Housing Affordability and Immigration:
An Exploratory Analysis in New York City

Jhonatan Medri[*]   Braden D. Probst[†]   Jürgen Symanzik [‡]

Abstract

The relationship between housing affordability and immigration has become relevant in the U.S., especially in a highly populated metropolis such as New York City (NYC). Determining whether immigration status affects home ownership, rent, or shelter cost could help understand the quality of life of NYC residents. Graphical exploration of housing affordability, immigration, and other social variables provides some insights about their relationships. Our exploration takes place at the borough and the sub-borough level.

Key Words: Visualization, Immigration, Home Ownership, Home Renting, Shelter Cost, EDA

1. Introduction

Immigration and housing affordability are two popular topics broadly considered when discussing public policies and human development. Renting or owning a home is the first step of getting out of poverty, and, also, a symbol of welfare accumulation. Unfortunately, sometimes having access to a place to live can involve some obstacles.

In the case of immigration, new difficulties could appear. The fact that immigrants are new to a place could imply they have to pay a symbolic "admission fee" to adapt. This "admission fee" can come in the form of lower household income, lower credit access, or just a cultural barrier. The first generations of immigrants are usually the ones that pay this "admission fee," so that the following generations can have better life conditions.

The 2019 Data Challenge Expo of the Sections on Statistical Computing, Statistical Graphics, and Government Statistics of the American Statistical Association (ASA) provided an opportunity to explore the relationship between housing affordability and immigration in New York City (NYC). Our analyses focused on the following motivating question:

How are housing affordability and immigration status related in NYC?

In addition, our analyses addresses the following specific questions:

- How are home ownership and immigration status related in NYC?

- How are home renting and immigration status related in NYC?

- How are shelter cost and immigration status related in NYC?

- How do these relationships differ among NYC sub-boroughs?

[*]Department of Mathematics and Statistics, Utah State University, 3900 Old Main Hill, Logan, UT 84322, `jmedri@aggiemail.usu.edu`

[†]Department of Mathematics and Statistics, Utah State University, 3900 Old Main Hill, Logan, UT 84322, `braden.probst@gmail.com`

[‡]Department of Mathematics and Statistics, Utah State University, 3900 Old Main Hill, Logan, UT 84322, `juergen.symanzik@usu.edu`
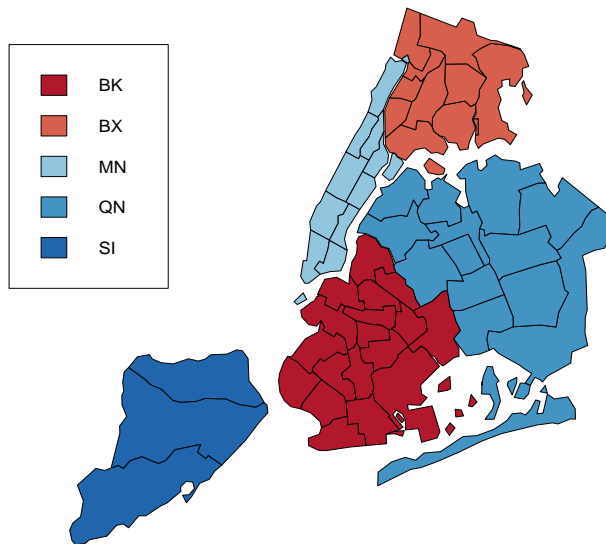
Figure 1: NYC Sub-boroughs

This article is organized as follows. In Section 2, we describe the data used and the variables selected in our analyses. We then proceed to address home ownership in Section 3, home renting in Section 4, and shelter cost in Section 5. We provide some conclusions and suggestions for future research in Section 6.

2. Exploratory Data Analysis (EDA) at the Borough Level

The data set used for the 2019 Data Challenge Expo is the New York City Housing and Vacancy Survey 2017 (NYCHVS). The NYCHVS is a representative survey of the New York City housing stock and population, sponsored by the New York City Department of Housing Preservation and Development (HPD). This survey takes place every three years and collects information from both vacant and occupied housing units. The data set is available at the following URL:

`http://www1.nyc.gov/site/hpd/about/nychvs-asa-data-challenge-expo.page`.

In this article, we work with the data from 2017 only and do not make any temporal assessments.

2.1 Variable Selection

Table 1 indicates the variables considered in our analyses. We decided to focus our analyses on the variables from the survey relating to housing affordability and immigration.

From the housing affordability perspective, we are including home ownership and renting variables. Home ownership variables provide information about how many tenants own a home and the conditions of their mortgage, in case the home is not loan-free. In addition, home renting variables indicate how many people live for free or pay rent and, in the second case, how much that amount is.

We constructed a shelter cost index that estimates monthly contract rent as a proportion of household income to get a relative perspective of housing affordability in NYC. Previous work made some estimations regarding housing costs (Elmelech, 2004); however, because of the complexity in the definition of these costs (Belsky et al., 2005), we will just consider the monthly contract rent in our analyses.

Table 1: List of housing and immigration variables included in our analyses. The sub-borough range relates to "1" for variables 1-5.

Variable Name	Description	Sub-borough Range
1. Householder's Birthplace	1 if outside the U.S	[14.1% - 74.4%]
2. Householder Parents' Birthplace	1 if outside the U.S	[35.5% - 90.9%]
3. Tenure Owner / Renter	1 if tenant owns home	[4.3% - 73.6%]
4. Tenure Cash Rent	1 if tenant pays cash rent	[21.5% - 93.5%]
5. Mortgage Status	1 if home has a mortgage	[27.3% - 87.5%]
6. Mortgage Interest Rate	Percentage	[3.2% - 5.0%]
7. Total Household Income	Annual amount in USD	[$20,696 - $145,000]
8. Contract Rent	Annual amount in USD	[$9,444 - $36,300]
9. Shelter Cost	Rent as a proportion of income	[12.0% - 54.7%]

From the immigration perspective, we are mainly considering the primary householder's origin. We first look at the place of the householder's birth country and classify this group as immigrants or US citizens. We then look at the place of the householder parents' birth country to classify the household as first or second generation immigrant or citizen. We also use that information to see any differences within each immigration group.

We run our analyses on the 55 sub-boroughs within the five main boroughs in NYC: Brooklyn (BK), Bronx (BX), Manhattan (MN), Queens (QN), and Staten Island (SI). Boroughs are county-level administrative divisions of NYC, while sub-boroughs are groups of census tracts summing to at least 100,000 residents. Figure 1 shows the sub-boroughs for each NYC borough.

2.2 Descriptive Statistics

We can make some general comments for each borough according to the descriptive statistics displayed in Tables 2, 3, and 4. New York City residents, compared to the rest of U.S. cities, tend to rent a place rather than own one. In addition, they report a higher household income and rent, but also a higher shelter cost (U.S. Census Bureau, 2019a). Moreover, NYC has a relatively high immigration and first generation concentration in the country (U.S. Census Bureau, 2019b).

As shown in Table 2, less than a third of the population owns a home in New York City. Staten Island is the borough with most owners (64%) and Bronx the one with least (20%). About 61% of those who own a home have a mortgage loan at an average interest rate of 4%. However, there are no major differences at the five boroughs.

In contrast, as shown in Table 3, about 67% of New Yorkers tend to rent a place. Bronx

Table 2: Home ownership descriptive statistics for 2017

Borough	% Owners	% Debtors	Interest Rate* (%)
Brooklyn	28.8%	64.6%	4.07%
Bronx	19.8%	50.1%	4.05%
Manhattan	21.1%	56.4%	3.65%
Queens	43.2%	61.8%	4.00%
Staten Island	63.6%	68.0%	4.00%
Total	30.8%	61.2%	4.00 %

* Annual Median value.

Table 3: Home renting descriptive statistics for 2017

Borough	% Renters	Rent* (USD)	Income* (USD)	Shelter Cost (%)
Brooklyn	68.9%	$15,600	$55,480	28.1%
Bronx	77.6%	$13,440	$35,000	38.4%
Manhattan	77.5%	$19,800	$72,000	27.5%
Queens	54.7%	$17,400	$62,800	27.7%
Staten Island	33.3%	$13,200	$75,000	17.6%
Total	67.1%	$15,600	$58,000	26.9%

* Annual Median value.

Table 4: Immigration descriptive statistics for 2017

Borough	% Immigrants	% First Generation
Brooklyn	38.6%	63.5%
Bronx	35.3%	67.2%
Manhattan	23.1%	50.0%
Queens	53.1%	75.2%
Staten Island	24.0%	47.3%
Total	36.3%	62.6%

and Manhattan are the boroughs with the highest percentage (78%) and Staten Island with the lowest (33%). Similarly, the annual median household rent and income are $15,600 and $58,000, respectively. Staten Island is the borough with the highest income ($75,000) and lowest rent ($13,200). Manhattan reported the highest median rent ($19,800) and Bronx the lowest median household income ($35,000).

The numbers in Table 4 support the claim that New York City has a relatively high immigrant concentration. More than a third of the population come from a foreign country and about 63% of the population has at least a parent that was born outside of the U.S. or, in other words, is considered a First Generation U.S. citizen. In the following sections, we will use graphical tools to explore with more detail any relationships between housing affordability and immigration at the sub-borough level.

The graphical tools we will use in the following sections are choropleth maps, linked micromap (LM) plots, box plots overlaid with dot plots, smoothed scatterplots, and common angle plots. Choropleth maps use color for shading the NYC map to represent sub-borough values of a certain variable. Unfortunately, choropleth maps are limited to the depiction of only one variable at a time (Symanzik and Carr, 2008). Therefore, we will also use LM plots. In general, in LM plots, a column of small maps is linked via color to one or more columns with statistical information (Gebreab et al., 2008). In our case, the maps show the sub-boroughs of NYC. The rows in LM plots arranged according to an increasing (or decreasing) order of one of the statistical variables.

The box plots overlaid with dot plots will allow us to have a general perspective of the distribution at a borough and sub-borough level. In the case of the scatterplots, we will use a LOESS (locally estimated scatterplot smoothing) smoother, a non-parametric approach that fits multiple regressions in a local neighborhood to notice a possible association between two variables. Finally, we will use common angle plots, which are a variation of a parallel coordinate plots for categorical data (Hofmann and Vendettuoli, 2013). Some correlations were calculated to provide supporting quantitative information.
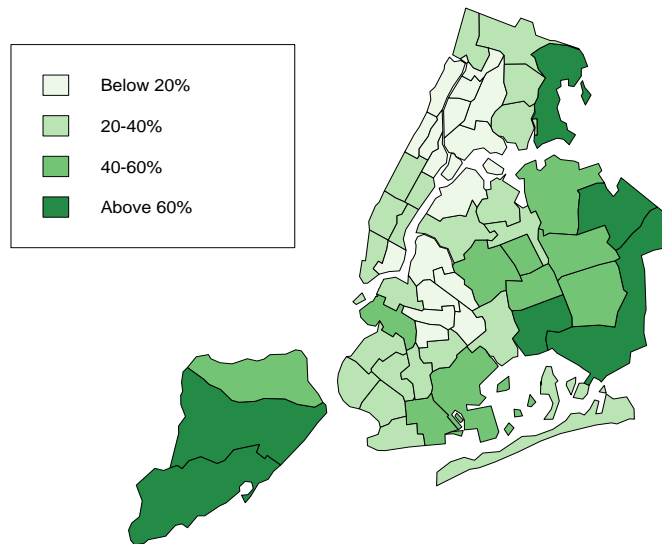
Figure 2: NYC home ownership choropleth map for 2017

3. Home Ownership Exploration at the Sub-Borough Level

Figure 2 shows the home ownership for NYC for 2017. Most of the Staten Island and eastern Queens sub-boroughs have the highest home ownership percentages (above 60%). In addition, the correlation between owning a home and getting a mortgage is strong in New York City ($r = 0.72$), especially in the Brooklyn ($r = 0.75$) and Manhattan ($r = 0.71$) boroughs. In other words, the culture of getting a mortgage to own a home is strong in NYC. Only in the Bronx borough is household income a variable that is moderately associated with owning a home ($r = 0.33$) or having a mortgage ($r = 0.39$).

Regarding the mortgage interest rates, Figure 3a suggests that most sub-boroughs have similar rates and spread. The Manhattan borough has the lowest average rate, which could be explained by a higher household income compared to the other boroughs.
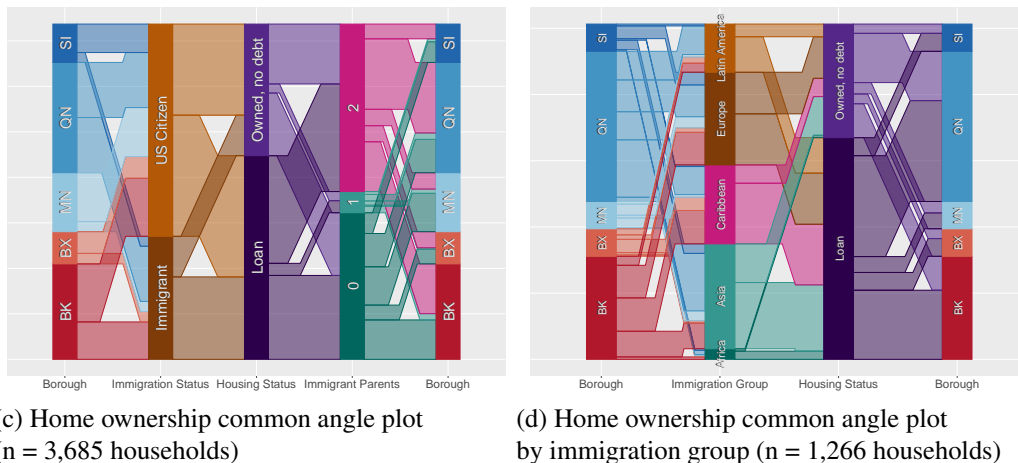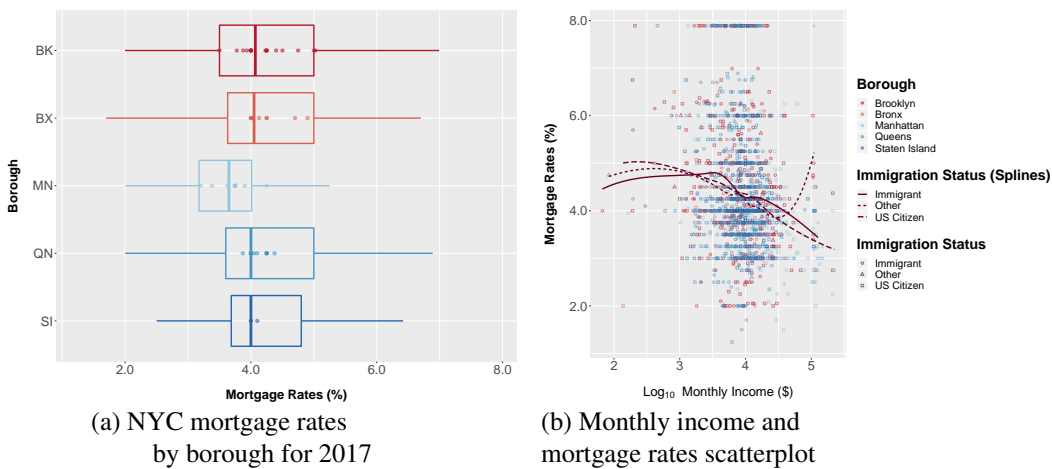
Figure 3b shows, however, that the immigration proportions among New York City report weak association with home ownership variables. Both the US citizens and immigrants splines don't show a noticeable difference.

In addition, Figures 3c and 3d give another general perspective among immigration groups. In Figure 3c, when we compare the immigration and housing status columns, we notice that the immigration proportion of U.S. citizens that own their own home, with or without debt, is greater than the proportion of immigrants. In addition, in Figure 3d, the interaction between the immigration group and the housing status indicates that the Asian and European groups tend to own a home more often than the other groups.

Figure 4 also indicates that immigrant households do not show a clear association with owning a home or having a mortgage. In other words, other variables, such as household income, may be playing a more important role when determining access to a mortgage.

The consideration of other controls could help explain a possible relationship between home ownership and immigration. Factors not considered in our analyses, but in previous research, include intergenerational transfers (Mulder and Smits, 1999) and ethnic variation on investment preferences (Owusu, 1998). Additionally, when looking at the rest of the country, immigrants observed (e.g., education) and unobserved attributes (e.g., ambition), in combination with other socioeconomic and spatial variables, could also influence home ownership (DeSilva and Elmelech, 2012).

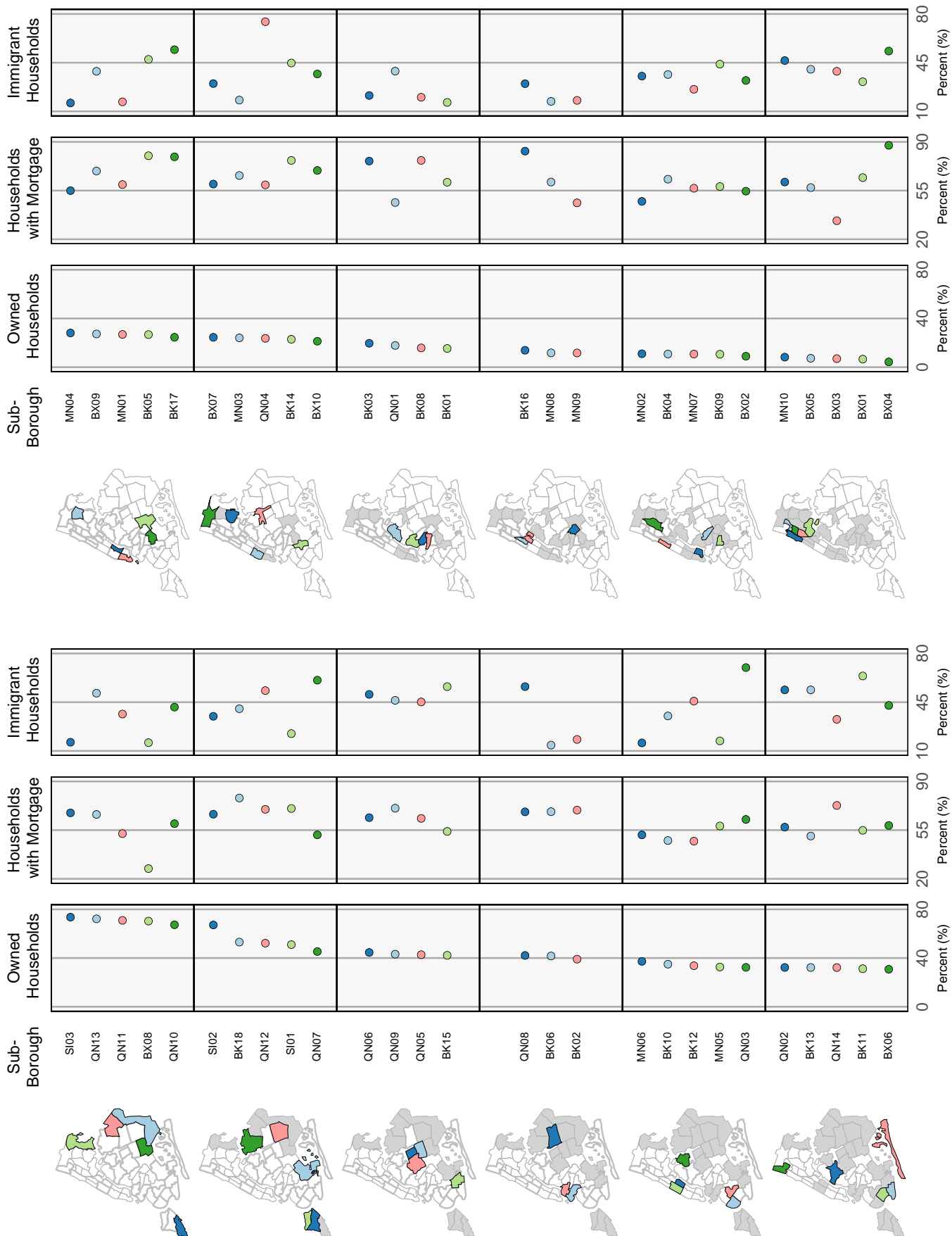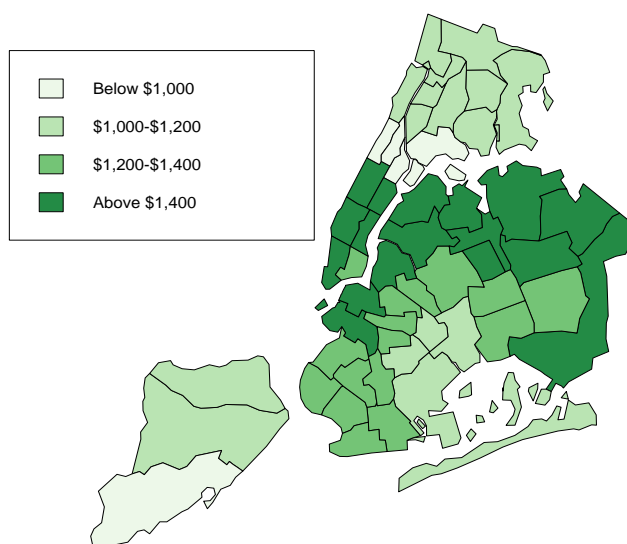**Figure 3**: NYC home ownership visualizations for 2017

(a) NYC mortgage rates
by borough for 2017

(b) Monthly income and
mortgage rates scatterplot

(c) Home ownership common angle plot
(n = 3,685 households)

(d) Home ownership common angle plot
by immigration group (n = 1,266 households)

Figure 4: NYC home ownership LM plot for 2017

Figure 5: NYC household monthly median rent ($) choropleth map for 2017

4. Home Renting Exploration at the Sub-Borough Level

Considering the amount of people that live for free is less than 2.5%, Figure 2 also gives a general perspective of home renting in NYC. Most of the western Bronx and northern Manhattan sub-boroughs tend to rent a place, above 60%. However, Figure 5 indicates that household rent is relatively higher in middle and southern Manhattan sub-boroughs compared to other places in NYC, reaching a median monthly rent above $1,400.

Figure 6a gives a general perspective of the rent spread within each borough. In the case of Manhattan, this borough has the largest sub-borough spread, a bimodal distribution, considering both the lowest- and highest-rent places are in northern and southern Manhattan, respectively. The Bronx borough, in contrast, reports both the smallest spread and average rent.

However, immigration shows weak or no noticeable relation with household rent or income. In Figure 6b we can't notice a clear difference between the US citizen or the immigration group splines. Inclusively, in Figure 7, immigration status apparently do not associate with household income or rent. However, as reviewed previously, because household income may have a stronger association with rent, NYC residents may pay rent based on their earnings rather than their immigration status.

In addition, Figure 7 shows a moderate association between household income and rent. This noticeable relation makes sense when looking at the individual correlations of the Brooklyn ($r = 0.5$), Bronx ($r = 0.36$), Manhattan ($r = 0.46$), and Queens ($r = 0.36$) boroughs. In other words, the culture of renting a place rather than owning, is strong in NYC, and that rent is moderately determined by the household income.

Previous research supports the fact that, even for immigrants, being employed and having sustainable income secure not only stable, but positive housing situations (Shier et al., 2016). Other variables such as social and human capital, adequate social networks support, and foreign-earned credentials and education are also associated with household income and rent (George and Chaze, 2009; Nuesch-Olver, 2002).

Figure 6: NYC home renting visualizations for 2017
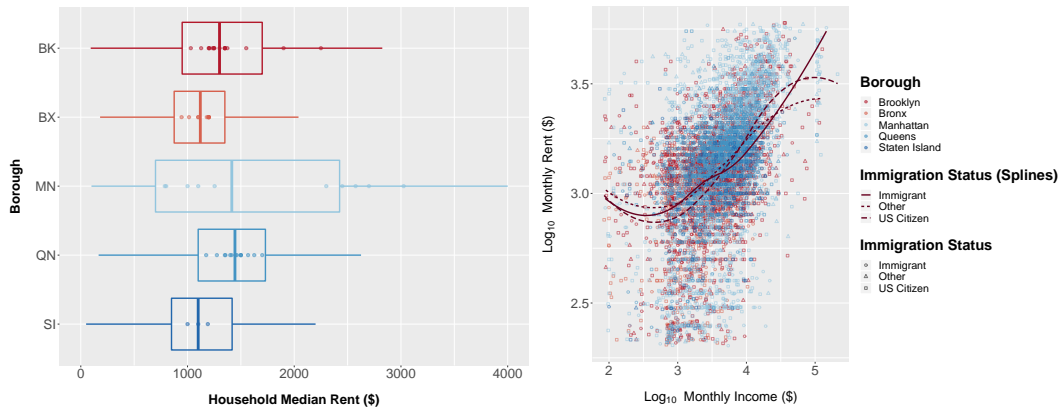
(a) NYC household median monthly rent by borough for 2017

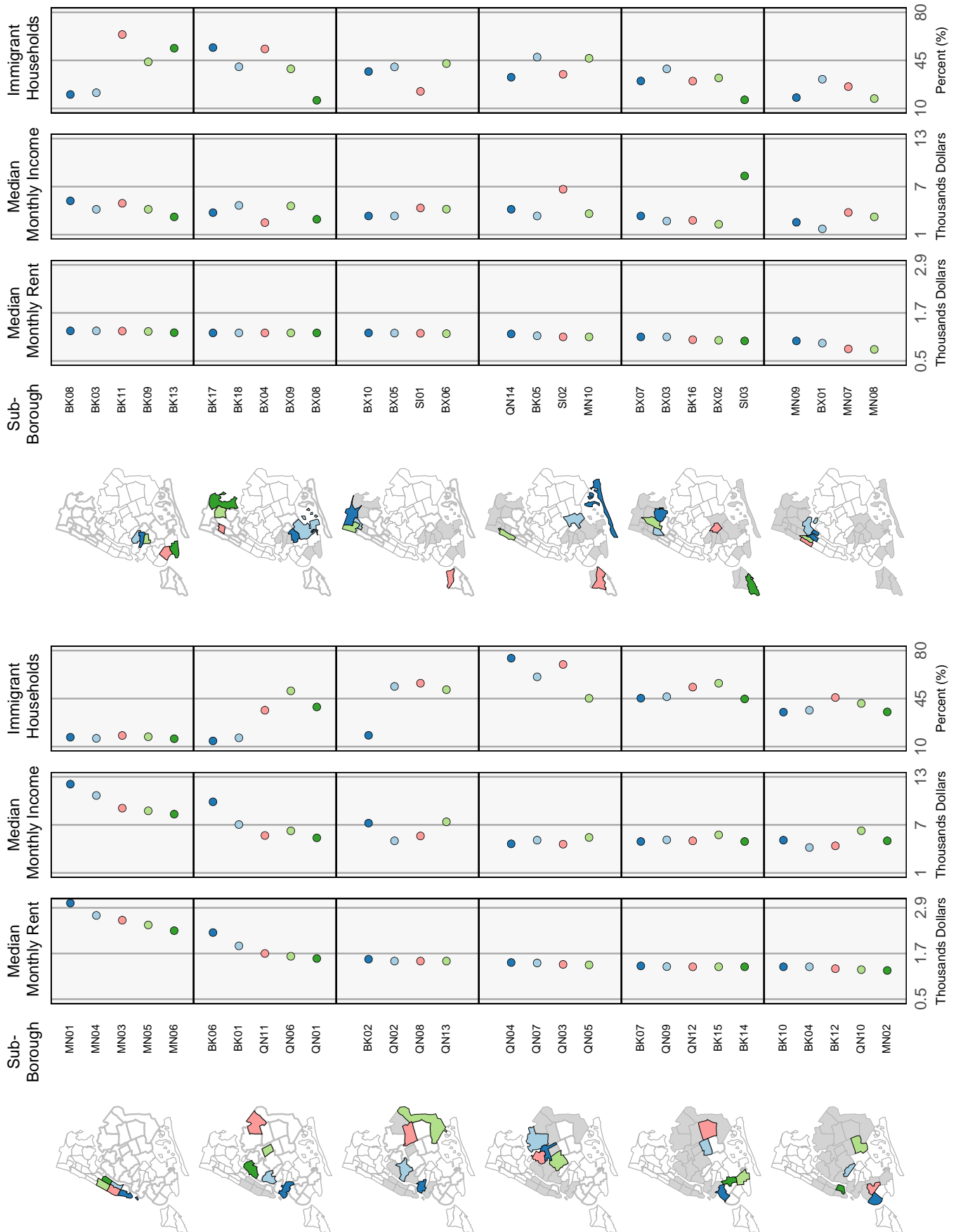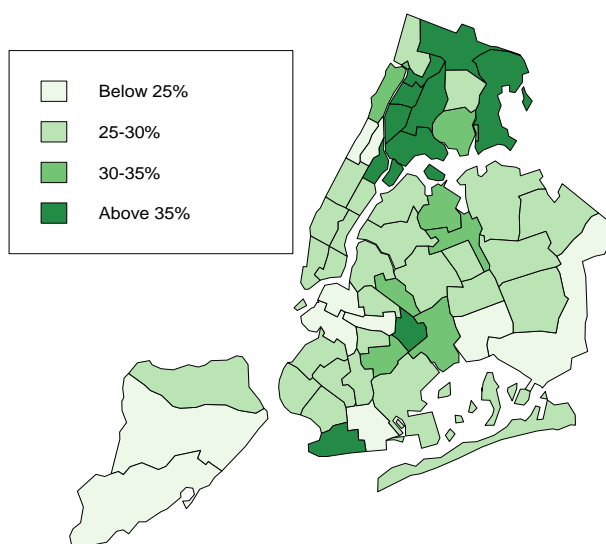(b) NYC monthly income and rent in scatterrplot for 2017

Figure 7: NYC home renting LM plot for 2017

Figure 8: NYC shelter cost (%) choropleth map for 2017

5. Shelter Cost Exploration at the Sub-Borough Level

How much of your income should go to pay rent? The United States Department of Housing and Urban Development (HUD) proposed a percent rule to determine which families are cost-burdened. If a family spend more than 30 percent of their income in rent, these families may report a moderate shelter cost-burden. This means they may have difficulties having access to basic needs such as food, clothing, etc. If this number reaches 50%, the family reports a severe shelter cost-burden.

In New York City, we can draw a general perspective from the data. Figure 8 shows that many Bronx sub-boroughs are the most cost-burdened, with a shelter cost of over 35%. In contrast, most Staten Island and western Queens sub-boroughs report shelter costs below 20%.

Figure 9a shows the shelter cost spread within each main borough. The Bronx borough, in addition to having the highest shelter average cost, also has the highest spread with a noticeable skewness to the right. The Manhattan borough, in contrast, displays both a relatively low average cost and spread compared to the other boroughs, explained most probably by a higher household income.

Figure 9b depicts the relation between shelter cost and rent by both borough and migration status. However, the high variability in the data doesn't allow us to notice an important distinction between US citizens or immigrants. In addition, Figure 10 also reveal a weak association when comparing the statistical panels of shelter cost, household rent, and immigration status.

Figure 9c shows that about 50% of NYC report either moderate or severe shelter cost-burden. Within immigration groups, Figure 9d shows that the Asian group reports the relatively lowest shelter cost when comparing the immigration group and shelter-cost burden columns.

Previous research on this topic indicates that racial/ethnic differentials are too complex to be analyzed generally because a high variability within groups. However, during times of large-scale migration and a shortage of affordable housing, immigrants could experience a higher shelter cost-burden (Elmelech, 2004). In addition, human capital characteristics, stage in life, traditional assimilation, and contextual variables are also associated with a high immigrant shelter cost-burden in the United States (McConnell and Akresh, 2010).
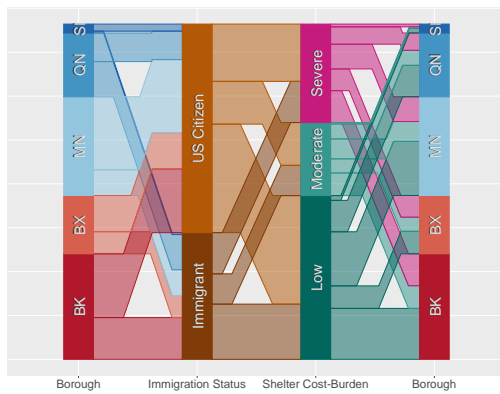
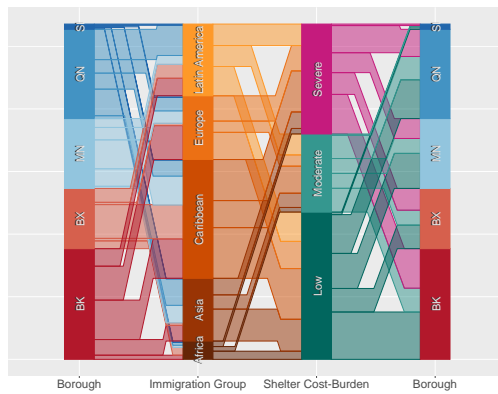Figure 9: NYC shelter cost visualizations for 2017
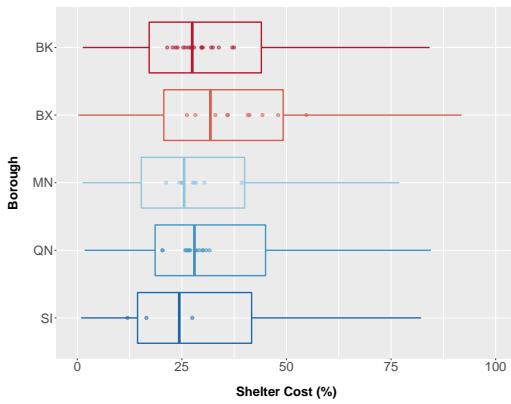
(a) NYC household shelter cost
by borough for 2017
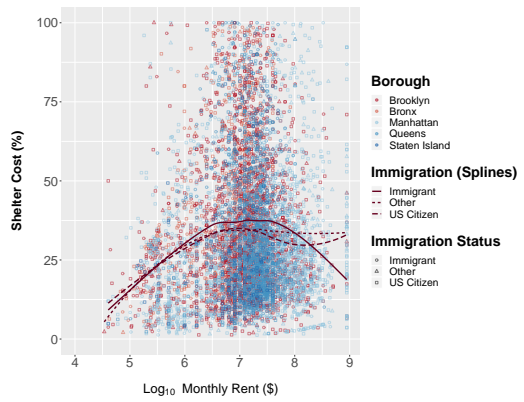
(b) NYC shelter cost and
monthly rent in scatterrplot

(c) Shelter cost common angle plot
(n = 7,641 households)

(d) Shelter cost common angle plot by
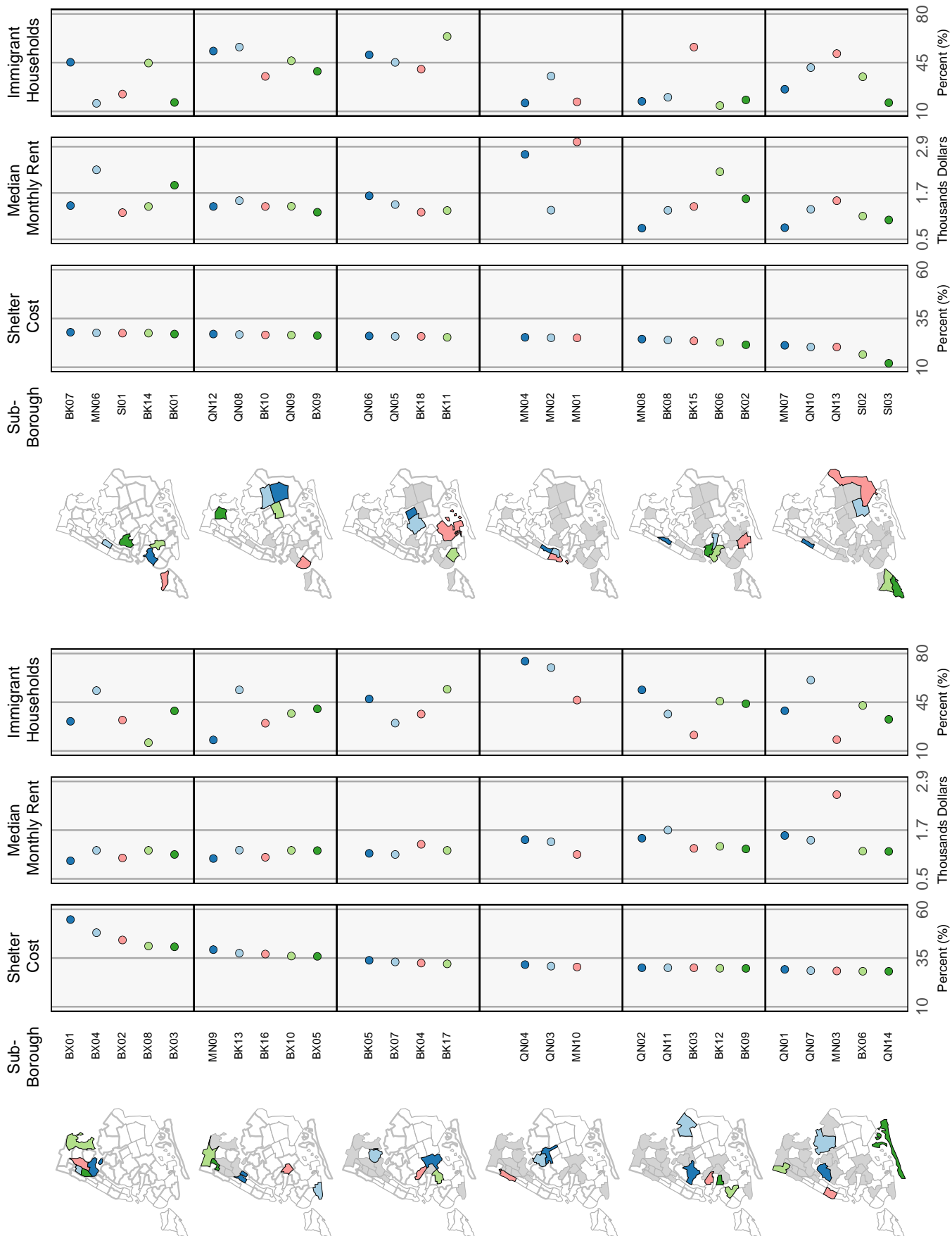immigration group (n = 2,678 households)

Figure 10: NYC shelter cost LM plot for 2017

6. Conclusions

This article provided exploratory analyses of housing affordability variables in New York City. We analyzed the five different boroughs and the 55 sub-boroughs, and considered the residents' immigration status. We selected home ownership, home renting, and shelter cost as the housing affordability variables of interest.

In the case of home ownership, residents who own a home tend to have a mortgage. When comparing immigrants with first generation citizens, the latter group, as expected, is more likely to own a home free of debt. However, mortgage rates are independent from immigration status and have a stronger, positive association with household income. This correlation is stronger in some boroughs, such as Manhattan.

New York City, as a city where renting a home is more popular than owning, shows a moderate association between household rent and income. In contrast, the association of immigration status with household rent or income is weak in most boroughs. The only scenario where immigration shows some association with household rent is in the high-income sub-boroughs in Manhattan which represent the upper quantiles.

Finally, shelter cost, as a social development variable, also provided some general insights. Unfortunately, most boroughs reported a high variability which came with a low correlation with rent, income, and migration status.

Suggestions for further research include modeling these housing affordability variables based on specific income and immigration groups at a sub-borough level. Including more control variables could help identify pattern within certain groups. For example, we notice throughout our visualizations that the Asian group tends to own a home more often and has lower shelter cost compared to other immigration groups.

7. Tools and Acknowledgments

The authors would like to thank the Sections on Statistical Computing, Statistical Graphics, and Government Statistics of the ASA for providing the data used in this analyses. Data manipulations and visualizations made use of the R packages "rgdal" (Bivand et al., 2019), "rgeos" (Bivand and Rundel, 2014), "geojsonio" (Chamberlain and Teucher, 2019), "rmapshaper" (Teucher and Russell, 2018), "sp" (Bivand et al., 2013), "dplyr" (Wickham et al., 2019), "ggplot2" (Wickham, 2016), "micromap" (Payton and Olsen, 2015), "grid" (R Core Team, 2019), and "ggparallel" (Hofmann and Vendettuoli, 2013).

References

Belsky, E., Goodman, J., and Drew, R. (2005). Measuring the Nation's Rental Housing Affordability Problems. Joint Center for Housing Studies of Harvard University (`https://www.jchs.harvard.edu/sites/default/files/rd05-1_measuring_rental_affordability05.pdf`).

Bivand, R., Keitt, T., and Rowlingson, B. (2019). *rgdal: Bindings for the 'Geospatial' Data Abstraction Library*. R package version 1.4-3 (`https://CRAN.R-project.org/package=rgdal`).

Bivand, R. S., Pebesma, E., and Gomez-Rubio, V. (2013). *Applied Spatial Data Analysis with R, Second Edition*. Springer, New York, NY.

Bivand, R. S. and Rundel, C. (2014). *rgeos: Interface to Geometry Engine — Open Source (GEOS)*. R package version 0.3–4 (`http://CRAN.R-project.org/package=rgeos`).

Chamberlain, S. and Teucher, A. (2019). *geojsonio: Convert Data from and to 'GeoJSON' or 'TopoJSON'*. R package version 0.7.0 (`https://CRAN.R-project.org/package=geojsonio`).

DeSilva, S. and Elmelech, Y. (2012). Housing Inequality in the United States: Explaining the White-Minority Disparities in Homeownership. *Housing Studies*, 27(1):1–26.

Elmelech, Y. (2004). Housing Inequality in New York City: Racial and Ethnic Disparities in Homeownership and Shelter-Cost Burden. *Housing, Theory and Society*, 21(4):163–175.

Gebreab, S. Y., Gillies, R. R., Munger, R. G., and Symanzik, J. (2008). Visualization and Interpretation of Birth Defects Data Using Linked Micromap Plots. *Birth Defects Research (Part A): Clinical and Molecular Teratology*, 82:110–119.

George, U. and Chaze, F. (2009). Social Capital and Employment: South Asian Women's Experiences. *Affilia*, 24(4):394–405.

Hofmann, H. and Vendettuoli, M. (2013). Common Angle Plots as Perception–True Visualizations of Categorical Associations. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2297–2305. (`https://cran.r-project.org/web/packages/ggparallel/index.html`).

McConnell, E. D. and Akresh, I. R. (2010). Housing Cost Burden and New Lawful Immigrants in the United States. *Population Research and Policy Review*, 29(2):143–171.

Mulder, C. H. and Smits, J. (1999). First-Time Home-Ownership of Couples the Effect of Inter-Generational Transmission. *European Sociological Review*, 15(3):323–337.

Nuesch-Olver, D. (2002). Thank You for Asking: The Experiences of Latina Immigrant Professional Women. *Social Work and Christianity*, 29(1):31–53.

Owusu, T. Y. (1998). To Buy or Not to Buy: Determinants of Home Ownership among Ghanaian Immigrants in Toronto. *Canadian Geographer/Le Géographe Canadien*, 42(1):40–52.

Payton, Q. C. and Olsen, A. R. (2015). *micromap: Linked Micromap Plots*. R package version 1.9.1 (`http://CRAN.R-project.org/package=micromap`).

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. (`https://www.R-project.org/`).

Shier, M. L., Graham, J. R., Fukuda, E., and Turner, A. (2016). Predictors of Living in Precarious Housing among Immigrants Accessing Housing Support Services. *Journal of International Migration and Integration*, 17(1):173–192.

Symanzik, J. and Carr, D. B. (2008). Interactive Linked Micromap Plots for the Display of Geographically Referenced Statistical Data. In Chen, C., Härdle, W., and Unwin, A., editors, *Handbook of Data Visualization*, pages 267–294 & 2 Color Plates. Springer, Berlin, Heidelberg.

Teucher, A. and Russell, K. (2018). *rmapshaper: Client for 'mapshaper' for 'Geospatial' Operations*. R package version 0.4.1 (`https://CRAN.R-project.org/package=rmapshaper`).

U.S. Census Bureau (2019a). 2013-2017 American Community Survey 5-Year Estimates. Table DP04; generated using American FactFinder; (`http://factfinder2.census.gov`); (25 September 2019). (`https://www.census.gov/quickfacts/newyorkcitynewyork`).

U.S. Census Bureau (2019b). 2013-2017 American Community Survey 5-Year Estimates. Table S0501; generated using American FactFinder; (`http://factfinder2.census.gov`); (25 September 2019). (`https://www.census.gov/quickfacts/newyorkcitynewyork`).

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer, New York, NY. (`https://ggplot2.tidyverse.org`).

Wickham, H., François, R., Henry, L., and Müller, K. (2019). *dplyr: A Grammar of Data Manipulation*. R package version 0.8.3 (`https://CRAN.R-project.org/package=dplyr`).