Report No: AUS7808

World
Measuring Living standards within cities

May 1, 2016

GSUGL

MLSC

Measuring Living Standards within Cities[1]

Department for International Development

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

WORLD BANK GROUP

[1] Please cite this report as World Bank (2016). "Measuring Living Standards within Cities". Washington, DC.

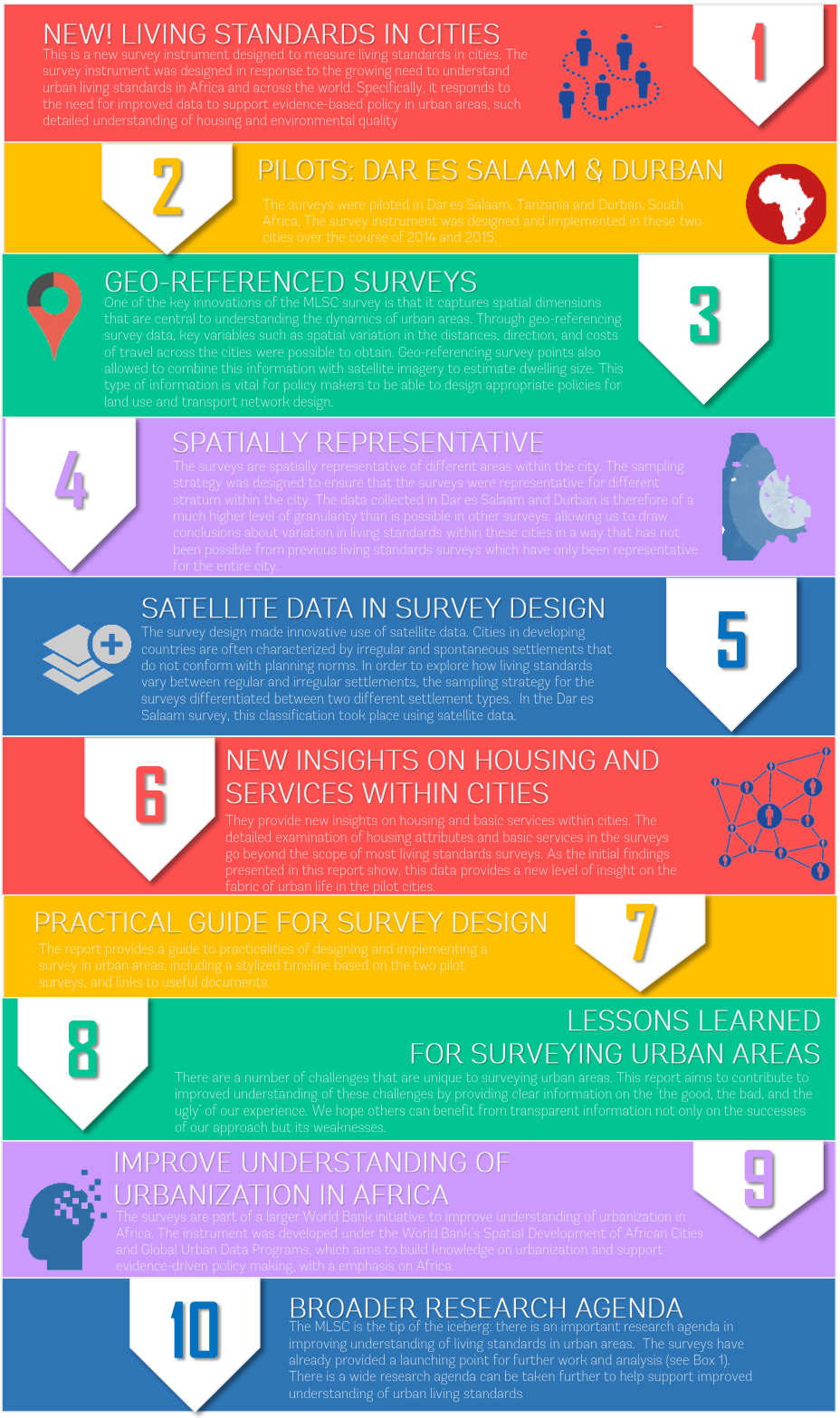
Table of Contents

Table of Figures

Acknowledgements

The Measuring Living Standards in Cities (MLSC) survey was developed under the World Bank's Spatial Development of Cities Program and under the overall guidance of Somik Lall (Lead Economist and TTL of the program, GPSURR)) and Sameh Wahba (GPSURR, Practice Manager). The survey instrument was designed and implemented by a team of World Bank staff and consultants led by Nancy Lozano Gracia (Senior Economist and TTL, GPSURR), and consisting of: Alexei Abrahams, Juliana Aguilar Restrepo, Ana Aguilera de Llano (Urban Specialist, GPSURR), Sarah Antos, Kshitij Batra, Juan Muñoz, and Alexandra Panman. The team is grateful for extensive technical guidance and support from Kathleen Beegle and Talip Kilic, as well as inputs from members of the project advisory committee, peer reviewers, and other colleagues: Joao Pedro Acevedo, Andre Bald, Judy Baker, Andrew Dabalen, Uwe Deichmann, Marianne Fay, Isis Gaddis, Jonathan Kastelic, Kate Owens, Peter F. Lanjouw, Megha Mukim, Harris Selod, and Tara Vishwanath. The two surveys were fielded by teams of Oxford Policy Management (OPM) under the leadership of Patrick Ward, and including Shafique Arif, Stephanie Brockerhoff, Ayanda Hlatshwayo, Maja Jackobsen, Ignatus Jacob, Fidelis Hove, Erick Kilala , Andreas Kutka, Andrej Kveder, Marta Moratti, and Jeje William. This work was funded by UK DFID and Switzerland SECO through the MDTF for Sustainable Urban Development TF071544.

The Quick Read: Ten Things to Know About the MLSC Survey Instrument and Report



Introduction: the Good, the Bad, and the Ugly of Survey Design

Across the world, urban development has been linked with economic growth and rising living standards. But throughout Africa, urbanization has been puzzling policymakers and researchers alike. Cities are growing in size and numbers, but low quality living conditions seem to persist, worker productivity does not appear to be rising, and economic density remains low in most urban areas (UN Habitat 2014; Collier 2013).

Despite a recent wealth of studies that aim at understanding the challenges faced by African cities, to date, very little is known about urban households in Africa. There is very little information on the trade-offs that households make in deciding where to live and where to work in African cities, on the factors that influence their housing investment decisions, and even on how living standards vary within cities. This information is needed to support evidence-based policy design: policy makers require information on housing quality, amenities, travel times/costs, and environmental qualities in their cities to design appropriate policies for land use, transport network design, and other public service improvements. The reason for this information gap? Censuses are infrequently updated and most national level household surveys – such as the Living Standard Measurement Survey – cannot measure these dynamics accurately, because they are rarely representative at the city level, let alone for areas within cities.

The Measuring Living Standards in Cities (MLSC) survey is a new instrument designed to enhance understanding of cities in Africa and support evidence based policy design. The instrument was developed under the World Bank's Spatial Development of African Cities Program, and was piloted in Dar es Salaam (Tanzania) and Durban (South Africa) over the course of 2014/15. These geo-referenced surveys provide information on urban living standards at an unprecedented level of granularity: they can be compared across different geographic levels within the cities, and between areas of 'regular' and 'irregular' settlement patterns.[2] They also respond to the need to increased understanding of specifically 'urban' dimensions of quality of living: housing attributes, access to basic services, and commuting patterns, among others.

This report is divided into three main chapters. In this first chapter, we explore 'the good, the bad, and the ugly' of survey design for urban areas. The objective is to contribute to best practice and knowledge sharing on surveying urban areas through a frank discussion of successes and failures of the pilots, drawing particular attention to challenges associated with working in urban areas. The second chapter provides detailed information on the survey methodology. This section will act as a guide on the sampling strategy and weighting of the survey that is likely to be useful for analysts that work with data. The final chapter draws on descriptive statistics from the two pilot surveys to provide a first look at the kinds of insights that can be supported by this new survey instrument. Although this is only a preliminary analysis, the descriptive statistics provided in this chapter already demonstrate the advantages of looking at cities through a more focused lens. The chapter sheds new light on the challenges of access to basic services, informality, urban mobility, and housing in Dar es Salaam and Durban.

[2] The sampling strategy used satellite data to differentiate between regular and irregular settlement patterns, called 'shanty' and 'non-shanty' areas in Dar es Salaam and 'slum' and 'non-slum' areas in Durban. Please see Chapter 2 for further information.

The remainder of this introduction provides an overview of key innovations, challenges, and lessons learned from the experience of piloting the instrument in two different cities. The aim is to help build and share knowledge on the design and implementation of surveys based on in-field experience. It is divided into three parts. The first part focuses on innovations and lessons learned for surveying urban areas, which is likely to be particularly useful for researchers working in the urban context. The second explores some additional lessons learned on general issues related to questionnaire design and implementation, and is likely to be of specific interest to researchers developing new instruments that nonetheless aim to retain a high degree of comparability with existing living standards measurement surveys. The concluding section is a one-page summary of our 'wishful thinking': reflections that combine our thoughts on improvements from the current surveys that we hope that teams implementing new surveys can learn from our experience and aim even higher. It can also be interpreted as a tentative research agenda to develop knowledge on the specific challenges to collecting reliable data on living standards in urban areas – an issue that is likely to become increasingly important as more and more of the world's population continues to urbanize.

Key Innovations and Lessons Learned for Surveying Urban Areas

Bringing in the spatial dimension: innovations in questionnaire design, sampling, and listing. One of the key innovations of the MLSC survey is that it captures spatial dimensions that are central to understanding the dynamics of urban areas. Spatial considerations were incorporated into questionnaire design, by including questions that help identify the location of key centres of activity, such as jobs or schools. Given that postal addresses are not universal or widely known in African cities, respondents were given the option to identify areas by landmarks as well. Initial analysis of the data suggests that this was a very successful way to gain a first approximation of the location of key activities in urban areas.

The sampling strategy for the surveys was designed with spatial considerations in mind. A separate sampling strategy was developed for each city in light of the specific conditions of the city. In both cases, however, the first stage of the sampling strategy involved selecting and sorting 200 Enumeration Areas (EAs) into several strata, depending on their distance to the city center reflecting the center of the city, the consolidated city, and peripheral areas; the likelihood of belonging to irregular settlements was also considered to define additional strata. The major advantage of this approach is that it ensures the data is representative for different areas of the city, allowing for important insights to be drawn regarding differences in living conditions across the city. The MLSC survey therefore allows for much granularity in the analysis of living conditions within the surveyed cities than nationwide living standards surveys, which are, at best, only representative for the entire area of the largest city in the country. For a more in-depth discussion of the sampling strategy, please see Chapter 2.

Four spatial innovations were introduced in listing: recording building height (Dar es Salam), measuring the size of buildings, the use of electronic bar codes as markers, and the use of satellite-imagery supported CAPI platforms to ensure the completeness of the listing operation. The first of these innovations was highly successful: building height is an important dimension of urban development that is difficult to gauge from satellite imagery. The addition of a simple instruction to record the number of floors in buildings at the listing stage proved to be a cost effective means to gather information that provides insights for the analysis of the fabric of a city. Beyond this, such information can be used to validate semi-automatic algorithms that estimate 3-D models of cities based on satellite imagery. The team coordinated with partners and shared

such information to improve the quality of such estimations in African cities. Attempts to measure length of all sides of the properties were met with resistance from residents, making it impossible on practice to gather reliable information in this characteristics. Households did not allow enumerators to collect such measures in many cases and the field had to be dropped from the listing exercise. Given that both surveys were georeferenced, however, a spatial variable could nonetheless be calculated by combining the survey data with other sources of information after implementation. Specifically, in Dar es Salaam satellite imagery was used to estimate the total roof area in square meters. In the case of multiple occupancy buildings, the roof area estimation was refined with information on the number of households living in the building, collected at listing. In the case of Durban, a similar task was performed by combining information from municipal land use maps with data on multi-family occupancy.

In terms of barcode stickers, these were replaced with chalk marking after an initial period of testing. The use of stickers proved to be time consuming and unreliable: they were prone to scanning/technological failure and did not stick equally well to all surfaces (see **Table 1** for additional information). Lastly, the household listing operation was conducted using tablets. The CAPI application used Google My Maps –a platform supported by Google Earth– that aimed at ensuring the completeness of the listing operation. Exclusion of households during listing may bias the representativeness of the survey, typically against the poorest and vulnerable populations, ethnic and religious minorities, and households located in remote areas. This innovation was led by OPM, and while field issues did not allow for full implementation, further consideration on this idea may substantially improve the quality of survey data.

Putting a spotlight on housing: residential mobility, house prices, and other market dynamics. The surveys introduced initiatives to capture insights on dynamics of housing in urban areas that are not often addressed in household surveys. The first is that the surveys introduced a 'residential history' module. This module aims at addressing the gap in understanding of households' residential moves within cities, which is not captured in conventional surveys and is partly absent in panel surveys where residential mobility is a common cause of attrition.

There are a number of lessons learned from the piloting of the residential history module. Although interviewers reported that households often enjoyed telling their residential history, the length and complexity of the module made it a challenge for implementation, with significant time dedicated to training enumerators in this part of the questionnaire. Households sometimes provided information about a past house out of chronological order, which meant that additional care had to be taken when cleaning the data in recording their move-out date. Future iterations of this module could also benefit from explicitly recognizing that households may move in and out of the city on more than one occasion to ensure these data is consistently collected. Lessons learned include the importance of training, checks, and flexibility of roster entry to cross-check at the time of data collection that the ordering of the residential histories is accurate.

The survey also collected information on house value and rents that are often omitted in surveys. Collecting information on housing values has been problematic in the past.[3] Rental rates are

[3] There is a long tradition of research that uses household survey data to conduct housing market analysis (Malpezzi 2000). Given the absence of reliable housing market transaction data in many developing countries, self-reported house prices are needed for this analysis. While collection of rental prices from renters is relatively straight forward (although consideration must be taken into whether the prices reported include features such as utility bills, subsidies, and the period of collection) it is also important to gather equivalent data for owner-occupied housing. These results can be cross-checked by imputing rents for owner-occupied housing based on a hedonic regression that estimates rental values as a function of housing characteristics from the rental housing in

easier to ascertain, since this is a regular payment made by households that rent. Some LSMS surveys attempt to record a similar value for owner-occupied housing, by asking how much households would have to pay "if, instead of owning your home, you had to rent it" (Haughton and Khandker 2009 p 27). However, this phrasing requires some knowledge of rental markets, and is a question that can arouse suspicion.[4] After extensive consultation with colleagues experienced in housing surveys, and with further refinement through survey piloting, the wording for the MLSC was modified to: "If a friend of yours wanted to buy a property like this in the same neighborhood, how much would he/she have to pay?". Drawing on lessons from behavioral economics and psychology, respondents tend to provide a more objective valuation when confronted to this wording. People often ascribe more value to things merely because they own them –a bias that is commonly referred in the literature as the endowment effect. The wording used in these surveys aims at minimizing the size if this effect when collecting self-reported information on housing values.

The identification of subletters in the housing market dynamics is one further innovation that should be noted. Subletters are often 'lost' in household surveys and census: neither recorded as members of the household nor specifically identified as a separate household. Enumerators were therefore instructed to identify all households living within dwellings at the listing stage, so that they would be included in the sample of households from which respondents were selected. In addition to this, the potential significance of this activity for household incomes and house market dynamics was probed with a small section in the questionnaire exploring how many households rented out rooms within their dwelling, the purpose they were rented out for, and the income that they earn from this activity.

Working in the urban field: security, sensitisation, and low response rates. There are a number of specific implementation challenges for surveys in urban areas. Crime is a major concern: both in terms of the security of field teams, and because households in areas with high crime rates may display greater suspicion and caution in engaging with enumerators. In addition to this, urban households often spend a large portion of time outside of the household, adding to the challenge of successfully interviewing specific members.[5] As such, all urban surveys are marked by low response rates. Methods to help create trust and ensure the safety of fieldworkers in both Dar es Salaam and Durban included community sensitisation and dialogue with local leaders. In Durban, where access was a particular challenge, additional efforts had to be made, including: adjusting of the listing strategy to lower the burden of access[6] and engaging with building trustees and organisers.

---

the sample. Another approach is to confirm rental values in community questionnaires, or with informal and formal real estate agents (e.g. Dowall 1995).

[4] Consultation with colleagues that have worked extensively with household surveys across the world revealed that respondents commonly perceived questions such as this as an expression of the enumerators interest in renting the house.

[5] For further details, see following section on Survey Methodology.

[6] In Durban enumerators had major problems in gaining access to building to conduct listing, particulalry in areas where there were gated communities and suspicion of outsiders was very high. The total nominal sample of 2400 households in Durban was, selected in four stages rather than two. These were: (i) selection of 200 EAs with probability proportional to size; (ii) large EAs were segmented into area units of roughly the same size (using GIS data), and one segments was selected randomly with equal probability; (iii) following listing of buildings, 15 were selected using systematic equal probability sampling; (iv) households in the 15 selected buildings were listed so that 12 households could then be selected per EA by systematic equal probability sampling. This approach reduced the need to enter as many buildings as would otherwise have been necessary, without reducing the representativeness of the sample.

Lessons learned on Questionnaire Design and Implementation

Balancing comparability and pragmatism: trade-offs in the collection of consumption data. Accurate consumption data is central to the measurement of living standards. As Beegle and others (2010) have shown, however, consumption data is very sensitive to small changes in the collection module. The consumption module of the MLSC survey was designed to match the best practice module used in the Tanzania LSMS surveys, in order to minimize uncertainty about the degree of variation in measurement of living standards that may emerge as a result of methodological difference.

A number of changes were made to the consumption module in the MLSC in response to respondent's fatigue and city-specific adaptations. During piloting stages, it was found that the survey length was such that both interviewers and interviewees were becoming fatigued – a factor that can be detrimental to the quality of the data collected. To address these concerns, some questions relating to expenditures on Food Outside of the Household, Education, and Health were condensed from the individual to the household level. This adjustment did not compromise the quality of the data,[7] but it did reduce the level of detail of information collected, a fundamental trade-off that all surveys face.

Small changes were also introduced to the collection of food items consumed within the house. Although these modifications reduced the length and complexity of the module in the field, they resulted in considerable additional challenges in the data analysis and some were not successful. The first change was that the 'total value' of purchased food was not collected, which led to marginal timesaving at the expense of an important verification tool for quantities and unit-price observations. Secondly, based on feedback from interviewer training, non-standard units were introduced for the collection of specific items. However, these changes were not fully documented: no systematic record was made as to how non-standard units were defined or explained to interviewers, resulting in considerable additional burden on the data analysis phases of the project to ensure the validity of the data collected.[8] Drawing from the lessons learned in Dar es Salaam, some of these issues were identified and addressed during survey preparation in Durban.

Pre-testing and interviewer training are vital. As with most large-scale surveys, the MLSC underwent important refinement and modification through feedback from pre-testing and interviewer training. As the LSMS Handbook stresses, field testing is an essential part of the work of survey design, but important knowledge sharing gains could be achieved if the lessons

[7] Based on initial analysis on the distribution of the observations. These are currently being applied to more rigorous tests by comparing the observations with predicted values from survey-to-survey imputations from the TZNPS.

[8] Non-standard units were defined during pilot and interviewer training phases. These definitions were not, however, properly recorded, adding additional uncertainty in the conversion values to kilograms during the data analysis phase. As such, the analytical work had to rely on kilogram conversions collected by the LSMS team in neighboring countries. A second source of uncertainty arose as a result of poorly documented interviewer training on food items. Modules that allow for collection of food observations in non-standard units typically allow interviewers to choose and record which unit they are using for all items. Post-field work analysis of the data revealed trends indicative of the fact that interviewers observations were not fully consistent in their use of units. As such, the trends in the recording of data were carefully analysed using statistical methods that helped to identify and correct variation in observations that was caused by problems with use of units across interviewers and the months in which the data was collected. A further check of the data is being conducted to validate the results through survey-to-survey value predictions using the TZNPS.

learnt from field testing were better documented and disseminated. To this end, **Table 1** outlines key changes that were made and the rationale behind them.

The two MLSC pilot surveys are among the first large-scale surveys that were implemented with the World Bank's Survey Solutions. There was thus a steep learning curve for the implementing firm. The relatively new nature of the software also provided some additional challenges, including: (i) constraints to CAPI development as certain options were not available at the time of the survey;[9] and (ii) frequent updates to the software while fieldwork was in progress. These challenges are to be expected at this stage of the product development, and mitigated by the close support provided by the World Bank's Survey Solutions team throughout. The close engagement of the OPM data management teams in solving these challenges was critical for the successful implementation of these surveys in Dar es Salaam and Durban.

Table 1: lessons learnt from pre-testing and interviewer training

Modification	Rationale
Removal of questions to shorten the survey	Interviewer and respondent fatigue. Main changes made occurred in the consumption module (see paragraphs above) as well as on social networks (see section on 'wishful thinking' below).
Phrasing of questions	Lessons from behavioural sciences and methodological literature on survey design, as well as the adaptation of questions to city-specific contexts were made to capture local realities more accurately.
Reduction in the number of households per Enumeration Area	Sample size had to be slightly sacrificed given the budget envelope, since the length of the questionnaire placed unanticipated cost burdens to the firm.
Use of electronic barcode tags for listing	An innovation to use electronic barcode tags in household listing was unsuccessfully tested in Dar - tags would not stick to the materials of doors and walls, so the traditional method of chalk marking was used instead.
Measure GPS coordinates of plot sizes	Efforts to measure the size of plots was met with extreme suspicion by households, and was therefore limited to recording the number of floors of the dwelling at listing.
Job categories	Improved training for enumerator to recording jobs by category would have resulted in time savings at data cleaning phase.

Fieldwork monitoring, data validation, and data management. The single most important lesson learned in terms of data management is that the importance of documentation cannot be overstated. Micro data must be clearly and consistently labelled. Lack of clarity can cause headaches in the long run; for example, lack of clear documentation at each stage of the selection process complicates the calculation of household weights (see Annex 1 for further details). It is also vital that there is complete transparency on the validation checks undertaken during data collection, and any changes that are made to data or interviewer instructions in response to

[9] The MLSC survey was one of the first large scale surveys implemented using Survey Solutions. Some features of the survey – such as selection of random respondents – could not been managed through survey solutions at that time. Internal cross-checks were also limited.

feedback during the survey implementation are fully recorded and communicated. These expectations should be clearly stated in the Terms of Reference of the implementing firm.

Weekly monitoring of data quality indicators during listing and survey implementation were critical to ensure the quality of the fieldwork. Both surveys relied on a set of quality control indicators that were monitored on a weekly basis. Typical indicators included summary statistics such as the average household size, average number of children per household, average number of people attending school per household, number of jobs recorded, average number of food items purchased or consumed per household, among others. These indicators were summarized by the week of data collection and enumerator, in order to address data quality concerns in real time. For instance, flags were raised when an enumerator systematically reported lower average number of food items consumed in a certain week. While sometimes this trend responded to variations across different EAs, in other cases increased supervision was needed to avoid shortcomings in survey implementation. While quality control indicators were produced weekly by HQ, specific responses relied mostly on the implementing firm, which performed field verifications to corroborate these perceptions and report to HQ.

Wishful Thinking

In preparing this report we have looked back over the experience of the two pilot surveys to highlight lessons learned and insights that can help refine the instrument. As we look ahead, we also see opportunities to go further: to take additional steps to improve our understanding of living standards in cities. We have called this our 'wishful thinking', because they are steps that we would have taken in the pilot surveys, had we not been constrained by budgets and timelines – something that we acknowledge is a reality of all survey work.

There are 'urban' dimensions to the collection of consumption data and identification of the urban poor that merit additional investigation. As highlighted above, small differences in the design of consumption modules can lead to significantly different poverty headcount and inequality rates (Beegle and others, 2010). As more people live in cities, there is need for additional experimental survey research to explore whether module design should factor in specifically urban conditions. This may include investigating: the impact of expanding the list of food items recorded, the level of detailed information collected on food away from home (FAFH) questions (for further information on recent advances in research on FAFH at the World Bank, see Farfan et al. 2015). This is because urban residents tend to consume a wider range of food items, and as such it is likely that optimal number and type of items listed should be different (see, for example, Smith and others 2014). Urban residents are also likely to eat more food outside of the home – but most modules for collection of food outside the home are very limited in the level of detail and accuracy for measuring basic needs food-consumption. This could have important implications for the accuracy of measures of consumption in urban areas.[10]

There is a need for more rigorous understanding of the best questionnaire design to ascertain house prices and secondary market activity. As noted above, the survey introduced new phrasing to elicit information from owner-occupied housing on the value of their dwelling. By using hedonic pricing methods, we will be able to use this data to explore the value of specific

[10] There are a range of FAFH modules used in LSMS and other household surveys. The level of detail in the module can have dramatic implications for poverty estimates, as demonstrated by Farfan and others (2015) in the case of Peru. In urban areas, FAFH is often a large portion of consumption – and, indeed, for many single person households, often the main source of food. Further research is needed to understand the implication of this on measures of consumption.

housing attributes in the cities, and explore the accuracy of self-reported prices.[11] Ideally, however, a field experiment would be conducted: there would be significant value added for future housing survey work if the impact of the phrasing of questions on reported prices could be better understood.

The survey went to new lengths to reveal the extent of 'secondary' housing markets: letting of rooms within owner occupied housing, as well as unofficial letting of entire buildings. Very little, however, is known about how people find rental housing, and the social and information networks that drive housing dynamics. Additional efforts could be dedicated to better understanding these dynamics better for example by asking about how households found their current house, and how they would go about looking for a new house. A detailed community questionnaire could also have contributed to capturing 'neighbourhood' level characteristics – trust, perceptions of safety, and reliance on neighbours for services – that are interrelated with of residential choices.

As we look ahead to the future, we believe additional research into the challenges outlined above could provide high value added insights to the analysis of urban areas. The MLSC survey has already provided the foundation for additional research (see Box 1), and it is our hope that the instrument can provide the basis for continued efforts to improve our understanding of living standards in cities.

Box 1: The Sensors are Here! A High-Resolution Application on Understanding Individual Travel Patterns in African Cities

Land use and transport policies can have a large impact on welfare and economic growth in cities. In uncoordinated and poorly integrated cities, households have to sacrifice access to jobs or essential services for the chance to get decent housing. When land use planning and transport policies are coordinated with better understanding of travel patterns, households and firms can make better location decisions. Reliable information about how individuals move around the city and the key constraints they face is vital: without this, policymakers are forced to take uninformed decisions that can prove to be costly over the long run.

The Sensors project builds on the MLSC survey in Dar es Salaam by combining face-to-face interviews with follow-up phone interviews and Big Data instruments to deliver unique insights on individual travel patterns and the economic costs of congestion. Households were randomly selected from the MLSC survey, and provided with a sensor-embedded GPS-enabled smartphone that collected and transmitted time and GPS locations of the individual every 1 minute in a daily 16 hour window, for the period of one month. The participants provided additional information on their movements through an initial face-to-face interview, follow up phone calls of a maximum of 15 minutes every three days to discuss the origin, destination, route, purpose and of trips undertaken in the intervening time.

The survey methodology was designed to collect information on urban mobility that did not suffer from the same degree of recall error associated with the traditional travel diary, while also avoiding the selection biases associated with previous Big Data applications in the transportation sector (as noted in World Bank, 2014).

[11] The survey collects information on services, tenure, distance to city center, and building attributes.

MLSC Survey Methodology

This chapter provides details of key components of the survey methodology. It is divided into four sections: sampling strategy, listing methodology, fieldwork, and weights. The surveys were implemented over the course of 2014 and 2015. Table 4 provides a stylized timeline for the full survey preparation, implementation, and analysis that is based on the experience of both surveys. The table provides a guide on key activities, common sources of 'unanticipated' delays, and links to documents and templates that will be useful to future survey design.

Sampling Strategy

The following section provides detailed information about the sampling strategy of the two surveys. As outlined below, the sampling for each survey was conducted in two main stages. There are some important differences between the methodology applied in Dar es Salaam and Durban, as the approach was adapted to reflect different conditions on the ground in these cities. Differences in the sampling strategy of these surveys are discussed throughout the following section. A condensed overview of the differences is visible in Table 2.

Table 2: Sampling Strategy – Dar es Salaam and Durban Compared

	Dar es Salaam	Durban
Sample frame	16,000 EAs generated by the Tanzania National Bureau of Statistics (NBS) for the 2012 Census.[12]	4,789 EAs (or Small Areas-SSA) in the Ethekhwini municipality published by Statistics South Africa in the 2011 National Population Census.[13]
Stage one	200 EAs sorted into four strata. The central strata was divided into 'central core, shanty' and 'central core, non-shanty'. Two EAs were replaced with reserve EAs as the original EAs were found to be inaccessible.	200 EAs sorted into five strata. The periphery stratum was divided into 'slum' and 'non-slum'. Large EAs selected in the first phase were segmented in two.
Stage two	12 households randomly selected by systematic equal-probability from updated listing of each EA	15 buildings were randomly selected by systematic equal probability in each EA, and 12 households were selected at random from within those buildings.

In the first stage, 200 Census Enumeration Areas (EAs) were selected. The EAs in the sample frame were sorted into strata depending on their distance to the city center and the likelihood they were slums/shanty areas. The methodology used to identify whether strata were slums/shantytowns was adapted to the information available in the city. In Durban, EAs in the periphery were identified as being either 'slum' or 'no slum' through land use maps and satellite

[12] The database included information on: Region, District, Ward, Sub-ward (or village), and numeric code within the sub-ward. The number of households and the population of each EA, as well as a digital description of the EA boundaries were also indicated.

[13] Each EA is identified by a geo-code that follows its administrative affiliation – Province, Metropolitan municipalities, District Municipalitis, Local Municipalities, and Wards. The database reports the total population in each EA, its racial distribution, number of dwellings and households, as well as its urban/rural stratum. The EA cartographic boundaries are described digitally in a linked Geographic Information System shape file.

imagery as reported in the 2006-2011 Informal Settlements Change Assessment provided by the South African National Space Agency (SANSA). In Dar es Salaam, EAs were divided into 'shanty' and 'non-shanty' based on satellite imagery: image pixels were first individually classified as "shanty" or "non shanty", and then the EA as "shanty" or "non shanty" depending on the predominant category of its pixels. This approach allowed for settlements to be distinguished as either 'shanty' or 'non-shanty' in the center of the Dar es Salaam where most EAs one of the two pixel categories predominated. For a more detailed explanation, please see Annex 1.

In both surveys, the distribution of the sample of 200 EAs was determined in accordance to Markward's rule (also known as the "50/50 equal/proportional allocation"). This approach was selected because it was considered the most appropriate means of ensuring accurate estimation at both the metropolitan and stratum level.[14] The allocation of EAs, corresponding sample sizes (in households), and the expected maximum margins of error are provided in Annex 1: Sampling Strata, Selection Probabilities, and Sampling Weights.

Within each stratum, the EAs in the expanded sample were selected randomly with probability proportional to size (*pps*), using the number of households as a measure of size, and with implicit stratification by the administrative units (region, district, ward, etc.). The EAs in the target sample were then selected at random with equal probability from the expanded sample. It is important to note that in Dar es Salaam, two EAs were found to be inaccessible during the listing, and therefore replace with two reserve EAs in the second stage of the sampling.[15]

In Durban, additional measures were taken to facilitate listing, including further segmentation of large EAs. Some of the selected sample EAs were found to be too large to be manageable for implementing a complete dwelling or household listing operation. Specifically, EAs with 250 dwellings or less were considered to be within the manageable threshold for conducting the field listing operation. Based on the 2011 Census information, 66 out of the 200 selected EAs reported over 250 dwellings. Large EAs selected were splited into 'segments' of roughly equal size, using a quick count of dwellings (QCD) as a measure of size. The number of dwellings in each segment was estimated based on human observation of 2015 ESRI World Imagery[16] and 2015 Google Earth imagery, and taking into account the estimated height of buildings and population density in each area. Within each of the 66 large EAs, one sample segment was selected randomly with equal probability.

Four EAs were further segmented on a case-by-case basis, as they were found to be larger than expected even after the first segmentation. This segmentation was carried out using OPM estimates on the number of dwellings based on field observation that was complemented with

[14] The alternatives would have been either by proportion to the population of the strata – which would deliver nearly optimal estimates for Metropolitan Dar es Salaam as a whole – or selecting the same number of EAs in each stratum, which would deliver estimates of nearly the same quality for each stratum.

[15] In Dar es Salaam ten additional EAs were identified as *reserve* in the first phase of the sampling to be mobilized in case of justified need. These reserve EAs were assigned to groups of 10 selected EAs. During the listing exercises the implementing firm found that two EAs were found to be in need of replacement: one of the EAs was found to be a prison and the other police quarters to which the enumerators could not gain access. These EAs were thus replaced with their assigned reserve EA. In one of the two cases, however, the assigned reserve EA was also found to be inaccessible: the EA was in the area of an airport area and it would require at least 2 weeks to get permission to start listing. The decision was taken to therefore use the 'next' reserve EA on the list of 10.

[16] Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AEX, Getmapping, Aerogrid, IGN, IGP, swisstopo, and the GIS User Community.

World Bank estimates based on detailed examination of the areas using Google Street View.[17] In addition to this, unpopulated areas were excluded in five rural EAs to facilitate fieldwork. The excluded areas were identified as either large bodies of water or green areas by reports from the field combined with analysis of recent satellite imagery.

In the second stage of the sampling strategy, 12 households were selected from each of the 200 EAs. The second stage of the sampling took place after listing. The second stage proceeded differently in the two surveys, as a reflection of differences in the process of listing (see below for further details). In Dar es Salaam 12 households from the 200 listed EAs were then selected by systematic equal-probability sampling. In Durban, 15 buildings were first selected in each of the 200 EA, and then 12 households were randomly selected from within these buildings. The total nominal sample for both surveys was thus 2,400 households. In Dar es Salaam, a small reserve of additional EAs were also identified, to be mobilized if there was justified need. For addition information on selection probabilities, please see Annex 1: Sampling Strata, Selection Probabilities, and Sampling Weights.

Listing Methodology

The listing exercise took place between the first and the second stage of sampling. The household listing operations were implemented with computer assisted paperless interviewing (CAPI) techniques, which generates electronic files directly. Enumerators collected basic information about household: the name of the household head name, phone number and total number of household members living in the dwelling. Enumerators also recorded the GPS location of all structures,[18] defined the type of structure, and aimed to provide measurement of structure size. Enumerators also recorded the number of floors in each building, and in Durban, information on the colour of roof tops.

Listing was preceded by community sensitisation in both cities. In Dar es Salaam, enumerators visited the local chief (Mjumbe) of their assigned EA two days in advance of listing and on the day of listing. In Durban, permission to conduct fieldwork was sought from Councilors, who were called and informed about the fieldwork and the likely start date. Councilors were also asked for information on the security situation in the ward; where necessary, meetings were also held with the local police or private security firms. Additional steps were also required to gain access to gated communities, such as meeting with trustees or other relevant stakeholders, who at times indicated access would be conditional on the profile of the fieldworker (e.g. female Indian fieldworkers).

In both cities enumerators were equipped with maps created on Google My Maps to display shapefiles for the listing exercise. Hardcopies of their respective EA maps were also provided to be use in case of network failure. The listing approach in Durban differed from that of Dar es Salaam because of the difficulties of gaining access to buildings. In Dar es Salaam, enumerators conducted a listing of all households in each of the selected EAs. In contrast, the listing in Durban was conducted in two phases: in the first phase, enumerators listed all buildings in each selected

[17] Three EAs were further segmented into two sub-segments of approximately equal size, with one of the two sub-segments randomly selected with equal probability. One EA was further segmented into three sub-segments of approximately equal size, selecting one of the three with equal probability.

[18] GPS coordinates of every building were recorded twice using the GPS capability of Survey Solutions and the tablet's Android location triangulation that resorts to GPS and mobile phone network mast location.

EA. In each EA, households resident in 15 buildings –selected by systematic equal probability- were then listed. In both cities, listed houses were marked in order to avoid duplication. The Dar es Salaam survey trialed a new marking system using electronic barcodes tags, but returned to chalk markings for pragmatic reasons (see Chapter 1 for further details).

Quality assurance was undertaken to check the validity of listing data. The Durban survey proved to be an example of good practice in this respect. Supervisors conducted checks, for example revisiting buildings marked as non-inhabited to ensure field workers did not misuse this classification as a means of speeding up their work. Missing information, naming conventions, and location of GPS coordinates were check using Stata. Errors in listing forms were either cleaned directly in Stata (if the information was available), or the forms were re-issued to teams. In addition to this, the Durban team introduced a further innovation: enumerators recorded the full count of dwellings in each segment using the Google My Maps interface to ensure the completeness of the listing operation.

Team composition differed significantly across the two surveys, largely as a reflection of the increased difficulty and risks faced by enumerators working in Durban. In Dar es Salaam, the listing exercise was conducted by 30 enumerators, each of which was assigned between 3 and 9 EAs for listing (enumerators were selected on the basis of performance from a group of 35 that were trained for listing). Enumerators were allocated EAs based on: (i) distance from enumerators' homes in order to minimize transport time and cost; (ii) distance between the EAs; and (iii) safety and response rate considerations. In Durban, The Durban team initially consisted of 8 enumerators and 1 supervisor, but was later reduced to 4 enumerators and 1 supervisor. Although enumerators initially worked individually – in order to avoid attracting unnecessary attention – the approach was modified in favour of groups of two or more following an armed attack on the team.

Survey Implementation

The surveys were fielded over the course of several months. The Dar es Salaam survey was implemented between November 2014 and January 2015, while the Durban survey was fielded between February and July 2015. The extended duration of the fieldwork in Durban was largely the result of: (i) constraints in finding a greater number of interviewers with the aptitude to implement the relatively complex questionnaire; and (ii) security concerns in the field, which caused delays in accessing specific EAs.

Cases were assigned to interviewers using Survey Solutions. Interviewers were provided with both an electronic and hardcopy map, as well as a printed completion form, and could contact the listing manager through email, WhatsApp, or google hangouts if they were unable to find the assigned house.[19]

Completing the survey often required repeat visits. This is because the survey required input from up to three separate respondents: the main respondent, who could be any present household member, and answered questions on household composition, basic information on members, assets, remittances, grants, housing, properties and consumption; the household head, who

[19] In Durban, there were 27 cases of listing errors across 14 EAs. Houses were classified as listing errors if they could not be found after repeated efforts by the interviewer, supervisor, and listing manager. These cases were replaced with the next dwelling on the sample list (sorted by random number).

answered questions on residential history, satisfaction, employment, time use and commuting;[20] and a random respondent, who was randomly selected from household members over the age of 12 (not including the head), who responded questions on satisfaction, employment, time use and commuting. Enumerators visited each house at least twice before a component could be marked as unavailable – in many cases, however, more than two visits were conducted.

Quality assurance procedures included: (i) In-interview feedback from CAPI, which provided a check that modules or questions were not missing, and alerted interviewers to mistakes and inconsistencies in given answers, so that these could be addressed while the interviewer was still with the respondent; (ii) Aggregate checks conducted using the Survey Solutions Supervisor application, which allows supervisors to identify common mistakes (applied to all initial interviews, and then through spot checks); interviewer performance and completion monitoring conducted by the implementing firm, through interviewer and EA level summaries of response rates, interview completion, and progress; (iii) weekly summaries of key indictors provided by the World Bank team (following each data delivery); (iv) direct observation of fieldwork; and (v) back check interviews. A key lesson learned is that the portion of back check interviews should be agreed in advance with the implementing firm: in Dar es Salaam back checks were conducted on 5% of the sample; this was considered inadequate by the team and therefore a higher portion (16.5%) was agreed and implemented in the Durban survey.

Overall, the refusal and non-completion rates for the two cities were relatively high as it was expected: 41 percent in Durban and 13 in Dar es Salaam. While these rates are high by international standards, they are not unusual for the two specific cities. High response rates are common in urban areas, given that urban residents may be away from their houses for much of the day or night, as they may commute long distances for daily work or other activities. In addition to this, higher crime rates can make urban residents particularly suspicious of strangers and therefore more reluctant to participate in surveys. Low response rates are particularly pronounced in cities such as Durban, where gated communities are common, and residents are cautious to speak with strangers.

Table 3: Summary Survey Implementation Statistics

	Dar es Salaam	**Durban**
Time in field	9 weeks	20 weeks
Team composition	30 interviewers	3 supervisors, between 18 and 9 interviewers (working in groups of two)
Number of households visited	2,397	2,407
Non-response rate (refusal, long term unavailable, uninhabited dwelling, or partially completed)	13%	41%
Total completed interviews	**2,083**	**1,420**

[20] In single person households or when the head was present, head and main respondent could be the same person

Sampling weights

Sample weights were designed to deliver unbiased estimates from the sample. The 'raw sampling weight' is a raising factor applied to each household that is equal to the inverse of its selection probability.[21] Any given household's selection probability is the product of the probability of selection at each stage of the sampling. The Dar es Salaam weights thus reflect the probability of selecting the EA in the first stage of the sampling, and the conditional probability of selecting the household in the second stage. The approach in Durban is the same, but the weights account for the additional steps after selecting the EA of: selecting that segment of the EA, if it was subdivided; selecting the building within the EA; and selecting the household within the building (for further details, please see Annex 1).

The raw sampling weights are then further adjusted to account for non-response rates. To account for non-response rates, the number of 'usable households' in each EA is calculated.[22] In Durban, non-response rates were so high that EA groupings had to be made before a weighting could be calculated. The weights were further adjusted through benchmarking to allow for post-stratification by ethnicity, in light of high levels of non-response among white households.

[21] The sample is approximately self weighted within each stratum: if the number of households in the EA, as per the sample frame and the number of households in the EA or segment, as per the household listing operation were always the same in all EAs, the formula would simplify to a constant.

[22] 'Usable households' are defined as those where at least one of the three respondents (main, household head, and random) could be interviewed, at least partially. This differs from the nominal sample of 12 per EA.

Table 4: Survey Roadmap – timelines, activities, road bumps, and resources

	Timeline	Key activities	Potential Sources of Delays	Useful links
Questionnaire preparation	2-3 months	Recruitment of consultants; Consultations. The MLSC team received extensive comments and inputs within the Bank, including comments from a panel of advisers, survey experts from the LSMS teams, and local experts for each of the cities. The team also met with policy makers, research institutions, academics, and private sector representatives in both cities.	Negotiation with firm: not having a final questionnaire at time of firm procurement gave room to additional cost negotiations further down the line	Dar es Salaam Questionnaire Durban Questionnaire
Procurement of firm	3rd month	Competitive selection of firm through econsult2		EOIs TORs
Sampling & listing	Listing took 3 weeks in Dar, and nearly 3 months in Durban	Research permits; permission from local authorities; recruitment and training of enumerators; design of sampling strategy.	- Access to census frame data - Challenge of listing households in gated communities (Durban)	Durban & Dar es Salaam Sampling Reports
Pre-test & interviewer Training	5 days pre-tests, 10 days training (minimum)	Sensitization of local community; training; modification of questionnaire		
Fieldwork	Estimated 2 interviews per interviewer per day (given length of MLSC questionnaire	Interviews Quality Assurance (real time validation) Back checks (in person & by phone)	- Holidays and festivals: fieldwork interrupted 2 weeks in December - Civil unrest: xenophobic violence outbreak in Durban delayed access to several EAs - number of respondents specified, for example, use of randomly selected	Dar (C:\Users\WB458187\Box Sync\HH surveys - data & analysis\Dar HH survey\QA indicators)

	and commuting/ daylight hours)		respondents for two modules placed additional burdens of repeat visits.	Durban (C:\Users\WB458187\ Box Sync\HH surveys - data & analysis\Durban HH survey\Quality Assurance\WB QA indicators).
Data analysis & cleaning		Assigning household weights; cleaning data; analysis.	- Specify documentation requirements & plan for time with firm after data delivery to answer questions that may arise - Allow for sufficient time between questionnaire completion and piloting for development of adequate quality assurance techniques - Low response rates may complicate household weighting	Munoz, J. 2005 'A Guide to Data Management of Household Surveys' In Household Sample Surveys in Developing and Transition Countries, chapter 15.

Survey Insights: What have we Learnt in Dar es Salaam and Durban?

Access to Basic Services

Key findings in Dar es Salaam:

- Dar es Salaam has a severe deficit of basic services, and many households have even 'lost' access over time.
- Migration into the city does not improve access to services: migrants have similar levels of access to services as non-migrants within the city, but access rates are not much higher than before migration – in fact, 15 percent report reduced access to improved water.
- There is spatial variation in the constraints households face in accessing basic services. Spatially targeted policies – for example, that look at in-house connections in the center of the city, or address use of unprotected natural sources in the periphery – could yield important improvements in access.

Key findings in Durban:

- Overall, higher rates of access to basic services but challenges remain in terms of connectivity to the public network and quality of service
- Variation in access to services persist across city areas despite better average access
- Access to improved services is highest in the center of the city; and the majority of those who have access are serviced by the public water and sewage network

Access to basic services is declining in Dar es Salaam. There is evidence to suggest that access to basic services in cities have failed to keep pace with urbanisation across Sub-Saharan Africa.[23] Dar es Salaam is no exception to this trend: the proportion of the population with access to basic services has declined as the urban population has expanded.[24] As the following chapter highlights, however, the MLSC survey provides new insight on the spatial distribution of these challenges within the city. The results show that declining access is not primarily driven by lower services in rapidly expanding peri-urban areas or among migrants. As the findings below indicate, reduced access is first and foremost a reflection of the decline and degradation of existing urban infrastructure in the center and consolidated areas of the city.

Only a fraction of houses in Dar es Salaam have piped water, despite the fact that at least half live in areas that are serviced by trunk infrastructure. Only 11 percent of households in Dar es Salaam have piped water in their houses. The majority of households in the city rely on their neighbors for access to water: over 40 percent of households in Dar say that their main source of water is piped water from a neighbor. In the center's irregular settlements, this number climbs to more than half (51 percent). This suggests that a major challenge to improving access to piped water is connection charges

[23] Urban infrastructure has not kept pace with urbanization across the region. Recent work includes analysis of satellite data of night time lights in urban areas (see Tanzania Urban Policy Note).
[24] NBS 2009 shows access to piped drinking water fell from 93 percent in 1991 to 61 percent in 2007.

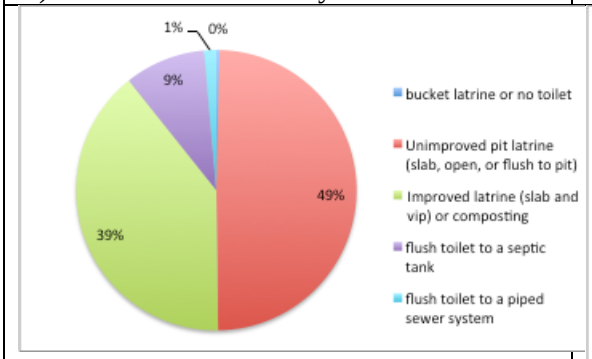
and policies. Indeed, when asked about barriers to water access, half of households suggest that they cannot connect to the water system because they rent their house.[25] An additional 20 percent cite high connection costs as the main hurdle.

Water services are unreliable, and access rates are declining. More than one third (37 percent) of household in the survey do not currently have access to improved water stated that they have previously lived in a house in Dar es Salaam that did have access. Furthermore, even among households that do have piped water, there is a major challenge around reliability of water services. On average, households with piped water reported that in the last week they had less than five days and 14 hours per day of service. Twenty one percent of households reported as few as five hours a day.

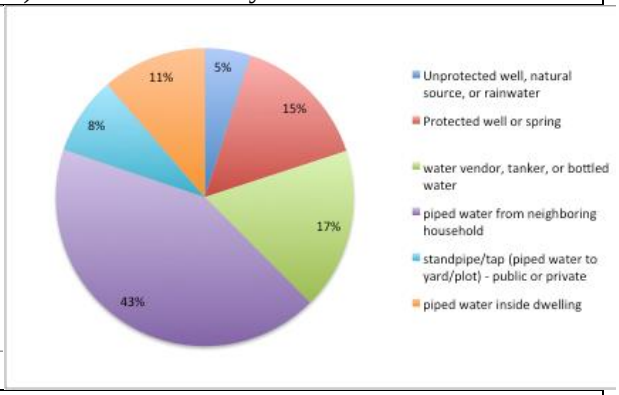
Half of households in Dar es Salaam lack access to improved sanitation; and 18 percent of those without access report that they did have it at some time in the past. As highlighted in **Figure 1** the most common form of sanitation in Dar es Salaam is an improved pit latrine, which is used by nearly forty percent of households. Other forms of improved sanitation are rare: only around ten percent have a flush toilet, and nine out of ten of these toilets are connected to a septic tank rather than a piped system. Pit latrines are the most common form of unimproved sanitation. Overall, about two thirds of households in the city share their toilet facilities, and there is some indication that access rates are 'relapsing': of the households in the survey that do not currently have access to sanitation, eighteen percent said that they did have access to this basic services in a house they previously lived in within Dar es Salaam.

Figure 1: Many Households Rely on Unimproved Basic Services in Dar es Salaam

A) Access to Sanitation by Source



B) Access to water by source



Improved latrine combines 33% slab latrine and 6% vip latrine; unimproved pit latrine is made up of 26 percent unimproved, 16 percent open slab, and 7 percent pour flush to pit.

The story is quite different in Durban, where over 66 percent of households have access to piped water in their dwellings. Despite overall access being considerably higher in Durban, within city variations in access are large. While almost 87 percent of households in the urban core have piped

[25] It is worth noting that almost 80 percent of those who said that the reason they did not have access was because they rent have a written agreement of rentership. This is a similar to the overall proportion of renters that have a written agreement in the sample.

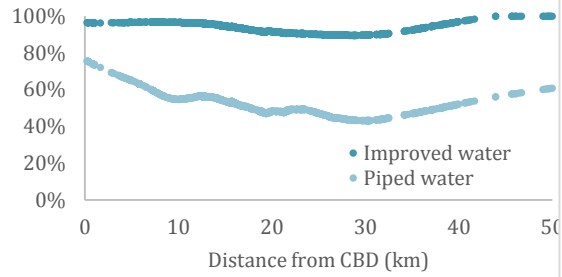
water in their dwelling, less than 20 in the outskirts of the city and less than 15 percent in shanty areas do so as well. In shanty areas, households do not rely as much on neighbors as they do in Dar. Only 4.6 percent of households in shanty areas and 7.7 percent of those in rural areas cite their neighbors dwelling as their main source for drinking water. Instead, almost 42 percent of households in shanty areas and 52 in rural areas rely on yard taps. Street standpipes are also an important source in shanty areas with 35.5 percent of households relying on them for drinking water.

Durban is also considerably ahead of Dar in terms of sanitation on average, but spatial variation in access is also stark. While almost 92 percent of households in the urban core are connected to the sewage system, only 18.7 in the outskirts of the city and 45.3 in the shanty areas are. In shanty areas, which are not very distant to the core of the city, over 44 percent of households still rely on pit latrines as their main form of sanitation compared to 15 percent in peri-urban formal areas and over 76 percent in rural areas in the outskirts of the city.

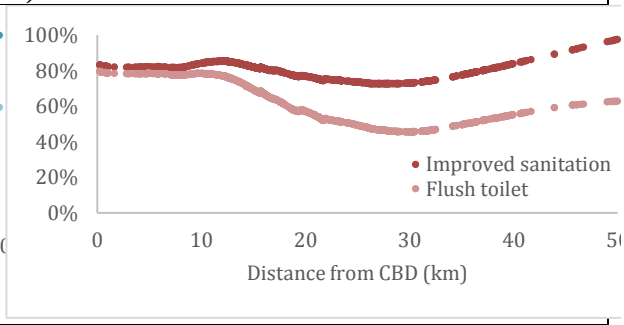
Figure 2 below shows how access to piped water and flush toilet varies with distance to the city center. A locally weighted smoothing scatterplot suggests that overall, access decreases as distance to the city center increases. The average increase in access beyond 35 km from the city center should be interpreted with care as it is driven by the higher access found in peri-urban areas further away from the city center. Further exploration using a local polynomial estimation suggest some over-smoothing is happening that hides away the low access values observed in rural areas and mentioned above.

Figure 2: Access to Services in Durban Decreases with Distance from the City Center

A) Piped Water in House



B) Flush toilet in House

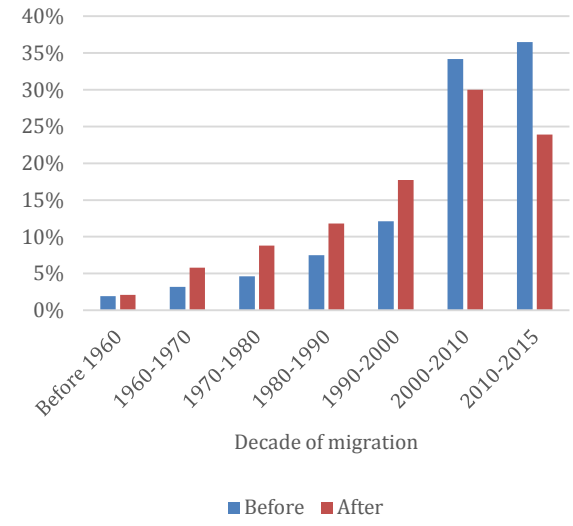


Source: Authors calculations using lowess regression and data from WB's 2015 Measuring Living Standards within Cities Survey in Dar es Salaam and Durban

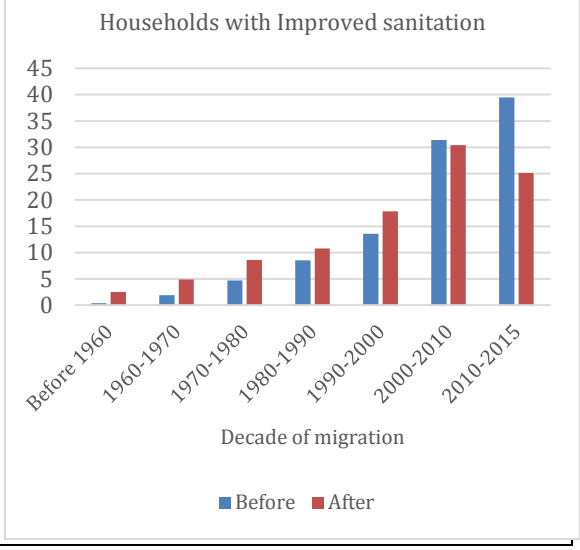
A history of access and migration also sheds light on the challenges faced today and how they vary with the stage of development of the city, and the country as a whole. Early migrants who came into Durban between1960 and 2000, enjoyed improvements in access to basic services, compared to their previous residences. Today, the challenges are deeper with recent migrants facing a deterioration of their living standards as they move into the city. They are sacrificing living standards in search for the opportunities that the city brings.

Figure 3: Migrants No Longer Improve Access to Services in Durban with Migration

A) Households with Improved water



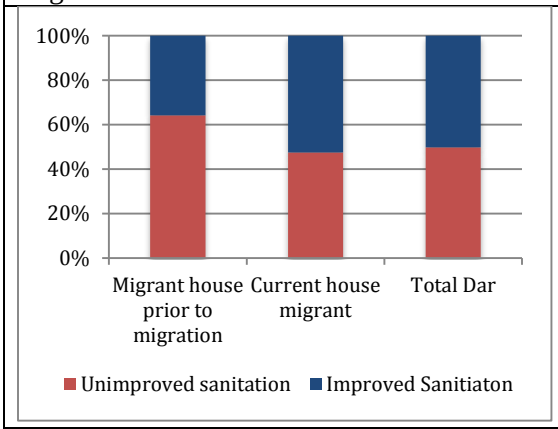
B) Households with Improved Sanitation



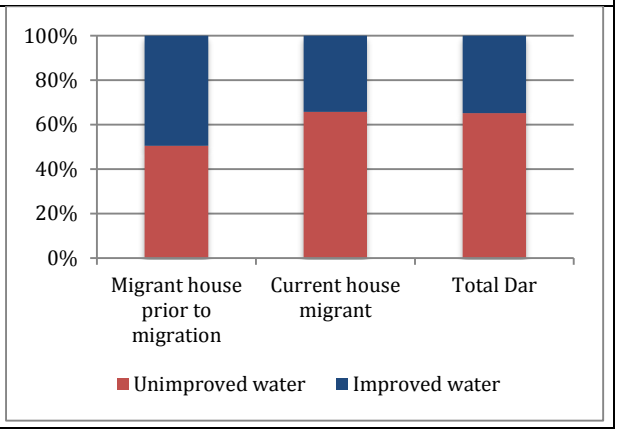
Migrants have similar rates of access to services as non-migrants in Dar es Salaam; migration does not, however, appear to improve access rates. The MLSC survey asked households to recall features of houses they have lived in the past. As highlighted in **Figure 4** half of migrant households recorded that the house they lived in before they moved to Dar es Salaam had access to improved water – compared with about one third of houses that migrants live in today. Access to sanitation improves with migration, from just over 30 percent to more than half of households. It should, however, be noted that, for about half of migrant (and non-migrant) households, even improved sanitation is most commonly an outdoor and shared facility – as will be discussed further in Chapter 4 on Density. These findings warrant further investigation: they raise questions about drivers of migration to Dar es Salaam, since access to services is commonly conceived of as being one of the many potential "pull" factors in rural-to-urban migration.

Figure 4: Fifteen Percent of Migrants Had Better Access to Water Before Migrating to Dar Es Salaam

A) Access to Improved Sanitation before migration and in current house



B) Access to Improved Water before migration and in current house

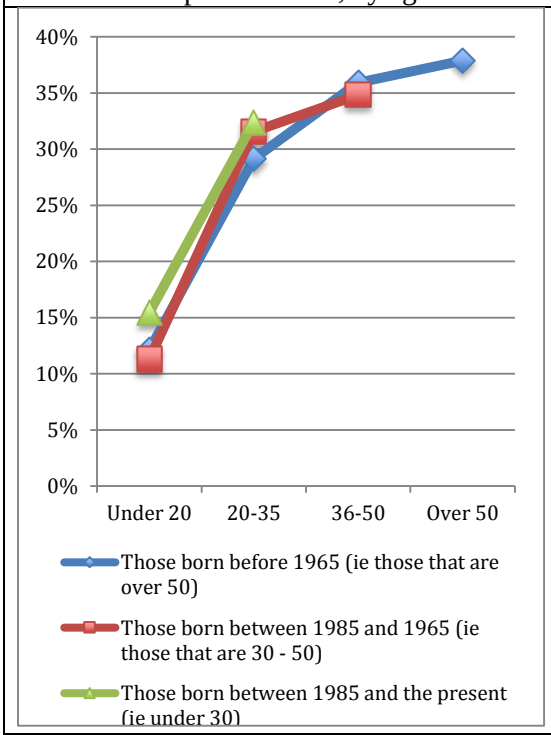


Note that for migrants who have remained in their first house in Dar current house is first house in Dar.

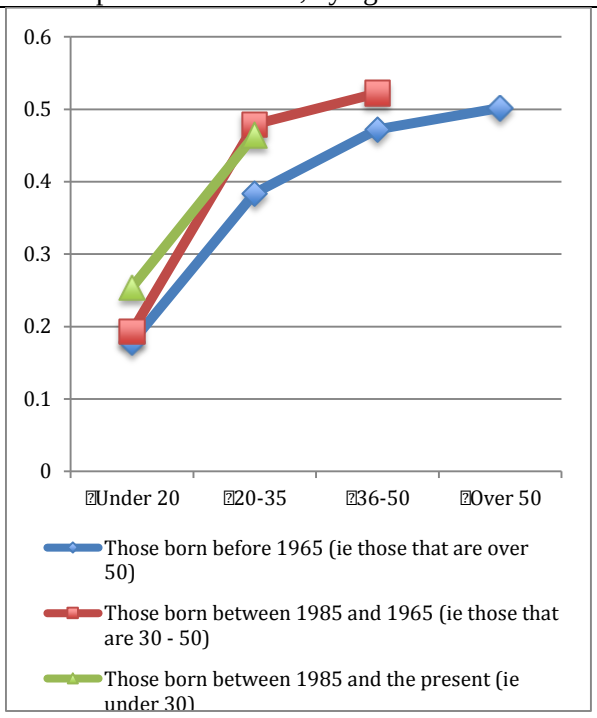
Access rates in Dar es Salaam are relatively consistent by age cohort and gender of household head. Figure 5 investigates access to services by age cohorts, breaking down the survey sample into three groups according to the date of birth of the household head. The two graphs chart the age at which households that currently have access to basic services first moved into a house in Dar es Salaam with access to improved water and sanitation (respectively), by age cohort. The graph shows improvement in the rate at which households gain access to sanitation among the younger generations, but only marginal advances for water. It does not, however, account for rates of relapse. Gender does not appear to be a major cause of disparity either: the proportion of female headed households with access to water and sanitation is one percent less than male headed households, for both water and sanitation.

Figure 5: Younger generations may access sanitation earlier than older generations, but there is little evidence of divergence on water services.

A) Age by which households lived in a house with improved water, by age cohort



A) Age by which households lived in a house with improved sanitation, by age cohort

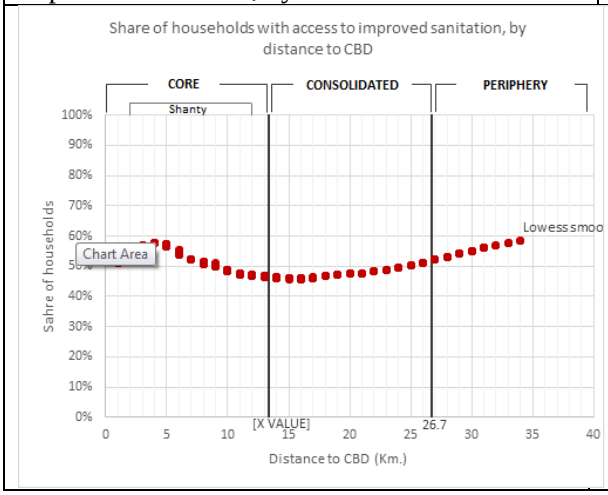


NB: these graphs mark the age at which household heads first moved into a house with basic services. It does not account for relapses in access.

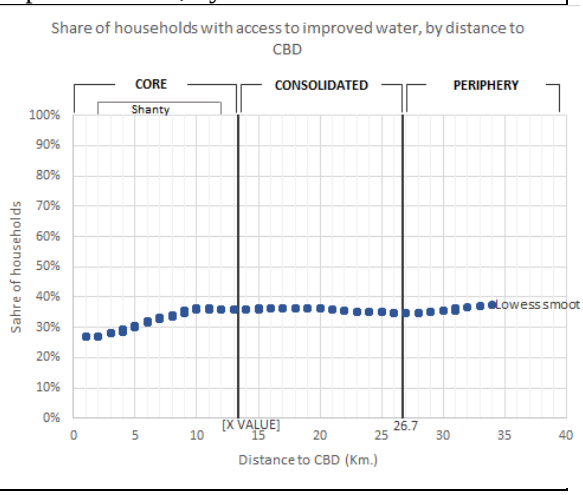
There is a strong spatial dimension in access to basic services, with peri-urban areas demonstrating comparatively high levels of access. It is often thought that it is the rapid urbanisation of the periphery of Sub-Saharan African cities that drive service deficits – as the city extends outwards at a pace that urban infrastructure expansion cannot match. As Figure 6 highlights, however, rates of access to improved sanitation are lowest in areas that make up the core and consolidated areas of Dar es Salaam (between 5 and 20 kilometers from the city center), and actually improve towards the periphery of the city. In terms of water, the area of greatest deficit is the city center, with access levels otherwise remaining fairly consistent across the consolidated and periphery of the city.

Figure 6: The deficit in access to improved water in highest in the center of the city, and sanitation access also shows strong spatial trends

A) Share of households with access to improved sanitation, by distance to CBD



B) Share of households with access to improved water, by distance to CBD

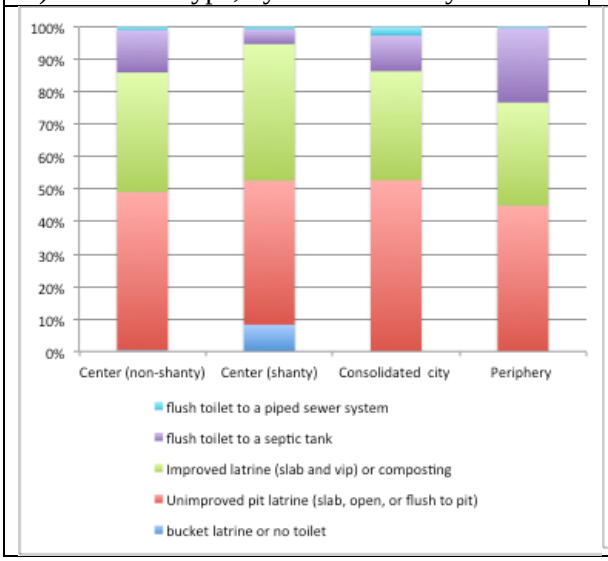


There is, however, striking variation in types of provision that households rely on for access to sanitation and water across the city. The distribution of services across the city is shown in **Figure 7**. In terms of sanitation, households across the city rely on latrines, and less than one percent are connected to the piped sewage system. The major difference between sanitation services between the periphery and other parts of the city is that there are a greater number of septic tanks in the periphery. Households are also less likely to share sanitation facilities – 36 percent of households in the periphery share their toilet facilities, compared with 44 percent in the consolidated city. This rises to 62 percent and 80 percent in the regular and irregular areas of the city center, respectively.

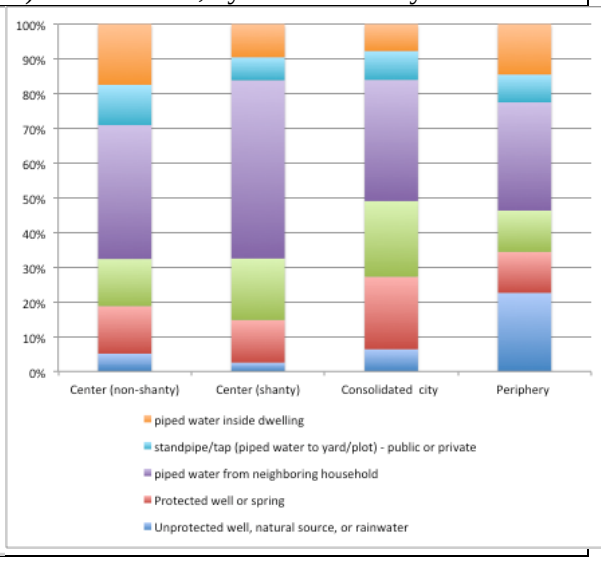
Differences in water sources across the city suggest that spatial dynamics should be factored into service improvement priorities. In the city center, short-term gains in access to water could be achieved through improvement in in-house connections. The majority of households live within reach of the existing piped water network but rely on indirect means to access it: only 17 percent have a connection in their house, but an additional 50 percent use piped water from neighbours or standpipes. The consolidated city, in turn, has the smallest portion of households with access to the piped network (53 percent, both directly and indirectly) and the highest that rely on trucks, vendors, or bottled water (42 percent). Given that these forms of water are often the most expensive for end users, this suggests that capacity to pay for piped water is not the major constraint to wider access in this area. In the periphery, in turn, there is an urgent need to address the fact that those who are not able to access piped water are most likely to use unprotected natural sources (23 percent) – a likely source of serious health hazards in urban areas.

Figure 7: Households in the Center of the City are More Likely to Rely on Neighbours for Access to Services

A) Sanitation type, by area of the city

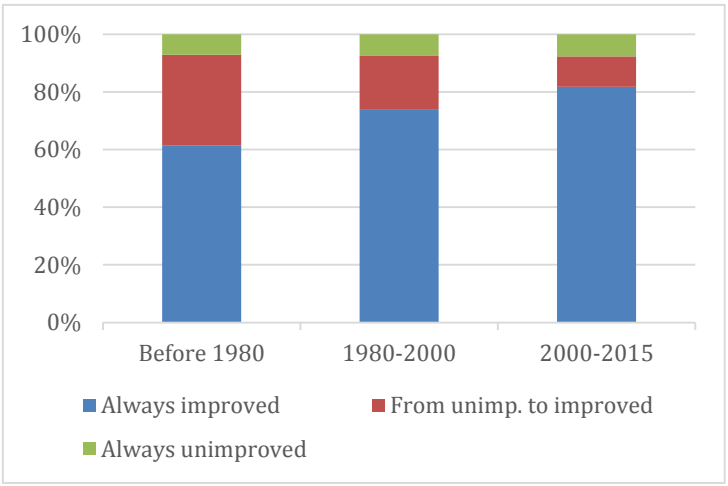


B) Water source, by area of the city



In Durban, looking at the history of household moves within the city it is also evident that the relation between moves and access to basic services has changed through time. While about 31 percent of within city moves before 1980 were from a house that did not have access to improved sanitation to one that had, this percentage has decreased over time, as access to improved sanitation has expanded. 18.8 percent of moves between 1980 and 2000 and 10.6 percent of moves between 2000 and 2015, reflected a move towards better sanitation. Across all periods, about 7 percent of households have moved between areas with no access to improved sanitation. This can be seen in Figure 8.

Figure 8: Mobility Within Durban is Often Accompanied by Improvement in Access to Services



Informality in Housing and Jobs

Key findings in Dar es Salaam

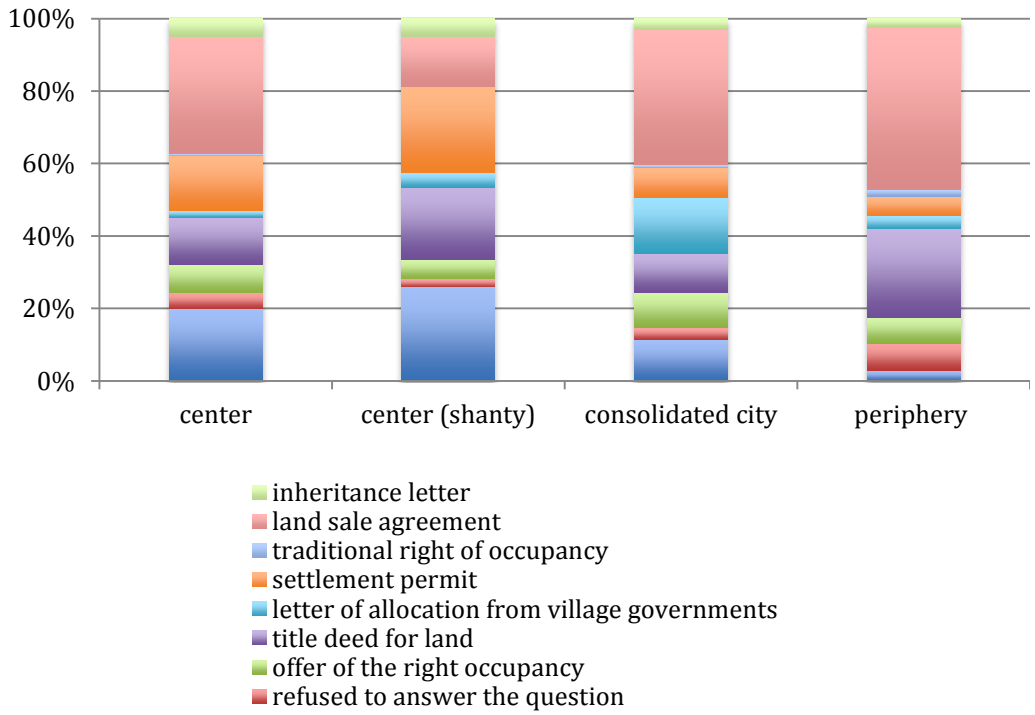
- Informality and insecurity of tenure is prevalent, but there is little overlap between them.
- Most jobs in Dar es Salaam are informal, low value-added services.
- While housing informality is higher in the center of the city, employment informality increases with distance to the center.

Key findings in Durban

- Informality of land is more prevalent in rural and peripheral slum areas wit. half of homeowners lacking formal documentation
- While low value- added services dominate the job distribution, different from Dar es Salaam, in Durban there is also an important share of jobs in manufacturing
- Job informality is spread out through the city, suggesting no clear spatial pattern

One quarter of home owners in Dar es Salaam lack any form of documentation of ownership. Just over one third of households in Dar es Salaam own the property they live in, but one quarter do not have a single document of proof of this ownership. In fact, when home owners where asked to identify ownership documents they had to prove their property status from a comprehensive list, 17 percent stated that they had never even heard of any of the documents (see **Figure 9**, below). This rises to 20 and 26 percent respectively in non-shanty and shanty areas in the center of the city. Among renters – who dominate the housing market, as highlighted in chapter 4 – 11 percent do not even have a written agreement with their landlords.

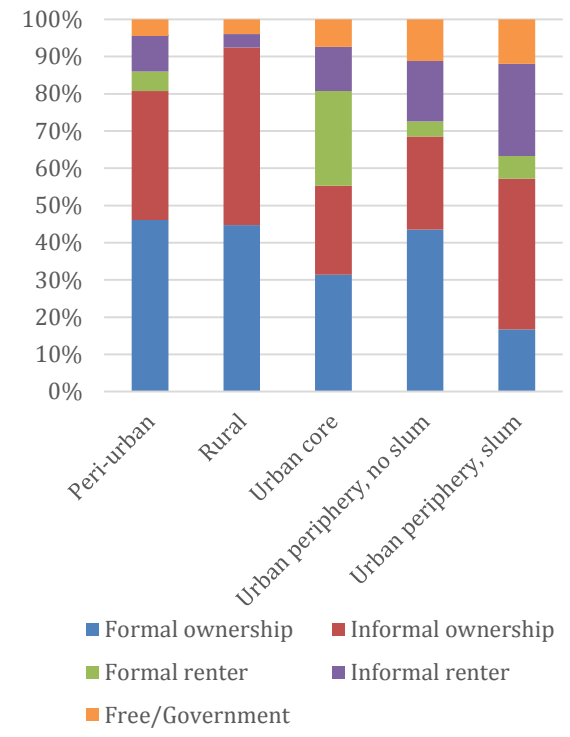
Figure 9: Awareness of Land and Property Ownership Documents is Limited



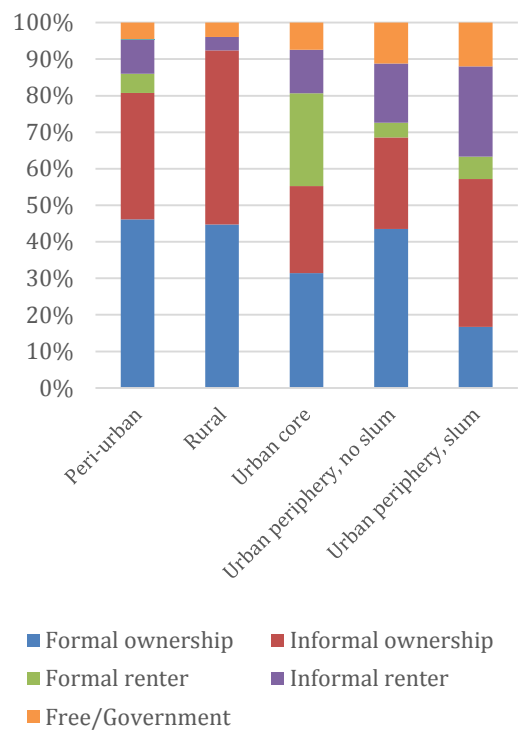
Housing informality is more prevalent in the center of the city and among non-migrants. Informal home ownership is most common in the center of the city (9.5 percent), while informal renting makes up a larger portion of the housing in the shanty areas of the center (13 percent). Migrants are more likely to have some document of tenure than non-migrants, whether they are renters of homeowners: the majority live in accommodation for which they have a written rental contract, but even among owner occupied housing they are slightly less likely to lack formal documentation, as seen in Figure 10. The proportion of migrants in informal housing does not vary considerably with the amount of time migrants have been in Dar es Salaam, although they are more likely to become home owners over time as outlined in Chapter 4.

Figure 10: Rates of Informal Ownership and Renting Vary Spatially and Among Demographic Subgroups

A) Dar es Salaam



B) Durban



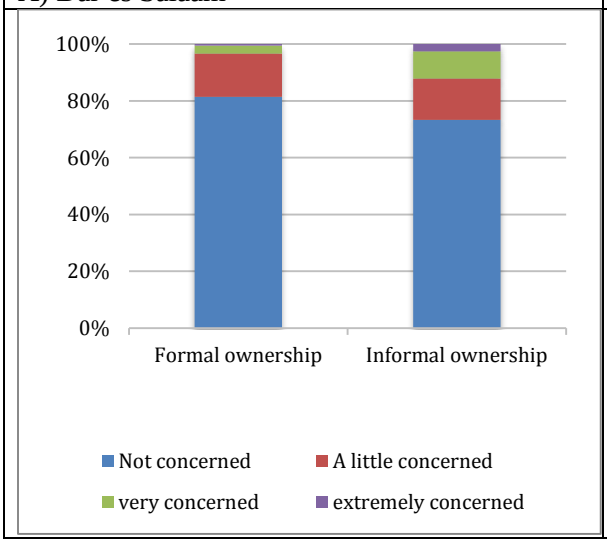
In Durban, housing informality is more prevalent in the rural and peripheral slum areas with 48 percent of households lacking formal documentation. Rural areas and urban peripheral slums areas have the higher percentage of owners with no documentation that supports their current status, with about 45 and 40 percent of houseohlds. Even in central areas, lack of ownership documents appears to be a common fenomenon with just over 20 percent of households with no proof of ownership. Formal rental markets seem much stronger in Durban than in Dar es Salam, with about 30 percent of households in the city core declaring to have a formal rental agreement.

Twenty percent of households in Dar es Salaam are worried about having their property taken away from them, and government expropriation is the greatest stated concern. Informality is often associated with insecurity of tenure, as lack of clear, formal tenure can make households vulnerable to conflicting claims on their property. Despite high rates of informality and the plurality of tenure documents that exist, insecurity of this form is not self-evident in Dar es Salaam: 80 percent of homeowners say that they do not worry about their property being taken away from them without their consent. Among the 170 households that said they were either a little, very, or extremely worried about their property being taken away, 154 said they worried about the government taking their property away from them without their consent, compared with

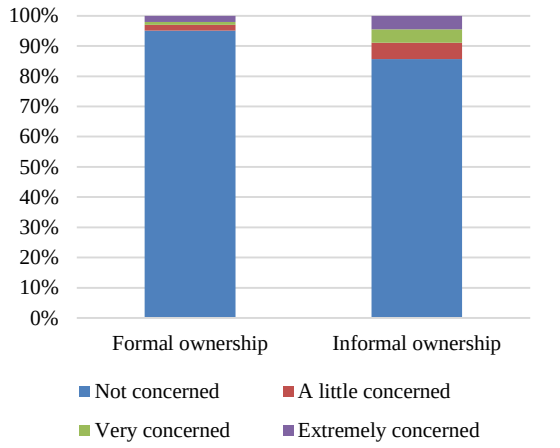
ten who cited relatives and one a money lender/bank.[26] It is also notable that households that are concerned about expropriation do not necessarily have informal tenure: seventy percent say they do possess one or two ownership documents. In Durban, the level of conern rises from 5 to 15 percent depending on whether or not households have documents of ownership

Figure 11: Households Are Not strongly Concerned about Having Their Property Taken Away From Them

A) Dar es Salaam



B) Durban



In Dar es Salaam, households move between formal and informal tenure over time. Households in the MLSC survey were asked to talk about their residential history, going over any previous houses in Dar es Salaam they had lived in. For migrants this included the last house they lived in before they moved to Dar es Salaam. In recording these residential histories, more than half of households gave information about the tenure status of past houses (60 percent). It is interesting to note that the majority of these households reported having held formal and informal tenure status at different times in the past. Even excluding their 'first house'[27] nearly half of current formal owners said they had informal tenure status in the past (39 percent as renters, and 4 percent as homeowners). A similar portion of informal owners said they held formal tenure previously (44 percent were formal renters, and 15 percent had documents of ownership). Renters had also moved between formal and informal status: 25 percent of current formal renters were informal renters in the past, while 32 percent of current informal renters had formal contracts in the past.

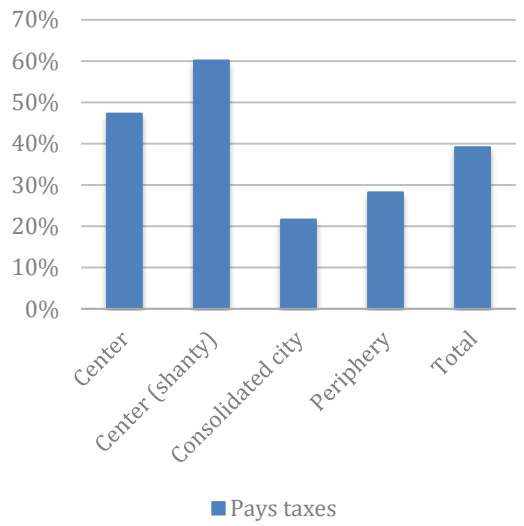
[26] It should be note that in Tanzania the government has extensive rights to expropriation – the Land Acquisition Act (1967) and Land Act (1999) maintain the President's role as custodian of all land with extensive powers to acquire land for public-interest uses (Kombe, 2010).

[27] i.e. the household head's parent's house for non-migrants and a house outside of Dar for migrants

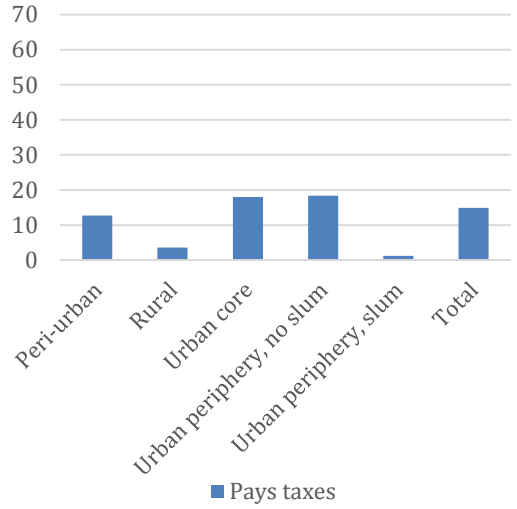
There is strong spatial variation in the payment of property taxes. Property informality is often associated with lower tax revenues for governments. In the two cities, however, the link between taxation and property formality is unclear. Forty percent of homeowners in Dar es Salaam pay taxes on their property while less than 15 percent do so in Durban. Three quarters of those in Dar es Salaam have at least one document of ownership, but one quarter are entirely informal. In addition to this, one quarter of households with formal ownership documents say they do not pay any taxes. There may, in fact, be a stronger spatial story that determines taxation payment rates: a greater proportion of households in the center of the city pay taxes than they do in the periphery for both cities (see Figure 12). This may reflect the reach of the property tax registry, which in many rapidly urbanising cities lags behind the growth of the physical extent of the city.

Figure 12: There is Strong Spatial Variation in Payment of Taxes

A) Dar es Salaam



B) Durban



Even among formal home-owners, access to housing finance is very low. Another key benefit of formal home ownership is access to finance: a formal property title is considered a prerequisite for the poor to unlock the value of their housing assets, and therefore can support poverty reduction and economic growth objectives. In Dar es Salaam, access to mortgage finance is low. Only three households in the MLSC survey said that they paid any kind of mortgage or financial payments in the last thirty days. This confirms findings form other work that suggest that property informality is one of a number of barriers to wider access to housing finance in Dar es Salaam.[28] These observed trends between informality and insecurity of tenure, payment of property taxes, and access to finance merit further investigation.

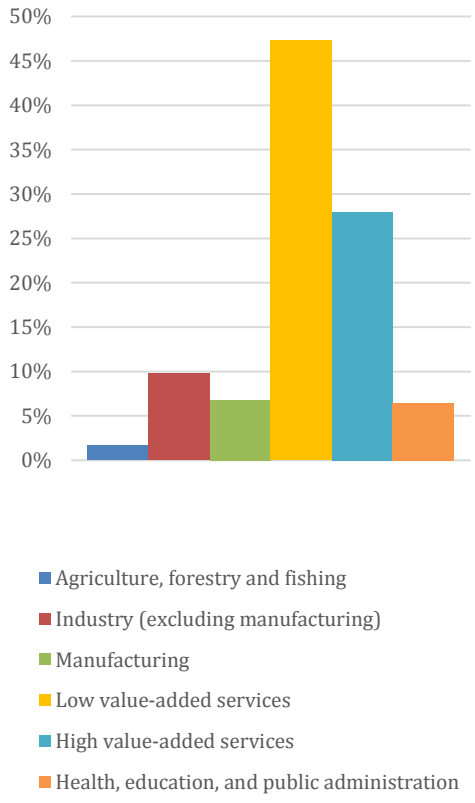
[28] Parsa et al (2011) and World Bank (Housing Finance Project Appraisal Document 2009), note that formal salary is often a requirement for access to loans, since loan repayment are often deducted directly from the payroll.

A look at the industries in which individuals work, suggests the dominance of services with over 47 percent of household heads working in low value-added services in Dar es Salaam and just below 40 percent in Durban. High value added services follow with 28 percent in Dar es Salaam and 18 percent in Durban. The most striking difference between Durban and Dar es Salaam is the percentage of household heads working in Manufacturing. Manufacturing is the second most important sector in Durban with about 20 percent. It is not surprising that low value added services are a higher percentage of jobs in the core-shanty areas than in the rest of the city with almost 55% of households heads being in this sector in Dar es Salaam and 51 percent in Durban.

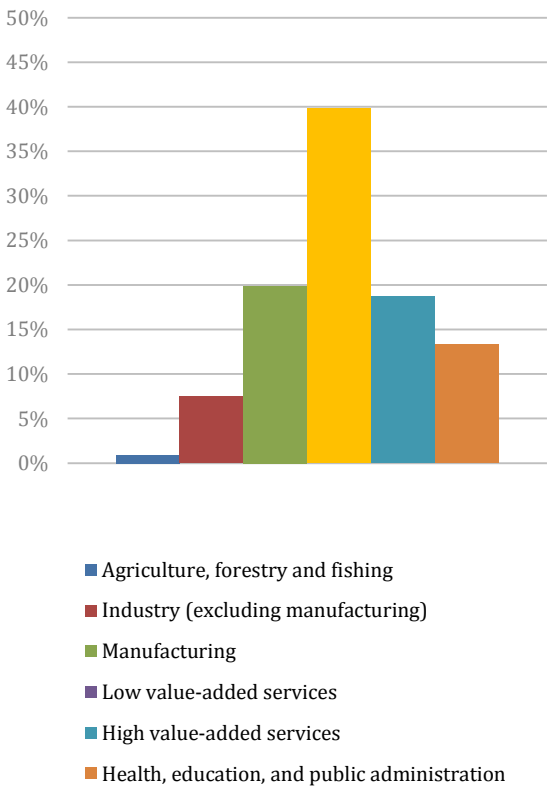
Job markets also seem to be largely informal with less than 40 percent of household heads reporting a formal contract or being formally registered if self-employed. Informality appears to be higher as one moves away from the city center. With 56 percent of heads being in the informal sector in the center and over 63 percent in all other areas of Dar es Salaam. In the consolidated city and the periphery, with 59 and 65 percent of heads being self-employed. For nearly half (47 percent) of household heads, this employment is in low value-added services. In Durban, 40 percent of households are also in low value added services, however, different from what was observed in Dar, a much larger percentage of individuals is involved in manufacturing activities (20 compared to just above 5 percent).

Figure 13: Most individuals are employed in low value added services

A) Dar es Salaam



B) Durban



Job informality is higher in the periphery than the centre of Dar es Salaam but seems to be evenly spread out in Durban. The percentage of households heads that are informally employed increases from 56 percent in the center of the city to over 63 percent in all other areas. In Durban, the rates of informal employment stay around 30 percent in all areas with the highest rates found in urban peripheral non- slum areas with 36% of informal employment. Rates of self employment also increase with distance from the city center in Dar es Salaam: 59 and 65 percent of household heads are self-employed in the consolidated and periphery of the city, respectively. This trend contrasts with housing informality, which is higher in the city center than the periphery. There is nonetheless, some overlap between housing and job informality: only twenty percent of those with formal jobs lack any form of tenure documentation, compared with thirty percent of informal workers.

Figure 14: Dar es Salaam, Informality and type of employment, by stratum (hh head)

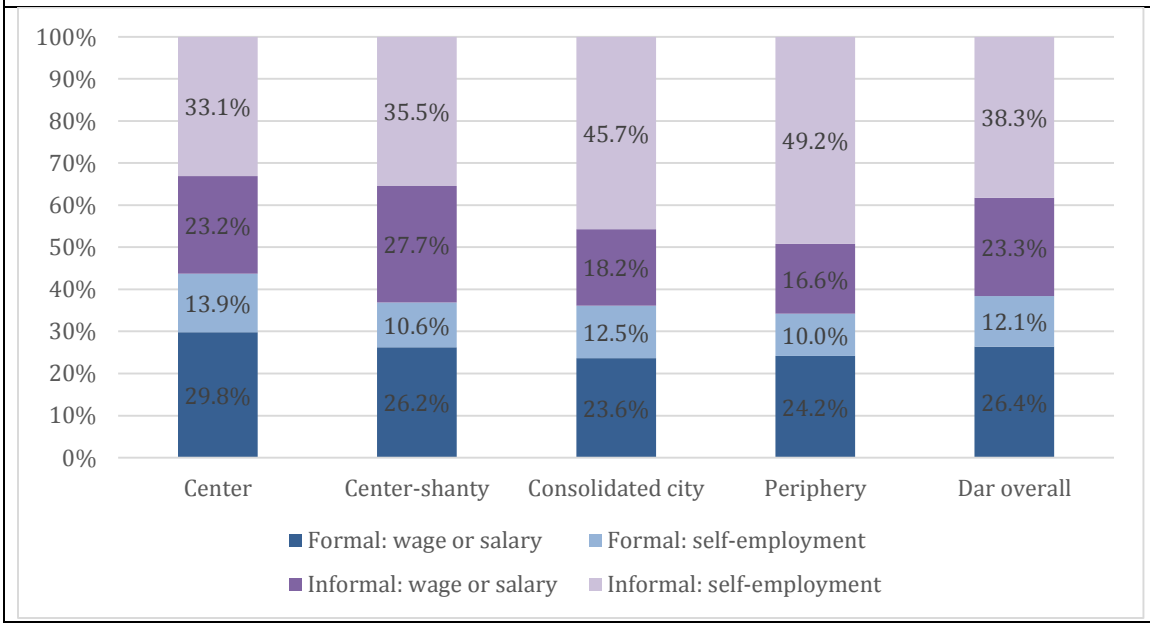
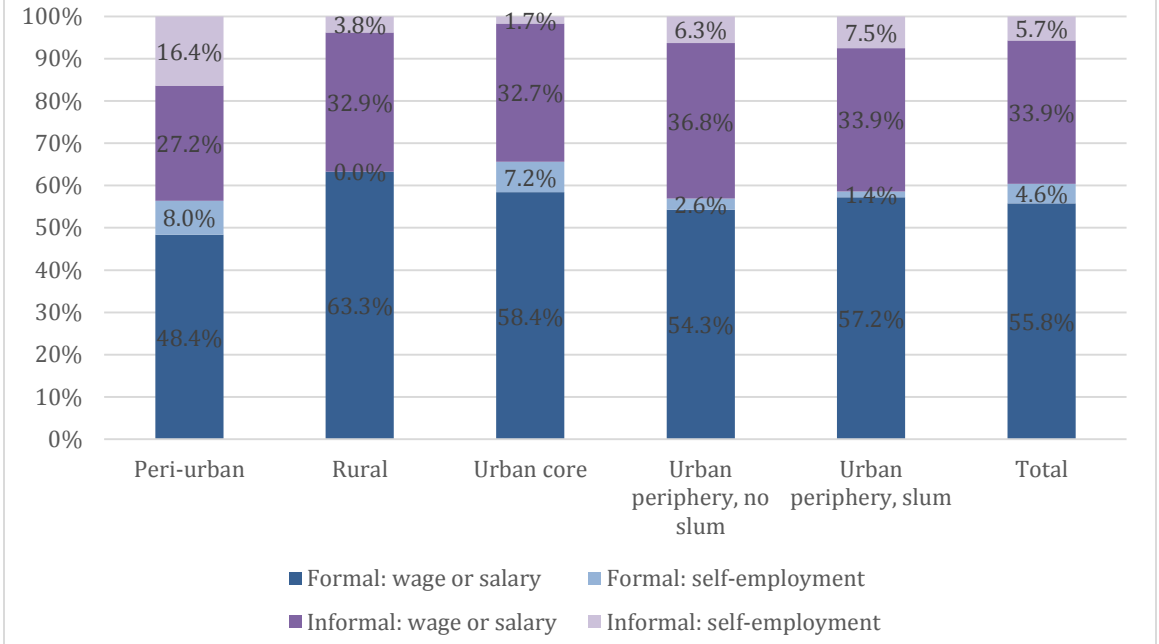


Figure 15: Durban, Informality and type of employment, by stratum (hh head)



Urban Density and Crowding

Key findings from Dar es Salaam

- Labor markets are highly localized, indicating that firms do not benefit from being able to recruit from a larger pool of workers across the city, and workers are not able to take advantage of a wider range of employment than that which is located in their immediate vicinity.
- Other potential social, economic, and environmental advantages of agglomeration are undermined by lack of basic services. Basic service deficits are highest in the most dense areas of the city.
- Nearly one third of the population lives in over-crowded living conditions

Key findings from Durban

- Diverse transport mode use and longer commutes suggest less fragmentation of labor markets in Durban compared to Dar es Salaam
- Slum dwellers face deeper challenges with longer commutes and more walking trips

Dar es Salaam and Durban are populous cities, and population density can be a desirable feature in urban areas. Density is associated with 'benefits of agglomeration', such as economies of scale and scope. With density, product and input markets can grow, reducing costs for firms and increasing the basket of goods available for consumers to choose from. In addition to this, there can be environmental benefits to density, as it can support more efficient public transportation, waste treatment, and energy systems.[29] These conditions combined can further attract new firms and workers to the city, thereby supporting a flourishing urban economy.

Crowded cities, however, face many problems. Crowding is a term that can be used to denote situations when the costs of population density outweigh or even undermine the benefits of agglomeration. In crowded cities, movement of people and goods is constrained by traffic jams and congestion. The environment is often severely damaged by pollution from household refuse as well as transport and industry. Economies of scale and scope are undermined, and economic productivity may be damaged. It is therefore important to identify what distinguishes beneficial density from costly crowding.

As the following chapter highlights, Dar es Salaam and Durban displays many of the symptoms of crowding; and improved basic services and connective infrastructure are needed to leverage the benefits of density. Basic services such as water and sanitation are vital because high population density can magnify the health risks and environmental costs associated with polluted water sources and untreated waste. Connective infrastructure is necessary because congestion effectively adds 'distance' – in terms of the cost and time needed for goods and people to move in the city – undermining the benefits of agglomeration and increasing pollution.

[29] For further information on the environmental benefits of density, see the World Bank's Cities and Climate Change agenda, as well as UNEP Livable Cities.

Effective urban management includes the need to ensure that land is efficiently allocated and reallocated over time to productive uses.

Results from the MLSC survey reveal three new insights that suggest labour markets in Dar es Salaam and –to a lesser extent Durban, are highly localized. This is important because localized labour markets limit agglomeration economies: it means that firms do not benefit from being able to recruit from a larger pool of workers across the city, and workers are not able to take advantage of a wider range of employment than that which is located in their immediate vicinity. In effect, they suggest that the city operates more as a series of villages than one urban area.

The first indication of very local employment markets is that commuting distances are small. The MLSC survey collected information on the location of work for both the household head and one random respondent in each household. This information shows that in Dar es Salaam households travel comparatively small distances to get to work: household heads' commuting distances is less than six kilometres, and random respondents travel only 4 kilometers to get to work. As one might expect, distances do vary by sector - with household heads in formal employment and higher value added sectors showing greater willingness to travel for jobs[30] - and increase in less built-up areas of the city (see Figure 16). Even at their peak, however, these commuting distances are comparatively low: in Nairobi, for example, the average commute in 2008 was estimated at between 30 and 40 kilometers.[31]

Most people walk to work, and for those that use motorized transport, distances are constrained by congestion. No matter where you live in Dar es Salaam, the most common way of getting to work is walking. Overall, 43 percent of household heads in Dar es Salaam walk to work, and, as indicates, walking is consistently the most common form of mobility across the city. Buses and minibuses are the second most common form of transportation. It is particularly prevalent in irregular areas in the center of the city where 46 percent of household heads travel to work by bus. Initial analysis of the survey results indicate that congestion limits distances travelled even by motorized means. These patterns are being investigated further through a follow up mobile-phone survey (see Box 1 in the introduction).

[30] For example, workers employed in high value added services travel about one and a half kilometres further than those in low-value added services (5.3 km and 6.75 km respectively). Formal wage or salaried workers travel nearly three kilometres further informal wage/ salaried workers, and the formal self employed travel more than a kilometre more than their informal counterparts (6 km compared 4.5 km; see chapter xxx for further details).

[31] City council of Nairobi http://www.thegef.org/gef/sites/thegef.org/files/gef_prj_docs/GEFProjectDocuments/Climate%20Change/Regional%20-%20%283461%29%20-%20Promoting%20Sustainable%20Transport%20Solutions%20for%20East/06-09-2009%20ID3461%20Endorsement%20letter%20from%20govt.pdf

Figure 16 Commuting distances are Low in Dar es Salaam

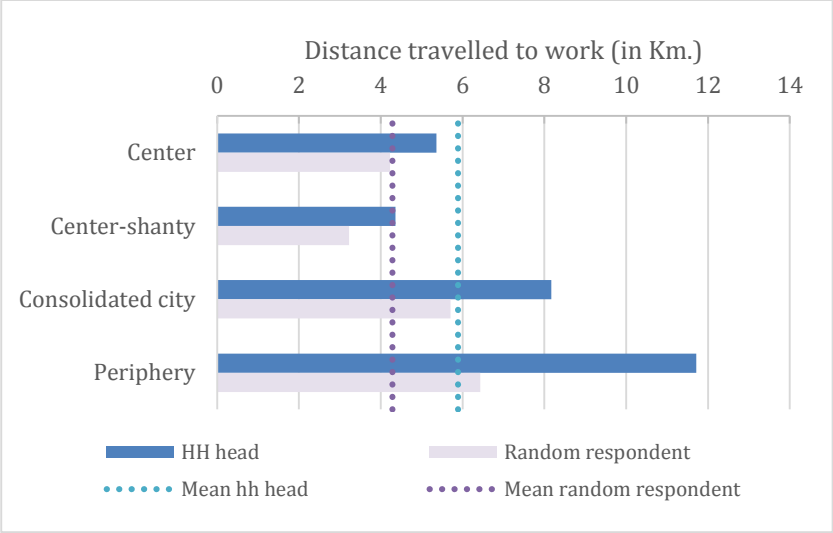
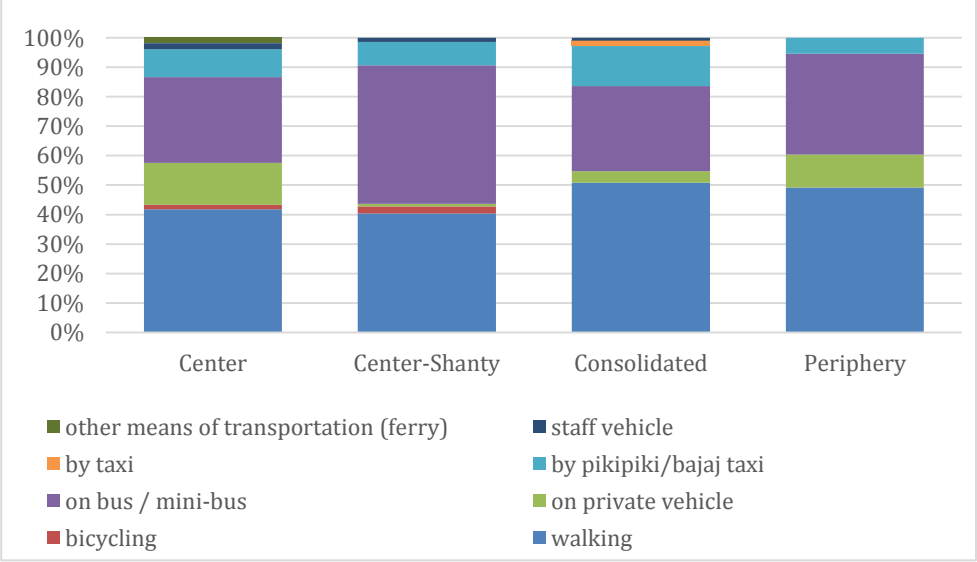
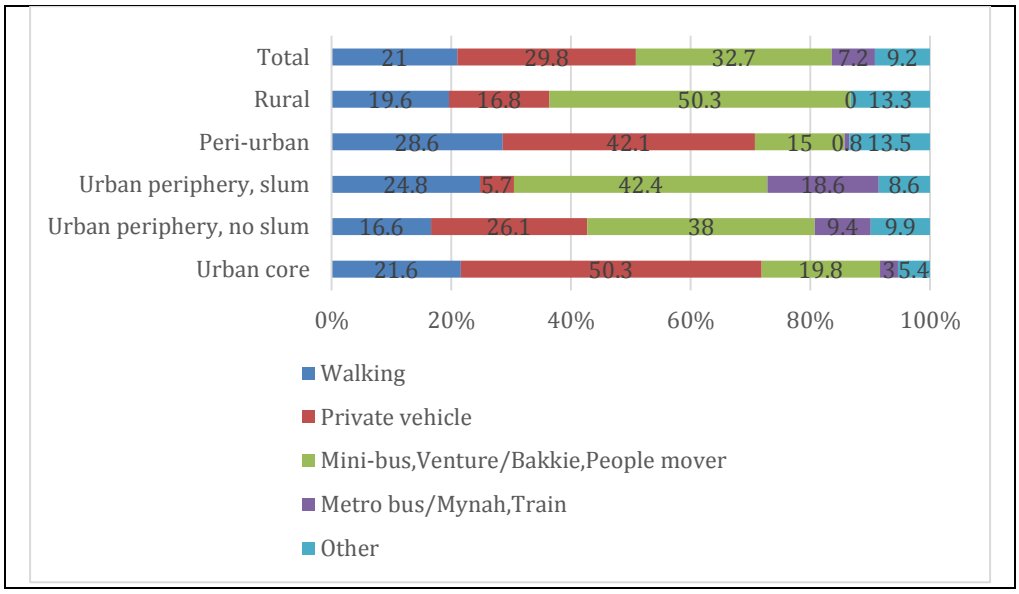


Figure 17: Mobility Patterns indicate Localized Labor Markets

A) Most People Walk To Work in Dar es Salaam



B) While in Durban the mode distribution suggest a wider use of mass transport and private vehicles



In Durban, labor markets do not appear to be as fragmented with most people travelling between 30 and 40 minutes from home to work and transport modes used for commuting being more diverse. A large proportion of the population using mini busses, Bakkie and People mover as their main mode of transport (32.7 percent), while almost 30 percent use private vehicles and a still considerable 21 percent walk to work. In the urban core and peri-urban areas, private vehicles are the dominating mode of commuting with 50.3 and 42.1 percent of households respectively. Mini buses help most people in the urban periphery and rural areas commute to work, with 38 percent of households in the urban periphery (42.4 for informal areas in the urban periphery) and 50.3 percent of households in rural areas using this form of mass transport.

Despite connectivity being better than in Dar, people living in slums also face important challenges getting to work. Of those that travel by foot only, the average trip time is over 45 minutes, whereas for people living in non-slum urban peripheral areas, average walking trips times remain considerably lower and around 30 minutes. More telling is the fact that for people living in slums in Durban, for almost 43 percent of residential moves "better opportunities" cited as the main reason for moving. For other residents in Durban, better opportunities are the main reason for 24 percent of the moves and live events for a much larger 30 percent.

Secondly, although households often change jobs when they move to a new house, very few say that a new job was the motivation for the move. Household mobility patterns provide further insight into the localized nature of labor markets.[32] Although less than 10 percent of households said that they moved house within Dar es Salaam for 'a new job or school', there is reason to believe that households frequently change jobs when they move house. The MLSC survey collected data on the broad category of employment that household heads held in each house that

[32] There is a theoretical case to be made that – despite low commuting distances – firms and households may be able to take advantage of large pools of work and workers if workers know about work opportunities in other areas of the city and are able to easily move house for new jobs. The MLSC survey suggests that this is not happening in Dar es Salaam.

they lived in.[33] In fifty five percent of moves, the household head's job category changed. Among those moves for which households also provided a stated motivation for moving house, it is notable that the highest proportion of job category changes occurred among those who said they moved in response to some form of stress (for more in depth discussion of stated reasons for moving, see chapter 4).[34] Similarly, 40% of the moves within Durban where paired with a change in job category. Among households that changed activity when moving house and provided a stated motivation, 19.2% moved to search for better job opportunities or to be closer to their place of work or better schooling.

[33] Households were asked what their primary occupation was in each house they lived in, selecting from a list of "wage earner: self employed; family farm/livestock/fishing; student; not working; and other".

[34] In 32 percent of 'stress' moves, the household head's job changed. The broad categories of job types used means that it is likely that the true degree of job changes that occurred with residential mobility is understated. Indeed, it is notable, that in only 30 percent of jobs where the household head the move was motivated by a new job or school did the job category change.

Access to Quality Housing

Key findings for Dar es Salaam

- The majority of households live in low quality rental housing.
- There is a broad trend of households moving 'up the property ladder' and into the periphery as they age. There is nonetheless a large degree of stress-induced mobility within the city, particularly among renters in the city center.
- One third of homeowners rent out rooms in the property that they live in.

Key findings for Durban

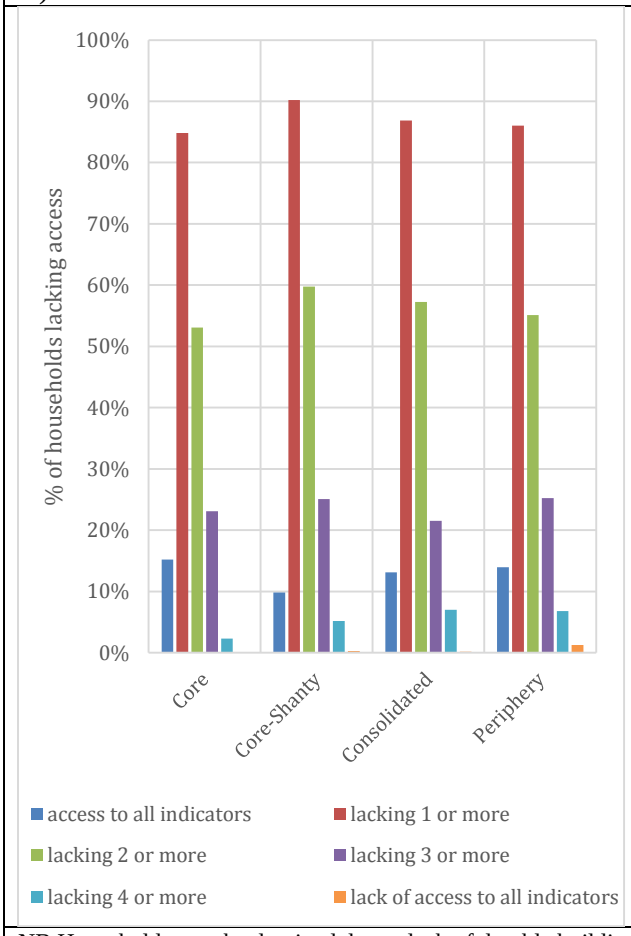
- Low quality housing is still a challenge in Durban
- Most recent migrants in Durban rent but overalownership rates are considerably higher than in Dar es Salaam

Most households in Dar es Salaam and Durban live in housing that is lacking in at least one dimension of durable materials. Renters make up 57 percent of the Dar es Salaam. Just over one third (37 percent) of houses are owner occupied. As highlighted in **Figure 18**, renters are more common in the center of the city than they are in the consolidated and periphery areas, and the highest proportion of renters is found in shanty areas. In Durban instead, ownership is much stronger with ownership rates falling below 60 percent only for the core of the city and slum areas where rental rates are around 40 percent. The majority of housing in both cities lacks some form of durable material. Non-durable roofing material is particularly common in Dar es Salaam: 95 percent of dwellings have metal sheet roofs.[35] The proportion of housing that lacks at least one durable material for floor, roof, or walls is very consistent between rented and owner-occupied housing, as well as between formal and informal housing (96 percent in all cases).

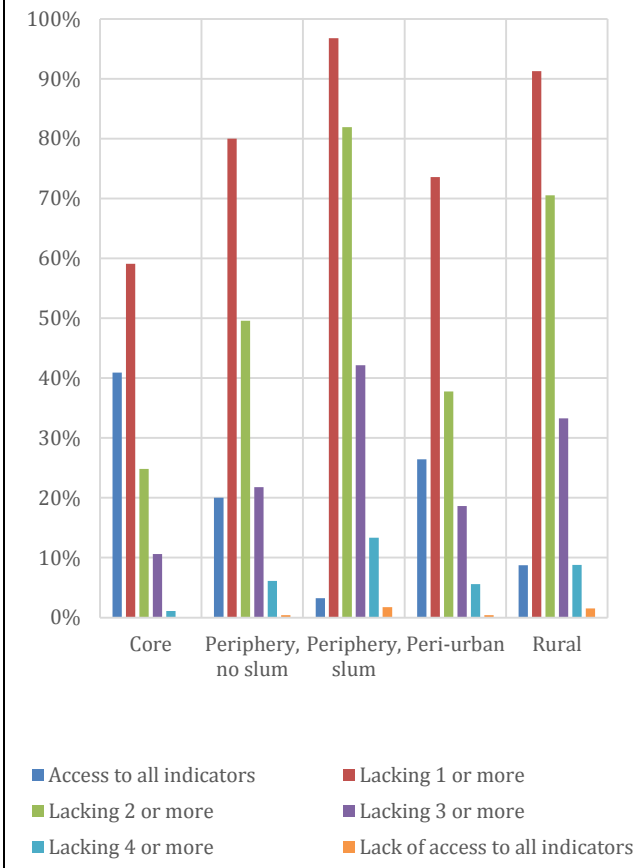
[35] Durable materials are concrete, cement, or brick. The majority of houses in Dar have durable walls (98 percent) and floor (97 percent); only three percent however, have a durable roof (95 percent is metal sheets).

Figure 18: A large proportion of the households live in areas with at least one slum-conditions indicators

A) Dar es Salaam



B) Durban

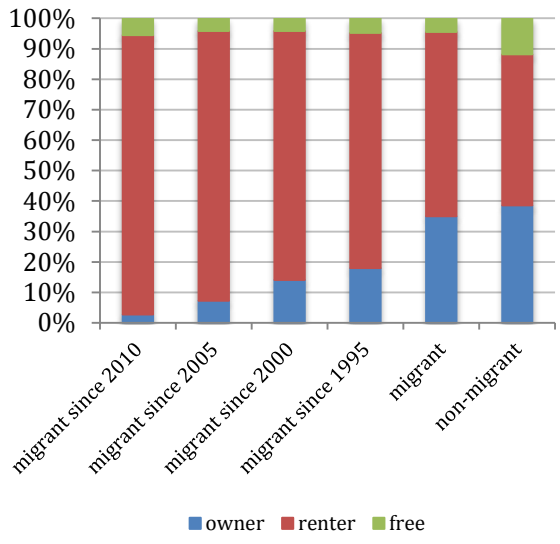


NB Households may be deprived due to lack of durable building materials, improved water and sanitation, overcrowding, or insecure tenure. Durable materials are defined as walls, floor, and roof made of concrete or cement. Improved water and sanitation are determined in keeping with the UNICEF Joint Monitoring Program of water supply and sanitation definition. Overcrowding is calculated on the basis of three or more members per room. Tenure is considered informal if households do not report having any form of formal document of ownership.

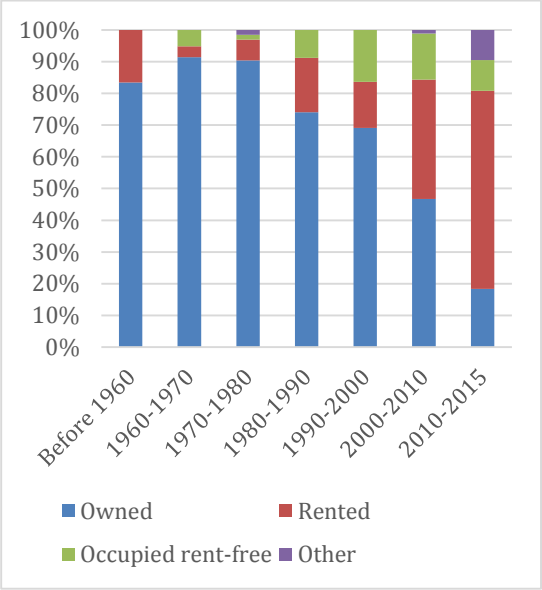
Renting is also more common among younger households and more recent migrants. Tenure does not vary greatly by gender of household head, and a similar proportion of migrants own their houses as non-migrants. It is nonetheless notable that renting is by and large the most important form of tenure for recent migrants. As **Figure 19** shows, around 90 percent of migrants that migrated to Dar es Salaam in the last ten years are renters; 62 percent of those who migrated into Durban in the last 5 years rent. The proportion of renters to owners changes over time in both cities, suggesting that migrant households may be moving 'up the property ladder' to home ownership over time.

Figure 19: New migrants are particularly likely to be renters, but over time migrant tenure choices normalize in line with general trends for the city

Dar es Salaam



Durban

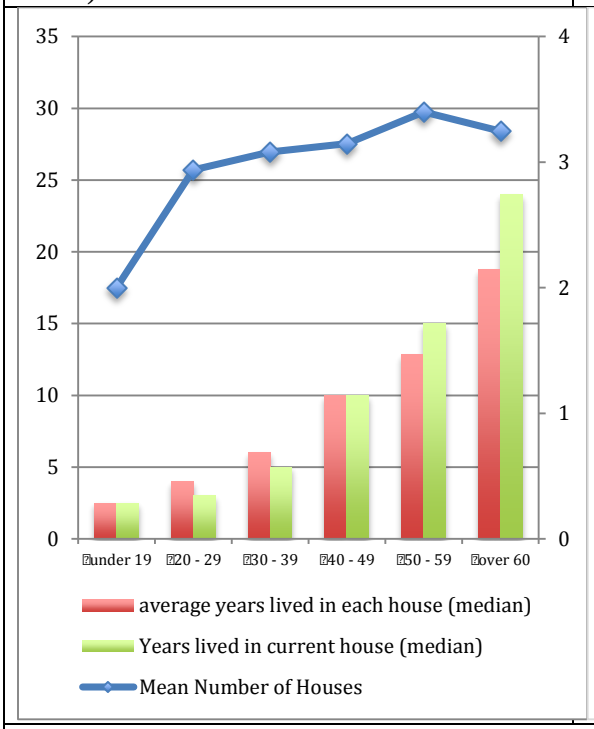


Residential mobility within the Dar es Salaam may be increasing. On average, households in the survey recorded having lived in three houses: for migrants, these records are of two houses in Dar es Salaam and one house prior to migration;[36] for households born in Dar es Salaam, the residential history covers their entire lifetime. As such, both migrants and non-migrants move into houses in Dar es Salaam twice: after their first move (either into the city or out of their parents house within the city) they will generally make one further move within the city over their lifetime. As the figure below indicates, however, the residential mobility rates do not alter greatly with the age of the household head, although the length of time spent in their current house increases substantially as households get older. This trend indicates that households tend to move more within the city when they are younger. .

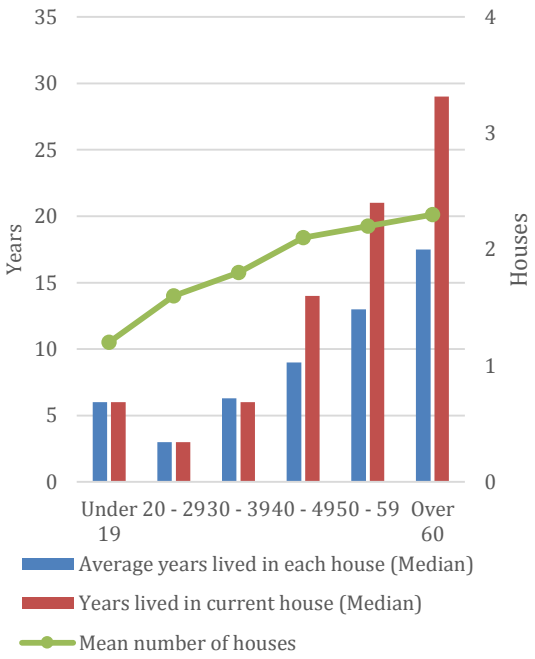
[36] The survey did not ask respondents to provide additional information on additional moves that may have taken place prior to migrating to the city, beyond providing information on the house immediately prior to moving.

Figure 20: Household Mobility in Dar es Salaam and Durban

A) Dar es Salaam[37]



B) Durban

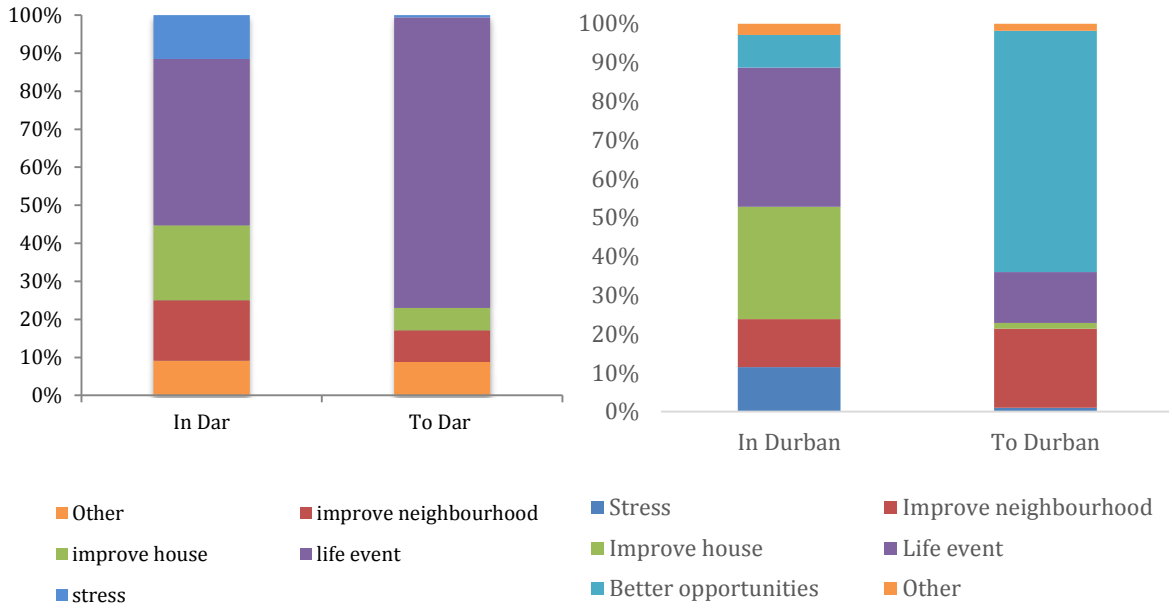


The number of years lived in each house is shown in the primary Y access, and the number of moves is in the secondary Y axis.

Residential movement within the city is motivated by life events and the desire to improve living conditions but moves to Durban are clearly motivated by opportunity. When asked what factors influenced their decision to move within Dar es Salaam, the majority of respondents in the MLSC survey highlighted 'life-events': being old enough to live on their own, a change in family circumstance (such as birth or marriage), getting a new job, or attending a new school. Over 70 percent of moves to Dar are motivated by life events compared to 44 percent of the moves within Dar. For Durban instead, over 60 percent of moves are in search of better opportunities while moves within the city are mostly due to life events (35.9 percent) and the search for a better home (29 percent). The desire to improve house or neighbourhood conditions is a much less cited reason for moving in Dar es Salaam (7.4 percent).

Figure 21: household's accounts of past moves confirm the importance of life events and the desire to improve living conditions in residential mobility

[37] On average households of all ages record having lived in three houses. The length of time that they lived in each house increases considerably with the age of the household head. This graph should be interpreted with caution: the impact of recall biases, as well as whether households that move frequently are more likely to leave the city permanently after a few years are just two factors that need to be understood better.



Improve neighbourhood collates responses indicating that moves were motivated by desire for safer or better neighbourhood conditions, or to be closer to work/school/transport/family/kin. Improve house refers to answers that indicated that the move was motivated the desire to improve the house itself: answers relating to the size and quality of house, the services available, or the opportunity to purchase land. Life events are responses that indicated the head moved because they were 'old enough to live on their own', or a change in family circumstances. In Dar es Salaam, this category also included 'a new job or school' responses (less than 10% of responses). The category of stress refers to eviction, natural disasters, or the need to move to either a cheaper or smaller house. In the Durban survey an additional category exists, 'better opportunities' which is to find work/ better opportunities to find work.

NB: these graphs are representative of the surveys, not the city (ie they are not weighted).

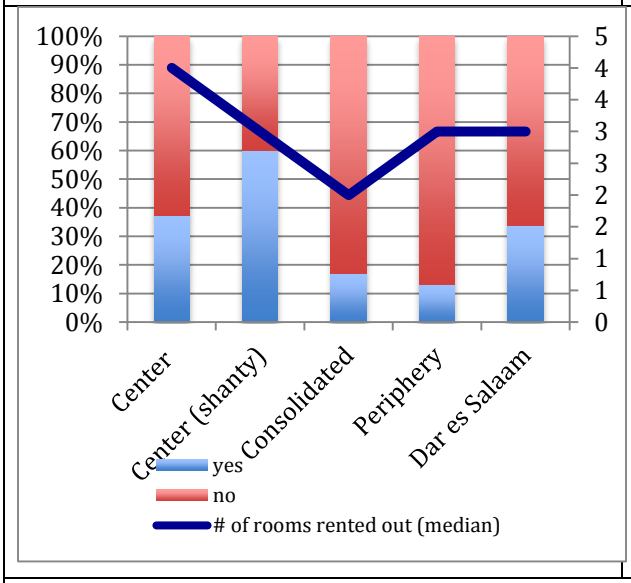
In addition to this, households living in the center of the Dar es Salaam are more likely to say they would like to move, and to indicate that if they were to move they would move neighbourhood. Among those households that said they were either possibly or likely to move house, 72% of those who, said they would move to a different neighbourhood in Dar es Salaam. Likelihood of moving was greatest in the center of the city (about one third of households), and most likely to say they would move neighbourhood (75%). Those in the periphery were less likely to say they would move (21%), and more likely to say they would stay in the same neighbourhood if they did (30%). Instead in Durban, most people say they are not likely to move; the percentage is lower but still very high in slum areas with 46.7 percent of households saying they would "definitely not move". The reluctance to move is highest in rural areas at over 87 percent.

Homeowners in Dar es Salaam rent out rooms in their property. This is not the case in Durban where less than 5 percent rent part of their property. In Dar, one third of all homeowners rent out part of their home for additional income. Households in central

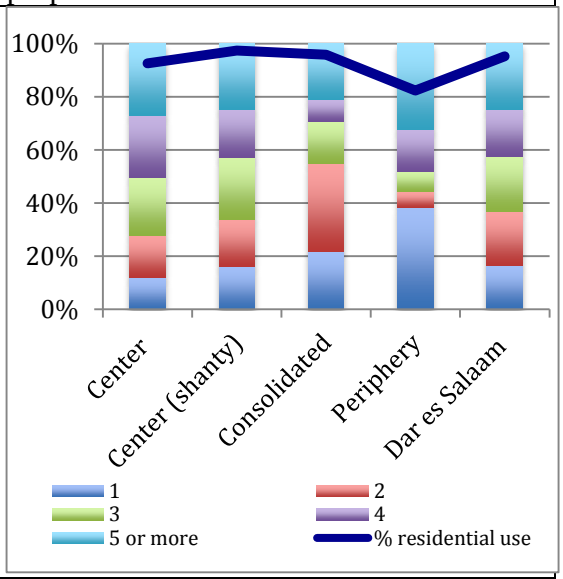
shanty areas are particularly likely to rent out part of the property they live in: fifteen percent of all households, and sixty percent of homeowners in central shanty areas rent out rooms (see Figure 22, A). The vast majority (95 percent) of these rooms are rented out for residential rather than commercial purposes, and the number of rooms rented can be large: as indicated in Figure 22 (B) nearly a quarter of households rent out more than five rooms in their property.

Figure 22: the majority of home-owners rent out part of their property for residential purposes

A) Percentage of homeowners that rent out part of their property, and average number of rooms rented out



B) Number of rooms rented out, and percentage rented for residential purposes



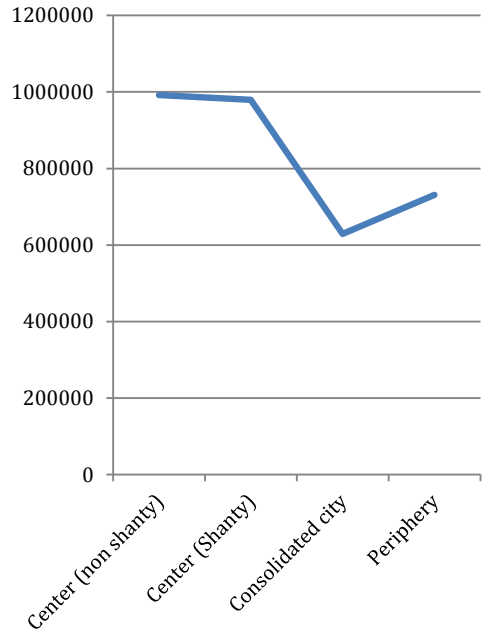
Note, 14 percent of households overall rent out part of their home. However, only 1 percent of renters do so, compared with 37 percent of homeowners and twenty percent of households that get their housing for free.

House prices follow a broadly mono-centric pattern. As discussed in the introductory chapter to this report, the MLSC survey collected information on house values. Renters were asked how much they pay per month in rent, and home-owners were asked how much a friend who wanted to buy a similar house in the same neighbourhood would have to pay. While during the pilots the teams tried to collect information on property area, the survey team found that individuals were not able to report area in an accurate way. Areas were calculated following two different procedures. Given that for Dar es Salaam information on building footprints was available, areas were assigned using such information and following simple assumptions that also used the information at the listing level. For Durban, parcel area was available from official sources and hence the value is attributed to the parcel. This may suggest that values are less accurate for rural areas,

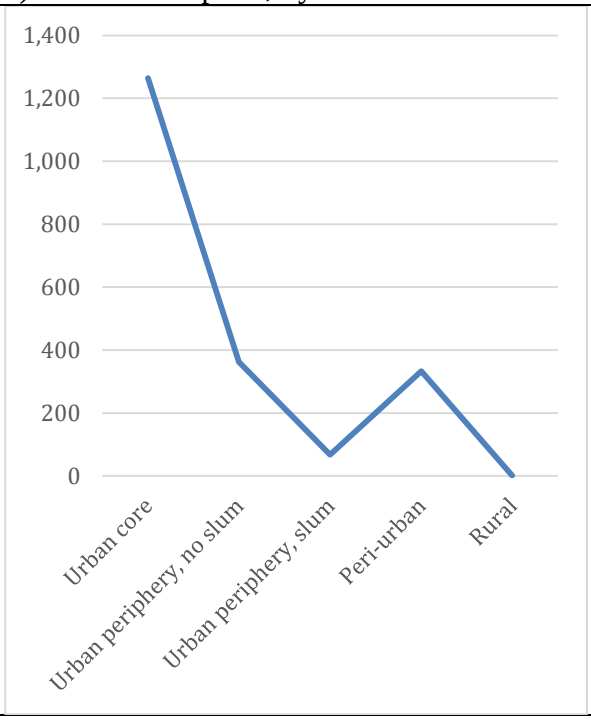
where parcels are larger and it is harder to identify what the individual areas are. The graphs below suggest that additional work is needed on these two area based variables..

Figure 23: Reported House Values Increase with Proximity to the City Center (Mean house price per m2)

A) Mean house price per m2



B) Mean house price, by stratum



References

Beegle, K., De Weerdt, J., Friedman, J, and Gisbon, J. 2010. Methods of Household Consumption Measurement through Surveys Experimental Results from Tanzania. *World Bank Policy Research Working Paper 5501*, (December).

Collier, P., 2013. Building African cities that work. Centre for Development and Enterprise, Working Paper, 27(June), pp.0–18.

Dowall, D. E. 1995. The Land Market Assessment: A New Tool for Urban Management. Washington, D.C.: The World Bank.

Farfan, G. M. Genoni, and R. Vakis. 2015. "You Are What (and Where) You Eat: Capturing Food Away from Home in Welfare Measures." Policy Research Working Paper 7257, World Bank, Washington, DC

Kombe, W. 2010. Land Conflicts In Dar Es Salaam: Who Gains? Who Loses? Working Paper no. 82 - Cities and Fragile States – Crisis States Working Papers Series 2, ISSN 1749-1800 (online)

Malpezzi, S. 2000. "Housing". *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 Years of the Living Standards Measurement Study, Volume 1*. Washington, DC: World Bank. P293-314

Parsa, Ali, et al. "Impact of formalisation of property rights in informal settlements: Evidence from Dar es Salaam city." *Land Use Policy* 28.4 (2011): 695-705.

Smith, L.C., Dupriez, O. & Troubat, N., 2014. Assessment of the Reliability and Relevance of the Food Data Collected in National Household Consumption and Expenditure Surveys,

UN Habitat, 2014. The State of African Cities. Nairobi, Kenya. ISBN 9789211333978

World Bank, 2015. A Measured Approach to Ending Poverty and Boosting Shared Prosperity: Concepts, Data, and the Twin Goals, Washington, D.C. doi:10.1596/978-1-4648-0361-1. License: Creative Commons Attribution CC BY 3.0 IGO

Annex 1: Sampling Strata, Selection Probabilities, and Sampling Weights

Dar es Salaam Survey

The sample frame for the first stage was the list of 15,262 EAs generated by the Tanzania National Bureau of Statistics (NBS) for the 2012 Census in Dar es Salaam. The EAs were sorted into four strata depending on their distance to the city center and on their likelihood of belonging to a shantytown.

Table 5: Dar es Salaam Sampling strata

Stratum	2012 Census			DESLSS 2015		
	No. of EAs	No. of HHs	Population	No. of Eas	Sample Size (HHs)	Margin of Error
1 Central core, non-shanty	4,943	329,477	1,317,095	56	672	5.3%
2 Central core, Shanty	6,073	422,123	1,586,762	64	768	5.0%
3 Middle Ring	3,692	275,620	1,146,747	51	612	5.6%
4 Periphery	554	41,388	166,360	29	348	7.4%
Total	**15,262**	**1,068,608**	**4,216,964**	**200**	**2,400**	**2.8%**

NB the last column of this table shows the maximum expected margins of error for the estimation of a household-level prevalence P (such as the percent of households with children, the percent of households reporting illnesses, etc.), at the 95% confidence level. These are given by $ME = 1.96\ [\ Deff\ P\ (1\text{-}P)\ /\ n\ ]^{0.5}$, where n is the sample size and $Deff$ is the *design effect*, basically due to the tendency of neighboring households to behave similarly in regards the indicator being observed. The column was computed for $Deff = 2$ (a value found in practice for many indicators of interest) and $P = 0.5$ (for which ME is maximum).

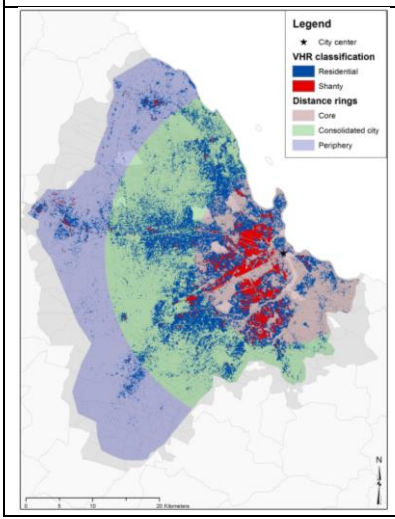
EAs in the center of the city were classified according to a 'shanty-ness' indicator derived from satellite imagery classification. Image pixels are first individually classified as "shanty" or "non shanty". Then the Enumeration Areas (EAS) were classified as "shanty" or "non shanty" depending on the predominant category of its pixels, based on an indicator of 'shanty-ness' I, calculated as follows:

$$I = \frac{area\ of\ \text{"shanty"}\ pixels}{area\ (shanty + non - shanty\ pixels)}$$

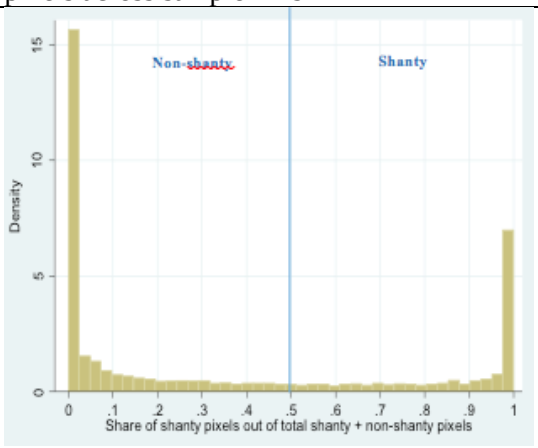
As shown in **Figure 24**, in most EAs one of the two pixel categories predominates (shanty vs. non-shanty), while few are mixed. Based on the distribution of the 'shanty-ness' indicator I across EAs, a threshold of $I = 0.5$ was chosen to classify EAs as shanty or non-shanty. EAs where $I \geq 0.5$ were classified as predominantly 'shanty', while $I < 0.5$ indicates the EA is predominantly 'non-shanty'.

Figure 24: Dar es Salaam – stratification by distance to city center and satellite imagery classification

A) Stratification by distance to city center



B) Distribution of 'shanty' and 'non-shanty' pixels across sample EAs



Dar es Salaam survey selection probabilities. In the Dar es Salaam sample design, the probability p_{hij} of selecting household *hij* in Enumeration Area *hi* of stratum *h* is given by

$$phij \quad = \frac{k_h n_{hi}}{\sum_i n_{hi}} \times \frac{m_{hi}}{n'_{hi}}$$

where the two fractions on the right hand side respectively represent the probability of selecting the EA in the first stage, and the conditional probability of selecting the household in the second stage, and

k_h is the number of EAs selected in the stratum (the fourth column in Figure 1),

n_{hi} is the number of households in the EA, as per the sample frame (the column headed 'private_ho' in Figure 2),

m_{hi} is the number of households selected in the EA (normatively always 15), and

n'_{hi} is the number of households in the EA, as per the household listing operation.

In the Dar es Salaam sample design, the probability p_{hij} of selecting household *hij* in Enumeration Area *hi* of stratum *h* is given by

$$p_{hij} = \frac{k_h n_{hi}}{\sum_i n_{hi}} \times \frac{m_{hi}}{n'_{hi}}$$

where the two fractions on the right hand side respectively represent the probability of selecting the EA in the first stage, and the conditional probability of selecting the household in the second stage, and

k_h is the number of EAs selected in the stratum (the fourth column in Figure 1),

n_{hi} is the number of households in the EA, as per the sample frame (the column headed 'private_ho' in Figure 2),

m_{hi} is the number of households selected in the EA (normatively always 15), and

n'_{hi} is the number of households in the EA, as per the household listing operation.

Durban Survey

The sample frame for the first stage is the list of 4,789 EAs (or Small Areas-SSA) in the Ethekhwini municipality published by Statistics South Africa in the 2011 National Population Census. The EAs were sorted into five strata depending on their distance to the city center and on their characterization of unplanned vs. planned settlements. The spatial distribution of unplanned settlements is based on ground information and satellite imagery as reported in the 2006-2011 Informal Settlements Change Assessment provided by the South African National Space Agency (SANSA). The distance stratification was done with official data provided by the Ethekhwini Municipal Demarcation Board as used in the 2012/13 Municipal Services and Living Conditions Survey. Large EAs selected in the first stage were segmented into area units of roughly the same size, with one of the segments selected randomly with equal probability. All dwellings in the EA or segment were listed individually, with 12 dwellings selected using systematic equal probability sampling. A total of 12 households per EA were selected with systematic equal probability sampling in the last stage.

Table 6: Durban Sampling strata

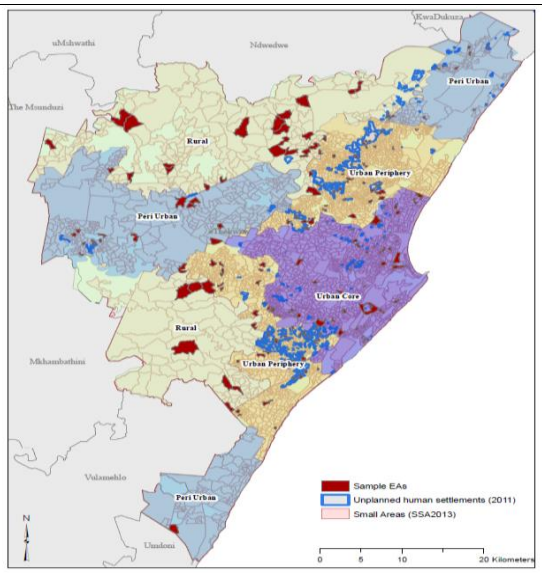
		2011 Census			Durban LSMS 2015		
Stratum		**No. of EAs**	**No. of HHs**	**Population**	**No. of Eas**	**Sample Size (HHs)**	**Margin of Error**
1	Peri Urban	641	108,867	397,635	32	384	7.1%
2	Rural	340	59,301	291,867	28	336	7.6%
3	Urban Core	1,366	282,126	879,714	46	552	5.9%
4	Urban Periphery, No Slum	2,088	415,161	1,603,143	66	792	4.9%
5	Urban Periphery, Slum	354	90,276	269,694	28	336	7.6%
	Total	**4,789**	**955,731**	**3,442,053**	**200**	**2,400**	**2.8%**

NB the last column of this table shows the maximum expected margins of error for the estimation of a household-level prevalence P (such as the percentage of households with children, the percent of households reporting illnesses, etc.) at the 95% confidence level. These are given by $ME = 1.96 [Deff\ P\ (1\text{-}P) / n]^{0.5}$, where n is the sample size and *Deff* is

the *design effect*, basically due to the tendency of neighboring households to behave similarly in regards the indicator being observed. The column was computed for *Deff* = 2 (a value found in practice for many indicators of interest) and *P* = 0.5 (for which *ME* is maximum).

Figure 25: Durban – stratification by distance to city center and satellite imagery classification

A) Geographical Distribution of Selected EAs



B) Example of segementation procedure for large EAs

Source: South African National Statistical Agency, Durban Municipality, World Bank

Durban survey weights.

Given the sampling design in Durban, the probability p_{hizsj} of selecting household *hij* in Enumeration Area *hi* of stratum h is given by:

$$p_{hizsj} = \frac{k_h n_{hi}}{\sum_i n_{hi}} \times \left(\frac{1}{T_{hi}} \times \frac{1}{G_{hi}} \right) \times \frac{15}{m'_{hisj}} \times \frac{12}{n'_{hisj}}$$

where the five fractions on the right hand side respectively represent the probability of selecting the EA in the first stage, and the conditional probabilities of selecting the household in the second stage (segmentation of larger EAs), third stage (selection of 15 buildings), and fourth stage (selection of 12 households).

k_h is the number of EAs selected in the stratum (the last column in the table above),

n_{hi} is the number of households in the EA, as per the sample frame (the column headed 'HHs' in the table above).

T_{hi} is the total number of segments in the EA, as per the segmentation procedure (the last column in the table above).

G_{hi} is the total number of sub-segments in the EAs that required additional segmentation.

m'_{hizsj} is the number of dwellings in the EA or segment, as per the dwelling listing operation.

n'_{hizsi} is the number of households in the EA or segment, as per the household listing operation.

To deliver unbiased estimates from the sample, the data from each household *hij* should be affected by a sampling weight (or raising factor) w_{hij}, equal to the inverse of its selection probability (i.e. $w_{hij} = p_{hij}^{-1}$). If n_{hi} and n'_{hi} were always the same in all EAs, the formula would simplify to a constant, and the sample would be self-weighted within each stratum. In practice, n_{hi} and n'_{hi} will seldom be identical, but often similar, meaning that the sample will not be exactly self-weighted, but approximately so.

There were a number of challenges in setting the weights for Durban. Firstly, poor documentation of the number of households in each building posed problems for estimation of the probability of selection in the fourth stage – the records had to be carefully cleaned. Secondly, the number of usable households was on average 7.3 per EA, and in four EAs, none of the selected households could be interviewed[38]. We assume that non-respondents are similar to respondents within an EA – this is a debatable but common place assumption. It is equivalent to declaring that m_{hi} is the number of usable households in the EA (rather than the number of households targeted for contact in the EA). In cases where non-response was very high, EAs were grouped:[39] EAs with four or fewer usable households were joined with a neighboring EA within the same stratum and in which there was a similar predominant ethnicity. In total, 161 EA groups were created.

The weights were further adjusted to reflect the fact that non-response rates are not random. Specifically, the survey reflects a trend in survey in Durban that non-response rates are particularly high among white households. We therefore use 'benchmarking' to weight the sample post-stratification: ethnicity data from the census sampling frame was used to create six post-strata. The table below gives the sample size (number of households) and the sum of weights in the resulting six post-strata.

[38] 'Usable households' are defined as those where at least one of the three respondents (main, household head, and random) could be interviewed, at least partially. This differs from the nominal sample of 12 per EA.

[39] The weight will be too large if the number of usable households is small: m_{ih} is too small, P_{hij} becomes too small, and W_{hij} becomes too large.

Post-stratum	N(weight~s)	sum(weight~s)
Peri-urban non-white	163	81903
Rural non-white	239	55275
Urban core non-white	286	191169
Urb periph no slum non-white	471	367419
Urb periph slum non-white	245	85446
White	57	174522
Total	1,461	955734