The Housing Suitability Model:
Measuring Accessibility and Opportunity in Affordable Housing

March 2017

Shimberg Center for Housing Studies
University of Florida
P.O. Box 115703
Gainesville, Florida 32611-5703

The work that provided the basis for this publication was supported by funding under a grant with the U.S. Department of Housing and Urban Development. The substance and findings of the work are dedicated to the public. The author and publisher are solely responsible for the accuracy of the statements and interpretations contained in this publication. Such interpretations do not necessarily reflect the views of the Government.

Introduction

The Housing Suitability Model (HSM) is a GIS-based tool for analyzing the suitability of locations for the development or preservation of affordable housing. The HSM incorporates a series of data layers superimposed within a GIS environment to assess the suitability of locations across a region for a particular type of use.

The HSM is an outgrowth of the Land-Use Conflict Identification Strategy (LUCIS) model developed at the University of Florida (Zwick et al., 2015). The HSM classifies locations spatially based on evaluation of a series of criteria, referred to as objectives. The criteria are grouped according to goals that reflect distinctive characteristics. The model is built from layers of spatial data that represent each objective. Within each spatial layer, locations are ranked or scored according to how well they represent the characteristic that is being evaluated by that layer.

The unit of analysis for the HSM is the land parcel. In some cases, such as measures of distances from housing locations to service destinations, data are created at the unit level and summarized between destinations associated with their respective parcels. In others, such as variables based on American Community Survey data, values are already calculated for a Census geography level, such as block group or tract. In those cases, the same ranks or scores are assigned to all parcels within the boundaries of a particular block group or tract.

The Shimberg Center for Housing Studies originally developed the HSM as a local planning tool in three Florida counties for identifying suitable land for affordable housing development based on land use, driving and transit accessibility, physical characteristics, and affordability of existing housing. The model was then modified for use in HUD Sustainable Communities regional planning processes in the Orlando metropolitan area and a group of seven rural counties in southwest Florida.

Most recently, the Center used the model to evaluate existing Housing Choice Voucher (HCV) and Low Income Housing Tax Credit (LIHTC) locations in 16 Florida counties. For this version, the model includes two goals that identify suitable sites for affordable housing: maximizing transportation accessibility and maximizing tenants' access to educational and economic opportunity, based on HUD's Affirmatively Furthering Fair Housing indices. In this paper, we describe the methodology underlying this version of the model as an example of the use of the HSM to evaluate affordable housing locations.

Model Structure

The goals and objectives for the most recent version of the HSM are shown in Figure 1.

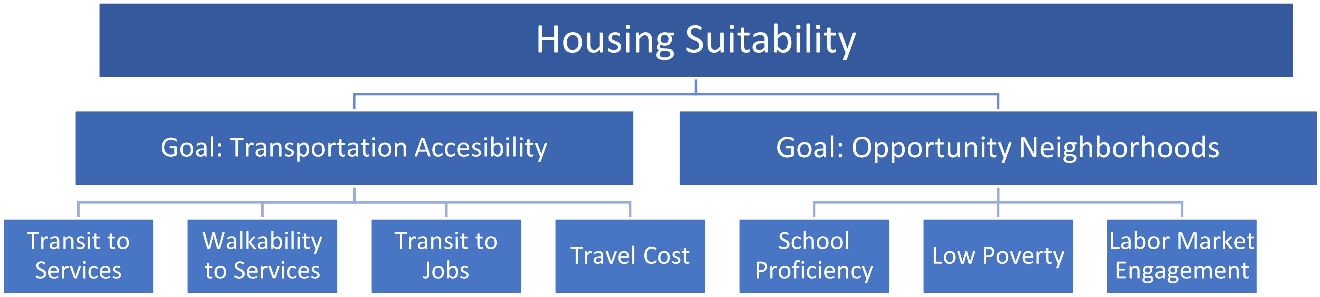

Figure 1. Housing Suitability Model Goals and Objectives

The Center developed new parcel-level measures of transit accessibility to services, walking access to services, and transit access to jobs. The methodology behind these new measures is described in detail below. We used existing block group-level measures of travel cost from HUD's Location Affordability Index (LAI) version 2 and of school proficiency, low poverty and labor market engagement from HUD's Affirmatively Furthering Fair Housing (AFFH) data.

Transportation Accessibility: Parcel-Level Measures

The three parcel-level accessibility scores measure the availability of destination types (jobs, services) from origins (all parcels) within a given distance or travel time. Service destinations include medical facilities, pharmacies, schools, libraries, community and senior centers, and grocery stores. These are chosen as essential services to support the health, education, and community and social involvement of affordable housing residents. More specifically, proximity to these destinations is incentivized under Florida Housing Finance Corporation's competitive process for funding from LIHTC and the state's affordable housing trust fund.

For each accessibility component, the general process was to create a matrix showing distance and associated time along the travel network between all origin parcels and all destinations. To do so, we chose random origin points, measured their distance and time to destinations, and then interpolated from those origin points to create a raster with estimates of origin-destination availability for a given travel time within each study county. The cell values from the raster data were then summarized to the parcel level so that accessibility could be measured from any parcel.

<u>Destination Data Sources</u>

Employment destinations came from the Transit Boardings Estimation and Simulation Tool (TBEST), available from the Florida Department of Transportation. TBEST is a transit simulation and demand modeling tool that includes data on business destinations (Center for Urban Transportation Research, 2015). Service destinations came from the Florida Geographic Data Library (FGDL), a source of statewide geospatial data collected from more than 35 federal, state, regional, local and private agencies (Florida Geographic Data Library, n.d.). FGDL is maintained by the GeoPlan Center at the Department of Urban and Regional Planning, University of Florida.

Walking Score

Walking times were measured based on network distance between two points. We found destinations within a five-minute walking distance of a particular origin by assuming a walking speed of 4.8 km (approximately three miles) per hour, resulting in a maximum distance of 400 meters. The street network was created from NAVTEQ data. ESRI's ArcGIS Network Analyst tool was used to create an origin-destination matrix, where random points were selected as origins and the service locations from FGDL were destinations. Each random point was assigned a count of destinations from each of the eight service categories that were within 400 meters of the origin. Next, inverse distance weighting (IDW) was used to interpolate values between the random points to create a raster layer with origin-destination counts for the entire county. IDW creates a matrix of cells whose values are determined by applying a weight that progressively decreases as the distance to a given origin point increases (ESRI, n.d. a).

Finally, the raster data were summarized to the parcel level to create a count of each type of service destination located within 400 meters of each origin parcel. Composite walking scores consist of a count of the total number of service destinations available from the parcel. For example, a parcel within 400 meters of one grocery store and two pharmacies would receive a score of 3.

Transit Access to Services and Jobs

We measured accessibility from origin parcels to service and job destinations by creating a series of time sheds; that is, by counting the number of destinations accessible to a parcel by transit within a 15, 30, 45, and 60 minute trip.

Information about transit stops, routes, and schedules came from the General Transit File Specification (GTFS) dataset. The GTFS data was used to produce a point feature class containing the transit stops and a line feature class containing the transit lines. We used ESRI's ArcGIS Network Analyst extension to create connector features between the transit lines/stops and the underlying street network. Transit stops were spatially adjusted to the street network using the connectors to create a transit network dataset that specifies actual travel routes between origins and destinations. Using the new transit network dataset, total transit time was calculated as the sum of all components of an origin-destination trip, including walking time to the origin transit stop, waiting time for transit, travel along the transit network to a destination stop, and walking time to destination.

Again, we created a series of origin-destination matrices, this time for employment destinations plus the eight categories of service destinations. Transit network service areas were created by generating 400m-radius buffers around each stop, which were then dissolved to create a single transit stop walking shed.[1] Random points were generated within the transit stop walking sheds to serve as the universe of origin points. Transit travel times between origins and destinations

[1] Four hundred meters (approximately 0.25 miles) has been found to be more representative with respect to job accessibility, in contrast to the widely-used standard of 800 meters for identifying walking-accessible transit stops (Guerra, et al., 2011; Houston, et al., 2013; Walter, et al., 2016).

were calculated based on distances and impedances depending on time of day and transportation network complexity. Origins and destinations were aggregated based on time step thresholds into 15, 30, 45 and 60-minute travel bins. The bins were cumulative; that is, a destination appearing in a shorter travel time shed would also appear in all of the longer time sheds. For example, a destination within 30 minutes of an origin would appear in the 30 minute time shed and, in addition, the 45 and 60 minute time sheds. Finally, as with the walking scores, point-based travel times were estimated for all origins in the transit shed using inverse distance weighting interpolation. This resulted in four raster layers for each of the destination categories, with origin-destination counts for the entire county within each of the time sheds. Again, raster values were summarized to the parcel level.

Next, we created two composite scores for each parcel—one for combined service accessibility and one for jobs. In each case, we used a decay function to combine the four transit shed counts into a single score, with destinations appearing in the shorter time sheds receiving more weight.

Transit to Services Score

The transit to services decay score for each parcel was higher if a variety of service types was accessible from the origin, if there were several instances of a single type of destination, and if travel times were shorter. First, for each time shed, each service category was assigned a point value based on the number of destinations of that type:

Table 1. Point Values for Service Counts, Transit to Services Score

Services Count		Points
0	=	0
1	=	0.75
2 or 3	=	1.5
3 +	=	3.125

For each parcel, the points for each destination type and time shed were entered into a matrix as shown in Table 2 (cell values in the table indicate maximum number of points available). The sum of the points for each time shed was multiplied by the weights in the second-to-last line of Table 2, which in turn were based on the percentage of transit trips of each length in Florida from the 2009 National Household Travel Survey (Center for Transportation Analysis, 2009). Finally, the weighted points from the four sheds were summed, generating a score between zero and 100.

Table 2. Scoring Matrix for Transit to Services Score

	15 min shed (max points)	30 min shed (max points)	45 min shed (max points)	60 min shed (max points)
Community Centers	3.125	3.125	3.125	3.125
Grocery Stores	3.125	3.125	3.125	3.125
Libraries	3.125	3.125	3.125	3.125
Medical Facilities	3.125	3.125	3.125	3.125
Parks	3.125	3.125	3.125	3.125
Pharmacies	3.125	3.125	3.125	3.125
Schools	3.125	3.125	3.125	3.125
Senior Centers	3.125	3.125	3.125	3.125
Max. Total Points	25	25	25	25
weights	**0.3324**	**0.8300**	**1.2692**	**1.5684**
Max. Weighted Points	8.31	20.75	31.73	39.21

(Maximum total score = 100; i.e. 8.31 + 20.75 + 31.73 + 39.21)

Because destinations within a shorter time shed also appear in all longer time sheds, a destination closer to the origin contributes more to the total score. For example, a single grocery store within the 15 minute time shed would contribute a total of 3 points to the score when its contribution to the 15, 30, 45 and 60 minute time shed scores are totaled. (0.75 * (0.3324 + .8300 + 1.2692 + 1.5684)). A single grocery store that is only located within the 60 minute time shed, however, would contribute 1.18 points (0.75 * 1.5684).

A zero score indicates that no destinations of any of the service types were accessible from the origin within a 60-minute time shed, while the maximum score of 100 indicates that three or more instances of each destination type were accessible within a 15-minute time shed.

Transit to Jobs Score

The transit to jobs score is calculated differently, since the number of jobs accessible by transit from origins can vary from zero into the hundreds of thousands. Therefore, the base scores were determined by the relative number of jobs in the time shed rather than the absolute number of destinations.

All parcels were placed in four master files, one for each of the time shed values. Within each file, parcels were classified into one of 10 possible quantiles based on the number of jobs accessible within that time shed, with equal numbers of parcels in each quantile (excluding parcels with zero accessible jobs). The result is that each parcel has four time shed attributes with associated values from 1-10, or zero if there were no jobs at all in that time shed. For example, TOT15 refers to this 1-10 value for the 15-minute time shed, and so forth for the other time sheds.

An overall employment accessibility score was derived by weighting each time shed quantile score and summing the weighted results. Weights for the time sheds were the same as used for service destinations. For each parcel, scores were calculated as follows:

([TOT15] * 0.3324) + ([TOT30] * 0.83) + ([TOT45] * 1.2692) + ([TOT60] * 1.5684)
where minimum value for each TOT(n) score is zero and maximum value is 10.

As with service destinations, employment destinations are cumulative, with destinations located within the 15-minute time shed also counted as being within all subsequent time sheds. Therefore, closer destinations contribute a higher value to the overall score. The maximum job score is 40, which would be obtained if the destination fell in the highest quantile of job counts in the county for all four time sheds. The lowest is zero, which would only occur if no jobs were accessible by transit within even a 60 minute transit trip from the origin.

Travel Cost

The travel cost measure comes from HUD's Location Affordability Index (LAI). The costs are measured in estimated dollars per year spent on driving and transit use.

The Location Affordability Index (LAI) was developed by HUD to estimate transportation and housing cost burdens based on location. Estimated transportation costs are calculated from a regression model by estimating the proportion of auto ownership, auto use and transit use, within a given Census block group, as a function of fourteen household, transit network and employment variables (median income, per capita income, average household size, average commuters per household, residential density, gross density, block density, intersection density, transit connectivity, transit frequency of service, transit access shed, employment access, job diversity, and average commute distance) (HUD, 2013). Variation in household characteristics and the subsequent impact on the cost estimates is controlled by establishing several representative household profiles. Base data for the LAI is compiled by the Center for Neighborhood Technology (Center for Neighborhood Technology, 2016) and sourced from the 2010 Census American Community Survey 5-year estimates, TIGER/Line files, Longitudinal Employment-Household Dynamics Origin-Destination Employment Statistics, as well as locally available transit data (HUD, 2013).

A travel cost model was developed for the study area, using HUD LAI transportation cost estimates for a single parent renter households at 50 percent of Area Median Income (AMI). This profile was chosen to best match the characteristics of affordable housing tenant households.[2]

The LAI dataset does not provide the cost numbers in dollars, but it provides an annual household income amount and a modeled percentage of income spent on transportation for each household type. We multiplied total household income by percentage of income spent on

[2] For example, in Florida, 59% of households residing in units subsidized by Florida Housing Finance Corporation and 92% of Housing Choice Voucher holders have incomes of 50% AMI or below (Shimberg Center for Housing Studies, 2016; HUD, 2015).

transportation to derive an estimated annual travel cost in dollars for each Census block group. The result was assigned to all parcels within the given block group.

Neighborhood Access to Opportunity

The remaining components of the model are indicators of neighborhood access to economic and educational opportunity, as measured by HUD's Affirmatively Furthering Fair Housing indices. We used three of the eight AFFH opportunity indices: the Low Poverty index, School Proficiency index, and Labor Market Engagement index. While the AFFH Jobs Proximity index is also a measure of neighborhood economic opportunity, this aspect is already included in the model through the transit to jobs score.

The AFFH indices are measured at the Census block group and tract level. As with the LAI, we applied the values for a block group or tract to all of the parcels within its boundaries.

Low Poverty

The Low Poverty index measures the percentage of the population below the poverty level in a Census tract compared to the national average. Possible score values are 0 to 100 based on a percentile ranking of tracts against others in the U.S.; higher values indicate a *lower* poverty rate. The data source for this index is the 2013 5-Year American Community Survey (Abt Associates and Mast, 2015).

School Proficiency

The School Proficiency index is a block group level measure of the presence of high-performing elementary schools, based on the percentage of fourth graders achieving proficiency for reading and math on standardized tests. Possible score values are 0-100 based on percentile ranking of the block group against others in the U.S. Data sources for this index include Great Schools, 2012; Common Core of Data (4th grade enrollment and school addresses), 2012; School Attendance Boundary Information System (SABINS), 2012 (Abt Associates and Mast, 2015).

Labor Market Engagement

The Labor Market Engagement index is measured at the Census tract level. It combines the percentage of residents who are employed, the percentage of residents participating in the workforce, and the percentage of residents with a bachelor's degree or higher compared to national means. Again, scores range from 0 to 100 based on a percentile ranking against other U.S. tracts. Data come from the 2013 5-Year American Community Survey (Abt Associates and Mast, 2015).

Model Summary

In sum, the current version of the Housing Suitability Model includes six scores for each parcel:

- Transit to services: 0-100 score based on accessibility to eight service destinations within 15-60 minute time shed
- Transit to jobs: 0-40 score based on accessibility to employment destinations within 15-60 minute time shed
- Walking to services: Count of service destinations with ¼ mile network walking distance
- Travel cost: Estimate of annual auto and transit costs based on LAI single commuter, renter, very low-income household profile for block group
- Low Poverty: 0-100 score with higher score indicating lower percentage of population in poverty for block group
- School Proficiency: 0-100 score with higher score indicating greater reading and math proficiency in neighborhood elementary schools for block group
- Labor Market Engagement: 0-100 score with higher score indicating higher employment, workforce participation, and educational attainment in block group

The model dataset also includes underlying counts of employment destinations and each of the eight categories of service destinations within ¼ mile walking distance and within each of the four transit time sheds. These underlying values can be used to create new weighted scores that emphasize particular destinations, according to the user's priorities.

Example: Evaluation of Affordable Housing Locations

The Center applied the HSM to affordable rental housing units subsidized by Housing Choice Vouchers (HCV), funds allocated by Florida Housing Finance Corporation (Florida Housing), or both. The major source of Florida Housing-allocated funds was the Low Income Housing Tax Credit (LIHTC). Many of these developments also received Florida Housing funding through private activity bonds and the state's affordable housing trust fund.

The purpose of the analysis was to determine how locations receiving either or both of these two types of funding compare in terms of HSM objectives, and in particular whether Florida Housing-sponsored units offer opportunities to voucher holders that are unavailable through more traditional market-rate multifamily and single family options. The full study is available as "Transportation, Services and Jobs: A Comparison of Accessibility for Housing Choice Voucher and Low Income Housing Tax Credit Locations in Florida."

<u>Geographic Scope</u>

Locations were assessed in in 16 Florida counties: Miami-Dade, the state's largest urban county; the counties making up the Jacksonville (Duval, Nassau, Clay, St. Johns), Tampa-St. Petersburg (Hillsborough, Pinellas, Pasco, Hernando), and Orlando (Orange, Osceola, Seminole, Lake) metropolitan areas; and three smaller metropolitan counties (Alachua, Volusia, and Polk). To simplify the analysis, the counties were grouped into four "functional regions" with similar characteristics.

Table 3. Functional Regions

Functional Region	Reason for Grouping	HUD Metro Fair Market area(s)	Counties
Miami-Dade	Largest urban county; accessibility much higher than all other counties and regions	Miami-Miami Beach-Kendall HMFA	Miami-Dade
Jacksonville/Tampa	Large metropolitan areas with moderate accessibility.	Jacksonville HMFA	Clay
			Duval
			Nassau
			St. Johns
		Tampa-St. Petersburg-Clearwater MSA	Hernando
			Hillsborough
			Pasco
			Pinellas
Medium Counties	Medium-sized counties with more suburban development patterns.	Gainesville MSA	Alachua
		Lakeland-Winter Haven MSA	Polk
		Deltona-Daytona Beach-Ormond Beach HMFA	Volusia
Orlando	Large metropolitan area with low accessibility compared to other large MSAs.	Orlando-Kissimmee-Sanford MSA	Lake
			Orange
			Osceola
			Seminole

<u>Household Dataset</u>

The analysis compared HSM objectives for locations of 139,210 households receiving a voucher, residing in a Florida Housing-sponsored unit, or both in 2013.

Tenant administrative datasets provided most of the household and unit-level information used in the study. HUD provided records for HCV participants from its Form 50058 database, which tracks address, demographic, income, housing structure type, and rent characteristics for voucher recipients. Location and household characteristics for tenants of Florida Housing-sponsored developments came from the Shimberg Center's Assisted Housing Inventory (AHI) database and administrative records provided by Florida Housing.

Both datasets included geocoding information for the residents' addresses. These were matched to parcel identifiers from the Florida Department of Revenue's tax roll data from county property appraisers. In this way, households could be matched to the appropriate parcel-level HSM scores.

When the parcel identifier in the HCV dataset matched parcels assigned to Florida Housing properties, the unit was identified as receiving both voucher and Florida Housing assistance. Using information from the matched HCV, Florida Housing, and property appraiser datasets, each household was classified in one of four housing types:

- Florida Housing units without a voucher: 86,822 households
- Florida Housing units with a voucher: 7,259 households
- Voucher in market-rate single family home: 18,489 households
- Voucher in market-rate multifamily unit: 26,640 households

Score Analysis

The six objective scores were summarized to two composite goal measures using principal components analysis (PCA). PCA can be used to weight and reduce correlated variables to a smaller number of indicators.

The four transportation accessibility scores were shown to be correlated using a Chi-Square Test of independence. PCA was used in SPSS to extract a single standardized factor, based on statistically estimated weights applied to the original scores. The standardized factor was converted to a composite accessibility score on a 0-10 scale by linear transformation, with zero representing the least accessible locations.

The comparison of accessibility scores within each region showed that overall, market-rate multifamily voucher locations were the most accessible housing type. In most regions, single family units with voucher holders were the least accessible housing type, with Florida Housing units either with or without voucher holders as a middle ground. However, the analysis also showed that results varied by building age. Two groups of Florida Housing units were as accessible or more accessible than the market-rate multifamily voucher locations: 1) pre-1989 units that were later rehabilitated and preserved using LIHTC and other Florida Housing funds, and 2) units built in 2004 and later, after Florida Housing instituted strong incentives for proximity to transit and services in its funding competitions.

Similarly, the three neighborhood opportunity index scores were found to be correlated and were combined into a single score using PCA, this time on a scale of 0-100 corresponding to the objective-level scores. Unlike with the transportation accessibility scores, there were no clear patterns across regions when comparing the neighborhood opportunity scores. For example, the highest opportunity scores were found in Florida Housing units without vouchers in the Jacksonville/Tampa and Orlando regions, but this was the lowest scoring housing type in medium counties. Therefore, these results were not included in the “Transportation, Services and Jobs” paper.

Conclusion and Next Steps

The Housing Suitability Model has gone through several iterations to arrive at the current version, which evaluates land parcels in terms of transportation accessibility and neighborhood opportunity goals. The model has generated a rich, parcel-level dataset with a variety of transportation cost, transit and walking time shed, and opportunity measures.

Our next step will be to use the model and underlying data to support regional and local affordable housing planning in Florida. Examples include a multi-county regional planning initiative to increase the affordable housing supply in Central Florida and data analysis for a potential Choice Neighborhoods grant application for an urban core neighborhood in North Florida. We are particularly interested in developing applications for the underlying transit and walking destination shed data to pinpoint locations for affordable housing development linked to specific goals, such as access to jobs for working-age adults or increasing seniors' access to health and social involvement destinations. We also plan to continue to develop opportunity neighborhood indicators based on local planners' priorities.

References

Abt Associates and Brent Mast (2015). *Affirmatively Furthering Fair Housing (AFFH) data documentation*. Retrieved from https://www.huduser.gov/portal/sites/default/files/docs/AFFH_Data_Documentation_12_31_2015.docx

Center for Transportation Analysis (2009). National household travel survey (NHTS). Retrieved from http://nhts.ornl.gov/introduction.shtml

Center for Urban Transportation Research (2015). Transit boardings estimation and simulation tool (TBEST). Retrieved from http://tbest.org

ESRI (n.d.-a). How IDW works. Retrieved from http://desktop.arcgis.com/en/arcmap/10.3/tools/3d-analyst-toolbox/how-idw-works.htm.

ESRI (n.d.-b) Understanding connectivity. ArcGIS Help 10.2. http://resources.arcgis.com/en/help/main/10.2/index.html#//004700000009000000

Florida Geographic Data Library (n.d.). Florida Geographic Data Library metadata explorer. Retrieved from http://www.fgdl.org/metadataexplorer/about.html

Guerra, E., Cervero, R. & Tischler, D. (2011). *The half-mile circle: does it best represent transit station catchments?* (Working Paper UCB-ITS-VWP-2011-5). Retrieved from http://www.its.berkeley.edu/sites/default/files/publications/UCB/2011/VWP/UCB-ITS-VWP-2011-5.pdf

Houston, D., Basolo, V. & Yang, D. (2013). Walkability, transit access, and traffic exposure for low-income residents with subsidized housing. *American Journal of Public Health, 103*(4), 673-678.

HUD. (n.d.-a). Housing Choice Vouchers fact sheet. Retrieved from https://portal.hud.gov/hudportal/HUD?src=/program_offices/public_indian_housing/programs/hcv/about/fact_sheet

HUD. (n.d.-b). Location Affordability Index data dictionary. Retrieved from http://lai.locationaffordability.info//lai_data_dictionary.pdf

HUD. (n.d.-c). Location Affordability Portal version 2: understanding the combined cost of housing and transportation. http://locationaffordability.info/about_lai.aspx

HUD. (2013). Data and methodology: Location Affordability Index and My Transportation Cost Calculator. Retrieved from http://lai.locationaffordability.info//LAPMethods_v1.pdf

HUD. (2015). Assisted housing: national and local. Retrieved from https://www.huduser.gov/portal/datasets/assthsg.html

Infogroup (2017). InfoUSA. Retrieved from https://www.infousa.com

Mast. B. (2015). Measuring neighborhood opportunity with AFFH data. *Cityscape: A Journal of Policy Development and Research*, *17*(3), 221-230.

Patton. T. (2014). GTFS brings open data to public transit: A developer's first look at the spec. Retrieved from http://www.techrepublic.com/article/gtfs-brings-open-data-to-public-transit-a-developers-first-look-at-the-spec

Shimberg Center for Housing Studies. (2016). 2016 rental market study. Retrieved from http://www.shimberg.ufl.edu/publications/Full_RMS_final_rev09_16.pdf

Walter, R., Evans, A. & Atherwood, S. (2016). Addressing the affordable housing crisis for vulnerable renters: insights from Broward County on an affordable housing acquisition tool. *Housing Policy Debate*, *26*(1), 123-149.

Zwick. P, Patten, I., & Arafat. A. (2015). *Advanced land-use analysis for regional geodesign: using LUCISplus*. Redlands, CA: ESRI Press.