There Is No Free House: Ethnic Patronage in a Kenyan Slum*

Benjamin Marx
Thomas M Stoker
Tavneet Suri

August 2016

Abstract

We show evidence of ethnic patronage in the housing market in a large Kenyan slum. Slum residents pay higher rents, and live in lower quality housing (measured via satellite pictures) when the landlord and the locality chief belong to the same tribe. Conversely, rents are lower, and investments higher when households and chiefs are co-ethnics. These effects are partially offset in more ethnically diverse areas, and in areas with high youth unemployment, where gangs restrain the power of chiefs. Our identification relies on the exogenous appointment of chiefs and is supported by several tests, including a regression discontinuity design.

Keywords: Ethnic diversity, governance, housing markets, satellite imagery

JEL Classification: O17, O18, J15

*We are extremely grateful to Suleiman Asman, Eleanor Wiseman and Geoffrey Onyambu for exceptional research management and assistance in the field. We thank Ali Ahmed for his help with the data on elders and chiefs, Jeff Campbell for his help processing the satellite pictures, and Daniel Sheehan for his advice on GIS computations. We are grateful to Daron Acemoglu, Prashant Bharadwaj, Michael Boozer, Esther Duflo, Ben Olken, Elias Papaioannou, and seminar participants at Stanford, the Working Group in African Political Economy (WGAPE), Columbia, Oxford, the Urban Poverty Workshop at Duke University and the World Bank for very helpful comments. We are also thankful to Sandra Joireman for inspiring some of the ideas presented in this paper and for her feedback on our survey instruments. A previous version of this paper was circulated under the title "There Is No Free House: Ethnic Patronage and Property Rights in a Kenyan Slum". This research was funded by MIT Sloan. Institutional Review Board (IRB) approvals for the data collection were obtained from MIT. Marx is at the Department of Economics, MIT. Stoker and Suri are at the MIT Sloan School of Management. Suri is the corresponding author: E62-517, 100 Main Street, Cambridge MA 02142. Email: tavneet@mit.edu.

I Introduction

There is a large literature documenting the relationship between ethnic diversity and economic performance. A number of these studies (e.g. Easterly and Levine (1997) and Miguel and Gugerty (2005), among others) show that ethnic diversity correlates with poorer economic outcomes. Alesina and La Ferrara's (2005) excellent review concludes that this relationship holds whether looking across countries, cities or even at the very local level; in Sub-Saharan Africa as well as in other countries. Although this negative correlation between ethnic diversity and economic outcomes is empirically robust, there is less understanding of the mechanisms behind this relationship. Why do ethnically heterogeneous communities tend to exhibit lower levels of economic activity?

One mechanism highlighted in the empirical literature is ethnic favoritism in the provision of public goods: governments and bureaucrats systematically privilege their co-ethnic constituents when making allocation decisions. For example, Burgess et al. (2015) quantify the extent of ethnic favoritism using the example of public expenditure on roads construction in Kenya, and show that transition to democracy undoes the effects of ethnic favoritism. Franck and Rainer (2012) find that primary education and infant mortality across Africa are systematically affected by changes in the ethnicity of country leaders. A second mechanism posited in the literature is that collective action problems make it harder for heterogeneous groups to coordinate. Using experimental games to understand how ethnic diversity affects public goods provision, Habyarimana et al. (2007) try to disentangle the mechanisms highlighted in Alesina and La Ferrara (2005) and the related literature, in particular, ethnic preferences, technology (it is easier to interact with co-ethnics) and strategic concerns (the ability to adhere to norms within groups via sanctions). They find little role for preferences or technology, but highlight the importance of strategic concerns.

In this paper, we make two contributions to this literature. First, we go beyond the public goods mechanism to explain the relationship between ethnic diversity and economic activity, and instead study whether ethnicity matters in private markets. With a few notable exceptions (e.g. Fafchamps (2000), Hjort (2014) and Michelevitch (2015)), there is a dearth of studies providing evidence of ethnic favoritism in private markets in developing countries. We document how ethnic bias affects private transactions in the market for slum housing in Nairobi, Kenya, and look at the economic consequences of these distorsions.[1] In doing so, we also aim to shed light on the economics of slums, a topic of increasing policy relevance but for which rigorous academic evidence remains very scarce. Second, we show that the ethnicity of local arbitrators matters in these private markets. Arbitrators in our setup are chiefs appointed by the administration whose job it is to mediate rent-related disputes in the slums. There is a literature in the U.S. focusing on the ethnicity of arbitrators (examples include Anwar, Bayer and Hjalmarsson (2012), Abrams, Bertrand and Mulainathan (2012), Shayo and Zussman (2011) and Price and Wolfers (2010)), but no studies have

[1] Related, but in the U.S. context, Bayer, Casey, Ferreira and McMillan (2013) study racial discrimination in the home sales market using standard panel data methods.

looked at their role in developing economies.

We focus on a cross-sectional sample of approximately 30,000 individuals in the Kibera slum in Kenya, one of Africa's largest informal settlements. Over the past decades, the expansion of Kibera (which hosts approximately 5% of Nairobi's population) has paralleled that of other slums in the developing world. An estimated 62% of Sub-Saharan Africa's urban population currently lives in slums.[2] Considering the expansion of cities and slums globally, the UN estimates that at least 450 million new housing units will be needed to accommodate populations in direct need of shelter (UN-Habitat 2012). Understanding how slum land markets operate in developing world cities is therefore a question of central economic importance, and will remain so in the coming decades.

To understand how ethnic bias affects transactions in the housing market and the price of housing, we look at the interaction between local authorities (in this environment, location chiefs),[3] landlords and tenants. Our main results can be summarized as follows. First, we find that a match between the tribe of the chief and that of the landlord results in tenants paying rents between 6% and 11% higher (we use the terms "tribe" and "ethnicity" interchangeably throughout the paper). Conversely, we find that a match between the tribe of the chief and that of the tenants leads to lower rents paid by tenants, by a similar order of magnitude. We argue that these effects arise when chiefs are called upon to arbitrate a dispute between landlords and tenants on the rent amount due by tenants, and accordingly side with a co-ethnic. This interpretation is consistent with all the anecdotal evidence we collected in the field, and with a variety of qualitative accounts on the role ethnic networks play in the governance of Kibera.[4]

Second, we find that investments in the housing infrastructure are lower (though not significantly so) when a landlord and the chief belong to the same tribe, and significantly higher when the household and the chief belong to the same tribe. We measure investments by looking at a measure of the quality of a household's roof, using the roof's luminosity (brightness) from high resolution satellite images of the slum. This evidence reinforces our previous results: quality-adjusted rents are higher for households whose landlord is of the same tribe as the chief, and lower when the chief is of their own tribe.

To identify these effects, we use the fact that the Provincial Administration (exogenously) appoints chiefs to their positions within Nairobi, and that chiefs are regularly rotated around the city. Because the chiefs are frequently transferred, they have limited time to form relationships with local stakeholders and to establish rent extraction mechanisms. We look at ethnicity as a factor that

[2]Slums are broadly defined as overcrowded urban areas with poor-quality housing, a lack of public services, and large numbers of informal residents (UN-Habitat 2006).

[3]Chiefs in the context of our study are not traditional ethnic leaders imposed by local custom (Michalopoulos and Papaioannou (2013a)), but employees of the administration locally recognized as chiefs. Acemoglu, Reed and Robinson (2014) showed that the amount of competition faced by customary chiefs in Sierra Leone had a long-term impact on local economic performance and public goods provision.

[4]In particular, Joireman and Vanderpoel (2011) document that a chief will generally side with members of the same ethnic group in cases of rent disputes involving different ethnicities. They also highlight how permissions are needed from the chief to upgrade housing structures, interactions that often involve bribery and where ethnicity also plays a key role.

potentially facilitates the establishment of such mechanisms.

We conduct three separate checks to defend our identification assumption that the appointment of chiefs is exogenous to characteristics of the households in the slum. First, we show that a wide variety of socioeconomic characteristics of households measured in 2009 do not predict ethnic alignments in 2012, even though they do correlate strongly with rents. If households were sorting based on underlying characteristics, we would have expected at least some of the variables observed in 2009 to predict whether a household and a chief are of the same ethnicity in 2012. Second, we show that the match between the household and the chief's ethnicity and the match between the landlord and the chief's ethnicity are not correlated with various measures of the ethnic composition of the locality a household lives in. Third, we use the idea that households are very similar along observable and unobservable characteristics across the internal administrative boundaries of the slum to design a regression discontinuity (RD) test. For the subsample of households who have a co-ethnic chief operating in the slum, for whom the RD would be relevant, we show that our results hold, and are in fact stronger, within very small bandwidths around the internal (location) boundaries of the slum.[5]

Why does ethnic favoritism take place in the context of our study? We cannot disentangle whether the effects we see come from preferences or from strategic motives on the part of the chiefs. However, we can explore mechanisms that help understand the variation in these effects. In particular, we explore two sets of mechanisms: first, local ethnic diversity, and second, a lack of checks and balances on the power of arbitrators (chiefs) that allow them to extract rents from the ethnic capital they possess. We first look at heterogeneity in our main effects by the extent of local ethnic diversity in different areas of the slum. Consistent with the prediction in Alesina and La Ferrara (2005) that more diverse communities experience more efficient provision and production of private goods, we find that local ethnic diversity diminishes ethnic discrimination against tenants (in terms of rent charged) when the landlord and chief are of the same tribe. Second, we argue that the chiefs' ability to make ethnically biased decisions when they arbitrate rent disputes reflects the lack of checks and balances on their power. In the Kibera slum, a limitation to the power of chiefs is the presence of youth gangs, which provide an alternative arbitration system for rent disputes generally more tilted towards tenants (Joireman and Vanderpoel 2011). We use youth unemployment as a proxy for the presence of youth gangs and find that high youth unemployment in the area partially offsets the rent effects of a landlord-chief match, even when controlling for wealth and education.

These empirical results suggest that ethnic patronage surrounding housing rents have long term effects on welfare in the slum, though the sign of these welfare effects is unclear. On the one hand, ethnic discrimination on the housing market will be detrimental to some members of

[5]Our data supports the notion that the mobility of households within the slum is limited. Households have lived in the same structure and with the same landlords for long periods of time (8 and 7 years, respectively), much longer than the typical tenure of a chief within the slum, which averages 23 months in our data.

society, and the fact that rents keep changing as a result of chiefs moving in and out of the slum suggests an additional welfare cost in terms of consumption volatility. On the other hand, to the extent that it may help offset market failures in land markets, the ethnic bias we document could be welfare-improving if the landlords are local monopolists. Our result that ethnic patronage is less prevalent in more diverse areas suggests that it is less likely to be the case that landlords are local monopolists.[6]

The paper is structured as follows. Section 2 provides relevant historical and institutional background. Section 3 describes the data we collected and our estimates of roof quality based on high-resolution satellite imagery. Section 4 describes our empirical framework. Section 5 provides our main results, including identification and robustness checks. Section 6 concludes.

II Background

A A Brief History of Kibera

The history of Kibera is closely intertwined with that of Nairobi, a city that was founded *ex nihilo* by the British as a railway depot in 1899 and soon after chosen to become the capital of Kenya. Kibera was established nearby the new colony a decade later, in 1912, to accommodate veteran Sudanese soldiers from the King's African Rifles (KAR), a contingent of the British colonial army. For several decades, the KAR veterans and their families were the only Kibera residents with formal land permits that exempted them from hut taxes. However, the settlement also soon became a refuge for Kenyan rural migrants as Nairobi expanded to 120,000 inhabitants by the end of World War II. By that time, only a small fraction of the 3,000 Kibera residents were relatives of the initial Sudanese (also referred to as Nubian) settlers (Parsons 1997). The British authorities never recognized Kibera as a legal settlement, but failed to evict its inhabitants until Kenya's independence in 1963. After independence, the Kenyan government formally reclaimed property over the Kibera area and tried to discourage new settlements. In spite of this, the population of slum dwellers continued to increase steadily, from 17,000 in 1972 to 62,000 in 1979 (Amis 1984) and probably several hundred thousands in the 1990s. Current estimates range between 170,000 (the official census number released in 2009), to more than one million according to unofficial sources. Figure 1 shows a map of the Kibera slum, which nowadays spans approximately 8 square kilometers in the center of Nairobi.

A densely populated and ethnically diverse area, Kibera has experienced many episodes of

[6]Given the long history of conflict surrounding land in Kibera which led to the current uncertainty about occupation rights, our paper also adds to a large literature on the adverse effects of weak property rights on economic performance (e.g. Besley 1995; Banerjee, Gertler and Ghatak 2002; Goldstein and Udry 2008; Hornbeck 2010), including a more recent literature documenting the importance of strengthening property rights in informal urban areas (Field 2007; Galiani and Schargrodksy 2010; Lanjouw and Levy 2002). Contrary to the prevailing view in the literature, however, the results we present imply that strengthening property rights may be a mixed blessing if it contributes to reinforce the market power of landlords (92% of the households we surveyed are rent-paying tenants).

inter-ethnic violence in the past. It is plagued by organized crime and gangs, and was amongst the areas heavily affected by the post-election violence in early 2008. The ethnic makeup of Kibera is not reflective of that of Kenya, since Luos, Luhyas and Kambas, in that order, are the tribes most represented in the slum (representing 36%, 27% and 15% of the slum population, respectively), while Kikuyus (the most prevalent tribe in the country) represent only 6%. Kenyans of Nubian origin constitute only a small fraction of Kibera's current population, yet they remain well represented amongst the local administration. Figure 2 shows the distribution of household tribes across the slum.

B Land Issues and Landlords in Kibera

Land ownership and tenancy rights in Kibera have been ambiguous ever since the establishment of the settlement. Throughout the colonial period, the entirety of Kibera's land remained the property of the British Crown.[7] The Nubian settlers, considered "Tenants of the Crown", formally remained the slum's only residents with settlement permits until 1969, when the Kenyan government revoked their claims on the land (Joireman and Vanderpoel 2011; de Smedt 2011).

In the meantime, however, a *de facto* housing market had developed in which so-called "landlords", usually long-term residents of the settlement, allocated or rented structures to newcomers (Amis 1984; Temple 1974). Starting in 1974, new land titles were (illegally) re-allocated by local chiefs and bureaucrats in the Provincial Administration, engendering the creation of 1,400 new structures and a tripling of the Kibera population in less than a decade (Amis 1984). These political favors resulted in the current situation where Kibera land formally remains the government's property, but the housing market is effectively controlled by private landlords. From our data, there are, in fact, at least several hundred different landlords in Kibera.[8] These landlords mostly do not have official property rights on land in the slum,[9] but were allowed to build, to buy and to rent out habitable structures (often by the District Officer) (de Smedt 2011)). The property rights of landlords are informal in this context - legally the rights to the land lie with the government (more so under the new 2010 constitution which requires the government to repossess such lands). But the constitution is not enforced and formal arbitration (courts) are not accessible, which allows lower-level informal norms to be enforced instead. In the Kibera context, the only institution vested with informal arbitation power is therefore the chief (this is similar to the case of Bubb (2013)).

Since independence, the majority of landlords in Kibera have been members of either the Kikuyu or the Nubi tribe (Amis 1984; Joireman and Vanderpoel 2011). These two groups may have been better able to obtain land titles from the administration (a claim made by Amis (1984)) and appear to mostly have been well-to-do, absentee landlords. Syagga et al. (2002) reported on a

[7]This was under the Crown Lands Ordinance of 1902.

[8]Our data does not allow us to calculate the exact number of landlords, since some households did not provide us with the full name of their landlord.

[9]The Nubi landlords claim official property rights based on the permits allocated by the British to their ancestors in the early days of Kibera. These permits have all been revoked since Kenyan independence.

survey of Kibera landlords where more than 80% of landlords lived outside of the slum and 57% were public officials (Joireman and Vanderpoel 2011). Our own survey data also suggests that the majority of landlords (55%) are Kikuyus or Nubis living in estates outside Kibera.[10] The remaining half of landlords is predominantly Luo (17%), Luhya (11%), and Kamba (8%). Figure 3 shows the distribution of landlord tribes across the slum.

C Chiefs and Elders in Kibera

Kibera is administered through multiple layers of local government, mixing traditional and formal authorities. Between 1950 and 1987, Kibera was governed by one chief appointed by the Provincial Administration, assisted by community elders.[11] In 1987, the administration added one layer of government below the chief - the assistant chief. In 2002, the organization of the slum was redefined into its present format: the slum area and its surroundings were divided into 4 locations, each governed by a chief and 9 sublocations, each governed by an assistant chief who reports to the location chief.[12] The 9 sublocations are divided into villages, which are relevant for the purpose of customary governance but are not part of the provincial administration. In total, there are 17 villages, of which 5 larger ones are split into zones.

Between 2002 and 2012, 18 different chiefs (each responsible for one of 4 locations) and 21 different assistant chiefs (each responsible for one of 9 sublocations) served in Kibera. In Nairobi, appointments are advertised by the Provincial Administration and made subsequent to an interview process. In addition, the chiefs are rotated through the entire province approximately every two years. These transfers are made entirely at the discretion of the administration. In the empirical analysis below, we use the exogenous rotation of location chiefs as our main source of identification. Very rarely, Kibera chiefs serve multiple times in Kibera over their career.[13] At the time of data collection in August 2012, the current chiefs in Kibera had been active for an average of 18.8 months, though based on our data on chief tenures, the average tenure for a chief has been 23 months.

Based on our surveys of chiefs, assistant chiefs and elders in Kibera (described below), the role of the chiefs is to support the administration in a variety of ways, including maintaining law and order and resolving disputes, primarily rent and domestic disputes. The chiefs are also responsible for appointing community elders, although these elders are always chosen amongst people well known to the community and tend to stay much longer than the chiefs and assistant chiefs (14.6 years on average in our survey data). In our surveys of chiefs, all respondents mention conflict

[10]In our survey data, 33% of households reported dealing with a Kikuyu landlord and 22% with a Nubi landlord. The proportion reported by Amis (1984) in the early 1980s was 66% and 22% for the Kikuyus and Nubis, respectively.

[11]The community elders represent the most basic of local (informal) authority in the slum. Their area of responsibility usually encompasses a village or a zone.

[12]The 4 locations are Kibera, Laini Saba, Sarangombe and Mugumo-ini. The 9 sublocations are Kibera, Lindi, Makina, Silanga, Laini Saba, Nyao Highrise/Soweto East, Bomas, Gatwikira, and Olympic.

[13]In our data, 4 Kibera chiefs were active during at least 2 different periods since 2002.

resolution and arbitration as one of their three most important responsibilities. We provide more details on rent-related disputes and the arbitration role of chiefs in section 3.

D Youth Gangs and Unemployment

While there is virtually no police presence in most areas of the slum, Kibera is host to a number of youth gangs which participate in organized banditry and are notorious for their frequent robberies, violent crimes, and other forms of intimidation. The gangs active in Kibera are affiliated with larger groups operating nationally, such as the *Mungiki* who claim ties to the anti-colonial Mau Mau movement, and the *Kamkunji*, also known as "the Taliban". Kibera's youth constitute a primary recruitment target for these groups, since all Kenyan gangs tend to recruit amongst unemployed and disenfranchised youth in poor urban areas (International Crisis Group 2013). Like elsewhere in Kenya, gang membership in Kibera is typically ethnic-based - for example, Kikuyu for the *Mungiki* and Luo for the *Kamkunji*.

The gangs control and levy taxes over certain public goods and amenities (water, sanitation, electricity, transport), and they may provide "protection" to co-ethnic individuals in cases of inter-ethnic disputes. Joireman and Vanderpoel (2011) document the role of the *Kamkunji* gang, which claims to defend Luo interests in Kibera and systematically intervenes in favor of Luo tenants in cases of housing-related disputes. The group convenes once a week in a formal meeting where Luo residents are invited to express their grievances and report on disputes with their landlords. The *Kamkunji* are so influential in some areas of the slum that they may entirely substitute for the authority of the chief, particularly in Gatwekera village (a village where we had some difficulty interviewing community elders and chiefs, as we describe below).

III Data

Our main analysis combines household-level data collected in Kibera with high resolution satellite imagery data captured over the slum area. We also use data from the 2009 Population and Housing Census and draw on surveys that we conducted with current chiefs, assistant chiefs and village and zone elders.

A Listing (Census) and Household Survey Data, 2012

Our survey data was collected between February and December 2012. To constitute a sampling frame of households within Kibera, we first listed 31,765 households over the 9 sublocations that compose the slum area.[14] The listing involved two rounds of visits in each sector of the slum. GPS

[14]Since the average household size in our listing is 3.65, our listing covered approximately 115,942 individuals or 68% of the 2009 slum population based on the 2009 Census data. Since 2009 there has been some formalization on the edges of the slum so it is unclear what the true population of the slum for 2012 should be.

coordinates were collected for all structures inhabited by the households, resulting in 9,728 unique sets of coordinates and hence structures. For each structure we also collected the current number of households inhabiting the structure.[15] The household size in our sample was approximately 3.7 individuals per household, with just over 2 adults. 65% of household heads gave us valid phone numbers. Table 1A shows summary statistics from this listing data.

Based on this listing of Kibera households, we conducted a phone survey starting in July 2012. The survey itself collected data on the tribe of landlords, rents, renovation of roofs, and previous evictions within the slum. To construct this sample, we stratified the listing by whether the household reported a phone number. We attempted to contact all households that reported a phone number and reached 79% or 16,314 households on the phone. Of the 21% of households that could not be reached by phone, we sampled 20% (888 households) and collected these surveys in person in the slum. Finally, of the 35% of households that did not provide a phone number, we randomly sampled 14% (1,595 households) and conducted the survey in person. Throughout our analysis, we re-weight the data to create a sample of surveys that is representative of the whole 31,765 households initially listed. This sampling strategy gave us a total target sample size of 18,797 households,[16] of which we reached 18,254 (97%), giving us an attrition rate of 11% (weighting attrition by survey sampling weights). Table 1A shows summary statistics from our household survey.

As can be seen in Table 1A, approximately 92% of households pay rent and average monthly rents are KShs 1,715 (US$ 20). About 11% of households report having ever been evicted from their houses in Kibera. The most common reasons for evictions are households not paying rents (45%) or refusing to pay higher rents (10%), a unilateral decision from the landlord (19%), or the structure being demolished (9%). 23% of households have had their roofs renovated since they moved in and 18% in the last two years. Finally, households have spent on average 16 years in Kibera, 8 years in the same structure, and 7 years with the same landlord. Looking at the ethnic match variables, 22% of households have a landlord with the same tribe as the chief. About 14% of households belong to the same tribe as the chief and 28% belong to the same tribe as their landlord. Only in 5% of households do all three tribes (household, landlord and chief) coincide.

B Chief and Elder Surveys

In addition to the listing and household surveys, we obtained the history of chiefs and assistant chiefs who served in Kibera since 1950. We also conducted surveys of all current location chiefs, sublocation chiefs, and community elders who accepted to be interviewed. Our sample of respondents was composed of 3 location chiefs, 7 sublocation chiefs (assistant chiefs)[17] and 45 community

[15]The average number of households per structure was 6.7.

[16]The details of the sampling strategy are described in Online Appendix Figure A1.

[17]The one location chief and two sublocation chiefs we missed were all operating in the same location (Sarangombe location) where the *Kamkunji* gang is particularly active. These individuals, as well as the elders in that village all declined to be interviewed.

elders. Chiefs and assistant chiefs report earning about KShs 18,000 monthly (US$ 228), which is low by Nairobi standards even though 7 of the 10 officials we surveyed had college education. 80% of respondents lived outside of Kibera. Half of them said they sided with the landlord in most instances of rent disputes, and 20% reported siding with the tenants. All officials surveyed cited the tenants' inability to pay the rent agreed upon as the most common reason for disputes, and 60% gave rent increases as the second most common reason.

We also visited various chiefs' offices to obtain, wherever possible, administrative records on the amount of rent-related disputes, and the way in which disputes are typically settled by chiefs. Unfortunately, since consistent records of these disputes are not being kept, we cannot provide exact statistics on the number and the outcomes of these disputes. The anecdotal figures we could obtain are the following. In Sarangombe location (one of the four locations in our data), over a one-week period in February 2014, thirteen rent-related cases were reported to the chief and settled by him. All cases were initiated by the landlord as a result of tenants not paying their rents. Out of these thirteen cases, three ended with the tenant vacating the dwelling, eleven resulted in a rent increase with payment of some arrears, and two ended with the tenant agreeing to pay the rent but arrears were forgiven. Over a three-day period in April 2014, six out of eleven cases also resulted in the chief summoning tenants to pay their rents and arrears over a given timeframe. In Kibera location, over a one-week period in March 2014, fifty-five out of fifty-eight cases were settled by the chief, with the remaining three being settled in courts. The outcome of these disputes was not reported. In Langata location, we could find a large number of rent-related disputes having been reported to the chief and documented (158 and 413 over two different time periods, the length of which is undocumented). Again, the outcome of these disputes is unknown.

The data on community elders included information on individual careers and backgrounds, public goods and private sector activity in the respondent's area of responsibility (e.g. the number of kiosks and M-PESA agents), the role of landlords, and governance. The elders we surveyed belonged to the Luo (31%), the Kikuyu (24%), or the Nubi (20%) tribe. 84% were appointed (in most cases, by the location chief) while the remaining 16% were elected. Finally, 96% of elders report that permission is needed to upgrade the housing in the slum, in particular, to upgrade roofs, and 77% of the time, the permission of one of the chiefs is needed.

C Satellite Pictures

To study investments in slum structures in Kibera, we develop a new methodology that uses the luminosity reflected by metal roofs from daytime satellite pictures of the area. A recent literature uses the luminosity data from nighttime satellite images to gauge economic activity, showing that nighttime luminosity correlates with GDP and other economic indicators (Chen and Nordhaus 2011; Henderson, Storeygard and Weil 2012).[18] Two limitations of the nighttime data are its low

[18] Recent works using nighttime data include Bleakley and Lin (2012) and Michalopoulos and Papaioannou (2013b).

resolution[19] and the lack of data variation in areas of the world that are not well-electrified. For example, Henderson, Storeygard and Weil (2012) report that over 99% of the territories of Mozambique and Madagascar are completely unlit at night over the period 1992-2008. Although useful as a proxy for aggregate activity, the nighttime data is inadequate for measuring economic activity at the household level.[20] Using daytime pictures of Kibera with a panchromatic resolution of 0.5 meters, we are able to use the satellite data at a disaggregated level of analysis, down to the level of a dwelling. Our satellite images span a period of three years between July 2009 and August 2012. An example of the roofs identified from the satellite images for January 2011 is shown in Figure 4.

Our analysis uses the fact that metal roofing quality is a useful variable to understand the investment behavior of households in the slum. 96% of slum residents have corrugated iron roofs. There is little variation in the floors and walls of these households. Most slum inhabitants live a subsistence lifestyle, work in the informal sector and own few assets. Households with higher-quality roofs are likely to be those with more savings, more secure tenure, and/or who can afford to replace their roofs when needed or pay the higher rents associated with better shelter.

Metal roofs are shiny when new or recently renovated, and become dull with age through degradation and rust. Online Appendix Figure A2 shows two pictures of the same area of Kibera at two different dates (July 2009 and August 2012) with sections highlighted that illustrate roofs getting upgraded, or decaying and rusting over the period. Though we have survey data on roof renovations, there are two reasons we use satellite images as well. First, we may not believe the self reports – households may not remember exactly when their roofs were replaced. Second, the satellite images provide a short term panel on what happens over a two year period, with detailed variation in the quality of the roof, not just the replacement of the roof.

Throughout our analysis, we extract the luminosity data from the panchromatic pictures using the *Zonal Statistics (Spatial Analyst)* tool of the ESRI ArcGIS software. The data is analyzed and presented at two different levels - at the level of a roof, and at the level of a household. For the roof-level analysis, we extract luminosity reflected by the roofs in our sample. This process involved segmenting the satellite pictures into roofs inhabited by households, as shown in Figure 4.[21] The mean surface of a roof is 466 square meters, and 11.2 households on average live under one roof. For the household analysis, we extract luminosity values within a one-meter radius of the household's GPS coordinates. We use these two different methods for robustness purposes as both involve some measurement error.[22] Table 1B shows summary statistics on the luminosity variables

[19]This data is provided by the Defense Meteorological Satellite Program's Operational Linescan System (DMSP-OLS) and is a six-bit measure (ranging from 0 to 63) captured for every 30 arc-second pixel on the surface of the earth - approximately 0.86 square kilometers at the equator.

[20]Other recent papers, including Burgess et al. (2012), use daytime satellite imagery for economic analysis, but their level of analysis remains too large for the measurement of individual outcomes.

[21]The segmentation was done using a software designed for object-level picture segmentation, eCognition Developer. We provide more details on this segmentation process in the Online Appendix.

[22]In the first case, the roofs cannot be recognized with full accuracy (in spite of the high resolution of the pictures); and in the second case, measurement error in the collection of GPS coordinates may result in a slightly incorrect positioning of the household on our pictures of the slum.

extracted from these satellite images. In the Online Appendix, we provide some empirical validation that the luminosity data can be used as a proxy for the quality of housing. In particular, we show that these measures of luminosity are correlated with measures of income and wealth.

D KNBS Census data, 2009

In addition to our survey data, we use individual-level data from the National Population Census conducted by the Kenya National Bureau of Statistics (KNBS) in August 2009. The data we have covers an entire Division (an administrative unit) that is itself called Kibera, where the Kibera slum is located.[23] In 2009, the slum was divided into 643 Enumeration Areas (EAs), encompassing 64,588 households, of which 576 EAs (57,804 households) were listed as informal settlements. We digitized EA maps from KNBS to map the 2009 census data to the EA level, and to match our own survey data to EAs. Since the 2009 census, there have been a number of changes to the geographic extent of the slum, the most important being the construction of highrise buildings along two edge areas of the slum. We therefore use data for the 608 EAs covered during our 2012 listing.

In Table 1B, we show summary statistics from the KNBS census data aggregated at the level of the EA. We show the rates of youth unemployment (defined both within households and within EAs as the fraction of individuals aged 16-25 without a job) which we use later as a proxy for the local presence of youth gangs. The EA-level average of these two measures is 17%. As measures of economic activity, about 4% of household heads have no education, 42% have some secondary education, and household heads have 9 years of education on average (a year longer than primary school completion). To provide some context, we also report measures of asset ownership and wealth. For example, 36% of households have water piped into their dwelling, and about 8% of households report owning their house.

IV Empirical Framework

A Overview of Research Question

This section describes the empirical framework we use to test our hypotheses. We are interested in understanding how a specific form of ethnic collusion - the tenant, the landlord, and/or the location chief belonging to the same tribe - affects rents in the slum. Our identification strategy is based on the fact that chiefs are exogenously assigned by the Provincial Administration to their position within the slum. Chiefs employed by the Provincial Administration are regularly transferred across locations and neighborhoods. Because they have limited time to form local relationships, chiefs may rely on other forms of social capital, such as ethnicity, to establish local networks. In

[23]The Kibera Division is split into 7 locations and a further 16 sublocations. Of these, the Kibera slum spreads across 4 locations and 9 sublocations. From the 1999 Census, there were a total of 2,427 locations and 6,612 sublocations in Kenya.

disputes between landlords and tenants over amounts paid in rent and investments, a common occurrence in the slum, a chief may be more likely to side with a landlord if he belongs to the same tribe and the tenant does not. This collusion, if it occurs, will lead to more rent extraction at the expense of the tenant. Conversely, a chief might protect a tenant of the same tribe against a landlord belonging to a different tribe, leading to the tenant paying a lower rent. These hypotheses are in line with Joireman and Vanderpoel (2011) who describe Kibera chiefs as "bureaucratic entrepreneurs, exhibiting rent-seeking behavior in exploiting their formal office to secure personal gain" and further argue that

> "Since structure owners have a higher socioeconomic status than tenants, they can pay more, giving chiefs an incentive to decide in favor of a structure owner. An alternative outcome is possible if the chief and tenant are of the same ethnic group and the structure owners' tribe is perceived as adversarial. For example, a Luo chief would generally side with a Luo tenant against a Kikuyu structure owner despite the smaller payment offered by the tenant." (Joireman and Vanderpoel, 2011: pp. 140-142).

In addition, we look at whether these relationships have longer term effects by constraining investments in the housing infrastructure. Anecdotal evidence supported by our data indicates that permission from chiefs is needed for any investments in housing quality, including any upgrading of roofs, an interaction that involves rent seeking and can therefore also be subject to ethnic patronage. Alternatively, as the quality of housing degrades, the tenant is likely to demand improvements and fixes for leaky roofs. Some of these will generate disputes which are then resolved by the chief, with the chief siding with the landlord or the tenant depending on the specific ethnic configuration. We test for these investment effects using satellite data to proxy for roof quality.

In the empirical section below, we start by describing our main specification. Then, we defend the identification assumptions needed for these effects to be interpreted as causal. Finally, we present a number of robustness checks to support our main results.

B Main Empirical Specification

We first outline our baseline specification. We start with a simple regression that illustrates the hypotheses we are interested in, as follows:

$$y_{ij} = \alpha + \beta^{lc} m^{lc}_{ij} + \beta^{hc} m^{hc}_{ij} + \beta^{hl} m^{hl}_{ij} + \gamma_l + \gamma_h + \Omega X_{ij} + \delta_j + \epsilon_{ij} \quad (1)$$

where y_{ij} is our outcome measure of interest (such as housing rents or luminosity) for household i in village j; m^{lc} is a dummy variable that takes value one if the tribes of the landlord and the location chief match (henceforth LC) and is zero otherwise; m^{hc} is similarly a dummy variable for the tribes of the household and the chief matching (HC); m^{hl} a dummy for a household-landlord tribe match (HL); the γ_l and γ_h are dummies for the tribes of the landlord and household, respectively (note we do not include dummies for the tribe of the chief as the village fixed effects control for these); X_{ij} are a set of controls; and δ_j are village fixed effects.

C Identification

We now discuss identification for the two main coefficients of interest, β^{lc} and β^{hc}. In equation (1) above, the identification assumption is that $Cov(m_{ij}^{lc}, \epsilon_{ij}|W_{ij}) = 0$ and $Cov(m_{ij}^{hc}, \epsilon_{ij}|W_{ij}) = 0$, where W_{ij} includes the set of controls in X_{ij} as well as all the dummy variables included in equation (1). Conditional on the covariates, village fixed effects and all the relevant tribe main effects, the matching of the landlord-chief and the household-chief tribes should be exogenous to unobserved household characteristics that determine rents.

To defend this assumption, we provide both some relevant history and background on the appointments of chiefs as well as a number of empirical checks using the characteristics of households and landlords in the area. We first discuss how chiefs are assigned to their positions, and the way in which they get transferred within Nairobi. We then describe the low mobility of residents and landlords within the slum. Finally, we propose three empirical checks that support our identification assumptions.

The chiefs operating in Kibera are employees of the Nairobi Provincial Administration and as such get assigned by the administration to their area of responsibility. Once appointed to a given location, these chiefs have an average tenure of approximately two years before they get moved to another area in Nairobi. This transfer process is entirely determined by the Provincial Administration. Appointments are made directly by the Office of the President of Kenya, or by the Provincial Administration through the Provincial Commissioner. Formally, this appointment process (whereby chiefs are vetted by the administration) implies that chiefs cannot sort across parts of the slum based on the characteristics of the residents or the landlords in any given location in the slum. The data we collected (which comprises all location and sublocation chiefs having been active in Kibera) shows that chiefs in Kibera are indeed frequently transferred. On average, since 2002, each of the chiefs spends 23 months on average as a chief of a given location in the slum: 3 chiefs stayed 3 months or less, 9 chiefs stayed for a period between 3 months and 2 years, and 5 chiefs stayed longer than 2 years. At the time of data collection in August 2012, the location chiefs currently active (not included in this calculation) had been appointed in June 2010, August 2010, May 2011 and October 2011, respectively.

One set of concerns with estimating a specification such as equation (1) involves sorting. Households could sort into areas with co-ethnic location chiefs, or alternatively choose housing where the landlord is of the same tribe as the chief. For example, an individual might get better access to public goods in areas governed by a chief belonging to the same tribe. Moving would then come with lower quality housing if the only housing available in the area is low-quality, and this could explain a negative coefficient on the household-chief tribe match. Similarly, landlords might be able to invest in their structures more easily when the chief is the same tribe as them, so that the housing they own is, in fact, of better quality (which would explain higher rents when landlords and chiefs are of the same tribe). As we show below, this is counter to the investment effects we find. Using measures of luminosity, we find that housing quality seems, in fact, lower when the tribe of the landlord and that of the chief match (although this estimate is not significant), and higher when the tribes of the household and the chief match.

We look more systematically at these types of concerns by providing three sets of empirical checks. First, to test for the exogeneity of chief appointments, we look at a range of characteristics of households in 2009 (from the KNBS census) and test whether a wide variety of education, wealth and employment measures predict our 2012 ethnic match variables. Given the tenure of the chiefs is about 23 months on average, we would expect these 2009 variables to predict matches if there was sorting by the chiefs or any choice on their part of which locations to be posted in. Second, we show these matches are largely uncorrelated with various measures of the ethnic composition of the household relative to its locality and measures of the ethnic diversity of the locality itself.

Third, we construct a regression discontinuity (RD) test in which we look at households living within a small distance of the internal location boundaries of the slum. We assume that households and local area characteristics are continuous across these boundaries, while the ethnicity of the location chief (who has the power to interfere with rents) changes discretely at the boundary. We use standard RD methods to look at whether rents and investments show a change across these internal location boundaries of the slum.

Note that we often also show the estimate for the household-landlord tribe match variable. It may be relevant to control for this variable but we do not interpret this effect as being causal. Our identification strategy is based on the tribe of the chief being exogenous to characteristics of the households and landlords. The tribe match of the households and landlords does not have any exogenous variation that allows us to interpret this effect as causal. In most specifications, the coefficient on this variable is very small in magnitude and statistically not significant. In addition, we report some specifications where we do not include this variable - the results are unchanged.

V Results

We present our main results, starting with the results from the baseline specification in (1) using measures of rents and luminosity as outcomes. We look at heterogeneity in these effects by measures of ethnic diversity, local economic activity and gang presence, proxied by youth unemployment. We then present results from our identification checks and the RD specification.

A Main Results

For clarity of exposition, we rewrite equation (1):

$$y_{ij} = \alpha + \beta^{lc} m_{ij}^{lc} + \beta^{hc} m_{ij}^{hc} + \beta^{hl} m_{ij}^{hl} + \gamma_l + \gamma_h + \Omega X_{ij} + \delta_j + \epsilon_{ij}$$

Table 2A shows the results from this regression where log household rent is the outcome of interest, and for a range of specifications where we progressively add more controls. We should note that the number of observations in these specifications is not the number of households surveyed. Overall, we surveyed 18,254 households, of which approximately 10% did not report either their own tribe or their landlord's tribe. This leaves us with a sample of 16,262 households, of which just over 14,300 pay rent. Throughout (unless otherwise specified) we cluster our standard errors at the level of location-household tribe-landlord tribe, i.e. at the level of the three tribe (chief, household and landlord) combinations. There are 224 such clusters in our data. In addition, we report the test statistic from an F-test for the null hypothesis that $\beta^{lc} + \beta^{hc} = 0$.

Before we report results, it may be useful to comment on the variation in the two match variables of interest. About 22% of households have a landlord who is a co-ethnic of the chief. About 14% of households belong to the same ethnicity as the chief themselves, and about 5% of households belong to the same ethnicity as both their landlord and the chief. This implies that there is a lot of variation in the landlord-chief and household-chief matches that does not imply that all three of the landlord, household and chief are co-ethnics. The correlation between the landlord-chief and the household-chief matches is only 0.14.

In column (1) of Table 2A, we report results without any controls and without village fixed effects. Note that chief tribe dummies are included in that specification. Households whose landlords are of the same tribe as the chiefs (LC match) pay 11% higher rents. The coefficient on the household-chief tribe match (HC match) is negative, but not significant. The coefficient on the household-landlord tribe match (HL match) is small in magnitude and insignificant. This is not surprising given the household-landlord tribe match is itself endogenous, and may correlate with household-level or dwelling-level unobservables that are positively associated with rents. In column (2), we add village fixed effects and in column (3) we add a set of household-level controls, in particular, controls for the demographic composition of the household (the number of adults and household size). The results are statistically unchanged. In column (4) we add a number of EA level controls [24] and in column (5) we include both household-level and EA-level controls. Our results are unchanged. In column (6), we add a dummy for whether all three of the household, the

[24] These are: a dummy for whether the EA was listed as a slum, average age of the household head, whether the head works, whether the head runs a business, whether the head works for the private sector, the head's years of education, a dummy for the head having no education, a dummy for the head having some secondary education, whether the household owns each of a number of assets (TV, radio, phone, bicycle), dummies for whether the walls of the house are made of mud and wood, whether the floor of the house is made of earth, whether the households main source of water is a water vendor, whether the households main waste disposal is an uncovered pit latrine, whether the households main cooking fuel is paraffin and whether the households main source of light is electricity.

landlord and the chief have the same tribe. The results on the landlord-chief tribe match remain significant with a coefficient of 8% as does the coefficient on the household and the chief belonging to the same tribe (-6%). The coefficient on all three tribes matching is small and not significant. In column (7), we include our data on the tribes of community elders. In particular, we control for dummies for an elder-landlord tribe match as well as an elder-household tribe match. Since elders are not exogenously rotated in the same way that chiefs are, we do not interpret these effects as causal but we show them for comparison.[25] In column (8), we include village dummies interacted with the fraction of the most common tribe in the zone. Finally, in column (9), we show results where we do not include the LH tribe match and the results are unchanged (this is true for all specifications).

In Table 2B, we show additional results on rents. In columns (1) and (2), we show the unweighted results, without and with controls. In columns (3) and (4), we show results where we do not cluster the standard errors. In columns (5) and (6), we report results for a sample obtained after trimming the top percentile of (log) rents. Iin columns (7) and (8) we report results obtained after dropping one village (Laini Saba) from the sample. We present this for comparison with Table 4A, where we also drop this village in some specifications. Finally, in column (9), we show results when we do not include the LH tribe match. The results in Table 2B are robust and similar to those in Table 2A.

In Table 3, we provide evidence on alternative measures of rents and on the extensive margin of the rental market. In column (1) we look at unconditional levels of rents (instead of logs) as our outcome. The rent value is set at zero if the household does not pay any rent. In column (2), we drop the LH tribe match variable. In columns (3) and (4), we use as our dependent variable the log of rents adjusted for luminosity (measured at the roof level), as a proxy for quality-adjusted rents. The LH tribe match dummy is again removed from the estimation in column (4). In columns (5) and (6), we use data on rents collected during the listing exercise for a (non-random) sub-sample of about 11,000 households. In columns (7) and (8), we look at the difference between the rents collected in the listing exercise and those collected during the household survey. These two measures were collected from a sub-sample of slighly more than 10,000 households.[26] Rents increased significantly (by 126 KShs on average) over the five-month period between the listing and the survey, and the correlation between the two rent measures is high (0.76). The sign of our estimates in columns (7)-(8) are consistent with our previous results: tenants facing a landlord-chief match report a higher increase in rent, and tenants with a co-ethnic chief report a decrease (non-significant). In column (9) we look at the extensive margin of rents by analyzing the effect of our tribe matches on whether the household pays rent. These columns exclude the landlord-chief match since for the subset of households that do not pay rent, we do not observe the landlord's tribe. The results are consistent

[25] Since some elders refused to be interviewed, column (7) has only about 11,800 observations.

[26] Rent data was only collected during the second wave of the listing exercise from approximately 18,600 households, and during the household survey. Of these 18,600, approximately 11,000 were also reached in the household survey, and 10,000 provided the rent amount they pay in both exercises.

with our earlier findings, though not significant: households with a co-ethnic chief are 3% less likely to pay rent.

B Investments

In Table 4A, we focus on the luminosity data from satellite images of the slum as our outcome. There is no simple way to measure housing investments and quality in a slum environment. In our sample, based on the 2009 census data, 96% of households have roofs made of corrugated iron. The luminosity data provides a proxy for housing quality that varies across households in the slum, and correlates with various socioeconomic characteristics including housing rents.

For luminosity, we look at three different specifications, and in each case we report results for the two different measures of luminosity (see Section 3.3 for a description of these two measures). We first look at the following regression specification:

$$L_{ijt} = \alpha + \beta^{lc} m_{ijt}^{lc} + \beta^{hc} m_{ijt}^{hc} + \beta^{hl} m_{ij}^{hl} + \gamma_c + \gamma_l + \gamma_h + \Omega X_{ij} + \delta_j + \mu_t + \epsilon_{ijt} \quad (2)$$

where L_{ijt} is the level of luminosity for household i in village j from period t. Note that we include dummies for the tribes of the chiefs, γ_c, in this specification as we have multiple years of data and thus the village dummies do not subsume the chief tribe dummies. There are four periods in the data spanning July 2009 to August 2012. In this analysis, we implicitly assume that very few households move within the period considered - households stay in the same structure and with the same landlord. This is consistent with the data we collected on households' duration of residence (households have lived in the same structure for 8.5 years, and with the same landlord for 6.8 years on average). We make this assumption as we do not have panel data on structures to capture how tribes change within a given structure. All the results presented in Table 4A are robust to dropping households that have lived in the same structure for less than two years. However, note that the tribe of the chief changes over time, so the match variables are now time varying and given by m_{ijt}^{lc} and m_{ijt}^{hc}. We include our regular set of controls, as well as the GPS coordinates of households (latitude and longitude) to capture other relevant spatial characteristics of the dwelling. We present results for the full sample (columns (2) and (4)) as well as the sample excluding one village, Laini Saba (columns (1) and (3)), for which the luminosity data is extremely different (there is a very high within-roof standard deviation in the luminosity measures for this village).

In columns (1) and (2), we focus on luminosity measured at the roof level. These regressions are two-way clustered at the roof and tribe combinations levels. In column (1), we find that roof quality is significantly higher when the household and the chief are of the same tribe. The coefficient on the landlord-chief tribe match is negative but not significant. In columns (3) and (4), we report similar results for our second measure of luminosity, the household measure. Here the regressions are two-way clustered at the tribe combination and at the longitude-latitude level (the latter to account for multiple households with the same set of GPS coordinates). Our estimates have the expected sign and are significant only in the case of a household-chief match.

In columns (5) through (8), we look at a different DGP for luminosity. We use the measures of the change in luminosity as a measure of investment in the housing infrastructure. In particular, we use this change between every pair of periods as our outcome and look at whether our ethnic match variables affect this measure of investment in the following specification:

$$DL_{ijt} = \alpha + \beta^{lc} m_{ijt}^{lc} + \beta^{hc} m_{ijt}^{hc} + \beta^{hl} m_{ij}^{hl} + \gamma_c + \gamma_l + \gamma_h + \Omega X_{ij} + \delta_j + \mu_t + \epsilon_{ijt} \tag{3}$$

where DL_{ijt} is the change in luminosity for household i in village j measured between period t and period $(t-1)$. In columns (5) and (7), we again exclude Laini Saba village from the sample.

We find similar results: the household-chief tribe match has large and significant positive effects on luminosity. The coefficient on the landlord-chief tribe match is not always significant. Comparing columns (5) and (6), we can see the role Laini Saba plays - the coefficients change, although the sample size only falls by about 5%. In column (7), we find that the two tribe match variables of interest have significant effects on investments: when the chief and the household belong to the same tribe, investments are higher and when the chief and the landlord belong to the same tribe, investments are lower. There is therefore no evidence that households whose landlord is a co-ethnic of the chief live in better housing. Similarly, for households who belong to the same tribe as the chief, investments are higher (with lower rents).

In columns (9) through (12), we take advantage of the panel nature of our data and report results from the following household fixed effects specification:

$$L_{ijt} = \alpha_i + \beta^{lc} m_{ijt}^{lc} + \beta^{hc} m_{ijt}^{hc} + \gamma_c + \mu_t + \epsilon_{ijt} \tag{4}$$

where α_i is a household fixed effect. Note β^{hl} is not identified as the tribes of the household and landlord do not change over time. This specification is identified from the tribes of the chiefs changing over time as they rotate in and out of locations in the slum. The chiefs do not rotate at the same time in each location - the rotations are staggered. Our panel of satellite images allows us to measure the effect of at least one change of chief in each location during the period considered. The results in columns (9) through (12) are very similar to the OLS specifications. We find strong effects of the household-chief tribe match on luminosity throughout. For the landlord-chief tribe match, we consistently find negative effects, though these are not significant.

In Table 4B, we use self-reported measures of roof renovations to corroborate our results on investments. In columns (1) through (6) we show that renovations reported by households predict the luminosity variables used in Table 4A.[27] We include our regular controls and GPS coordinates, and report results for the two different measures of luminosity. In all specifications, luminosity

[27] Households in our survey sample were asked when the roof of their dwelling was last renovated. We use this to construct a set of dummy variables indicating whether the household had a roof renovation since they moved in, in the past two years (since 2010), and in the past five years (since 2007).

correlates positively and significantly with the self-reported measures of renovations. In columns (7) through (9), we run our main specification (equation (2)) using self-reported renovations as the outcome. The sign of the estimate on the landlord-chief match is, as in our main results, negative and significant. The estimate on the household-chief match is close to zero and not significant.

C Heterogeneity

In Table 5, we look at the heterogeneity in our estimated effects along measures of ethnic diversity, local economic activity and youth unemployment. First, following Alesina and La Ferrara (2005), we look at the impacts of ethnic diversity on local economic activity. In the first 3 columns, we focus on a measure of ethnic diversity computed using a Hirschman-Herfindahl index (HHI) of the form: $H = 1 - \Sigma_{i=1}^{N} s_i^2$, where s_i is the fraction of the area population belonging to tribe i (a high index here stands for high ethnic diversity). However, ethnic diversity and youth unemployment correlate with various other socio-economic indicators (such as poverty) which may themselves play a role in these ethnic interactions. Therefore, throughout, we control for the interaction of the tribe match variables with the average level of education of the household head, the average wealth in the EA and the poverty rate in the EA.[28] As described below, our results are robust to adding these interactions.

Columns (1), (2) and (3) correspond to the three different area levels at which we compute this index: EA, zone and village, respectively. For ease of interpretation, the variable we use for the relevant interaction is a dummy for the diversity variable being greater than the median value in the sample. In columns (2) and (3), we find that in areas that are more ethnically diverse, residents where the landlord and the chief belong to the same tribe pay rents about 12% to 17% lower relative to less diverse areas. Local ethnic diversity therefore undoes the ethnic discrimination in rents that otherwise results from the landlord and the chief having the same tribe.

To add to these results, in columns (4) and (5) we look at the interaction of our match variables with measures of local private sector activity. The two variables we use are the number of retail kiosks and the number of M-PESA (mobile money) agents in the zone. Ethnic discrimination against tenants is reduced in areas where the local economy is more dynamic. Thus, the intensity of the ethnic bias in the rental market is *reduced* by local ethnic diversity and economic activity. These results suggest that an important component of the link between ethnic diversity and performance is private sector activity. Diverse areas seem to exhibit better functioning markets for private goods (housing). Correlations in our data also support the notion that the provision of public goods is lower in more ethnically diverse areas of the slum.

In columns (6) and (7), we test whether our results hold in areas with high youth unemploy-

[28] For wealth, we use a principal components analysis for the asset and economic variables from the 2009 census to create a wealth index. For poverty, we use the poverty mapping methodology in Elbers, Lanjouw and Lanjouw (2003) and data from Jack and Suri (2014) to create EA level poverty rates, i.e. the fraction of households in the EA that live on less than $2 a day.

ment. We use youth unemployment in the EA as a proxy for the local presence of youth gangs, since unemployed youth constitute the primary source of recruits for these gangs. In particular, we use a dummy for the EA-level youth unemployment rate being above the median. We find that high unemployment partially offsets the effects of the landlord-chief tribe match: households in this case no longer pay higher rents. The coefficient on the interaction with the household-chief tribe match is also negative, but mostly small and not significant. Since high youth unemployment may indicate a stronger role of gangs locally, and since gangs have been documented to protect co-ethnic tenants, we interpret this as evidence that the rent extraction organized by chiefs and landlords is hampered when local gangs are active in the area. In column (7), we see that the results are robust to controling for the interaction of all the match variables with education, poverty and wealth, illustrating that these heterogeneous effects do not come from the fact that EAs with higher youth unemployment are poorer.

D Identification Checks: Correlates of ethnic matches

In this section we present two sets of checks to support our identification assumptions.[29] In our first robustness check, we analyze whether our tribe match variables are correlated with observables from the 2009 census data. Our match variables come from data collected in 2012 and all the chiefs serving in 2012 were appointed after 2009. This allows us to check if chiefs were sorting into locations based on the 2009 socio-economic characteristics of the slum residents. Since these measures were collected in 2009 - before any of the current chiefs began their terms - we use them as a measure of "baseline" data. This data is aggregated at the level of an EA because the census data is de-identified. Table 6A reports the results. Each cell in the table represents a separate regression where the dependent variable is a tribe match dummy. We show results for 24 different right-hand side variables, correlating each with the relevant match dummy. We control for village fixed effects, main effects of the tribes and we cluster the regressions at the EA level. Column (1) reports the mean of each variable. Column (2) uses the dummy for a landlord-chief tribe match (m^{lc} in our notation) as the dependent variable, column (3) uses the dummy for the household-chief tribe match (m^{hc}), column (4) the dummy for the household-landlord tribe match (m^{hl}) and column (5), log rents.

In column (2), we see that only three of the 24 variables predict a match between the landlord and chief tribes (one is significant at the 5% level and two at the 10% level). In column (3), we find only one significant predictor of the match between household and chief tribes. In column (4), four variables predict the household-landlord match, in particular whether the household head works for pay, whether the household head owns a business, whether the household works for the private sector, and the poverty rate, with the first of these significant at 1%. It seems that households may be able to sort into the landlord of their choice at least to some extent, and that this sorting is based

[29] In the Online Appendix, we also show that our results are robust to survey attrition.

on tribe as well as socioeconomic variables. Finally, in column (5) of Table 6A, we test whether the 2009 characteristics are correlated with rents in 2012 in case the census data are just noisy. We find that a number of the census characteristics are significantly correlated with rents, especially those that describe housing quality. Thus, these variables do have predictive power.

As a second check, in Table 6B, we look at whether the tribe match variables correlate with measures of local ethnic composition. Here we regress ethnic composition variables on the tribe match dummies and tribe main effects. In columns (1), (2) and (3), we look at a dummy for whether the household belongs to the majority tribe. Majority is defined at the EA level in column (1), at the zone level in column (2) and at the village level in column (3). In all three columns, the landlord-chief and household-chief tribe matches do not correlate with whether the household is part of the majority tribe (while the landlord-household match does). Similarly, in columns (4) and (5) we look at a measure of ethnic diversity, measured at the EA and zone levels. Again, the landlord-chief tribe match does not correlate with local ethnic diversity. The household-chief tribe match exhibits a correlation when diversity is measured at the zone level. However, we should note that we do not correct for multiple comparisons in this table, so we would expect one of the ten coefficients to be significant.

E Regression Discontinuity Specification

As a second check, we look at a regression discontinuity (RD) specification where a household-chief tribe match represents the "treatment" and the distance to a location boundary represents the running variable - the distance to the treatment cutoff. The identifying assumption behind this RD design is that households living on either side of a location boundary are indistinguishable in terms of both observables and unobservables. In particular, household and local area characteristics (including accessibility and public goods or amenities) should be continuous across internal boundaries of the slum, while the ethnicity of chiefs who can interfere in rent disputes in favor of co-ethnic tenants changes discretely at these boundaries. One concern may be that public goods distribution may be different across these boundaries. However, the slum is very densely packed with public goods that are shared by everyone in the slum. Location chiefs have, in fact, very little power to influence the distribution of public goods in their areas of jurisdiction. Administratively, the country is split into districts which would be the relevant level for differences in public good provision. Here, the entire slum belongs to one district.

To implement this RD, we limit the sample to the set of households that can explicitly benefit from living in certain areas of the slum - areas governed by a co-ethnic chief. The relevant sample for this exercise is thus composed of all households who have at least one co-ethnic individual active as chief: in 2012, Kalenjins, Luos, Luhyas and Nubis. We exclude households who do not have a co-ethnic chief anywhere in Kibera. Of 31,765 households visited in the listing exercise, our sample for the RD analysis contains 20,271 households. Then, for any given tribe, the relevant “cutoff” is the boundary of the only location that is governed by a co-ethnic chief. For example, for Kalenjins, the relevant boundary is the border of the Sarangombe location (governed by a Kalenjin), and the running variable is the distance to that particular boundary.[30] We focus on this RD on the effect of the household-chief tribe match.

For this restricted sample of households, we use the local linear specification below and report results for a variety of bandwidths. Following common practice, we also report results for a third order polynomial specification without restricting the sample to bandwidths.

$$y_{ij} = \alpha + \beta m_{ij}^{hc} + \beta^R m_{ij}^{hc} * D_{ij} + \beta^L (1 - m_{ij}^{hc}) * D_{ij} + \Omega X_{ij} + \delta_j + \epsilon_{ij} \tag{5}$$

where y_{ij} is either rent or luminosity for household i in village j, and D_{ij} is the distance from the household to the relevant location boundary. For the local linear specifications, we look at four different bandwidths: 25m (close to the optimal bandwidth for all our outcomes), 50m, 100m and 150m from the boundary. In addition, we look at a polynomial specification that allows for up to third order polynomials in the distance, and their interactions with the match dummy.

Figure 5 shows the graphical version of our RD results, the top panel for log rents and the bottom panel for luminosity at the roof level. Table 7 reports the regression results. In columns (1) through (4), we look at local linear specifications for our four bandwidths. We do not include weights in these specifications as the weights are not applicable in the RD specification given the different sample. We find negative coefficients on the household-chief match, consistent with the results in Tables 2 and 3. In column (5), we look at the polynomial specification and find similar results. If anything, the magnitude of these effects is larger than in Table 2A. In columns (6) and (7), we show results for the smallest bandwidth and the polynomial specification where we include weights, purely for comparison. Finally, in columns (8) through (11), we present results for the change in luminosity for the smallest bandwidth and the polynomial specifications, restricting the analysis to just one period of data, separately for our two measures of luminosity (at the roof level and at the household level). We find similar results to Table 4. Overall, the RD specifications support our earlier results: having a co-ethnic chief allows households to pay lower rents and have better quality housing.

VI Conclusion

The mechanisms underlying the relationship between ethnic diversity and performance are not well understood in the context of developing economies. In this paper, we study two outcomes of

[30]Within the slum, Kalenjins have an average a distance of 495 meters to the boundary of Sarangombe location, and 14% of all Kalenjins are “treated” (i.e., they live in Sarangombe). Luhyas have an average distance of 530 meters to the location governed by a Luhya (Laini Saba), and 26% of all Luhyas live in Laini Saba. Luos live on average 811 meters away from the location governed by a Luo (Mugumo-ini), and 5% of Luos live there. Nubis live on average 321 meters away from the boundary of Kibera location, the one governed by a Nubi, and 89% of all Nubis live there.

ethnic fragmentation in a Kenyan slum: housing rents, and investments in the housing infrastructure. We show that ethnic patronage affects private markets in developing economies, in ways that are consistent with the existing literature on patronage in public goods and ethnic diversity.

In particular, we analyze the extent to which ethnic patronage affects housing rents and investments in the slum. We find that ethnicity plays an important role in determining the level of rents: households pay higher rents when the area chief is of the same tribe as the landlord, and lower rents when the household is of the same tribe as the chief. The primary responsibility of the chiefs is to arbitrate conflicts in the slum, a large fraction of which are rent disputes. Our findings reflect how chiefs side with landlords in these conflicts when belonging to the same tribe, and conversely with households when they share their tribe with the latter.

In addition, we find that ethnic relationships affect longer term outcomes in the slum via their impact on investments. We use satellite image data to compute measures of roof luminosity which proxy for investments in housing quality. We show that the landlord-chief ethnic match variable is not positively correlated with luminosity (as may have been expected given that rents are higher for these cases), but, in fact, negatively correlated (though not significantly so). Similarly, the household-chief match is not negatively correlated with luminosity, but instead has positive effects on investments. This implies that the ethnic patronage involving chiefs, landlords and tenants also has consequences for investments and welfare in the slum.

We find that the intensity of this ethnic bias is affected by the extent of local ethnic diversity. In particular, more diverse areas experience less discrimination against tenants in terms of rent paid when the chief and the landlord belong to the same tribe. The effect of ethnic diversity in this case is exactly similar to that of a more dynamic local economy, which also reduces ethnic bias on the rental market. We highlight this result because even in the context of our study, an analysis of the relationship between ethnic diversity and economic activity based solely on public goods provision would have pointed to the standard result in the literature - that more diverse areas exhibit lower access to public goods. However, the effect of ethnic diversity on the efficiency of local private goods markets seems to go in the opposite direction.

In addition, we find that our main results are weaker in areas with high youth unemployment, where youth gangs are likely to be more influential. In the context of Kibera, where conflicting claims on the land and landlord-tenant disputes are common, these gangs play an important role for price determination by limiting the extractive power of arbitrators - chiefs - and their ability to collude with landlords.

Finally, these results have important implications for our understanding of urban poverty across the developing world. The ethnic fragmentation and extensive systems of informal land rights prevalent in Kenyan slums are also common to other slums. This paper shows that such typical dimensions of urban slum living give rise to extensive systems of rent extraction that have important implications for investment, welfare and the functioning of markets. A more rigorous understanding of these extractive institutions will be necessary to design effective policies to tackle urban poverty.

VII References

Abrams, David, Marianne Bertrand and Sendhil Mulainathan. 2012. "Do Judges Vary in Their Treatment of Race." *Journal of Legal Studies*, 41(2): 347-383.

Acemoglu, Daron, Tristan Reed, and James Robinson. 2014. "Chiefs: Elite Control of Civil Society and Economic Development in Sierra Leone." *Journal of Political Economy* 122(2): 319-368.

Alesina, Roberto and Eliana La Ferrara. 2005. "Ethnic Diversity and Economic Performance." *Journal of Economic Literature* 43(3): 762-800.

Amis, Philip. 1984. "Squatters or Tenants: The Commercialization of Unauthorized Housing in Nairobi." *World Development* 12(1): 87-96.

Anwar, Shamena, Patrick Bayer, and Randi Hjalmarsson. 2012. "The Impact of Jury Race in Criminal Trials." *Quarterly Journal of Economics* 127(2): 1017-1055.

Banerjee, Abhijit, Paul Gertler and Maitreesh Ghatak. 2002. "Empowerment and Efficiency: Tenancy Reform in West Bengal." *Journal of Political Economy* 110(2): 239-280.

Bayer, Patrick, Marcus Casey, Fernando Ferreira and Robert McMillan. 2013. "Estimating Racial Price Differentials in the Housing Market." NBER Working Paper no. 18069.

Besley, Timothy. 1995. "Property Rights and Investment Incentives: Theory and Evidence from Ghana." *Journal of Political Economy* 103(5): 903-937.

Bleakley, Hoyt and Jeffrey Lin. 2012. "Portage and Path Dependence." *Quarterly Journal of Economics* 127(2): 587-644.

Bubb, Ryan. 2013. "State Law or Informal Norms." *Journal of Law and Economics* 56(3).

Burgess, Robin, Rémi Jedwab, Edward Miguel, Ameet Morjaria and Gerard Padro i Miquel. 2015. "The Value of Democracy: Evidence from Road Building in Kenya". *American Economic Review*, forthcoming.

Burgess, Robin, Matthew Hansen, Benjamin Olken, Peter Potapov and Stefanie Sieber. 2012. "The Political Economy of Deforestation in the Tropics." *Quarterly Journal of Economics* 127 (4): 1707-1754.

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

Easterly, William, and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112(4): 1203–1250.

Elbers, Chris, Jean Lanjouw and Peter Lanjouw. 2003. "Micro-Level Estimation of Poverty and Inequality." *Econometrica* 71(1): 355-364.

Fafchamps, Marcel. 2000. "Ethnicity and Credit in African Manufacturing." *Journal of Development Economics* 61: 205-235.

Field, Erica. 2007. “Entitled to Work: Urban Tenure Security and Labor Supply in Peru.” *Quarterly Journal of Economics* 122(4): 1561-1602.

Fitzgerald, John, Peter Gottschalk and Robert Moffitt. 1998. “An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics.” *Journal of Human Resources* 33: 251-299.

Franck, Raphael and Ilia Rainer. 2012. “Does the Leader’s Ethnicity Matter? Ethnic Favoritism, Education, and Health in Sub-Saharan Africa.” *American Political Science Review* 106(2): 294-325.

Galiani, Sebastian and Ernesto Schargrodsky. 2010. “Property Rights for the Poor: Effects of Land Titling.” *Journal of Public Economics* 94(9-10): 700-729.

Goldstein, Markus and Christopher Udry. 2008. “The Profits of Power: Land Rights and Agricultural Investment in Ghana.” *Journal of Political Economy* 116(6): 981-1022.

Habyarimana, James, Macartan Humphreys, Daniel Posner, and Jeremy Weinstein. “Why Does Ethnic Diversity Undermine Public Goods Provision?” *American Political Science Review* 101(4): 709-725.

Henderson, Vernon, Adam Storeygard and David Weil. 2012. “Measuring economic growth from outer space”. *American Economic Review* 102 (2): 994-1028.

Hjort, Jonas. 2014. “Ethnic Divisions and Production in Firms.” *Quarterly Journal of Economics*, forthcoming.

Hornbeck, Richard. 2010. “Barbed Wire: Property Rights and Agricultural Development.” *Quarterly Journal of Economics* 125(2): 767-810.

International Crisis Group. 2013. *Kenya’s 2013 Elections. Africa Report no. 197*. Brussels: International Crisis Group.

Jack, William and Tavneet Suri. 2014. “Risk Sharing and Transaction Costs: Evidence from Kenya’s Mobile Money Revolution.” *American Economic Review* 104(1): 183223.

Joireman, Sandra and Rachel Sweet Vanderpoel. 2011. “In Search of Order. State Systems of Property Rights and Their Failings” in *Where There is No Government: Enforcing Property Rights in Common Law Africa.*, ed. Sandra Joireman. New York: Oxford University Press.

Lanjouw, Jean and Philip Levy. 2002. “Untitled: A study of formal and informal property rights in Urban Ecuador.” *The Economic Journal* 112: 986-1019.

Michalopoulos, Stelios and Elias Papaioannou. 2013a. “On the Ethnic Origins of African Development: Traditional Chiefs and Pre-colonial Political Centralization.” Working Paper.

Michalopoulos, Stelios and Elias Papaioannou. 2013b. “Pre-Colonial Ethnic Institutions and Contemporary African Development.” *Econometrica* 81(1): 113-152.

Michelitch, Kristin. 2015. “Does Electoral Competition Exacerbate Interethnic or Interpartisan Economic Discrimination? Evidence from a Field Experiment in Market Price Bargaining.” *American Political Science Review* 109)1: 43–61.

Miguel, Edward, and Mary Kay Gugerty. 2005. “Ethnic Diversity, social sanctions, and public goods in Kenya.” *Journal of Public Economics* 89(11-12): 2325–2368.

Parsons, Timothy. 1997. ""Kibra is our blood": The Sudanese military legacy in Nairobi's Kibera location, 1902-1968." *The International Journal of African Historical Studies* 30(1): 87-122.

Price, Joseph, and Justin Wolfers. "Racial Discrimination Among NBA Referees." *Quarterly Journal of Economics* 125(4): 1859-1887.

Shayo, Moses and Asaf Zussman. 2011. "Judicial Ingroup Bias in the Shadow of Terrorism." *Quarterly Journal of Economics* 126: 14471484.

de Smedt, Johan Victor Adriaan. 2011. "The Nubis of Kibera: a social history of the Nubians and Kibera slums." Doctoral thesis, University of Leiden.

Syagga, Paul, Winnie Mitullah and Sarah Karirah-Gitau. 2002. "Nairobi Situation Analysis Supplementary Study: A Rapid Economic Appraisal of Rents in Slums and Informal Settlements'.' Contribution to the Preparatory Phase (January-November 2002) of the Government of Kenya & UN-HABITAT Collaborative Nairobi Slum Upgrading Initiative.

Temple, Nelle. 1974. "Housing preferences and policy in Kibera, Nairobi." Nairobi: University of Nairobi, Institute for Development Studies, Discussion Paper no. 196.

UN-Habitat. 2006. "State of the Worlds Cities 2006/2007."

UN-Habitat. 2012. "Sustainable Housing for Sustainable Cities: A Policy Framework for Developing Countries."

ONLINE APPENDIX: NOT FOR PUBLICATION

Appendix 1: Attrition

The survey we use had some attrition. Here, we provide two checks on attrition, both of which illustrate that attrition does not seem to play an important role.

The attrition in our survey comes from the households that could not be reached by phone and those that did not have phones. In particular, of the households that could not be reached by phone, we sampled 20% and focused on finding this subset of households in the field to complete the survey. We were able to locate 80% of target respondents in this subsample. In addition, of the households without phone numbers, we sampled 14% to conduct field surveys and were able to find 77% of these. As shown in Appendix Figure A1, the total target sample size was 18,797 households, and the number of households we reached was 18,254, giving us an unweighted attrition rate of 3% and a weighted attrition rate of 11% (sampling weights are used throughout all of our analysis).[31]

These attrition rates are obviously not negligible, although they are not strikingly high for urban areas (for example, Jack and Suri (2014) find higher attrition rates in Nairobi). The context of Kibera makes tracking households harder for a number of reasons. First, we conducted our surveys during weekends, but not in the evenings because of safety concerns for our staff. Households living the slum are often out looking for work or travelling long distances for work and therefore not easily available during the day, including on weekends. Second, even though our surveys were conducted shortly (6 months) after the initial listing exercise, there are reasonably high mortality rates in Kibera which arguably made it more difficult to locate our entire target sample. In our survey data, 4% of respondents reported that at least one member of their household died in the past 6 months.

However, since this does not explain all the attrition in the sample, we present two checks to show that attrition is not driving our results, which we report in Table A1. The first check is a specification where we drop the higher attrition EAs. We keep the EAs that have attrition rates of 5% or less. Results for our two main specifications are reported in columns (1) and (2) of Table A1. As can be seen, our results are robust to dropping the higher attrition EAs. The second check we conduct is to weight the regressions as proposed by Fitzgerald, Gottschalk and Moffitt (1998). Once again, our results are mostly robust to this re-weighting as seen in columns (3) and (4).

[31] By definition of our sampling strategy, attrition was null in the subsample of households that we surveyed by phone. Attrition is therefore highest in the subset of households without phone, and the subset of households who could not be reached on their phones, which also have the largest sampling weights by way of the sampling strategy.

Appendix 2: Geospatial Methods and Satellite Imagery Analysis

A2.1. Data collection

Our initial listing exercise in Kibera involved collecting GPS coordinates for every household in our sample. Coordinates were received and recorded at the center of each dwelling using Garmin GPS receivers, which are usually accurate within 15 meters or less (`http://www8.garmin.com/aboutGPS/`). Households whose coordinates could not be placed within Kibera after data cleaning were dropped from our sample. To improve the geo-positioning of our satellite pictures, we also collected GPS coordinates for all major installations in Kibera (schools, churches, etc.) and major arteries and boundaries, including the railway that cuts through the slum through a West-East axis.

A2.2. Mapping of Kibera

After the end of the first wave of listing, we used various sources of information to create our own map of the Kibera area, including location, sublocation, village and enumeration area boundaries. The villages are the finest level of governance we could map with accuracy (some villages are divided into zones, but our data does not allow us to precisely demarcate these zones). Kibera is divided into 17 villages, which are not administrative entities but whose boundaries are well-known by residents of the slum. To construct the map of villages, we used the information reported by households (who were asked to name their village and zone in the listing exercise), a land use shapefile made publicly available by the Center for Sustainable Urban Development (CSUD) at Columbia University, various maps created by the Map Kibera project (`http://mapkibera.org/`), our GPS coordinates of major installations in Kibera, and our own satellite pictures of the area which helped us determine relevant boundaries when those are marked by a street or a river. The Enumeration Area boundaries (drawn as part of the 2009 Population Census) were manually digitized from maps given to us in paper format by the Kenyan National Bureau of Statistics.

A2.3. Description of the Luminosity Data

We use four high-resolution (0.5 meters) panchromatic pictures captured over the Kibera area in July 2009, January 2011, December 2011 and August 2012. The imagery was acquired from two different remote sensing operators, GeoEye and DigitalGlobe, and processed through an independent remote sensing/GIS consultant. For each date of acquisition, the raw data generally consisted of two pictures, one with a panchromatic resolution of .5 meters and one with a multispectral resolution of 2 meters, over an area of interest (AOI) comprised between 36.77 and 36.89 longitude and -1.298 and -1.324 latitude. This AOI covers about 35 square kilometers within the city of Nairobi. To optimize the accuracy of the geo-positioning and the superposition of the pictures over time, the consultant ortho-rectified the raw data using regional geo-reference points and a shapefile of land use in Nairobi released to the public by the Center for Sustainable Urban Development at

Columbia University. The pictures were later trimmed to fit the boundaries of the Kibera area, comprised between 36.77 and 36.81 longitude and -1.305 and -1.321 latitude (approximately 7.91 square kilometers).

To extract luminosity data from the satellite pictures, we use the *Zonal Statistics (Spatial Analyst)* tool provided by ESRI ArcGIS. The ZS tool extracts statistics over areas defined by a zone dataset, based on the underlying value in the image or raster dataset. On our pictures, the dynamic range of the luminosity data (the value raster dataset) is 11-bit radiometric (2,048 levels). This roughly corresponds to a grey scale with 2,048 unique values, where low values correspond to dark areas and high values to bright areas. The statistics computed through the ZS tool include the mean, median, standard deviation, minimum, and maximum luminosity (pixel value) over any given area of interest. Since all of our pictures have a different average luminosity (corresponding mostly to different levels of solar radiation), all of our specifications that span different periods include picture fixed effects. Below we describe the two different levels (roof and household) at which we compute the luminosity data.

A2.4. Extraction of the luminosity data at the roof level

To demarcate dwelling roofs on our pictures, we used a different software specifically designed for picture segmentation, eCognition Developer 8 (Trimble). eCognition segments pictures into homogenous objects by aggregating neighboring pixels with similar values on the radiometric scale, yielding objects that can be as small as one pixel. The user can alter the segmentation algorithm by entering values for three different parameters: a scale parameter, a shape parameter and a compactness parameter. The scale parameter defines the maximum color gap between objects produced by the segmentation algorithm. A higher scale parameter will result in the segmentation producing larger objects, and vice versa. The shape parameter determines how much the shape of objects (as opposed to color) influences the segmentation process. The compactness defines how compact (as opposed to smooth) the produced objects will be. The output of the eCognition algorithm can be exported into ArcGIS in the form of a shapefile.

Our delimitation of roofs was done in three steps. First, we segmented the pictures into objects corresponding to the 17 Kibera villages, using the Chessboard Segmentation tool of eCognition with our shapefile of village boundaries as the underlying thematic layer. Second, we used the Multi-Resolution Segmentation (MRS) algorithm to segment these objects into roofs or blocks or roofs. The appropriate values of the parameters of interest were determined through trial-and-error - we compared the segmented pictures obtained after defining different parameter values, and chose the unique combination of parameter values that produced the best segmentation. We use a scale parameter of 50, a shape parameter of 0.8, and a color parameter of 0.2. Third, we manually rectified the roof objects in ArcGIS based on simple visual checks.

A2.5. Extraction of the luminosity data at the household level

In addition to the roof-level analysis, we present the luminosity data reflected by the roof surface situated just above the households in our sample. The main goal of this particular method is to obtain a measure that is specific to each household. One advantage of this technique is that it reduces noise due to large roofs accommodating multiple households. We do this by computing, for each household, the same set of luminosity statistics as above, but this time only over an area comprised within one meter of each household. Since our GPS devices are not accurate within one meter, we expect some measurement error in this computation.

A2.6. Empirical validation of the methodology

We now provide evidence that the data from our satellite images are indeed correlated with a range of socioeconomic outcomes and housing characteristics, using data from the 2009 census. Since EAs are the geographical unit of reference in the KNBS census, all the luminosity data available at a 0.5 meters resolution is aggregated to the level of the EA for this empirical validation. Throughout all specifications, we use the sample of only informal EAs (576 out of 643 EAs), we weight the regression by the size of the EA, and we control for sublocation fixed effects.

Our first specification looks at the correlation between luminosity and a variety of individual level demographic, education and occupation characteristics. The regression is as follows:

$$L_{ij} = \alpha + \beta_w WP_{ij} + \beta_b B_{ij} + \beta_s Sec_{ij} + \beta_u Univ_{ij} + \beta_a Age_{ij} + \eta_j + \epsilon_{ij}$$

where L_{ij} is the mean luminosity reflected by EA i in sublocation j, WP_{ij} is the average fraction of household members working for pay in the EA, B_{ij} is the average fraction of household members running a business, Sec_{ij} is the fraction of household heads that have some secondary education, $Univ_{ij}$ is the fraction of household heads that have some university education, Age_{ij} is the average age of the household head, η_j is a set of sublocation fixed effects, and the coefficients of interest are the five β coefficients.

We report the results of this specification in the first column of Appendix Table A2. We find that higher luminosity is associated with higher secondary education and (not significantly) with higher university education. It is also correlated with the fraction of household heads that run a business and with the fraction who report having some form of paid employment.

We then report on a specification that includes a number of standard wealth measures:[32]

$$L_{ij} = \alpha + \beta_m WM_{ij} + \beta_f F_{ij} + \beta_c C_{ij} + \beta_a WA_{ij} + \beta_e E_{ij} + \eta_j + \epsilon_{ij}$$

where WM_{ij} is the fraction of households in the EA with walls made of mud mixed with cement,

[32] These measures are those used in the poverty mapping literature (Elbers et al. (2003)) and in Demographic and Health Surveys (DHS) conducted across the developing world.

F_{ij} is the fraction of households with cement floors, C_{ij} is the fraction of households whose main cooking fuel is charcoal, WA_{ij} is the fraction of households whose main water source is a water vendor, and E_{ij} is the fraction of households whose main light source is electricity.

Appendix Table A2, column (2) reports the results from this specification. Luminosity correlates significantly and positively with the floor being made of cement (the highest-quality floor material) and with the main source of drinking water being a water vendor.[33] Luminosity is also significantly negatively associated with a mix of mud and cement being used as the material to construct their walls and with charcoal being used as cooking fuel - two indicators of relative deprivation within the slum. The coefficient on access to electricity is insignificant and negative - possibly because segments of the slum are illegally connected to the city electricity grid.

Finally, we look at correlations of luminosity with measures of household expenditures and poverty. The census does not collect these measures, instead we estimate them using the poverty mapping methods described in Elbers et al. (2003) and data from Jack and Suri (2014). We look at household expenditures and two different measures of poverty: the fraction of households in the EA that live on less than \$1.25 (PPP) a day and the fraction that live on less than \$2 (PPP) a day. Since these are generated regressors, we report bootstrap standard errors. In addition, we look at correlations of luminosity with retailer presence and road access in the EA. Since these five variables are highly correlated with each other, column (3) reports coefficients from separate regressions of each of these variables on luminosity, using the following specification:

$$L_{ij} = \alpha + \beta_x X_{ij} + \eta_j + \epsilon_{ij}$$

where X_{ij} is one of the economic variables of interest, i.e. one of the following: (i) the fraction of households living on less than \$1.25/capita/day; (ii) the fraction of households living on less than \$2/capita/day; (iii) average per capita consumption in the EA; (iv) whether there is a retail business in the EA; and (v) the fraction of the area of the EA that is within 15 meters of a road.

We find that the EA luminosity correlates positively with estimated household expenditures and negatively with the two poverty measures. Appendix Figure A2 presents a visual analysis of these results, comparing average luminosity and average consumption across EAs. To further support these results, we find that EAs with higher luminosity are more likely to have a retail business and are likely to be closer to roads.

[33]These households typically pay for clean water, as opposed to using poorer quality sources such as public pumps.

Table 1A: Survey Summary Statistics

	Mean	SD	N
Listing (Census) Data			
Number of adults in the household	2.16	2.03	31718
Household size (adults and children)	3.65	2.11	31717
Amount paid in rent, census sample	1678.55	2105.90	18679
Kalenjin tribe	0.01	0.12	28890
Kamba tribe	0.15	0.36	28890
Kikuyu tribe	0.06	0.23	28890
Kisii tribe	0.06	0.24	28890
Luhya tribe	0.27	0.44	28890
Luo tribe	0.36	0.48	28890
Nubi tribe	0.05	0.23	28890
Other tribe	0.02	0.14	28890
Majority Tribe	0.44	0.50	28890
Household Survey Data			
Household pays rent	0.92	0.28	17251
Amount paid in rent (levels, KShs)	1715.30	1784.06	15473
Difference in rent between census and survey	101.56	1209.98	11019
Household ever evicted in Kibera	0.11	0.32	17216
Roof renovated since tenant moved in	0.23	0.42	17196
Roof renovated in last two years	0.18	0.38	17196
Roof renovated in last five years	0.21	0.41	17196
Landlord Kalenjin tribe	0.00	0.06	15620
Landlord Kamba tribe	0.08	0.27	15620
Landlord Kikuyu tribe	0.33	0.47	15620
Landlord Kisii tribe	0.05	0.22	15620
Landlord Luhya tribe	0.11	0.31	15620
Landlord Luo tribe	0.17	0.37	15620
Landlord Nubi tribe	0.22	0.41	15620
Number of years in Kibera	15.55	11.12	17067
Number of years with the same landlord	6.91	6.82	15884
Number of years in the same structure	8.47	8.97	17153
Match Variables			
Landlord-Chief (LC) tribe match	0.22	0.41	15620
Household-Chief (HC) tribe match	0.14	0.34	17742
Landlord-Household (LH) tribe match	0.28	0.45	14814
All tribes match (ALL)	0.05	0.21	14814

Note: The only census data we use throughout the paper is the tribe of the household.
The household survey is a stratified random sample taken from the census data.
The data on the tribe of the landlord is collected in the household survey data.
The tribe of the chief comes from the administrative history of chief lineage.
The exchange rate at the time of the household survey was 80 KShs to the US $.

Table 1B: Summary Statistics: Other Data

	Mean	SD	N
Satellite Image Data			
Luminosity, roof level	307.33	70.11	75062
Luminosity, household level	304.69	78.71	75188
Luminosity, roof level, trimmed	304.23	51.39	73663
Luminosity, household level, trimmed	301.23	57.19	73782
Luminosity change, roof level	20.12	85.88	56293
Luminosity change, roof level, trimmed	20.55	56.93	55208
Luminosity change, household level, trimmed	19.91	64.23	55331
2009 Census Data, EA Level			
EA listed as informal (slum)	0.92	0.27	608
Age of household head	35.43	2.11	608
Household head works for pay	0.68	0.16	608
Household head owns a business	0.15	0.13	608
Household head works for private sector	0.41	0.21	608
Household head has no education	0.04	0.04	608
Household head has some secondary education	0.42	0.10	608
Household head, years of education	9.29	1.44	608
Hours worked in last 7 days by household head	55.68	9.08	608
TV	0.42	0.20	608
Radio	0.76	0.10	608
Mobile phone	0.83	0.10	608
Bicycle	0.06	0.07	608
Walls are made of mud/wood	0.29	0.30	608
Floor is made of earth	0.37	0.29	608
Main water source is water vendor	0.23	0.39	608
Main waste disposal is uncovered pit	0.22	0.34	608
Main cooking fuel is paraffin	0.53	0.21	608
Main light source is electricity	0.52	0.29	608
Fraction of youth unemployed in household	0.17	0.12	608
Fraction of youth unemployed in household, trimmed	0.17	0.11	602
Fraction of youth unemployed in EA	0.17	0.12	608
Fraction of youth unemployed in EA, trimmed	0.17	0.11	602
Tenure is individual	0.87	0.21	608
Tenure is purchased	0.06	0.17	608
Tenure is inherited	0.02	0.06	608
Household owns housing (purchase or inherit)	0.08	0.18	608
Dummy for no ownership	0.16	0.36	608
Household has piped water into its dwelling	0.36	0.48	608
Wealth index (principal components)	-0.00	2.26	608
Poverty ($1.25/day)	0.14	0.08	608
Poverty ($2/day)	0.42	0.14	608
Community Elder Survey and Zone Level Data			
Was the area visited by a government official	0.67	0.48	33
No of kiosks in zone	81.47	94.17	32
No of M-PESA agents in zone	10.30	12.58	33
Ethnic fractionalization index by zone	0.69	0.07	33
Village Level Data			
Ethnic fractionalization index by village	0.69	0.09	16

Note: Luminosity data is from satellite images for Jul 2009, Jan 2011, Dec 2011, Aug 2012. The trimmed luminosity data drop the top 1% and the bottom 1% of observations.

Table 2A: Rents

	(1) Log Rent	(2) Log Rent	(3) Log Rent	(4) Log Rent	(5) Log Rent	(6) Log Rent	(7) Log Rent	(8) Log Rent	(9) Log Rent
Landlord-Chief (LC) tribe match	0.108***	0.071***	0.070***	0.069**	0.067**	0.077***	0.084**	0.060**	0.066**
	[0.040]	[0.026]	[0.026]	[0.027]	[0.026]	[0.025]	[0.033]	[0.028]	[0.026]
Household-Chief (HC) tribe match	-0.047	-0.081***	-0.080***	-0.074***	-0.073***	-0.064***	-0.069**	-0.066***	-0.072***
	[0.041]	[0.027]	[0.026]	[0.026]	[0.025]	[0.022]	[0.028]	[0.025]	[0.024]
Landlord-Household (LH) tribe match	-0.034	0.006	0.006	0.005	0.006	0.011	0.020	0.014	
	[0.026]	[0.019]	[0.019]	[0.019]	[0.018]	[0.019]	[0.021]	[0.021]	
All tribes match (ALL)						-0.038			
						[0.057]			
Landlord-Elder (LE) tribe match							-0.023		
							[0.020]		
Household-Elder (HE) tribe match							-0.021		
							[0.024]		
Test Stat for LC+HC=0	0.061	-0.010	-0.010	-0.005	-0.007	0.014	0.015	-0.007	-0.006
SE	[0.063]	[0.043]	[0.042]	[0.042]	[0.041]	[0.037]	[0.049]	[0.043]	[0.041]
Test Stat for LC+HC+LH+ALL=0						-0.013			
SE						[0.055]			
Dependent Variable Mean	7.255	7.255	7.255	7.254	7.254	7.254	7.313	7.313	7.254
Village FE		X	X	X	X	X	X	X	X
HH Controls			X		X	X	X	X	X
EA Controls				X	X	X	X	X	X
Zone Tribe Controls								X	
R-squared	.09	.4	.405	.419	.426	.426	.329	.333	.426
Observations	14311	14311	14277	14236	14202	14202	11823	11823	14202

Note: Standard errors clustered at the location-household tribe-landlord tribe level (reported in brackets). * $p<0.1$, ** $p<0.05$, *** $p<0.01$.
Rents are monthly rents paid. Observations vary across columns due to occasional missing observations for the various covariates.
Community elders in one village refused to be interviewed - specifications with zone controls or elder matches therefore have less observations.
Results are identical when the sample is restricted to the observations with non-missing data.
Throughout the paper, data on rents and household controls are from the household survey.
EA controls are from the 2009 census data. The zone dummies and elder match variables are from the community elder survey.

Table 2B: Rents: Alternative Specifications

	(1) Log Rent	(2) Log Rent	(3) Log Rent	(4) Log Rent	(5) Log Rent	(6) Log Rent	(7) Log Rent	(8) Log Rent	(9) Log Rent
Landlord-Chief (LC) tribe match	0.065***	0.064***	0.071***	0.067***	0.076***	0.072***	0.053**	0.049**	0.049**
	[0.020]	[0.019]	[0.021]	[0.020]	[0.027]	[0.027]	[0.022]	[0.022]	[0.022]
Household-Chief (HC) tribe match	-0.073***	-0.068***	-0.081***	-0.073***	-0.075***	-0.069**	-0.072***	-0.066**	-0.066**
	[0.019]	[0.018]	[0.021]	[0.020]	[0.028]	[0.027]	[0.026]	[0.026]	[0.026]
Landlord-Household (LH) tribe match	-0.008	-0.004	0.006	0.006	0.012	0.012	-0.004	-0.002	
	[0.015]	[0.014]	[0.015]	[0.015]	[0.020]	[0.019]	[0.018]	[0.017]	
Test Stat for LC+HC=0	-0.007	-0.004	-0.010	-0.007	0.000	0.002	-0.019	-0.017	-0.017
SE	[0.026]	[0.026]	[0.028]	[0.027]	[0.045]	[0.043]	[0.038]	[0.038]	[0.038]
Dependent Variable Mean	7.255	7.254	7.255	7.254	7.232	7.231	7.248	7.247	7.247
Village FE	X	X	X	X	X	X	X	X	X
HH Controls		X		X		X		X	X
EA Controls		X		X		X		X	X
No Weights	X	X							
No Clustering			X	X					
Trim Top Percentile					X	X			
Drop Laini Saba							X	X	X
R-squared	.39	.421	.4	.426	.36	.389	.409	.436	.436
Observations	14311	14202	14311	14202	14174	14065	13516	13410	13410

Note: Standard errors clustered at the location-household tribe-landlord tribe level (reported in brackets). * $p<0.1$, ** $p<0.05$, *** $p<0.01$.
The standard errors reported in columns (3) and (4) are heteroskedasticity robust standard errors.

Table 3: Other Measures of Rents

	(1) Rent	(2) Rent	(3) Log Adjusted Rent	(4) Log Adjusted Rent	(5) Log Listing Rent	(6) Log Listing Rent	(7) Change in Rent	(8) Change in Rent	(9) Pay Rent
LC tribe match	106.049**	106.006**	0.072***	0.072***	0.071***	0.071***	26.010*	25.931*	
	[46.652]	[46.276]	[0.026]	[0.026]	[0.023]	[0.023]	[13.437]	[13.340]	
HC tribe match	-81.916*	-82.149*	-0.073***	-0.073***	-0.071***	-0.071***	-18.788	-18.777	-0.033
	[46.451]	[46.177]	[0.024]	[0.024]	[0.023]	[0.023]	[17.739]	[17.638]	[0.023]
LH tribe match	7.882		-0.001		0.007		-4.295		
	[34.174]		[0.017]		[0.016]		[13.299]		
Test Stat LC+HC=0	24.134	23.857	-0.001	-0.001	-0.000	0.000	7.222	7.154	
SE	76.106	75.303	0.039	0.039	0.029	0.029	21.177	20.892	
Dep Var Mean	1627.389	1627.389	1.450	1.450	7.157	7.157	126.107	126.107	0.924
All Controls	X	X	X	X	X	X	X	X	X
R-squared	.399	.399	.423	.423	.533	.533	.015	.015	.155
Observations	14089	14089	14189	14189	11006	11006	10444	10444	16262

Note: Standard errors clustered at the location-household tribe-landlord tribe level (reported in brackets). * $p<0.1$, ** $p<0.05$, *** $p<0.01$.
All controls refers to village fixed effects, household controls and EA controls. These are all included across all specifications.
Rent levels (columns (1), (2), (5), (6)) are measured in Kenyan Shillings (KShs 80 to the dollar). The top percentile of rents is trimmed.
In columns (3)-(4) the dependent variable is rent adjusted by luminosity calculated at the roof level.
Census rents (columns (5)-(6)) were not collected for all households hence the smaller sample size. These regressions are not weighted.
The change in rent (columns (7)-(8)) is the difference between the census and household survey rents. These regressions are not weighted.
The pay rent dummy (columns (9)) indicates whether the household paid any rent (those that do not pay do not have landlords).
The standard errors in the last column are clustered at the location-household tribe level.

Table 4A: Investments in the Slum: Luminosity and Self Reports

	Luminosity Level, OLS				Luminosity Change				Luminosity Level, Panel			
	(1) Roof	(2) Roof	(3) HH	(4) HH	(5) Roof	(6) Roof	(7) HH	(8) HH	(9) Roof	(10) Roof	(11) HH	(12) HH
LC tribe match	-1.079	-0.729	-1.890	-1.449	-0.340	1.083	-3.328***	0.281	-0.583	-0.255	-1.570	-0.815
	[1.477]	[1.346]	[1.662]	[1.550]	[1.551]	[1.470]	[1.269]	[2.107]	[1.152]	[1.106]	[1.472]	[1.548]
HC tribe match	2.246**	2.184**	2.361**	1.926*	3.604**	2.145	4.463***	1.597	2.936***	2.541**	2.868**	2.056
	[1.008]	[0.878]	[1.132]	[1.064]	[1.731]	[1.488]	[1.602]	[1.975]	[1.034]	[0.979]	[1.234]	[1.320]
LH tribe match	0.945	1.465*	1.879*	1.899*	-0.189	-0.038	-0.526	-0.343				
	[0.869]	[0.888]	[1.010]	[0.988]	[0.235]	[0.229]	[0.491]	[0.470]				
Test Stat LC+HC=0	1.167	1.454	0.471	0.478	3.263	3.228	1.135	1.878	2.353*	2.286*	1.298	1.241
SE	[1.336]	[1.223]	[1.650]	[1.475]	[2.246]	[1.987]	[1.644]	[1.913]	[1.211]	[1.161]	[1.593]	[1.404]
Dep Var Mean	303.342	303.629	300.733	300.811	20.371	20.306	20.175	20.024	303.342	303.629	300.733	300.811
Drop Laini Saba	X		X		X		X		X		X	
Roof Clustering	X	X			X	X						
GPS Clustering			X	X			X	X			X	X
Tribe Clustering	X	X	X	X	X	X	X	X	X	X	X	X
All Controls	X	X	X	X	X	X	X	X				
R-squared	.374	.361	.301	.295	.179	.177	.151	.15	.485	.473	.411	.408
Observations	54621	57704	54695	57756	41017	43290	41057	43352	54621	57704	54686	57744

Note: Standard errors clustered at the levels indicated in the table (reported in brackets). * p<0.1, ** p<0.05, *** p<0.01.
All controls refers to village fixed effects, household and EA controls. All specifications also control for period fixed effects and GPS coordinates.
The luminosity data is extracted from satellite images of Kibera for four periods, July 2009, January 2011, December 2011, August 2012.
All results for luminosity use the trimmed luminosity data (top and bottom one percentiles removed).
Roof represents the roof level measure of luminosity and HH represents the household measure. See text for details on how these are computed.
The panel regressions cannot be clustered at the roof level since roofs cannot be matched across periods (or pictures).

Table 4B: Investments in the Slum: Luminosity and Self Reports

	(1) Luminosity Roof	(2) Luminosity HH	(3) Luminosity Roof	(4) Luminosity HH	(5) Luminosity Roof	(6) Luminosity HH	(7) Roof Renovated Ever	(8) Roof Renovated 2 Years	(9) Roof Renovated 5 Years
Roof renovated since tenant moved in	2.556**	2.173**							
	[1.072]	[1.014]							
Roof renovated in last two years			2.013**	2.259**					
			[0.950]	[1.124]					
Roof renovated in last five years					1.653*	2.019*			
					[0.852]	[1.041]			
LC tribe match							-0.037**	-0.030**	-0.034**
							[0.017]	[0.014]	[0.017]
HC tribe match							0.005	0.003	0.001
							[0.017]	[0.016]	[0.015]
LH tribe match							0.013	0.002	0.003
							[0.013]	[0.012]	[0.013]
Dep Var Mean	332.186	329.357	332.186	329.357	332.186	329.357	0.224	0.179	0.207
Roof Clustering	X		X		X				
GPS Clustering		X		X		X			
Tribe Clustering							X	X	X
All Controls	X	X	X	X	X	X	X	X	X
R-squared	.104	.066	.104	.066	.103	.066	.019	.019	.017
Observations	16994	17019	16994	17019	16994	17019	14693	14693	14693

Note: Standard errors clustered at the levels indicated in the table (reported in brackets). * $p<0.1$, ** $p<0.05$, *** $p<0.01$.
All controls refers to village fixed effects, household and EA controls. In addition, all specifications control for GPS coordinates.
All results for luminosity use the trimmed luminosity data (top and bottom one percentiles removed).
Roof represents the roof level measure of luminosity and HH represents the household measure. See text for details on how these are computed.

Table 5: Heterogeneity along Measures of Ethnic Diversity and Youth Unemployment

Dependent Variable is Log Rent

	(1) Diversity, EA	(2) Diversity, Zone	(3) Diversity, Village	(4) Kiosks in Zone	(5) M-PESA Agents in Zone	(6) High UE	(7) High UE
LC tribe match*Variable	-0.012	-0.119***	-0.181***	-0.104***	-0.087**	-0.075***	-0.089***
	[0.040]	[0.035]	[0.040]	[0.038]	[0.040]	[0.019]	[0.019]
HC tribe match*Variable	-0.070	-0.082	-0.075	-0.119**	-0.060	-0.016	-0.009
	[0.052]	[0.060]	[0.056]	[0.058]	[0.052]	[0.028]	[0.026]
LH tribe match*Variable	-0.009	0.036	0.010	0.045	0.000	0.024	0.023
	[0.039]	[0.032]	[0.038]	[0.046]	[0.045]	[0.020]	[0.022]
LC tribe match	0.059	0.116***	0.172***	0.089**	0.107**	0.109***	0.109***
	[0.045]	[0.040]	[0.046]	[0.044]	[0.049]	[0.029]	[0.033]
HC tribe match	-0.018	-0.031	-0.011	-0.061	-0.046	-0.064**	-0.058
	[0.056]	[0.052]	[0.065]	[0.040]	[0.048]	[0.031]	[0.040]
LH tribe match	0.036	0.006	0.023	0.072**	0.055	-0.005	0.013
	[0.040]	[0.034]	[0.039]	[0.034]	[0.034]	[0.021]	[0.039]
Dependent Variable Mean	7.254	7.254	7.254	7.313	7.313	7.254	7.254
Education Interac	X	X	X	X	X		X
Wealth Interac	X	X	X	X	X		X
Poverty Interac	X	X	X	X	X		X
All Controls	X	X	X	X	X	X	X
R-squared	.427	.429	.429	.33	.332	.43	.43
Observations	14202	14202	14202	11823	11823	14202	14202

Note: Standard errors clustered at the location-household tribe-landlord tribe level (reported in brackets). * $p<0.1$, ** $p<0.05$, *** $p<0.01$

All controls refers to village fixed effects, household and EA controls.

Column titles refer to the variable that is interacted with the match dummies.

The diversity indices are based on a Hirschman-Herfindahl index, as defined in the paper.

The interactions in all columns are with dummy variables for the relevant variable being greater than the median in the sample.

The Education, Wealth and Poverty Interac controls include levels of the variable and its interactions with all the match variables.

Wealth is an EA level principal components index using 2009 data.

The poverty rate (less than $2 a day) is at the EA level and estimated using the methodology in Elbers, Lanjouw and Lanjouw (2003).

Table 6A: How do Observables Correlate with Matches and Rents

	(1) Mean	(2) LC	(3) HC	(4) HL	(5) Rent
EA listed as informal (slum)	.9349	.0053	.0047	.08	-.2758**
		[.109]	[.022]	[.1005]	[.1332]
Age of household head	35.21	-.0027	.0022	.0002	-.0019
		[.0033]	[.0022]	[.003]	[.0059]
Household head works for pay	.6724	-.0214	-.0196	-.0852***	-.1293*
		[.0392]	[.0221]	[.0301]	[.0684]
Household head owns a business	.1548	.006	.0166	.0751**	.1415
		[.0481]	[.026]	[.0383]	[.0947]
Household head works for private sector	.4003	.0238	.0043	-.0539**	.0216
		[.0281]	[.0184]	[.0244]	[.0387]
Household head has no education	.0365	.0451	.0475	.0436	-.3481*
		[.1623]	[.0748]	[.1178]	[.1788]
Household head has some secondary educ	.4301	-.0893	-.0216	-.0113	.2102**
		[.0619]	[.0406]	[.0592]	[.1067]
Household head, years of education	9.237	-.0032	-.0093**	.0004	.0666***
		[.0078]	[.0038]	[.0063]	[.0118]
Household head, hours worked	55.91	-.0007	-.0002	.0004	.0021*
		[.0008]	[.0004]	[.0006]	[.0011]
TV	.4274	.0949*	-.0048	.0165	.5066**
		[.0544]	[.0319]	[.0452]	[.0741]
Radio	.7583	.1064	-.0425	.0794	.2145**
		[.0748]	[.0434]	[.0528]	[.0907]
Mobile phone	.8279	.0281	-.0644	.0164	.4206***
		[.0663]	[.0538]	[.0497]	[.0858]
Bicycle	.06	.1509*	.0046	.0846	.0479
		[.083]	[.0436]	[.0629]	[.1182]
Walls are made of mud/wood	.2815	.0087]	.0109	-.0101	-.0734**
		[.0215]	[.0144]	[.0169]	[.0292]
Floor is made of earth	.3324	.0001	-.0017	.0108	-.2143***
		[.0287]	[.0208]	[.0223]	[.0381]
Main water source is water vendor	.2424	-.0221	-.0016	-.0114	.0442*
		[.0148]	[.0093]	[.0137]	[.0228]
Main waste disposal method is uncovered pit	.2182	.0093	.0023	.0095	.009
		[.018]	[.0104]	[.0143]	[.0242]
Main cooking fuel is paraffin	.5161	-.0837**	-.0264	.0293	-.052
		[.0375]	[.0219]	[.0262]	[.0546]
Main light source is electricity	.5399	.0207	-.0131	-.0159	.2791***
		[.0324]	[.0193]	[.0252]	[.0433]
Fraction of youth unemployment in the household	.1638	.0295	.0415	-.0109	.0388
		[.0591]	[.0339]	[.0447]	[.0922]
Household owns housing	.0726	.0019	.0277	-.0495	-.0095
		[.0527]	[.0256]	[.0331]	[.0571]
Household has piped water into dwelling	.0462	-.0335	.0276	-.0131	-.0301
		[.0519]	[.02]	[.0266]	[.0647]
Wealth index	.0483	.0041	-.0024	.002	.0497***
		[.0048]	[.003]	[.004]	[.0063]
Poverty rate ($2/day)	.4315	-.0831	.0455	-.0858*	-.5063***
		[.055]	[.0354]	[.0456]	[.0895]

Note: Standard errors are clustered at the EA level (reported in brackets). * p<0.1, ** p<0.05, *** p<0.01.
Each cell is from a separate regression, controling for village fixed effects and tribe dummies.

Table 6B: How does Tribe Composition and Diversity Correlate with Matches

	(1) Majority Tribe, EA	(2) Majority Tribe, Zone	(3) Majority Tribe, Village	(4) Diversity, EA	(5) Diversity, Zone
Landlord-Chief (LC) tribe match	0.022	0.015	-0.018	0.014	0.011
	[0.101]	[0.099]	[0.112]	[0.012]	[0.008]
Household-Chief (HC) tribe match	-0.107	-0.005	-0.158	0.017	0.012**
	[0.120]	[0.116]	[0.134]	[0.014]	[0.005]
Landlord-Household (LH) tribe match	0.144**	0.111*	0.077	-0.027***	-0.007*
	[0.061]	[0.065]	[0.071]	[0.009]	[0.004]
Dependent Variable Mean	0.544	0.496	0.490	0.611	0.689
R-squared	.24	.3	.319	.28	.352
Observations	14814	14814	14814	14814	12390

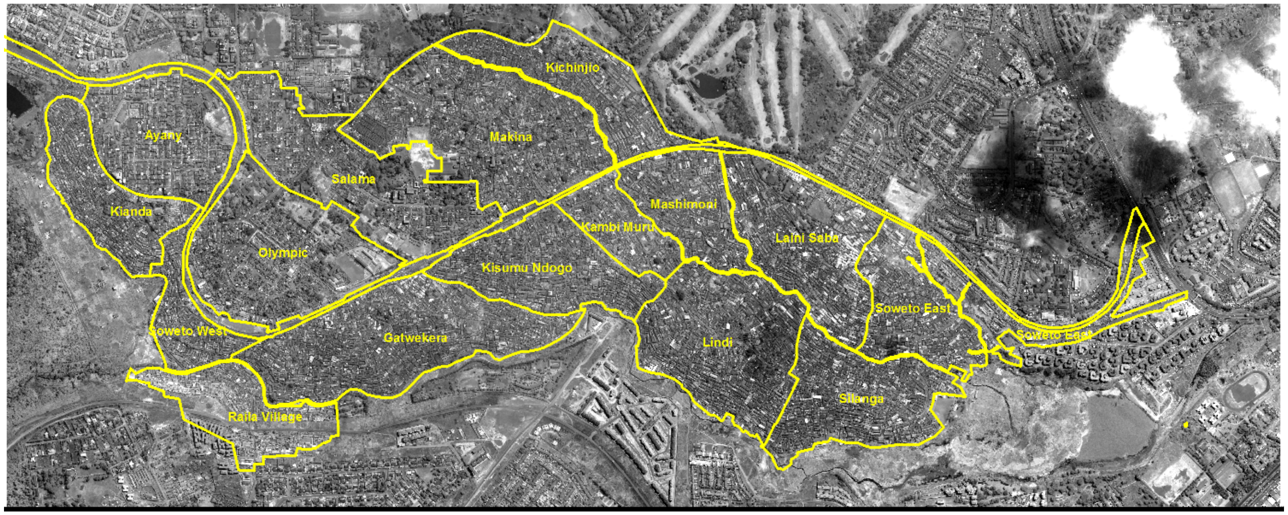
Note: Standard errors clustered at the location-household tribe-landlord tribe level (reported in brackets). * p<0.1, ** p<0.05, *** p<0.01.
Standard errors in column (5) clustered at the zone level.

Table 7: Rents for Households Close to Administrative Boundaries: RD Specification

	(1) Log Rent	(2) Log Rent	(3) Log Rent	(4) Log Rent	(5) Log Rent	(6) Log Rent	(7) Log Rent	(8) Luminosity Change Roof	(9) Luminosity Change Roof	(10) Luminosity Change HH	(11) Luminosity Change HH
HC tribe match	-0.352**	-0.216**	-0.183**	-0.121**	-0.102**	-0.359**	-0.158***	48.822**	11.087**	19.431	10.901**
	[0.150]	[0.103]	[0.076]	[0.048]	[0.039]	[0.138]	[0.048]	[19.968]	[4.683]	[16.718]	[5.285]
Bandwidth	1	2	3	4		1		1		1	
Polynomial					X		X		X		X
Dep Var Mean	7.116	7.137	7.156	7.178	7.218	7.116	7.218	7.116	7.218	7.114	7.218
No Weights	X	X	X	X	X			X	X	X	X
All Controls	X	X	X	X	X	X	X	X	X	X	X
R-squared	.3	.25	.33	.32	.44	.36	.45	.44	.06	.38	.05
Observations	307	710	1368	1956	9996	307	9996	312	10233	311	10247

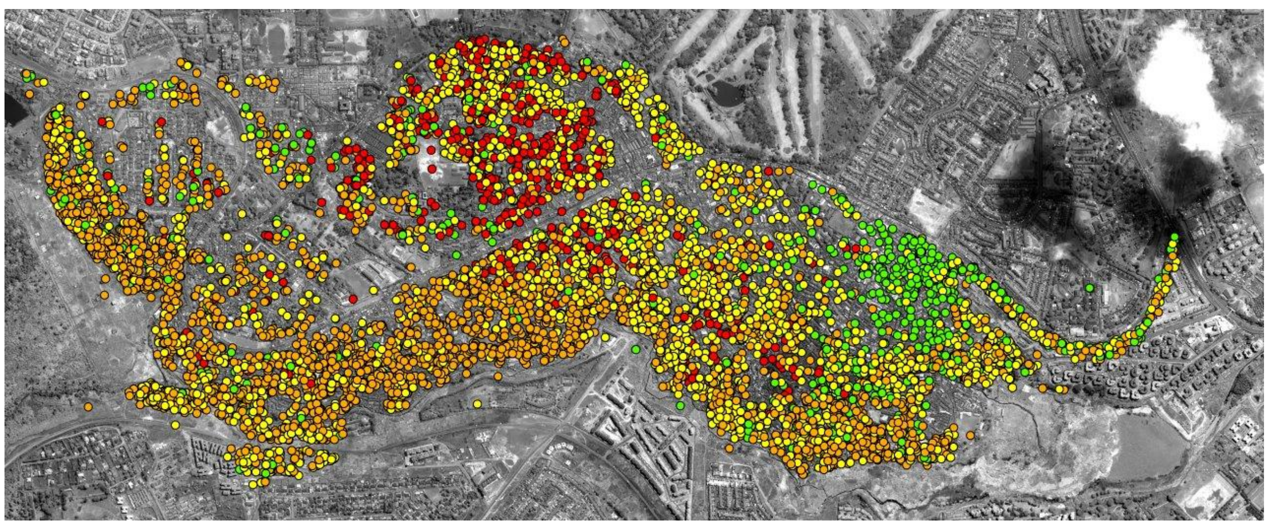
Note: Standard errors clustered at the location-household tribe-landlord tribe level (reported in brackets). * p<0.1, ** p<0.05, *** p<0.01.
Bandwidths 1, 2, 3 and 4 are for local linear regressions with bandwidths of 75m, 100m, 125m and 150m respectively.
The polynomial specification includes third order polynomials, estimated separately on either side of the boundary.
All controls refers to village fixed effects, household and EA controls.
The specifications for luminosity change are based on the two most recent pictures (i.e they measure the investment between late 2011 and mid 2012).

Figure 1: Map of Kibera



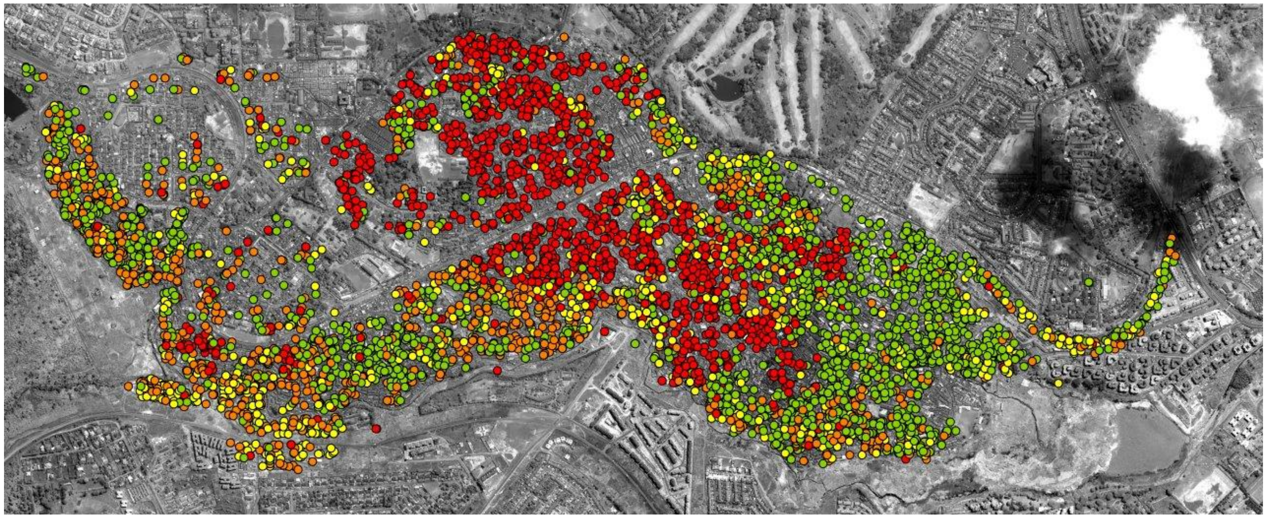
Note: This figure shows a map and satellite image of the Kibera area. The villages in Kibera are outlined and labelled in yellow.

Figure 2: Distribution of Household Tribes in Kibera



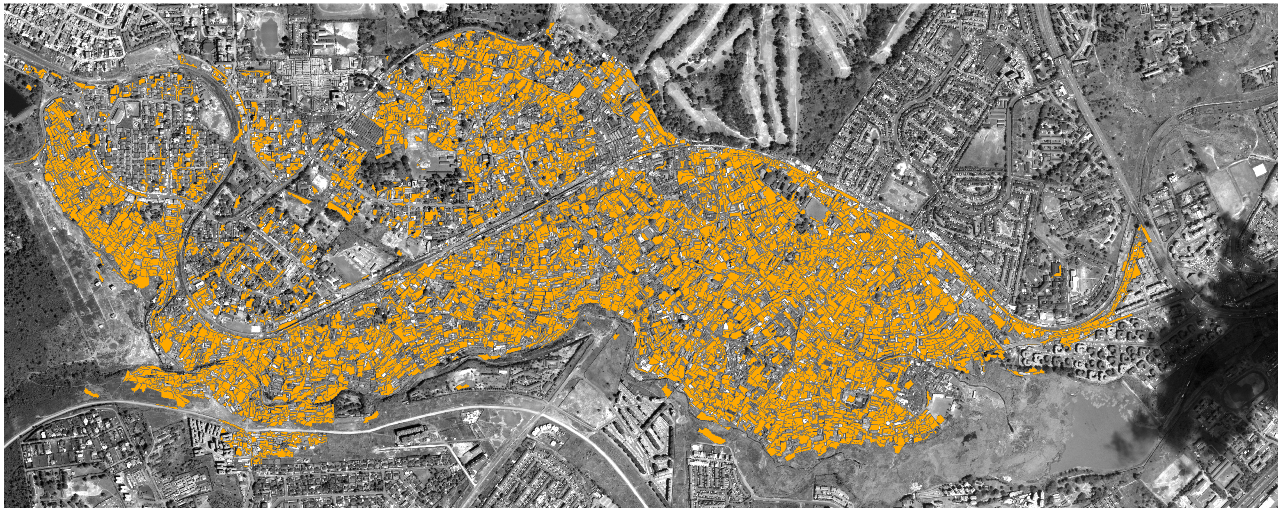
Note: For confidentiality reasons and under IRB requirements, we do not provide a key as to which color is which tribe.

Figure 3: Distribution of Landlord Tribes in Kibera



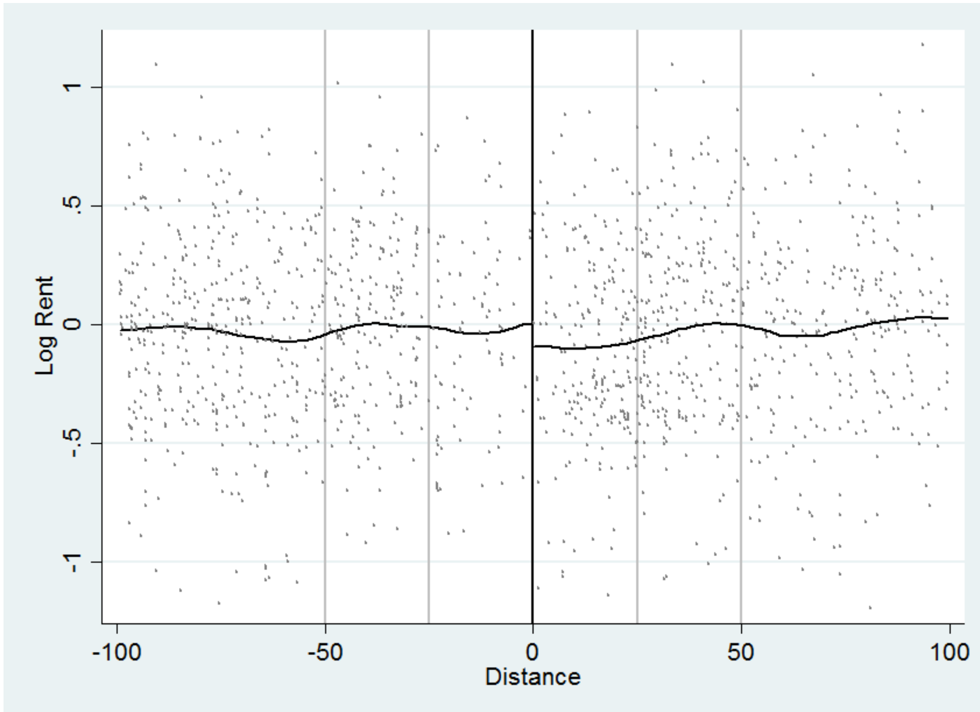
Note: For confidentiality reasons and under IRB requirements, we do not provide a key as to which color is which tribe.

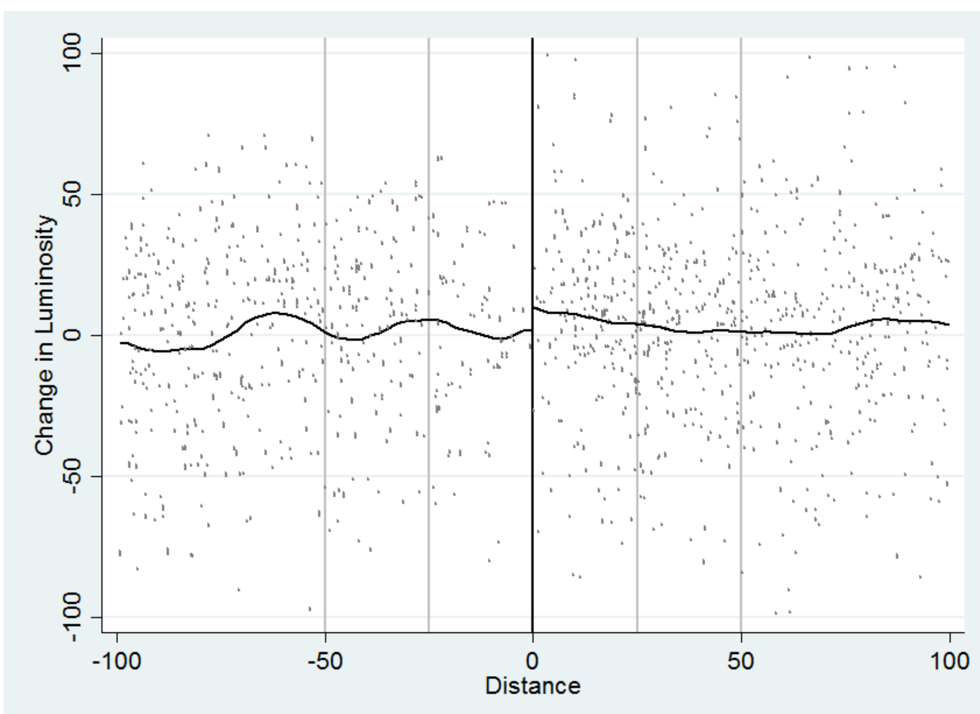
Figure 4: Roofs from Satellite Images in Kibera (January 2011)



Note: This figure shows a satellite image of the Kibera area. Roofs in our sample are indicated in orange. See Online Appendix for methods.

Figure 5: RD Graphs For Log Rent and Luminosity





Appendix Table A1: Attrition Checks

Dependent Variable is Log Rent

	(1) Low Attrition Sample	(2) Low Attrition Sample	(3) Attrition Weights	(4) Attrition Weights
Landlord-Chief (LC) tribe match	0.093***	0.128***	0.064**	0.080**
	[0.023]	[0.032]	[0.027]	[0.034]
Household-Chief (HC) tribe match	-0.073***	-0.071**	-0.075***	-0.073**
	[0.026]	[0.030]	[0.028]	[0.030]
Landlord-Household (LH) tribe match	-0.008	0.008	0.008	0.020
	[0.019]	[0.023]	[0.019]	[0.022]
Landlord-Elder (LE) tribe match		-0.055**		-0.023
		[0.025]		[0.022]
Household-Elder (HE) tribe match		0.002		-0.009
		[0.029]		[0.026]
Dependent Variable Mean	7.267	7.318	7.247	7.306
Village FE	X	X	X	X
HH Controls	X	X	X	X
EA Controls	X	X	X	X
R-squared	.433	.337	.417	.321
Observations	11233	9371	13990	11631

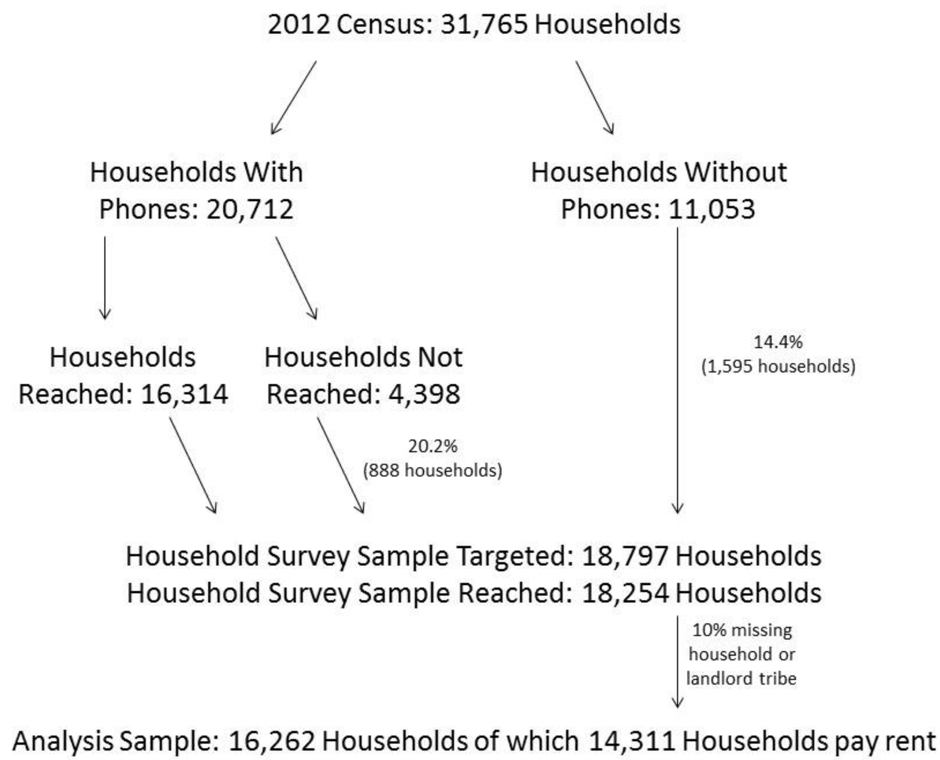
Note: Standard errors clustered at the location-household tribe-landlord tribe level (reported in brackets). * $p<0.1$, ** $p<0.05$, *** $p<0.01$.
Low attrition sample covers the EAs with a 5% attrition rate or less.
The attrition weights are computed using the methods in Fitzgerald, Gottschalk and Moffitt (1998).

Appendix Table A2: Correlations of Luminosity with Wealth, Consumption and Poverty

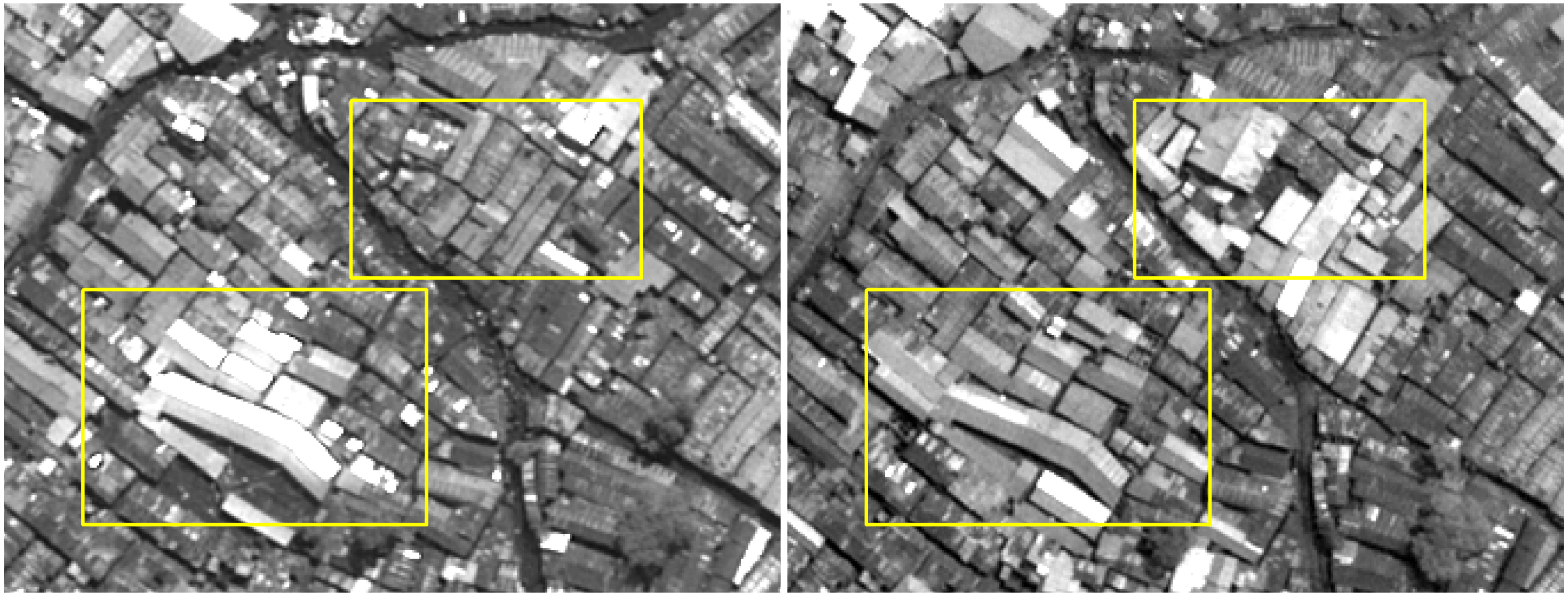
Relationship with Demographics		Relationship with Wealth		Relationship with Other Indicators	
Fraction of HH members working	14.782*	Walls are made of mud or cement	-5.955**	Poverty ($1.25 per day)	-51.992***
	[7.840]		[2.507]		[18.497]
Fraction of HH members in business	38.907***	Floor is made of cement	16.963***	Poverty ($2 per day)	-39.862***
	[10.341]		[3.581]		[11.938]
HH head has some secondary educ	17.584**	Main cooking fueld is charcoal	-9.690**	Log consumption per capita	24.503***
	[8.218]		[4.061]		[6.554]
HH head has some univ educ	21.124	Main water source is vendor	5.863***	% of EA within 15m of road	8.360***
	[13.993]		[1.943]		[2.861]
Age of HH head	-1.540***	Main light source is electricity	-3.007	Presence of business in EA	2.937**
	[0.403]		[3.759]		[1.460]
Observations	576		576		576

Note: Heteroskedasticity robust standard errors in brackets. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.
HH stands for household; EA stands for Enumeration Area.
All regressions are weighted by the size of the EA and include eight sublocation fixed effects.
The first two columns are results from one regression of luminosity on the reported variables in the column.
In the third column, each cell reports results from a separate regression as all these variables are extremely highly correlated.

Appendix Figure A1: Household Sampling Strategy

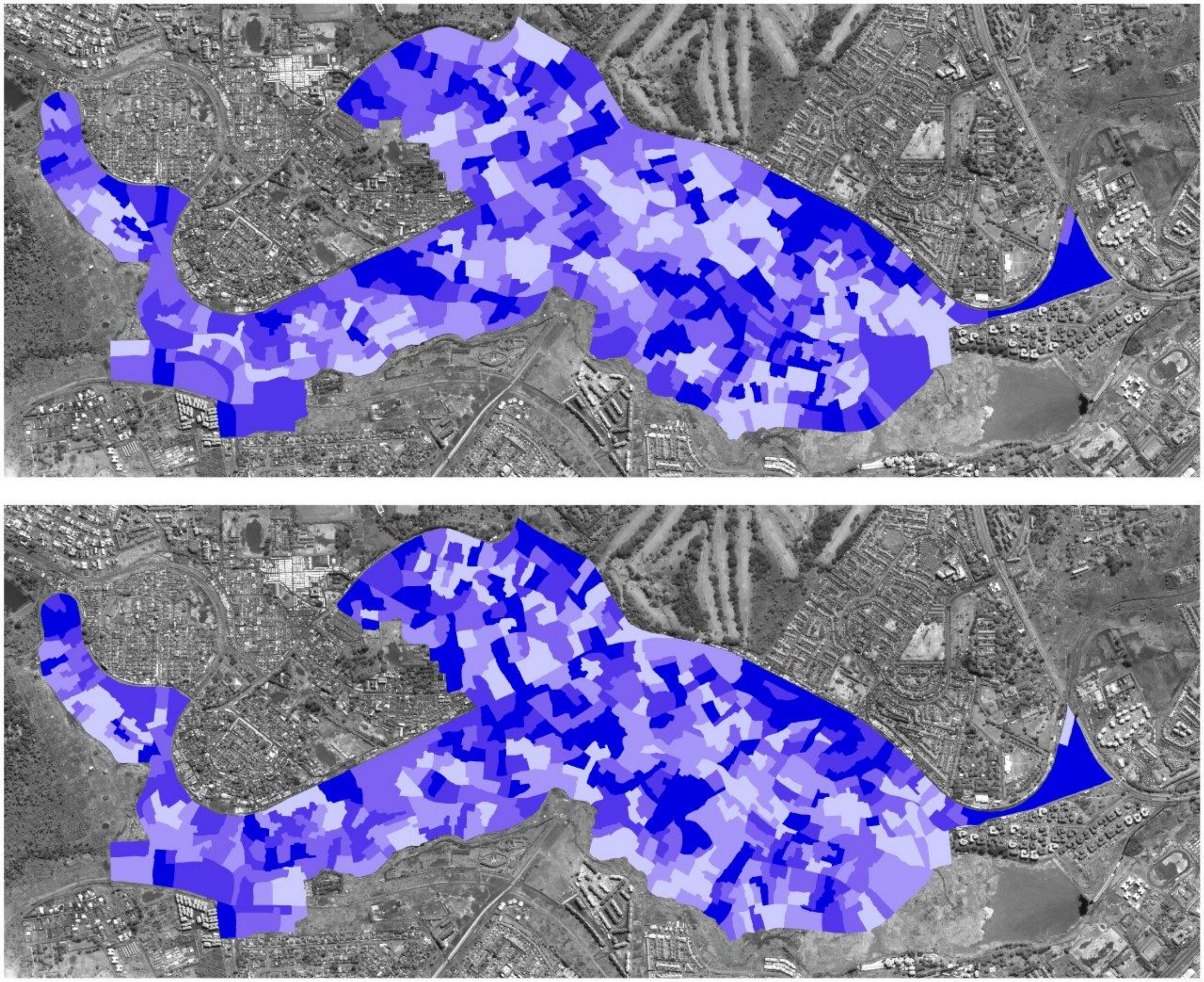


Appendix Figure A2: Old and New Roofs in Kibera



Note: Both pictures are taken over the same area of the slum with the same resolution (0.5 meters panchromatic).
The picture in the left panel was taken in July 2009 and that in the right panel in August 2012.
The yellow rectangles highlight clusters of roofs that markedly evolved over the period.
Roofs highlighted in the bottom rectangle degraded while roofs within the top rectangle were upgraded in the same timeframe.
The picture area is approximately 175 meters long and 140 meters wide.

Appendix Figure A3: Roof Luminosity and Household Consumption Across Kibera



Note: The top panel displays mean luminosity by EA and the bottom mean (estimated) household consumption by EA.
The data used are residualized values after accounting for the sublocation fixed effects in a regression framework.
The blue color scale is divided into the 5 quintiles in the data. The picture only includes the 546 informal EAs in the 2009 census.