APPLYING UTILITY-BASED ACCESSIBILITY MEASURES IN NAIROBI, KENYA: COMPARING ACCESSIBILITY IMPACTS OF TRANSPORTATION AND LAND USE IMPROVEMENT WITH LOGSUMS

Louis Alcorn
Graduate Research Assistant
The University of Texas at Austin
loualcorn@utexas.edu

Kara M. Kockelman
(Corresponding Author)
Dewitt Greer Professor in Engineering
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
kkockelm@mail.utexas.edu
Tel: 512-471-0210

Word Count: 6346 words + 4 tables (250 words) = 7346 word equivalents (+ 10 figures)

Under review for publication in *Journal of the Transportation Research Forum*

ABSTRACT

Accessibility is an important consideration in improving networks, modes, and/or land use patterns, in any transport setting. In countries like Kenya, many travelers do not have access to cars, and transit service is key to access, economic development, and community welfare. This paper uses monetized differences in nested logsums to value changes in Nairobi's residents across destination zones and three travel modes. Results show how transit versus land use improvement policies affect residents of formal and informal (slum-area) housing differently. Welfare changes are compared when assuming independent error terms in before-after cases and allowing for perfect correlation in error terms (to recognize the same persons and modes are present in both settings). Under both access-improvement scenarios, residents of formal housing tend to benefit more than those in Nairobi's informally developed areas (largely slum settlements).

BACKGROUND

The ability to reach opportunities is widely referred to as accessibility, which fundamentally differs from more traditional transportation performance measures (like speed and flow volumes) that emphasize mobility, or ease of movement (Litman, 2003). Accessibility can be monetized, and it is embodied in real estate values (which are readily observable [Srour et al. 2002]) as well as people's willingness-to-pay for changes in access (often using Hicksian or income-compensated demand functions, which are not directly observable [Kockelman et al., 2013; Varian, 1992]).

Network improvements, new modes, and land use changes all impact access. Using accessibility measures to compare such investments or policy measures, rather than looking solely at travel times or speeds, is important and valuable. This perspective is particularly useful in the context of transportation systems in the Global South where most residents rely on walking, non-motorized transport, and public transit to meet their daily needs.

The academic literature on accessibility began with Hansen's (1959, p. 74) gravitational-force-based measure as "the *potential* of opportunities for interaction". Over time, various other metrics have been defined and used to predict travel choices (see, e.g., Kockelman [1997] and Cervero and Kockelman [1997]). Researchers like Handy and Niemeier (1997) and Bhat et al. (2000) have identified cumulative opportunities, gravity-based, utility-based, and space-time measures as separate access metric categories. Published applications of accessibility in sub-Saharan Africa are rare, but they do exist. For example, Campbell et al. (2019) employed both cumulative opportunities and gravity-based measures to assess accessibility to health care facilities for walking, matatu minibus, and driving modes. This study found the highest levels of access to health care facilities could be found proximate to the CBD and poor areas actually have comparatively better walking access to health facilities than wealthier ones. Medium-low income (~$37USD/capita/month) areas located in formal tenement apartment building districts achieved the highest overall accessibility ratings while informal slum settlements achieved the highest access by walking mode metrics (Campbell et al., 2019). The World Bank (2016) simulated cumulative-jobs-access measures with Monte Carlo draws and a hill-climbing optimization procedure. While Nairobi's current layout outperformed all of their 10,000 random, counterfactual scenarios, they estimated alternative, coordinated land-use patterns could increase overall accessibility by 15% for those using cars and by 100% within an hour's travel time via matatus (The World Bank, 2016). Avner and Lall (2016) also examined Nairobi's jobs-housing balance. However, no published research has yet quantified accessibility differences across Nairobi and across new land use-transportation settings for Nairobi using utility-based metrics. This paper uses monetized and nested logsum differences in access to jobs, health care, and education opportunities via three travel mode alternatives to anticipate what policies or practices may best support Nairobians in their access needs and aspirations.

Public transit services in Nairobi, the capital of Kenya, are provided almost exclusively by *matatus*, or privately-owned and operated 14- and 25-seater vans and buses. Aside from a handful of commuter rail lines, Nairobi generally lacks formal transit services (Salon and Aligula, 2012). To fill this gap, an extensive network of matatus provide the region with a semi-formal transit system. Cervero (2001, p. 1) calls this type of operation "laissez-faire transit [whereby] through the invisible hand of the marketplace, those who are willing to pay for transport services hook up with those who are willing to provide them." Impressively, 99% of Nairobi's residents indicate matatu services are available to them, and two-thirds of adult travelers use matatus every day (Salon and Aligula, 2012). While the matatu system fills a transport need not formally provided in Nairobi, this informal transit sector generates some negative externalities, like high levels of noise and air pollution, increased traffic congestion, dangerous driving behaviors, mafia-style management practices, and a general lack of accountability within the system (Cervero and Golub, 2007; Kaltheier, 2002). Additionally, the free-market nature of these entrepreneurial operators tends to lead to "cream-skimming" or concentrating frequent services during peak travel times to maximize ridership (and therefore fare revenues) while often requiring off-peak riders to wait until their bus is filled before departing the stop (Cervero and Golub, 2007). Unlike formal transit systems in democratic nations that may have a mandate to address spatial equity concerns in their service provision, matatu operators can choose to serve wherever they want. Due to informal operators' profit-maximizing motivations, residents of the informal 'slum' settlements that house the city's most vulnerable population, with arguably the strongest need for transit access, may not be seen as a priority by operators.

This study investigates the accessibility differences between those living in informal settlements and those living in conventional residential units across Nairobi, or informally housed vs formally housed residents: IHR vs FHR. Over 1.8 million of Nairobi's 4M residents live in slums (Kenya ICT Authority, 2020; The World Bank, 2016). Many of these persons live in Kibura, East Africa's second largest slum, with certain areas housing population densities over 100,000 people per square kilometer (Marras, 2012) – which is four times higher than what one sees on the island of Manhattan, New York, with its much taller (multi-level) building stock. While many slum dwellers lack resources to pay for transit fares and rely on walking to access employment and other opportunities, between 50 and 60 percent of traveling adults below the poverty line in Nairobi report use public transit regularly (Salon and Aligula, 2012). This work estimates changes in different Nairobians' consumer surplus (CS) or welfare under several land use-transportation scenarios, as described below.

METHODS

There are two basic ways to improve accessibility: 1) improve the transportation system so people can travel longer distances in less time and/or at lower cost and thereby access more opportunities of interest, and 2) add more useful land uses/destination opportunities close to where travelers are located. The scenarios modeled here include a base case (business-as-usual) vs. two distinct access improvement initiatives: 1) improving Nairobi's transit system by prioritizing road space for buses and matatus, and 2) adding jobs, schools, and healthcare facilities closer to people's homes. The first improvement scenario is labeled "TI" for transit improvement and the second "LUI" for land use improvement.

One set of welfare results (labeled Scenarios 1 & 2) assumes random-utility error terms (for mode and destination choices across Nairobi) are independent between the base case and new access settings, while with the second set of results (labeled Scenarios 3 & 4) assumes perfect correlation between unobserved components in the same mode and destination zone options after transit systems or land uses are modified. Such correlation helps better reflect the unmodeled features affecting traveler preferences, like past bad experiences onboard a bus, riding or walking with friends each day to work or school, the presence or absence of a special shop or garden in one's preferred destination zones. All access impacts, embodied in the CS changes, are computed and then compared across the two user groups identified here: the informally housed and formally housed residents (IHR versus FHR).

UTILITY-BASED ACCESS MEASURES FOR CONSUMER SURPLUS CALCULATIONS

This study relies on a destination and mode-choice nested logit model, as illustrated in Figure 1. It is similar to the one presented by Lemp and Kockelman (2011)..

Scale parameters for mode of 1.39 ($\boldsymbol{\mu}_1$) and destination of 1.20 ($\boldsymbol{\mu}_2$) represent the inverse of the nested-logit model's inclusive value coefficients.

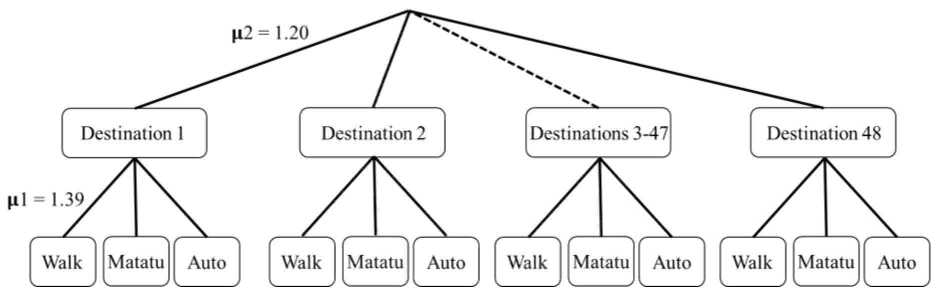


Figure 1: Destination/Mode-choice Nested Logit Model with Nesting Parameters Applied to Nairobi Case Study Includes 48 Potential Destination Zones and 3 Modes

Based on existing studies, the Income and Household Characteristics and ASC values are displayed in Table 1 and Table 2, respectively (Campbell et al., 2019; The World Bank, 2016; Walker et al. 2010).

Table 1: Income and Household Characteristics by Residence Type in Nairobi

Residence Type	Indicator	Value	Source
Informal (IHR)	Avg. Monthly Income ($USD equivalent)	$16.85	Campbell et al., 2019
	Avg. HH Size	8.0	The World Bank, 2016
	VOTT ($USD/hour)	$0.25	The World Bank, 2016
Formal (FHR)	Avg. Monthly Income ($USD equivalent)	$110.42	Campbell et al., 2019
	Avg. HH Size	3.4	The World Bank, 2016
	VOTT ($USD/hour)	$0.70	The World Bank, 2016

Table 2: Alternative Specific Constants by Mode for Developing Cities

Alternative Specific Constants (ASC's) by mode	**ASC** (Walker et al, 2010)	**ASC** fit to data for this analysis
Walk	0	2.199
Matatu Minibus	-2.22	-2.611
Auto	-1.11	-1.969

Destination attractiveness represents a measure of utility based upon a combination of access to employment, education, healthcare, and proximity to the CBD embodied in the following equation:

$$Attr_d = \beta_E * E + \beta_S * S + \beta_H * H + \beta_D * D_{CBD} + \varepsilon_i$$

where β_E, β_S, β_H, and β_D, are the coefficients on employment (jobs), schools, healthcare facilities, and distance to the CBD. Here, E represents employment density in jobs per square kilometer, S equals education density [schools/km^2], H represents the density of healthcare facilities [facilities/km^2] in zone d, and D_{CBD} represents the network travel distance between zone d's centroid and the intersection of Kenyatta Ave. and Moi Ave. at the center of the CBD (1°17'00.1"S 36°49'24.7"E). ε_i are independent and identically distributed (iid) Gumbel error terms (Walker et al., 2010; Zhao et al., 2012; Ma et al., 2015).

According to the Japan International Coordination Agency (2014), this work uses OLS regression and multi-criteria optimization with the Nairobi datasets available and described herein to arrive at the following β parameters: β_E = 0.00003, β_S = 0.07602, β_H = 0.01510, and β_D = -0.30430, with an adjusted R_2 value of 0.5109 (N = 48). The ε_i error term was omitted for the first set of results presented below ("Scenario 1: TI iid Gumbel" and "Scenario 2: LUI iid Gumbel", respectively) and then incorporated as a random value from the Gumbel distribution with a mode (μ') of 0 and scale (σ) of 1 (see Figure 2) for the second set of comparative results ("Scenario 3: TI with Gumbel" and "Scenario 4: LUI with Gumbel", respectively). The ε_i term accounts for random variation in individuals' preferences and is often assumed to be iid before versus after the policy or design change, which simplifies welfare calculations. To reflect the perfect-correlation case, 10,000 vectors of Gumbel error terms were simulated and used here, along with random utility maximization for evaluation of 10,000 individuals' CS impacts in each before-after comparison.

$$f(x) = \frac{1}{\sigma} exp(-z - exp(-z))\,, \qquad s.t.\ z = \frac{x - \mu'}{\sigma'}$$

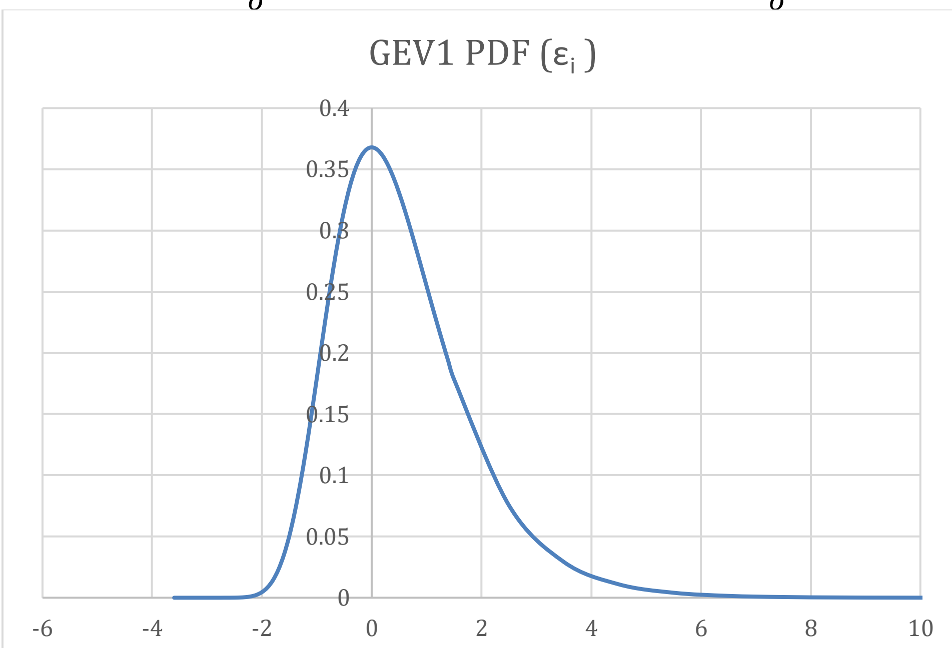


Figure 2: Generalized Extreme Value Distribution for ε_i

Finally, the following the cascade of equations is used to calculate the logsum differences embodying changes in consumer surplus for individuals in the two residential groups across different scenarios where a superscript of 0 represents the base scenario (existing conditions) and a superscript of 1 represents the scenario of interest.

Equation 2 (Value of Travel Time):

$$VOTT_i = HourlyWage_i * HHsize_i * 0.3$$

Equation 3 (Generalized Trip Cost):

$$GC_{i,dm} = VOTT_i * t_{dm} + OC_{dm} + Fare_{dm}$$

Equation 4 (Systematic Utility):

$$V_{i,dm} = [ln(Attr_d) - ln(Attr_B)] + ASC_m - GC_{i,dm} + \varepsilon_i$$

Equation 5 (Inclusive Value/Expected Maximum Utility for an Upper Level Alternative):

$$\Gamma_{i,d} = \frac{1}{\mu_1} * ln\,[exp\,(\mu_1 * \,V_{i,d,walk}) + exp\big(\mu_1 * V_{i,d,matatu}\big) + \,exp(\mu_1 * V_{i,d,drive})]$$

Equation 6 (Consumer Surplus Logsum):

$$CS_i = \frac{1}{\mu_2} ln\,(\sum_{k\epsilon D} \exp\,(\mu_2 * \Gamma_{i,k}))$$

Equation 7 (Change in Consumer Surplus):

$$\Delta CS_i = \frac{1}{\mu_2} l\,n\left(\sum_{k\epsilon D} \exp\,(\mu_2 * \Gamma^1_{i,k})\right) - l\,n\left(\sum_{k\epsilon D} \exp\,(\mu_2 * \Gamma^0_{i,k})\right)$$

where *i* denotes user group (IHR vs. FHR), *d* denotes destination of interest (zones 1-48), *m* denotes mode of interest (walk, matatu, or auto), *D* is the set of destination alternatives, *t* denotes time (in hours) and *OC* stands for out-of-pocket operating expenses (for autos only). $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_2$ represent scaling parameters for mode and destination nests, respectively and *ASC* refers to the alternative specific constants for mode alternatives. *Attr* refers to attractiveness of each destination, which is the output of a combination of the three destination opportunity types consolidated in the model: education, employment, and healthcare, as described in equation 1 above (see Gwilliam, 1997; Kockelman and Lemp, 2011).

DATA SETS

A variety of data are required to compute these measures, namely, a multimodal transportation network and information about origin and destination zones. The following section details the methods for data wrangling and geospatially processing a variety of different data sources in order to produce reliable travel time skims for walking, matatu+walk, and driving between all 48 zones as well as aggregating data about the contents of the zones themselves in terms of employment, education, and healthcare.

CONSTRUCTING THE TRANSPORTATION NETWORK

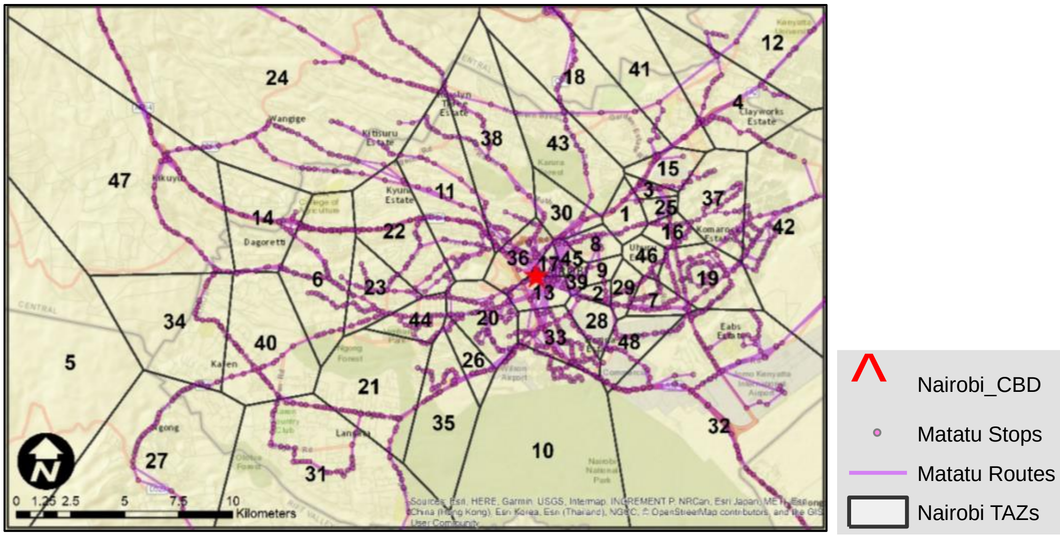


Figure 3: January 2017 Matatu Routes and Stops Network Dataset in ArcMap

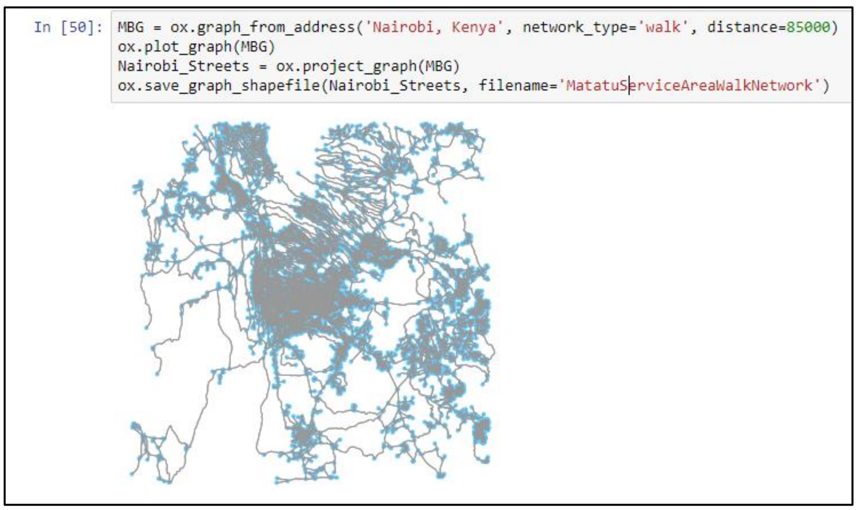


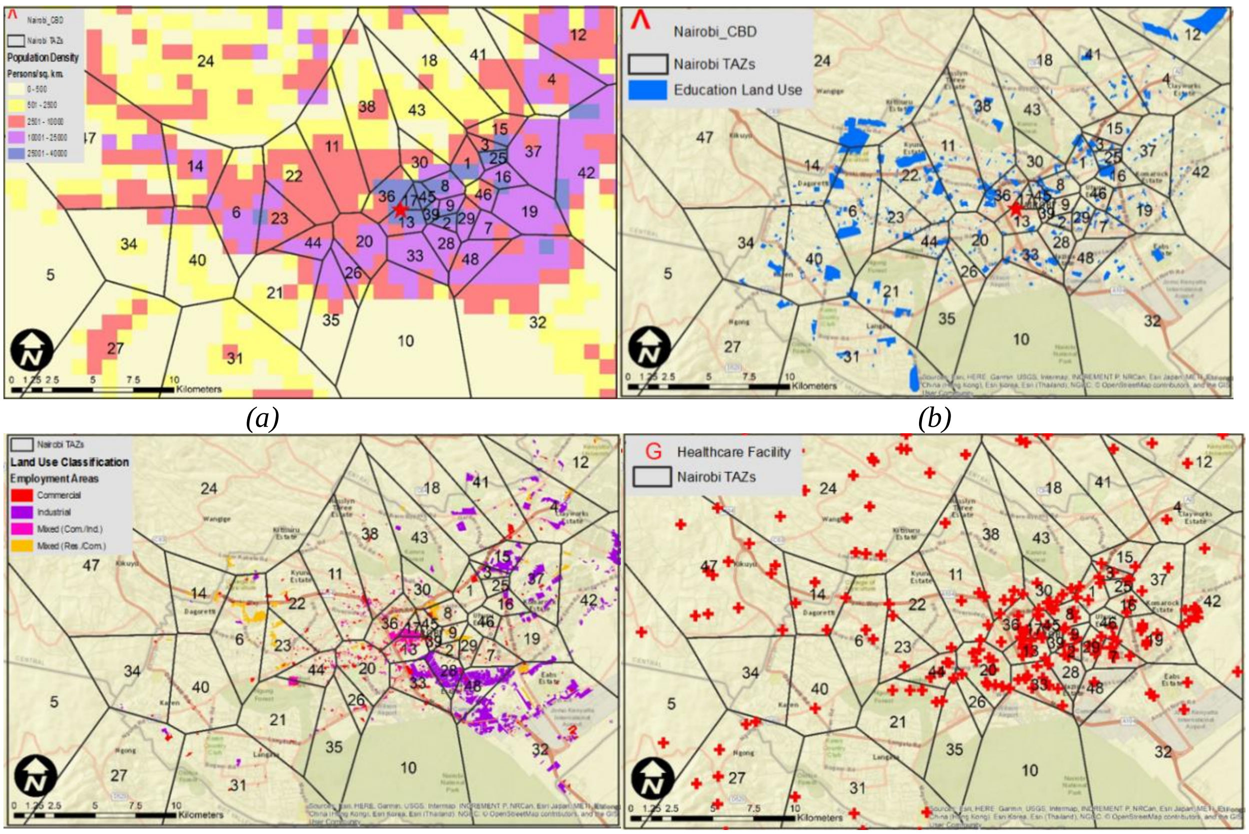
Figure 4: OpenStreetMap Pedestrian Network generated using OSMnx

The transportation network is manufactured using a combination of python code and ArcMap processing. Matatu routes and schedules come from a modified semi-formal transit GTFS feed from January 2017 produced using mobile phones by a collaborative research project involving MIT, Columbia University, University of Nairobi and a design group called Groupshot (Williams et al., 2015). ESRI's Add GTFS to a Network Dataset tool within ArcMap generates a transit network for analysis purposes (see Figure 3). Since geospatial data about walking paths and sidewalks is not available in the study area, the walking network is sourced from OpenStreetMap (see Figure 4).

ZONE CREATION: ORIGINS & DESTINATIONS

By using Thiessen polygon technique using the Voronoi method (Brassel and Reif, 1979) in the Kenyan government's data portal (Kenya ICT Authority, 2020), total 48 transportation analysis zones (TAZs) are given in Figure 4a. School location data is sourced from a 2010 Nairobi Land Use survey, shown in Figure 4b (Edwin, 2010). Employment location data is sourced from World

Bank and JICA (Japan International Cooperation Agency (JICA) et al., 2014; The World Bank, 2016), which can be cross-referenced with the commercial and industrial land-use classifications from the land use survey data (see Figure 4c). Finally, healthcare data is sourced from ESRI Eastern Africa (2017), which published a dataset representing healthcare institutions by type and by institution (see Figure 4d).



(a) *(b)*

(c) *(d)*

Figure 4 Data sets (a: 48 TAZs in Nairobi, Kenya and Population Density by Square Kilometer; b: Nairobi Education Land Use; c: Nairobi Employment-related Land Use; d: Nairobi Healthcare Facilities)

MODELED SCENARIOS

Travel skims were calculated using ArcMap's Network Analyst functionality for interzonal travel. Intrazonal travel approximates the TAZs area as a circle and assumes intrazonal trips involve traversing the equivalent "radius" of this circle at the average system speed of the respective mode of travel. Thus, intrazonal travel times increase in proportion to the area of the TAZ. These travel skims and zonal attraction data are input into and excel- and R-based logsum calculation program, which computes changes in consumer surplus as the difference in logsum calculations between a base condition (business as usual) and various scenarios for two user types (IHR and FHR). The scenarios tested in this analysis include the following and are described in more detail individually in the *Results* section:

The first two scenarios presented have uncorrelated before-after terms (ε_i) incorporated in the utility equation and assume these terms will cancel out when aggregated over a large dataset as many logsum accessibility analyses in literature and practice do. The latter two scenarios take a critical look at this assumption (as recommended by Zhao et al. (2012)) by incorporating the random error term (ε_i) into the systematic utility equation at an individual level for every origin-destination pair as random iid pulls from the GEV1 distribution described above.

- Scenario 1 - "Transit Improvement without correlated epsilons" (TI no correlated epsilons): Transit-dedicated lanes for matatus, which results in 10% faster transit travel times and 20% slower travel times for the driving mode.
- Scenario 2 - "Land Use Improvement without correlated epsilons" (LUI no correlated epsilons): Building marketplaces (employment), schools and health centers in relatively underserved areas.
- Scenario 3 - "Transit Improvement with correlated epsilons" (TI w/ correlated epsilons): Same as "Scenario 1", but incorporating ε_i term
- Scenario 4 - "Land Use Improvement with correlated epsilons" (LUI w/ correlated epsilons): Same as "Scenario 2", but incorporating ε_i term

RESULTS

This section presents the results of the welfare analysis calculations described above. First a base scenario is computed and presented for the TI and LUI scenarios that don't have correlated epsilons (Scenarios 1 & 2). Subsequently, maps describing the change in consumer surplus (in $USD) between the base case and the respective scenarios for IHR and FHR by TAZ are displayed and discussed. Then an alternative base case incorporating the ε_i error term for the calculation of Scenarios 3 & 4 is presented followed by maps describing the change in consumer surplus between this base case and the respective modeled scenarios.

After all the data are aggregated to the TAZ level, one can pull out the relative attractiveness of each zone with respect to access to jobs, healthcare, and educational opportunities as well as a variable representing distance to the CBD. Recalling our attractiveness equation (1) from above, Table 3 depicts the attractiveness values (in utils) of each TAZ. Opportunities are normalized by area of each TAZ and presented as density per square kilometer in this calculation.

Table 3: Zone Attractiveness by TAZ

TAZ #	Zone Name	Job Density (Jobs/km2)	Health Facility Density (Building/km2)	Education Density (School/km2)	Distance to CBD (km)	Zone Attractiveness (utils)
1	Huruma	2,687	1.84	3.94	7.74	5.91
2	Bahati	14,245	2.55	4.25	3.22	7.61
3	Dandora	12,391	0.7	3.49	7.94	6.04
4	Githurai	3,664	0.1	0.69	12.82	4.12
5	Karen	41	0.04	0.23	25.90	0.01
6	Kawangware	1,526	0.16	2.54	11.64	4.56
7	Makadara	1,522	1.67	3.34	5.95	6.38
8	Eastleigh North	26,509	2.04	4.42	3.72	7.77
9	Eastleigh South	11,460	2.56	1.92	3.04	7.42
10	Embakasi	78	0	0.08	6.11	6.02
11	Highridge	711	0.34	1.57	3.95	6.81

12	Kahawa	680	0.19	0.23	17.00	2.73
13	Kamukunji	32,126	2.75	2.47	0.95	8.62
14	Kangemi	670	0.28	0.91	15.05	3.38
15	Kariobangi	18,154	0.69	0.69	9.03	5.64
16	Kariobangi South	10,096	0.69	1.38	8.25	5.73
17	Kariokor	72,870	5.41	2.32	1.07	9.64
18	Kasarani	82	0.08	0.2	9.78	4.91
19	Kayole	4,267	0.67	2.97	7.62	5.89
20	Kenyatta	3,633	1.34	2.68	4.31	6.87
21	Kibera	175	0.24	0.58	7.36	5.68
22	Kileleshwa	643	0.13	3.3	6.91	6.03
23	Kilimani	5,685	0.24	4.38	7.80	5.98

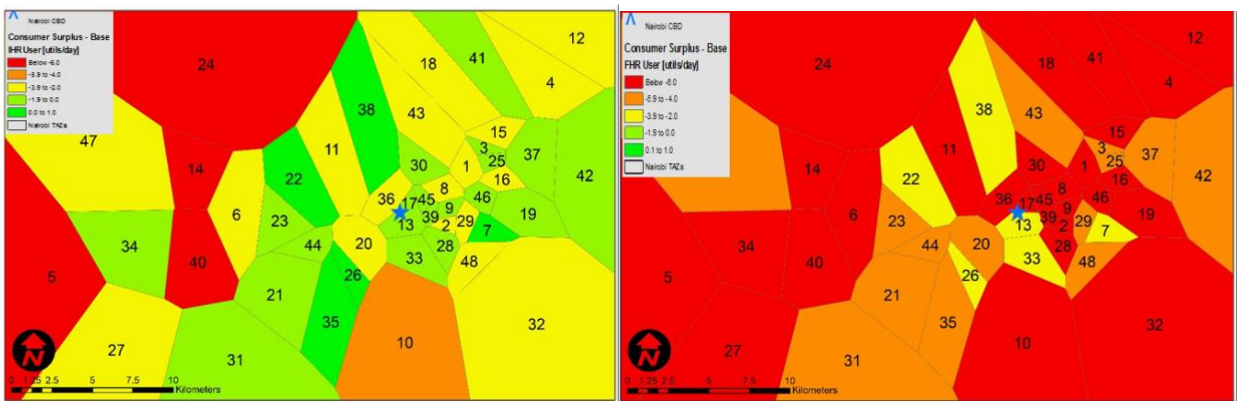
Table 3: Zone Attractiveness by TAZ continued.

TAZ #	Zone Name	Job Density (Jobs/km2)	Health Facility Density (Building/km2)	Education Density (School/km2)	Distance to CBD (km)	Zone Attractiveness (utils)
24	Kitisuru	3	0.15	0.08	14.83	3.37
25	Korogocho	2,904	1.93	9.01	7.66	6.33
26	Laini Saba	1,136	0.26	0.78	5.85	6.18
27	Langa’ta	21	0.17	0.04	23.44	0.74
28	Makongeni	35,521	0	1.07	4.07	7.61
29	Maringo	18,683	1.96	5.1	4.84	7.29
30	Mathare	1,106	1.06	1.23	4.04	6.78
31	Mugumoini	17	0.07	0.48	20.43	1.69
32	Mukuru Kwa Njenga	2,847	0.05	0.22	10.91	4.64
33	Mukuru Nyayo	12,887	1.52	3.32	3.07	7.54
34	Mutuini	110	0	0.61	16.85	2.79
35	Nairobi West	162	0.06	0.79	6.76	5.88
36	Ngara	13,447	0.53	10.61	1.41	8.60
37	Njiru	7,171	0.17	2.13	9.88	5.21
38	Parklands	559	0.21	1.38	3.57	6.91
39	Pumwani	23,661	1.01	3.03	2.41	7.98
40	Riruta	18	0.05	1.29	14.62	3.52
41	Roysambu	246	0.05	2.42	11.24	4.64
42	Ruai	2,894	0.3	0.97	13.46	3.92
43	Ruaraka	50	0.05	0.22	5.34	6.26
44	Serangombe	2,239	0.82	6.37	7.44	6.16
45	Starehe	18,844	2.43	8.11	2.16	8.34
46	Umoja	1,191	0.73	0.73	7.06	5.82
47	Uthiru-Ruthimitu	0.3	0.16	0.07	17.00	2.70
48	Viwandani & Waithaka	26,309	0.15	0.62	7.20	6.39

	TOTAL	395,942	38.6	109.26	412.70	271.06
	MEDIAN	2,463	0.27	1.38	7.4	6.0
	MEAN	8,249	0.8	2.3	8.60	5.65
	STD. DEV.	13,400	1.1	2.4	5.80	2.05

BASE CASE FOR SCENARIOS 1 AND 2 ASSUMING INDEPENDENT ERROR TERMS

Inputting pedestrian, matatu and driving skims, the relative attractiveness values, and then calibrating the aforementioned parameters using OLS regression and multi-criteria optimization with Excel's solver function produces outputs of consumer surplus (in utils) after iterating through equations 2-6. Figure 5a and Figure 5b illustrate the relative consumer surplus values for the base case (existing conditions for travel in January 2017), for IHR and FHR individuals, assuming a zero error term. Zone 1, "Huruma", is the most middle-of-the-pack zone in terms of the four independent variables shown above in Table 3 and was therefore selected as the 'base' zone for systematic utility calculation (4) and used as a benchmark for other zones.

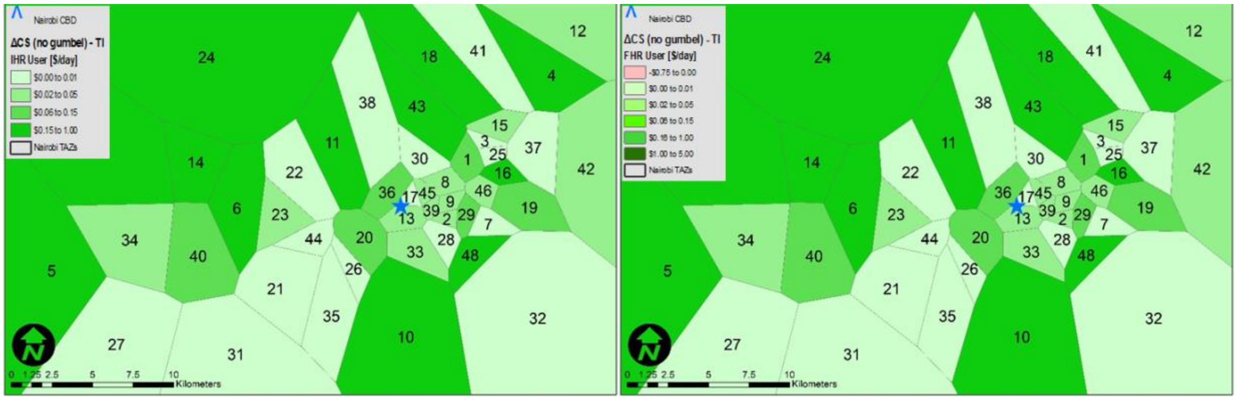


(a) *(b)*

Figure 5:– Base Case (a: Informal Slum Dweller Consumer Surplus; b: Resident of Formal Housing Consumer Surplus)

SCENARIO 1 TRANSIT IMPROVEMENT WITH IID ERROR TERMS

The first scenario aims to increase accessibility by lowering impedance on the roadway for the mode carrying the majority of travelers in Nairobi. This scenario models a matatu-prioritization program through which road space and infrastructure is devoted to accommodating faster and more efficient matatu services. Thus, this scenario models a 10% reduction in travel times for matatu users. These matatu-only lanes will also come at the expense of other roadway users (drivers) and thus auto drivers will see their travel times will increase by 20% (during peak commute times) due to the decreased capacity available for general traffic. These impedance improvements apply to both inter- and intrazonal travel. Figure 6a and Figure 6b depict the change in consumer surplus values (in $USD/day) for IHR and FHR users, respectively.



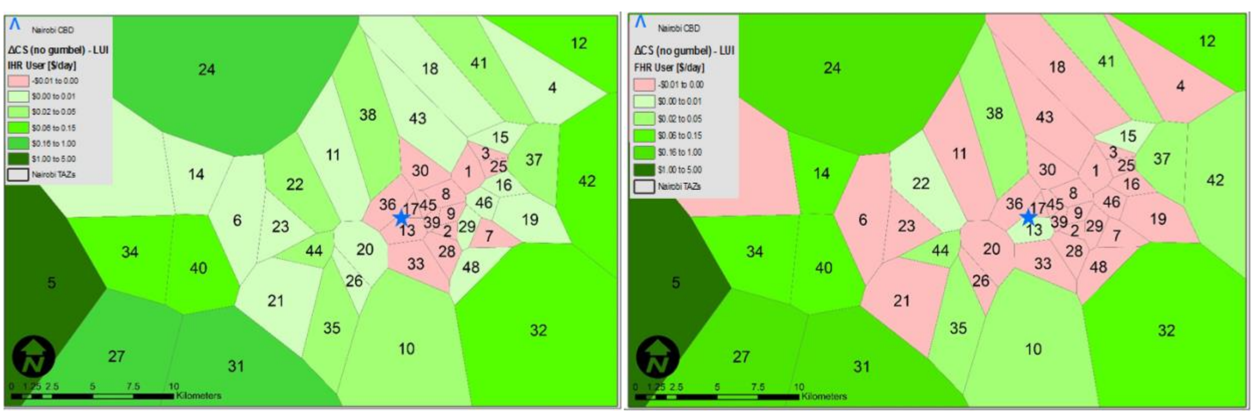
(a) *(b)*

Figure 6:– Change in CS (a: Informal Slum Dweller under Transit Improvement Scenario with No Error Terms; b: Informal Slum Dweller under Transit Improvement Scenario with No Error Terms)

As evidenced by the larger prevalence of darker green colors in Figure 6b, it appears residents of formal housing benefit more than slum dwellers from Scenario 1's transit improvements. It also appears residents of formal housing in TAZ 8, "Eastleigh North", experience a negative change in relative accessibility as a result of the addition of matatu-only lanes.

SCENARIO 2 LAND USE IMPROVEMENT WITH IID ERROR TERMS

The second scenario aims to improve accessibility by modeling land use changes in the form of constructing schools, health centers, and employment opportunities. As such, this scenario goes about trying to improve accessibility without making any transportation improvements to the network at all. This scenario seeks to test marginals by adding one standard deviation of health centers and schools across all TAZs. This amounts to the construction of 962 health centers and 2,158 new schools across Nairobi. Additionally, it is assumed the building of a new health center comes with 150 new jobs and a new school comes with 50 new jobs in TAZs where these additions were made. and depict the change in consumer surplus values (in $USD/day) for IHR and FHR users, respectively for Scenario 2.



(a) *(b)*

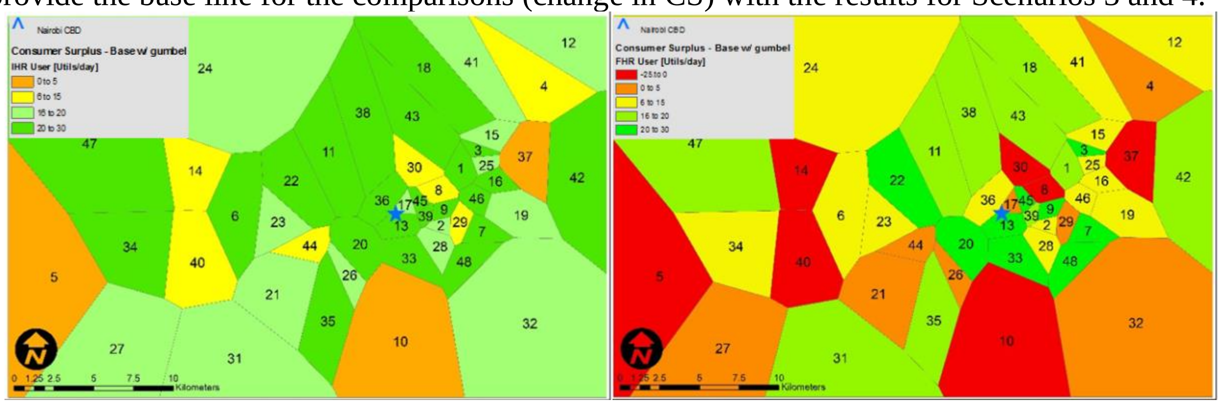
Figure 7:– Change in CS (a: Informal Slum Dweller under Land Use Improvement Scenario with No Error Terms; b: Informal Slum Dweller under Land Use Improvement Scenario with No Error Terms)

Compared to Scenario 1, Scenario 2 results display a higher prevalence of lighter greens and reds on Figure 7a and Figure 7b. This indicates overall changes in consumer surplus are more modest than the transit-improvement scenario and are even slightly negative in several zones. One important distinction to point out, though, is IHR in aggregate tend to reap more accessibility benefits than FHR in Scenario 2. Overall, it also appears zones peripheral to the CBD tend to experience more positive changes in CS as a result of the LUI scenario, a trend that does not appear so much in the TI scenario.

BASE CASE SCENARIOS 3 & 4 WITH CORRELATED ERROR TERMS

The following scenarios incorporate a randomly generated epsilon error term into the equation for indirect utility (4). This ε_i term varies by mode and by destination and is approximated as a random draw from the GEV1 distribution described in the *Methods* section above. This term helps us capture unobserved (and unmeasurable) utility associated with the scenarios described above. With any change in the urban landscape, there will be winners and losers. The Monte Carlo simulation of this error term should help to provide insight into a distribution of individual winners and losers who may or may not appear through the conventional accessibility analysis logsum calculations presented in the previous two scenarios.

Figure 8a and Figure 8b display the base CS values (in utils/day) for IHR and FHR, respectively. Again, since utils are not easily comparable across individuals, these values should not be compared with the utils presented in the base case maps for Scenarios 1 and 2. These values simply provide the base line for the comparisons (change in CS) with the results for Scenarios 3 and 4.

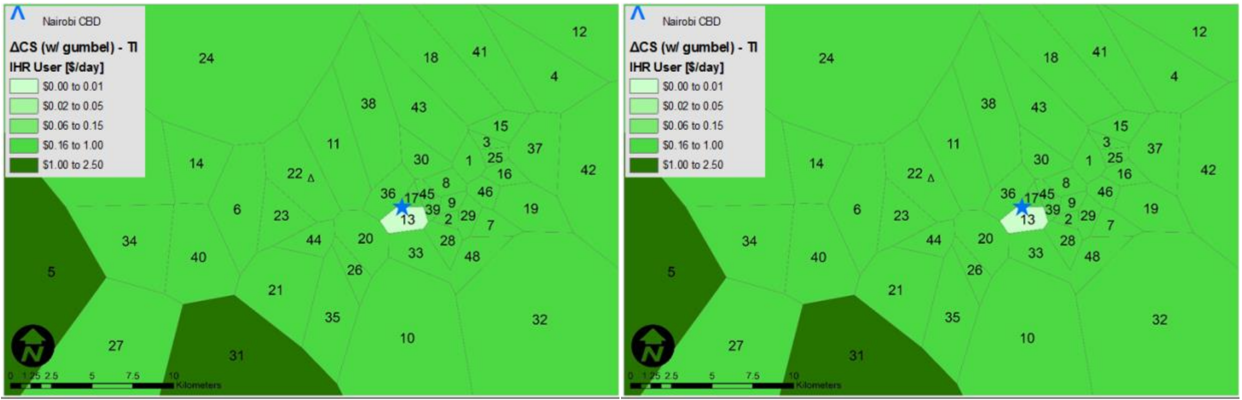


(a) *(b)*

Figure 8:– Base Case w/ Gumbel Error Term (a: Resident of Formal Housing Consumer Surplus; b: Resident of Formal Housing Consumer Surplus)

SCENARIO 3 TRANSIT IMPROVEMENT WITH CORRELATED ERROR TERMS

Scenario 3 mimics Scenario 1 presented above with decreased matatu network travel times and increased travel times for the driving mode. Figure 9a and Figure 9b depict the change in consumer surplus experienced based on the mean value of 10,000 random individual's preferences.



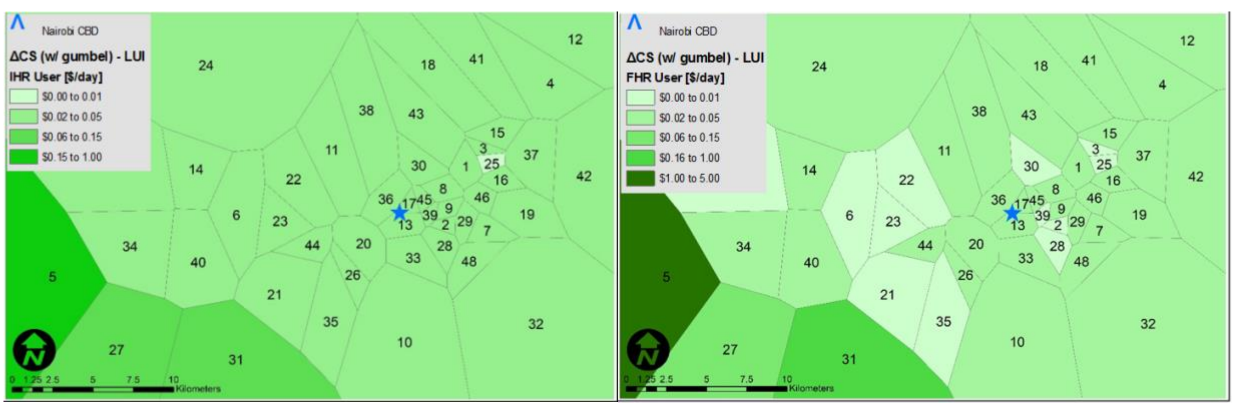
(a) *(b)*

Figure 9:– Change in CS for (a: Resident of Formal Housing under Transit Improvement Scenario with Gumbel Error Terms; b: Resident of Formal Housing under Transit Improvement Scenario with Gumbel Error Terms)

The darkest shade of green in these maps represents a category that only showed up in Zone 5 in the results for Scenarios 1 and 2, which represents an individual change in consumer surplus that exceeds $1USD/day (~equal to 4 hours of travel time for IHR and 1.4 hours of travel time for FHR). Overall, we can see the higher prevalence of darker shades of green in Figure 9a compared to Figure 9b indicates FHR reap more monetary benefit from this scenario when compared to IHR.

SCENARIO 4 LAND USE IMPROVEMENT WITH CORRELATED ERROR TERMS

Scenario 4 mimics Scenario 2 presented above by constructing various opportunities across Nairobi's TAZs. Figure 10a and Figure 10b depict the change in consumer surplus experienced based on the mean value of 10,000 random individual preference sets.



(a) *(b)*

Figure 10:– Change in CS for (a: Informal Slum Dweller under Land Use Improvement Scenario with Gumbel Error Terms; b: Resident of Formal Housing under Scenario 4)

Again, compared to Scenario 3, Scenario 4 results display a higher prevalence of lighter greens on Figure 10a and Figure 10b thereby indicating overall changes in consumer surplus are more modest for the land use improvement than the transit-improvement scenario. One important distinction to point out, though, is IHR in nearly any TAZ (with the exception of TAZ #5) would tend to reap more accessibility benefits than residents of formal housing developments in the land-use improvement scenario compared to the transit improvement alternative. Results of Scenario 4 do not differ substantially from the values presented in Scenario 2 without correlated epsilons. Again, FHR are more likely to perceive a loss of accessibility; however, the mean value of this distribution of 10,000 simulations (Figure 10a & Figure 10b) shows no TAZ with a negative ΔCS for IHR or FHR.

COMPARING SCENARIOS

We can aggregate the values presented in the above figures by multiplying the mean change in CS experienced by a user in a given zone by the number of residents in that zone: These results are presented in Table 4. Looking at these results, firstly, the TI scenario vastly outperforms the LUI scenario with or without correlated epsilons. Secondly, the mean values of the 10,000 sample Monte Carlo Simulation incorporating the Gumbel error term provide significantly higher values of ΔCS for both IHR and FHR. Additionally, FHR benefit more than IHR in all scenarios except for the LUI scenario without correlated epsilons (Scenario 2).

$$Total\ \Delta CS_{i,d} = \sum_{d \epsilon D} Population_{i,d} * \Delta CS_{i,d} \quad (8)$$

Table 4: Aggregated Citywide Change in CS by User Type by Scenario

Scenario Type	Total City ΔCS – for IHR	Total City ΔCS – for FHR
1) TI no correlated epsilons	$123,035	$629,980
2) LUI no correlated epsilons	$36,755	$87,445
3) TI w/ correlated epsilons	$538,491	$1,656,082
4) LUI w/ correlated epsilons	$23,935	$90,020

DISCUSSION

Overall, the results display four distinct trends: 1) the TI scenario vastly outperforms the LUI scenario with or without correlated epsilons in aggregate and for 47 out of 48 zones for both IHR and FHR, 2) the mean values of the 10,000 sample Monte Carlo Simulation incorporating the Gumbel error term provide significantly higher values of ΔCS for both IHR and FHR, but more importantly provide more insight into the potential variation of these values at an individualized level that can be helpful for policy design, 3) peripheral zones tend to benefit more from the LUI than TI enhancements without correlated epsilons (Scenario 2), but this finding is not apparent when including the Gumbel error term (Scenario 4), and 4) in aggregate, FHR benefit more than IHR in all scenarios except for the LUI scenario without correlated epsilons (Scenario 2). These aggregations do hide some nuances that are important to be aware of. To investigate this, we will take a deeper dive into two very different TAZs and their respective results.

ZONE 5 KAREN (WEALTHIER ZONE)

Comparing the aggregated results of ΔCS across scenarios overlooks important pieces of data. For example, $50,328.74 of the $90k in ΔCS for FHR can be attributed to a single zone – "Karen" (Zone 5) – displaying a ΔCS/capita value of $3.76/day when $0.05/day was the largest achieved by any other zone in Scenario 4. Karen is a mostly affluent peripheral community to Nairobi with a population of 13,403 FHR and just 238 IHR. This zone receives the lowest zone attractiveness level of any zone, despite its massive size – 56.9 square kilometers. Karen is the only zone in this analysis lacking a matatu stop within its borders and thus we see more positive changes in ΔCS in the LUI than the TI scenarios, which opposes the trend for most every other TAZ. Karen sports extremely low levels of employment, health care and education density compared to the rest of Nairobi's TAZs. Increasing these values by 1 SD in the LUI scenarios imparts huge accessibility benefits to the TAZ alongside the modeled construction of 60 healthcare facilities, 136 schools, and the addition of over 15,000 jobs. While the magnitude of these benefits are more pronounced in aggregate for FHR in Karen, the area's small share of IHR also benefit significantly according to the models for both the TI ($2.21/day) and LUI ($0.54/day) scenarios, with most other TAZs hovering around $0.01 to $0.05/.

Though the results of the scenarios modeled with and without correlated Gumbel error terms display similar results, the distribution of potential ΔCS values gives more insight into the potential winners and losers when modeling each scenario. Having an idea of where this variation exists can be very helpful for impactful policy design that can help to target underserved areas and achieve more equitable accessibility outcomes. For example, in Zone 5, "Karen" (wealthier zone), the TI Scenario 3 resulted in an average $0.05/day ΔCS value for FHR, but with an SD of +/-$1.75, indicating many FHR are likely to experience some disbenefits from this scenario, presumably due to the increase in auto-mode impedance as a result of the scenario. Meanwhile, under the LUI Scenario 4, IHR in Zone 5, "Karen" (wealthier zone) experience the highest expected ΔCS value of any zone ($0.54/day), but with an SD of $1.30. Nearly 70% of the 10,000 simulated values occur between -$0.02 and $0.02. As is true in general: the average or expected value can be quite misleading, thanks to great variation in actual values.

ZONE 6 KAWANGWARE (LOW-INCOME ZONE)

To provide a counter example to Zone 5, the low-income Zone 6 ("Kawangware") has 111,057 IHR and just 2,069 FHR. Despite being more centrally located (but still several kilometers from the CBD), this area features below-average access to employment, schools, and health care opportunities, when compared to the other 48 TAZs. Table 7 below displays the results for Zone 6, "Kawangware" (low income zone). As one can see again, the TI scenarios vastly outperform the LUI scenarios for both user groups. Zone 6, "Kawangware" (low income zone) is one of many examples in this study whereby the LUI scenarios benefit IHR more so than FHR. Still, this is only by a very modest margin – a few hundred dollars' worth of benefit spread across over 100,000 IHR each day, but it is still a positive finding. More analysis should be done in order to determine what elements of a TAZ influence it to have a greater ΔCS for IHR versus FHR in order to design policies and target investments to improve these outcomes.

POTENTIAL POLICY SOLUTIONS

Several places in the developing world have enacted social housing programs and transit formalization programs aimed at increasing accessibility and engendering economic development and overall community welfare. Still, political leadership can be reticent to champion projects offering better services to informal (and usually illegal) slum settlements. In Bogota, Colombia, a project called *Metrovivienda* combined speculative land-banking with poverty alleviation alongside the build out of the Transmilenio BRT System. Starting in 1999, the organization purchased cheap agricultural land in close proximity to planned BRT terminuses and later on sold these plots to developers to construct affordable housing units targeted at *clandestinos*, or informal slum dwellers (Hidalgo and Huizenga, 2013). The transit agency also provides free feeder bus services to these areas and surrounding informal settlements to afford peripheral poor urban residents of Bogota access to the BRT trunk lines (Ferro, 2018). The solution is not perfect – the affordable housing units only cater to the top tier of those living in informal settlements and the 3,200 peso ($1.39 USD) round trip bus fare to access the CBD is inhibitory to many peripheral residents – but it's a positive start (Cervero, 2005).

Nothing of the sort has gone past a preliminary planning phase in Nairobi to date. The general public sentiment reflects more regulation, reliability and accountability is better for the transit system (Behrens et al., 2017; Salon and Aligula, 2012). The government has made significant strides to better enact and enforce ordinances to bring more structure to the matatu system (Behrens et al., 2017); this is apparent in that the system was able to be coherently explained in a GTFS feed and map by a group of researchers (Williams et al., 2015). The greatest impact will likely come with a multi-pronged accessibility improvement approach implementing land-use improvements in some places as well as improving the transportation network to improve accessibility from the transportation side as well.

CONCLUSION

Much data was generated as a result of this study, and many nuanced stories emerge for the two access improvement policies (transit vs. land use changes). This work highlights the fact that blanket policies affect everyone differently. Taking the extra effort to simulate and correlate before-after Gumbel error terms (rather than assuming independence between the two logsum terms) can help planners and policymakers tailor system design and decision-making more effectively than presuming the before and after populations (and the unobserved components of their choice alternatives) are independent.

Overall, this analysis confirms the original hypothesis: informal slum settlements in Nairobi tend to experience poorer access to basic daily needs - like education, health care and employment – than those living in formal housing. More importantly, under both access- improvement scenarios, residents of formal housing tend to benefit more. This benefits gap (between the informally and formally housed) was more pronounced (with greater CS changes) for the transit improvement (TI) scenarios, for both the types of residents. The results differ substantially by neighborhood, however. Some peripheral zones benefit more from land use improvements (LUI) than TI, despite the overall trend of TI outperforming LUI. Nairobi is a large region, with over 4 million people, and this data should be leveraged to analyze local conditions on a place by place basis. Policy decisions should incorporate public feedback from members of each locality in order to design the best policies to equitably benefit the entire city. One cannot design and operate an equitable, efficient, and effective transportation system without the input of those using this urban system day in and day out.

AUTHOR CONTRIBUTION STATEMENT

The authors confirm contribution to the paper as follows: writing-original draft preparation: L. Alcorn; conceptualization and design: K. Kockelman, L. Alcorn; methodology: K. Kockelman, L. Alcorn; data assembly and analysis: L. Alcorn; writing-reviewing and editing: L. Alcorn & K. Kockelman. All authors have reviewed the results and approved the final version of the manuscript.

REFERENCES

Avner, P., Lall, S., 2016. Matchmaking in Nairobi: The Role of Land Use, Policy Research Working Papers. The World Bank. https://doi.org/10.1596/1813-9450-7904

Behrens, R., McCormick, D., Orero, R., Ommeh, M., 2017. Improving paratransit service: Lessons from inter-city matatu cooperatives in Kenya. Transp. Policy 53, 79–88. https://doi.org/10.1016/j.tranpol.2016.09.003

Bhat, C., Handy, S.L., Kockelman, K., Mahmassani, H.S., Chen, Q., Weston, L., 2000. Urban Accessibility Index: Literature Review (No. TX-01/7-4938-1). Texas Department of Transportation, Austin, TX.

Boeing, A., 2016. OSMnx: Python for Street Networks. Geoff Boeing. URL http://geoffboeing.com/2016/11/osmnx-python-street-networks/ (accessed 5.8.18).

Campbell, K.B., Rising, J.A., Klopp, J.M., Mbilo, J.M., 2019. Accessibility across transport modes and residential developments in Nairobi. J. Transp. Geogr. 74, 77–90. https://doi.org/10.1016/j.jtrangeo.2018.08.002

Cervero, R., 2005. Progressive Transport and the Poor: ACCESS Mag. 1, 8.

Cervero, R., 2001. Informal Transit: Learning from the Developing World. ACCESS Mag. 18.

Cervero, R., Golub, A., 2007. Informal transport: A global perspective. Transp. Policy 14, 445–457. https://doi.org/10.1016/j.tranpol.2007.04.011

Cervero, R., Kockelman, K. 1997. Travel Demand and the Three D's: Density, Diversity, and Design. *Transportation Research Part D* 2(3): 199-219.

Edwin, W., 2010. 2010 Nairobi Land Use Data [WWW Document]. Wiki Sch. Nairobi GIS Maps. URL https://nairobigismaps.wikischolars.columbia.edu/file/view/landuse2010memo.pdf/360496680/landuse2010memo.pdf. (accessed 3.15.20).

ESRI, 2020. Add GTFS to a Network Dataset | GTFS in ArcGIS [WWW Document]. Using GTFS Data ArcGIS. URL https://esri.github.io/public-transit-tools/AddGTFStoaNetworkDataset.html (accessed 3.15.20).

ESRI Eastern Africa, 2017. Health Facilities in Kenya [WWW Document]. ArcGIS.com. URL https://www.arcgis.com/home/item.html?id=6ada9cfb67114d67977a266875800994 (accessed 4.4.20).

Farber, S., Morang, M.Z., Widener, M.J., 2014. Temporal variability in transit-based accessibility to supermarkets. Appl. Geogr. 53, 149–159. https://doi.org/10.1016/j.apgeog.2014.06.012

Ferro, P.S., 2018. Phone Interview with Dr. Pablo Salazar Ferro.

Google, n.d. GTFS Static Overview [WWW Document]. Google Transit APIs. URL https://developers.google.com/transit/gtfs (accessed 3.15.20).

Gwilliam, K.M., 1997. The Value of Time in Economic Evaluation of Transport Projects (No. OT-5). World Bank, Washington, DC.

Handy, S.L., Niemeier, D.A., 1997. Measuring Accessibility: An Exploration of Issues and Alternatives. Environ. Plan. Econ. Space 29, 1175–1194. https://doi.org/10.1068/a291175

Hansen, W.G., 1959. How Accessibility Shapes Land Use. J. Am. Inst. Plann. 25, 73–76. https://doi.org/10.1080/01944365908978307

Hidalgo, D., Huizenga, C., 2013. Implementation of sustainable urban transport in Latin America. Res. Transp. Econ. 40, 66–77. https://doi.org/10.1016/j.retrec.2012.06.034

Japan International Cooperation Agency (JICA), Nippon Koei Co., Ltd., IDCJ Inc., EJEC Inc., 2014. The Project on Integrated Urban Development Master Plan for the City of Nairobi in the Republic of Kenya. Nairobi, Kenya.

Kaltheier, R.M., 2002. Urban Transport and Poverty in Developing Countries. Dtsch. Ges. Für Tech. Zusammenarbeit, Environmental Management, Water, Energy, Transport 44, 58.

Kenya ICT Authority, 2020. Formal and Informal Settlements Population in Nairobi by Location 2009 [WWW Document]. Kenya Open Data. URL https://www.opendata.go.ke/datasets/formal-and-informal-settlements-population-in-nairobi-by-location-2009-2 (accessed 3.11.20).

Kockelman, K. 1997. Travel Behavior as a Function of Accessibility, Land Use Mixing, and Land Use Balance: Evidence from the San Francisco Bay Area. *Transp Research Record* 1607: 117-125.

Kockelman, K., Chen, T.D., Larsen, K., Nichols, B., 2013. Economics of Transportation Systems: A Reference for Practitioners (No. 0-6628-P1). Texas Department of Transportation, Austin, TX.

Kockelman, K.M., Lemp, J.D., 2011. Anticipating new-highway impacts: Opportunities for welfare analysis and credit-based congestion pricing. Transp. Res. Part Policy Pract. 45, 825–838. https://doi.org/10.1016/j.tra.2011.06.009

Litman, T., 2003. Measuring Transportation: Traffic, Mobility and Accessibility. Inst. Transp. Eng. J. 73, 28–32.

Ma, S., Kockelman, K.M., Fagnant, D.J., 2015. Welfare Analysis Using Logsum Differences vs. Rule of Half: A Series of Case Studies. Transp. Res. Rec. 78–83.

Marras, S., 2012. Mapping the Unmapped. Map Kibera Proj. URL http://mapkiberaproject.yolasite.com/resources/Kibera_mapping_the_unmapped.pdf (accessed 3.11.20).

Omulo, C., 2019. City Hall doubles parking fees. Dly. Nation.

Salon, D., Aligula, E.M., 2012. Urban travel in Nairobi, Kenya: analysis, insights, and opportunities. J. Transp. Geogr. 22, 65–76. https://doi.org/10.1016/j.jtrangeo.2011.11.019

Srour, I., Kockelman, K., Dunn, T. 2002. Accessibility Indices: A Connection to Residential Land Prices and Location Choice. *Transp Research Record* 1805: 25-34.

The World Bank, 2016. Republic of Kenya Urbanization Review (No. AUS8099). Washington, DC.

Varian, H.R., 1992. Microeconomic Analysis, Third Edition. ed. W.W. Norton & Company, Inc., New York, NY.

Walker, J., Li, J., Srinivasan, S., Bolduc, D., 2010. Travel demand models in the developing world: correcting for measurement errors. Transp. Lett. 2, 231–243. https://doi.org/10.3328/TL.2010.02.04.231-243

Williams, S., White, A., Waiganjo, P., Orwa, D., Klopp, J., 2015. The digital matatu project: Using cell phones to create an open source data for Nairobi's semi-formal bus system. J. Transp. Geogr. 49, 39–51. https://doi.org/10.1016/j.jtrangeo.2015.10.005

Yu, H., Shaw, S.-L., 2008. Exploring potential human activities in physical and virtual spaces: a spatio-temporal GIS approach. Int. J. Geogr. Inf. Sci. 22, 409–430. https://doi.org/10.1080/13658810701427569

Zhao, Y., Kockelman, K., Karlström, A., 2012. Welfare calculations in discrete choice settings: An exploratory analysis of error term correlation with finite populations. Transp. Policy 19, 76–84. https://doi.org/10.1016/j.tranpol.2011.09.002