Mental Health in the National Health Interview Study: 2007-2013

Vincent La

June 12, 2017

Introduction

The consequences of neglecting mental health can be just as serious as neglecting one's phyisical health. For example, the National Institute of Mental Health esetimated in 2015 that 6.7% of US adults suffered an episode of major depression, with 2.4% of adults suffering an episode which severely impaired their daily lives. ("Major Depression with Severe Impairment Among Adults," n.d.) However, mental illness is often referred to as an invisible illness because its afflictions are not as obvious as those caused by physical illnesses. While the quality of mental health services has risen throughout the years, people remain reticent on their mental health. Stigma associated with being labelled "mentally ill" is well-documented as a primary reason for this reluctance in both the general American population, (Corrigan 2004) and in combat veterans returning from duty in Iraq and Afghanistan. (Hoge et al. 2004)

Data and Methodology

This study utilized logistic regression, decision trees, and random forests to analyze the *National Health Interview Survey* data from 2007-2013. The National Health Interview Survey is one of the largest and most comprehensive surveys of Americans. Organized by the Centers for Disease Control and carried out with Census Bureau employees, this annual survey interviews a random but representative sample of American households every year. Each year's survey data is divided into several files, each containing a specific set of variables. For example, in every household the interviewer may ask a random adult more detailed questions–answers which are then recoded in the Sample Adult file. The majority of this study's analysis was done by linking the person, family, and sample adult level files.

Aggregating Data

In order to gain higher precision estimates, and to be able to analyze changes over time, data from different survey years were aggregated. However, this was not straightforward. Due to sampling design and non-response, the National Health Interview Survey oversamples certain groups of the American population. As a result, survey weights–the sum of which equals the total population in that point in time–were attached to each year's data.

For the analysis to be statistically valid, these weights also must be modified when rearranging the data. In addition to joining data across files, it is also possible to join data cross years. This is done to increase the sample size available and increase the precision of our estimates. However, two issues arise when doing so. Some variable names and definitions may change, which can be rectified by reading the attached documention. Secondly, the weights will need to be recalculated. Per the documentation, multi-year analyses are conducted by dividing the weights by the number of years being analyzed.

Missing Data

A small but significant portion of the observations were removed due to missing values in either a predictor or response variable. As a result, only 158,803 out of 175,001 (90.74%) possible observations were used.

Model Building

Because of the large number of potential covariates being studied, a strategy was devised to build the final model. First, covariates were considered on a univariate basis. If they were part of a group of dummy variables, e.g. *Black*, *Asian*, etc., then the dummies were regressed in as a group. Possible interaction effects were then attached to the univariate regressions. Plots of predictors over time were used to assess if there were variations over time. Finally, multiple logistic regression models were constructed by including predictors that were significant at the 80% level.

How Mental Distress Was Measured

This study assessed the prevalence of psychological distress in the American population through responses to three questions contained in the Sample Adult file:

1. How often a person felt worthless in the past 30 days
2. How often he or she felt hopeless in the past 30 days
3. On average, how much sleep a person gets in a 24-hour period

These attributes were chosen since they are useful on their own for evaluating mental distress in an individual. For example, one study found that feelings of hopelessness provided a link between severity of depression and suicide, and that degree of hopelessness predicted suicide more strongly than severity of depression. (Kovacs and Garrison 1985) Furthermore, these are part of the criteria used for diagnosing major depression in the fifth edition of the *Diagonistic and Statistical Manual of Mental Disorders.* ("Depression Symptoms (Major Depressive Disorder)" 2017)

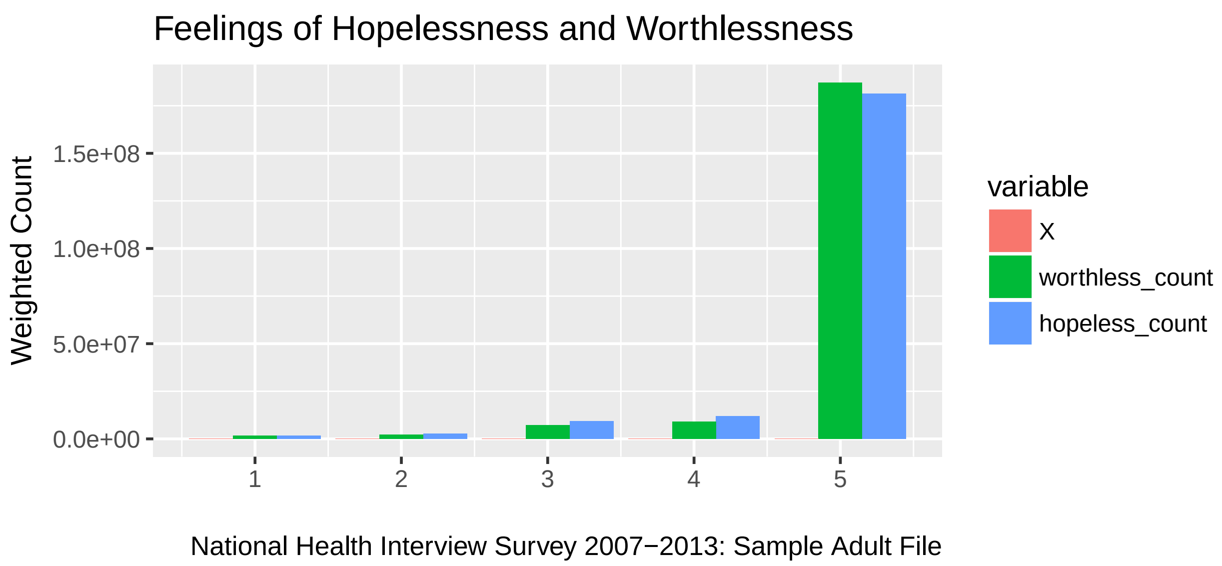
Feelings of Worthlessness and Hopelessness

The responses to the questions regarding worthlessness and hopelessness were coded on a 1-5 scale:

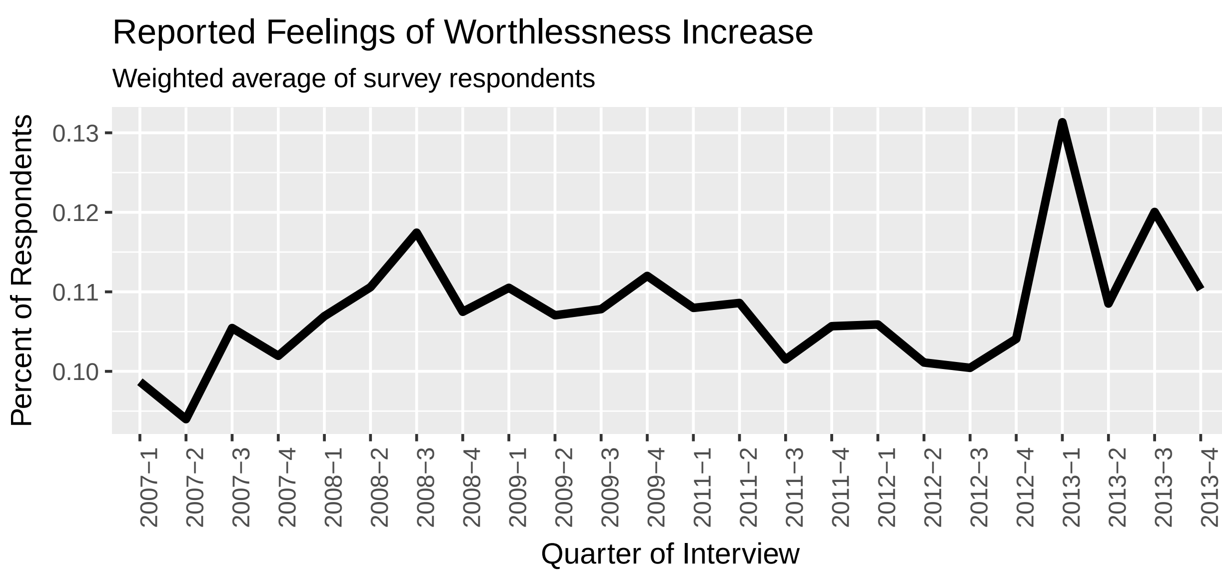
1. All of the time
2. Most of the time
3. Some of the time
4. A little of the time
5. None of the time

Because only a small proportion of respondents affirmed that they had feelings of worthlessness or hopelessness, a dichtonomous variable was created by separating responses into two groups–those that felt worthless or hopeless at least a little of the time (1-4) and those that never felt worthless or hopeless (5).

As a result of this coding change, I was reduced from being able to assess the severity of mental distress to merely being able to detect its presence. However, this change also simplified my problem as regressions could be estimated with binomial–rather than more complicated multinomial–logistic regressions. Furthermore, I still encountered difficulty building a reliable model for predicting binary outcomes. These difficulties would only be magnified if my model had to distinguish between levels of mental distress as well. Lastly, but most importantly, this coding change ameliorates the possibility that survey respondents would underreport their levels of mental distress to an unfamilar Census Bureau employee conducting the survey.



Reported feelings of worthlessness in the NHIS is on a positive trend. Whether this is a result of incresed mental distress or greater honesty about internal feelings is undetermined.



Quantity of Sleep

A new feature used in the following regressions was constructed from the amount of sleep a person had–abnormal sleep–which was defined as having less than 6 or more than 10 hours of sleep.

Unweighted Summary Statistics for the NHIS 2007-2013 Sample Adult File

Time of Survey

Survey Year	Survey Quarter	n
2007	1	5196
2007	2	5881
2007	3	3042
2007	4	6138
2008	1	4974
2008	2	5729
2008	3	5734
2008	4	3039
2009	1	2776
2009	2	6588
2009	3	6507
2009	4	9669
2011	1	7466
2011	2	8138
2011	3	7571
2011	4	7382
2012	1	6985
2012	2	8512
2012	3	8080
2012	4	8260
2013	1	7482
2013	2	7519
2013	3	8239
2013	4	7896

Demographic Control Variables

Male	n
0	88210
1	70593

Age

Average	Median	Standard Deviation
47.61	46	17.96

Income

Family Income	n
$0 - $34,999	68435
$35,000 - $49,999	23282
$50,000 - $74,999	26318
$75,000 - $99,999	15861
$100,000 and over	24907

Race and Ethnicity

Table 5: Table continues below

	Hispanic, any race	White (Non-Hispanic)	Black (Non-Hispanic)
n	28357	94890	24647
Percent	0.1786	0.5975	0.1552

Asian (Non-Hispanic)	Other Race (Non-Hispanic)
9386	1523
0.0591	0.00959

Response Variables

Original Survey Variables

Table 7: Table continues below

worthless_mean	hopeless_mean	sleep_mean	worthless_sd	hopeless_sd
4.806	4.754	7.151	0.6434	0.7093

sleep_sd
1.414

Modified Variables

These are the frequencies (in terms of percentages) of the features I created from existing survey attributes.

worthless_once	hopeless_once	abnormal_sleep
0.105	0.1376	0.1038

Logistic Regression

Age

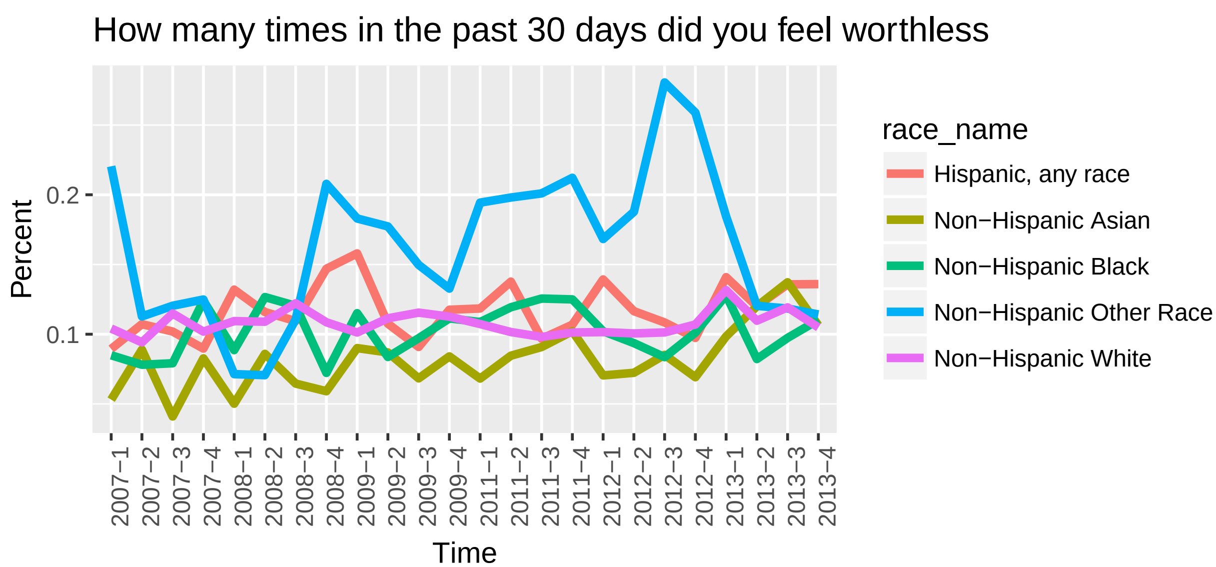
Age was not significant as a predictor of feelings of worthlessness. It was however, found to be significant at the 99% level to be negatively associated with feelings of hopelessness but positive associated with levels of "abnormal sleep". While the former finding suggests that older Americans are less likely to be depressed, the latter may simply be a result of old age rather than any underlying mental issues.

Table 10: Age and Symptoms of Depressions

	Dependent variable:		
	Felt Worthless	Felt Hopeless	Abnormal Sleep
	(1)	(2)	(3)
age	0.0002	-0.002^{***}	0.005^{***}
	(0.001)	(0.001)	(0.001)
Constant	-2.241^{***}	-1.829^{***}	-2.474^{***}
	(0.035)	(0.031)	(0.035)
Note:			*p<0.1; **p<0.05; ***p<0.01

Race

Contrary to my expections, there was no significant interaction effect between race and time throughout the recession. The estimates of feelings of worthlessness for "Other Race" varies erratically throughout time, although this may simply be an artifact of their relatively low sample size.



Sex

There was no noticeable interaction between sex and time in regards to symptoms of worthlessness. However, a higher proportion of women consistently reported feelings of worthlessness. This parallels the findings of the 2015 National Survey on Drug Use and Health which estimated 8.5% of women having a major depressive episode (5.5% experiencing a depressive episode without severe impairment) but only 3.0% of men having a major depressive episode (3.0% without severe impairment). ("Major Depression with Severe Impairment Among Adults," n.d.)

Table 11: Relationship Between Race and Symptoms of Depression

	Dependent variable:		
	Felt Worthless	Felt Hopeless	Abnormal Sleep
	(1)	(2)	(3)
Black	−0.041	0.038	0.542***
	(0.035)	(0.031)	(0.030)
Asian	−0.258***	−0.205***	−0.299***
	(0.057)	(0.051)	(0.059)
Hispanic	0.075**	0.318***	−0.033
	(0.032)	(0.027)	(0.033)
Other Race	0.404***	0.396***	0.626***
	(0.112)	(0.102)	(0.102)
Constant	−2.232***	−1.991***	−2.301***
	(0.015)	(0.014)	(0.016)
Note:			*p<0.1; **p<0.05; ***p<0.01

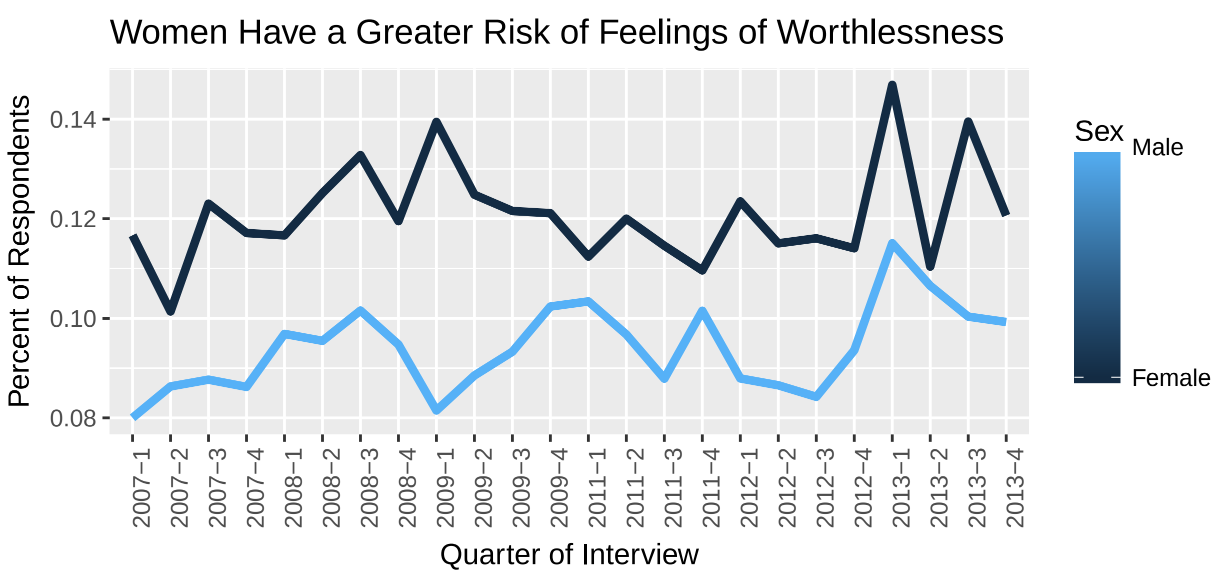


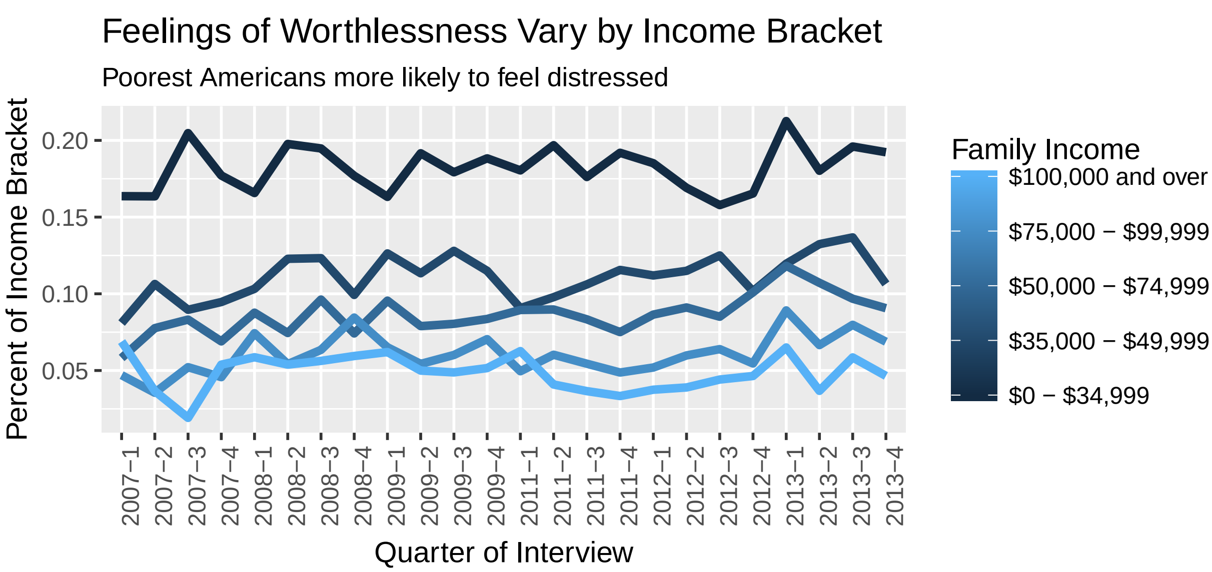
Table 12: Sex and Symptoms of Depression

	Dependent variable:		
	Felt Worthless	Felt Hopeless	Abnormal Sleep
	(1)	(2)	(3)
male	−0.261***	−0.294***	−0.088***
	(0.025)	(0.022)	(0.024)
Constant	−2.113***	−1.807***	−2.192***
	(0.016)	(0.014)	(0.016)
Note:			*p<0.1; **p<0.05; ***p<0.01

Family Income: Does Money Buy Happiness

Family income was once of the strongest predictors of feelings of worthlessness. While the link between happiness and family income may be unclear in the literature, the analyses below that additional income seems to guard against feelings of psychological distress. One explanation, suggested by the paper "Spending Money on Others Promotes Happiness", is that wealthier familers are more able and willing to spend money on others, thus leading to greater happiness. In a representative survey of 632 Americans, the authors analyzed the relationship between income, pro-social spending, and happiness. In one regression, they found the effect of extra income was just as powerful as the effect of spending more money on others. (Dunn, Aknin, and Norton 2008)

The authors then followed up with an experimental design where participants were paid either \$5 or \$20 and told to either spend this money on themselves or on others. Resultantly, those who were instructed to spend the money on others were found to have been significantly happier. However, when they asked the respondants whether they thought personal or prosocial spending was better for their happiness, the majority answered with personal spending. Although this is speculative, it is possible that richer households–having accomodated their basic needs and then some–may be more inclined to spend their wealth on others, thus leading to greater psychological wellbeing.



Employment Status

One of the questions I was interested in was the relationship between employment status and depression, especially if the person had a disability. In the most recent *Current Population Reports* published by the US Census Bureau, the occurence of disability was measured via the 2008 Survey of Income and Program Participation (SIPP), conducted from May 2010 through August 2010. (Kaye 2010) The study found that disabled workers were more likely to lose their jobs during the recession and less likely to be hired as the economy recovered.

Unfortunately, to contain the length of this paper, interactions between disability and unemployment were not measured. However, regressions against employment status were ran.

Looking for Work

Those actively looking for a job were found to have a slightly higher risk of symptoms of psychological distress in preliminary estimates. For example, when regressing against feelings of hopelessness, the coefficient on

Table 13: Association Between Income and Symptoms of Depression

	Dependent variable:		
	Felt Worthless	Felt Hopeless	Abnormal Sleep
	(1)	(2)	(3)
fam_income35_50k	-0.482^{***}	-0.475^{***}	-0.276^{***}
	(0.035)	(0.032)	(0.035)
fam_income50_75k	-0.744^{***}	-0.733^{***}	-0.444^{***}
	(0.037)	(0.033)	(0.034)
fam_income75_100k	-1.133^{***}	-1.061^{***}	-0.619^{***}
	(0.053)	(0.045)	(0.045)
fam_income101k	-1.316^{***}	-1.302^{***}	-0.856^{***}
	(0.046)	(0.040)	(0.041)
Constant	-1.709^{***}	-1.421^{***}	-1.893^{***}
	(0.015)	(0.014)	(0.016)
Note:			*p<0.1; **p<0.05; ***p<0.01
	When controlling for income, the lowest income bracket was the left out group		

looking for work was $\hat{\beta} = 0.82$. This translates to a $\frac{exp(-2+0.82)}{1+exp(-2+0.82)} = 23.5\%$ chance of feeling hopeless if looking for a job, but only 11.9% otherwise. Since, one study on alcohol use during the recession found that middle aged men were more likely to drink as a result of economic loss,(Mulia et al. 2014) I was motivated to see if there was an interaction effects between employment status, age and sex. However, the interaction coefficients were not statistically significant.

Table 14: Employment Status and Symptoms of Depression

	Dependent variable:			
	Felt Worthless		Felt Hopeless	Abnormal Sleep
	(1)	(2)	(3)	(4)
looking_for_work	0.777^{***}	0.714^{***}	0.821^{***}	0.180^{***}
	(0.042)	(0.127)	(0.038)	(0.048)
male		-0.270^{***}		
		(0.026)		
age		0.002^{**}		
		(0.001)		
looking_for_work:male		-0.009		
		(0.084)		
looking_for_work:age		0.003		
		(0.003)		
Constant	-2.289^{***}	-2.237^{***}	-2.000^{***}	-2.244^{***}
	(0.013)	(0.039)	(0.011)	(0.012)
Note:				*p<0.1; **p<0.05; ***p<0.01

Time Fixed Effects

Survey year dummies were included to account for time fixed effects. Quarter dummies were also added to account for any potential seasonality, but none of them were statistically significant.

Table 15: Time Fixed Effects and Seasonality

	Dependent variable:		
	Felt Worthless	Felt Hopeless	Abnormal Sleep
	(1)	(2)	(3)
factor(srvy_yr)2008	0.107**	0.152***	0.038
	(0.048)	(0.042)	(0.047)
factor(srvy_yr)2009	0.105**	0.138***	0.057
	(0.046)	(0.041)	(0.046)
factor(srvy_yr)2011	0.060	0.109***	0.123***
	(0.043)	(0.039)	(0.043)
factor(srvy_yr)2012	0.035	0.062	0.097**
	(0.044)	(0.040)	(0.043)
factor(srvy_yr)2013	0.165***	0.252***	0.137***
	(0.043)	(0.039)	(0.043)
factor(intv_qrt)2	−0.048	−0.063**	0.051
	(0.034)	(0.030)	(0.034)
factor(intv_qrt)3	−0.018	−0.010	0.038
	(0.035)	(0.031)	(0.036)
factor(intv_qrt)4	−0.051	−0.026	0.045
	(0.035)	(0.031)	(0.035)
Constant	−2.285***	−2.039***	−2.345***
	(0.040)	(0.036)	(0.040)
Note:			*p<0.1; **p<0.05; ***p<0.01

Language Spoken

I was interested in if there was a relationship between one's primary language and their risk factor for mental distress. Because English is the de facto language of the United States, those who cannot speak English may feel "left out." The language the NHIS interview was conducted in was used as a proxy for the individual's main language. Initially, no such association existed but there appears to be a spike every four years, occuring roughly during and right after the election. This may be a result of the extensive rhetoric and debate around elections, especially concerning immigration.

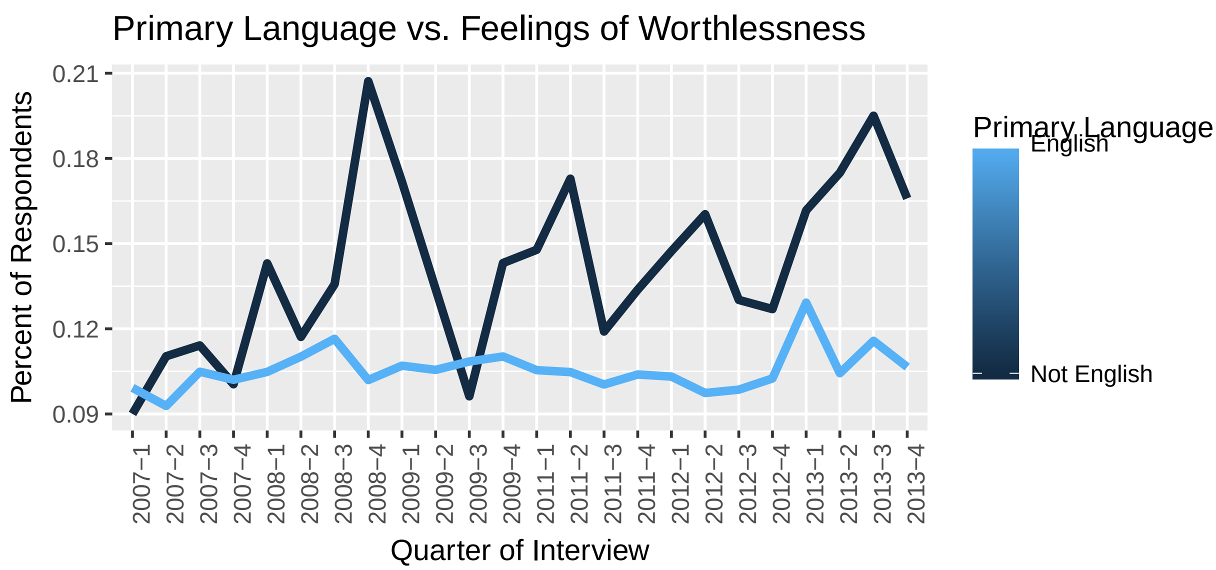


Table 16: Language Spoken and Symptoms of Depression

	Dependent variable:			
	Felt Worthless		Felt Hopeless	Abnormal Sleep
	(1)	(2)	(3)	(4)
english	-0.289^{***}	-0.019	-0.551^{***}	0.277^{***}
	(0.041)	(0.116)	(0.034)	(0.048)
factor(srvy_yr)2008		0.448^{***}		
		(0.160)		
factor(srvy_yr)2009		0.273^{*}		
		(0.149)		
factor(srvy_yr)2011		0.297^{**}		
		(0.136)		
factor(srvy_yr)2012		0.366^{***}		
		(0.142)		
factor(srvy_yr)2013		0.517^{***}		
		(0.138)		
english:factor(srvy_yr)2008		-0.365^{**}		
		(0.167)		
english:factor(srvy_yr)2009		-0.179		
		(0.157)		
english:factor(srvy_yr)2011		-0.254^{*}		
		(0.144)		
english:factor(srvy_yr)2012		-0.357^{**}		
		(0.149)		
english:factor(srvy_yr)2013		-0.379^{***}		
		(0.146)		
Constant	-1.963^{***}	-2.296^{***}	-1.430^{***}	-2.495^{***}
	(0.039)	(0.110)	(0.032)	(0.046)
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Marriage

With the talk of high divorce rates, many people may be hesitant to marry. However, a univariate regression reveals a strong negative association between being married and feelings of worthlessness. The coefficient on marriage of $\hat{\beta} = -0.51$ translates to a $\frac{exp(-1.99095)}{(1+exp(-1.99095)}$ $= 12\%$ chance of feeling worthless for unmarried individuals but only a $\frac{exp(-1.99095-0.50677)}{(1+exp(-1.99095-0.50677))} = 7.6\%$ chance of feeling worthless for married individuals–a difference of 4.4%. Based on feelings of worthlessness, compared to the regressions against family income, being married has roughly the same protective effect as jumping two income brackets. Similar results were found for the other two response variables. In the final model, after controlling for other demographic variables, the protective impact of marriage was diminished but still strong and statistically significant.

I was also curious to see if any sex benefitted more from marriage than the other, at least in terms of mental health. However, in the regression against feelings of worthlessnesss, the interaction term was trivial and not statistically significant.

Table 17: Relationship Between Marriage and Symptoms of Depression

	Dependent variable:			
	Felt Worthless		Felt Hopeless	Abnormal Sleep
	(1)	(2)	(3)	(4)
married	-0.507^{***}	-0.493^{***}	-0.530^{***}	-0.371^{***}
	(0.025)	(0.033)	(0.022)	(0.024)
male		-0.236^{***}		
		(0.032)		
married:male		-0.008		
		(0.051)		
Constant	-1.991^{***}	-1.887^{***}	-1.687^{***}	-2.051^{***}
	(0.015)	(0.019)	(0.014)	(0.015)
Note:				*p<0.1; **p<0.05; ***p<0.01

Neck and Back Pain

One study of chronic back pain in the Canadian Community Health Survey–with a sample size of over 100,000–found that those with chronic bank pain were at high risk for developing major depression. (Currie and Wang 2004) Furthermore, they also tested for the sigificance of the interaction between back pain, age, and gender–with only the gender interaction being statistically significant.

Unlike the aforementioned study, severity of neck or back pain could not be considered as–in the National Health Interview Survey–they were simply coded as "Yes" or "No" if the person being interviewed suffered either in the last three months. However, the regressions below obtained similar results, with the coefficients obtained being the greatest in magnitude in this study (aside from those on income). For example, in the single regression of back pain against feelings of worthlessness, the fitted coefficient was $\hat{\beta} = 1.09$ with an intercept of -2.64. This translates to a probability of 17.3% of feeling worthless for those who had back pain compared to a just 6.6% for those without.

Furthermore, I also found statistically significant mixed results between the interaction effects and the three response variables. For example, back pain was more likely to hinder sleep for older males. Likewise, males with back pain were also more likely to experience feelings of worthlessness than others afflicted.

Table 18: Association Between Neck and Back Pain and Symptoms of Depression

	Dependent variable:					
	Felt Worthless		Felt Hopeless	Abnormal Sleep		abnormal_sleep
	(1)	(2)	(3)	(4)	(5)	(6)
neck_pain	1.124***	1.124***	0.832***		1.124***	0.832***
	(0.027)	(0.027)	(0.028)		(0.027)	(0.028)
back_pain				1.087***		
				(0.025)		
Constant	−2.471***	−2.471***	−2.395***	−2.643***	−2.471***	−2.395***
	(0.015)	(0.015)	(0.014)	(0.017)	(0.015)	(0.014)
Observations	127,274	127,274	127,274	127,274	127,274	127,274
Log Likelihood	−38,584.490	−38,584.490	−39,059.990	−38,282.010	−38,584.490	−39,059.990
Akaike Inf. Crit.	77,172.980	77,172.980	78,123.990	76,568.010	77,172.980	78,123.990
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 19: Neck and Back Pain Interaction Effects

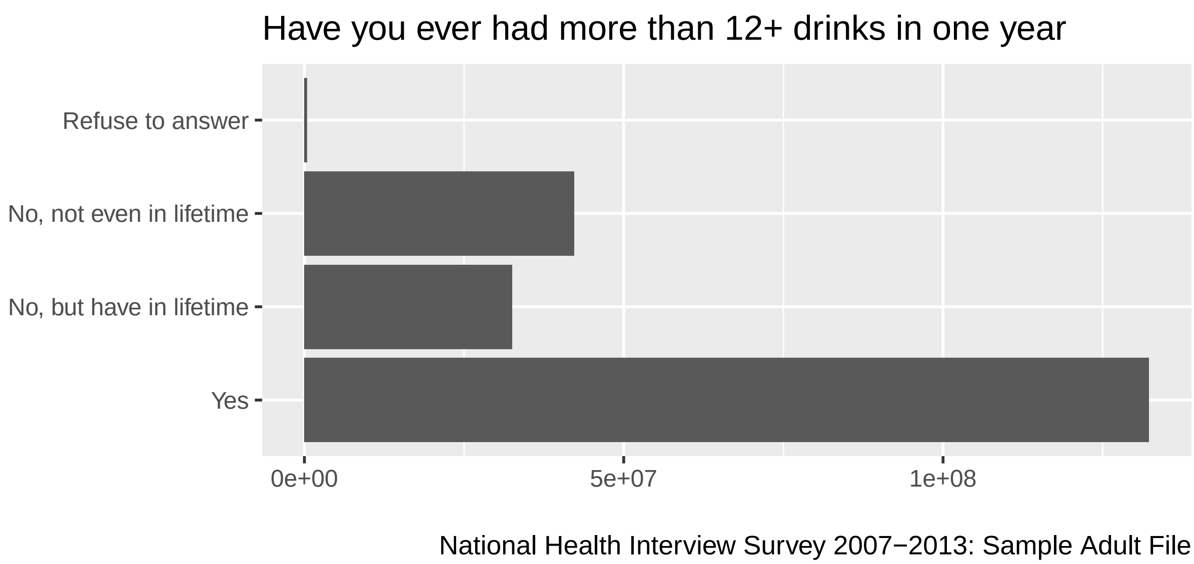
	Dependent variable:		
	Felt Worthless	Felt Hopeless	Abnormal Sleep
	(1)	(2)	(3)
neck_pain	0.726***	0.670***	0.621***
	(0.102)	(0.093)	(0.109)
back_pain	0.883***	0.876***	0.958***
	(0.085)	(0.076)	(0.083)
male	−0.160***	−0.229***	0.026
	(0.036)	(0.031)	(0.033)
age	−0.004***	−0.007***	0.005***
	(0.001)	(0.001)	(0.001)
neck_pain:back_pain	−0.093	−0.094*	0.058
	(0.063)	(0.057)	(0.066)
neck_pain:male	0.091	0.044	0.084
	(0.060)	(0.055)	(0.062)
back_pain:male	−0.127**	−0.042	−0.178***
	(0.055)	(0.049)	(0.054)
neck_pain:age	0.0003	0.002	−0.004**
	(0.002)	(0.002)	(0.002)
back_pain:age	0.001	−0.0003	−0.005***
	(0.002)	(0.001)	(0.001)
Constant	−2.440***	−1.975***	−2.827***
	(0.054)	(0.044)	(0.049)
Observations	127,274	127,274	127,274
Log Likelihood	−37,761.520	−44,797.990	−38,557.190
Akaike Inf. Crit.	75,543.040	89,615.990	77,134.380
Note:			*p<0.1; **p<0.05; ***p<0.01

Alcohol Use

Alcohol use is commonly linked with depression (why). It is also doubly important to this paper since the time frame of our data includes the recession. Intuitively, we can imagine (or have seen) some turn to the bottle as a result of losing a job or even worse, a home.Economic loss and alcohol consumption and problems during the 2008 to 2009 US recession. In one paper, found evidence that severe economic losses were associated with increased alcohol consumption and its associated consequences. (Mulia et al. 2014)

Frequency and Prevalence of Drinking

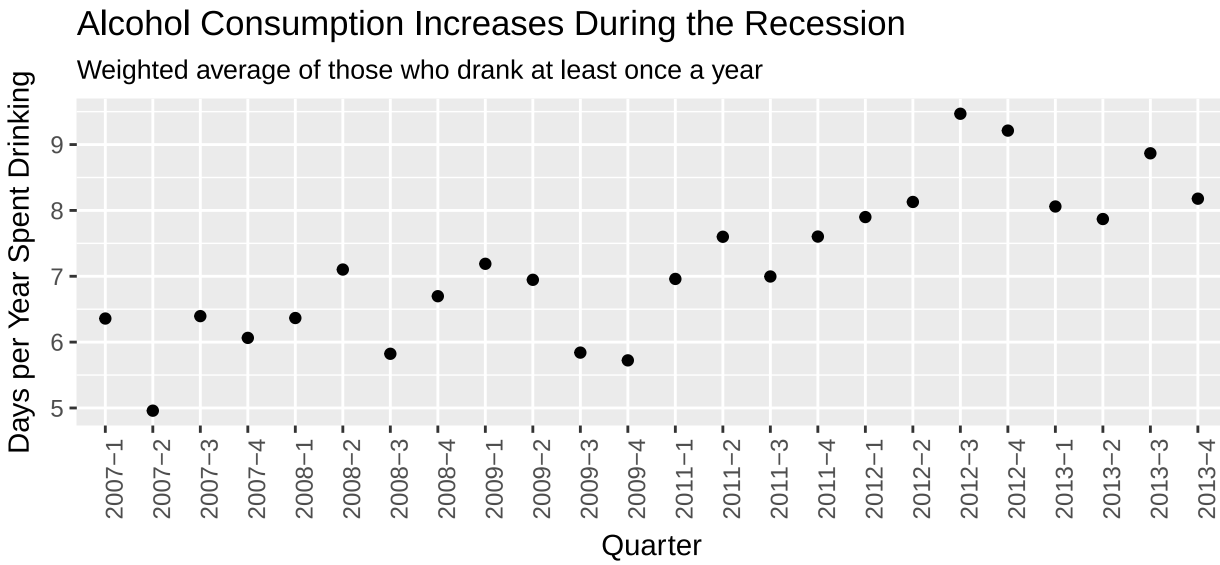
The vast majority of survey respondants resported that they have had at least 12 drinks in a single year.

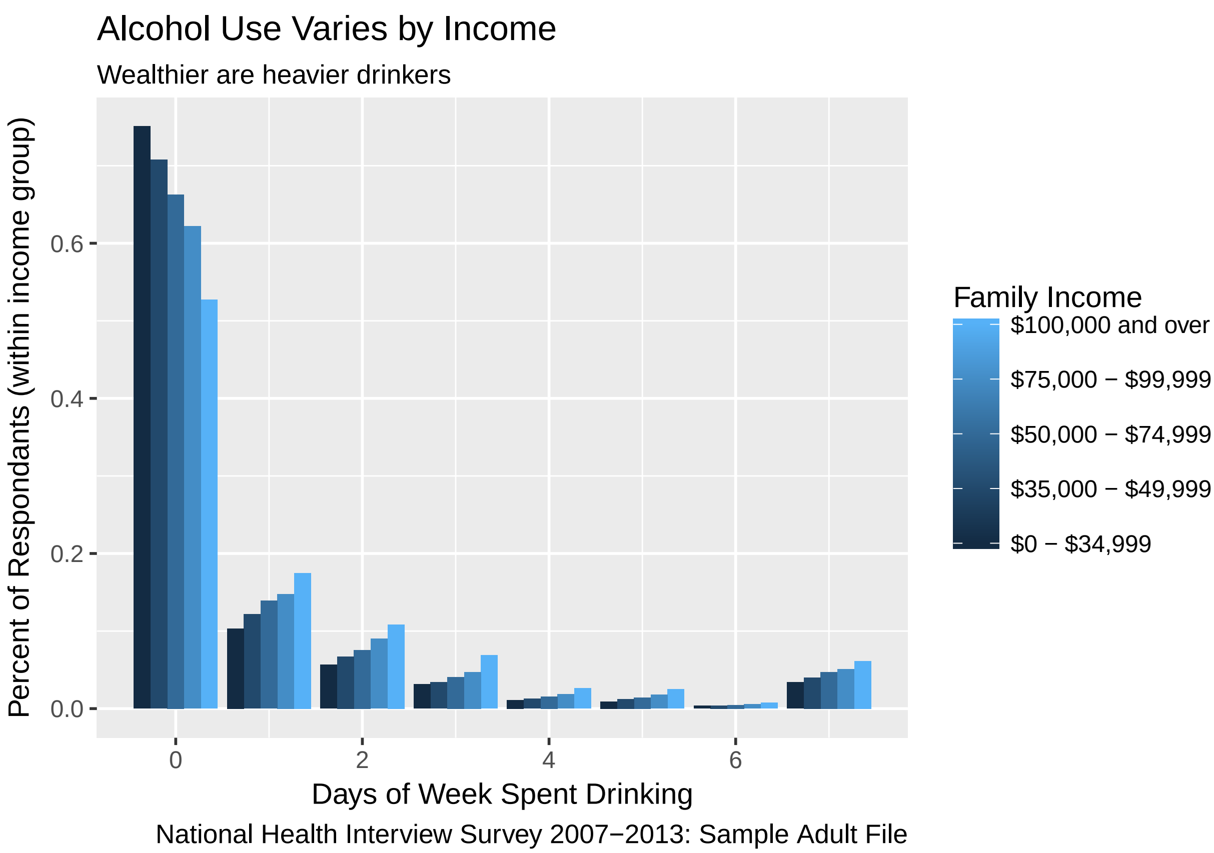


However, the distribution of drinking is skewed. with most respondants report drinking less than once a week or not at all.

Drinking and the Recession

On average, Americans drank more during the recession, but the plot below also suggests that this may be driven by just a small fraction of people.



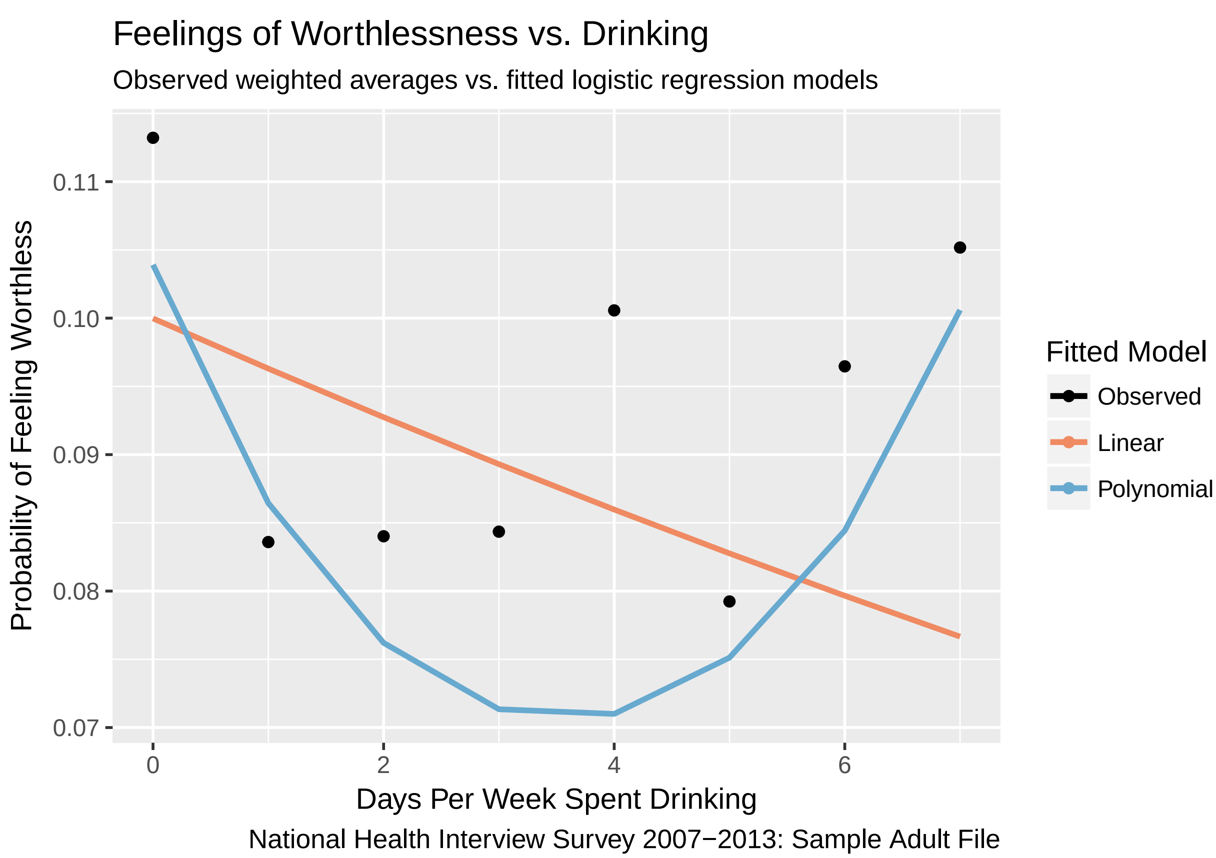


A Surprising Finding

Surprisingly, the group at highest risk for mental distress were those who either did not drink at all or drank–on average–less than once a week. The heaviest drinkers (those who drank all days of the week) were the seccond highest risk group. It was found that a polynomial relationship between alcohol consumption (days per week spent drinking) and mental distress was more accurate than a linear one. While the figure above suggests that this may simply due to poorer Americans not being able to afford alcohol, the regression in the final model still found a significant curvilinear relationship after controlling for demographic variables (although the magnitude of the coefficients was reduced).

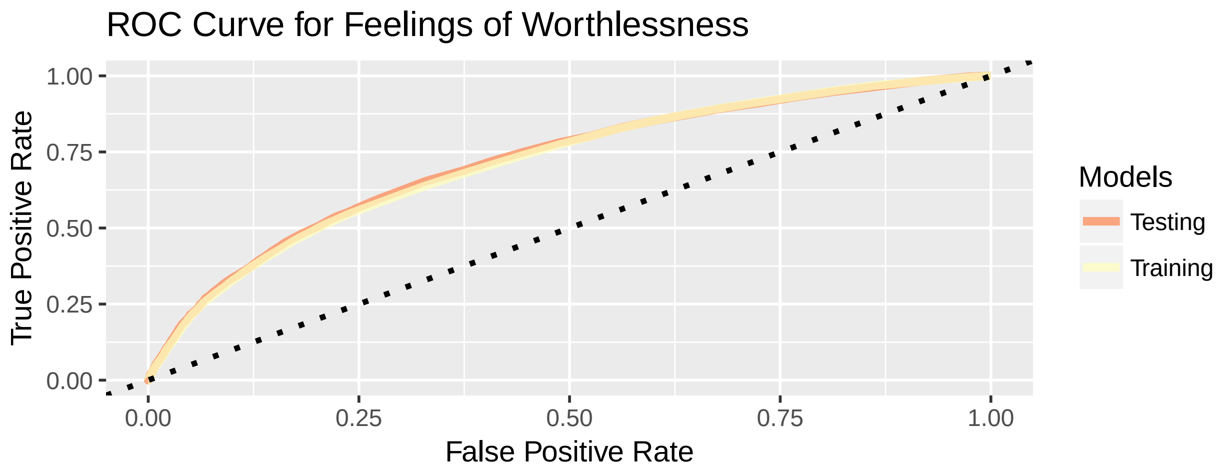
Table 20: Effect of Alcohol on Symptoms of Depression

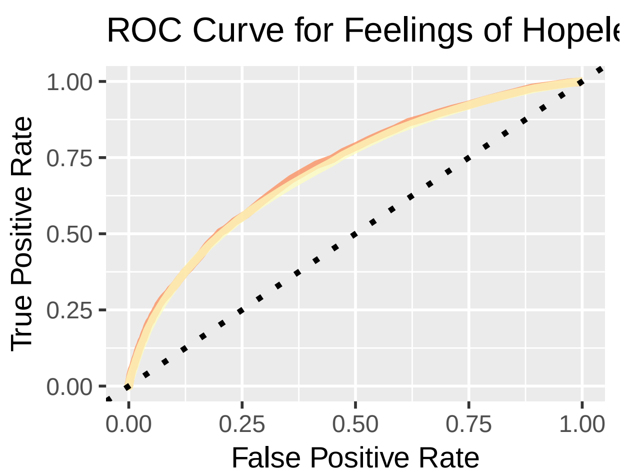
	Dependent variable:					
	Felt Worthless		Felt Hopeless		Abnormal Amount of Sleep	
	(1)	(2)	(3)	(4)	(5)	(6)
Days/Week Drinking	−0.046***	−0.260***	−0.056***	−0.225***	−0.071***	−0.275***
	(0.008)	(0.023)	(0.007)	(0.020)	(0.008)	(0.024)
$(\text{Days/Week Drinking})^2$		0.036***		0.029***		0.035***
		(0.004)		(0.003)		(0.004)
Constant	−2.185***	−2.139***	−1.885***	−1.848***	−2.183***	−2.140***
	(0.014)	(0.014)	(0.012)	(0.012)	(0.013)	(0.014)
Observations	127,509	127,509	127,509	127,509	127,509	127,509
Log Likelihood	−40,058.720	−39,967.950	−47,479.880	−47,410.740	−39,408.460	−39,335.030
Akaike Inf. Crit.	80,121.440	79,941.900	94,963.760	94,827.470	78,820.920	78,676.050
Note:						*p<0.1; **p<0.05; ***p<0.01

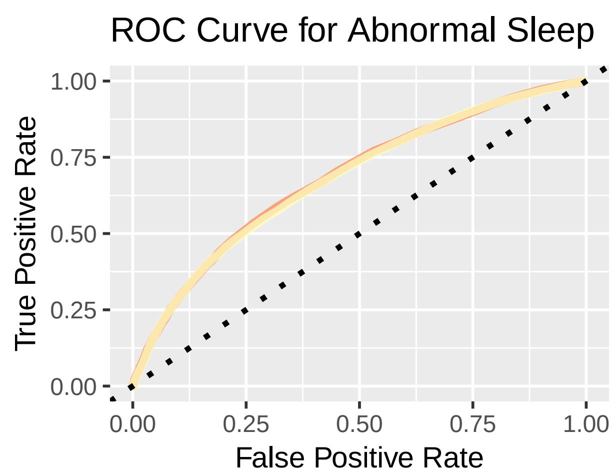


Logistic Regression: Full Model

ROC curves are shown below for the models predicting feelings of worthlessness and abnormal sleep. The closeness of the curves for the testing and training sets suggest that our model will generalize well to outside data.







Decision Trees and Random Forests

Alongside logistic regression, I also decided to fit decision trees to my data. While the logic of decision trees is different from that of logistic regression, they have nice properties which make them an attractive model for detecting symptoms of depression. Like logistic regression, they are readily interpretable–a useful quality especially when it comes to explaining illnesses. Furthermore, the branching structure of decision trees at first glance appears to be a decent approximation to how mental illness operates. Each node gives new probabilities of having a symptom of mental distress as a function of that node and its parent node.

Random forests are an extension of decision trees. Because decision trees can vary randomly in what nodes appear, random forests attempt to reduce this variance by training a large number of random forests and then choosing an answer based on a majority vote.

Table 21: Full Model: Socioeconomic Factors

	Dependent variable:		
	worthless_once	hopeless_once	abnormal_sleep
	(1)	(2)	(3)
age	−0.0005	−0.004***	0.009***
	(0.001)	(0.001)	(0.001)
black	−0.304***	−0.198***	0.430***
	(0.038)	(0.033)	(0.033)
fam_income35_50k	−0.410***	−0.362***	−0.201***
	(0.037)	(0.033)	(0.036)
fam_income50_75k	−0.634***	−0.556***	−0.328***
	(0.039)	(0.035)	(0.037)
fam_income75_100k	−0.983***	−0.822***	−0.453***
	(0.056)	(0.048)	(0.048)
fam_income101k	−1.136***	−1.010***	−0.651***
	(0.051)	(0.044)	(0.046)
looking_for_work	0.596***	0.640***	0.001
	(0.046)	(0.042)	(0.051)
male	−0.102***	−0.196***	0.126***
	(0.036)	(0.031)	(0.033)
factor(srvy_yr)2008	0.419**	0.335**	0.266
	(0.164)	(0.133)	(0.194)
factor(srvy_yr)2009	0.168	0.320**	0.367**
	(0.153)	(0.126)	(0.182)
factor(srvy_yr)2011	0.162	0.329***	0.402**
	(0.140)	(0.115)	(0.158)
factor(srvy_yr)2012	0.263*	0.236**	0.282*
	(0.146)	(0.120)	(0.166)
factor(srvy_yr)2013	0.483***	0.456***	0.350**
	(0.144)	(0.119)	(0.162)
family_size	0.086***		0.090***
	(0.012)		(0.010)
english:factor(srvy_yr)2008	−0.335*	−0.195	−0.239
	(0.172)	(0.141)	(0.200)
english:factor(srvy_yr)2009	−0.115	−0.240*	−0.346*
	(0.161)	(0.134)	(0.188)
english:factor(srvy_yr)2011	−0.190	−0.312**	−0.328**
	(0.148)	(0.123)	(0.165)
english:factor(srvy_yr)2012	−0.292*	−0.235*	−0.208
	(0.154)	(0.128)	(0.172)
english:factor(srvy_yr)2013	−0.367**	−0.235*	−0.240
	(0.151)	(0.127)	(0.168)
male:back_pain	−0.054	0.025	−0.118**
	(0.051)	(0.046)	(0.050)
age:back_pain			−0.007***
			(0.001)
Observations	127,274	127,274	127,274
Log Likelihood	−36,342.650	−43,091.520	−37,713.600
Akaike Inf. Crit.	72,745.300	86,241.040	75,489.200
Note:			*p<0.1; **p<0.05; ***p<0.01
			When controlling for race, whites were the left out group
			When controlling for income, the lowest income bracket was the left out group

Table 22: Full Model: Lifestyle and Medical Factors

	Dependent variable:		
	worthless_once	hopeless_once	abnormal_sleep
	(1)	(2)	(3)
neck_pain	0.693***	0.684***	0.498***
	(0.030)	(0.028)	(0.031)
back_pain	0.803***	0.763***	1.030***
	(0.035)	(0.031)	(0.074)
alc12mwk2	−0.099***	−0.100***	−0.168***
	(0.025)	(0.022)	(0.025)
I(alc12mwk2ˆ2)	0.017***	0.015***	0.021***
	(0.004)	(0.003)	(0.004)
male:back_pain	−0.054	0.025	−0.118**
	(0.051)	(0.046)	(0.050)
age:back_pain			−0.007***
			(0.001)
Observations	127,274	127,274	127,274
Log Likelihood	−36,342.650	−43,091.520	−37,713.600
Akaike Inf. Crit.	72,745.300	86,241.040	75,489.200
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 23: Full Model: English as a Primary Language

	Dependent variable:		
	worthless_once	hopeless_once	abnormal_sleep
	(1)	(2)	(3)
english	0.104	−0.106	0.650***
	(0.124)	(0.103)	(0.137)
english:factor(srvy_yr)2008	−0.335*	−0.195	−0.239
	(0.172)	(0.141)	(0.200)
english:factor(srvy_yr)2009	−0.115	−0.240*	−0.346*
	(0.161)	(0.134)	(0.188)
english:factor(srvy_yr)2011	−0.190	−0.312**	−0.328**
	(0.148)	(0.123)	(0.165)
english:factor(srvy_yr)2012	−0.292*	−0.235*	−0.208
	(0.154)	(0.128)	(0.172)
english:factor(srvy_yr)2013	−0.367**	−0.235*	−0.240
	(0.151)	(0.127)	(0.168)
Observations	127,274	127,274	127,274
Log Likelihood	−36,342.650	−43,091.520	−37,713.600
Akaike Inf. Crit.	72,745.300	86,241.040	75,489.200
Note:			*p<0.1; **p<0.05; ***p<0.01

Research Questions

In particular, I had some questions about how decisions would fit the data compared to logistic regression.

1. What is the effect of class imbanace? How can it be handled?
2. How sensitive are decision trees to sampling method?

Class Imbalance and Downsampling

For decision trees, I focused my research mainly on feelings of worthlessness–which only 10.5% of the sample reported having at any time. As a result, both decision trees and random forests simply classified every individual as not having any symptoms of depression. This was unsuprisingly as a “model” which blindly classifies every individual as being perfectly fine would have a 89.5% accuracy rate. Likewise, the observation that random forests were no different than individual decision trees is unsurprising. If a single decision tree is highly likely to classify every observation as “not depressed,” then so would a multitude of them.

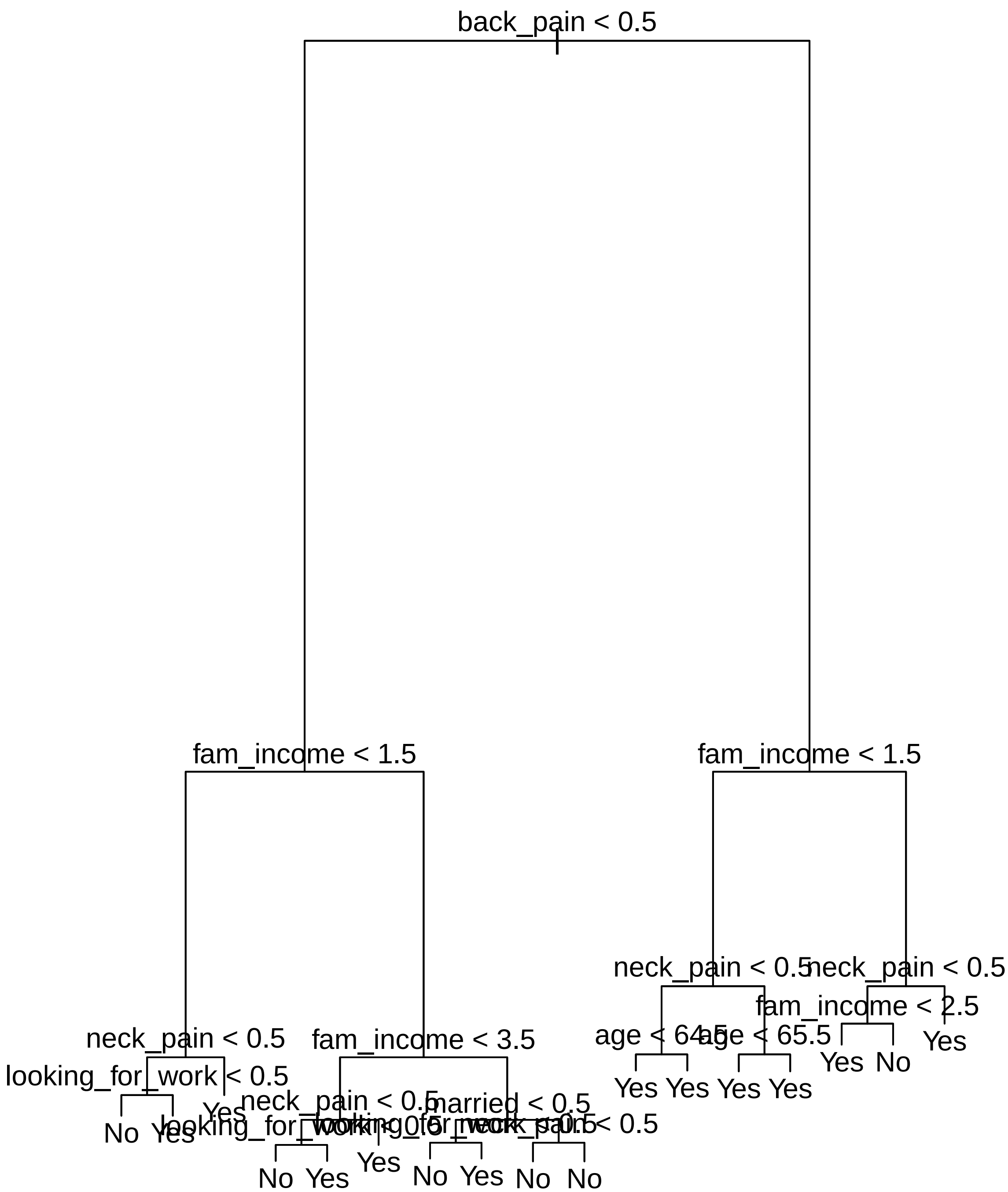
To deal with this issue, I performed downsampling on my original training set. Specifically, I created a new training set composed with all elements of the minority class (all individuals who reported feeling “worthless”) and randomly sampled an equally numerous set of observations from the majority class. As a result, in the new training set, there were equal numbers of both classes.

Decision Tree Training

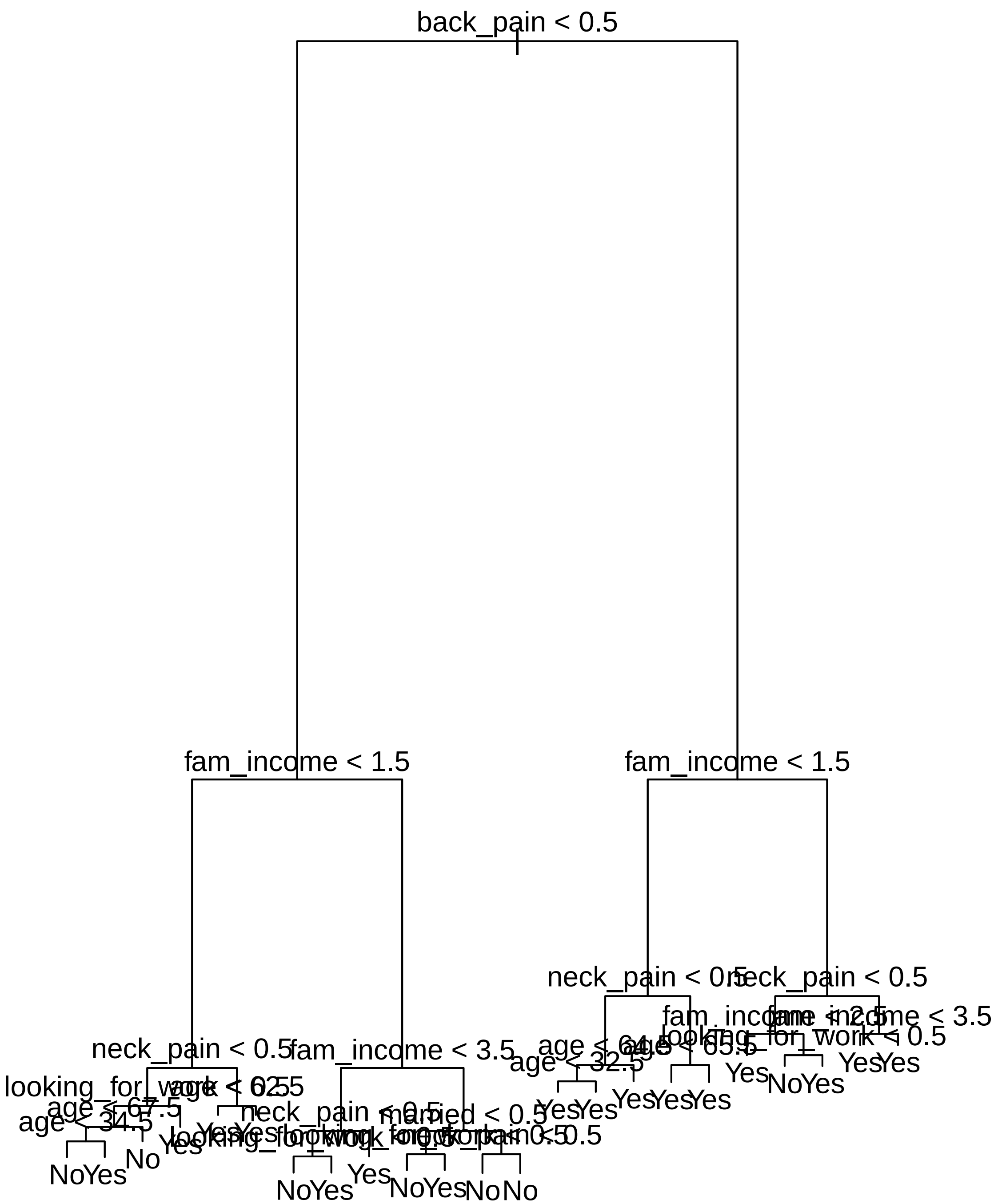
To answer my research questions, I trained four decision trees. Two pairs of these decision trees were trained with the same minimum deviance parameter: either 0.001 or 0.0005. With each pair, I trained either tree with one of two training sets. Both training sets were created using the same downsampling procedure, but with a differnet random seed. This was done to test the trees’ sensitivity to different training data.

As we can see, there is some deviation between the pairs of decision trees created by the different training sets. Lastly, I also trained a random forest with 100 trees using the same predictors as for the decision trees above. The results of the classification metrics comparing decision trees, random forests, and logistic regression can be found below.

Decision Tree 1: Training Set 1 – Mindev 0.001

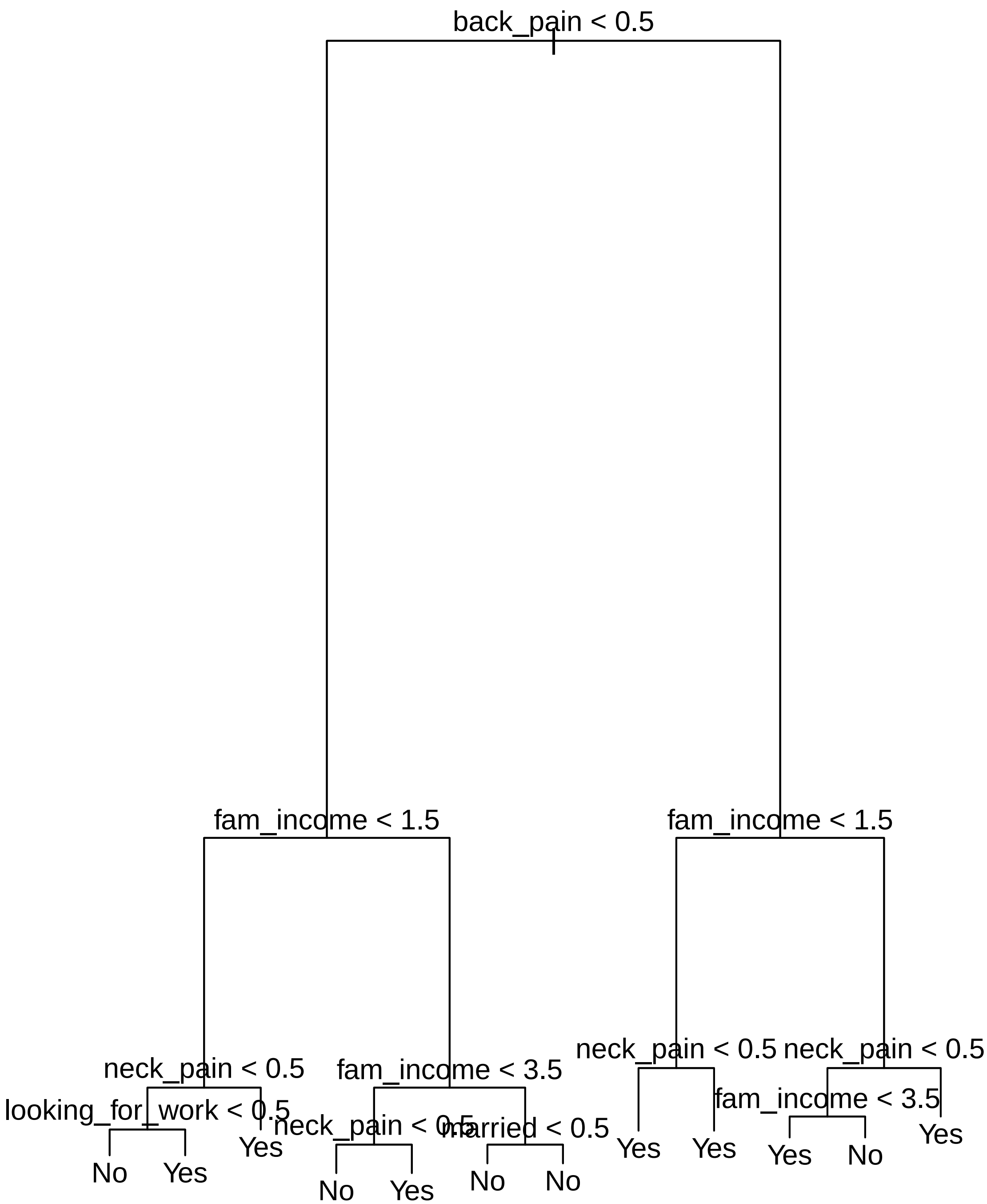


Decision Tree 2: Training Set 1 – Mindev 0.0005



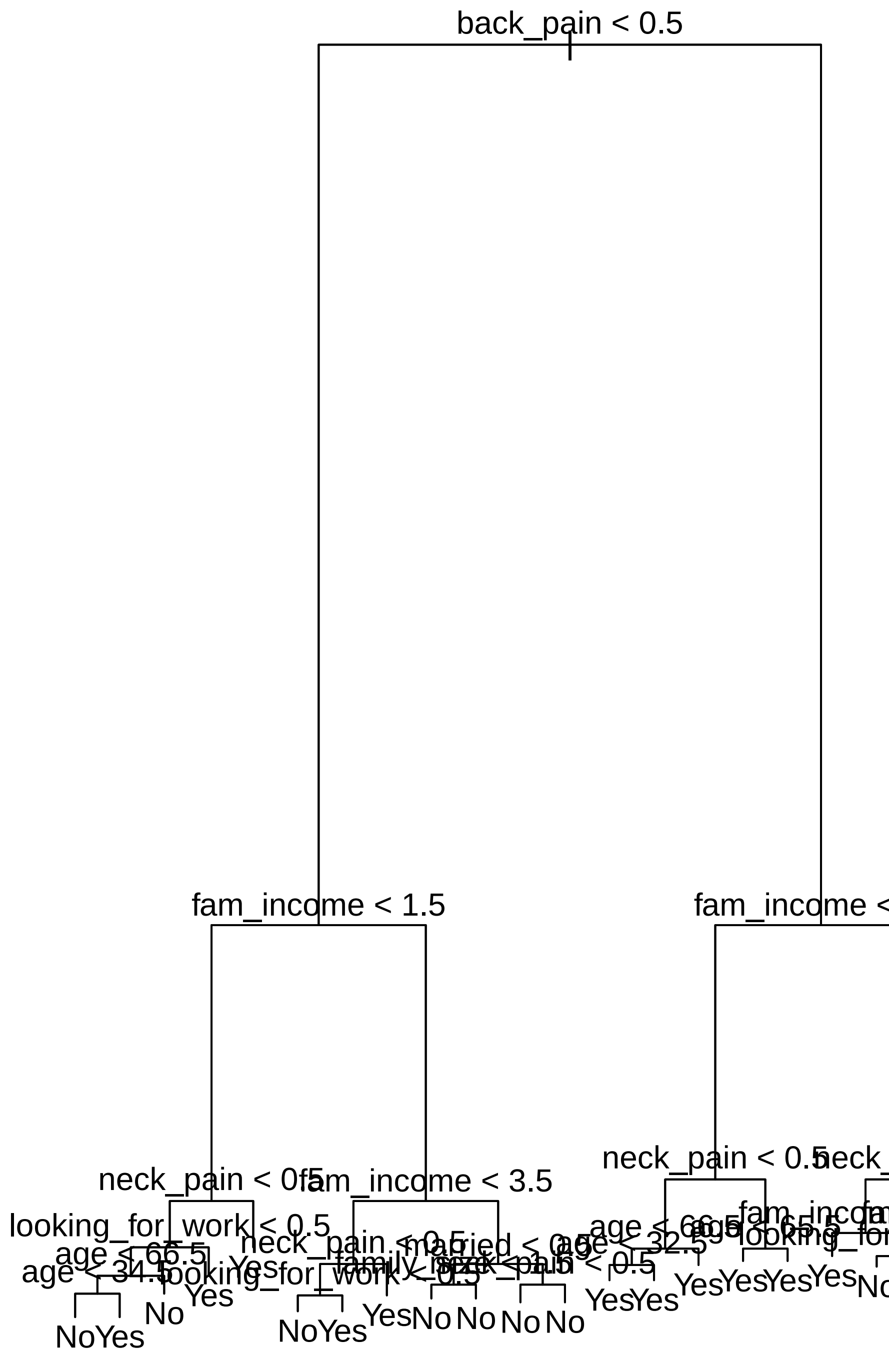
Decision Tree 3: Training Set 2 – Mindev 0.001

Although this tree was grown with the same parameters as Decision Tree 1, and its training set was sampled using the same sampling method, the seed used was different. The different configuration of the tree shows that decision trees are sensitive to sampling method.


back_pain < 0.5
fam_income < 1.5
fam_income < 1.5
neck_pain < 0.5
fam_income < 3.5
neck_pain < 0.5
neck_pain < 0.5
looking_for_work < 0.5
neck_pain < 0.5
married < 0.5
fam_income < 3.5
No
Yes
Yes
No
Yes
No
No
Yes
Yes
Yes
No
Yes

Decision Tree 4: Training Set 2 – Mindev 0.0005

Likewise, this decision tree also differs from Decision Tree 2–which was trained with the same minimum deviance parameter but a different training set.

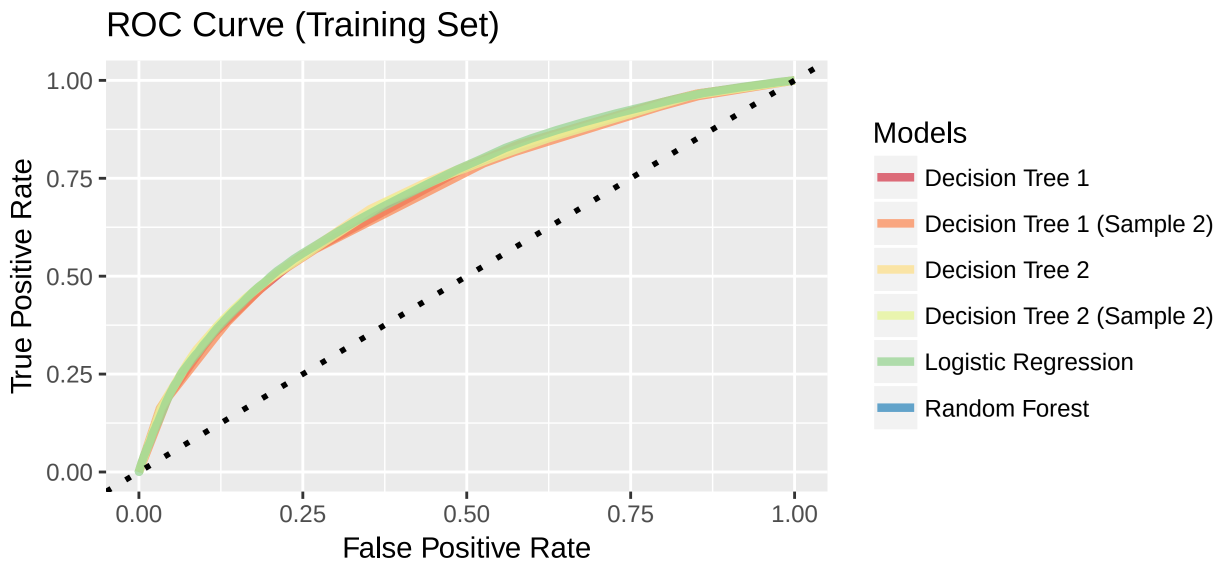


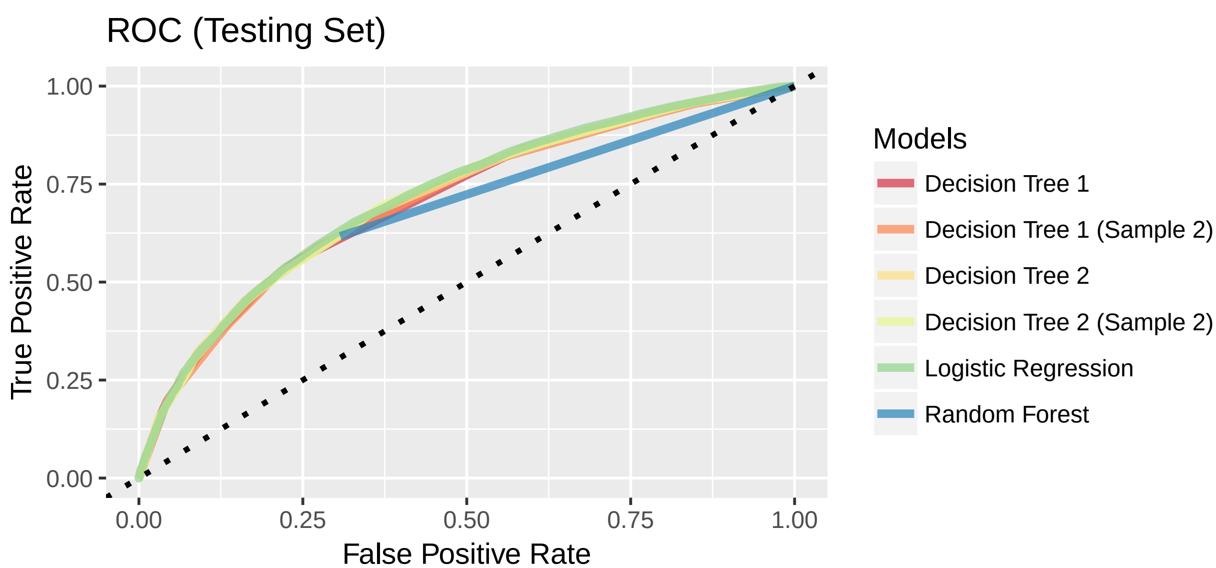
Logistic Regression vs. Trees and Forests: Measures of Accuracy

Finally, after training all of my models I used various classification metrics to compare them.

ROC Curves

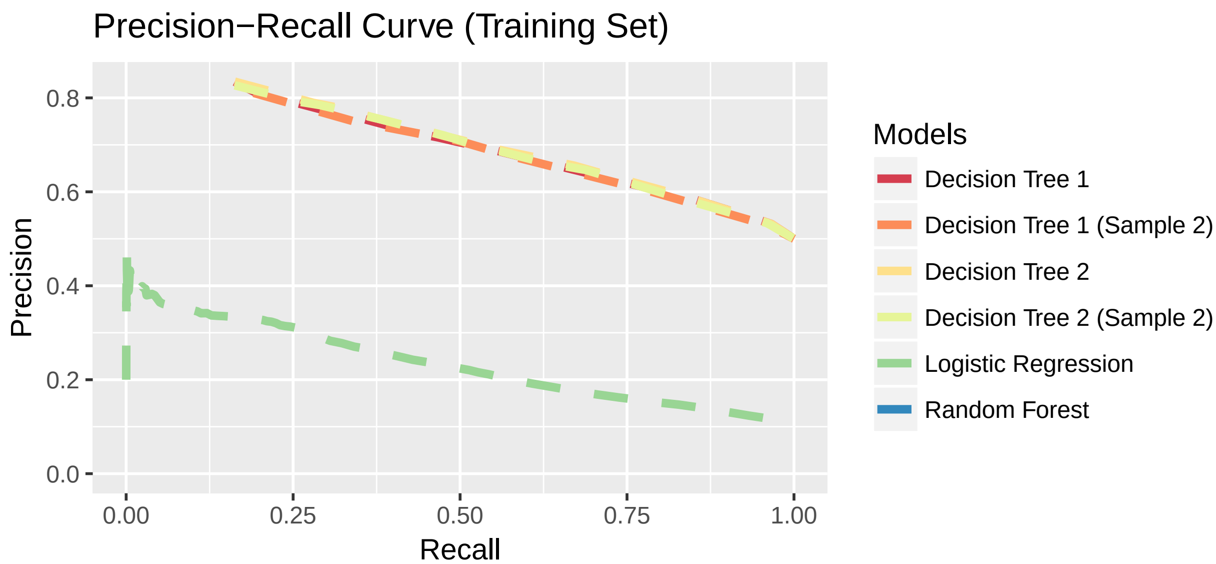
While the logistic regression was the best fitting curve in both the training and testing sets, it is hard to tell unless you look closely. The most noticeable feature of the ROC curve graphs is the drop-off of performance of the random forest in the validation set. It seems that a multitude of trees does worse than individual decision trees.

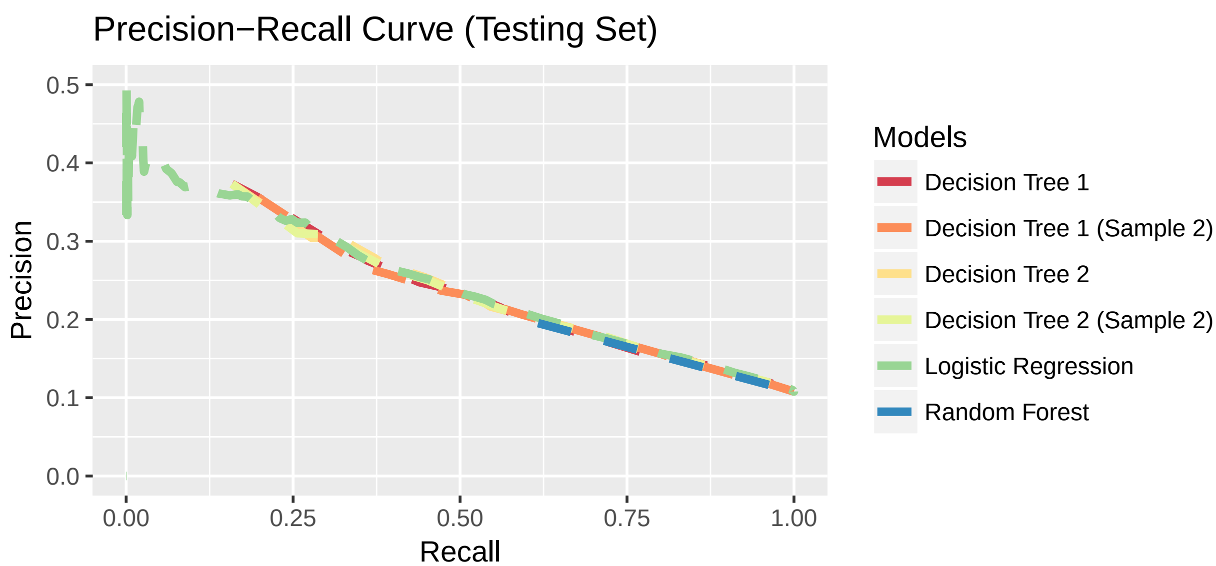




Precision-Recall Curve

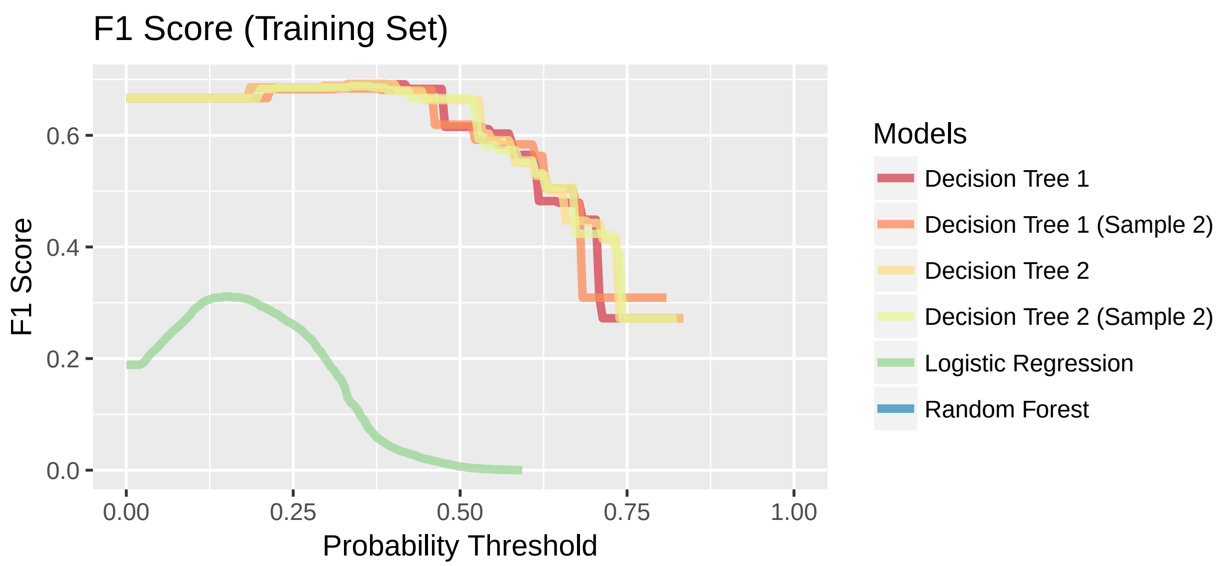
The precision-recall curve is another classification metric, which measures the trade-off between how accurate a model is on the data it retrieves (precision), and how many positive cases out of the data it retrieves overall (recall). Unlike the ROC graphs above, these pair of graphs show a stunning difference. While decision trees fit very well to the training data, their performance in the testing set is no better than the logistic regression. On the other hand, the logistic regression is the only model which has similar performance in both sets–suggesting that its coefficient estimates are very stable while decision trees overfit the data.

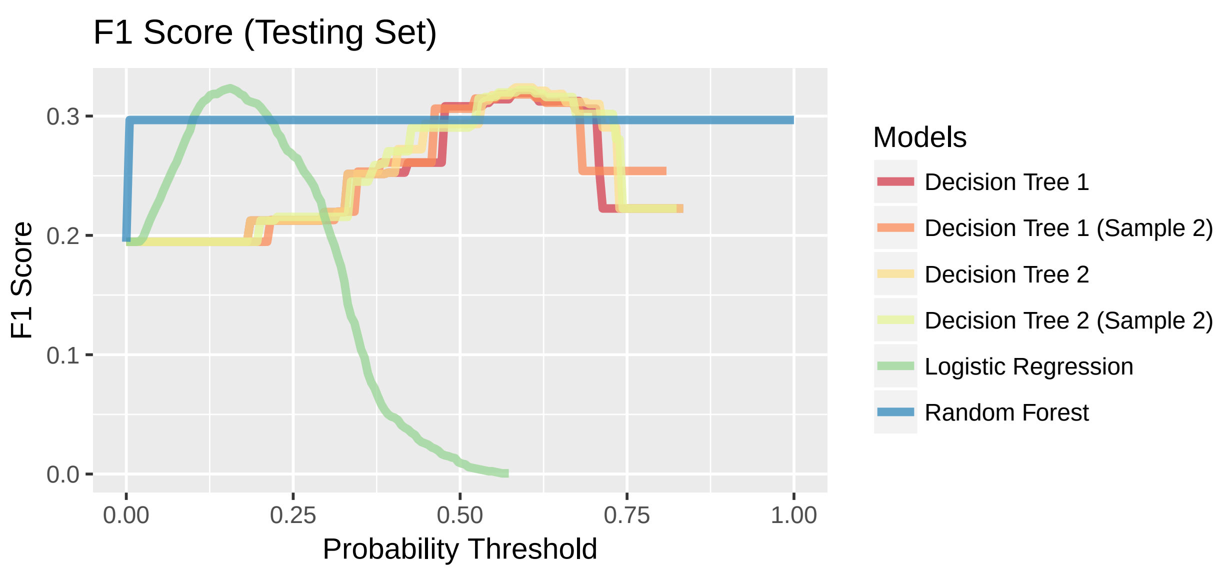




F1 Score

Lastly, I also considered the F1 score since it is a useful alternative to the ROC curve in the case of severe class imbalance. These graphs tell a similar story to the precision-recall curves. While the logistic regression initially seems worse when looking at the training set, it is the only model which has consistent performance in both the training and testing set.





Summary

While mental health is a difficult public health problem, this paper has shown that it is possible to build a decent predictive model with just a handful of easily attainable socioeconomic and lifestyle characteristics. For feelings of worthlessness, the probability threshold closest to a perfect classifier for a logistic regression model (TPR = 1, FPR = 0) gives a 58.4% true positive and a 35.1% false positive rate. Although this is far from perfect, this model is significantly better than guessing at random. Others may wish to build upon this model to achieve better rates, working with the predictors identified to build a model more suitable for medical use. However, this model may also be put to immediate use. For example, those trying to raise awareness about mental health issues, or trying to offer services, should target areas with high unemployment rates, high poverty rates, and a larger proportion of unmarried individuals.

Best Model: Logistic Regression

In conclusion, I choose logistic regression as the best model presented in this paper. Firstly, it simply performs better than the other models. While when looking at the testing set, it performances similarly to other models in terms of the F1 score and ROC curve, it has higher precision. Likewise, the logistic regression model is the only model which has similar behavior in both the training and testing sets. This suggests that it will behave similarly if applied to outside data, unlike the other models which had a sharp drop of performance (when looking at the F1 score or precision-recall curve) when applied to the testing data.

Furthermore, most of the papers cited by this project also utilized logistic regression models. There are several reasons why this is the case. First, while decision trees and logistic regression models are both readily interpretable, logistic regressions are much more concise. While a fitted logistic regression can be expressed as a single mathematical formula, a complicated decision tree can take up a page or more. Moreover, decision trees did not include some significant predictors identified by the logistic regression model–especially sex. As shown in earlier graphs, women were much more likely to develop symptoms of depression–a fact which is corroborated by outside research. Given the sample size of the data set, this is not a random fluke. Yet despite its importance, none of the decision trees included sex as a significant predictor, suggesting that they may not be a good method of identifying important risk factors.

Limitations

This problem this paper deals with is classification. As such, the purpose was to include was many covariates as possible in order to increase the accuracy of predictive models. This method allows us to analyze the relative importance of various socioeconomic, medical, and lifestyle factors as risk factors for mental distress. However, the main drawback is that the relationships between specific covariates and the response were not examined very throughly. While some nonlinear and interaction effects were considered, many more were not due to the scope of this paper. Many of the covariates in this model were included in linear form and interested researchers may wish to examine these relationships more thoroughly.

Furthermore, due to time constraints, some interesting attributed contained in the NHIS were not examined. For example, while this paper examined the relationship between being married and mental distress, it did not examine the relationship between divorce and mental distress. While the relationship between disability and depression was briefly examined through the inclusion of not being able to work due to disability as a predictor, others may wish the study the relationship between the nature and severity disability and depression. Lastly, sexual orientation was not studied in this paper because it was only recently introduced as an item in the 2013 National Health Interview Survey.

Works Cited

Corrigan, Patrick. 2004. "How Stigma Interferes with Mental Health Care." *American Psychologist* 59 (7). American Psychological Association: 614.

Currie, Shawn R, and JianLi Wang. 2004. "Chronic Back Pain and Major Depression in the General Canadian Population." *Pain* 107 (1). Elsevier: 54–60.

"Depression Symptoms (Major Depressive Disorder)." 2017. *Psych Central*. Psych Central. https://psychcentral.com/disorders/depression/depression-symptoms-major-depressive-disorder/.

Dunn, Elizabeth W, Lara B Aknin, and Michael I Norton. 2008. "Spending Money on Others Promotes Happiness." *Science* 319 (5870). American Association for the Advancement of Science: 1687–8.

Hlavac, Marek. 2015. "Stargazer: Well-Formatted Regression and Summary Statistics Tables. R Package Version 5.2." http://CRAN.R-project.org/package=stargazer.

Hoge, Charles W, Carl A Castro, Stephen C Messer, Dennis McGurk, Dave I Cotting, and Robert L Koffman. 2004. "Combat Duty in Iraq and Afghanistan, Mental Health Problems, and Barriers to Care." *New England Journal of Medicine* 351 (1). Mass Medical Soc: 13–22.

Hosmer Jr, David W, Stanley Lemeshow, and Rodney X Sturdivant. 2013. *Applied Logistic Regression*. Vol. 398. John Wiley & Sons.

Kaye, H. Stephen. 2010. "The Impact of the 2007-09 Recession on Workers with Disabilities." *Monthly Labor Review*. Bureau of Labor Statistics, U.S. Department of Labor, 19–30. http://www.jstor.org/stable/monthlylaborrev.2010.10.019.

Kovacs, Maria, and Betsy Garrison. 1985. "Hopelessness and Eventual Suicide: A 10-Year Prospective Study of Patients Hospitalized with Suicidal Ideation." *American Journal of Psychiatry* 1 (42). Am Psychiatric Assoc: 559–63.

"Major Depression with Severe Impairment Among Adults." n.d. *National Institute of Mental Health*. U.S. Department of Health and Human Services. https://www.nimh.nih.gov/health/statistics/prevalence/major-depression-with-severe-impairment-among-adults.shtml.

Mulia, Nina, Sarah E Zemore, Ryan Murphy, HuiGuo Liu, and Ralph Catalano. 2014. "Economic Loss and Alcohol Consumption and Problems During the 2008 to 2009 Us Recession." *Alcoholism: Clinical and Experimental Research* 38 (4). Wiley Online Library: 1026–34.

"National Health Interview Survey, 2011. Icpsr36145-V1." 2017. Ann Arbor, MI: United States Department of Health and Human Services. Centers for Disease Control and Prevention. National Center for Health Statistics. Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/ICPSR36145.v1.

"National Health Interview Survey, 2012. Icpsr 36146-V1." 2016. Ann Arbor, MI: United States Department of Health and Human Services. Centers for Disease Control and Prevention. National Center for Health Statistics. Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/ICPSR36146.v1.

"National Health Interview Survey, 2013. Icpsr 36147-V1." 2015. Ann Arbor, MI: United States Department of Health and Human Services. Centers for Disease Control and Prevention. National Center for Health Statistics. Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/ICPSR36147.v1.