MEMORANDUM

SUBJECT: Demographic and Social Patterns in Housing Units Near Large Highways and other Transportation Sources

FROM: Chad Bailey, Office of Transportation and Air Quality/Assessment and Standards Division

TO: Docket # EPA-HQ-OAR-2011-0135

I. Executive Summary

I. A. Background

The purpose of this memo is to examine the issue of environmental justice as it relates to housing near roads. EPA defines environmental justice as the fair treatment and meaningful involvement of all people regardless of race, color, national origin, or income with respect to the development, implementation, and enforcement of environmental laws, regulations, and policies. EPA seeks to provide the same degree of protection from environmental health hazards for all people.

Concentrations of many air pollutants are elevated near high-traffic roadways. If minority populations and low-income populations disproportionately live near such roads, then an issue of environmental justice may be present.

This memo explores demographic and socioeconomic differences that exist on a national level between populations near roads and populations that are not near roads. Characteristics of housing units near roads are compared to those not near roads in order to determine if there is a statistically significant difference in demographic or socioeconomic traits between the two groups.

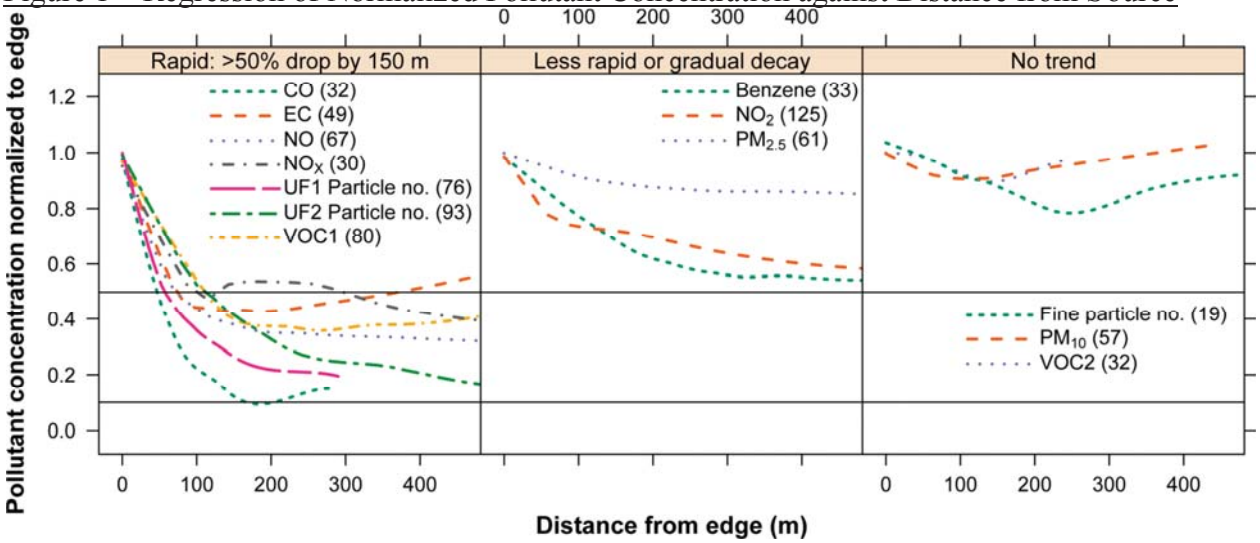
I. B. Method

Data employed in this analysis are from the American Housing Survey (AHS). Every two years, the U.S. Census Bureau conducts the AHS on a national level. The AHS is sponsored by the United States Department of Housing and Urban Development and gathers data about housing units in the United States. In this repeated survey, the same set of housing units is surveyed every odd-numbered year in order to track the changes in households and their housing. Basic information about the unit, resident characteristics, housing and neighborhood quality, size of living space, and financial information are among categories of data gathered in the AHS.[1] The 2009 AHS used in this analysis contained responses from 73,222 households. The Census Bureau makes the data available in several different groups called modules. These modules group the data by type, with similar variables put together into one module. Data from two modules are used in this analysis: "*newhouse*"[2], which contains general information about the households and housing units surveyed; and "*mortg*", which contains mortgage information for different households.

Most statistical analyses were performed using SPSS 10.1.

The variable that will be focused on in the analysis is coded as *etrans* and contains each household's response to the question "How about any railroads, airports or highways with at least 4 lanes – any of these within a half block of this building [your housing unit]?" or 300 feet[3] The Census Bureau reports the presence of these features within 300 feet (about 91 meters) of each housing unit.[4] Concentrations of many pollutants are elevated with reduced distance to roads, with the steepest declines often occurring within about 100 meters. Select pollutants, including NOx and CO, have been found to drop to below 50% of at-road concentration outside 100 meters (m).[5] As seen in Figure 1, for numerous pollutants, the highest concentrations and largest gradients occur within 100 m of roads. As a result, the half block (300 feet) specified in the variable *etrans* will represent a difference in pollutants from on-road emissions. All the statistical analysis detailed in this memo is done with respect to the binary variable *etrans*. The other variables used in the analyses are detailed at the outset of each section.

Figure 1 – Regression of Normalized Pollutant Concentration against Distance from Source



Source: Karner, A. A., Eisinger, D. S., & Niemeier, D. A. (2010). Near-roadway air quality: Synthesizing the findings from real-world data. *Environmental Science & Technology, 44*(14), 5334-5344.

Within SPSS, we weighted the sample data from the American Housing Survey according to the specifications of the Census Bureau.[6] As a result of the weighting, the data shows a representation of the 130,111,607 households in the United States in 2009. The weighting accounts for survey design and non-responses, ensuring all housing units nationally are properly represented in summary statistics.[A] The statistical tests used are two-sample t-tests and logistic regression.

Of the 130,111,607 households represented in the 2009 AHS, the sampling data characterizes households near four-lane highways, railroads, and airports in the following way:

- 22,075,098 households were within a half-block (17.0%)

[A] While these analyses factor in survey weights recommended by the Census Bureau, the information required to incorporate the survey design into the estimate of variances was unavailable in the AHS dataset. We contacted the preparers of the codebook and dataset about this issue, but received no responses. Without the ability to perform a redistribution sample on the data, there is no way to recalibrate the variance for the population. The variance is thus lower than it should be to properly represent the national housing data. Consequently, summary statistics may underestimate standard errors and underestimate heterogeneity in national housing stock.

- 104,936,084 were not (80.7%)
- 284,953 were represented by samples that didn't know (0.2%)
- 391,903 were represented by samples that refused to answer (0.3%)
- 2,423,569 were represented by samples that were not reported (1.9%)

The sum of the weight for each category was used to determine the population for each value of the variable.

I. C. Key Findings

The statistical analyses described below indicate significant differences between average characteristics of households residing within a half block of railroads, airports, or four-lane highways and characteristics of those residing outside that range. Specifically, minority householders, householders with lower educational attainment, and lower-income households are all more likely to be within a half block of railroads, airports, or four-or-more-lane (4+ lane) highways than households not fitting those characteristics.

II. T-tests For Significance

II. A. Introduction

We used two-sided Student's t-tests assuming independent samples to determine the significance of pairwise contrasts in various demographic and economic characteristics between households within or outside 300 feet from railroads, airports, or 4+lane highways (the *etrans* response). [7]

The raw results of the t-tests performed in this section are in Appendix A.

II. B. Variables

The variables which will be examined using a t-test are *ammort*, *ammrt2*, *lot*, *unitsf*, *cars*, *value*, *lmed*, *zinc2*, and *zinc*. These coded names represent the following variables:

- *ammort* – amount of first mortgage when acquired
- *ammrt2* – amount of second mortgage when acquired
- *lot* – area of lot in square feet
- *unitsf* – area of the unit in square feet
- *cars* – number of cars kept for use of members of the household
- *value* – current market value of the unit
- *lmed* – average area median family income
- *zinc2* – household income
- *zinc* – family income

We examined these variables to separately control for differences in housing unit characteristics (*unisf, value*), site characteristics (*lot*), metropolitan statistical area economic patterns (*lmed*), family and household income characteristics (*zinc2, zinc*), and household financing (*ammort, ammrt2*).

II. C. Results

Detailed statistics and t-test results for each variable are presented in Appendix A.

Overall, households with a positive *etrans* response, that is, those located within a half block of a railway, airport, or four-lane highway, had worse economic situations. Households with positive *etrans* responses, in univariate comparisons, have lower average household income, family income, smaller homes and lots. If they are own their homes, they have smaller first and second mortgages. Lastly, housing units with a positive *etrans* response live in metropolitan areas (MSAs) with slightly higher average median incomes.

The results suggest that personal economic situations and regional economic trends affect the odds that a housing unit is located within a half block of a railway, airport, or four-lane highway (*etrans* response).

Reflecting personal economic situations, households that reside away from such sources have, on average, higher incomes, more durable assets (e.g., cars), larger homes, larger lots, and, if they own their homes, have greater mortgage debt. This suggests that people living near railways, airports, and 4+ lane highways have lower socioeconomic status (SES).

Regional economic trends also appear to influence the likelihood of living near a transportation sources, with housing units within half a block having greater average median income at the MSA level. These differences will be examined further in the following sections.

I. A. Summary

II. Univariate Logistic Regression

II. A. Introduction

In order to investigate the significance of binary variables in relation to the variable *etrans*, the focus of this analysis, we have applied logistic regression to examine whether various qualitative factors change the odds of living within a block of a railway, airport, or 4+ lane highway. These results are expressed in terms of how each variable affects the "odds ratio" (OR) of an affirmative *etrans* response. An odds ratio reflects how the chance that a given outcome happens depends an independent variable (e.g., race, income). [B]

II. B. Variables

[B] An odds ratio (OR) is a value which indicates how much more likely the success of a binary response variable is when the value of a certain variable increases by one. For the purposes of this analysis, a "success" in the variable *etrans* is when a household is located within a half block of an airport, railway, or four-lane highway. This is a convenient method to use with binary explanatory variables; as the only possible values are 1 and 0, an increase of one represents the presence of that characteristic. That is, in this setting, the odds ratio will relate how much more likely a success in the variable *etrans* is when a given variable characteristic is present. The odds ratio will indicate how the corresponding characteristic affects the odds of a household being within a half block of a four-lane highway, railway, or airport by multiplying the original odds when that characteristic is present. Thus, if the odds ratio is greater than one, greater odds are produced, and vice versa. The raw SPSS output for this section can be located in Appendix B.

The binary variables examined using this method are described below:

- Householder race and ethnicity
  - *hhspan* – whether or not the householder is Hispanic
  - *white* – whether or not the householder is white
- Household socioeconomic status
  - *hhhsgrad* – whether or not the householder graduated from high school
  - *unigrad* – whether or not the householder graduated from a four-year university
  - *hhwlineq* – whether or not the householder worked in the last week
  - *qfs1* – whether or not the household received foods stamps in the past year
  - *qwelf* – whether or not the household received welfare in the past year
  - *qdiv* – whether or not the household received stock dividends in the past year
- Housing unit characteristics
  - *proj* – whether or not the unit is owned by a public housing authority
  - *garage* – whether or not a garage or carport is included with the unit

For the most part, these variables represent demographic and economic characteristics of the householder, or the house itself.

II. C. Results

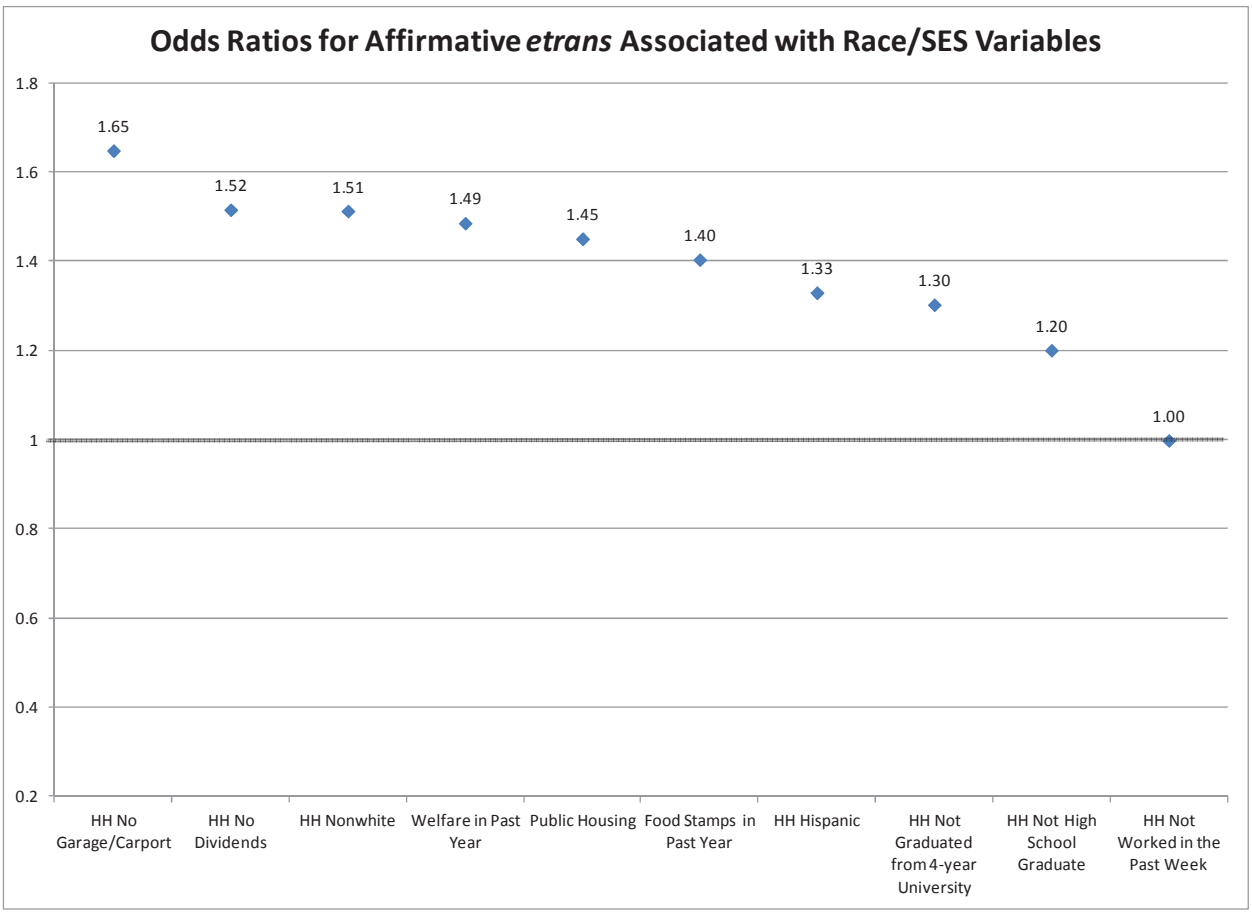
The detailed tables for each variable examined are presented in Appendix B of this memorandum, along with SPSS output tables with logistic regression results.

Overall, every variable described in Section II.B had a statistically-significant OR (OR), suggesting that the demographic and economic characteristics analyzed affect whether a household lives near a major transportation source.

Figure II-1 summarizes the ORs for univariate regressions, ranking them by magnitude.[C] As shown, nearly every EJ-related variable results in a substantial increase in the likelihood that a housing unit will be located in a location where traffic-related air quality is a concern. The only SES-related variable that results in little change in the odds of a positive *etrans* response is *hhwlineq*, which indicates whether the householder worked in the past week. This pattern suggests that living in a location with greater potential traffic-related air pollution is associated more with individual factors (e.g., race and ethnicity) and economic factors (e.g., educational attainment, housing unit features) that reflect longer-term issues.

Figure II-1

[C] In Figure 1, odds ratios are displayed on the basis of an affirmative response to EJ-related questions (e.g., race, ethnicity, and SES). As such, a variable with an odds ratio >1 indicates potentially greater traffic-related air quality problems for those responders.



III. Multivariate Logistic Regression Model

III. A. Introduction

Multivariate logistic regression is a method of finding a relationship between an OR (OR) of a binary variable resulting from a number of explanatory variables. To examine the effect of any particular variable, logistic regression allows for the control of other variables, aiding in interpretation.

This section presents the results of four logistic regression models: one model employing a small number of EJ-related variables, one including a moderate number of variables, one including many variables that relate to both demographic and socioeconomic characteristics, and the last one readjusted to remove cases of multicollinearilty from the large model. The raw SPSS output for this section can be located in Appendix C.

III. B. Minimal EJ Logistic Regression Model

Out of the 73,222 households surveyed, 52,217 indicated an answer to *etrans.* Executive Order 12898 stipulates that all federal agencies consider the impacts of their actions affecting "minority populations and low-income populations."[8] Thus, adhering strictly to the characteristics

specifically enumerated in the E.O., this model only includes the variables *hhspan* (householder Hispanic status); *white* (householder white/nonwhite); and *zinc2* (household income).[D]

In performing the regression, SPSS identifies which cases are valid to use. Due to the use of survey weights to make the data representative of the general population, cases with a missing or zero weight value were removed. Of the 73,222 households surveyed, 53,354 have a useable weight value. 9,385 of those cases were removed for having missing values in at least one of the explanatory variables or the response variable, leaving 43,969 cases to be weighted and analyzed as part of the regression.

According to the model, the three explanatory variables included all had significant effect on whether or not a living unit is within a half block of a four-lane highway, a railway, or an airport.

We assessed fit using the Hosmer and Lemeshow test. However, a known limitation of this test is that for large sample sizes, most of these tests result in a significant "bad fit." As the analysis contain in this section has a total weighted population of 130,111,607 households, this large sample size may contribute to the significance found in the Hosmer-Lemeshow tests.[9]

Table III-1 shows correlation coefficients between the variables. For correlation between two binary variables, *hhspan* and *white*, a phi coefficient was used. For correlation between *zinc2* and each of the two binary variables, a point-biserial correlation coefficient was used.

Table III-1

Correlation	Household Income	White Householder
Hispanic Householder	-0.074	0.103
White Householder	0.092	

The largest correlation was between the variables white householder and Hispanic householder, but all correlations were small and unlikely to affect the model's outputs due to their low multicollinearity.

The results of the logistic regression, that is, the variables and their corresponding ORs, are shown in Table III-2.

Table III-2

Variable	OR	95% Confidence Interval
Constant	0.283	-
Hispanic Householder	1.332	1.331-1.334

[D] *zinc2* was included in the model in a modified form. First, as *zinc2* is a variable expressed in dollars, coefficients produced for this variable would represent the odds ratio for a change in income of $1. In order to have observable differences, the values of the variable were divided by 10,000 so that the variable was expressed in terms of $10,000. Also, to make the constant more representative, the variable was centered at the average value of $66,200 by subtracting 6.62 from each value in the variable. The value of the constant, which represents the odds of a success if each variable is 0, then represents the odds of success in the case of the family income being $66,200 instead of $0.

White Householder	0.678	0.677-0.679
Household Income (10,000s)	0.962	0.961-0.962

As described above, the income-based OR is based on how each household's income differs from $66,200, the average level reported in the AHS. The "constant" represents odds of a household with a non-Hispanic, nonwhite householder with a average household income living within a half block of a four-lane highway, a railway, or an airport (i.e., an affirmative *etrans* response). These odds are 0.283, which equates to roughly a 22% chance of an affirmative *etrans* response. If the householder is white, then the odds decrease to 0.192, with probability of an affirmative *etrans* response of about 16%. If the householder is Hispanic and nonwhite, the odds increase to 0.377, with a probability of *etrans* of about 27%. For housing units with Hispanic householders, the odds of living within a half block of an affirmative *etrans* response are about twice that of a non-Hispanic white householder. Also, a change of $10,000 in household income results in a proportional change in odds of 4%.

Compared with the univariate analysis above, the OR for housing units with Hispanic and nonwhite householders are nearly identical, suggesting that they are confounded by neither one another nor by income.

The results of this model suggest that race/ethnicity and income status both have a significant effect on the odds of a household living within a half block of a four-lane highway, a railway, or an airport.

III. C. Medium-level Logistic Regression Model

To address the potential for confounding by other factors and to determine which factors contribute to the probability of living near railway, airport, or 4+lane highway, we developed a regression model including more variables. In addition to the variables included in the above model, the following variables were used: whether the household had received welfare in the past year, whether the household had received stock dividends in the past year, number of bedrooms in the housing unit (difference from 3), and indicator variables representing the type of unit that the household lives in (*nunit2*). The Hosmer and Lemeshow test produced a chi-square statistic of 1598.2 for this model, indicating significance in the test.

Table III-3 shows correlation coefficients between the variables. For correlation between two binary variables, a phi coefficient was used. For correlation between quantitative variables and binary variables, a point-biserial correlation coefficient was used. The correlation between the binary variables and the nominal variable *nunit2*, reports Cramer's V as the measure. The significance of contrasts in household income (*zinc2*) among unit types in the nominal housing unit type variable *nunit2* is computed using the Scheffe method as a part of post-hoc analysis, and is reported below the table.

Table III-3

Correlation	Unit Type	Received Stock Dividends in Past Year	Received Welfare in Past Year	Household Income	White Householder

Hispanic Householder	0.111	0.086	0.038	-0.074	0.103
White Householder	0.175	0.085	0.072	0.092	
Household Income	*	-0.080	0.190		
Received Welfare in Past Year	0.085	0.038			
Received Stock Dividends in Past Year	0.105				

*Scheffe post-hoc test indicates significant differences for each contrast between unit types.

The largest correlation was between the variables household income and Welfare in Past Year. However, 0.190 is a very low correlation coefficient. None of the variables in this model are very correlated with the others.

The variables and their corresponding ORs are as shown in Table 36.

Table III-4

Variable	OR	95% Confidence Interval
Constant	0.200	-
Hispanic Householder	1.166	1.164-1.168
White Householder	0.783	0.782-0.784
Household Income (10,000s)	0.983	0.983-0.983
Received Welfare in Past Year	1.022	1.019-1.026
Received Stock Dividends in Past Year	0.864	0.862-0.865
Bedrooms	0.865	0.865-0.866
Unit Type		
One-unit building, attached	1.364	1.361-1.367
Building with 2+ apartments	2.028	2.025-2.031
One-unit Mobile Home	1.158	1.156-1.161

The value of the constant represents the odds of living within a half block of a four-lane highway, a railway, or an airport for a household with a non-Hispanic, nonwhite householder with a household income of $66,200 that hasn't received welfare or stock dividends in the past year, in a one-unit, detached building with three bedrooms. These odds are 0.198, which equates to roughly a 16.5% probability of an affirmative *etrans* response. The odds change depending on the differing characteristics present and their corresponding ORs, above.

III. D. Summary of Logistic Models

Comparisons between the univariate ORs, the "minimal" multivariate model, and the larger model containing more variables, there are several notable commonalities. First, the most directly EJ-related variables, race, Hispanic status, and income, maintain statistically significant ORs. Second, between the univariate and "minimal" multivariate models, the ORs (and associated confidence intervals) for race and ethnicity change very little in magnitude,

suggesting that they are confounded neither by one another, nor by income. Third, comparing the expanded model, including housing unit characteristics and added economic indicators, with the "minimal" and univariate models, the ORs for Hispanic status and race are attenuated to some extent, though they remain significant. Fourth, comparing the "minimal" model with the expanded model, the OR for household income remains very similar, despite the inclusion of indicators of high-income and low-income status (e.g., receiving dividend income and welfare).

The above observations suggest that income may be a non-linear effect on the probability of residence near an airport, railway, or 4+ lane highway, with larger effects at the "tails" of the income distribution. Though nonlinear income terms were not examined in this analysis, we plan to examine them in forthcoming analyses.

Furthermore, it appears that accounting for very high and very low income indicators, in addition to housing unit characteristics (e.g., bedroom number and unit type), account for the attenuation in race and ethnicity indicators. This phenomenon will be examined in greater detail in subsequent analyses.

IV. Discussion

The above analyses generally support a conclusion that households near airports, railways, and 4+ lane highways are more likely to be of a racial or ethnic minority and lower in income. People on welfare are also more likely to live near such large transportation sources. Housing units near facilities named in *etrans* are also smaller and more often attached structures (e.g., multi-unit dwellings).

In subsequent analyses, we plan to address other issues that affect how demographics and SES affect responses to *etrans*.

.

[1] U.S. Census Bureau, Housing and Household Economic Statistics Division. (2005, February 18). Overview. In *American Housing Survey (AHS)*. Retrieved from http://www.census.gov/hhes/www/housing/ahs/overview.html

[2] It should be noted that the name "newhouse" refers to a new grouping of housing variables, not data on newly constructed or newly included housing units in the survey.

[3] Etrans. (2011). In *Codebook for the American Housing Survey, public use file: 1997 and later* (p. 576).

[4] U.S. Department of Housing and Urban Development, & U.S. Census Bureau. (n.d.). Age of other residential buildings within 300 feet. In *American Housing Survey for the United States: 2009* (pp. A-1). Retrieved from http://www.census.gov/hhes/www/housing/ahs/ahs01/appendixa.pdf

[5] Karner, A. A., Eisinger, D. S., & Niemeier, D. A. (2010). Near-roadway air quality: Synthesizing the findings from real-world data. *Environmental Science & Technology, 44*(14), 5334-5344.

[6] Appendix I: AHS data users FAQs. (2011). In *Codebook for the American Housing Survey, public use file: 1997 and later* (p. 1255).
It should be noted that this instruction only included readjustment due to weights, and did not include a system for replication of the survey. This could result in lower variances than is accurate and thus findings may be for significant than they are in reality.

[7] When there is an indication that one sample may have a higher mean value for a certain variable than another sample, the t-test is preformed looking for that difference in particular. This indication is typically tied into the hypothesis, and can result in statistical significance being easier to achieve.

[8] Exec. Order. No. 12898, 59 Fed. Reg. 32 (1994, February 16), http://www.archives.gov/federal-register/executive-orders/pdf/12898.pdf.
[9] Garcon, G. D. (2011, April 27). Significance tests. In *Logistic Regression*. Retrieved July 25, 2011, from http://faculty.chass.ncsu.edu/garson/PA765/logistic.htm

Appendix A: T-test Data and Results

A.1 Data Summary

Mortgage

t-tests were run to determine if there were significant differences in the values of mortgages for households located within/outside of a half block of a railway, airport, or four-lane highway. One t-test examined the initial value of the first mortgage (variable *ammort*), and another t-test examined the initial value of the second mortgage (variable *ammrt2*). As not every household assumes one or more mortgage, there are missing values for each mortgage measurement. Households without mortgages may be renters or own their homes outright. Out of a total of 130,111,607 households represented in the AHS, 47,995,888 households have values for the first mortgage with 82,115,719 missing, and 5,517,534 values for the second mortgage with 124,594,073 values missing. Additional cases are left out due to missing values in the variable *etrans*. The summary statistics of these groups are shown in Tables A-1 and A-2 below:

Table A-1

First Mortgage	Original Sample Size	Weighted Sample Size	Weighted Mean	Weighted Standard Deviation
Within ½ block of airport/railway/ highway	2488	6,339,776	$147,719.59	$135,168.162
Outside ½ block of airport/railway/ highway	15,788	40,410,963	$168,015.23	$154,991.759
Difference (within-outside) (t-value)			-$20,295.63 (t=-344.23)	

Table A-2

Second Mortgage	Original Sample Size	Weighted Sample Size	Weighted Mean	Weighted Standard Deviation
Within ½ block of airport/railway/ highway	286	741,548	$52,521.57	$70,785.283

Outside ½ block of airport/railway/ highway	1874	4,733,338	$57,022.88	$72,479.515
Difference (within-outside)			-$4,501.30 (t=-50.75)	

Households that are not within a half block of an airport, railroad, or four-lane highway assume a larger first and second mortgage, on average. Table A-1 shows the difference in first mortgages between the two groups (95% CI: -$20,180 - -$20,411). Table A-2 shows the difference in second mortgage values (95% CI: $4327-$4675).

Lot and Unit Area

The variable measuring total lot area (*lot*) has 95,876,348 valid households analyzed as part of the mean of either near road households or non-near road households. Values of this variable less than 200 ft^2 were listed as 200 ft^2, and values over 999,997 ft^2 were recorded as 999,997 ft^2.[E,10]

Unit area for each household (*unitsf*), which gives a measure of total interior livable area, had 119,404,448 valid responses and 10,707,159 missing. Values under 99 ft^2 were coded at 99 ft^2, and values at over 24,870 ft^2 were recorded at 24,870 ft^2.

The summary statistics of the groups of the two different types of area are shown below in Tables A-3 and A-4:

Table A-3

Lot Area	Original Sample Size	Weighted Sample Size	Weighted Mean	Weighted Standard Deviation
Within ½ block of airport/railway/ highway	4983	12,532,783	42,437.64 ft^2	130,819.283 ft^2
Outside ½ block of airport/railway/ highway	31,890	81,021,724	76,202.05 ft^2	191,377.678 ft^2
Difference (within-outside)			-33,764 ft^2 (t=-791.98)	

[E] Setting lower and upper limits for a variable's value is known as bottom-coding and top-coding, respectively. It is done to protect the anonymity of the households taking the survey. Outliers could potentially be identified using location data also available as part of the AHS, and thus the confidentiality of their survey responses would be compromised.

Table A-4

Unit Area	Original Sample Size	Weighted Sample Size	Weighted Mean	Weighted Standard Deviation
Within ½ block of airport/railway/ highway	8378	20,345,461	1548.24 ft^2	2072.338 ft^2
Outside ½ block of airport/railway/ highway	38,609	96,458,011	1906.16 ft^2	2240.453 ft^2
Difference (within-outside)			-357.92 ft^2 (t=-697.77)	

t-tests were run again with the same hypothesis and variables, only with the top-coded and bottom-coded values removed. Qualitatively, results were similar and statistically significant, with mean lot size difference of -18,389 ft^2 and unit area difference of 335.13 ft^2.

Vehicle Ownership

A t-test was performed to determine whether there was a significant difference in number of cars (*cars* variable) for households with positive and negative *etrans* responses.

Of the 130,111,607 households in the United States in 2009, there were 18,250,711 households represented by the response "Not Applicable" and thus were left out of the analysis. The means were calculated from a split sample of a total of 111,860,896 households. The variable cars is top-coded at 5, indicating that 5 or more cars are available for household use.

Table A-5 shows summary statistics by proximity to transportation sources.

Table A-5

Cars	Original Sample Size	Weighted Sample Size	Weighted Mean	Weighted Standard

				Deviation
Within ½ block of airport/railway/ highway	8004	19,621,627	1.123	0.872
Outside ½ block of airport/railway/ highway	35,965	89,422,881	1.265	0.930
Difference (within-outside)			0.142 (t=-645.31)	

Housing Unit Value

A t-test was performed to determine whether there was a significant difference in *value*, the variable indicating the unit value for households with different *etrans* responses.

The variable value contains the estimated market value of the unit at the time of the survey. This variable only includes housing units which are owner-occupied or is otherwise non-rental. 87,944,404 households are represented by the responses in the survey. Table A-6 presents the summary statistics:

Table A-6

Market Value of the Unit	Original Sample Size	Weighted Sample Size	Weighted Mean	Weighted Standard Deviation
Within ½ block of airport/railway/ highway	4784	11,844,689	$201,147.67	$224,696.916
Outside ½ block of airport/railway/ highway	30,082	76,099,715	$243,271.65	$293,497.539
Difference (within-outside)			-$42,123.98 (t=573.53)	

Income

Separate tests were performed to examine differences in responder-level income (family and household income) and metropolitan statistical area-level income. These tests reflect economic issues at different geographic scales.

The summary statistics and t-test results are displayed in Table A-7, Table A-8, and Table A-9 for each of the three types of income (household, family, and median MSA):

Table A-7

Household Income	Original Sample Size	Weighted Sample Size	Weighted Mean	Weighted Standard Deviation
Within ½ block of airport/railway/ highway	8004	19,621,627	$53,236.85	$56,373.439
Outside ½ block of airport/railway/ highway	35,965	89,422,881	$67,882.66	$68,911.201
Difference (within-outside)			-$14,645.81 (t=-998.68)	

Table A-8

Family Income	Original Sample Size	Weighted Sample Size	Weighted Mean	Weighted Standard Deviation
Within ½ block of airport/railway/ highway	9305	22,075,098	$44,896.56	$54,558.41
Outside ½ block of airport/railway/ highway	42,912	104,936,084	$55,792.51	$67,048.362
Difference (within-outside)			-10,895.95 (t=-817.42)	

Table A-9

Average Area Median Income	Original Sample Size	Weighted Sample Size	Weighted Mean	Weighted Standard Deviation
Within ½ block of airport/railway/ highway	9305	22,075,098	$65,765.70	$11,537.298
Outside ½ block of airport/railway/ highway	42,912	104,936,084	$64,801.81	$11,712.817
Difference (within-outside)			$963.89 (t=355.85)	

The difference in the average incomes of individual households (family or household income) within a half block of a railway, airport, or 4+ lane highway from those outside half a block suggests that personal economic situations affect the likelihood of living near transportation sources.

As indicated in Table A-9, MSA-level median incomes tend to be higher for housing units within half a block of these transportation sources. This finding suggests that regional economic patterns influence the likelihood of homes being built near these transportation sources.

A.2 Summary Statistics and T-Tests

Table A-12

Group Statistics

	Within 1/2 Block of Transport	N	Mean	Std. Deviation	Std. Error Mean
mort1	Yes	6339776	147719.59	135168.162	53.683
	No	40410963	168015.23	154991.759	24.381
mort2	Yes	741548	52521.57	70785.283	82.200
	No	4733338	57022.88	72479.515	33.314
Lot Area	Yes	12532783	42437.64	130819.283	36.953
	No	81021724	76202.05	191377.678	21.261
Unit Area	Yes	20345461	1548.24	2072.338	.459
	No	96458011	1906.16	2240.453	.228
Lot Area Recode	Yes	12368995	31235.08	80157.897	22.792
	No	78703093	49623.61	111339.848	12.550
Unit Area Recode	Yes	20200023	1404.81	961.483	.214
	No	95657432	1739.93	1081.621	.111
Cars	Yes	19621627	1.12	.872	.000
	No	89422881	1.27	.930	.000
Market Value	Yes	11844689	201147.67	224696.916	65.288
	No	76099715	243271.65	293497.539	33.644
Household Income	Yes	19621627	53236.85	56373.439	12.726
	No	89422881	67882.66	68911.201	7.287
Area Median Income	Yes	22075098	65765.70	11537.298	2.456
	No	1.0E+08	64801.81	11712.817	1.143
Family Income	Yes	22075098	44896.56	54558.410	11.612
	No	1.0E+08	55792.51	67048.362	6.545

For Table A-2 and Table A-3, "Levene's Test for Equality of Variances" refers to a significance test performed to determine whether it can be assumed that the variances for each subgroup (the explanatory variable divided based on *etrans*) are equal. The null hypothesis being tested is that they are equal. As each test found a significant difference in the values of the variances, the second row ("Equal variances not assumed") in each variable should be the one referenced.

Table A-2

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
mort1	Equal variances assumed	62273.623	.000	-311.641	46750738	.000	-20295.63	65.125	-20451.8	-20139.4
	Equal variances not assumed			-344.225	9163808	.000	-20295.63	58.960	-20411.2	-20180.1
mort2	Equal variances assumed	2858.160	.000	-49.883	5474884	.000	-4501.30	90.237	-4678.165	-4324.441
	Equal variances not assumed			-50.751	1000927	.000	-4501.30	88.695	-4675.141	-4327.464
Lot Area	Equal variances assumed	935099.8	.000	-603.168	93554505	.000	-33764.41	55.978	-33903.9	-33624.9
	Equal variances not assumed			-791.982	2.2E+07	.000	-33764.41	42.633	-33848.0	-33680.9
Unit Area	Equal variances assumed	82619.319	.000	-663.230	1.2E+08	.000	-357.92	.540	-359.268	-356.581
	Equal variances not assumed			-697.770	3.1E+07	.000	-357.92	.513	-359.044	-356.805
Lot Area Recode	Equal variances assumed	594759.1	.000	-558.546	91072086	.000	-18388.53	32.922	-18470.5	-18306.5
	Equal variances not assumed			-706.740	2.1E+07	.000	-18388.53	26.019	-18439.5	-18337.5
Unit Area Recode	Equal variances assumed	397453.9	.000	-1289.139	1.2E+08	.000	-335.13	.260	-335.775	-334.479
	Equal variances not assumed			-1391.600	3.2E+07	.000	-335.13	.241	-335.599	-334.655

Table A-3

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Cars	Equal variances assumed	415982.4	.000	-619.349	1.1E+08	.000	-.14	.000	-.143	-.141
	Equal variances not assumed			-645.306	3.0E+07	.000	-.14	.000	-.142	-.142
Market Value	Equal variances assumed	222350.7	.000	-472.856	87944402	.000	-42123.98	89.084	-42328.3	-41919.6
	Equal variances not assumed			-573.526	1.9E+07	.000	-42123.98	73.447	-42267.9	-41980.0
Household Income	Equal variances assumed	545056.8	.000	-879.101	1.1E+08	.000	-14645.81	16.660	-14687.3	-14604.3
	Equal variances not assumed			-998.681	3.4E+07	.000	-14645.81	14.665	-14676.4	-14615.2
Area Median Income	Equal variances assumed	7471.322	.000	352.358	1.3E+08	.000	963.89	2.736	957.086	970.695
	Equal variances not assumed			355.846	3.2E+07	.000	963.89	2.709	958.583	969.198
Family Income	Equal variances assumed	738039.1	.000	-715.337	1.3E+08	.000	-10895.95	15.232	-10933.8	-10858.1
	Equal variances not assumed			-817.419	3.8E+07	.000	-10895.95	13.330	-10922.1	-10869.8

Appendix B: Univariate Logistic Survey Data and Regression Results

This appendix presents the survey data from the 2009 AHS and the SPSS outputs from the logistic regression procedure. Section B.1 presents the survey data and ORs for each included variable. Section B.2 presents the logistic regression results.

B.1 Survey Data

B.1.1 Householder Race and Ethnicity

hhspan is a binary variable that measures whether the householder was Hispanic. Of the 73,222 households surveyed, there were 43,969 responses to both *hhspan* and *etrans*. Table B-1 describes the survey responses for these variables.

Table B-1

	Original Sample		Weighted Values	
Households	Householder Hispanic	Householder Not Hispanic	Householder Hispanic	Householder Not Hispanic
Within ½ block of airport/railway/ highway	1115	6889	2,743,936	16,877,692
Outside ½ block of airport/railway/ highway	4062	31,903	9,748,796	79,674,084

22.0% of Hispanic householders and 17.5% of non-Hispanic householders respond affirmatively to *etrans*. The OR for living "nearby" corresponding to a householder being Hispanic is 1.329 (95% CI 1.327-1.330), meaning that the a Hispanic householder is about 33% more likely to live near

B.1.2 Householder Race

Of the 73,222 households surveyed, 43,969 both answered whether the household was within a half block of a four-lane highway, railway, or airport and indicated the race of the householder. For this analysis, we coded a variable *white*, indicating whether the householder was white or nonwhite. Table B-2 describes the survey responses for these variables.

Table B-2

	Original Sample		Weighted Values	
Households	Householder White	Householder Not White	Householder White	Householder Not White
Within ½ block of airport/railway/ highway	6,137	1,867	14,937,801	4,683,826
Outside ½ block of airport/railway/	29,807	6,158	74,064,319	15,358,562

highway				

16.8% of housing units with white householders and 23.4% of those with nonwhite householders responded affirmatively to *etrans*. The OR corresponding to a householder being white is 0.661 (95% CI .661-.662). The odds of a household being within a half-block of a four-lane highway, railway, or airport decrease by 34% if the householder is white.

B.1.3 Householder High School Graduation

The variable *hhhsgrad* indicates whether the householder graduated high school. Of the 73,222 households surveyed, 43,969 responded to both *hhhsgrad* and *etrans*. Table B-3 describes the survey responses for these variables.

Table B-3

	Original Sample		Weighted Values	
Households	Householder Graduated	Householder Not Graduated	Householder Graduated	Householder Not Graduated
Within ½ block of airport/railway/ highway	6,746	1,258	16,592,399	3,029,228
Outside ½ block of airport/railway/ highway	31,259	4706	77,593,559	11,829,321

17.6% of housing units with a high school graduate as householder responded affirmatively to *etrans*, as did 20.4% of those with householders who were not graduates. The OR corresponding to a householder having graduated from high school is 0.835 (95% CI 0.834-0.836). The odds of a household being within a half-block of a four-lane highway, railway, or airport decrease by 16.5% if the householder graduated high school.

B.1.4 Householder University Graduation

The second variable examined in relation to education is *unigrad*, the binary variable of whether the householder graduated from a university. Of the 73,222 households surveyed, 43,969 answered whether the household was within a half block of a four-lane highway, railway, or airport and indicated the householder's level of education. Table B-4 describes the survey responses for these variables.

Table B-4

	Original Sample		Weighted Values	
Households	Householder Graduated	Householder Not Graduated	Householder Graduated	Householder Not Graduated
Within ½ block of airport/railway/	2,112	5,892	5,004,499	14,617,128

highway				
Outside ½ block of airport/railway/ highway	11,340	24,625	27,586,625	61,836,256

15.3% of units whose householders graduated from university and 19.1% of other units responded affirmatively to *etrans*. The OR corresponding to a householder having graduated from a university is 0.768 (95% CI 0.767-0.769), meaning that a college degree reduces one's chance of being in a "nearby" housing unit by 23.2%.

B.1.5 Current Householder Employment

hhwlineq designates whether the householder worked in the past week. Of the 73,222 households surveyed, 43,412 responded to both *etrans* and *hhwlineq*. Table B-5 describes the survey responses for these variables.

Table B-5

	Original Sample		Weighted Values	
Households	Householder Worked	Householder Did Not Work	Householder Worked	Householder Did Not Work
Within ½ block of airport/railway/ highway	4,664	3,259	11,830,420	7,592,285
Outside ½ block of airport/railway/ highway	21,089	14,400	53,717,301	34,528,774

18.05% of housing units with householders working in the past week and 18.02% of those with householders that did not work in the last week responded to *etrans* affirmatively. The OR corresponding to a householder having worked in the past week is 1.002 (95% CI 1.001-1.003). Relative to the ORs of longer-term economic indicators, this short-term does not address most of the *etrans* responses.

IV. A. 1. Public Housing

proj is the binary variable indicating whether the household resided within a building owned by a public housing authority. Of the 73,222 households surveyed, 4721 both answered both *proj* and *etrans*. Table B-6 describes the survey responses for these variables.

Table B-6

	Original Sample		Weighted Values	
Households	Public Housing	Not Public Housing	Public Housing	Not Public Housing
Within ½ block of	243	958	559,723	1,693,293

airport/railway/ highway				
Outside ½ block of airport/railway/ highway	536	2984	1,192,188	5,231,271

Households in units owned by public housing authorities had a 32.0% chance of living within a half block of an airport, railroad, or 4+ lane highway (i.e., affirmative *etrans* response). Households living in units not owned by public housing authorities had a 24.5% chance of an affirmative *etrans* response. The OR corresponding to a household having resided within a building owned by a public housing authority is 1.450 (95% CI 1.445-1.456).

IV. A. 2. Food Stamps

Another socioeconomic variable examined is *qfs1*, the binary variable of whether the household received food stamps in the past year. Of the 73,222 households surveyed, 12,286 both answered whether the household was within a half block of a four-lane highway, railway, or airport and whether the household had received food stamps in the past year. Table B-8 describes the survey responses for these variables.

Table B-8

	Original Sample		Weighted Values	
Households	Receive Food Stamps	Do Not Receive Food Stamps	Receive Food Stamps	Do Not Receive Food Stamps
Within ½ block of airport/railway/ highway	727	2102	1,815,348	5,115,085
Outside ½ block of airport/railway/ highway	1917	7540	4,766,553	18,844,023

27.5% of households receiving food stamps lived in areas within a half block of an airport, railroad, or 4+ lane highway (i.e., affirmative *etrans* response). 19.8% of households not receiving food stampes had affirmative *etrans* responses. The OR corresponding to a household having received food stamps in the past year is 1.403 (95% CI 1.400-1.406).

IV. A. 3. Welfare

qwelf indicates whether the household received welfare in the past year. Of the 73,222 households surveyed, 43,969 both answered whether the household was within a half block of a four-lane highway, railway, or airport (*etrans*) and whether the household had received welfare in the past year (*qwelf*). Table B-9 describes the survey responses for these variables.

Table B-9

	Original Sample		Weighted Values	
Households	Received Welfare	Not Received Welfare	Received Welfare	Not Received Welfare
Within ½ block of airport/railway/ highway	199	7805	482,598	19,139,029
Outside ½ block of airport/railway/ highway	601	35,364	1,492,707	87,930,173

17.9% of households not receiving welfare in the past year responded affirmatively to *etrans*. 24.5% of households receiving welfare in that time frame responded affirmatively. The OR corresponding to a household having received welfare in the past year is 1.485 (95% CI 1.480-1.490), meaning that the odds of living near an airport, railroadd, or 4+lane highway were 48.5% greater for households receiving welfare in the past year.

IV. A. 4. Stock Dividends

qdiv reports whether the household received stock dividends in the past year. Of the 73,222 households surveyed, 43,969 both answered *qdiv* and *etrans*. Table B-10 describes the survey responses for these variables.

Table B-10

	Original Sample		Weighted Values	
Households	Stock Dividends	No Stock Dividends	Stock Dividends	No Stock Dividends
Within ½ block of airport/railway/ highway	550	7454	1,269,987	18,351,641
Outside ½ block of airport/railway/ highway	3643	32,322	8,492,672	80,930,208

Among households not receiving dividends, 18.5% responded affirmatively to *etrans*; 13.0% of those receiving dividends did as well. The OR corresponding to a household having received stock dividends in the past year is 0.660 (95% CI 0.659-0.661), meaning that the odds of a household being within a half-block of a four-lane highway, railway, or airport decrease by 34% if the household received stock dividends in the past year.

IV. A. 5. Garage

Of the 73,222 households surveyed, 52,177 both answered *etrans* and whether the household had had a garage or carport available for their use. Table B-11 describes the survey responses for these variables.

Table B-11

	Original Sample		Weighted Values	
Households	Garage	No Garage	Garage	No Garage
Within ½ block of airport/railway/ highway	4,846	4,455	11,827,620	10,240,417
Outside ½ block of airport/railway/ highway	27,753	15,123	68,740,753	36,115,500

22.1% of households with no garage or carport available for their use responded affirmatively to *etrans*. 14.7% of households with access to a garage or carport responded affirmatively. The OR for a garage or carport is 0.607 (95% CI 0.606-0.607), meaning odds of a household with a garage or carport was 39% lower than if the household had no garage or carport available for their use.

B.2 Logistic Regression Results

Table B-12

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	HHSPAN	.284	.001	149803.8	1	.000	1.329	1.327	1.330
	Constant	-1.552	.000	3.4E+07	1	.000	.212		

a. Variable(s) entered on step 1: HHSPAN.

Table B-13

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	WHITE	-.413	.001	475827.4	1	.000	.661	.661	.662
	Constant	-1.188	.001	5061853	1	.000	.305		

a. Variable(s) entered on step 1: WHITE.

Table B-14

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	HHHSGRAD	-.180	.001	66561.327	1	.000	.835	.834	.836
	Constant	-1.362	.001	4475443	1	.000	.256		

a. Variable(s) entered on step 1: HHHSGRAD.

Table B-15

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	UNIGRAD	-.264	.001	218089.6	1	.000	.768	.767	.769
	Constant	-1.442	.000	2.5E+07	1	.000	.236		

a. Variable(s) entered on step 1: UNIGRAD.

Table B-16

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	HHWLINEQ	.002	.001	9.686	1	.002	1.002	1.001	1.003
	Constant	-1.515	.000	1.4E+07	1	.000	.220		

a. Variable(s) entered on step 1: HHWLINEQ.

Table B-17

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	PROJ	.372	.002	40588.801	1	.000	1.450	1.445	1.456
	Constant	-1.128	.001	1627601	1	.000	.324		

a. Variable(s) entered on step 1: PROJ.

Table B-18

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	QFS1	.339	.001	113631.9	1	.000	1.403	1.400	1.406
	Constant	-1.304	.000	6840804	1	.000	.271		

a. Variable(s) entered on step 1: QFS1.

Table B-19

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	QWELF	.396	.002	55771.053	1	.000	1.485	1.480	1.490
	Constant	-1.525	.000	3.7E+07	1	.000	.218		

a. Variable(s) entered on step 1: QWELF.

Table B-20

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	QDIV	-.415	.001	177605.5	1	.000	.660	.659	.661
	Constant	-1.484	.000	3.3E+07	1	.000	.227		

a. Variable(s) entered on step 1: QDIV.

Table B-21

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	GARAGE	-.500	.000	1111785	1	.000	.607	.606	.607
	Constant	-1.260	.000	1.3E+07	1	.000	.284		

a. Variable(s) entered on step 1: GARAGE.

[10] Appendix I: AHS data users FAQs. (2011). In *Codebook for the American Housing Survey, public use file: 1997 and later* (p. 1260).

Appendix C – Multivariate Logistic Regression Models

For the strict model, shown first, there are three tables shown. The first two tables shown are the results of the Hosmer and Lemeshow test, and the third table shows the results of the logistic regression. In the other three models, there are four tables displayed for each. Before the table showing the results of the logistic regression, there is a table indicating the representations of the categorical variables.

Strict Model

Table C-1

Contingency Table for Hosmer and Lemeshow Test

		Within 1/2 Block of Transport = No		Within 1/2 Block of Transport = Yes		
		Observed	Expected	Observed	Expected	Total
Step 1	1	3348	3314.129	0	34.317	3348
	2	4405	4348.431	0	56.984	4405
	3	4770	4696.664	0	73.207	4770
	4	4841	4762.109	0	79.012	4841
	5	3098	3046.254	0	52.215	3098
	6	3151	3097.380	0	53.820	3151
	7	2678	2631.885	0	45.989	2678
	8	19403	19068.592	0	333.946	19403
	9	640	4770.590	4226	94.599	4865
	10	87927883	8.8E+07	19297919	1.9E+07	1.1E+08

Table C-2

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	184717.609	8	.000

Table C-3

Variables in the Equation

								95.0% C.I.for EXP(B)	
		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Step 1[a]	HHSPAN	.287	.001	145819.2	1	.000	1.332	1.331	1.334
	WHITE	-.388	.001	395413.7	1	.000	.678	.677	.679
	XKINC266	-.039	.000	622793.0	1	.000	.962	.961	.962
	Constant	-1.264	.001	5343423	1	.000	.283		

a. Variable(s) entered on step 1: HHSPAN, WHITE, XKINC266.

Medium-level Model

Table C-4

Contingency Table for Hosmer and Lemeshow Test

		Within 1/2 Block of Transport = No		Within 1/2 Block of Transport = Yes		
		Observed	Expected	Observed	Expected	Total
Step 1	1	4448	4329.759	0	118.444	4448
	2	5083	4917.812	0	165.124	5083
	3	4770	4602.723	0	167.628	4770
	4	3959	3807.268	0	151.572	3959
	5	4644	4461.908	0	181.701	4644
	6	4231	4065.139	0	166.055	4231
	7	5269	5062.106	0	207.209	5269
	8	5036	4833.102	0	203.232	5036
	9	1685	4481.945	2997	200.332	4682
	10	87935091	8.8E+07	19299147	1.9E+07	1.1E+08

Table C-5

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	42200.882	8	.000

Table C-6

Categorical Variables Codings

			Parameter coding		
		Frequency	(1)	(2)	(3)
Unit Type	One-unit building, detached	27856	.000	.000	.000
	One-unit building, attached	2725	1.000	.000	.000
	Building with 2+ apartments	10707	.000	1.000	.000
	One-unit mobile home	1975	.000	.000	1.000

Table C-7

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	HHSPAN	.154	.001	39470.829	1	.000	1.166	1.164	1.168
	WHITE	-.245	.001	145425.4	1	.000	.783	.782	.784
	XKINC266	-.017	.000	116767.6	1	.000	.983	.983	.983
	QWELF	.022	.002	163.553	1	.000	1.022	1.019	1.026
	QDIV	-.147	.001	20444.050	1	.000	.864	.862	.865
	BEDRMS3	-.145	.000	194146.9	1	.000	.865	.865	.866
	NUNIT2			938428.3	3	.000			
	NUNIT2(1)	.311	.001	75098.801	1	.000	1.364	1.361	1.367
	NUNIT2(2)	.707	.001	927208.1	1	.000	2.028	2.025	2.031
	NUNIT2(3)	.147	.001	17562.546	1	.000	1.158	1.156	1.161
	Constant	-1.611	.001	6368956	1	.000	.200		

a. Variable(s) entered on step 1: HHSPAN, WHITE, XKINC266, QWELF, QDIV, BEDRMS3, NUNIT2.

Wide-scope Model

Table C-8

Contingency Table for Hosmer and Lemeshow Test

		Within 1/2 Block of Transport = No		Within 1/2 Block of Transport = Yes		
		Observed	Expected	Observed	Expected	Total
Step 1	1	2523	2497.508	0	25.665	2523
	2	3681	3634.666	0	46.099	3681
	3	4233	4168.201	0	64.482	4233
	4	5362	5260.583	0	101.719	5362
	5	4575	4482.677	0	92.244	4575
	6	5247	5135.374	0	111.417	5247
	7	5507	5382.064	0	125.389	5507
	8	4651	4542.498	0	108.932	4651
	9	4411	4305.582	0	105.336	4411
	10	87934027	8.8E+07	19302144	1.9E+07	1.1E+08

Table C-9

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	797.555	8	.000

Table C-10

Categorical Variables Codings

			Parameter coding															
		Frequency	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Householder Education	Less than 1st Grade	114	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	1st, 2nd, 3rd, or 4th Grad	337	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	5th or 6th Grade	644	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	7th or 8th Grade	953	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	9th Grade	739	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	10th Grade	1011	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	11th Grade	1082	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	12th Grade, no diploma	936	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000
	HS Grad	11784	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000
	Some college, no degree	7449	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000
	Diploma/cert. from vocational/tech./business trade school	1438	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000
	Associate degree - Occupational/vocational program	1753	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000
	Associate degree - Academic program	1720	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000
	Bachelors degree	8286	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000
	Master's degree	3593	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000
	Progessional School degree	764	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000
	Doctorate degree	660	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Metro Area	Central city of MSA	12491	.000	.000	.000	.000												
	Inside MSA, but not in central city - urban	16499	1.000	.000	.000	.000												
	Inside MSA, but not in central city - rural	5659	.000	1.000	.000	.000												
	Outside MSA, urban	3039	.000	.000	1.000	.000												
	Outside MSA, rural	5575	.000	.000	.000	1.000												
Region	Northeast	9862	.000	.000	.000													
	Midwest	11214	1.000	.000	.000													
	South	13872	.000	1.000	.000													
	West	8315	.000	.000	1.000													
Unit Type	One-unit building, detached	27856	.000	.000	.000													
	One-unit building, attached	2725	1.000	.000	.000													
	Building with 2+ apartments	10707	.000	1.000	.000													
	One-unit mobile home	1975	.000	.000	1.000													
Type of Ownership	Owned or being bought by someone in household	29177	.000	.000														
	Rented	13427	1.000	.000														
	Occupied without payment of rent	659	.000	1.000														

Table C-11

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
								Lower	Upper
Step 1[a]	BEDRMS3	-.066	.000	26959.074	1	.000	.936	.936	.937
	BATHS2	-.228	.001	184547.9	1	.000	.796	.795	.797
	HALFB1	-.064	.001	12739.193	1	.000	.938	.937	.939
	PER3	.009	.000	1259.345	1	.000	1.009	1.009	1.010
	ZADULT2	-.011	.000	602.005	1	.000	.989	.988	.990
	CARS1	.010	.000	961.278	1	.000	1.011	1.010	1.011
	QWELF	-.015	.002	74.443	1	.000	.985	.982	.988
	QDIV	-.087	.001	6859.794	1	.000	.917	.915	.919
	XKINC266	.004	.000	442.547	1	.000	1.004	1.004	1.005
	XKZINC63	-.012	.000	3374.339	1	.000	.988	.987	.988
	XKLMED65	-.016	.000	3095.654	1	.000	.984	.983	.985
	NATIVE	.353	.003	19590.919	1	.000	1.424	1.416	1.431
	PACIFIC	.022	.005	20.643	1	.000	1.022	1.013	1.032
	BLACK	.161	.001	45041.265	1	.000	1.174	1.172	1.176
	ASIAN	-.182	.001	15084.193	1	.000	.834	.831	.836
	MIXED	.071	.002	1045.210	1	.000	1.074	1.069	1.079
	METRO3			906816.6	4	.000			
	METRO3(1)	.085	.001	17643.041	1	.000	1.088	1.087	1.090
	METRO3(2)	-.567	.001	326667.4	1	.000	.567	.566	.568
	METRO3(3)	-.265	.001	58919.413	1	.000	.767	.765	.769
	METRO3(4)	-.780	.001	520990.2	1	.000	.458	.457	.459
	TENURE			39634.697	2	.000			
	TENURE(1)	.153	.001	39580.637	1	.000	1.165	1.163	1.167
	TENURE(2)	.045	.002	429.220	1	.000	1.046	1.041	1.050
	HHGRAD			92387.138	16	.000			
	HHGRAD(1)	.148	.005	864.812	1	.000	1.159	1.148	1.171
	HHGRAD(2)	.084	.004	562.031	1	.000	1.088	1.080	1.095
	HHGRAD(3)	-.182	.003	3284.076	1	.000	.834	.829	.839
	HHGRAD(4)	-.206	.003	4749.734	1	.000	.814	.809	.819
	HHGRAD(5)	.070	.003	542.384	1	.000	1.072	1.066	1.079
	HHGRAD(6)	.118	.003	1701.780	1	.000	1.125	1.118	1.131
	HHGRAD(7)	-.032	.003	132.085	1	.000	.968	.963	.973
	HHGRAD(8)	.071	.003	599.939	1	.000	1.074	1.068	1.080
	HHGRAD(9)	.026	.002	112.172	1	.000	1.026	1.021	1.031
	HHGRAD(10)	.018	.002	56.648	1	.000	1.018	1.014	1.023
	HHGRAD(11)	.095	.003	1205.377	1	.000	1.099	1.094	1.105
	HHGRAD(12)	.082	.003	940.950	1	.000	1.085	1.079	1.091
	HHGRAD(13)	.021	.003	62.320	1	.000	1.021	1.016	1.027
	HHGRAD(14)	-.038	.002	253.949	1	.000	.962	.958	.967
	HHGRAD(15)	-.225	.003	7863.146	1	.000	.798	.794	.802
	HHGRAD(16)	-.115	.003	1282.646	1	.000	.891	.886	.897

	REGION			445695.2	3	.000			
	REGION(1)	.405	.001	214676.7	1	.000	1.499	1.496	1.502
	REGION(2)	.565	.001	438900.2	1	.000	1.759	1.756	1.762
	REGION(3)	.329	.001	138231.5	1	.000	1.389	1.387	1.391
	NUNIT2			456433.1	3	.000			
	NUNIT2(1)	.260	.001	49050.331	1	.000	1.297	1.294	1.300
	NUNIT2(2)	.560	.001	408526.8	1	.000	1.751	1.748	1.754
	NUNIT2(3)	.367	.001	98622.114	1	.000	1.444	1.440	1.447
	Constant	-2.172	.003	737715.0	1	.000	.114		

a. Variable(s) entered on step 1: BEDRMS3, BATHS2, HALFB1, PER3, ZADULT2, CARS1, QWELF, QDIV, XKI XKZINC63, XKLMED65, NATIVE, PACIFIC, BLACK, ASIAN, MIXED, METRO3, TENURE, HHGRAD, REGIO

Wide-scope Model, Adjusted

Table C-12

Contingency Table for Hosmer and Lemeshow Test

		Within 1/2 Block of Transport = No		Within 1/2 Block of Transport = Yes		
		Observed	Expected	Observed	Expected	Total
Step 1	1	3681	3638.708	0	42.058	3681
	2	4334	4281.212	0	53.100	4334
	3	4812	4731.725	0	80.539	4812
	4	4258	4176.162	0	81.556	4258
	5	4881	4782.565	0	98.258	4881
	6	5076	4968.036	0	107.730	5076
	7	4909	4798.606	0	110.465	4909
	8	4062	3963.712	0	98.137	4062
	9	5062	4936.645	0	124.952	5062
	10	89381807	8.9E+07	19621627	2.0E+07	1.1E+08

Table C-13

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	813.427	8	.000

Table C-14

Categorical Variables Codings

			Parameter coding															
		Frequency	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Householder Education	Less than 1st Grade	114	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	1st, 2nd, 3rd, or 4th Grade	337	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	5th or 6th Grade	644	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	7th or 8th Grade	953	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	9th Grade	739	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	10th Grade	1011	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	11th Grade	1082	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	12th Grade, no diploma	936	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000
	HS Grad	11784	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000
	Some college, no degree	7449	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000
	Diploma/cert. from vocational/tech./business/ trade school	1438	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000
	Associate degree - Occupational/vocational program	1753	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000
	Associate degree - Academic program	1720	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000
	Bachelors degree	8286	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000
	Master's degree	3593	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000
	Progessional School degree	764	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000
	Doctorate degree	660	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Metro Area	Central city of MSA	12491	.000	.000	.000	.000												
	Inside MSA, but not in central city - urban	16499	1.000	.000	.000	.000												
	Inside MSA, but not in central city - rural	5659	.000	1.000	.000	.000												
	Outside MSA, urban	3089	.000	.000	1.000	.000												
	Outside MSA, rural	5575	.000	.000	.000	1.000												
Region	Northeast	9862	.000	.000	.000													
	Midwest	11214	1.000	.000	.000													
	South	13872	.000	1.000	.000													
	West	8315	.000	.000	1.000													
Unit Type	One-unit building, detached	27856	.000	.000	.000													
	One-unit building, attached	2725	1.000	.000	.000													
	Building with 2+ apartments	10707	.000	1.000	.000													
	One-unit mobile home	1975	.000	.000	1.000													
Type of Ownership	Owned or being bought by someone in household	29177	.000	.000														
	Rented	13427	1.000	.000														
	Occupied without payment of rent	659	.000	1.000														

Table C-15

Variables in the Equation

								95.0% C.I.for EXP(B)	
		B	S.E.	Wald	df	Sig.	Exp(B)	Lower	Upper
Step 1[a]	HALFB1	-.080	.001	21371.635	1	.000	.923	.922	.924
	ZADULT2	-.008	.000	540.080	1	.000	.992	.991	.993
	CARS1	.008	.000	586.319	1	.000	1.008	1.007	1.009
	QWELF	-.006	.002	10.161	1	.001	.994	.991	.998
	QDIV	-.084	.001	6410.036	1	.000	.920	.918	.922
	XKLMED65	-.018	.000	4043.759	1	.000	.982	.981	.982
	NATIVE	.346	.002	19353.645	1	.000	1.414	1.407	1.421
	PACIFIC	.014	.005	7.869	1	.005	1.014	1.004	1.023
	BLACK	.139	.001	35135.869	1	.000	1.149	1.148	1.151
	ASIAN	-.169	.001	13593.654	1	.000	.844	.842	.847
	MIXED	.053	.002	601.454	1	.000	1.055	1.050	1.059
	METRO3			931190.3	4	.000			
	METRO3(1)	.088	.001	19390.393	1	.000	1.092	1.091	1.093
	METRO3(2)	-.567	.001	329048.7	1	.000	.567	.566	.569
	METRO3(3)	-.270	.001	62122.771	1	.000	.763	.762	.765
	METRO3(4)	-.788	.001	537678.8	1	.000	.455	.454	.456
	TENURE			50729.353	2	.000			
	TENURE(1)	.168	.001	50054.124	1	.000	1.183	1.181	1.184
	TENURE(2)	.014	.002	41.252	1	.000	1.014	1.010	1.018
	HHGRAD			93735.059	16	.000			
	HHGRAD(1)	.135	.005	746.954	1	.000	1.145	1.134	1.156
	HHGRAD(2)	.044	.004	158.808	1	.000	1.045	1.038	1.053
	HHGRAD(3)	-.193	.003	3753.909	1	.000	.824	.819	.829
	HHGRAD(4)	-.222	.003	5578.574	1	.000	.801	.797	.806
	HHGRAD(5)	.066	.003	492.578	1	.000	1.068	1.062	1.074
	HHGRAD(6)	.099	.003	1229.211	1	.000	1.104	1.098	1.110
	HHGRAD(7)	-.058	.003	423.700	1	.000	.944	.939	.949
	HHGRAD(8)	.058	.003	398.902	1	.000	1.059	1.053	1.065
	HHGRAD(9)	.012	.002	26.945	1	.000	1.013	1.008	1.017
	HHGRAD(10)	.017	.002	51.683	1	.000	1.017	1.013	1.022
	HHGRAD(11)	.092	.003	1153.131	1	.000	1.097	1.091	1.103
	HHGRAD(12)	.077	.003	847.596	1	.000	1.080	1.075	1.086
	HHGRAD(13)	.019	.003	52.879	1	.000	1.019	1.014	1.025
	HHGRAD(14)	-.040	.002	278.069	1	.000	.961	.956	.965
	HHGRAD(15)	-.230	.003	8316.292	1	.000	.794	.790	.798
	HHGRAD(16)	-.120	.003	1400.240	1	.000	.887	.882	.893
	REGION			456316.4	3	.000			
	REGION(1)	.408	.001	221992.7	1	.000	1.504	1.501	1.506
	REGION(2)	.566	.001	449908.5	1	.000	1.761	1.758	1.764
	REGION(3)	.333	.001	144754.5	1	.000	1.395	1.392	1.397
	NUNIT2			649952.1	3	.000			
	NUNIT2(1)	.281	.001	59844.057	1	.000	1.325	1.322	1.328
	NUNIT2(2)	.612	.001	586148.5	1	.000	1.843	1.840	1.846
	NUNIT2(3)	.394	.001	116177.4	1	.000	1.482	1.479	1.486
	BATHS2	-.264	.000	305066.6	1	.000	.768	.767	.769
	XKZINC63	-.008	.000	17752.795	1	.000	.993	.992	.993
	Constant	-2.203	.003	773867.8	1	.000	.110		

a. Variable(s) entered on step 1: HALFB1, ZADULT2, CARS1, QWELF, QDIV, XKLMED65, NATIVE, PACIFIC, BLACK, ASIA MIXED, METRO3, TENURE, HHGRAD, REGION, NUNIT2, BATHS2, XKZINC63.