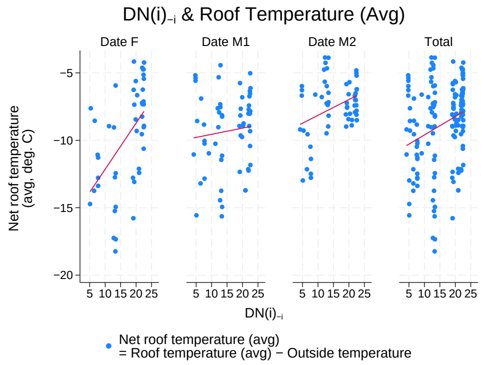
Before proceeding to the estimation results, we briefly consider the relationship between roof temperature and the cluster size for each date in Figure 4.[41] In Figure 4a, the horizontal axis is the number of houses in a cluster $N(i)_{-i}$, and the vertical axis is the net roof temperature, defined as the roof temperature minus the outside temperature. As explained in Section 2.4, some houses have missing temperature values, so only three clusters appear in the panel for Date F. Panels of the other dates display four clus-

[40]For the grass-coverage variable, we assume that grass covers the land hidden by trees in the aerial images.

[41]Online Appendix Figure A7 displays the relationship between $ST(i)_{-i}$ ($DST(i)_{-i}$) and the roof temperature. A casual observation of the graph reveals a positive correlation between the cluster-affluence and the roof temperature.



(a) $N(i)_{-i}$



(b) $DN(i)_{-i}$

Figure 4: Cluster Size & Roof Temperature (Avg.)

Both in Figures 4a and 4b, the vertical axis expresses "Net roof temperature" defined by "Roof temperature (avg.) recorded by the drone" minus "Outside temperature on the record time". Red lines are linear fits.

ters; the largest cluster, with $N(i) = 29$, is houses 18–46 (cluster 2), the next largest, with $N(i) = 17$, is houses 53–63 (cluster 1), the third largest, with $N(i) = 11$, is houses 53–63 (cluster 4), and the smallest, with $N(i) = 6$, is houses 47–52 (cluster 3). In all panels in Figure 4a, the cluster size and roof temperature show a positive relationship, indicating that belonging to a larger cluster is correlated with higher current living standards. Taking $DN(i)_{-i}$ as the cluster-size variable in Figure 4b, an analogous tendency can be found (i.e., roof temperatures in a larger cluster measured by $DN(i)_{-i}$ tend to be higher).

Table 2 shows the results for the baseline model.[42] Odd-numbered columns present the uninstrumented results with cluster-size variables ($N(i)_{-i}$ and $DN(i)_{-i}$) and cluster-affluence variables ($ST(i)_{-i}$ and $DST(i)_{-i}$).[43] As expected, the coefficients of all four social-interaction variables are significantly

[42] Full results are presented in Online Appendix Table A3.

[43] We clustered standard errors by houses rather than by clusters, as wild cluster bootstrap is unsuitable with only four clusters (Canay et al., 2021). The method assumes similar cluster sizes and within-cluster variation in $N(i)_{-i}$—both potentially violated in our case. Instead, we use Conley (1999) standard errors in a robustness check, which account for spatial correlation within a cut-off distance and allow infinite serial correlation. For more explanation and a robust result, see Online Appendix Section E.1.

Table 2: Baseline Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Model:	RE	REIV	RE	REIV	RE	REIV	RE	REIV
Sample:	Full							
Dependent variable:	Roof temperature (deg. C)							
$N(i)_{-i}$	0.092**	0.195***						
	(0.037)	(0.067)						
$DN(i)_{-i}$			0.121**	0.267***				
			(0.051)	(0.096)				
$ST(i)_{-i}$					0.017***	0.124**		
					(0.005)	(0.063)		
$DST(i)_{-i}$							0.030***	0.132**
							(0.008)	(0.063)
Base control	Y	Y	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y	Y	Y
Administrative division FE	Y	Y	Y	Y	Y	Y	Y	Y
# of observations	160	160	160	160	160	160	160	160
Within R-sq	0.82		0.82		0.82		0.83	
Between R-sq	0.65		0.64		0.62		0.62	
Breusch and Pagan LM χ^2	24.79		24.66		30.11		31.69	
Sargan-Hansen stat. χ^2	5.61		6.52		9.21		13.29	
First stage								
Inner/outer curve		9.443***		6.886***		14.697**		13.833***
		(1.766)		(1.392)		(5.749)		(4.860)
Kleibergen-Paap F-stat.		28.60		24.46		6.53		8.10

Standard errors clustered by houses are in parentheses. Base control = Outside temperature, Distance to toilet/tap water, Forest and grass coverage (5-m buffers), Distance to the river shore.
*** p<0.01, ** p<0.05, * p<0.1

positive at the 5% level. When the (quality-adjusted) social interaction size is large, each house has a higher roof temperature, indicating that residents with a larger within-cluster network may be better off.[44] One may have concerns regarding the positive relationship between the roof temperature and cluster size from a viewpoint of thermal radiation from adjacent houses. In other words, a roof temperature may be high because heat radiating from

[44] The finding that homeless persons located in larger clusters tend to experience higher current living standards does not result from the logic used in Culhane (2010), Corinth and Lucas (2018), and Lee and Price-Spratlen (2004), which sheds light on the gravitation of homeless individuals to services offered to them. In their research on homeless agglomeration, homeless persons in areas with homeless concentration benefit by agglomeration of services that target them, such as shelters, free meals, and soup kitchens. However, this mechanism does not apply to our case, because there are no facilities that offer services to homeless persons in the study area. Moreover, our interview with the division of the medial welfare services in the municipality that covers our study area revealed that there were only three free meal services offered there per month.

other houses raises its temperature, which is more likely in a larger cluster. However, our thermal experiment reveals that the possibility of such a situation is low. As we saw in Section 3.1, thermal radiation from a replica house is limited within only a geographically narrow range, and it is unlikely that heat radiating from a house affects the temperatures of houses nearby.

4.6 Endogeneity Concerns

In Section 4.5, we confirmed that roof temperature, a proxy for how well-off each homeless person is, is positively associated with the social interaction size within the same cluster. However, we suspect there are two sources of endogeneity related to the cluster-size/affluence variable: unobservable individual-specific characteristics and unobservable location-specific characteristics.[45] In terms of the endogeneity due to unobservable individual-specific characteristics,[46] we expect that the coefficient of the cluster-size variable suffers from downward bias. Homeless persons who are less able or more disadvantaged may prefer to reside in a larger cluster to seek greater support from other homeless persons. Those who cannot consistently earn enough money may want to stabilize their lives by sharing food or daily necessities with others.

In contrast, those with sufficiently high abilities do not need to rely on others and may prefer to live in isolation. Instead of seeking greater mutual support in a larger cluster, such able homeless individuals may want to live

[45] In assessing the agglomeration effect (measured by the population density in a city) on wages, two sources of endogeneity are recognized (Combes et al., 2010, 2011; Moretti, 2004). The first is unobservable worker-specific characteristics such as ability, which may cause an upward bias of the estimated coefficient of the density variable due to workers' positive sorting to larger (denser) cities, i.e., unobservably more competent workers sort into larger (denser) cities. With an access to individual panel data, it is recommended to eliminate individual fixed effects to resolve the endogenous quality of labor bias (Combes et al., 2010, 2011). The second source of endogeneity is unobservable city-specific characteristics. Cities with unobservably better amenities may attract workers, which raises population density in those cities. Employing instruments for the density variable is encouraged to resolve the endogenous quantity of labor bias (Combes et al., 2010, 2011).

[46] Unobservable heterogeneity among homeless persons does not largely come from demographic differences, as most of the homeless persons in our study area are single older males. Thus, we focus on heterogeneity of ability and how disadvantaged a homeless person is as unobservable characteristics.

in a smaller cluster to avoid competition for cans or metal, which are limited income sources. Approximately 15,000 garbage collection stations for neighborhood non-homeless citizens are evenly distributed near the study area, and homeless persons must visit various garbage stations frequently to collect a lot of cans or metals. If there are many competitors in the neighborhood (i.e., many homeless persons in the same cluster), one can collect only a small amount of cans, which leads to lower earnings. A homeless person who can manage to live on their own may prefer to live in a smaller cluster to circumvent such a situation on their earnings. In short, homeless persons unobservably less able may sort into larger clusters, while those with unobservable high abilities may sort into smaller clusters.

We do not deny completely the possibility that the cluster-size variable suffers from an upward bias (i.e., that more able homeless persons may choose to live in larger clusters). However, as is already discussed, the income sources in the geographic scope they can reach, such as disposed cans and metals, are limited. Such income sources do not increase even when homeless persons locate in larger clusters, because their knowledge exchange itself does not expand the pie of their income source (i.e., the number of disposed cans and metals by non-homeless citizens does not increase even when able homeless persons live in a large cluster). The merit in terms of knowledge spillover via homeless persons' social interaction mainly comes from the access to information about which scrap firms informally transact with them, as well as physical and mutual supports within clusters.

It is, hence, reasonable to consider that more competent homeless persons, who already know a lot about the area, may not need more homeless persons nearby, given their advantage from circumventing the homeless rivals who take the income sources available in the area. In addition, physically stronger individuals may prefer to live in smaller clusters, as they are less likely to need assistance from others nearby. By contrast, less able homeless persons who have limited knowledge about their residential area are more urged by access to information about the scrap firms that make illicit transactions with them. Also, homeless persons with declining physical strength may prefer to live with other homeless persons nearby, seeking for mutual support from them. Thus, the negative-sorting story seems to be more reasonable than the positive-sorting story.

Endogeneity caused by unobserved individual characteristics should ideally be addressed by introducing individual fixed effects into the estimation model. Unfortunately, this option is not possible in our dataset, as our featured variable, cluster size, is time invariant in the study period. Nevertheless, if the estimated coefficient of cluster size is positive, then its real impact on roof temperature should also be positive, because, following the above discussion of negative sorting into larger clusters, the naively estimated coefficient is considered to have a downward bias (the positive effect of cluster size is underestimated).

Regarding the second source of endogeneity coming from unobservable location-specific characteristics—locations unobservably beneficial for their current economic activities may attract more homeless persons—is addressed using an instrumental variable. An appropriate instrumental variable should affect the current living standards of a homeless person only through the channel of house distribution. One possibility is a variable that expresses only the suitability for house construction in a particular area on the riverbank, as it is more likely that a larger cluster of houses is generated in areas where building is easier. Specifically, a variable indicating whether a location is on the inner or outer curve of the meandering river can serve this purpose. The relevance of this instrument is supported by geological reasoning.

Inner river curves are typically flatter due to sedimentation, as water flow slows along these bends. In contrast, outer curves are steeper and narrower due to erosion. As a result, inner curves tend to be more suitable for building, often leading to larger clusters of houses. This suggests that riverside areas on inner curves are more likely to accommodate housing than those on outer curves, primarily because of the flatter terrain. Indeed, inspecting house distributions on both sides (our study area side and the other) of the river in 2020,[47] we find a tendency for there to be more houses are located along inner curves than along outer curves.[48] There are almost no houses on the

[47]See Online Appendix Figure A9c.

[48]Although we collected house distribution data in 2020, these data are inadequate

right-opposite side of the shore along which a house cluster is generated. This is natural because the two shores of a river can be usually approximated as a pair of nearly parallel curves, implying that the right-opposite side of an inner curve is an outer curve, and vice versa. This observation aligns with the idea that there are more houses located on inner curves than on outer curves.

While the exogeneity of the instrument cannot be directly tested, we argue that the inner/outer curve variable affects roof temperature only through its impact on cluster size, once control variables are included. In other words, the exogeneity condition is said to hold if it is reasonable to assume that a homeless person's location on the inner or outer curve of the river affects their current economic activities—such as collecting scrap metal or cans, or mutual support like food and second-hand goods sharing—only through the channel of the cluster size, after controlling for neighborhood amenities that may also influence their economic activity.

As noted in Section 4.4, our control variables include distance to the nearest toilet or tap water, forest and grass coverage, and distance to the riverbank—all of which capture neighborhood amenities that may affect the economic activity of homeless individuals. Access to water and sanitation, along with more stable outdoor temperatures from surrounding vegetation, can reduce physical strain and support physical tasks like collecting cans or scrap metal. If inner curve areas offer better work-related amenities than outer curves, this could explain higher living standards there. However, after conditioned on these amenity-related variables, this "better-amenity"

for our formal analysis, as the house distribution is unstable and changed considerably in a short period in the winter of 2020. In October 2019, a large typhoon hit the study area, and most of the houses were washed away. This was broadcasted as a rare event that had not happened in the past 50 years (`https://www.sankei.com/article/20191013-DHD6RFIVIBNCZDIY4BPXT25SXY/` in Japanese, accessed at January 21, 2023). After a few months, we collected aerial images of houses built in January 2020, as depicted in Online Appendix Figure A9c. When we tried to collect the roof temperature data in February 2020, however, the house distribution appears to have changed further, possibly because this area was still in recovery. Moreover, due to leaves and edges carried away by the flood and piled up near houses, we were unable to construct vegetation variables. The data in the winter of 2020 is only employed for a casual observation. This is why we did not exploit the data in 2020.

channel is closed. Thus, the inner/outer curve variable can reasonably be considered exogenous with respect to current economic activities.

One might still suspect differences between the residential areas of homeless individuals on the inner versus outer curves of the river. While inner curves generally offer more flat inland space and may appear more favorable, this space is often used by non-homeless residents for playgrounds or sports fields. As a result, even on inner curves, only narrow, less desirable strips along the riverbank are available for homeless settlement—similar to outer curves. This suggests that, regardless of curve location, homeless individuals tend to occupy equally marginal riverside areas.

Another potential concern is that the risk of flooding may differ depending on whether houses are located on the inner or outer curve of the river. Since river flow is faster on outer curves, homeless individuals in those areas might face greater flood risk, potentially lowering their economic activity. However, Murata (2015) reports that the dwellings are typically built in areas that are rarely flooded. This suggests there is no significant difference in flood risk between inner and outer curve locations. Overall, the inner/outer-curve dummy appears to be a valid instrument.

A rough idea to construct the inner/outer-curve dummy variable along the meandering river is as follows: if the straight line connecting two sufficiently distant points along the river shore passes through a water area, houses located in that area are classified as being on the outer curve. If the line passes through a land area, houses in that area are considered to be on the inner curve. Based on this, we construct an inner/outer curve dummy variable that takes 1 if a house is located on an inner curve and 0 if located on an outer curve. For the detailed description on how to construct the instrument, see Online Appendix Section D.2.

Columns (2) and (4) in Table 2 exhibit the instrumented results for specifications based on the choice of $N(i)_{-i}$ and $DN(i)_{-i}$, respectively. In the first-stage estimation, the instrumental variable shows sufficient significance (the sufficiently large first-stage F-statistics for the weak instrument), with the expected sign in all specifications, which supports the relevance of the instrumental variable. On the other hand, the first-stage F-statistics in columns

(6) and (8), corresponding to the results for $ST(i)_{-i}$ and $DST(i)_{-i}$, are not sufficiently high, exhibiting the weak instrument issue. Thus, this section mainly focuses on the quality-unadjusted cluster-size regressions.[49]

In the second-stage estimation, the coefficients of the cluster-size variables, $N(i)_{-i}$ and $DN(i)_{-i}$, exhibit strong positive significance, indicating that living among more nearby homeless persons may improve present living standards. Moreover, the coefficient value in the $N(i)_{-i}$ specification can be interpreted in a way that an additional house in the same cluster increases 0.195°C of roof temperature (column (2)). In the same vein, an additional increase of a distance-discounted house raises roof temperature by 0.268°C (column (4)).

Next, to consider the direction of the bias, we compare the instrumented and uninstrumented results. By comparing the coefficients of cluster-size variables in columns (1) and (2) for the $N(i)_{-i}$ specification and in columns (3) and (4) for the $DN(i)_{-i}$ specification, we see that the uninstrumented estimates are negatively biased in both specifications. In fact, this downward bias is unexpected, because unobservably beneficial locations were expected to attract homeless persons, presumably causing an upward bias of the estimated coefficient of the social-interaction variable. We interpret the downward bias found in this instrumented result as follows: unobservably preferable places with attractive first nature are strictly managed and controlled by local administrators so as not to be inhabited by homeless persons informally. Unobservably less advantageous locations are relatively freely available for them, and such disadvantageous locations are easier for them to illicitly occupy, which can lead to the downward bias of the uninstrumented coefficient of the cluster-size variable.

[49] We add some comments on the undesired aspects of $ST(i)_{-i}$ and $DST(i)_{-i}$ as a measure of the cluster affluence. The summation of roof temperatures over a cluster is subject to the replacement of missing roof temperatures $T_{j\in C(i)\setminus C'(i)}$ with $\bar{T}_{C'(i)}$. However, this replacement may not be ideal for clusters with the relatively large number of houses lacking temperature values. Thus, we utilize simpler but more-secure measures for the cluster-size variable, such as $N(i)_{-i}$ and $DN(i)_{-i}$, in the following section. In Section 4.8, however, we conduct several additional regressions focusing on $ST(i)_{-i}$ and $DST(i)_{-i}$, all of which show sufficiently large first-stage F-statistics.

Table 3: Roof Footprint vs Roof Temperature

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Model:	OLS	OLS	OLS	RE	REIV	RE	REIV	RE	RE
Dependent variable:	Log (Footprint, m^2)			Roof temperature (deg. C)					
Oldcomer	0.592***	0.481***	0.483***	-0.292	-0.664	-0.292	-0.695	-0.196	0.109
	(0.172)	(0.176)	(0.179)	(0.580)	(0.625)	(0.575)	(0.639)	(1.569)	(1.621)
$N(i)_{-i}$		0.031***		0.099***	0.219***			0.101**	
		(0.010)		(0.038)	(0.069)			(0.044)	
$DN(i)_{-i}$			0.041***			0.131**	0.302***		0.139**
			(0.015)			(0.052)	(0.102)		(0.057)
Oldcomer$\times N(i)_{-i}$								-0.005	
								(0.070)	
Oldcomer$\times DN(i)_{-i}$									-0.026
									(0.095)
Base control	Y	Y	Y	Y	Y	Y	Y	Y	Y
Date FE	-	-	-	Y	Y	Y	Y	Y	Y
Administrative division FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
# of observations	63	63	63	160	160	160	160	160	160
Adj R-sq	0.11	0.16	0.16						
Within R-sq				0.82		0.82		0.82	0.82
Between R-sq				0.65		0.65		0.65	0.65
Breusch and Pagan LM χ^2				24.60		24.45		24.56	24.24
Sargan-Hansen stat. χ^2				5.69		6.61		5.78	6.70
First stage									
Inner/outer curve					9.100***		6.602***		
					(1.799)		(1.381)		
Kleibergen-Paap F-stat.					25.58		22.86		

Standard errors clustered by houses are in parentheses. Base control for roof footprint regressions = Distance to toilet/tap water, Forest and grass coverage (5-m buffers), Distance to the river shore. Base control for roof temperature regressions = Outside temperature, Distance to toilet/tap water, Forest and grass coverage (5-m buffers), Distance to the river shore.
*** p<0.01, ** p<0.05, * p<0.1

4.7 Footprint and Temperature Revisited: Oldcomers vs Newcomers

In Section 3.2, we showed that roof footprint and roof temperature capture different dimensions of economic status by comparing oldcomer and newcomer houses.[50] We revisit this relationship in the current section. The control variables used in the regressions, where the log of roof footprint is the dependent variable, are the same as in the baseline regression (3), except that outside temperature and date fixed effects are excluded.[51]

[50] The newcomer share for each cluster is around 70% (65%–80%) and a clear tendency of the share regarding the cluster size was not observed.

[51] For the full results, see Online Appendix Table A4.

Columns (1)–(3) of Table 3 report results from the roof footprint regressions, while columns (4)–(9) show those for roof temperature. All regressions include a dummy variable for oldcomers versus newcomers. We begin by examining columns (1)–(7) to compare differences in footprint and temperature. Focusing on the oldcomer variable, a clear distinction emerges: in all footprint regressions (columns (1)–(3)), the coefficient on the oldcomer dummy is significantly positive, indicating that oldcomer houses are larger than those of newcomers. This suggests that oldcomers tend to possess greater wealth.[52] By contrast, in the roof temperature regressions shown in columns (4)–(7), the oldcomer variable is often negative and statistically insignificant. This suggests that oldcomers do not necessarily enjoy higher current living standards than newcomers. The contrasting results between the roof footprint and roof temperature regressions highlight that these two measures capture different aspects of the economic conditions of homeless individuals.

Another observation from columns (2) and (3) is that houses in larger clusters tend to have larger footprints. Similar to what can be found with the roof temperature regressions presented in columns (4)–(7), not only the current living standards, expressed by the roof temperature, but also the amount of wealth, captured by the roof footprint, tend to be larger in more populated clusters.

Lastly, motivated by the oldcomer–newcomer classification, we examine whether the benefits of social interaction differ between the two groups. Columns (8) and (9) present results that include an interaction term between cluster size and the oldcomer indicator.[53] Both $N(i)_{-i}$ and $DN(i)_{-i}$

[52]We do not report the instrumented results for columns (2) and (3), even though they include cluster-size variables. This is because it is conceptually inappropriate to consider that current social interaction influences house sizes that were built in the past. Moreover, since the primary aim of this table is to compare roof footprint and temperature, we focus on this contrast and omit the instrumented estimates. That said, we did conduct instrumental variable regressions, and the findings remain consistent: the instrumented coefficient of $N(i)_{-i}$ is estimated to be 0.058 and is significant at the 1% level, with a first-stage F-statistic of 29.79.

[53]In the interaction term regressions, there are two endogenous variables: $N(i)_{-i}$ ($DN(i)_{-i}$) and Oldcomer$\times N(i)_{-i}$ (Oldcomer$\times DN(i)_{-i}$). weakivtest2 Stata command does not support the calculation of the Generalized First-Stage Statistic proposed by Lewis and Mertens (2022) when using an RE model. Therefore, we report only the unin-

are significantly positive at the 5% level, while the interaction terms are statistically insignificant and close to zero with negative coefficients. This suggests that both oldcomers and newcomers benefit similarly from being in larger clusters, likely through increased social interaction within the same community. What might explain this? One plausible interpretation is that older, less physically active oldcomers may receive support from less old, more physically capable newcomers.[54] At the same time, oldcomers—having lived in the area longer—likely possess more knowledge about local conditions, including informal job information, which they may share with newcomers. In this way, both groups benefit from mutual support and information exchange within their clusters.

4.8 Robustness Checks

We conducted several robustness checks, all of which support the results in this section (see Online Appendix E).

5 Conclusion

Reaching vulnerable populations is vital for both academic and humanitarian reasons, as they often require social welfare support. However, when respondents face issues like addiction or cognitive impairments, obtaining reliable socio-economic quantitative data through surveys becomes extremely challenging.

To overcome this, the study proposed a novel, non-survey method using drone-mounted infrared cameras to measure winter roof temperatures as a proxy for the current household living standards. Higher roof temperatures indicate greater use of heat-generating appliances and thus higher living standards. Validation through a replica house experiment confirmed

strumented results.

[54]For instance, oldcomer residents often suffer from poor health, making it difficult for them to consistently collect cans and scrap metal. In such cases, physically able newcomers can take over these activities, helping to stabilize income through informal risk-sharing arrangements.

that roof temperature reliably reflects indoor temperatures and captures current economic activity, differing from roof footprint which reflects the amount of wealth. Applying this method, the study found that urban squatters in larger clusters have higher roof temperatures, suggesting social interaction benefits.

This study pioneers the use of drone-based thermal imagery to measure living standards in urban squatter communities, introducing a promising interdisciplinary approach that overcomes traditional data collection challenges. By bridging natural and social sciences, this method offers new avenues for researching hard-to-access populations and advancing our understanding of welfare dynamics.

References

Akiyama, Yuki, Kotaro Iizuka, Osamu Yachida, and Satoru Sugita. 2019. "Fundamental Research on Estimating the Vacant Houses using Thermal Infrared and Visible Imagery Observed from Unmanned Aerial Vehicle (UAV)." *Proceedings of the 28th Annual Meeting on Geographic Information Systems Association of Japan (in Japanese)* F-2-3.

Balaras, C Argitiou, and AA Argiriou. 2002. "Infrared Thermography for Building Diagnostics." *Energy and Buildings* 34 (2): 171–183.

Baragwanath, Kathryn, Ran Goldblatt, Gordon Hanson, and Amit K Khandelwal. 2021. "Detecting urban markets with satellite imagery: An application to India." *Journal of Urban Economics* 125 103173.

Barnhardt, Sharon, Erica Field, and Rohini Pande. 2017. "Moving to opportunity or isolation? Network effects of a randomized housing lottery in urban India." *American Economic Journal: Applied Economics* 9 (1): 1–32.

Bayer, Patrick, Stephen L Ross, and Giorgio Topa. 2008. "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes." *Journal of Political Economy* 116 (6): 1150–1196.

Brueckner, Jan K. 2013. "Slums in Developing Countries: New Evidence for Indonesia." *Journal of Housing Economics* 22 (4): 278–290.

Calínski, Tadeusz, and Jerzy Harabasz. 1974. "A Dendrite Method for Cluster Analysis." *Communications in Statistics-theory and Methods* 3 (1): 1–27.

Canay, Ivan A, Andres Santos, and Azeem M Shaikh. 2021. "The Wild Bootstrap with a " Small " Number of " Large " Clusters." *Review of Economics and Statistics* 103 (2): 346–363.

Chen, Xi, and William D Nordhaus. 2011. "Using luminosity data as a proxy for economic statistics." *Proceedings of the National Academy of Sciences* 108 (21): 8589–8594.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2011. "The Identification of Agglomeration Economies." *Journal of Economic Geography* 11 (2): 253–266.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux. 2010. "Estimating Agglomeration Economies with History, Geology, and Worker Effects." In *Agglomeration Economics*, edited by Glaeser, Edward L. 15–66, Oxford, UK: Oxford University Press.

Conley, Timothy G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92 (1): 1–45.

Corinth, Kevin, and David S Lucas. 2018. "When Warm and Cold Don't Mix: The Implications of Climate for the Determinants of Homelessness." *Journal of Housing Economics* 41 45–56.

Culhane, Dennis P. 2010. "Tackling Homelessness in Los Angeles' Skid Row: The Role of Policing Strategies and the Spatial Deconcentration of Homelessness." *Criminology & Public Policy* 9 (4): 851–857.

DHWKC. 2019. "Project for Supporting the Homeless to be Independent (2019–2023) (in Japanese)." *Report of Division of Health & Welfare in Kawasaki City.*

Early, Dirk W. 2005. "An Empirical Investigation of the Determinants of Street Homelessness." *Journal of Housing Economics* 14 (1): 27–47.

Feler, Leo, and J Vernon Henderson. 2011. "Exclusionary Policies in Urban Development: Under-servicing Migrant Households in Brazilian Cities." *Journal of Urban Economics* 69 (3): 253–272.

Fisher, Walter D. 1958. "On Grouping for Maximum Homogeneity." *Journal of the American Statistical Association* 53 (284): 789–798.

Gendron-Carrier, Nicolas, Marco Gonzalez-Navarro, Stefano Polloni, and Matthew A Turner. 2022. "Subways and urban air pollution." *American economic journal: Applied economics* 14 (1): 164–196.

Go, Eugenia, Kentaro Nakajima, Yasuyuki Sawada, and Kiyoshi Taniguchi. 2022. "On the use of satellite-based vehicle flows data to assess local economic activity: The case of philippine cities."Technical report, ADB Economics Working Paper Series.

Halpin, Brendan. 2016. "Cluster Analysis Stopping Rules in Stata." *University of Limerick Working Paper* WP2016-01.

Hellerstein, Judith K, Melissa McInerney, and David Neumark. 2011. "Neighbors and Coworkers: The Importance of Residential Labor Market Networks." *Journal of Labor Economics* 29 (4): 659–695.

Henderson, J Vernon, Tanner Regan, and Anthony J Venables. 2021. "Building the city: from slums to a modern metropolis." *The Review of Economic Studies* 88 (3): 1157–1192.

Henderson, J Vernon, Adam Storeygard, and David N Weil. 2012. "Measuring economic growth from outer space." *American economic review* 102 (2): 994–1028.

Hodler, Roland, and Paul A Raschky. 2014. "Regional favoritism." *The Quarterly Journal of Economics* 129 (2): 995–1033.

Huang, Luna Yue, Solomon M Hsiang, and Marco Gonzalez-Navarro. 2021. "Using satellite imagery and deep learning to evaluate the impact of anti-poverty programs."Technical report, National Bureau of Economic Research.

Huebner, Gesche, David Shipworth, Ian Hamilton, Zaid Chalabi, and Tadj Oreszczyn. 2016. "Understanding electricity consumption: A comparative contribution of building factors, socio-demographics, appliances, behaviours and attitudes." *Applied energy* 177 692–702.

Ito, Kae, Suimei Morikawa, Tsuyoshi Okamura, Kentaro Shimokado, and Shuichi Awata. 2014. "Factors Associated with Mental Well-being of Homeless People in Japan." *Psychiatry and Clinical Neurosciences* 68 (2): 145–153.

Lall, Somik V, Mattias KA Lundberg, and Zmarak Shalizi. 2008. "Implications of Alternate Policies on Welfare of Slum Dwellers: Evidence from Pune, India." *Journal of Urban Economics* 63 (1): 56–73.

Lee, Barrett A, and Townsand Price-Spratlen. 2004. "The Geography of Homelessness in American Communities: Concentration or Dispersion?" *City & Community* 3 (1): 3–27.

Lee, Yong Suk. 2018. "International isolation and regional inequality: Evidence from sanctions on North Korea." *Journal of Urban Economics* 103 34–51.

Lewis, Daniel J, and Karel Mertens. 2022. "A Robust Test for Weak Instruments with Multiple Endogenous Regressors." *FRB of New York Staff Report* (1020): .

Marx, Benjamin, Thomas M Stoker, and Tavneet Suri. 2019. "There is no free house: Ethnic patronage in a Kenyan slum." *American Economic Journal: Applied Economics* 11 (4): 36–70.

Marx, Benjamin, Thomas Stoker, and Tavneet Suri. 2013. "The economics of slums in the developing world." *Journal of Economic perspectives* 27 (4): 187–210.

Michalopoulos, Stelios, and Elias Papaioannou. 2013. "Pre-colonial ethnic institutions and contemporary African development." *Econometrica* 81 (1): 113–152.

Michell, D, and KL Biggs. 1979. "Radiation Cooling of Buildings at Night." *Applied Energy* 5 (4): 263–275.

Moretti, Enrico. 2004. "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-Sectional Data." *Journal of Econometrics* 121 (1-2): 175–212.

Morikawa, Suimei, Ritei Uehara, Koji Okuda, Yuko Shimizu, and Yoshikazu Nakamura. 2011. "Morbidity of Mental Illness among Homeless People in One Area of Tokyo (in Japanese)." *Japanese Journal of Public Health* 58 (5): 331–339.

Murata, Ramu. ed. 2015. *Homeless Superstar Legends (in Japanese)*. Loft Books.

Nishio, Akihiro, Ryo Horita, Tadahiro Sado, Seiko Mizutani, Takahiro Watanabe, Ryosuke Uehara, and Mayumi Yamamoto. 2017. "Causes of Homelessness Prevalence: Relationship Between Homelessness and Disability." *Psychiatry and Clinical Neurosciences* 71 (3): 180–188.

Oakes, Peter M, and Ros C Davies. 2008. "Intellectual Disability in Homeless Adults: A Prevalence Study." *Journal of Intellectual Disabilities* 12 (4): 325–334.

Okamoto, Yoshihiro. 2007. "A Comparative Study of Homelessness in the United Kingdom and Japan." *Journal of Social Issues* 63 (3): 525–542.

Okamura, Tsuyoshi, Kae Ito, Suimei Morikawa, and Shuichi Awata. 2014. "Suicidal Behavior Among Homeless People in Japan." *Social Psychiatry and Psychiatric Epidemiology* 49 (4): 573–582.

Okamura, Tsuyoshi, Tadashi Takeshima, Hisateru Tachimori, Ken Takiwaki, Yuki Matoba, and Shuichi Awata. 2015. "Characteristics of Individuals with Mental Illness in Tokyo Homeless Shelters." *Psychiatric Services* 66 (12): 1290–1295.

Philippot, Pierre, Catherine Lecocq, Fanny Sempoux, Hilde Nachtergael, and Benoit Galand. 2007. "Psychological Research on Homelessness in Western Europe: A Review from 1970 to 2001." *Journal of Social Issues* 63 (3): 483–503.

Salize, Hans Joachim, Cornelia Dillmann-Lange, Gerlinde Stern, Beate Kentner-Figura, Klaus Stamm, Wulf Rössler, and Fritz Henn. 2002. "Alcoholism and Somatic Comorbidity Among Homeless People in Mannheim, Germany." *Addiction* 97 (12): 1593–1600.

Santamouris, M, SM Alevizos, L Aslanoglou et al. 2014. "Freezing the poor—Indoor environmental quality in low and very low income households during the winter period in Athens." *Energy and Buildings* 70 61–70.

Schmutte, Ian M. 2015. "Job Referral Networks and the Determination of Earnings in Local Labor Markets." *Journal of Labor Economics* 33 (1): 1–32.

Storeygard, Adam. 2016. "Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa." *The Review of economic studies* 83 (3): 1263–1295.

Sugita, Sanae, Nobuhiro Kobayashi, and Masato Dohi. 2010. "The Movement of the Sleeping Places of Homeless People and Exclusion of those People (in Japanese)." *Journal of the City Planning Institute of Japan* 45 751–756.

Suzuki, Wataru. 2008. "What Determines the Spatial Distribution of Homeless People in Japan?" *Applied Economics Letters* 15 (13): 1023–1026.

Takano, Takehito, Keiko Nakamura, Sachiko Takeuchi, and Masafumi Watanabe. 1999. "Disease Patterns of the Homeless in Tokyo." *Journal of Urban Health* 76 (1): 73–84.

Takeuchi, Akie, Maureen Cropper, and Antonio Bento. 2008. "Measuring the Welfare Effects of Slum Improvement Programs: The Case of Mumbai." *Journal of Urban Economics* 64 (1): 65–84.

Umishio, Wataru, Toshiharu Ikaga, Yoshihisa Fujino et al. 2020. "Disparities of indoor temperature in winter: A cross-sectional analysis of the Nationwide Smart Wellness Housing Survey in Japan." *Indoor air* 30 (6): 1317–1328.

Van Straaten, Barbara, Carola TM Schrijvers, Jorien Van der Laan, Sandra N Boersma, Gerda Rodenburg, Judith RLM Wolf, and Dike Van de Mheen. 2014. "Intellectual Disability Among Dutch Homeless People: Prevalence and Related Psychosocial Problems." *PLoS One* 9 (1): .

Wolch, Jennifer R, Afsaneh Rahimian, and Paul Koegel. 1993. "Daily and Periodic Mobility Patterns of the Urban Homeless." *The Professional Geographer* 45 (2): 159–169.

Yamakita, Teruhiro. 2006. "Comradeship within Communities of Homeless People (in Japanese)." *Japanese Sociological Review* 57 (3): 582–599.

Online Appendix to Heat as a Measure of Living Standards: Exploring Social Interaction Benefits in Squatter Communities

Mariko Nakagawa Kotaro Iizuka

A Thermal Experiment

Building on Section 3.1.1, this section provides a detailed explanation of the thermal experiment we conducted. The experiment to replicate the thermal data collected in the study area was conducted in the evening of December 9, 2022. The experiment site was a private drone field with a ground covered by grass, similar to the natural conditions on the riverbank. We temporarily built a simplified replica of a house similar to those found in our study area. The replica house had sides equal to 3 m and a height equal to 2.2 m.[A1] We covered the framework of the house with blue plastic sheets for its exterior and duckboards on the ground as a floor. Also, the interior thin walls and floor were covered by cardboard.[A2] For a heat origin inside the replica house, we put two portable electric heaters and two volume mobile batteries. For pictures of our replica house, see Figure A3.

[Figure A3 around here]

Analogous to the data collection method in the study area, we conducted drone flights and took aerial and thermal images of the replica house. Figure A4a is the aerial picture of the house taken by the drone before the sunset, and Figure A4b is the thermal image we took at 19:15.

[A1] Due to the limited rental time of the drone field, we put up a ready-made tent framework whose size was similar to the average house size observed in the study area, and then modified it.

[A2] We thank Nairu Kimura for kindly sharing his experience when he visited a house of a homeless person on the riverbank of our study area. Referring to his video and pictures printed in Murata (2015), we built the replica.

[Figure A4 around here]

We conducted drone flights, capturing thermal images at 3-minute intervals between 18:12 and 19:24. One image, at 18:51, is missing due to a battery replacement. The equipment used—the Matrice 210 drone and Zenmuse XT2 thermal camera—was the same as in the main study area. Flights were conducted at an altitude of 140 meters. From the thermal and aerial images, we calculated the average temperature within each house polygon, which we define as the roof temperature recorded by the drone.

We collected the indoor and outdoor temperatures of the replica house with five portable thermometers that can digitally log temperature values every 10 second. Two thermometers were set inside the house; one was set near the ceiling, and the other on the floor. The other three thermometers were set at different locations outside; immediately outside the exterior wall of the replica house, 1.5 m apart from the wall, and 3.0 m apart from it. The recorded temperatures were converted to 3-minute-level data by averaging 18 values of a 10-second interval.

B Identification of Oldcomer and Newcomer Houses

This section explains: (i) the selection of the past daytime imagery, (ii) the method used to delineate house polygons in the past imagery, (iii) how we identify oldcomer and newcomer houses by comparing past and present images, and (iv) why polygons representing the same houses in the two time periods do not perfectly align—necessitating some allowance for discrepancies between identical houses in the past and present imagery data.

B.1 Selecting Past Daytime Imagery Data

Since our 2019 data was collected in winter, we needed to use winter imagery for the historical data as well. Among the daytime aerial photographs available from the Geospatial Information Authority of Japan (GIAJ) for our study area, only those from 2009 and 2010 were taken in winter.[A3] GIAJ provides 2009 winter images only for the southern part of the study area, and 2010 winter images only for the northern part. For this reason, we used both the 2009 and 2010 winter aerial photographs. Some overlap exists between the two in the central area.

B.2 Identifying Oldcomer and Newcomer Houses

We began by delineating house polygons in the 2009 and 2010 imagery separately, using the same criteria applied to the 2019 drone daytime imagery: (i) a rectangular or box shape, (ii) covered with blue, green, or gray plastic sheets, and (iii) sides longer than approximately 1.5 m and height greater than approximately 1 m. Because the 2009 and 2010 aerial images overlap in the central part of the study area, some houses appear in both sets of imagery. However, as discussed in the next section, polygons drawn from the 2009 and 2010 images do not always align perfectly, even when representing the same house in terms of location and general shape. Therefore, we treat two polygons—one from 2009 and one from 2010—as representing the same house if they overlap by more than 70%:[A4]

$$\frac{\text{“2010 house polygon area”} \cap \text{“2009 house polygon area”}}{\text{“2009 house polygon area”}} > 0.7.$$

For houses that appear in both the 2009 and 2010 imagery and are identified as the same, we retain the 2010 polygons over the 2009 ones. In this way, we construct the past house polygon dataset by combining the 2009 and 2010 daytime imagery.

Using the past house polygons constructed in the previous step, we identify whether a house in the present data (2019) corresponds to one in the past data (2009–2010) by comparing their locations. As in the earlier step, we consider two polygons—one from the past and one from the present—to

[A3] While 2017 aerial photographs are also available, they were taken in summer and are therefore unsuitable for this analysis.

[A4] The results remain similar even when using an 80% overlap threshold to identify matching house polygons from 2009 and 2010.

represent the same house if they overlap by more than 70%:[A5]

$$\frac{\text{“Present house polygon area”} \cap \text{“Past house polygon area”}}{\text{“Past house polygon area”}} > 0.7.$$

After matching identical houses in the past and present data, we classify each house in the present data as either an oldcomer or a newcomer. Houses that appear in both the past and present datasets are labeled oldcomers, while those only present in the current data are classified as newcomers. The geographic distribution of old- and newcomer houses is shown in Figure A8b.

[Figure A8b around here]

B.3 Discrepancies Between Identical Polygons in Past and Present Data

Due to the combined effects of imagery data acquisition methods, resolution disparities, and processing limitations, even identical objects may not appear perfectly aligned across different datasets. Then, assuming 100% spatial overlap as a criterion for object correspondence is unrealistic.[A6] This study used a threshold of 70% overlap between polygons from 2009, 2010, and 2019 to determine whether they are identical houses. This approach relies not on strict absolute positional matching but on relatively consistent spatial relationships and is regarded as both technically valid and practically robust.

The reasons behind the misalignment of objects can occur due to the following explanations: “Absolute coordinates” refer to geospatial positions defined by a globally consistent reference system, such as a national geodetic control network. The GEONET (GNSS Earth Observation Network System) serves as a standard for precise positioning in Japan. However, the aerial photographs used in this analysis between 2009 and 2010 were not georeferenced using rigorous methods, such as GNSS-based surveying or the precise placement of ground control points (GCPs). For instance, GNSS data

[A5] The results remain similar even when using an 80% overlap threshold to identify matching house polygons from 2009–2010 and 2019.

[A6] In addition, if a house was extended, 100% spatial overlap is impossible.

recorded by aerial platforms or drones typically contains inherent positional errors ranging from one to several meters, influenced by positioning satellite geometry during acquisition, ionospheric and tropospheric conditions, and device-specific errors. Since making temporal comparisons across different years was a central objective in this study, GCPs were placed using satellite imagery, such as Google Earth, as a reference. However, this base imagery may contain geometric distortions or misalignments between real-world features and the image data, introducing further uncertainty in georeferencing.

In addition, the 3D reconstruction of structures using Structure from Motion (SfM) relies on identifying and matching visual features across multiple images to estimate object geometry and position. This method fundamentally depends on the quality of image matching and the acquisition conditions, such as camera angles, lighting, and resolution. As a result, achieving precise spatial alignment is nearly impossible. Even in cases where GCPs are accurately placed and surveyed on-site, residual georeferencing errors of approximately 1–3 cm are often observed. This indicates that even with highly rigorous processing, perfect spatial alignment cannot be guaranteed.

Furthermore, differences in the spatial resolution of imagery contribute to positional discrepancies. For example, drone images from 2019 typically have a high resolution of 3–5 cm per pixel, while aerial photographs from 2009–2010 often have a coarser resolution of 25–50 cm per pixel. Consequently, the delineation of the same structure may vary based on the image resolution, particularly for small objects, leading to noticeable differences in shape overlap even if the features represent the same object.

The reasons explain the error contents; however, it is important to note that within a single dataset (e.g., only 2009, only 2010, or only 2019), where the processing conditions are consistent, measurements of spatial relationships and area are reliable. The aforementioned sources of error must be carefully accounted for only in comparisons across multiple time points.

C Determination of Clusters

C.1 Cluster Analysis

This Online Appendix section details how to classify houses into clusters, based on the dissimilarity matrix whose (i, j)-th entry is d_{ij}. First, we run single-linkage clustering to obtain the dendrogram in Figure A5a. The number of clusters differs depending on the height at which the dendrogram is horizontally cut.

[Figure A5 around here]

The next step is to determine the optimal number of clusters, accompanied by the determination of the grouping of the houses. To do so, we calculate the Calinski-Harabasz index (CH-index) under the possible numbers of clusters to find the number of clusters that returns the maximum CH-score. The CH-index is the ratio of the between-cluster variation (which measures how spread out the clusters are from each other) to the within-cluster variation (an inverse measure of how tightly the clusters are grouped). For the CH-index used in this paper, see Online Appendix Section C.2. With this definition, the number of clusters that gives the maximum CH-score should be the optimal number of clusters.

Figure A5b shows the CH-scores under different numbers of clusters for $K = 2, \cdots, 30$.[A7] From Figure A5b, the CH-index calculated with $K = 4$ achieves the local maximum but not the global maximum. The CH-score generally increases for $K \geq 10$ in the figure. In such a case, the clusters are not considered well-separated, because the CH-index does not have a global maximum. However, if an optimal number of clusters exists, the likeliest number is $K = 4$, according to Calínski and Harabasz (1974). Thus, we adopt $K = 4$ as our optimal number of clusters.

From the dendrogram shown in Figure A5a, when the number of clusters is 4, the houses are grouped as follows: cluster 1 = [1–17], cluster 2 = [18–

[A7] The CH-scores can be obtained for $K = 2, \cdots, 63$, although Figure A5b shows only $K = 2, \cdots, 30$.

46], cluster 3 = [47–52], and cluster 4 = [53–63], where the numbers in the brackets indicate the house IDs.

C.2 CH-index

This appendix supplements the CH-index used in this paper. Occasionally, for a given number of the clusters, K, the CH-index is defined as

$$CH(K) \equiv \frac{B(K)/(K-1)}{W(K)/(n-K)}, \tag{A1}$$

where n is the total number of houses. The within-cluster variation, $W(K)$, is given by $W(K) \equiv \sum_{k=1}^{K}\sum_{i=1}^{n_k}(x_i - \bar{x}_k)^2$, where n_k is the number of houses in cluster k, x_i is the position of house i, and $\bar{x}_k \equiv \frac{1}{n_k}\sum_{i=1}^{n_k} x_i$. The between-cluster variation, $B(K)$, is given by $B(K) \equiv \sum_{k=1}^{K} n_k(\bar{x}_k - \bar{x})^2$, where $\bar{x} \equiv \frac{1}{n}\sum_{i=1}^{n} x_i$. Because our cluster analysis is not directly based on the locations of houses (x_i) but on the distance between houses i and j, d_{ij}, it is useful to rewrite the CH-index using d_{ij} and circumvent x_i, $\bar{x}$, and $\bar{x}_k$ appearing in the expression. Thus, we use the following version of the CH-index, which is equivalent to (A1),

$$CH(K) \equiv \frac{[T - W(K)]/(K-1)}{W(K)/(n-K)},$$

where $T \equiv \frac{1}{n}\sum_{i=1}^{n}\sum_{j=i+1}^{n} d_{ij}^2$, and $W(K) = \sum_{k=1}^{K}\frac{1}{n_k}\sum_{i=1}^{n_k}\sum_{j=i+1}^{n_k} d_{ij}^2$. This version of the CH-index is implemented by Halpin (2016).

C.3 Segmentation Analysis in One-Dimensional Space

One might interpret the layout of houses in our study area as being arranged along a one-dimensional space. Based on this view, some may argue that interval segmentation is more appropriate than cluster analysis. To address this, we also conducted a segmentation analysis that groups houses in a way that minimizes the total within-group sum of squared deviations from the group mean, given a number of groups, $\hat{K}$, following Fisher (1958). In a segmentation analysis, each house is on a line that has a measurement.

Figure A6a shows $d_{1i}, i \in \{2, \cdots, 63\}$ on the vertical axis and the house ID on the horizontal axis.

[Figure A6 around here]

We consider a line starting from house 1. All houses are on the line and ordered by house ID. Stated differently, houses are ordered in terms of d_{1i} from small to large. By choosing $\hat{K} - 1$ boundaries from 62 intervals to minimize the sum of the within-group sum of squared deviation from the group means, houses are optimally grouped into $\hat{K}$ groups.[A8]

Next, we investigate the optimal number of groups using the elbow method. The vertical axis in Figure A6b is the sum of the within-group sum of squared deviation from the group mean for a given number of clusters (segments) $\hat{K}$. From Figure A6b, $\hat{K} = 3$ or 4 appears to be the elbow. That is, $\hat{K} = 3$ and 4 are candidates for the optimal number of segments.

Inspection of Figure A6a reveals that the houses may be grouped into four groups rather than three: the first consists of houses 1 to 17, the second consists of houses 18 to 46, the third consists of houses 47 to 53, and the fourth consists of houses 53 to 63. This is the same classification as in the single-linkage clustering, showing the robustness of classification obtained via single-linkage clustering.

D How to Construct Variables Used in Estimations

This section outlines the construction of the variables used in the empirical analysis. Specifically, it explains how we create the outside temperature variable (Section 4.4) and the inner/outer curve dummy variable used as an instrument (Section 4.6).

[A8] This procedure is implemented by a Stata user-written command "group1d."

D.1 Outside Temperature Calculation

This section describes the detailed calculation of the outside temperature variable used in the application analysis in Section 4.4. $\text{OutsideTemperature}_{itd}$ is the outside temperature in the neighborhood of house i at time t on date d. As mentioned in Section 2.4, the roof temperature data collection time for each house was at the minute level. Ideally, the outside temperature data would be collected at exactly the same location of the houses at the same time, but this was not possible due to the outside temperature data availability. We use outside temperature data recorded at 10-minute intervals from the two nearest Automated Meteorological Data Acquisition System (AMeDAS) stations, between which our study area is located.[A9]

To fit temperature data recorded by the AMeDAS to our dataset, we construct the outside temperature data via the following procedure. First, from the set of reported times in the AMeDAS data, we choose the closest time to time t, which is the time at which house i's temperature was collected by the drone. we calculate the outside temperature at the selected time as a distance-weighted average based on the direct distances from house i to the two AMeDAS stations. Specifically,

$$
\begin{aligned}
&\text{OutsideTemperature}_{itd} \\
&= \frac{\text{OutsideTemperature}_{\text{AMeDAS1},td} * \text{Distance}(\text{House}_i, \text{AMeDAS2})}{\text{Distance}(\text{House}_i, \text{AMeDAS2}) + \text{Distance}(\text{House}_i, \text{AMeDAS1})} \\
&+ \frac{\text{OutsideTemperature}_{\text{AMeDAS2},td} * \text{Distance}(\text{House}_i, \text{AMeDAS1})}{\text{Distance}(\text{House}_i, \text{AMeDAS2}) + \text{Distance}(\text{House}_i, \text{AMeDAS1})}.
\end{aligned}
$$

D.2 Calculation of the Inner/Outer-Curve Variable

This section describes how to construct the inner/outer-curve dummy variable used in Section 4.6. To construct the instrumental variable of inner/outer curves, we take the following four steps. An illustrative image of constructing the instrumental variable is depicted in Figure A10.

[A9]The AMeDAS is a collection of automatic weather stations operated by the Japan Meteorological Agency for automatic observation of precipitation, wind direction/speed, temperature, and sunshine duration to support real-time monitoring of weather conditions.

[Figure A10 around here]

In the first step, we locate points every 1 m along the river shore in our study area.[A10] In the second step, we draw straight lines connecting some of the points generated in the first step. In doing so, we admit that the newly drawn straight lines can depart from the actual river shore by 100 m.[A11] In the third step, we identify whether the points generated in the first step are located on an inner or an outer curve: (i) if the straight line connecting two points drawn in the second step runs through land area, the points between these two points (two ends of the straight line) are located on an inner curve, and (ii) if the straight line connecting two points generated in the first step runs through water, the points are on an outer curve. In the fourth step, we find the nearest point among those generated in the first step for each house and assign the nearest point's characteristic identified in the third step (i.e., whether the point is located in an inner curve or an outer curve) to each house. Finally, we obtain a dummy variable equal to 1 if a house is located on an inner curve and equal to 0 if it is located on an outer curve as the instrumental variable to implement the cluster size.

E Detailed Discussions for Section 4.8

This Online Appendix section presents the robustness checks.

[A10]Instead of the line data of the river shore provided by the National Land Numerical Information Download Service, we use data that we constructed by delineating a boundary between the water area and the land area based on the aerial photographs provided by the Geospatial Information Authority of Japan (GIAJ). This is because the GIAJ line data are more convenient due to the smoothness of the shorelines.

[A11]We also constructed an instrumental variable based on 50-m and 10-m divergences from the shoreline. However, instrumental variables based on such smaller divergence did not show sufficient relevance (i.e., the first stage F-statistics are small). In addition, simultaneously including all instrumental variables is not supported by the overidentification test. Therefore, we used only an instrumental variable based on a 100-m divergence from the river shore.

E.1 Spatial Correlation

As is briefly explained in Section 4.5, we simply clustered the standard errors by houses, not by clusters because the wild cluster bootstrap for few clusters (we only have four clusters) is considered inapplicable in our situation, according to Canay et al. (2021). First, for the wild cluster bootstrap to work properly, relative sizes of clusters should not differ dramatically. This assumption, however, may not be satisfied in our case since we exploit the variation of sizes of clusters. The second assumption for the wild cluster bootstrap is that $N(i)_{-i}$ should vary within the same cluster, which is not obviously satisfied in our case as the cluster size variable is identical within the same cluster.

Instead of clustering the standard errors by cluster, we use Conley (1999) standard errors, allowing for spatial correlation within a certain cut-off distance and infinite serial correlation. For the cut-off distance, we use $0.183km$, distance between houses 47 and 52, shortest cluster geographic length among four clusters, and $0.992km$, distance between houses 53 and 63, the longest cluster geographic length. Unfortunately, the option to estimate the RE model is not available in the 'acreg" Stata package that is used to perform the Conley (1999) correction, so we report the POLS and IV results in Table A5.

[Table A5 around here]

The results on $N(i)_{-i}$ and $DN(i)_{-i}$ are robust.

E.2 Houses with Roofs Fully Visible

As is already mentioned in Section 2.4, some of the houses have roofs covered by tree edges and leaves. In the baseline analysis, to maintain the sample size, we constructed a roof temperature variable by averaging thermal information over the visible area of the house polygon from the sky, and we did not drop houses whose roofs are partly covered by vegetation from the sample. In this fashion, however, the invisible roof area may show much higher or lower temperature than the visible area does, leading to an imprecise measure of the roof temperature. To overcome this, we only exploit the houses whose roofs are fully visible in the daytime aerial pictures, at the expense of losing observations.

Table A6 shows the result.

[Table A6 around here]

After limiting the sample to houses with fully visible roofs, the number of observations reduces to 111 from 160, which is a substantial decline. Despite this drop of the number of observations, the result displayed in Table A6 corroborates the significantly positive effect of the cluster size (affluence) on the roof temperature.

E.3 Different Choice of Roof Temperature Variable

Despite its instability as a measure of roof temperature compared to the average value of the temperatures within house polygons, the maximum value of the roof temperature for each house polygon, such as the most reddish part in Figure 1b, may better represent the inside temperature.[A12] To see if the results remain unchanged, we adopt the highest value of roof temperature instead of the average value as our dependent variable. To alleviate the concern that the part not covered by edges and leaves shows the highest temperature over the roof polygon, we conduct the same regression based on the sample of houses with fully visible roofs.

Table A7 shows the results.

[Table A7 around here]

[A12]For instance, if heat-producing objects such as birds and small animals are on the roof, the maximum inside roof temperature measured from outside may capture these outside heat sources. Moreover, when the analysis is based on the full sample, employing the highest temperature over the house polygon may be less adequate because of the possibility that a roof part hidden by leaves and edges reveals the highest roof temperature of a house. To circumvent such mistakes in collecting roof temperature data, we use the average temperature value rather than an extreme value, such as the highest temperature within a house polygon.

Columns (1)–(4) and (9)–(12) show results using the full sample, while columns (5)–(8) and (13)–(16) focus on houses with fully visible roofs. The cluster-affluence variables, $ST(i)_{-i}$ and $DST(i)_{-i}$, are recalculated on the basis of the maximum roof temperatures. In the visible-roof sample, roof temperatures are set as missing if vegetation partially covers the roof. While results from the full sample are occasionally noisy as expected, those from the fully visible subset yield consistently positive and significant coefficients, supporting the robustness of the findings.

We further consider another possibility of the roof temperature measure that takes into account the house size. In Section 4.7, we separately ran regressions based either on the roof temperature or roof footprint as the dependent variables. This was to illuminate what different economic aspects these two measures capture. However, some may think that a roof temperature measure that takes into account the house size would express the current living standards, because larger houses are more difficult to heat up, needing more use of heat-generating products.

To answer this possibility, we calculate two different measures of the house temperatures: (i) "Roof temperature (°C)"×"Roof footprint (m^2)" and (ii) "Roof temperature (°C)"×"Roof footprint (m^2)"×"House height (m)". Columns (1)–(4) show the results based on the roof temperature variable taken into account the house floor area (roof footprint), and columns (5)–(8) display those based on the roof temperature variable taken into account the house volume.

[Table A8 around here]

The results shown in Table A8 are robust because estimated coefficients of $N(i)_{-i}$ and $DN(i)_{-i}$ are significantly positive.

E.4 Number of Supermarkets

This robustness check accounts for the effect of the number of supermarkets near the houses. As mentioned in Section 2, homeless persons in our study area rarely beg for food or money. Instead, they earn money and buy food on their own. Moreover, it is rare for free meals to be provided by the riverbank in our study area. DHWKC (2019) report that one of the worries for homeless persons is food deficiency. Better access to supermarkets (especially those selling food at discounted prices) in their neighborhood improves their living standards by decreasing food expenditures.[A13] We consider the aspect of access to (inexpensive) food by adding a variable representing the number of supermarkets in a 1-km radius from the house centroids.

[Table A9 around here]

In all specifications in columns (1)–(4), the number of supermarkets near houses does not affect the current living standards of homeless persons. This may be because our target area is highly urbanized, and there is an abundance of supermarkets, which leads to an insufficient variation in the number of supermarkets. Moreover, homeless persons may purchase food in supermarkets that sell items at discounted prices. The number of supermarkets alone does not capture which supermarkets homeless persons choose to buy food from at low prices. For the cluster-size variables, the positive coefficients of $N(i)_{-i}$ (column (2)) and $DN(i)_{-i}$ (column (4)) are corroborated.

E.5 Number of Scrap Firms

On the income side, we include the number of scrap firms within a 2-km radius from the house centroids. As mentioned in Section 2, homeless persons' incomes mainly come from illegally collecting cans and metal scrap and conducting illicit transactions with scrap firms. We expect that the number of scrap firms near the houses of homeless persons to have a positive association with living standards.

Columns (5)–(8) in Table A9 display the results with an additional control variable, the number of scrap firms within a 2-km radius from each house centroid. The number of scrap firms has almost no impact on the living standards of homeless persons. This insignificance may be caused by the

[A13]Indeed, Suzuki (2008) interpret a higher density of supermarkets as better access to food for homeless persons in Osaka, Japan.

closed transactions of used cans and metal scraps. Even though homeless persons earn money by selling the cans and metal they illegally collect to scrap firms, such informal transactions may be conducted with only a small number of scrap firms. In addition, information about which firms conduct informal deals with homeless persons is not open. Therefore, we cannot detect the effect of access to such illicit scrap buyers. As for the cluster-size effect, even after controlling for the number of scrap firms, the significantly positive effect of larger clusters on roof temperature remains.

E.6 Access to Jobs in the Construction Sector

We consider another possible type of earnings. The majority of homeless persons rely on earnings from can or metal scrap collection, but some earn money by temporarily working in the construction sector. To take into account this possibility, we include the distance to X Town in the model ("X" is a dummy location name to keep the anonymity of the study area). X Town provides informal job opportunities for daily construction workers and is approximately 7 km from the southeast end of our study area. We expect a negative sign for the variable distance to X Town because better access (shorter distance) to X Town should lead to more job opportunities for homeless persons engaged in the construction sector.

Columns (9)–(12) in Table A9 displays the results of this specification. The coefficients of the distance to X Town (inverse access to jobs in the construction sector) are positive, against our expectation, although they are insignificant in the IV results. The positive sign of the coefficient means that the farther from X Town homeless persons' houses are, the better off the individuals are. There are two explanations for this unexpected sign. First, homeless persons in the study area may not find daily jobs in X Town. Inexpensive hotels for homeless daily construction workers are located near X Town, and daily jobs offered around X Town are supplied to homeless persons staying in such hotels, not to those living in our study area. Thus, homeless persons in our study area may not find job opportunities in X Town.

Second, there are gambling facilities located near X Town. Some of the homeless persons in Japan suffer from gambling addiction, and being close to an area with abundant gambling facilities may worsen homeless persons' situations because they waste money in an uneconomical manner. Thus, being close to X Town may negatively affect roof temperatures. Our featured variables, $N(i)_{-i}$ in column (10) and $DN(i)_{-i}$ in column (12), have a significant positive effect on roof temperature in this robustness check.

E.7 Vegetation Coverage

In the baseline model, we choose a 5-m buffer range for vegetation coverage. In this robustness check, we change the buffer range to 10 m and 15 m. Table A10 shows the results for different vegetation buffers (10-m buffer for columns (1)–(4) and 15-m buffer for columns (5)–(8)).

[Table A10 around here]

Unlike in the case of a 5-m buffer, the forest coverage variable with the 15-m buffer is insignificant, but the forest coverage of the 10-m buffer is still significant. In terms of the cluster-size variables, the positive impact of $N(i)_{-i}$ and $DN(i)_{-i}$ (columns (2) and (4), respectively) remains effectively unchanged.

E.8 Wind Speed Effect

We consider the effect of wind speed at the time of the collection of roof temperature data. As with the calculation of outside temperature, we compute wind speed as follows:

$$\begin{aligned}
&\mathrm{WindSpeed}_{itd} \\
&= \frac{\mathrm{WindSpeed}_{\mathrm{AMeDAS1},td} * \mathrm{Distance}(\mathrm{House}_i, \mathrm{AMeDAS2})}{\mathrm{Distance}(\mathrm{House}_i, \mathrm{AMeDAS2}) + \mathrm{Distance}(\mathrm{House}_i, \mathrm{AMeDAS1})} \\
&+ \frac{\mathrm{WindSpeed}_{\mathrm{AMeDAS2},td} * \mathrm{Distance}(\mathrm{House}_i, \mathrm{AMeDAS1})}{\mathrm{Distance}(\mathrm{House}_i, \mathrm{AMeDAS2}) + \mathrm{Distance}(\mathrm{House}_i, \mathrm{AMeDAS1})}.
\end{aligned}$$

We expect a negative sign of the wind speed coefficient on roof temperature. Table A11 displays the results of estimations with the baseline variables and the wind speed variable.

[Table A11 around here]

As expected, the wind speed variable is negative but insignificant in both specifications. This result is reasonable given the drone flight conditions. As drone flights are prohibited on windy days, and the flights to collect roof temperature data were conducted under light wind, it is natural that the wind speed variable does not affect roof temperatures. Turning to the coefficients on the social interaction effects within a house cluster, the results remain unchanged: larger cluster sizes ($N(i)_{-i}$ in column (2) and $DN(i)_{-i}$ in column (4)) raise roof temperatures, which suggests that social interaction among homeless persons benefits them in large clusters.

E.9 Effect of Bridges

As noted in Section 2.4, we were unable to collect data for houses located beneath bridges, as drone flights are not permitted under bridge structures.[A14] However, the houses of homeless persons may be located underneath bridges and are thus invisible in the aerial images. To address this point, we exclude houses close to bridges and repeat the estimation. In addition, we conduct an analogous estimation by dropping all houses in a cluster that contains houses sufficiently close to bridges.

The distance between adjacent houses differs across pairs of houses, and the average distance between adjacent houses within a cluster also differs across clusters. Thus, it is not appropriate to set the same distance to create a bridge buffer for all houses and clusters. Instead, we employ the following distance measure to create buffers around bridges:

$$d_{C(i)}^{\max} \equiv \max_{k \in C(i) \setminus \{\bar{k}_{C(i)}\}} d_{k,k+1}, \tag{A2}$$

[A14]Figure A9d shows bridges (both pedestrian and non-pedestrian bridges) in the study area.

where $\bar{k}_{C(i)}$ is the maximum house ID in cluster $C(i)$, e.g., $\bar{k}_1 = 17$ for cluster 1, $\bar{k}_2 = 46$ for cluster 2, and so on. In (A2), $d_{C(i)}^{\max}$ expresses the maximum distance between two adjacent houses belonging to the same cluster $C(i)$. The choice of the buffer (A2) is explained as follows. By definition and the merging procedure of the cluster analysis, for every cluster $C(i)$, the distance between two arbitrarily chosen adjacent houses belonging to cluster $C(i)$ is shorter than or equal to $d_{C(i)}^{\max}$. Therefore, if the distance between a bridge edge and the house closest to the bridge among houses in cluster $C(i)$, house $i_{C(i)}^{\text{bridge}}$, is longer than $d_{C(i)}^{\max}$, there should be no house located underneath the bridge, given the clusters optimally detected in the cluster analysis. By contrast, if the distance between house $i_{C(i)}^{\text{bridge}}$ and the bridge edge is shorter than $d_{C(i)}^{\max}$, there is a possibility that a house is hidden underneath the bridge. In such a case, the cluster to which house $i_{C(i)}^{\text{bridge}}$ belongs may contain a larger number of houses than originally observed because of the hidden houses underneath the bridge. Thus, we drop such cluster(s) and repeat the analysis using the baseline model. Specifically, we run regressions by dropping cluster $C(i)$ such that

$$\min_{k \in C(i)} d_{k,\text{bridge}} < d_{C(i)}^{\max}, \tag{A3}$$

where $\min\limits_{k \in C(i)} d_{k,\text{bridge}}$ is the distance between the bridge edge and the house closest to the bridge among houses in cluster $C(i)$.

We also conduct a less straightforward but similar analysis to the one stated above. Instead of dropping all houses belonging to the cluster(s) satisfying (A3), we drop houses that are sufficiently close to the bridge edge. Namely, we drop house i, a member of cluster $C(i)$, such that

$$d_{i,\text{bridge}} < d_{C(i)}^{\max}. \tag{A4}$$

Columns (1)–(4) in Table A12 show the results when houses satisfying (A4) are dropped, and columns (5)–(8) are the results when the cluster(s) satisfying (A3) is dropped.[A15]

[A15] In our dataset, only cluster 4 satisfies (A3), so the number of clusters dropped from the analysis is one. Other clusters belong to the same administrative division, so the administrative fixed effect is not included in the model in Table A12.

[Table A12 around here]

The results are effectively the same for both approaches to dropping observations near bridges; that is, cluster size has a positive effect on the current living standards of homeless persons.

E.10 Effect of Houses on the Opposite Shore

As displayed in Figure A9c, houses of homeless persons are located on opposite shores of the river, implying that an interaction with homeless persons located on the opposite shore can occur. Because meeting with individuals located on the opposite shore is possible only by crossing a pedestrian bridge, we exclude houses sufficiently close to pedestrian bridges in the study area. As shown in Figure A9d, there are three pedestrian bridges in the study area, so we drop samples within 500 meters of each pedestrian bridge, which leaves 137 observations.[A16]

Columns (9)–(12) in Table A12 report the results. Although in the specification with $N(i)_{-i}$, the first-stage F-statistic is not sufficiently high in columns (10) and (12), we find a positive effect of cluster size on roof temperature. Excluding observations that may have interaction with homeless persons on the other side of the river does not change the main result.

E.11 Definition of Oldcomers and Newcomers

In Section 4.7, we classified houses as oldcomers or newcomers based on whether past and present house polygons overlapped by more than 70%. As

[A16]In this robustness check, the house distribution on the shore of our study area and that on the other side of the river are important. Unfortunately, we did not collect house location data on the opposite shore in 2019. To construct buffer lengths from pedestrian bridges, as we did in Section E.9, house location data on the other side of the river, which we do not have, are indispensable. Instead, we use a fixed value for the buffer length for all houses. As the river width in our study area is approximately 300–400 m, we set the buffer length to 500 m in this robustness check. In contrast to Section E.9, we conduct a robustness check by dropping houses that are located near pedestrian bridges, but we cannot implement a robustness check by dropping clusters that contain at least one house that is within 500 m of a pedestrian bridge. This is because the latter robustness check based on dropping clusters as a whole leaves only cluster 3, which contains 6 houses.

a robustness check, we raise this threshold to 80%, redefine the classifications accordingly, and re-estimate the same set of regressions.

Table A13 displays the result.

[Table A13 around here]

In the roof footprint regressions shown in columns (1)–(3), the coefficient on the oldcomer variable remains significantly positive, consistent with the findings in Table 4.7. Additionally, the coefficients on the cluster-size variables, $N(i)_{-i}$ and $DN(i)_{-i}$, are also significantly positive, further confirming the results reported in Table 4.7.

Turning to the roof temperature regressions without interaction terms, shown in columns (4)–(7), the oldcomer variable is negative and marginally significant at the 10% level. Although the coefficients were insignificantly negative in Table 4.7, the overall message remains consistent: roof footprint and roof temperature capture different economic dimensions, particularly when distinguishing between oldcomers and newcomers. Regarding the cluster-size variables, the coefficients on $N(i)_{-i}$ and $DN(i)_{-i}$ are significantly positive across columns (4)–(9), indicating strong robustness. Finally, the interaction terms are negative and insignificant in the $N(i)_{-i}$ model, and negative but marginally significant at the 10% level in the $DN(i)_{-i}$ model—patterns that closely align with those reported in Table 4.7.

Figures to Online Appendix

Figure A1: Typical Houses of Homeless Individuals in the Study Area

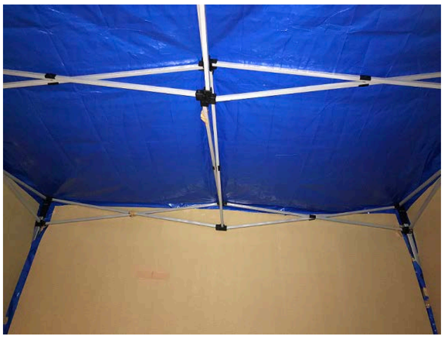
Figure A1 is from an article "There is a struggle to find a safe place to sleep for the homeless" (in Japanese), Murata (2019) published in Toyo Keizai Online, `https://toyokeizai.net/articles/-/285068?page=7`, accessed on January 12, 2023.

Figure A2: Toilet & Tap on the River Bank

Figure A2 is a picture taken by the authors in the study area in February, 2020.

(a) Exterior of Replica House

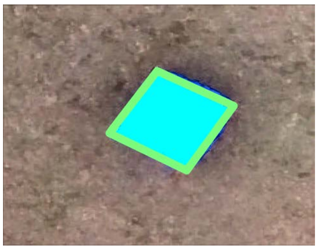
(b) Interior of Replica House



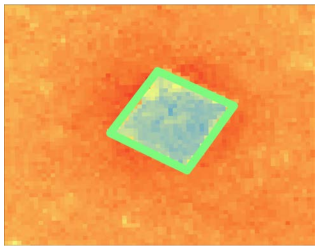
(c) Ceiling of Replica House

Figure A3: Pictures of Thermal Experiment

Figure A3a is a picture of the house taken by the authors in the thermal experiment on December 9, 2022. Figure A3b is a picture inside the house taken by the authors in the thermal experiment on December 9, 2022. Two mobile volume batteries and small heaters were the heat origin. Figure A3c is a picture of the ceiling of the house taken by the authors in the trial thermal experiment on February 28, 2022. Note that the house structure in the trial experiment is the same as in the thermal experiment conducted on December 9, 2022.



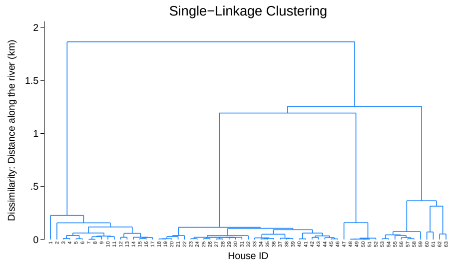
(a) Aerial Picture in Experiment



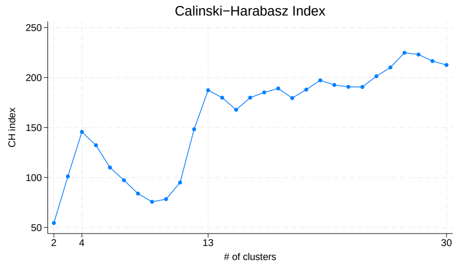
(b) Thermal Image in Experiment

Figure A4: Aerial & Thermal Images in Experiment

Figure A4a: Polygon marked with light green lines is the replica house built for the thermal experiment on December 9, 2022. Figure A4b: Polygons marked with light green lines correspond to the replica house in Figure A4a. Reddish (bluish) area indicates higher (lower) temperatures. Temperature is measured in °C. Data was collected at 19:15 on December 9, 2022.

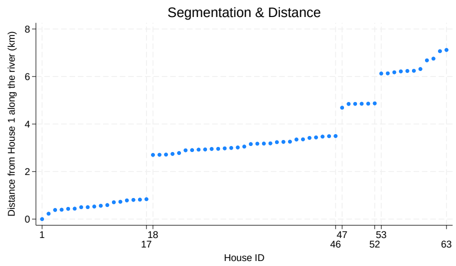


(a) Single-Linkage Clustering: Dendrogram

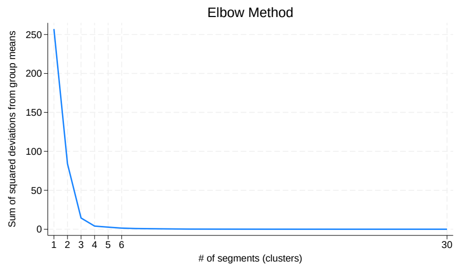


(b) Single-Linkage Clustering: CH-Index

Figure A5: Single-Linkage Clustering & the Optimal Number of Clusters

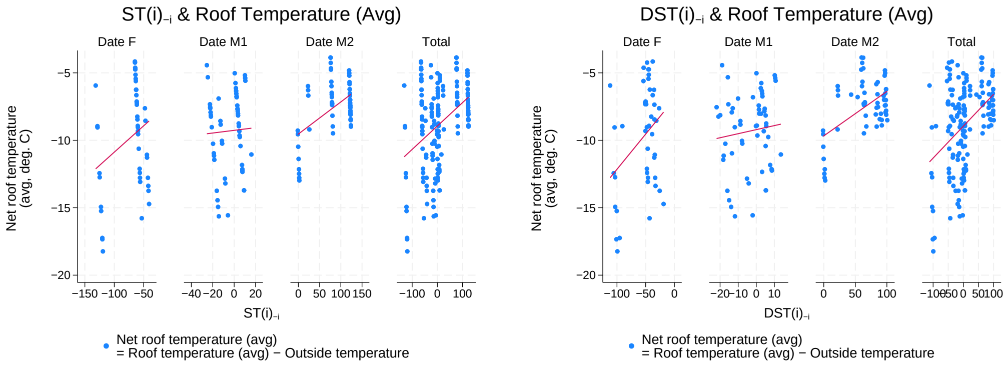


(a) Segmentation: Distance to House 1



(b) Segmentation: Elbow Method

Figure A6: Segmentation & the Optimal Number of Clusters

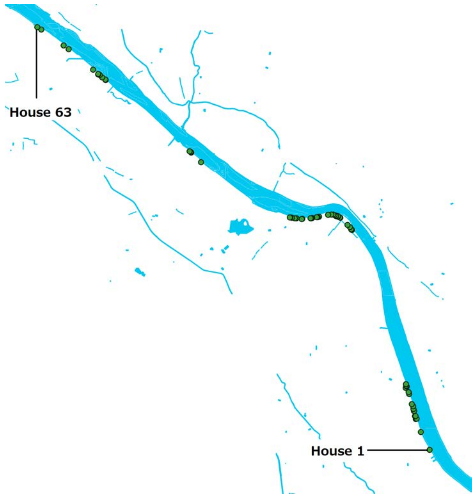


(a) $ST(i)_{-i}$

(b) $DST(i)_{-i}$

Figure A7: Cluster Affluence & Roof Temperature

Both in Figures A7a and A7b, the vertical axis expresses "Net roof temperature" defined by "Roof temperature recorded by the drone" minus "Outside temperature on the record time". Red lines are linear fits.

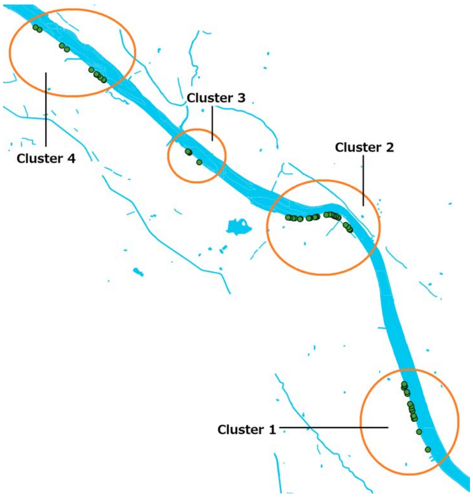


(a) House Centroid Distribution

(b) House Distribution: Oldcomers vs Newcomers

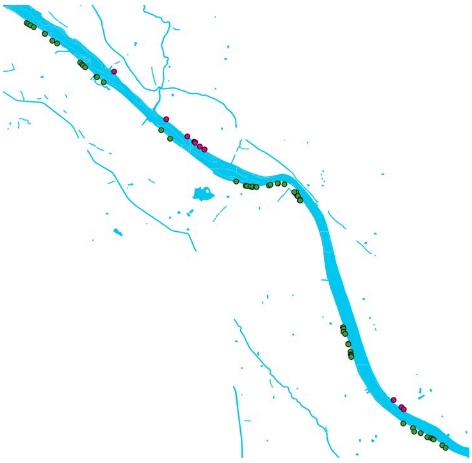
Figure A8: Maps of the Study Area

In all figures, light blue ribbon running from the southeast to the northwest is the river. Other light blue lines are ditches. Figure A8a: Green points depict houses of the homeless persons on the side of the study area in February and March 2019. Figure A8b: Pink crosses depict oldcomer houses. Green points depict newcomer houses.

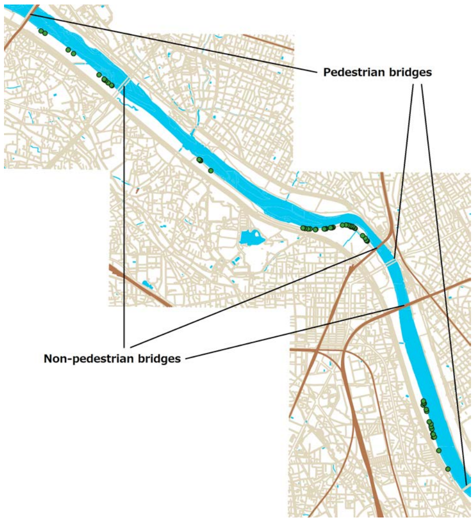


(a) House Centroids & Clusters

(b) Houses & Toilets/Taps



(c) Houses & Inner/Outer Curves: Both Sides of the River in 2020



(d) Houses & Bridges

Figure A9: Maps of the Study Area: Application Analysis

In all figures, light blue ribbon running from the southeast to the northwest is the river. Other light blue lines are ditches. Green points depict houses of the homeless persons on the side of the study area in February and March 2019. Figure A9a: House centroids depicted by green points are enclosed by orange circles form each cluster. Figure A9b: Yellow points depict toilets/taps. Figure A9c: Pink points depict houses of the homeless persons on the other side. Figure A9d: Brown lines depict train rails. Beige lines depict roads.

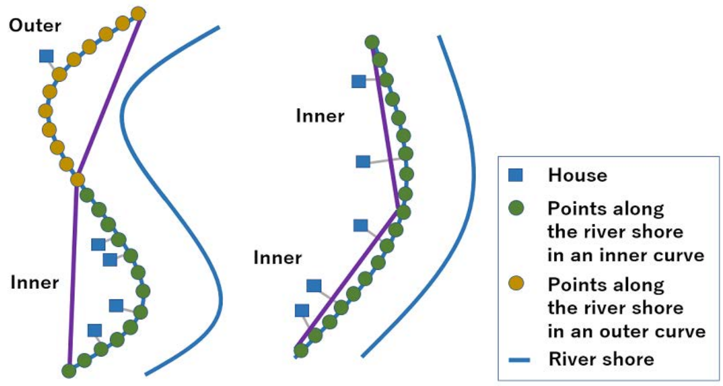


Figure A10: Image of Constructing the Instrumental Variable

Tables to Online Appendix

Table A1: Data Source

Variable	Source	
House Temperature (2019)	Own collection and calculation	
House location (2019)	Own collection	
Old-newcomer houses (2009-2010)	Own detection from the aerial photos of the Geospatial Information Authority of Japan	Purchased from Japan Map Center (*Nihon Chizu Sentaa*) `https://net.jmc.or.jp/`
Experimental temperature data	Own collection and calculation	
Outside temperature:		
Air temperature (2019)	Japan Meteorological Agency (JMA)	`http://www.data.jma.go.jp/obd/stats/etrn/index.php`
Distance from the AMeDAS to houses	Own calculation	
$N(i), N(i)_{-i}$	Own calculation from the house location data	
$DN(i), DN(i)_{-i}$	Own calculation from the house location data	
$ST(i), ST(i)_{-i}$	Own calculation from the house location and temperature data	
$DST(i), DST(i)_{-i}$	Own calculation from the house location and temperature data	
Inner/outer curve (1 if inner, 0 if outer)	Own calculation from Geospatial Information Authority of Japan (GIAJ)	`https://www.gsi.go.jp/ENGLISH/index.html`
Distance to toilet/tap (m)	Own calculation from the house location data and the GIAJ data	
Forest coverage	Own collection and calculation	
Grass coverage	Own collection and calculation	
Distance to the river shore (m)	Own calculation from the house location data and the line data of the shore of Tama River provided by the National Land Numerical Information Download Service	
Distance to X Town (km)	Own calculation from the house location data and the GIAJ data	
House floor area	Own collection and calculation	
Past wind speed:	Japan Meteorological Agency	
Past wind speed	Japan Meteorological Agency (JMA)	
Distance from the AMeDAS to houses	Own calculation	
Number of supermarkets ($1km$):	Own data transformation and calculation from the data of NTT Hello Page	
Data transformation	Address matching service provided by Center for Spatial Information Science (CSIS), University of Tokyo	`http://newspat.csis.u-tokyo.ac.jp/geocode/`
Number of scrap firms ($2km$):	Own data transformation and calculation from the data of NTT Hello Page	
Data transformation	Address matching service provided by CSIS	
Distance from bridges	Own calculation from the line data of the rail roads and roads provided by the National Land Numerical Information Download Service	`http://nlftp.mlit.go.jp/ksj-e/index.html`

Table A2: Summary Statistics

	Obs.	Mean	Std.Dev	Min	Max
Individual house variables					
House temperature (avg, deg. C)	160	0.219	4.301	-12.800	7.500
House temperature (max, deg. C)	160	2.112	3.864	-11.100	9.300
House temperature (house area taken into account)	160	1.760	52.722	-176.582	131.860
House temperature (house area + height taken into account)	111	-13.540	112.540	-356.802	270.313
House floor area (m^2)	160	11.594	7.415	2.437	35.638
House height (m)	111	1.762	0.372	1	2
Oldcomer (height > 0.9m, overlap > 70%)	160	0.356	0.480	0	1
Cluster size variables					
$N(i)_{-i}$	160	20.850	7.995	5	28
$DN(i)_{-i}$	160	15.873	6.174	4.228	22.574
$ST(i)_{-i}$ (avg)	160	8.079	70.811	-131.590	123.857
$DST(i)_{-i}$ (avg)	160	6.526	54.721	-111.898	100.694
$ST(i)_{-i}$ (max)	160	45.026	67.164	-93.700	158.171
$DST(i)_{-i}$ (max)	160	34.285	51.880	-80.230	127.621
$ST(i)_{-i}$ (avg, roof fully visible)	160	-7.511	67.654	-146.748	100.147
$DST(i)_{-i}$ (avg, roof fully visible)	160	-5.326	51.163	-113.003	79.873
$ST(i)_{-i}$ (max, roof fully visible)	160	27.347	62.614	-107.112	131.136
$DST(i)_{-i}$ (max, roof fully visible)	160	20.642	47.468	-81.167	104.229
Other variables					
Inner/outer curve (1 if inner, 0 if outer)	160	0.487	0.501	0	1
Outside temperature (deg. C)	160	9.022	2.218	5.438	11.796
Distance to toilet/tap water (m)	160	310.229	153.581	58.327	632.955
Forest coverage (5m, %)	160	22.184	23.554	0	100
Forest coverage (10m, %)	160	20.914	18.582	0	84.753
Forest coverage (15m, %)	160	16.727	14.788	0	63.718
Grass coverage (5m, %)	160	62.391	22.968	0	100
Grass coverage (10m, %)	160	64.465	19.465	9.280	97.844
Grass coverage (15m, %)	160	64.460	18.999	8.735	97.835
Distance to the river shore (m)	160	18.745	8.681	4.875	42.867
Distance to X Town (km)	160	7.344	1.738	4.594	10.969
Wind speed (m/s)	160	3.964	0.748	2.275	5.116
Number of supermarkets (1km)	160	10.444	2.601	4	15.000
Number of scrap firms (2km)	160	8.919	3.715	3	13.000

Table A3: Baseline: Full Result

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Model:	POLS	RE	REIV	POLS	RE	REIV	POLS	RE	REIV	POLS	RE	REIV
Dependent variable:	Roof temperature (deg. C)											
$N(i)_{-i}$	0.100**	0.092**	0.195***									
	(0.038)	(0.037)	(0.067)									
$DN(i)_{-i}$				0.134**	0.121**	0.267***						
				(0.053)	(0.051)	(0.096)						
$ST(i)_{-i}$							0.014**	0.017***	0.124**			
							(0.006)	(0.005)	(0.063)			
$DST(i)_{-i}$										0.027***	0.030***	0.132**
										(0.009)	(0.008)	(0.063)
Outside temperature (deg. C)	3.402***	2.947***	2.863***	3.486***	2.982***	2.925***	2.464***	1.617***	-7.262	1.902***	0.996*	-5.810
	(0.645)	(0.521)	(0.505)	(0.672)	(0.532)	(0.520)	(0.666)	(0.525)	(5.345)	(0.607)	(0.535)	(4.357)
Distance to toilet/tap water (m)	-0.005**	-0.005***	-0.007***	-0.004**	-0.004**	-0.006***	-0.003**	-0.004**	-0.008***	-0.003**	-0.003**	-0.005***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.001)	(0.002)
Forest coverage ($5m$, %)	0.072***	0.072***	0.069***	0.071***	0.072***	0.067***	0.075***	0.075***	0.073***	0.074***	0.074***	0.071***
	(0.016)	(0.016)	(0.015)	(0.016)	(0.016)	(0.015)	(0.017)	(0.017)	(0.020)	(0.017)	(0.017)	(0.018)
Grass coverage ($5m$, %)	0.026	0.027	0.025	0.027	0.028	0.025	0.029	0.029	0.026	0.028	0.029	0.025
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.018)	(0.018)	(0.021)	(0.018)	(0.018)	(0.019)
Distance to the river shore (m)	-0.067**	-0.064**	-0.100**	-0.072**	-0.068**	-0.111**	-0.041	-0.041	-0.103	-0.048	-0.048*	-0.105*
	(0.031)	(0.028)	(0.039)	(0.032)	(0.029)	(0.045)	(0.032)	(0.029)	(0.065)	(0.031)	(0.029)	(0.061)
Date FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Administrative division FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
# of observations	160	160	160	160	160	160	160	160	160	160	160	160
Adj R-sq	0.72			0.71			0.71			0.71		
Within R-sq		0.82			0.82			0.82			0.83	
Between R-sq		0.65			0.64			0.62			0.62	
Breusch and Pagan LM χ^2		24.79			24.66			30.11			31.69	
Sargan-Hansen stat. χ^2		5.61			6.52			9.21			13.29	
First stage												
Inner/outer curve			9.443***			6.886***			14.697**			13.833***
			(1.766)			(1.392)			(5.749)			(4.860)
Kleibergen-Paap F-stat.			28.60			24.46			6.53			8.10

Standard errors clustered by houses are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A4: Roof Footprint vs Roof Temperature: Full Result

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Model:	OLS	OLS	OLS	RE	REIV	RE	REIV	RE	RE
Dependent variable:	Log (Footprint, m^2)			Roof temperature (deg. C)					
Oldcomer	0.592***	0.481***	0.483***	-0.292	-0.664	-0.292	-0.695	-0.196	0.109
	(0.172)	(0.176)	(0.179)	(0.580)	(0.625)	(0.575)	(0.639)	(1.569)	(1.621)
$N(i)_{-i}$		0.031***		0.099***	0.219***			0.101**	
		(0.010)		(0.038)	(0.069)			(0.044)	
$DN(i)_{-i}$			0.041***			0.131**	0.302***		0.139**
			(0.015)			(0.052)	(0.102)		(0.057)
Oldcomer$\times N(i)_{-i}$								-0.005	
								(0.070)	
Oldcomer$\times DN(i)_{-i}$									-0.026
									(0.095)
Outside temperature (deg. C)				2.942***	2.855***	2.979***	2.922***	2.937***	2.977***
				(0.520)	(0.503)	(0.531)	(0.519)	(0.518)	(0.532)
Distance to toilet/tap water (m)	0.000	-0.000	-0.000	-0.005***	-0.008***	-0.005***	-0.007***	-0.005***	-0.005**
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Forest coverage ($5m$, %)	-0.000	-0.002	-0.003	0.071***	0.065***	0.070***	0.063***	0.070***	0.069***
	(0.004)	(0.004)	(0.005)	(0.016)	(0.015)	(0.016)	(0.015)	(0.017)	(0.017)
Grass coverage ($5m$, %)	-0.007	-0.007	-0.007	0.028	0.026	0.028*	0.026	0.028	0.027
	(0.005)	(0.005)	(0.005)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Distance to the river shore (m)	0.000	-0.012	-0.013	-0.067**	-0.110***	-0.071**	-0.123***	-0.066**	-0.070**
	(0.009)	(0.009)	(0.008)	(0.028)	(0.040)	(0.029)	(0.046)	(0.029)	(0.029)
Date FE	-	-	-	Y	Y	Y	Y	Y	Y
Administrative division FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
# of observations	63	63	63	160	160	160	160	160	160
Adj R-sq	0.11	0.16	0.16						
Within R-sq				0.82		0.82		0.82	0.82
Between R-sq				0.65		0.65		0.65	0.65
Breusch and Pagan LM χ^2				24.60		24.45		24.56	24.24
Sargan-Hansen stat. χ^2				5.69		6.61		5.78	6.70
First stage									
Inner/outer curve					9.100***		6.602***		
					(1.799)		(1.381)		
Kleibergen-Paap F-stat.					25.58		22.86		

Standard errors clustered by houses are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A5: Standard Errors Allowing for Spatial Correlation (Conley)

	(1)	(2)
Model:	POLS	IV
Dependent variable:	Roof temperature (deg. C)	
$N(i)_{-i}$ (Coef.)	0.100	0.207
(Conley Std. Err.; cut-off = 0.183 *km*)	(0.052)*	(0.089)**
(Conley Std. Err.; cut-off = 0.992 *km*)	[0.054]*	[0.092]**
Base control	Y	Y
Date FE	Y	Y
Administrative division FE	Y	Y
# of observations	160	160
R-sq	0.73	
First stage Kleibergen-Paap F-stat. (weak instrument)		
(Conley; cut-off = 0.183 *km*)		13.99
(Conley; cut-off = 0.992 *km*)		21.15
	(3)	(4)
Model:	POLS	IV
Dependent variable:	Roof temperature (deg. C)	
$DN(i)_{-i}$ (Coef.)	0.134	0.287
(Conley Std. Err.; cut-off = 0.183 *km*)	(0.066)**	(0.124)**
(Conley Std. Err.; cut-off = 0.992 *km*)	[0.075]*	[0.132]**
Base control	Y	Y
Date FE	Y	Y
Administrative division FE	Y	Y
# of observations	160	160
R-sq	0.73	
First stage Kleibergen-Paap F-stat. (weak instrument)		
(Conley; cut-off = 0.183 *km*)		15.60
(Conley; cut-off = 0.992 *km*)		16.61

Conley (1999) standard errors in parentheses (in brackets, resp.), allowing for spatial correlation within a 0.183-*km* (0.992-*km*) radius and for infinite serial correlation, are reported. Base control = Outside temperature, Distance to toilet/tap water, Forest and grass coverage (5-*m* buffers), Distance to the river shore. Because the option to estimate the RE model is not available in the 'acreg" Stata package that is used to perform the Conley (1999) correction, we report the POLS and IV results in the table.
*** p<0.01, ** p<0.05, * p<0.1

Table A6: Houses With Roofs Fully Visible

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Model:	RE	REIV	RE	REIV	RE	REIV	RE	REIV
Sample:	Houses with roofs fully visible							
Dependent variable:	Roof temperature (deg. C)							
$N(i)_{-i}$	0.212***	0.284***						
	(0.055)	(0.082)						
$DN(i)_{-i}$			0.295***	0.392***				
			(0.086)	(0.121)				
$ST(i)_{-i}$ (roof fully visible)					0.023***	0.171**		
					(0.008)	(0.079)		
$DST(i)_{-i}$ (roof fully visible)							0.042***	0.195**
							(0.016)	(0.089)
Date FE	Y	Y	Y	Y	Y	Y	Y	Y
Administrative division FE	Y	Y	Y	Y	Y	Y	Y	Y
# of observations	111	111	111	111	111	111	111	111
Within R-sq	0.84		0.83		0.84		0.84	
Between R-sq	0.60		0.60		0.49		0.50	
Breusch and Pagan LM χ^2	12.54		11.92		18.71		19.78	
Sargan-Hansen stat. χ^2	3.52		6.72		18.92		14.39	
First stage								
Inner/outer curve		10.054***		7.306***		16.576***		14.572***
		(1.416)		(1.119)		(5.035)		(4.238)
Kleibergen-Paap F-stat.		50.42		42.64		10.84		11.83

Standard errors clustered by houses are in parentheses. Base control = Outside temperature, Distance to toilet/tap water, Forest and grass coverage (5-m buffers), Distance to the river shore.
*** p<0.01, ** p<0.05, * p<0.1

Table A7: Maximum Roof Temperature Within a House Polygon

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Model:	RE	REIV	RE	REIV	RE	REIV	RE	REIV	RE	REIV	RE	REIV	RE	REIV	RE	REIV
Sample:	Full				Houses with roofs fully visible				Full				Houses with roofs fully visible			
Dependent variable:	Roof temperature (maximum, deg. C)															
$N(i)_{-i}$	0.064	0.149**			0.180***	0.255***										
	(0.040)	(0.067)			(0.057)	(0.076)										
$DN(i)_{-i}$			0.065	0.204**			0.242***	0.352***								
			(0.059)	(0.097)			(0.083)	(0.114)								
$ST(i)_{-i}$ (max)									0.015***	0.050*						
									(0.004)	(0.026)						
$DST(i)_{-i}$ (max)											0.026***	0.059*				
											(0.008)	(0.032)				
$ST(i)_{-i}$ (max, roof fully visible)													0.027***	0.085**		
													(0.010)	(0.034)		
$DST(i)_{-i}$ (max, roof fully visible)															0.046**	0.104**
															(0.018)	(0.043)
Outside temperature (deg. C)	3.754***	3.695***	3.770***	3.728***	4.151***	4.019***	4.255***	4.142***	2.451***	-0.586	1.986***	-0.355	1.990*	-3.229	1.216	-2.874
	(0.541)	(0.528)	(0.549)	(0.537)	(0.588)	(0.587)	(0.610)	(0.599)	(0.551)	(2.356)	(0.656)	(2.268)	(1.016)	(3.210)	(1.355)	(3.173)
Distance to toilet/tap water (m)	-0.004	-0.005**	-0.003	-0.005**	-0.008***	-0.009***	-0.007***	-0.009***	-0.003*	-0.006**	-0.003*	-0.004**	-0.006***	-0.010***	-0.006***	-0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Forest coverage ($5m$, %)	0.057***	0.054***	0.057***	0.053***	0.058***	0.056***	0.054***	0.050***	0.058***	0.055***	0.057***	0.054***	0.062***	0.060***	0.061***	0.059***
	(0.015)	(0.014)	(0.016)	(0.014)	(0.020)	(0.018)	(0.021)	(0.018)	(0.016)	(0.016)	(0.016)	(0.016)	(0.023)	(0.021)	(0.022)	(0.021)
Grass coverage ($5m$, %)	0.013	0.011	0.014	0.012	-0.005	-0.009	-0.005	-0.009	0.014	0.012	0.013	0.011	0.000	-0.006	-0.001	-0.005
	(0.016)	(0.016)	(0.016)	(0.016)	(0.015)	(0.014)	(0.016)	(0.015)	(0.016)	(0.017)	(0.017)	(0.017)	(0.017)	(0.016)	(0.017)	(0.016)
Distance to the river shore (m)	-0.055**	-0.085**	-0.052*	-0.093**	-0.068*	-0.078	-0.077**	-0.092*	-0.048*	-0.084*	-0.057**	-0.088*	-0.046	-0.050	-0.049	-0.056
	(0.025)	(0.039)	(0.027)	(0.045)	(0.035)	(0.049)	(0.037)	(0.055)	(0.025)	(0.048)	(0.025)	(0.050)	(0.035)	(0.058)	(0.037)	(0.060)
Date FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Administrative division FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
# of observations	160	160	160	160	111	111	111	111	160	160	160	160	111	111	111	111
Within R-sq	0.80		0.80		0.82		0.82		0.81		0.81		0.82		0.83	
Between R-sq	0.59		0.58		0.60		0.60		0.57		0.56		0.55		0.55	
Breusch and Pagan LM χ^2	21.76		22.51		9.81		9.95		25.23		26.50		13.25		14.56	
Sargan-Hansen stat. χ^2	4.01		4.47		2.53		4.47		21.51		6.61		17.01		12.71	
First stage																
Inner/outer curve		9.444***		6.893***		10.046***		7.303***		28.084***		23.458***		30.192***		24.602***
		(1.766)		(1.394)		(1.419)		(1.120)		(8.504)		(6.957)		(6.943)		(5.866)
Kleibergen-Paap F-stat.		28.60		24.46		50.11		42.51		10.90		11.37		18.91		17.59

Standard errors clustered by houses are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A8: Roof Temperature (House Size Taken Into Account)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Model:	RE	REIV	RE	REIV	RE	REIV	RE	REIV
Sample:	Full				Houses with roofs fully visible			
Dependent variable:	Roof temperature (deg. C) × Roof footprint (m^2)				Roof temperature (deg. C) × Roof footprint (m^2) × House height (m)			
$N(i)_{-i}$	1.418***	2.705***			4.773***	6.246***		
	(0.497)	(0.837)			(1.482)	(2.103)		
$DN(i)_{-i}$			1.638**	3.715***			5.921***	8.590***
			(0.646)	(1.255)			(2.068)	(3.109)
Outside temperature (deg. C)	25.485***	24.306**	26.048***	25.253***	69.703***	65.397***	74.655***	70.557***
	(9.571)	(9.541)	(9.682)	(9.673)	(21.758)	(20.783)	(21.896)	(21.084)
Distance to toilet/tap water (m)	-0.053**	-0.077***	-0.045**	-0.070***	-0.140**	-0.169***	-0.129**	-0.165***
	(0.022)	(0.024)	(0.023)	(0.023)	(0.055)	(0.060)	(0.057)	(0.062)
Forest coverage ($5m$, %)	0.649***	0.606***	0.645***	0.583**	1.876***	1.837***	1.792***	1.697***
	(0.239)	(0.234)	(0.241)	(0.232)	(0.603)	(0.572)	(0.616)	(0.557)
Grass coverage ($5m$, %)	0.314	0.281	0.323	0.288	0.570	0.502	0.592	0.504
	(0.272)	(0.275)	(0.274)	(0.281)	(0.580)	(0.563)	(0.589)	(0.572)
Distance to the river shore (m)	-0.542	-0.995*	-0.533	-1.152*	-2.150**	-2.358*	-2.335**	-2.724*
	(0.386)	(0.508)	(0.405)	(0.596)	(0.950)	(1.211)	(1.010)	(1.399)
Date FE	Y	Y	Y	Y	Y	Y	Y	Y
Administrative division FE	Y	Y	Y	Y	Y	Y	Y	Y
# of observations	160	160	160	160	111	111	111	111
Within R-sq	0.63		0.63		0.64		0.64	
Between R-sq	0.45		0.44		0.50		0.48	
Breusch and Pagan LM χ^2	6.95		7.61		1.01		1.34	
Sargan-Hansen stat. χ^2	6.96		7.68		3.19		5.97	
First stage								
Inner/outer curve		9.383***		6.828***		9.918***		7.214***
		(1.755)		(1.382)		(1.461)		(1.157)
Kleibergen-Paap F-stat.		28.59		24.40		46.09		38.87

Standard errors clustered by houses are in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table A9: Access to Food & Informal Job Opportunities

Model:	(1) RE	(2) REIV	(3) RE	(4) REIV	(5) RE	(6) REIV	(7) RE	(8) REIV	(9) RE	(10) REIV	(11) RE	(12) REIV
Dependent variable:	Roof temperature (deg. C)											
$N(i)_{-i}$	0.093***	0.182***			0.082**	0.179***			0.083**	0.175***		
	(0.035)	(0.062)			(0.034)	(0.065)			(0.033)	(0.065)		
$DN(i)_{-i}$			0.122**	0.250***			0.108**	0.246***			0.113**	0.235***
			(0.048)	(0.087)			(0.048)	(0.091)			(0.050)	(0.088)
Number of supermarkets	-0.138	-0.144	-0.138	-0.145								
	(0.117)	(0.121)	(0.117)	(0.122)								
Numbero of scrap firms					0.113	0.088	0.116	0.091				
					(0.073)	(0.089)	(0.074)	(0.092)				
Distance to X Town (*km*)									0.380**	0.319	0.406**	0.372
									(0.174)	(0.232)	(0.182)	(0.227)
Base control	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Administrative division FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
# of observations	160	160	160	160	160	160	160	160	160	160	160	160
Within R-sq	0.82		0.82		0.82		0.82		0.82		0.82	
Between R-sq	0.65		0.65		0.65		0.65		0.66		0.66	
Breusch and Pagan LM χ^2	25.25		25.27		25.07		25.00		24.91		24.51	
Sargan-Hansen stat. χ^2	3.01		4.72		3.60		6.16		2.85		4.05	
First stage												
Inner/outer curve		9.628***		7.024***		9.377***		6.853***		9.507***		7.066***
		(1.773)		(1.399)		(1.946)		(1.558)		(1.957)		(1.521)
Kleibergen-Paap F-stat.		29.49		25.21		23.22		19.36		23.60		21.58

Standard errors clustered by houses are in parentheses. Base control = Outside temperature, Distance to toilet/tap water, Forest and grass coverage (5-m buffers), Distance to the river shore. "Number of supermarkets" is the number of supermarkets within a 1-km radius from each house centroid. "Number of scrap firms" is the number of scrap firms within a 2-km radius from each house centroid.

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

Table A10: Vegetation Coverage

Model:	(1) RE	(2) REIV	(3) RE	(4) REIV	(5) RE	(6) REIV	(7) RE	(8) REIV
Dependent variable:	Roof temperature (deg. C)							
$N(i)_{-i}$	0.095**	0.187***			0.111**	0.203**		
	(0.042)	(0.071)			(0.055)	(0.091)		
$DN(i)_{-i}$			0.131**	0.255***			0.155**	0.274**
			(0.055)	(0.099)			(0.071)	(0.124)
Forest coverage (10m, %)	0.070***	0.061***	0.070***	0.061***				
	(0.023)	(0.024)	(0.023)	(0.023)				
Grass coverage (10m, %)	0.031	0.022	0.032	0.024				
	(0.025)	(0.026)	(0.024)	(0.026)				
Forest coverage (15m, %)					0.030	0.014	0.030	0.016
					(0.028)	(0.031)	(0.027)	(0.031)
Grass coverage (15m, %)					0.002	-0.014	0.003	-0.011
					(0.026)	(0.028)	(0.024)	(0.028)
Base control excl. veg. cov.	Y	Y	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y	Y	Y
Administrative division FE	Y	Y	Y	Y	Y	Y	Y	Y
# of observations	160	160	160	160	160	160	160	160
Within R-sq	0.82		0.82		0.82		0.82	
Between R-sq	0.55		0.55		0.49		0.49	
Breusch and Pagan LM χ^2								
First stage								
Inner/outer curve		8.954***		6.557***		8.138***		6.024***
		(1.798)		(1.456)		(1.725)		(1.451)
Kleibergen-Paap F-stat.		24.79		20.29		22.26		17.24

Standard errors clustered by houses are in parentheses. Base control excl. veg. cov. = Outside temperature, Distance to toilet/tap water, Distance to the river shore.
*** p<0.01, ** p<0.05, * p<0.1

Table A11: Wind Speed

Model:	(1) RE	(2) REIV	(3) RE	(4) REIV
Dependent variable:	Roof temperature (deg. C)			
$N(i)_{-i}$	0.093**	0.203***		
	(0.039)	(0.070)		
$DN(i)_{-i}$			0.122**	0.277***
			(0.052)	(0.100)
Wind speed (m/s)	-0.368	-0.459	-0.345	-0.422
	(0.692)	(0.753)	(0.681)	(0.741)
Base control	Y	Y	Y	Y
Date FE	Y	Y	Y	Y
Administrative division FE	Y	Y	Y	Y
# of observations	160	160	160	160
Within R-sq	0.82		0.82	
Between R-sq	0.64		0.64	
Breusch and Pagan LM χ^2	24.87		24.68	
Sargan-Hansen stat. χ^2	7.48		23.97	
First stage				
Inner/outer curve		9.547***		6.982***
		(1.785)		(1.406)
Kleibergen-Paap F-stat.		28.59		24.65

Standard errors clustered by houses are in parentheses. Base control = Outside temperature, Distance to toilet/tap water, Forest and grass coverage (5-m buffers), Distance to the river shore.
*** p<0.01, ** p<0.05, * p<0.1

Table A12: Bridge Effects

Model:	(1) RE	(2) REIV	(3) RE	(4) REIV	(5) RE	(6) REIV	(7) RE	(8) REIV	(9) RE	(10) REIV	(11) RE	(12) REIV
Sample:	Excl. houses near bridges				Excl. a cluster near bridges				Excl. houses near pedestrian bridges			
Dependent variable:	Roof temperature (deg. C)											
$N(i)_{-i}$	0.089**	0.172***			0.076**	0.133**			0.091*	0.317*		
	(0.036)	(0.064)			(0.035)	(0.059)			(0.052)	(0.167)		
$DN(i)_{-i}$			0.119**	0.236***			0.108**	0.181**			0.125*	0.323**
			(0.049)	(0.091)			(0.051)	(0.079)			(0.066)	(0.150)
Base control	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Date FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Administrative division FE	Y	Y	Y	Y	-	-	-	-	Y	Y	Y	Y
# of observations	142	142	142	142	136	136	136	136	137	137	137	137
Within R-sq	0.82		0.82		0.83		0.83		0.83		0.83	
Between R-sq	0.67		0.67		0.66		0.66		0.69		0.69	
Breusch and Pagan LM χ^2	17.09		16.69		18.88		18.54		18.17		17.56	
Sargan-Hansen stat. χ^2	3.77		5.18		3.96		4.98		4.68		5.19	
First stage												
Inner/outer curve		9.584***		6.982***		9.765***		7.235***		5.934**		5.821***
		(1.804)		(1.459)		(1.921)		(1.533)		(2.556)		(1.981)
Kleibergen-Paap F-stat.		28.22		22.91		25.83		22.27		5.39		8.64

Standard errors clustered by houses are in parentheses. Base control = Outside temperature, Distance to toilet/tap water, Forest and grass coverage (5-m buffers), Distance to the river shore.
*** p<0.01, ** p<0.05, * p<0.1

Table A13: Roof Footprint vs Roof Temperature: Overlap > 80%

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Model:	OLS	OLS	OLS	RE	REIV	RE	REIV	RE	RE
Dependent variable:	Log (Footprint, m^2)			Roof temperature (deg. C)					
Oldcomer (height > 0.9m, overlap > 80%)	0.640***	0.574***	0.569***	-1.156*	-1.255*	-1.186*	-1.319*	0.458	0.951
	(0.166)	(0.152)	(0.158)	(0.606)	(0.652)	(0.626)	(0.677)	(1.599)	(1.655)
$N(i)_{-i}$		0.037***		0.099***	0.197***			0.122***	
		(0.010)		(0.036)	(0.060)			(0.040)	
$DN(i)_{-i}$			0.050***			0.134***	0.271***		0.166***
			(0.014)			(0.050)	(0.085)		(0.054)
Oldcomer×$N(i)_{-i}$								-0.085	
								(0.063)	
Oldcomer×$DN(i)_{-i}$									-0.147*
									(0.088)
Outside temperature (deg. C)				2.932***	2.853***	2.972***	2.920***	2.938***	2.983***
				(0.515)	(0.502)	(0.527)	(0.517)	(0.519)	(0.534)
Distance to toilet/tap water (m)	0.000	-0.001	-0.000	-0.006***	-0.008***	-0.005***	-0.007***	-0.006***	-0.006***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Forest coverage (5m, %)	-0.001	-0.002	-0.003	0.067***	0.064***	0.066***	0.062***	0.064***	0.062***
	(0.004)	(0.004)	(0.004)	(0.015)	(0.014)	(0.015)	(0.014)	(0.015)	(0.015)
Grass coverage (5m, %)	-0.007	-0.008*	-0.008*	0.029**	0.027*	0.030**	0.028*	0.028*	0.028*
	(0.004)	(0.004)	(0.004)	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)	(0.016)
Distance to the river shore (m)	0.003	-0.011	-0.013	-0.074***	-0.109***	-0.079***	-0.121***	-0.073***	-0.077***
	(0.009)	(0.008)	(0.008)	(0.027)	(0.036)	(0.027)	(0.040)	(0.026)	(0.027)
Date FE	-	-	-	Y	Y	Y	Y	Y	Y
Administrative division FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
# of observations	63	63	63	160	160	160	160	160	160
Adj R-sq	0.09	0.18	0.18						
Within R-sq				0.82		0.82		0.82	0.82
Between R-sq				0.67		0.67		0.68	0.69
Breusch and Pagan LM χ^2				21.72		21.46		20.59	19.37
Sargan-Hansen stat. χ^2				5.25		6.34		5.36	6.38
First stage									
Inner/outer curve					9.422***		6.863***		
					(1.748)		(1.361)		
Kleibergen-Paap F-stat.					29.07		25.42		

Standard errors clustered by houses are in parentheses. Base control = Outside temperature, Distance to toilet/tap water, Forest and grass coverage (5-m buffers), Distance to the river shore.
*** p<0.01, ** p<0.05, * p<0.1