Immigrant-Native Disparities in Perceived and Actual Met/Unmet Need for Medical Care: Analytical Replication and Critique

Introduction

This paper serves to replicate the study done by Stephanie Howe Hasanali in her article, *Immigrant-Native Disparities in Perceived and Actual Met/Unmet Need for Medical Care* [2]. Immigrants face many barriers in the United States. This disparity is clear in many different facets, especially within healthcare. From an individual standpoint, some immigrants have avoided receiving healthcare due to fear of deportation, language differences, lacking financial resources, stigma, and being unfamiliar with the healthcare system[1]. Many of these individual issues can be attributed to negative attitudes towards immigrants and lacking many of the resources and policies the country needs to welcome immigrants into the United States. Some of these external factors include lacking employee benefits, translators, and health care access[5]. As a result, there is a large divide between immigrants' and natives' healthcare access, significantly blocking immigrants from much needed healthcare. A literature review done in 2013 found that much research reported that immigrant women experienced a higher frequency of pregnancy complications, postpartum depression, and psychosis[4]. Both these physical and mental complications can be attributed to several factors, including negative healthcare experiences, cultural insensitivity, or not having access to health insurance, all leading to immigrant women to avoid receiving healthcare.

The purpose of Hasanali's article is to compare disparities, such as the example previously mentioned, through the lenses of subjective unmet need, and objective unmet need for both immigrants and citizens in the United States. Subjective unmet need refers to delayed or avoiding needed medical care, due to more individual or personal reasons. Objective unmet need indicates inadequate or poorly timed care. This focuses more on the treatment of the individual, while they are receiving healthcare.

The purpose of this paper is to determine if there had been a change in comparative percentage of objective and subjective unmet need in 2014, compared to Hasanali's sample from 2007-2009. Our group was drawn towards this article because we wanted to learn more about disparities within healthcare.

Given the various debates today surrounding immigration and globalization, our group wanted to learn more about the impacts on health, and how health disparities affect immigrants in the United States, as compared to citizens.

Additionally, we wanted to determine what variables contribute most significantly to objective and subjective unmet need. To analyze this relationship, our group ran multiple logistic regression in RStudio to determine the significance of the covariates to determine objective and subjective unmet need for both our immigrant sample and the total sample. The results will be described later in the paper.

Data and Creation of Target Population

The data for this analysis is from the 2014 Medical Expenditure Panel Survey (MEPS). More specifically, we used the Household Component file, which includes survey data from individual household members from panels 18 and 19. Hasanali utilized longitudinal data from 2007 to 2009, meaning she viewed data over time for a few years, while our group utilized data solely from 2014. To view the 2014 MEPS Household Component file, we downloaded the file into RStudio, using the load function. For our analysis, we subset these variables: yrsinus, bornusa, sex, intvlang, racethx, MARRY14X, INSCOV14, MDDLAY42, CHECK53, AGE14X, RTHLTH53, eduyrdg, POVCAT14, OCCCAT53, and REGION53. More information about the variables and their definitions is described in the Data Dictionary in the Appendix.

Immigrant-Specific Variables

The variables yrsinus, bornusa, intvlang, REGION53, and OCCCAT53 were considered immigrant-specific variables, and used when studying the immigrant sample. The yrsinus variable represents the length of time the individual resided in the United States. Hasanali used a continuous length of time measure from their data, which they squared and cubed, which we did not have access to. To adjust to this discrepancy, we took the midpoint of each category in the yrsinus variable and created a midpoint variable (YrsInUSAMid). We then made two more variables, which squared and cubed the midpoint values within each category (YrsInUSASqr and YrsInUSACub). It is important to note that for the last category of years within the yrsinus variable (15 or more), the midpoint used was 15. This could possibly impact the true average and influence the variable significance towards unmet needs. The "YrsInUSA" variables were determined to be insignificant in determining both subjective and objective unmet need (Table 2, Table 3). The variable bornusa, recoded as ForeignBorn, was used to determine if the individual was born in the United States. We used this variable to separate our immigrant sample from the total population for analysis. The intvlang variable determined English proficiency, based on whether an individual's MEPS interview was done in English. The intvlang variable was recoded as EngProf.

Region-Destination Type

REGION53 and OCCCAT53 were variables not found in Hasanali's study. Hasanali had access to state-specific data, in which she was able to use a state-level typology to characterize each state as new or traditional, based on their immigration history. She ended up with 4 state-level destination types: traditional, low skill; traditional, high/balanced skill; new, low skill; and new, high-balanced skill. As our group was unable to access this state-specific data, we instead classified our data using REGION53, eduyrdg, and OCCCAT53 to determine region-level destination type. REGION53 was coded with a 1 referring to the Northeast, 2 as the Midwest, 3 as the South, and 4 as the West. REGION53 was recoded as Region, with each level named for its corresponding region. OCCCAT53 was recoded as OccCat, with observations between codes 1 and 5 referring to new occupations (post-industrial occupations), and occupations between 6 and 9 referring to traditional occupations (agricultural and industrial occupations). The immigrant skill level ratio was determined by dividing the number of high-skill immigrants (those who obtained a college degree) in a region by the number of low-skill immigrants (those who did not obtain a college degree) in that region. The immigrant occupation type ratio was determined by dividing the number of new occupation immigrants in a region by the number of traditional occupation immigrants in that region. A comparative analysis of both skill level and occupation type ratios determined their categories. The two final regional categories determined were traditional, low-skill regions (South and West); and new, high-skill regions (Northeast and Midwest). This region-level destination type was used in place of Hasanali's state-level destination type throughout the study.

Sociodemographic Covariates

The variables sex, racethx, AGE14X, MARRY14X, INSCOV14, RTHLTH53, eduyrdg, and POVCAT14 were sociodemographic covariates. These variables were applied to the whole sample. The sex and racethx (recoded as Race) variables, were used to categorize the gender and race of the surveyed individual. AGE14X is a continuous variable of the ages of the surveyed individuals, of which we only included respondents 18 or older in our subset. AGE14X was also recoded as Age, including all non-missing observations in the sample, for Table 1 statistics. Terms for Age_2 and Age^3 were also included to control for non-linear age dependency for healthcare access and usage. MARRY14X determined the partnership status of the surveyed individual and was recoded as a factor variable called Marriage. The INSCOV14 variable was used to determine an individual’s insurance status. In terms of coding for our study, INSCOV14 was recoded as a factor variable, HealthInsCov, with levels for individuals who were never uninsured, and those who were ever uninsured. RTHLTH53 was used to determine the individual’s self-reported health status and was recoded as a factor variable called Healthstat, based on the levels within RTHLTH53. The eduyrdg variable was used to determine educational levels of the observations. This variable was recoded as a factor variable, called EducNew, the levels based on the eduyrdg levels. POVCAT14 was used to measure family income as a percentage of the poverty line. In 2014, the federal poverty level for a one-person household was $11,670, increasing about $4,000 for each family member [3]. Like previous variables, POVCAT14 was recoded as a factor variable called Income. In addition, interaction terms included all interactions between age, race/ethnicity, and sex. It is also to be noted that for the construction of Table1, RTHLTH53, eduyrdg, POVCAT14, and MARRY14X were split into four binomials each, based on their relative levels.

Outcome Variables

Lastly, MDDLAY42 and CHECK53 were the binary outcome variables of focus. MDDLAY42 represented subjective unmet need, and CHECK53 represented objective unmet need. MDDLAY42 referred to delayed medical care within the past year and was recoded as DelayedMedCare. This is determined by the respondent, who is asked to evaluate need based on both statements made by a medical provider and personal beliefs. Because this need must be self-recognized and self-reported, it is likely to be biased and potentially inaccurate. In comparison, CHECK53 determined the time since the last checkup. This was recoded as LastCheckup, with a 1 meaning “within the past year”, and greater than 1 being “not within past year”. This eliminated some of the uncertainty and variation in response occurring in subjective unmet need.

For the target population, we only included non-missing observations where AGE14X was greater than 18. After recoding as described, we ended with a sample population of 22,711 adults, 6,316 of which who were immigrants, as noted by Flowchart 1 found in the Appendix.

Data Description

The 2014 MEPS data was reduced from 34,875 observations to 22,711 observations, consisting of 16,395 U.S. born people and 6,316 foreign born people, as noted previously. Missing observations accounted for about 4.5% of the original sample and were removed. Pertaining to missing observations, if there was an NA in an essential question, then the observation would be deleted. If the question was non-essential to the observation, then that observation would be kept. For example, if an individual born in the United States skipped the question, “How many years have you lived in the USA?”, they would be included in the dataset. If an immigrant had skipped the question, their data would be deleted from the dataset. Approximately 28% of the adjusted sample was foreign born.

Compared to the original sample, this sample represented demographic groups similarly, apart from the race category. As shown in Table 1, 54% of our population was female, compared to Hasanali’s study which consisted of a 52.1% female population. In addition, mean age was approximately 45 for total, native, and immigrant samples; it ranged from 43.9 – 46.7 in Hasanali’s study. Because of Hasanali’s sample weighting, race categories varied substantially. Our study did not include any sample weighting. While Hasanali’s sample was 70.4% non-Hispanic white, ours was only 41.2% white, due to the overweighting of black and Hispanic populations in MEPS.

Table 1: Characteristics of Sample Respondents

Variable	Total N = 22,711	US N = 16,395	Immigrant N = 6,316
Subjective Unmet Need	3.4%	3.9%	2.0%
Objective Unmet Need	33.3%	31.4%	38.3%
Foreign Born (vs. Native)	28.0%	0.0%	100.0%
Age (mean)	45.53	45.48	45.66
Age (sd)	17.61	18.31	15.67
Female (vs. Male)	54.0%	54.2%	53.5%
Race/Ethnicity			
*White**	41.2%	54.4%	6.8%
Black	21.7%	27.0%	7.7%
Hispanic	29.3%	16.0%	63.9%
Asian	7.8%	2.5%	21.6%
Health Status			
Fair/Poor	12.1%	11.3%	14.3%
Good	29.8%	28.3%	33.5%
Very Good	31.1%	32.9%	26.2%
*Excellent**	23.8%	23.8%	23.7%
Educational Attainment			
Less Than High School	24.6%	17.7%	42.6%
High School Diploma	24.3%	26.1%	19.5%
Some College	29.3%	33.4%	18.4%
*Bachelor's Degree or Higher**	21.8%	22.7%	19.6%
Income (as percentage of poverty line)			
Poor	25.6%	23.8%	30.1%
Low	16.6%	15.0%	20.9%
Middle	29.6%	30.0%	28.7%
*High**	28.2%	31.2%	20.3%
Ever Uninsured (vs. Never Uninsured)	17.0%	11.8%	30.2%
Marriage Status			
*Married**	47.3%	43.4%	57.3%
Widowed	6.0%	6.6%	4.5%
Divorced	14.5%	14.9%	13.6%
Never Married	32.2%	35.1%	24.7%
Immigrant-Specific Variables			
Years in US (mean midpoint of ranges)			12.77
English Interview (vs. Non-English Interview)			48.8%
Region Destination Type			
New, high-skill			28.1%
Traditional, low-skill			71.9%

About 3.4% of all respondents reported subjective unmet need (delaying or forgoing needed medical care) in the past year. Natives reported unmet need at 3.9% while immigrants reported a level of unmet need at 2.0%. In comparison, about 33.3% of the total population reported objective unmet need. This consisted 31.4% of natives reporting objective unmet need and 38.3% of immigrants reporting

objective unmet need (Table 1). While the difference between immigrant and native experience is much greater for objective unmet need, differences in both subjective and objective unmet need are of considerable size.

Other important observations from the sample characteristics provided in Table 1 are the significant differences in perceived health status, education level, income level, and propensity to be uninsured between the native and immigrant sample groups. This is consistent with current knowledge of immigrant populations residing within the US.

Description of Methods

To describe the distribution of the demographic variables of the total, native, and immigrant samples, means and frequencies were calculated for each of the variables (Table 1). Then, twelve binary logistic regression models were created (Appendix, Flow Chart 2). These determined both the association between the independent variables and subjective unmet medical need (Table 2), and the association between the independent variables and objective unmet medical need (Table 3).

Within each of Table 2 and Table 3, the models were first split into two panels, Panel A and Panel B. Panel A applies to the entire sample and looks at the relationship between an unmet need and immigrant status, education, income, insurance status, and family type. Panel B applies to only the immigrant sample and looks at the relationship between an unmet need and some of the above factors along with immigrant specific variables. This split was included because some variables (time in US, English proficiency, and state destination type) apply only to immigrants. The specific impacts these variables have on immigrants was essential to understanding the components contributing to immigrants' levels of unmet need.

Panels A and B were then split into three models (Model 1, Model 2, and Model 3). The distinctions between these variables include the presence of control variables and the inclusion of non-immigrant-specific variables. Model 1 included only immigrant-specific variables and did not have any control variables. Model 2 included both immigrant-specific variables and the control variables. The control variables consisted of age, age_2, age^3, gender, race, self-reported health status, and all interaction terms between age, gender, and race. Model 3 included both immigrant-specific and non-immigrant-specific variables, along with the same control variables that were included in Model 2.

The coding practices utilized in R relied on the glm function. Each linear model was created with this function and used the binomial family with the logit link. The data used for each model was either the total data sample (Panel A) or the immigrant data sample (Panel B).

Results

Our analysis consists of models that have been previously described. Our results will be divided into 3 parts: Table 1 (characteristics of sample respondents), Table 2 (odds ratio (OR) and 95% confidence intervals (CI) for determinants of subjective unmet medical need), and Table 3 (odds ratio (OR) and 95% confidence intervals (CI) for determinants of objective unmet medical need). To do so, we used the odds.ratio function on our models in RStudio.

Table 1:

The biggest difference between our analysis and the original paper that we found was in race/ethnicity. In the original study, Non-Hispanic white consists about 70.4% of the total population sample, but in our analysis, it only consists of about 41.2% of the total population sample. Subsequently, Non-Hispanic white constitute only 54.4% of the US born population and 6.8% of foreign-born population, compared to 79.9% and 19.4% respectively in the original paper. As such, the MEPS data that our group uses contain significantly more individuals in minority groups compared to the original paper, because MEPS overweights minority populations.

Another significant difference that we found was the educational attainment between our analysis and the original paper. Our data contain more people with educational attainment of less than high school at 24.6%, compared to the original data at 15.6%. At 17.7%, we have a bigger percentage of the US born population with less than a high school education, compared with 12.1% in the original paper. The same goes for foreign born population, with 42.6% and 34% respectively.

In terms of income, our data indicate that more people are considered as poor or lower income compared to the original data. Our data showed that 25.6% and 16.6% of people are categorized as poor and low income respectively, compared with 15.4% and 13.3% in the original data. However, the foreign-born population has a greater probability to be categorized as poor or lower income, in line with the data in the original paper. However, our data says that there are less people who are ever uninsured for both the US born population and foreign-born population in comparison with the original data. Overall, there are many disparities between US born and foreign-born population in our sample. Immigrants have lower income, poorer health status, less education, and tend to be married relative to the US born population.

Table 2:

For our analysis of Table 2, we use the odds ratio function in R in order to determine the appropriate odds ratio for each variable. Panel A depicts the regression results for models predicting subjective unmet need among the total sample. When accounting for age, gender, race, health status, and interactions between age, gender, and race, our models predict that immigrants are almost 40% less likely to experience subjective unmet need compared to US born, and this experience is even less likely than in the original paper. The lower education levels of the total population tend to have lower odds of subjective unmet medical need when compared to bachelor's degree or higher, which is in line with the original paper. However, in terms of income, the overall population with poor/near poor are 36% more likely to experience subjective unmet medical need, in line with the original paper. In terms of family type, the overall population who are widowed and divorced/separated are also almost twice as likely as married in experiencing subjective unmet medical need, while those never married are 30% more likely to experience this unmet need.

In Panel B, which evaluates the immigrant sample, the only statistically significant variables are insurance status and family type. Immigrants who were ever uninsured were about 42% less likely to experience subjective unmet medical need compared to immigrants who were insured, which is the opposite of the results in the original paper. The overall population sample states that widows are the most likely to experience subjective unmet medical needs. However, immigrants who are divorced/separated are the only significant immigrant family type, at over two times as likely to experience subjective unmet need as married immigrants. This result is different from the original paper, where overall populations that are divorced are more likely to experience subjective unmet medical needs.

Table 2: Odds Ratios and 95% Confidence Intervals for Determinants of Subjective Unmet Need

Variables	Panel A Total Sample			Panel B Immigrant Sample		
	Model 1	*Model 2*	*Model 3*	*Model 1*	*Model 2*	*Model 3*
Foreign-born						-
Foreign-born	0.50*** (0.41, 0.60)	0.61** (0.47, 0.79)	0.65** (0.50,0.85)			
Length of time in US						
Midpoint				1.26 (0.53, 3.11)	0.95 (0.39, 2.38)	0.93 (0.38, 2.32)
Midpoint²				0.96 (0.85, 1.08)	1.00 (0.89, 1.12)	1.00 (0.89, 1.13)
Midpoint³				1.00 (1.00, 1.01)	1.00 (1.00, 1.00)	0.99 (0.99, 1.00)
English proficiency						
English interview				1.35 (0.94, 1.94)	1.49 (0.93, 2.37)	1.32 (0.80,2.15)
Region destination type (vs. new, high-skill region)						
Traditional, low-skill region				0.76 (0.52, 1.11)	0.86 (0.59, 1.27)	0.91 (0.62,1.34)
Educational attainment (vs. bachelor's degree or higher)						-
Less than high school			0.52*** (0.40,0.67)			0.80 (0.44,1.48)
High school diploma			0.54*** (0.43,0.68)			0.80 (0.42,1.52)
Some college			0.80* (0.65,0.98)			1.03 (0.58, 1.83)
Income (vs. high income)						
Poor/near poor			1.36** (1.08,1.72)			1.04 (0.58, 1.85)
Low income			1.24 (0.97,1.60)			0.85 (0.46, 1.56)
Middle income			0.92 (0.74,1.14)			0.72 (0.41, 1.23)
Insurance status						
Ever uninsured			0.96 (0.76,1.20)			0.58* (0.34, 0.97)
Family type (vs. married)						
Widowed			1.99*** (1.43,2.72)			1.78 (0.67, 4.17)
Divorced/separated			1.89*** (1.55,2.30)			2.05** (1.28, 3.24)
Never married			1.30* (1.04,1.62)			1.24 (0.72, 2.11)
Control variables	No	Yes	Yes	No	Yes	Yes

* P < .05

** P < .01

*** P < .001

Table 3:

For our analysis of Table 3, we use the odds ratio function in R in order to determine the appropriate odds ratio for each variable. Panel A depicts the regression results for models predicting objective unmet need among the total sample. When accounting for age, gender, race, health status, and interaction between age, gender, and race, our models predict that immigrants are more likely to experience objective unmet medical need. In terms of educational attainment, the overall population has higher odds of experiencing objective unmet medical needs if they have received less than some college education when compared to bachelor's degree or higher, in line with the original paper. The overall population who are categorized as poor/near poor, low, and middle income also have significantly higher odds of experiencing objective unmet medical need compared to the high-income people, in line with the original paper. The population who are ever uninsured have over three times higher odds of experiencing subjective unmet medical need compared to those insured, and those who belong to any family type other than married also have higher odds of experiencing objective unmet medical need relative to people who are married.

In Panel B, which only evaluates the immigrant sample, Model 1 and Model 2 explain that immigrants with English proficiency tend to experience significantly lower odds of objective unmet medical needs. However, all the models in Panel B prove that immigrants residing within a traditional, low skill region tend to have 41-52% higher odds of objective unmet medical needs. This result is opposite of the result in the original paper. In terms of educational attainment, immigrants who have less than a bachelor's degree have higher odds in experiencing objective unmet medical need; this result is also mostly opposite with the trending result in the original paper. In terms of insurance status, immigrants who are ever uninsured are over three times more likely compared to immigrants who are insured in experiencing objective unmet medical need, which is slightly higher than in the original paper. In terms of family status, immigrants who are never married are almost 40% more likely to experience objective unmet medical need, in line with the result in the original paper.

Table 3: Odds Ratios and 95% Confidence Intervals for Determinants of Objective Unmet Need

Variables	Panel A Total Sample			Panel B Immigrant Sample		
	Model 1	*Model 2*	*Model 3*	*Model 1*	*Model 2*	*Model 3*
Foreign-born						
Foreign-born	1.36*** (1.28, 1.44)	1.25*** (1.14, 1.35)	1.01 (0.93, 1.11)			
Length of time in US						
Midpoint				0.89 (0.71, 1.12)	1.06 (0.83, 1.35)	1.12 (0.87, 1.43)
Midpoint²				1.02 (0.99, 1.05)	0.99 (0.96, 1.02)	0.98 (0.95, 1.01)
Midpoint³				0.99* (0.99, 0.99)	1.00 (0.99, 1.00)	1.00 (0.99, 1.00)
English proficiency						
English interview				0.80*** (0.72, 0.89)	0.84* (0.73, 0.97)	1.11 (0.95, 1.23)
Region destination type (vs. new, high skill region)						
Traditional, low skill region				1.52*** (1.35, 1.72)	1.49*** (1.31, 1.70)	1.41*** (1.23, 1.61)
Education attainment (vs. bachelor's degree or higher)						
Less than high school			1.23** (1.10, 1.37)			1.49*** (1.21, 1.82)
High school diploma			1.23*** (1.12, 1.36)			1.24* (1.00, 1.53)
Some college			1.09 (0.99, 1.19)			1.25* (1.02, 1.54)
Income(vs. high income)						
Poor/near poor			1.19*** (1.08, 1.32)			1.20 (0.98, 1.47)
Low Income			1.33*** (1.20, 1.48)			1.32** (1.08, 1.63)
Middle income			1.25*** (1.15, 1.36)			1.21* (1.01, 1.46)
Uninsured (vs. insured)						
Ever uninsured			3.12*** (2.87, 3.39)			3.11*** (2.72, 3.57)
Family type (vs. married)						
Widowed			1.37** (1.13, 1.66)			1.16 (0.78, 1.71)
Divorced/separated			1.26*** (1.14, 1.38)			1.09 (0.92, 1.31)
Never married			1.34*** (1.23, 1.46)			1.38*** (1.18, 1.62)
Control Variables	No	Yes	Yes	No	Yes	Yes

* P < .05

** P < .01

*** P < .001

Bibliography

1. Hacker, Karen et al. "Barriers to health care for undocumented immigrants: a literature review." *Risk Management and Healthcare Policy*, vol. 8, pp.175-183. 30 Oct. 2015, doi:10.2147/RMHP.S70173

2. Howe Hasanali, Stephanie. "Immigrant-Native Disparities in Perceived and Actual Met/Unmet Need for Medical Care." *Journal of Immigrant and Minority Health*, vol. 17, no. 5, pp. 337-46. 1 Oct. 2016, doi:10.1007/s10903-014-0092-x

3. Rau, Jordan. "HHS Releases Poverty Guidelines For 2014." *Kaiser Health News*, 27 Jan. 2014, khn.org/news/hhs-releases-poverty-guidelines-for-2014/.

4. Santiago, Maria da Conceição and Maria H. Figueiredo. "Immigrant Women's Perspective on Prenatal and Postpartum Care: Systematic Review." *Journal of Immigrant and Minority Health*, vol. 17, 20 Sept. 2013, pp. 276-284, doi: https://doi.org/10.1007/s10903-013-9915-4

5. Wafula, Edith Gonzo and Shedra A. Snipes. "Barriers to Health Care Access Faced by Black Immigrants in the US: Theoretical Considerations and Recommendations." *Journal of Immigrant and Minority Health*, vol. 16, no. 4, 5 Sept. 2013, pp. 689-698, doi: https://doi.org/10.1007/s10903-013-9898-1

Additional Items:

Unique R Packages or Techniques: psych and questionr packages

Statement of original work: We verify that our work is original, and was completed independently with some assistance from Professor Rosenberg

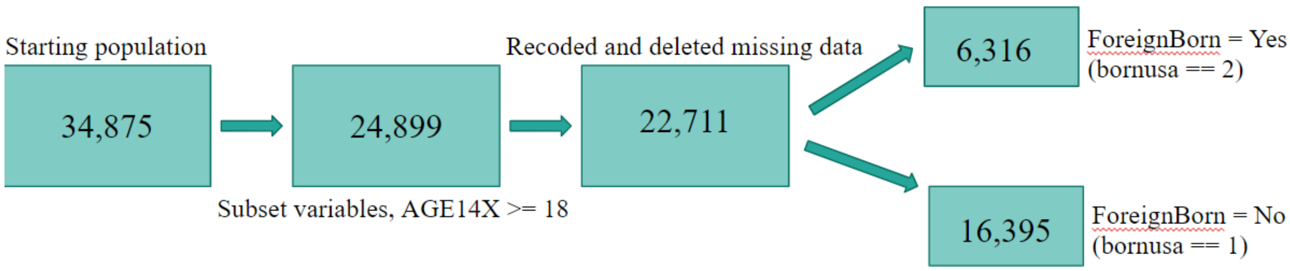
Log of hours for project:

- Choosing Article: 8 hours
- Recoding and cleaning data to create target population: 12 hours
- Statistical analysis: 10 hours
- Presentation preparation: 3 hours
- Writing paper and Literature Reviews: 11 hours

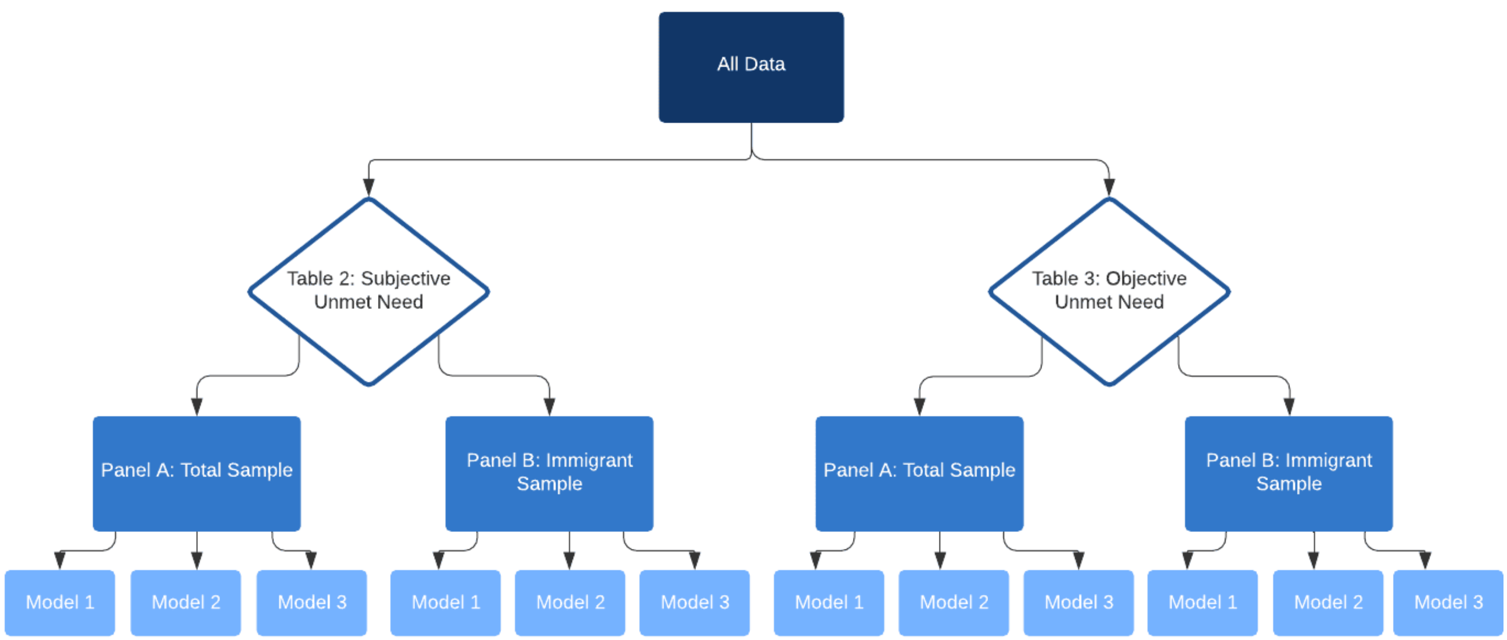
Future Use: This paper may be used for students for reference for future projects.

Appendix:

Flow Chart 1



Flow Chart 2



Appendix:

Data Dictionary

MEPS Variables	**Description**	**Variable Type**
Sociodemographic Covariates		Note – first level is reference category
sex	Sex	Categorical: Binary – Male/Female
racethx	Race	Categorical: Factor – Non-Hispanic white, Non-Hispanic Asian, Non-Hispanic black, Hispanic
MARRY14X	Partnership status	Categorical: Factor – Married, Divorced/Separated, Never married, Widowed Binomial for Table 1*
INSCOV14	Insurance status	Categorical: Binary – Never Uninsured/Ever Uninsured
AGE14X	Age	Continuous (Age)
RTHLTH53	Health Status	Categorical: Factor – Excellent, Very good, Good, Fair/Poor Binomial for Table 1*
eduyrdg	Educational attainment	Categorical: Factor – Bachelor's degree or higher, Some college, High school diploma, Less than HS Binomial for Table 1*
POVCAT14	Income measured as percentage of poverty line	Categorical: Factor – High income, Middle income, Low income, Poor/near poor Binomial for Table 1*
Immigrant-Specific Variables		
yrsinus	Duration of time living in United States	Categorical – Took midpoints of each category, then squared and cubed them
bornusa	Place of birth (United States or elsewhere)	Categorical: Binary – No/Yes
intvlang	English proficiency based on MEPS interview	Categorical: Binary – No/Yes
REGION53	Region	Categorical: Factor – Northeast/Midwest/South/West
OCCCAT53	Occupation	Categorical: Binary – New/Traditional
Outcome Variables		
MDDLAY42	Delayed medical care – Subjective unmet need	Categorical: Binary – No/Yes
CHECK53	Last checkup – Objective unmet need	Categorical: Binary – Within Past Year/Not Within Past Year
*reference category was all other category not equal to the current level – for MARRY14X married variable, reference category is "not married"		

Appendix:

R Code for Analysis:

```
#Clear out workspace

rm(list=ls())

#Load H171 data into workspace

load("C://Users//Health//H171.RData")

#Subset needed variables

keep <-
c("yrsinus","bornusa","sex","intvlang","racethx","MARRY14X","INSCOV14"
,"MDDLAY42","CHECK53","AGE14X", "RTHLTH53", "eduyrdg",
"POVCAT14","OCCCAT53","REGION53")

dat <- H171[keep]

dat <- subset(dat, AGE14X >= 18)

library(psych)

library(questionr)

#Years in US: only asked when the person is not born in the US, so
there are many NAs for this variable

#Years in US midpoint variable: “1” for less than a year, “2” for one
to four years, “3” for five to nine years, “4” for ten to fourteen
years, “5” for over fifteen years

dat <- within(dat, {

  YrsInUSAMid <- NA

  YrsInUSAMid[yrsinus == 1] <- .5

  YrsInUSAMid[yrsinus == 2] <- 2

  YrsInUSAMid[yrsinus == 3] <- 7

  YrsInUSAMid[yrsinus == 4] <- 12

  YrsInUSAMid[yrsinus == 5] <- 15

})
```

```
#Years in US midpoint cubed

dat <- within(dat, {

  YrsInUSACub <- NA

  YrsInUSACub[yrsinus == 1] <- .5^3

  YrsInUSACub[yrsinus == 2] <- 2^3

  YrsInUSACub[yrsinus == 3] <- 7^3

  YrsInUSACub[yrsinus == 4] <- 12^3

  YrsInUSACub[yrsinus == 5] <- 15^3

})


#Years in US midpoint squared

dat <- within(dat, {

  YrsInUSASqr <- NA

  YrsInUSASqr[yrsinus == 1] <- .5^2

  YrsInUSASqr[yrsinus == 2] <- 2^2

  YrsInUSASqr[yrsinus == 3] <- 7^2

  YrsInUSASqr[yrsinus == 4] <- 12^2

  YrsInUSASqr[yrsinus == 5] <- 15^2

})

#Born in the USA determines immigrant status

#Born in USA variable; "1" means born in the USA, "2" means born in other country

dat <- within(dat, {

  ForeignBorn <- NA

  ForeignBorn[bornusa == 1] <- "No"

  ForeignBorn[bornusa == 2] <- "Yes"

})
```

```r
dat$ForeignBorn <- factor(dat$ForeignBorn, levels = c("No","Yes"))


#English proficiency determined by language in which study was conducted – “1” means the study was in English, “2” means the study was not in English

dat <- within(dat, {

  EngProf <- NA

  EngProf[intvlang == 1] <- "Yes"

  EngProf[intvlang != 1] <- "No"

})

dat$EngProf <- factor(dat$EngProf,levels = c("No","Yes"))


#Sex – “1” for male, “2” for female

dat <- within(dat, {

  Sex <- NA

  Sex[sex == 1] <- "Male"

  Sex[sex == 2] <- "Female"

})

dat$Sex <- factor(dat$Sex, levels = c("Male","Female"))


#Race – “1” for Hispanic, “2” for Non-Hispanic white, “3” for Non-Hispanic Black, and “4” for Non-Hispanic Asian

dat <- within(dat, {

  Race <- NA

  Race[racethx == 1] <- "Hispanic"

  Race[racethx == 2] <- "Non-Hispanic white"

  Race[racethx == 3] <- "Non-Hispanic black"
```

```r
  Race[racethx == 4] <- "Non-Hispanic Asian"

})

dat$Race <- factor(dat$Race, levels = c("Non-Hispanic white","Non-Hispanic Asian","Non-Hispanic black","Hispanic"))


#Race binomials

#Binomial for Non-Hispanic white

dat <- within(dat, {

  Race1 <- NA

  Race1[Race == "Non-Hispanic white"] <- "Non-Hispanic white"

  Race1[Race != "Non-Hispanic white" & is.na(Race) == FALSE] <- "Not non-Hispanic white"

})

dat$Race1 <- factor(dat$Race1, levels = c("Not non-Hispanic white", "Non-Hispanic white"))


#Binomial for Non-Hispanic black

dat <- within(dat, {

  Race2 <- NA

  Race2[Race == "Non-Hispanic black"] <- "Non-Hispanic black"

  Race2[Race != "Non-Hispanic black" & is.na(Race) == FALSE] <- "Not non-Hispanic black"

})

dat$Race2 <- factor(dat$Race2, levels = c("Not non-Hispanic black", "Non-Hispanic black"))


#Binomial for Non-Hispanic Asian

dat <- within(dat, {

  Race3 <- NA
```

```
  Race3[Race == "Non-Hispanic Asian"] <- "Non-Hispanic Asian"

  Race3[Race != "Non-Hispanic Asian" & is.na(Race) == FALSE] <- "Not
non-Hispanic Asian"

})

dat$Race3 <- factor(dat$Race3, levels = c("Not non-Hispanic Asian",
"Non-Hispanic Asian"))


#Binomial for Hispanic

dat <- within(dat, {

  Race4 <- NA

  Race4[Race == "Hispanic"] <- "Hispanic"

  Race4[Race != "Hispanic" & is.na(Race) == FALSE] <- "Not Hispanic"

})

dat$Race4 <- factor(dat$Race4, levels = c("Not Hispanic", "Hispanic"))


#Health insurance coverage - only looking at insured vs. uninsured

#Health insurance coverage - "1" for private coverate and "2" for
public coverage(never uninsured), "3" for Uninsured (ever uninsured)

dat <- within(dat, {

  HealthInsCov <- NA #initialize variable within empty placeholders

  HealthInsCov[INSCOV14 == 1 | INSCOV14 == 2] <- "Never Uninsured"

  HealthInsCov[INSCOV14 == 3] <- "Ever Uninsured"

})

dat$HealthInsCov <- factor(dat$HealthInsCov, levels = c("Never
Uninsured", "Ever Uninsured"))


#Marriage categorical variable

dat <- within(dat, {
```

```
  Marriage <- NA #initialize variable within empty placeholders
  Marriage[MARRY14X == 1] <- "Married"
  Marriage[MARRY14X == 3 | MARRY14X == 4] <- "Divorced/Separated"
  Marriage[MARRY14X == 5] <- "Never married"
  Marriage[MARRY14X == 2] <- "Widowed"
  Marriage[MARRY14X == -9 | MARRY14X == -8 | MARRY14X == -7 ] <- NA
})
dat$Marriage <- factor(dat$Marriage, levels = c("Married",
"Divorced/Separated","Never married", "Widowed"))


#Binomials for marriage variable
#Binomial for married
dat <- within(dat, {
  Marriage1 <- NA #initialize variable within empty placeholders
  Marriage1[MARRY14X == 1] <- "Married"
  Marriage1[MARRY14X != 1] <- "Not Married"
  Marriage1[MARRY14X == -9 | MARRY14X == -8 | MARRY14X == -7 ] <- NA
})
dat$Marriage1 <- factor(dat$Marriage1, levels = c("Not Married",
"Married"))


#Binomial for divorced/separated
dat <- within(dat, {
  Marriage2 <- NA #initialize variable within empty placeholders
  Marriage2[MARRY14X == 3 | MARRY14X == 4] <- "Divorced/Separated"
  Marriage2[MARRY14X != 3 & MARRY14X != 4] <- "Not divorced/separated"
  Marriage2[MARRY14X == -9 | MARRY14X == -8 | MARRY14X == -7 ] <- NA
```

```
})

dat$Marriage2 <- factor(dat$Marriage2, levels = c("Not
divorced/separated", "Divorced/Separated"))


#Binomial for never married

dat <- within(dat, {

  Marriage3 <- NA #initialize variable within empty placeholders

  Marriage3[MARRY14X == 5] <- "Never married"

  Marriage3[MARRY14X != 5] <- "Not never married"

  Marriage3[MARRY14X == -9 | MARRY14X == -8 | MARRY14X == -7 ] <- NA

})

dat$Marriage3 <- factor(dat$Marriage3, levels = c("Not never married",
"Never married"))


#Binomial for widowed

dat <- within(dat, {

  Marriage4 <- NA #initialize variable within empty placeholders

  Marriage4[MARRY14X == 2] <- "Widowed"

  Marriage4[MARRY14X != 2] <- "Not widowed"

  Marriage4[MARRY14X == -9 | MARRY14X == -8 | MARRY14X == -7 ] <- NA

})

dat$Marriage4 <- factor(dat$Marriage4, levels = c("Not widowed",
"Widowed"))


#Outcome variables

#MDDLAY42 for delayed medical care – "1" for yes, "2" for no

dat <- within(dat, {

  DelayedMedCare <- NA #initialize variable within empty placeholders
```

```
  DelayedMedCare[MDDLAY42 == -8 | MDDLAY42 == -7 | MDDLAY42 == -1] <-
NA

  DelayedMedCare[MDDLAY42 == 1] <- "Yes"

  DelayedMedCare[MDDLAY42 == 2] <- "No"

})

dat$DelayedMedCare <- factor(dat$DelayedMedCare, levels = c("No",
"Yes"))


#LastCheckup for time since last checkup within last year – "1" for
within past year, "2" for not within past year

dat <- within(dat, {

  LastCheckup <- NA

  LastCheckup[CHECK53 == -9 | CHECK53 == -8 | CHECK53 == -7 | CHECK53
== -1] <- NA

  LastCheckup[CHECK53 == 1] <- "Within Past Year"

  LastCheckup[CHECK53 > 1] <- "Not Within Past Year"

})

dat$LastCheckup <- factor(dat$LastCheckup, levels = c("Within Past
Year", "Not Within Past Year"))


#Age – continuous variable for table 1 stats

dat <- within(dat, {

  Age <- AGE14X # initialize variable

  Age[AGE14X == -1] <- NA

})


#Education levels – categorical

dat <- within(dat, {

  EducNew <- NA  # note need this statement to initialize the variable
```

```
  EducNew[eduyrdg == 5 | eduyrdg == 6 | eduyrdg == 7] <- "Some
college"

  EducNew[eduyrdg == 3 | eduyrdg == 4] <- "High school diploma"

  EducNew[eduyrdg == 1 | eduyrdg == 2] <- "Less than HS"

  EducNew[eduyrdg == 8 | eduyrdg == 9] <- "Bachelor's degree or
higher"

})

dat$EducNew <- factor(dat$EducNew, levels = c("Bachelor's degree or
higher","Some college","High school diploma","Less than HS"))



#Education level - binomials

#Binomial for some college

dat <- within(dat, {

  EducNew1 <- NA  # note need this statement to initialize the
variable

  EducNew1[eduyrdg == 5 | eduyrdg == 6 | eduyrdg == 7] <- "Some
college"

  EducNew1[eduyrdg == 8 | eduyrdg == 9 | (eduyrdg < 5 & eduyrdg > 0)]
<- "Not some college"

})

dat$EducNew1 <- factor(dat$EducNew1, levels = c("Not some
college","Some college"))



#Binomial for high school diploma

dat <- within(dat, {

  EducNew2 <- NA  # note need this statement to initialize the
variable

  EducNew2[eduyrdg == 3 | eduyrdg == 4] <- "High school diploma"

  EducNew2[eduyrdg == 8 | eduyrdg == 9 | eduyrdg == 5 | eduyrdg == 6 |
eduyrdg == 7 | (eduyrdg < 3 & eduyrdg > 0)] <- "Not high school
diploma"
```

```
})

dat$EducNew2 <- factor(dat$EducNew2, levels = c("Not high school
diploma","High school diploma"))


#Binomial for less than high school

dat <- within(dat, {

  EducNew3 <- NA  # note need this statement to initialize the
variable

  EducNew3[eduyrdg == 1 | eduyrdg == 2] <- "Less than HS"

  EducNew3[eduyrdg > 2] <- "Not less than HS"

})

dat$EducNew3 <- factor(dat$EducNew3, levels = c("Not less than
HS","Less than HS"))


#Binomial for bachelor's degree or higher

dat <- within(dat, {

  EducNew4 <- NA  # note need this statement to initialize the
variable

  EducNew4[eduyrdg == 8 | eduyrdg == 9] <- "Bachelor's degree or
higher"

  EducNew4[eduyrdg < 8 & eduyrdg > 0] <- "Not Bachelor's degree or
higher"

})

dat$EducNew4 <- factor(dat$EducNew4, levels = c("Not Bachelor's degree
or higher","Bachelor's degree or higher"))


#Health Status (as perceived by participant) - categorical

dat <- within(dat, {

  HealthStat <- NA #initialize variable

  HealthStat[RTHLTH53 == -1] <- NA
```

```r
  HealthStat[RTHLTH53 == 2] <- "Very good"
  HealthStat[RTHLTH53 == 3] <- "Good"
  HealthStat[RTHLTH53 == 4 | RTHLTH53 == 5] <- "Fair/Poor"
  HealthStat[RTHLTH53 == 1] <- "Excellent"
})

dat$HealthStat <- factor(dat$HealthStat, levels = c("Excellent", "Very good", "Good","Fair/Poor"))

#Health status binomials
#Binomial for “very good” health
dat <- within(dat, {
  HealthStat1 <- NA #initialize variable
  HealthStat1[RTHLTH53 == -1] <- NA
  HealthStat1[RTHLTH53 == 2] <- "Very good"
  HealthStat1[RTHLTH53 > 0 & RTHLTH53 != 2] <- "Not very good"
})
dat$HealthStat1 <- factor(dat$HealthStat1, levels = c("Not very good", "Very good"))

#Binomial for “good” health
dat <- within(dat, {
  HealthStat2 <- NA #initialize variable
  HealthStat2[RTHLTH53 == -1] <- NA
  HealthStat2[RTHLTH53 == 3] <- "Good"
  HealthStat2[RTHLTH53 > 0 & RTHLTH53 != 3] <- "Not good"
})
```

```
dat$HealthStat2 <- factor(dat$HealthStat2, levels = c("Not good",
"Good"))


#Binomial for "fair/poor" health

dat <- within(dat, {

  HealthStat3 <- NA #initialize variable

  HealthStat3[RTHLTH53 == -1] <- NA

  HealthStat3[RTHLTH53 == 4 | RTHLTH53 == 5] <- "Fair/Poor"

  HealthStat3[RTHLTH53 > 0 & RTHLTH53 != 4] <- "Not fair/poor"

})

dat$HealthStat3 <- factor(dat$HealthStat3, levels = c("Not fair/poor",
"Fair/Poor"))


#Binomial for "excellent" health

dat <- within(dat, {

  HealthStat4 <- NA #initialize variable

  HealthStat4[RTHLTH53 == -1] <- NA

  HealthStat4[RTHLTH53 == 1] <- "Excellent"

  HealthStat4[RTHLTH53 > 1] <- "Not excellent"

})

dat$HealthStat4 <- factor(dat$HealthStat4, levels = c("Not excellent",
"Excellent"))


#Family income as percentage of poverty line

#Family income - categorical

dat <- within(dat, {

  Income <- NA

  Income[POVCAT14 == 1 | POVCAT14 == 2] <- "Poor/near poor"
```

```
  Income[POVCAT14 == 3] <- "Low income"

  Income[POVCAT14 == 4] <- "Middle income"

  Income[POVCAT14 == 5] <- "High income"

  })

dat$Income <- factor(dat$Income, levels = c("High income", "Middle
income","Low income","Poor/near poor"))


#Family Income binomials

#Binomial for poor/near poor

dat <- within(dat, {

  Income1 <- NA

  Income1[POVCAT14 == 1 | POVCAT14 == 2] <- "Poor/near poor"

  Income1[POVCAT14 > 0 & POVCAT14 != 1 & POVCAT14 != 2] <- "Not
poor/near poor"

})

dat$Income1 <- factor(dat$Income1, levels = c("Not poor/near poor",
"Poor/near poor"))


#Binomial for low income

dat <- within(dat, {

  Income2 <- NA

  Income2[POVCAT14 == 3] <- "Low income"

  Income2[POVCAT14 != 3 & POVCAT14 > 0] <- "Not low income"

})

dat$Income2 <- factor(dat$Income2, levels = c("Not low income","Low
income"))


#Binomial for middle income
```

```
dat <- within(dat, {

  Income3 <- NA

  Income3[POVCAT14 == 4] <- "Middle income"

  Income3[POVCAT14 != 4 & POVCAT14 > 0] <- "Not middle income"

})

dat$Income3 <- factor(dat$Income3, levels = c("Not middle
income","Middle income"))


#Binomial for high income

dat <- within(dat, {

  Income4 <- NA

  Income4[POVCAT14 == 5] <- "High income"

  Income4[POVCAT14 != 5 & POVCAT14 > 0] <- "Not high income"

})

dat$Income4 <- factor(dat$Income4, levels = c("Not high income","High
income"))


#Type of occupation - split into new (1-5) vs. traditional (6-9) types

dat <- within(dat, {

  OccCat <- NA

  OccCat[OCCCAT53 >= 1 & OCCCAT53 <= 5] <- "New"

  OccCat[OCCCAT53 >= 6 & OCCCAT53 <= 9] <- "Traditional"

})

dat$OccCat <- factor(dat$OccCat, levels = c("New","Traditional"))


#Region of residence – "1" for Northeast, "2" for Midwest, "3" for
South, "4" for West

dat <- within(dat, {
```

```r
  Region <- NA

  Region[REGION53 == 1] <- "Northeast"

  Region[REGION53 == 2] <- "Midwest"

  Region[REGION53 == 3] <- "South"

  Region[REGION53 == 4] <- "West"

})


dat$Region <- factor(dat$Region, levels =
c("Northeast","Midwest","South","West"))


#Remove negative vals for eduyrdg, marriage, and health stat

dat <- dat[dat$eduyrdg > 0,]

dat <- dat[dat$MARRY14X > 0,]

dat <- dat[dat$RTHLTH53 > 0,]


#Remove unnecessary variables

dat <- subset(dat, select = -
c(yrsinus,bornusa,sex,intvlang,racethx,MARRY14X,INSCOV14,MDDLAY42,CHEC
K53,AGE14X, RTHLTH53, eduyrdg, POVCAT14,OCCCAT53,REGION53))

#Remove data that has NA for any columns

#Did not include YrsInUSA because this is NA for US natives

#Also did not include the added columns - region & occupation type
(these are only used for defining their region type, and including
would cause loss of most of data)

dat <- dat[!(is.na(dat$ForeignBorn)) & !(is.na(dat$EngProf)) &
!(is.na(dat$Sex)) & !(is.na(dat$Race)) & !(is.na(dat$HealthInsCov)) &
!(is.na(dat$DelayedMedCare)) & !(is.na(dat$LastCheckup))
!(is.na(dat$Age)),]


#Determine skill ratio of each region (high/low)
```

```
#SR >= .25 is high and < .25 is low

NortheastSR <- sum(dat$ForeignBorn == "Yes" & dat$EducNew4 ==
"Bachelor's degree or higher" & dat$Region == "Northeast", na.rm =
TRUE)/sum(dat$ForeignBorn == "Yes" & dat$EducNew4 != "Bachelor's
degree or higher" & dat$Region == "Northeast", na.rm = TRUE)

MidwestSR <- sum(dat$ForeignBorn == "Yes" & dat$EducNew4 ==
"Bachelor's degree or higher" & dat$Region == "Midwest", na.rm =
TRUE)/sum(dat$ForeignBorn == "Yes" & dat$EducNew4 != "Bachelor's
degree or higher" & dat$Region == "Midwest", na.rm = TRUE)

SouthSR <- sum(dat$ForeignBorn == "Yes" & dat$EducNew4 == "Bachelor's
degree or higher" & dat$Region == "South", na.rm =
TRUE)/sum(dat$ForeignBorn == "Yes" & dat$EducNew4 != "Bachelor's
degree or higher" & dat$Region == "South", na.rm = TRUE)

WestSR <- sum(dat$ForeignBorn == "Yes" & dat$EducNew4 == "Bachelor's
degree or higher" & dat$Region == "West", na.rm =
TRUE)/sum(dat$ForeignBorn == "Yes" & dat$EducNew4 != "Bachelor's
degree or higher" & dat$Region == "West", na.rm = TRUE)



#Determine occupation ratio of each region (new/traditional)

#OR >= 2.25 is high and < 2.25 is low

NortheastOccR <- sum(dat$ForeignBorn == "Yes" & dat$OccCat == "New" &
dat$Region == "Northeast", na.rm = TRUE)/sum(dat$ForeignBorn == "Yes"
& dat$OccCat == "Traditional" & dat$Region == "Northeast", na.rm =
TRUE)

MidwestOccR <- sum(dat$ForeignBorn == "Yes" & dat$OccCat == "New" &
dat$Region == "Midwest", na.rm = TRUE)/sum(dat$ForeignBorn == "Yes" &
dat$OccCat == "Traditional" & dat$Region == "Midwest", na.rm = TRUE)

SouthOccR <- sum(dat$ForeignBorn == "Yes" & dat$OccCat == "New" &
dat$Region == "South", na.rm = TRUE)/sum(dat$ForeignBorn == "Yes" &
dat$OccCat == "Traditional" & dat$Region == "South", na.rm = TRUE)

WestOccR <- sum(dat$ForeignBorn == "Yes" & dat$OccCat == "New" &
dat$Region == "West", na.rm = TRUE)/sum(dat$ForeignBorn == "Yes" &
dat$OccCat == "Traditional" & dat$Region == "West", na.rm = TRUE)



#Determine region destination type

#Northeast: High SR, High OR

#Midwest: High SR, High OR
```

```
#South: Low SR, Low OR

#West: Low SR, Low OR

#So we will only have two levels: high SR, high OR and low SR, low OR

dat <- within(dat, {

  RegionDestType <- NA

  RegionDestType[Region == "Northeast" | Region == "Midwest"] <- "New, high-skill"

  RegionDestType[Region == "South" | Region == "West"] <- "Traditional, low-skill"

})

factor(dat$RegionDestType,levels = c("New, high-skill","Traditional, low-skill"))


#Add age^2 and age^3 to dat for glm

dat <- within(dat, {

  Age2 <- Age^2

  Age3 <- Age^3

})


#View basic stats

summary(dat)

psych::describe(dat)


#Table 1 US data

USdat <- dat[dat$ForeignBorn == "No",]

USSubUnmetNeed <- sum(USdat$DelayedMedCare == "Yes")/nrow(USdat)

USObUnmetNeed <- sum(USdat$LastCheckup == "Not Within Past Year")/nrow(USdat)
```

```
USForeignBorn <- sum(USdat$ForeignBorn == "Yes")/nrow(USdat)

USAgeMean <- mean(USdat$Age)

USAgesd <- sd(USdat$Age)

USFemale <- sum(USdat$Sex == "Female")/nrow(USdat)

USRaceWhite <- sum(USdat$Race == "Non-Hispanic white")/nrow(USdat)

USRaceBlack <- sum(USdat$Race == "Non-Hispanic black")/nrow(USdat)

USRaceHispanic <- sum(USdat$Race == "Hispanic")/nrow(USdat)

USRaceAsian <- sum(USdat$Race == "Non-Hispanic Asian")/nrow(USdat)

USHealthStatPoor <- sum(USdat$HealthStat3 == "Fair/Poor",
na.rm=TRUE)/nrow(USdat)

USHealthStatGood <- sum(USdat$HealthStat2 == "Good", na.rm =
TRUE)/nrow(USdat)

USHealthStatVGood <- sum(USdat$HealthStat1 == "Very good", na.rm =
TRUE)/nrow(USdat)

USHealthStatExcellent <- sum(USdat$HealthStat4 == "Excellent", na.rm =
TRUE)/nrow(USdat)

USEducLessThanHS <- sum(USdat$EducNew3 == "Less than HS", na.rm =
TRUE)/nrow(USdat)

USEducHS <- sum(USdat$EducNew2 == "High school diploma", na.rm =
TRUE)/nrow(USdat)

USEducSomeCollege <- sum(USdat$EducNew1 == "Some college", na.rm =
TRUE)/nrow(USdat)

USEducBachelors <- sum(USdat$EducNew4 == "Bachelor's degree or
higher", na.rm = TRUE)/nrow(USdat)

USIncPoor <- sum(USdat$Income1 == "Poor/near poor", na.rm =
TRUE)/nrow(USdat)

USIncLow <- sum(USdat$Income2 == "Low income", na.rm =
TRUE)/nrow(USdat)

USIncMid <- sum(USdat$Income3 == "Middle income", na.rm =
TRUE)/nrow(USdat)

USIncHigh <- sum(USdat$Income4 == "High income", na.rm =
TRUE)/nrow(USdat)
```

```
USUninsured <- sum(USdat$HealthInsCov == "Ever Uninsured", na.rm =
TRUE)/nrow(USdat)

USFamMarried <- sum(USdat$Marriage1 == "Married", na.rm =
TRUE)/nrow(USdat)

USFamWidowed <- sum(USdat$Marriage4 == "Widowed", na.rm =
TRUE)/nrow(USdat)

USFamDivorced <- sum(USdat$Marriage2 == "Divorced/Separated",na.rm =
TRUE)/nrow(USdat)

USFamNeverMarried <- sum(USdat$Marriage3 == "Never married", na.rm =
TRUE)/nrow(USdat)

USYrsInUSMid <- NA

USEnglishIntv <- NA

USRegionDestTypeNH <- NA

USRegionDestTypeTL <- NA



USTable1 <-
data.frame(USSubUnmetNeed,USObUnmetNeed,USForeignBorn,USAgeMean,USAges
d,USFemale,USRaceWhite,USRaceBlack,USRaceHispanic,USRaceAsian,USHealth
StatPoor,USHealthStatGood,USHealthStatVGood,USHealthStatExcellent,USEd
ucLessThanHS,USEducHS,USEducSomeCollege,USEducBachelors,USIncPoor,USIn
cLow,USIncMid,USIncHigh,USUninsured,USFamMarried,USFamWidowed,USFamDiv
orced,USFamNeverMarried,USYrsInUSMid,USEnglishIntv,USRegionDestTypeNH,
USRegionDestTypeTL)

USTable1 <- data.frame(t(USTable1))



#Table 1 Foreign Born/Immigrant data

Immdat <- dat[dat$ForeignBorn == "Yes",]

ImmSubUnmetNeed <- sum(Immdat$DelayedMedCare == "Yes")/nrow(Immdat)

ImmObjUnmetNeed <- sum(Immdat$LastCheckup == "Not Within Past
Year")/nrow(Immdat)

ImmForeignBorn <- sum(Immdat$ForeignBorn == "Yes")/nrow(Immdat)

ImmAgeMean <- mean(Immdat$Age)

ImmAgeSD <- sd(Immdat$Age)
```

```
ImmFemale <- sum(Immdat$Sex == "Female")/nrow(Immdat)

ImmRaceWhite <- sum(Immdat$Race == "Non-Hispanic white")/nrow(Immdat)

ImmRaceBlack <- sum(Immdat$Race == "Non-Hispanic black")/nrow(Immdat)

ImmRaceHispanic <- sum(Immdat$Race == "Hispanic")/nrow(Immdat)

ImmRaceAsian <- sum(Immdat$Race == "Non-Hispanic Asian")/nrow(Immdat)

ImmHealthStatPoor <- sum(Immdat$HealthStat3 == "Fair/Poor",
na.rm=TRUE)/nrow(Immdat)

ImmHealthStatGood <- sum(Immdat$HealthStat2 == "Good", na.rm =
TRUE)/nrow(Immdat)

ImmHealthStatVGood <- sum(Immdat$HealthStat1 == "Very good", na.rm =
TRUE)/nrow(Immdat)

ImmHealthStatExcellent <- sum(Immdat$HealthStat4 == "Excellent", na.rm
= TRUE)/nrow(Immdat)

ImmEducLessThanHS <- sum(Immdat$EducNew3 == "Less than HS", na.rm =
TRUE)/nrow(Immdat)

ImmEducHS <- sum(Immdat$EducNew2 == "High school diploma", na.rm =
TRUE)/nrow(Immdat)

ImmEducSomeCollege <- sum(Immdat$EducNew1 == "Some college", na.rm =
TRUE)/nrow(Immdat)

ImmEducBachelors <- sum(Immdat$EducNew4 == "Bachelor's degree or
higher", na.rm = TRUE)/nrow(Immdat)

ImmIncPoor <- sum(Immdat$Income1 == "Poor/near poor", na.rm =
TRUE)/nrow(Immdat)

ImmIncLow <- sum(Immdat$Income2 == "Low income", na.rm =
TRUE)/nrow(Immdat)

ImmIncMid <- sum(Immdat$Income3 == "Middle income", na.rm =
TRUE)/nrow(Immdat)

ImmIncHigh <- sum(Immdat$Income4 == "High income", na.rm =
TRUE)/nrow(Immdat)

ImmUninsured <- sum(Immdat$HealthInsCov == "Ever Uninsured", na.rm =
TRUE)/nrow(Immdat)

ImmFamMarried <- sum(Immdat$Marriage1 == "Married", na.rm =
TRUE)/nrow(Immdat)
```

```
ImmFamWidowed <- sum(Immdat$Marriage4 == "Widowed", na.rm =
TRUE)/nrow(Immdat)

ImmFamDivorced <- sum(Immdat$Marriage2 == "Divorced/Separated",na.rm =
TRUE)/nrow(Immdat)

ImmFamNeverMarried <- sum(Immdat$Marriage3 == "Never married", na.rm =
TRUE)/nrow(Immdat)

ImmYrsInUSMid <- mean(Immdat$YrsInUSAMid,na.rm=TRUE)

ImmEnglishIntv <- sum(Immdat$EngProf == "Yes",na.rm=TRUE)/nrow(Immdat)

ImmRegionDestTypeNH <- sum(Immdat$RegionDestType == "New, high-
skill",na.rm=TRUE)/nrow(Immdat)

ImmRegionDestTypeTL <- sum(Immdat$RegionDestType == "Traditional, low-
skill",na.rm = TRUE)/nrow(Immdat)



ImmTable1 <-
data.frame(ImmSubUnmetNeed,ImmObjUnmetNeed,ImmForeignBorn,ImmAgeMean,I
mmAgeSD,ImmFemale,ImmRaceWhite,ImmRaceBlack,ImmRaceHispanic,ImmRaceAsi
an,ImmHealthStatPoor,ImmHealthStatGood,ImmHealthStatVGood,ImmHealthSta
tExcellent,ImmEducLessThanHS,ImmEducHS,ImmEducSomeCollege,ImmEducBache
lors,ImmIncPoor,ImmIncLow,ImmIncMid,ImmIncHigh,ImmUninsured,ImmFamMarr
ied,ImmFamWidowed,ImmFamDivorced,ImmFamNeverMarried,ImmYrsInUSMid,ImmE
nglishIntv,ImmRegionDestTypeNH,ImmRegionDestTypeTL)

ImmTable1 <- data.frame(t(ImmTable1))



#Table 1 total data

total_bachelors <- sum(dat$EducNew4 == "Bachelor's degree or higher",
na.rm = TRUE)/nrow(dat)

total_somecollege <- sum(dat$EducNew1 == "Some college", na.rm =
TRUE)/nrow(dat)

total_highschool <- sum(dat$EducNew2 == "High school diploma", na.rm =
TRUE)/nrow(dat)

total_lessthanhs <- sum(dat$EducNew3 == "Less than HS", na.rm =
TRUE)/nrow(dat)

total_hs_excellent <- sum(dat$HealthStat4 == "Excellent", na.rm =
TRUE)/nrow(dat)
```

```
total_hs_verygood <- sum(dat$HealthStat1 == "Very good", na.rm =
TRUE)/nrow(dat)

total_hs_good <- sum(dat$HealthStat2 == "Good", na.rm =
TRUE)/nrow(dat)

total_hs_fairpoor <-  sum(dat$HealthStat3 == "Fair/Poor",
na.rm=TRUE)/nrow(dat)

total_inc_high <- sum(dat$Income4 == "High income", na.rm =
TRUE)/nrow(dat)

total_inc_poor <- sum(dat$Income1 == "Poor/near poor", na.rm =
TRUE)/nrow(dat)

total_inc_low <- sum(dat$Income2 == "Low income", na.rm =
TRUE)/nrow(dat)

total_inc_middle <- sum(dat$Income3 == "Middle income", na.rm =
TRUE)/nrow(dat)

total_age_mean <- mean(dat$Age)

total_age_sd <- sd(dat$Age)

total_sub_unmet <- sum(dat$DelayedMedCare == "Yes")/nrow(dat)

total_ob_unmet <- sum(dat$LastCheckup == "Not Within Past
Year")/nrow(dat)

total_foreign_born <- sum(dat$ForeignBorn == "Yes")/nrow(dat)

total_female <- sum(dat$Sex == "Female")/nrow(dat)

total_white <- sum(dat$Race == "Non-Hispanic white")/nrow(dat)

total_black <- sum(dat$Race == "Non-Hispanic black")/nrow(dat)

total_hispanic <- sum(dat$Race == "Hispanic")/nrow(dat)

total_asian <- sum(dat$Race == "Non-Hispanic Asian")/nrow(dat)

total_uninsured <- sum(dat$HealthInsCov == "Ever Uninsured", na.rm =
TRUE)/nrow(dat)

total_married <- sum(dat$Marriage1 == "Married", na.rm =
TRUE)/nrow(dat)

total_widowed <- sum(dat$Marriage4 == "Widowed", na.rm =
TRUE)/nrow(dat)
```

```
total_divorced <- sum(dat$Marriage2 == "Divorced/Separated",na.rm =
TRUE)/nrow(dat)

total_nevermarried <-  sum(dat$Marriage3 == "Never married", na.rm =
TRUE)/nrow(dat)

total_yrsinusmid <- NA

total_englishintv <- NA

total_regiondesttypeNH <- NA

total_regiondesttypeTL <- NA



TotTable1 <-
data.frame(total_sub_unmet,total_ob_unmet,total_foreign_born,total_age
_mean,total_age_sd,total_female,total_white,total_black,total_hispanic
,total_asian,total_hs_fairpoor,total_hs_good,total_hs_verygood,total_h
s_excellent,total_lessthanhs,total_highschool,total_somecollege,total_
bachelors,total_inc_poor,total_inc_low,total_inc_middle,total_inc_high
,total_uninsured,total_married,total_widowed,total_divorced,total_neve
rmarried,total_yrsinusmid,total_englishintv,total_regiondesttypeNH,tot
al_regiondesttypeTL)

TotTable1 <- data.frame(t(TotTable1))



#Combine tables to create full Table 1

Table1 <- cbind(TotTable1,USTable1,ImmTable1)

row.names(Table1) <-
c("UnmetNeed_Subjective","UnmetNeed_Objective","Foreign_Born","Age_Mea
n","Age_sd","Female","Race_White","Race_Black","Race_Hispanic","Race_A
sian","HealthStat_FairPoor","HealthStat_Good","HealthStat_VeryGood","H
ealthStat_Excellent","Education_LessThanHS","Education_HS","Education_
SomeCollege","Education_Bachelors","Inc_Poor","Inc_Low","Inc_Middle","
Inc_High","Ever_Uninsured","MarriageStat_Married","MarriageStat_Widowe
d","MarriageStat_Divorced","MarriageStat_NeverMarried","YrsInUSMidpoin
t","EnglishProficiency","RegionDestType_NH","RegionDestType_TL")



#Calculate model odds ratio and confidence intervals for Table 2

#Table 2, Panel A, Model 1

tab2_panA_mod1 <- glm(DelayedMedCare ~ ForeignBorn, family =
binomial(link = "logit"), data = dat)
```

```
summary(tab2_panA_mod1)

odds.ratio(tab2_panA_mod1)


#Table 2, Panel A, Model 2

tab2_panA_mod2 <- glm(DelayedMedCare ~ ForeignBorn + Age + Age2 + Age3
+ Sex + Race + HealthStat + Age:Sex + Age:Race + Sex:Race, family =
binomial(link = "logit"), data = dat)

summary(tab2_panA_mod2)

odds.ratio(tab2_panA_mod2)


#Table 2, Panel A, Model 3

tab2_panA_mod3 <- glm(DelayedMedCare ~ ForeignBorn + EducNew + Income
+ HealthInsCov + Marriage + Age + Age2 + Age3 + Sex + Race +
HealthStat + Age:Sex + Age:Race + Sex:Race, family = binomial(link =
"logit"), data = dat)

summary(tab2_panA_mod3)

odds.ratio(tab2_panA_mod3)


#Table 2, Panel B, Model 1

tab2_panB_mod1 <- glm(DelayedMedCare ~ EngProf + RegionDestType +
YrsInUSAMid + YrsInUSASqr + YrsInUSACub, family = binomial(link =
"logit"), data = Immdat)

summary(tab2_panB_mod1)

odds.ratio(tab2_panB_mod1)


#Table 2, Panel B, Model 2

tab2_panB_mod2 <- glm(DelayedMedCare ~ EngProf + RegionDestType + Age
+ Age2 + Age3 + Sex + Race + HealthStat + Age:Sex + Age:Race +
Sex:Race + YrsInUSAMid + YrsInUSASqr + YrsInUSACub, family =
binomial(link = "logit"), data = Immdat)

summary(tab2_panB_mod2)
```

```
odds.ratio(tab2_panB_mod2)


#Table 2, Panel B, Model 3

tab2_panB_mod3 <- glm(DelayedMedCare ~ EngProf + RegionDestType +
EducNew + Income + HealthInsCov + Marriage + Age + Age2 + Age3 + Sex +
Race + HealthStat + Age:Sex + Age:Race + Sex:Race + YrsInUSAMid +
YrsInUSASqr + YrsInUSACub, family = binomial(link = "logit"), data =
Immdat)

summary(tab2_panB_mod3)

odds.ratio(tab2_panB_mod3)


#Calculte odds ratios and confidence intervals for Table 3

#Table 3, Panel A, Model 1

tab3_panA_mod1 <- glm(LastCheckup ~ ForeignBorn, family =
binomial(link = "logit"), data = dat)

summary(tab3_panA_mod1)

odds.ratio(tab3_panA_mod1)


#Table 3, Panel A, Model 2

tab3_panA_mod2 <- glm(LastCheckup ~ ForeignBorn + Age + Age2 + Age3 +
Sex + Race + HealthStat + Age:Sex + Age:Race + Sex:Race, family =
binomial(link = "logit"), data = dat)

summary(tab3_panA_mod2)

odds.ratio(tab3_panA_mod2)


#Table 3, Panel A, Model 3

tab3_panA_mod3 <- glm(LastCheckup ~ ForeignBorn + EducNew + Income +
HealthInsCov + Marriage + Age + Age2 + Age3 + Sex + Race + HealthStat
+ Age:Sex + Age:Race + Sex:Race, family = binomial(link = "logit"),
data = dat)

summary(tab3_panA_mod3)
```

```r
odds.ratio(tab3_panA_mod3)


#Table 3, Panel B, Model 1

tab3_panB_mod1 <- glm(LastCheckup ~ EngProf + RegionDestType +
YrsInUSAMid + YrsInUSASqr + YrsInUSACub, family = binomial(link =
"logit"), data = Immdat)

summary(tab3_panB_mod1)

odds.ratio(tab3_panB_mod1)


#Table 3, Panel B, Model 2

tab3_panB_mod2 <- glm(LastCheckup ~ EngProf + RegionDestType + Age +
Age2 + Age3 + Sex + Race + HealthStat + Age:Sex + Age:Race + Sex:Race
+ YrsInUSAMid + YrsInUSASqr + YrsInUSACub, family = binomial(link =
"logit"), data = Immdat)

summary(tab3_panB_mod2)

odds.ratio(tab3_panB_mod2)


#Table 3, Panel B, Model 3

tab3_panB_mod3 <- glm(LastCheckup ~ EngProf + RegionDestType + EducNew
+ Income + HealthInsCov + Marriage + Age + Age2 + Age3 + Sex + Race +
HealthStat + Age:Sex + Age:Race + Sex:Race + YrsInUSAMid + YrsInUSASqr
+ YrsInUSACub, family = binomial(link = "logit"), data = Immdat)

summary(tab3_panB_mod3)

odds.ratio(tab3_panB_mod3)
```