Exploratory Regression Analysis on Adult Cognition

Ray Carpenter, Jason Jimenez

Dr. Gilleskie, Econ 573

November 2022

Abstract

The purpose of this study is to explore health, economic, and demographic predictors that may affect adult cognition. Using the University of Michigan's published Health and Retirement Study, we use regression to explore factors that might explain a total cognition score. We found many significant variables. Some variables found to have a positive average marginal effect on cognition were years of education, being female, earnings, total wealth, and walking time. A few factors that had negative effects were difficulty with different daily living activities and having diabetes. Our study didn't find significance in variables like sleep being restless, labor force status, region, smoking currently or before, and having heart disease. This study explores many factors that can give us better insight into understanding our cognition in adult life.

Introduction

As the United States faces a growing aging population with advances in healthcare and technology, it is essential to note which life factors play a role in affecting cognitive state. According to the U.S. Department of Health and Human Services, the population aged 65+ was 54.1 million in 2019, representing 16% of the U.S. population. That is expected to increase by 21.6% by 2040. Adulthood cognition is a significant indicator of general health as it can affect the elderly's ability to communicate, impact motor and physical functions, and affect income from labor force status.

The total cognition score is a measure of memory and reasoning performance. A low score could help indicate a mental decline, brain impairment, or other neurological health conditions. A high score does not write someone off from brain impairment but indicates a healthier state of mind. However, underlying brain functioning issues could still be present. The Center of Disease Control and Prevention reports Subjective Cognitive Decline (SCD) is a growing public health issue. The prevalence of SDC among adults aged 65 years and older is 11.7% or about 1 in 9 adults. Some of them may be unable to care for themselves or perform routine daily activities such as meal preparation, managing personal finances, or scheduling medical appointments. It is relevant to study cognition scores along with life variables because it can provide insight into preventative measures and increase awareness of predetermined susceptibilities older adults may experience.

Those with cognitive difficulties face challenges that can affect their well-being. Therefore, further research on the subject is relevant in order to educate aging adults to modify possible risk factors. Health professionals and researchers can work to mitigate future impacts of cognitive decline. Encouraging health-conscious decisions can lead to early intervention and assessment for treatment to improve quality of life.

Literature Review

A study conducted by Ramon et al. (2020) investigated predictors of cognitive decline using machine learning. Predictors included age, body mass index, gender, education, a specific gene type presence, cardiovascular disease, hypertension, diabetes, stroke, neighborhood socioeconomic status, and a certain risk gene of Alzheimer's dementia. Performance metrics of accuracy, sensitivity, and specificity were reported. Nonetheless, using the machine learning technique had limitations, such as the inability to account for time. The researchers derived cognitive trajectory groups and instead classified participants based on their cognitive trajectories over time. This altered the class-specific variances in intercept and slope, which might be necessary here. In the end, researchers found that nongenetic factors contribute more to cognitive decline than genetic factors.

Another study with researchers Levine et al. (2018) suggested cognitive decline differs by sex and race. The literature reviewed pre and post incident stroke cognitive trajectory. It showed that incident stroke alters a patient's cognitive outcome, and this effect is more significant with increasing age and cardioembolic stroke. Race, gender, geography, and hypertension status may modify the risk of poststroke cognitive decline. Part of their experiment focused on determining whether age is a predictor of cognitive decline. They were able to show that age is a predictor, as well as education and diabetes. One challenge with measuring what affects cognition is that older adults are more likely to have a neurodegenerative disease, atrophy, cerebrovascular disease, or vascular risk factors than younger adults have. These may amplify brain injury and cognitive deficits. This is among many articles that show how age and education affect one's cognitive state.

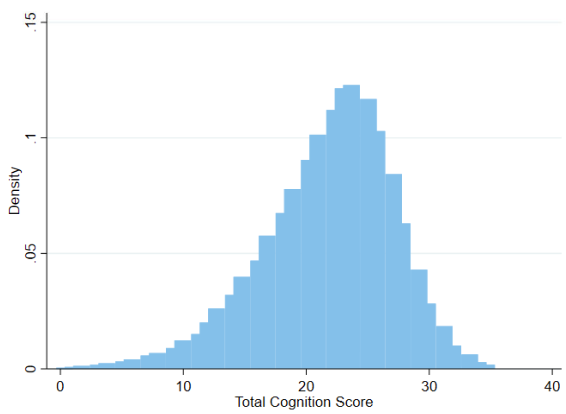
A recent paper by Smagula (2022) shows a relationship between activity level and cognition. Similar activity variables in the HRS longitudinal dataset are used in this paper, such as time walked and a breathing test. They show that different levels of activity will affect cognitive aging differently. Those who had daily physical activity showed better cognition tests than those who had less. The data indicates that activity pattern disruption may be ordinary in aging and may provide targets for tailored interventions. However, the study is only partially credible as future research is needed to investigate the biological processes of these behavioral phenotypes, including why earlier and robust activity patterns appear protective and whether modifying disrupted patterns improves outcomes.

Methodology

Data Overview

We used the data presented by the Health and Retirement Study (HRS), a longitudinal panel study that surveys a representative sample of approximately 20,000 people in America, supported by the National Institute on Aging and the Social Security Administration. The study, published by the University of Michigan, has followed respondents through in-depth interviews for 30 years. HRS began in 1992 when the first participants agreed to share their stories. It is one of the leading sources of information on the health and economic circumstances of adults over age 50 in the United States. The HRS collects information on physical and mental health outcomes, socioeconomic data such as an individual's income, and demographics. This dataset can be used to study many different aspects of adulthood, including one's cognition scores. (HRS Website)

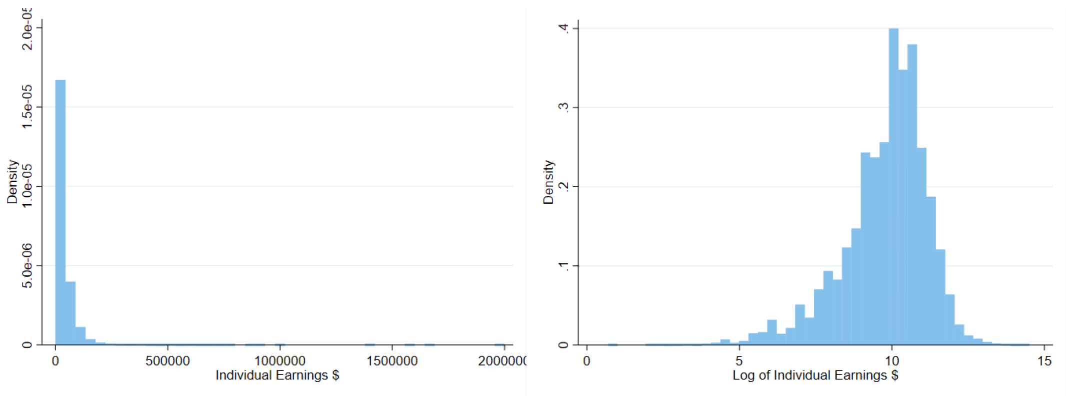
The variable of focus for this study is called the cognition score. Cognition score summarizes different tests to measure one's cognitive ability. These tests include immediate and delayed word recall, the serial 7s test, counting backward, naming tasks, and vocabulary questions. There were just over 130,000 observations of the cognition score kept for this study. It had a mean of 21.83 with a standard deviation of 5.28. It had a minimum score of 0 with a maximum of 35. A histogram of cognition can be seen below.



Data Cleaning Decisions

The raw dataset consisted of thousands of variables and rows. The file was 1.15 gigabytes. Due to the size of the data, we used UNC-Chapel Hill's Longleaf computing cluster. The parallel version of Stata was used for faster processing.

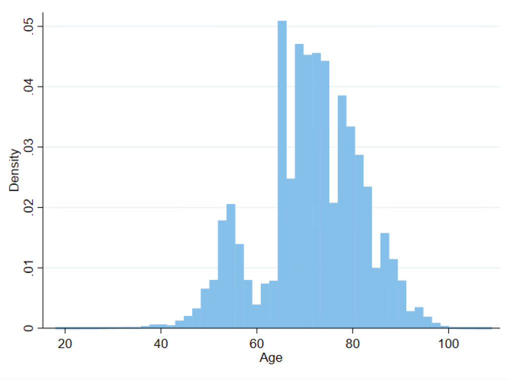
The full longitudinal dataset came in a wide format. The data had to be converted from wide to long. In this process, we had to select each variable we needed for the final models. We transformed over 500 unique variables from wide to long format. To narrow down on the number of predictors, we read through every unique variable and made a list of over 100 we thought might be pertinent to explaining one's cognition score. We also chose variables that were used in previous studies. This brought our list of predictors down to 150.



A cross-validated elastic net model was used to further slim down the number of predictors. This model is similar to a regression, except it drops predictors that are not statistically significant. After the elastic net model chose relative variables, we added variables back to our set that might have significance based on past studies. This process brought our number of unique predictors to 55. This was the predictor set that we would use to run our regressions.

The long dataset consisted of over 500,000 observations. To clean them, we removed all missing cognition score observations. This brought the dataset to 139,484 observations. Many variables contained missing values. We created two new variables for any variable that contained over 500 missing values. One variable was the same as the original, except the missing values were zero. The second variable is an indicator variable for each missing observation of the original variable. This prevented us from having to drop rows where the observation was missing. We also dropped every observation where the household wealth total was less than zero. The log variables of household total wealth, number of days in the hospital over the last two years, and earnings were created to be used. Next, we dropped all missing rows for variables that had less than 500 observations missing. This dropped around 8,000 observations. Finally, all rows of individuals with only one observation were dropped to allow for fixed and random effects. The final dataset consisted of 94,720 observations.

We decided to include all ages in our study. This was different than in our preliminary models. This is because a large number of observations in adulthood, 18 to 60, would be left out. We did run models that used only individuals from 58 to 100, and there was no meaningful change. The root mean squared error increased when we included all individuals. It also allowed us to gain insight into the average marginal effects of age on cognition scores from 18 to 108. If we were to use this dataset in the future, we would only look at individuals between the ages of 40 to 100. Another possibility is to run different regressions limited to each age group and explore how they differ.



Summary Statistics for the final dataset can be seen in tables 1 and 2 in the appendix. Note that many variables are categorical, and “(SR)” is self-reported.

Econometric Model

To explore cognition scores, we used different regressions to explore what factors may or may not affect cognition scores. We included predictors we believed were significant based on previous literature and preliminary elastic-net regressions. We searched for interactions and higher-order terms. We tried different models. Then we used the model with the highest root mean squared error (RMSE) for our final analysis.

To search for possible interactions in our model, we used a brute-force best subset selection method. The program glmulti in R was used here. This tries a different model for every combination of predictors and their interactions. After 2 million different models, the program recommended different interactions we would go on to test. Not all interactions were kept in our final model. One interaction of note that was found to be statistically insignificant was years of education with age. This was not included in our final model. Others include having high blood pressure and BMI and having high blood pressure with years of education. However, we only kept statistically significant interactions in our final model. The interactions that it found and we kept were between high blood pressure with age and body mass index with age. We included a quadratic of BMI and a 3rd-order polynomial of age.

The HRS is a panel dataset. The individuals included in our study all had more than one observation. We used fixed effects and random effects models to deal with the correlation between observations.

We ran different regression models to find which fit our data best. We used Root Mean Squared Error(RMSE) to measure which model best fits. We ran an ordinary least squares model to find a baseline RMSE to compare against. This produced an RMSE of 3.94. This is a rough estimate of how far off we predict cognition scores on average.

Next, we tried different family distributions using Generalized Linear Models(GLM). We employed the "xtgee" command to do this for panel data in Stata. This command takes the fixed effects of each model. This does leave out variables that are constant over time. This includes gender, race, Hispanic, and education. We first ran a gaussian GLM with an identity link with an RMSE of 3.97. Next, we used the same model except with a Poisson distribution, creating an RMSE of 19.47. We also tried a negative binomial distribution which had an RMSE of 19.47. The Gaussian distribution did fit best out of the GLMs so far. We used the gaussian distribution on all models going forward.

We wanted to try different links with our GLM models. The identity link had an RMSE of 3.97. The log link had an RMSE of 19.47. The power link produced 3.97 also. The negative binomial link had an RMSE of 22.5, while the reciprocal link yielded an RMSE of 22.41. At this point, none of these models outperformed OLS.

We used other models outside of GLM estimators also. The next one we tried was random effects. This would be our chosen model to investigate further. It had an RMSE of 2.95. We also tried quantile regression which produced an RMSE of 4.9. The Correia estimator was tried to allow for fixed effects that would allow for constant variables to be included, such as gender or race. This had an RMSE of 3.39. We also tried the Maximum Likelihood estimator with an RMSE of 3.91. We then tried fixed effects with OLS that produced the lowest RMSE of 2.52.

The fixed effects model with OLS produced the lowest RMSE. However, the R-squared for this model was 0.16 overall, while the R-squared for the random effects model was 0.43. The fixed effects model has to leave out constant factors over time to each individual and wouldn't allow for introspection into possibly important variables like race or gender. Therefore, we used the random effects model with OLS for further analysis of cognition score.

Results

We used a random effects model with Ordinary Least Squares for inference. It had an RMSE of 2.95. It had an overall R-squared of 0.43. The predictions for this model were off by an average cognition score of about 2.95. The between R-squared was 0.52. This means the model explains 52 percent of the variance between individuals. The within R-Squared was 0.27. We explained 27 percent of the variance within each individual observed over waves.

The positive significant marginal outcomes include female, reports lung disease, never married, years of education, has high blood pressure, log of household wealth. The negative significant marginal outcomes include age in years, has diabetes, has psychological conditions diagnosed, not in labor force, everything feels like effort, reports stroke, log of hospital stays, feeling depressed, retired, region: mountain, region: SE central, other (race), self-rate memory: fair, Hispanic, self-rate memory: poor, Black/African-American (race). The marginal effect of females was the strongest at 1.01, suggesting being a female impacts total cognition score by 1.01 points. This could be due to uncontrollable biological factors, such as different timelines in full brain development. Following suit is reporting lung disease at improving cognition score by 0.67 points. This could be due to a cautious diet and other life improvements that maintain healthy cognition. Next on the list is being never married, which has a marginal effect of 0.54 points. This could be explained by the possible avoidance of stress that comes from marriage. Another positive, significant marginal effect is an individual's years of education that improves cognition by 0.40 points. The phenomenon could be due to the higher use of mental power and strengthening of the mind through schooling tasks. Subsequently, high blood pressure has a marginal effect of 0.37 points on cognition, likely due to the well-fed diet someone with the disease will adapt to, which generally improves overall well-being. The log of household wealth holds a 0.17-point association with cognition giving the impression that comfortable assets relieve the individual of financial troubles and stress that may deteriorate their cognition. The change in self-reported mental health is correlated with a 0.12-point increase in cognition. It seems feasible since one's belief that their mental solidarity is improving can work as a pseudo effect to motivate further cognition use.

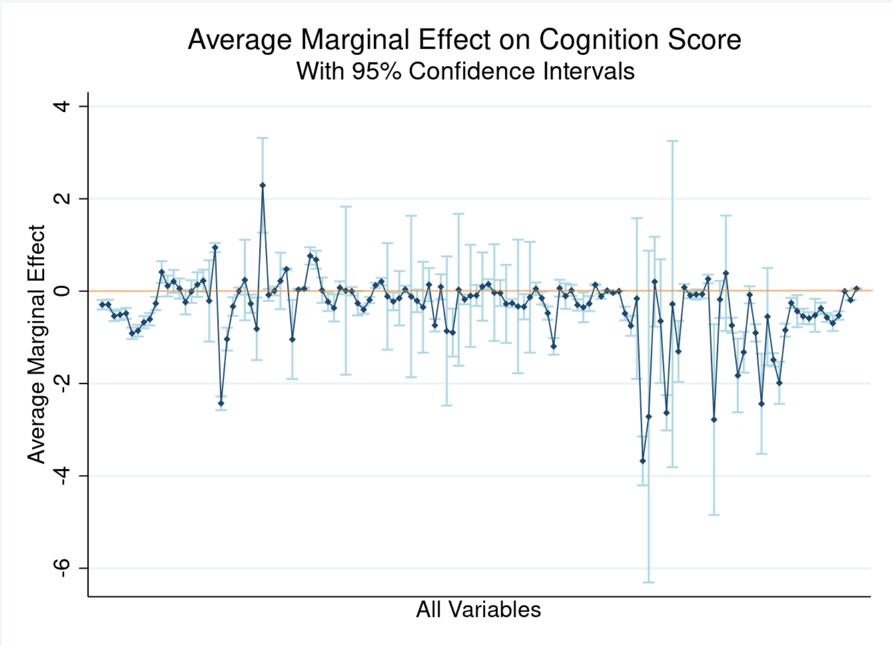
Additionally, age is shown to have a negative 0.14-point correlation with cognition, which is reasonable since cognition is expected to decline slowly with older age biologically. Those who reported having diabetes showed a negative 0.43-point relationship towards cognition, inferring that poor health conditions impact mental state, possibly due to the emotional and physical strain of living with a chronic disease. Similarly, respondents who reported having a condition diagnosed by a physician suffered a 0.16-point decrease in cognition score. Furthermore, the status of not being in the labor force resulted in a 0.46-point drop. Labor force participation can maintain cognition by practicing mental tasks and staying active. The variable everything feels like effort contained a 0.50-point decline, likely due to the poor mental state of the respondent, who lacked the mental energy to perform well during the cognition test. Those who reported a stroke struggled with 0.53 points less in cognition likely because respondents suffered from fully recovering from a dangerous health episode.

Moreover, the log of hospital stays returns a 0.56-point drop, and felt depressed experienced a 0.57 decline in cognition, which is consistent as both predictors place a mental toll on the respondent that positions them in a difficult situation to overcome. Retirees were associated with a 0.58-point decrease, most likely due to less engagement in mental exercises than workers. Also, geography plays a role in affecting cognition scores. For example, those living in the mountain and the southeast central region of the United States experienced a 0.58 and 0.61-point decrease, respectively. These regions often have most of their residents in rural areas that likely experience isolation and fewer social interactions that take a toll on cognition. Race and ethnicity are unmodifiable genetic factors that showed a 0.95-point drop in those races other than white or African/American, a 1.41 drop in those who reported a Hispanic ethnicity, and a 2.43 decrease in those who reported a Black/African American race. A mixture of genetic risk factors and socioeconomic factors could cause this. Lastly, the self-rated memory scores of fair were negative 0.97 and poor were negative 2.24. The results provide insight that self-rated memory scores play a significant role in how respondents perform in cognition tests.

The insignificant marginal outcomes were widowed (marital status), memory has worsened self report, memory has not changed self report, NW central region, reports heart problem, separated (marital status), NE central region, good self-rated memory, works part-time, divorced (marital status), has arthritis, South Atlantic region, number of drinks, SW central region, reports psych problems, sleep is restless, smoked ever, very good self-rated memory, partnered (marital status), Pacific region, Disabled, partly retired, unemployed, Mid-Atlantic (marital status), and interview begin date. These were not included in our final model.

The coefficient estimates of the random effects model can be found in the appendix under Best Model Part 1-4.

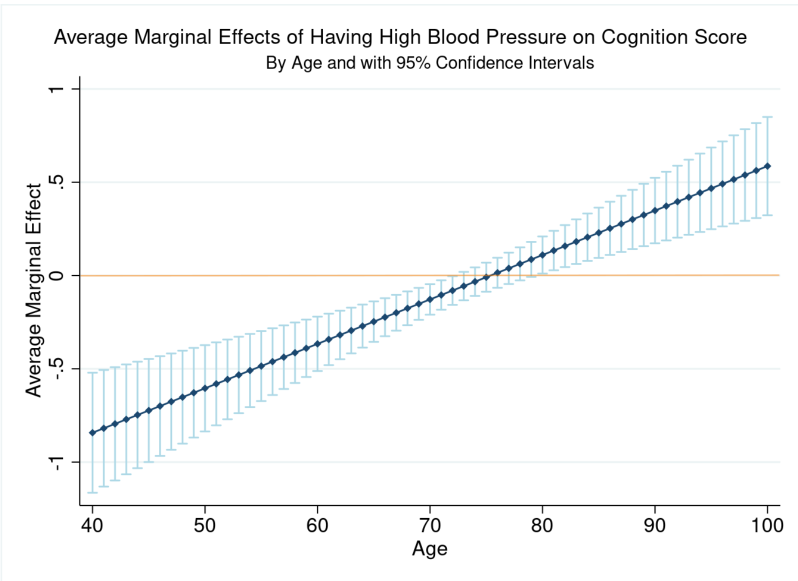
Discussion



Our best model explained over half of the between variation and a fifth of the within variation of cognition score. We identified variables that are and are not statistically significant. We also looked at higher-order terms and interactions between variables.

Many of the health questions about daily living activities were statistically significant. Another interesting find was that sleep being restless had a positive effect on cognition score. However, this is most likely a noisy zero. Sleep being restless was a self-reported question and may not truly measure the quality of one's sleep. Having arthritis, feeling sad, having cancer, memory getting worse, not getting going, marital status, and good memory are most likely noisy zeroes also. This is because our dataset consists of many individuals, which makes the standard errors smaller. This creates a higher chance of significance when there is none.

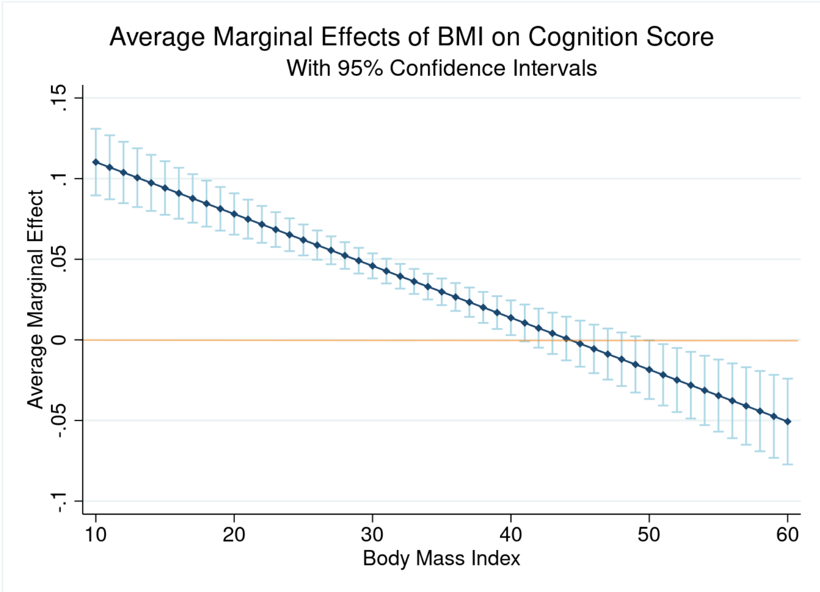
One of the more interesting findings was how having high blood pressure affects cognition score at different ages. According to a paper published by the American Heart Association, having high blood pressure is a risk factor for cognitive decline, which includes aspects such as memory, verbal fluency, attention, and concentration. A blood pressure elevation during middle and older age is linked to a faster decline in cognition. This can be seen in the graph below. This interaction was found using the glmulti program to test different interaction variables. This study supports the idea that high blood pressure negatively affects cognition scores from age 40 to about 72, decreasing negativity with age. It then suggests having high blood pressure positively affects cognition score as one ages, increasing with age. We initially thought that this variable might be picking up on how much nutrients one intakes. We do not account for diet in our study. Those with high blood pressure most likely intake enough nutrients for cognitive homeostasis. This is only our guess as to what explains this effect.



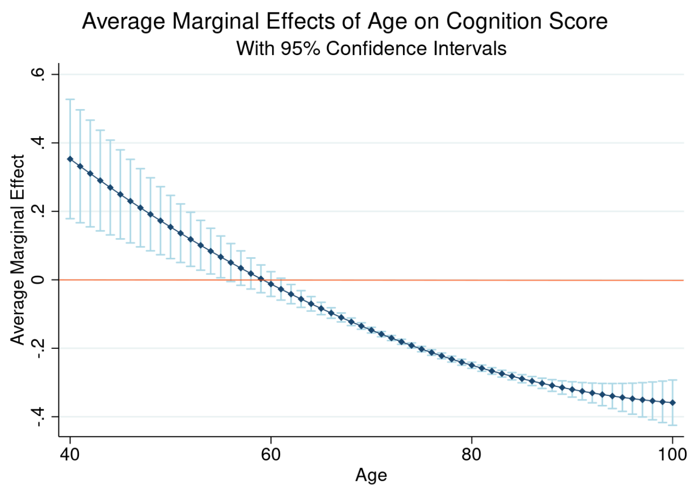
We also wanted to look at how BMI affects cognition over age.

According to the Center of Disease Control and Prevention, if your BMI is less than 18.5, it falls within the underweight range. If your BMI is 18.5 to 24.9, it falls within the normal or Healthy Weight range. If your BMI is 25.0 to 29.9, it falls within the overweight range. If your BMI is 30.0 to 39.9, it falls within the obese range. BMI is a weight and height function that varies by sex and ethnicity.

Our study suggests that the higher the BMI, the more negative it affects cognition score. Shifting from a positive average marginal effect to a negative effect around a BMI of 43 and after that. This can be seen in the image below.



We found that a 3rd-degree polynomial explained age the best in our model. It shows how being under the age of about 60. Age has a positive marginal effect on cognition score. It then becomes more negative as one ages at a decreasing rate. The below image shows the average marginal effects of age on cognition score for each age between 40 and 100.



Even though the fixed effects with OLS produced the lowest RMSE, we still didn't use it for inference. This is because it left out information on fixed variables throughout the individual's life. It is also due to the R-squared being 0.28 for the within and 0.06 for the between. By the R-squared parameter, fixed effects explained much less variance in our model than the random effects model. In a prediction scenario, we would use the fixed effects model. That is because it made better predictions and was a more simple model. However, this study is focused on inference and not a forecast.

There are several things we would like to explore or do differently in the future. We want to try a model with fewer non-statistically significant predictors to see how our RMSE and R-squared scores change. We would also like to clean some of the variables further where there may have only been a few instances of a categorical variable.

Our study also had many limitations. There were still hundreds of variables we still needed to explore. Diet is one variable we would explore in the future. We were also limited by the information given in the study. The study also only focused on those in adulthood and mainly retirement age. Specifically, adolescence and young adulthood were not included or focused on. We would like to look at other data containing cognition information for all ages. We could also try different clustered standard errors. However, we can still gain useful information from our current study.

Conclusion

Our exploratory analysis found many significant variables that give us insight into adult cognition. We found using random effects with Ordinary Least Squares produces the best fit model for inference. We also found that a fixed effects model might work better in a prediction scenario. This information can further our understanding and current research surrounding factors that affect our cognition in adulthood.

References

American Heart Association. "High blood pressure at any age, no matter how long you have it, may speed cognitive decline." *ScienceDaily*. Dec. 14, 2020. https://www.sciencedaily.com/releases/2020/12/201214090133.htm

Health and Retirement Study, (RAND HRS Longitudinal File 2018 (V2)) public use dataset. Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor, MI, July 2022. https://hrsdata.isr.umich.edu/sites/default/files/documentation/other/1658948491/randhrs1992_2018v2.pdf.

Morgan E. Levine, Amal Harrati & Eileen M. Crimmins (2018) Predictors and implications of accelerated cognitive aging, Biodemography and Social Biology, 64:2, 83-101, DOI: 10.1080/19485565.2018.1552513

Ramon Casanova, PhD, Santiago Saldana, MSc, Michael W Lutz, PhD, Brenda L Plassman, PhD, Maragatha Kuchibhatla, PhD, Kathleen M Hayden, PhD, Investigating Predictors of Cognitive Decline Using Machine Learning, *The Journals of Gerontology: Series B*, Volume 75, Issue 4, May 2020, Pages 733–742, https://doi.org/10.1093/geronb/gby054

RAND HRS Longitudinal File 2018 (V2). Produced by the RAND Center for the Study of Aging, with funding from the National Institute on Aging and the Social Security Administration. Santa Monica, CA, July 2022. https://hrsdata.isr.umich.edu/sites/default/files/documentation/other/1658948491/randhrs1992_2018v2.pdf

Smagula, Stephen F. PhD; Biggs, Mary L. PhD; Jacob, Mini E. MD, PhD; Rawlings, Andreea M. PhD; Odden, Michelle C. PhD; Arnold, Alice PhD; Newman, Anne B. MD; Buysse, Daniel J. MD. Associations of Modifiable Behavioral Risk Factor Combinations at 65 to 74 Years Old With Cognitive Health Span for 20 Years. Psychosomatic Medicine: September 2022 - Volume 84 - Issue 7 - p 785-792 doi: 10.1097/PSY.0000000000001100

"Subjective Cognitive Decline - A Public Health Issue." *Centers for Disease Control and Prevention*. Division of Population Health, National Center for Chronic Disease Prevention and Health Promotion: U.S. Department of Health and Services. Feb. 27, 2019. https://www.cdc.gov/aging/data/subjective-cognitive-decline-brief.html.

Appendix

Table 1: Summary Statistics 1

	Variable	Observations	Mean	Standard Dev.	Min	Max
Variable of Interest						
	Cognition Score	95,225	21.47	5.30	0	35
Demographic Variables						
	Age in Years	95,225	75.18	7.28	36	109
	Gender	95,225	1.59	0.49	1	2
	Race	95,225	1.20	0.48	1	3
	Is Hispanic	95,225	0.08	0.27	0	1
	Marital Status	95,225	3.36	2.80	1	8
Economic Variables						
	Labor Force Status	95,225	4.74	1.25	1	7
	Log of Earnings	95,225	1.46	3.46	0	14.29
	Log of Household Wealth	95,225	11.54	2.96	0	18.58
	Out of Pocket Medical Expenditures	41,873	3446.42	9208.31	0	634,821
	Out of Pocket Medical Expenditures Missing	95,225	0.56	0.50	0	1
	Region	95,225	5.02	2.34	1	11
	Total Medical Expenditures	24,487	3.9	2.12	0	11
	Total Medical Expenditures Missing	95,225	0.74	0.44	0	1
	Wave	95,225	8.97	3.02	4	14
	Years of Education	95,225	12.33	3.23	0	17
Health Variables						
	BMI	95,225	26.94	6.09	0	103.6
	BMI Missing	95,225	0.01	0.10	0	1
	Change in Health (SR)	94,597	0.08	0.88	-4	4
	Change in Health Missing (SR)	95,225	0.00	0.04	0	1
	Couldn't Get Going (SR)	95,225	0.22	0.41	0	1
	Difficulty Managing Money	95,225	0.39	1.68	0	9
	Difficulty Preparing Hot Food	95,225	0.49	1.88	0	9
	Difficulty Shopping for Food	95,225	0.35	1.49	0	9
	Difficulty Taking Meds	92,960	0.04	0.30	0	9
	Difficulty taking Meds Missing	95,225	0.02	0.13	0	1
	Difficulty Using a map	95,225	1.19	2.86	0	9
	Difficulty Using Phone	95,225	0.08	0.58	0	9

Note: (SR) stands for Self-Reported. This is Summary 1 of 2.

Table 2: Summary Statistics 2

	Variable	Observations	Mean	Standard Dev.	Min	Max
Health Variables Continued						
	Everything Feels Like Effort (SR)	95,225	0.23	0.42	0	1
	Feels Sad (SR)	95,225	0.18	0.39	0	1
	Has Alzheimer's	41,873	0.02	0.29	0	7
	Has Alzheimer's Missing	95,225	0.56	0.50	0	1
	Has Arthritis	95,225	0.76	0.63	0	5
	Has Cancer	95,225	0.20	0.46	0	5
	Has Dementia	41,873	0.03	0.25	0	4
	Has Dementia Missing	95,225	0.56	0.50	0	1
	Has Diabetes	95,225	0.26	0.53	0	5
	Has Had Stroke	95,225	0.10	0.37	0	5
	Has Heart Disease	95,225	0.38	0.70	0	6
	Has High Blood Pressure	95,225	0.70	0.62	0	5
	Has Lung Condition	95,225	0.15	0.48	0	5
	Has Psychological Condition	95,225	0.19	0.55	0	5
	Is Depressed (SR)	95,225	0.14	0.35	0	1
	Log of Days in Hospital Over 2 Years	94,720	0.28	0.47	0	4.62
	Memory Rating (SR)	95,225	3.08	0.92	1	5
	Memory Rating Change (SR)	95,225	2.22	0.46	1	3
	Number of Drinks per Day	95,225	0.54	1.09	0	45
	Breathing Test	61,165	144.61	186.56	0	870
	Breathing Test Missing	95,225	0.35	0.48	0	1
	Breathing Test Position	61,165	0.46	0.55	0	3
	Breathing Test Position Missing	95,225	0.35	0.48	0.00	1
	Sleep is Restless (SR)	95,225	0.28	0.45	0	1
	Smoked Ever	94,065	0.56	0.5	0	1
	Smoked Ever Missing	95,225	0.01	0.08	0	1
	Smokes Now	94,211	0.09	0.28	0	1
	Smokes Now Missing	95,225	0.01	0.07	0	1
	Walking Time Test	61,16s5	1.44	2.18	0	81
	Walking Time test Missing	95,225	0.35	0.48	0	1
	Was Happy (SR)	95,225	0.89	0.32	0	1

Note: (SR) stands for Self-Reported. This is Summary 2 of 2.

Best Model Coefficient Output Part 1

		Coefficient	Standard Error	Z	P>|Z|	95% CI	
Constant		-19.16	9.84	-1.95	0.05	-38.44	0.12
Female		0.94	0.05	18.55	0.00	0.84	1.04
Black/African American		-2.42	0.08	32.19	0.00	-2.57	-2.28
Other		-1.05	0.13	-8.33	0.00	-1.29	-0.80
Hispanic		-0.35	0.10	-3.41	0.00	-0.55	-0.15
Has High Blood pressure							
Yes		-1.81	0.35	-5.11	0.00	-2.50	-1.11
	3	-8.45	4.89	-1.73	0.08	-18.03	1.13
	4	1.09	1.03	1.06	0.29	-0.93	3.12
	5	2.85	3.25	0.88	0.38	-3.52	9.21
BMI Missing		2.02	0.27	7.43	0.00	1.48	2.55
Smokes Now		-0.08	0.07	-1.24	0.22	-0.21	0.05
Smoked Ever		0.00	0.05	-0.01	0.99	-0.10	0.10
Smoked Ever Missing		0.22	0.31	0.7	0.48	-0.39	0.83
Years of Education		0.48	0.01	55.01	0.00	0.46	0.49
Self-Reported Health		-1.02	0.44	-2.34	0.02	-1.88	-0.17
Change in Self-Reported Health		0.03	0.01	2.61	0.01	0.01	0.05
Number of Drinks Per Day		0.05	0.01	3.47	0.00	0.02	0.08
Memory Compared to Past (SR)							
Same		0.75	0.10	7.78	0.00	0.56	0.94
Worse		0.67	0.10	6.7	0.00	0.47	0.87
Married, Spouse Absent		0.02	0.14	0.14	0.89	-0.26	0.30
Partnered		-0.23	0.10	-2.25	0.02	-0.43	-0.03
Separated		-0.37	0.15	-2.5	0.01	-0.67	-0.08
Divorced		0.07	0.07	1.02	0.31	-0.07	0.21
Separated/Divorced		0.00	0.92	0	1.00	-1.81	1.81
Widowed		0.00	0.05	-0.05	0.96	-0.09	0.09
Never Married		-0.26	0.13	-1.96	0.05	-0.53	0.00
Is Depressed (SR)		-0.40	0.04	-8.94	0.00	-0.49	-0.31
effort							
Everything Feels Like Effort		-0.19	0.03	-5.59	0.00	-0.26	-0.13
Sleep Is Restless		0.13	0.03	4.35	0.00	0.07	0.18
Has Arthritis							
Yes		0.20	0.04	5.02	0.00	0.12	0.28
	3	-0.12	0.59	-0.2	0.84	-1.27	1.03
	4	-0.23	0.10	-2.34	0.02	-0.42	-0.04
	5	-0.16	0.30	-0.53	0.59	-0.75	0.43

Note: Variables that had 2, 3, 4, 5, and 6 in place of an indicator variable were missing and not sure responses. These include cases such as heart disease, has had a stroke, a psychological conditions, lung disease, diabetes, breathing test position, Alzheimer's, and dementia.

Best Model Coefficient Output Part 2

	Coefficient	Standard Error	Z	P>|Z|	95% CI	Coefficient
Has Heart Disease						
Yes	0.03	0.04	0.85	0.39	-0.04	0.11
3	-0.12	0.89	-0.13	0.90	-1.86	1.63
4	-0.21	0.12	-1.7	0.09	-0.46	0.03
5	-0.34	0.50	-0.67	0.50	-1.31	0.64
Has had Stroke						
Yes	-0.74	0.07	10.74	0.00	-0.88	-0.61
TIA/Possible Stroke	0.09	0.14	0.63	0.53	-0.19	0.36
3	-0.87	0.83	-1.06	0.29	-2.49	0.75
4	-0.90	0.27	-3.41	0.00	-1.42	-0.38
5	0.02	0.84	0.02	0.98	-1.63	1.67
Has Psychological Condition						
Yes	-0.18	0.05	-3.31	0.00	-0.28	-0.07
3	-0.10	0.57	-0.17	0.87	-1.21	1.01
4	-0.09	0.12	-0.82	0.41	-0.32	0.13
5	0.09	0.38	0.24	0.81	-0.65	0.83
Has Lung Disease						
Yes	0.15	0.06	2.68	0.01	0.04	0.26
3	0.01	0.53	0.01	0.99	-1.03	1.04
4	-0.04	0.14	-0.28	0.78	-0.32	0.24
5	-0.28	0.43	-0.64	0.52	-1.12	0.57
Has Diabetes						
Yes	-0.26	0.04	-5.92	0.00	-0.35	-0.18
3	-0.34	0.74	-0.46	0.65	-1.78	1.11
4	-0.34	0.14	-2.39	0.02	-0.61	-0.06
5	-0.14	0.62	-0.22	0.83	-1.34	1.07
Memory (SR)						
Very Good	0.05	0.07	0.71	0.48	-0.09	0.19
Good	-0.16	0.07	-2.23	0.03	-0.29	-0.02
Fair	-0.49	0.07	-6.55	0.00	-0.63	-0.34
Poor	-1.20	0.09	13.06	0.00	-1.39	-1.02
Employed	0.06	0.09	0.62	0.54	-0.12	0.24
Unemployed	-0.11	0.14	-0.8	0.43	-0.39	0.16
Partly Employed	0.01	0.06	0.18	0.86	-0.11	0.13
Retired	-0.31	0.06	-4.92	0.00	-0.43	-0.19
Disabled	-0.37	0.16	-2.27	0.02	-0.69	-0.05
Not in Labor Force	-0.28	0.08	-3.4	0.00	-0.45	-0.12
Log of Total Household Wealth	0.14	0.01	18.58	0.00	0.12	0.15
Log of Time in Hospital Over the Last 2 Years	-0.12	0.03	-4.26	0.00	-0.17	-0.06
Log of Earnings	0.01	0.00	2.88	0.00	0.00	0.02
Walk Time Test	-0.04	0.01	-3.58	0.00	-0.05	-0.02

Note: Variables that had 2, 3, 4, 5, and 6 in place of an indicator variable were missing and not sure responses. These include cases such as heart disease, has had a stroke, a psychological conditions, lung disease, diabetes, breathing test position, Alzheimer's, and dementia.

Best Model Coefficient Output Part 3

		Coefficient	Standard Error	Z	P>|Z|	95% CI	
Breathing Test Score		0.00	0.00	8.86	0.00	0.00	0.00
Breathing Test Position							
	1	-0.49	0.08	-6.4	0.00	-0.63	-0.34
	2	-0.75	0.11	-6.7	0.00	-0.97	-0.53
	3	-0.13	0.89	-0.15	0.88	-1.88	1.61
Has Alzheimer's							
Yes		-3.65	0.27	13.46	0.00	-4.18	-3.12
	3	-2.67	1.81	-1.48	0.14	-6.21	0.87
	4	0.22	0.50	0.43	0.66	-0.76	1.19
	7	-0.64	0.68	-0.93	0.35	-1.98	0.70
Has Dementia							
Yes		-2.63	0.20	13.47	0.00	-3.02	-2.25
	3	-0.27	1.82	-0.15	0.88	-3.83	3.30
	4	-1.27	0.34	-3.76	0.00	-1.93	-0.61
Felt Sad		0.08	0.04	2.17	0.03	0.01	0.15
Couldn't Get Going		-0.09	0.03	-2.92	0.00	-0.16	-0.03
Enjoyed Life		-0.07	0.06	-1.27	0.21	-0.18	0.04
Was Happy		-0.07	0.05	-1.44	0.15	-0.16	0.02
Has Cancer		0.26	0.05	5.53	0.00	0.17	0.35
	3	-2.80	1.06	-2.65	0.01	-4.87	-0.73
	4	-0.18	0.21	-0.88	0.38	-0.58	0.22
	5	0.36	0.64	0.57	0.57	-0.89	1.61
Trouble Using Phone							
Yes		-0.73	0.08	-8.71	0.00	-0.89	-0.56
Can't		-1.81	0.41	-4.44	0.00	-2.60	-1.01
Don't		-1.31	0.23	-5.79	0.00	-1.75	-0.87
Trouble taking Meds Missing		-0.08	0.09	-0.82	0.41	-0.26	0.11
Trouble Taking Meds							
Yes		-0.90	0.10	-9.38	0.00	-1.09	-0.71
Can't		-2.45	0.55	-4.44	0.00	-3.54	-1.37
Don't		-0.56	0.54	-1.04	0.30	-1.62	0.50
Trouble Managing Money							
1.yes		-1.48	0.07	20.29	0.00	-1.63	-1.34
2.can't		-1.97	0.23	-8.41	0.00	-2.43	-1.51
9.don't		-0.84	0.07	11.9	0.00	-0.98	-0.70
Trouble Shopping							
1.yes		-0.25	0.06	-4.43	0.00	-0.36	-0.14
2.can't		-0.43	0.18	-2.43	0.02	-0.78	-0.08
9.don't		-0.54	0.08	-6.44	0.00	-0.71	-0.38

Note: Variables that had 2, 3, 4, 5, and 6 in place of an indicator variable were missing and not sure responses. These include cases such as heart disease, has had a stroke, a psychological conditions, lung disease, diabetes, breathing test position, Alzheimer's, and dementia.

Best Model Coefficient Output Part 4

	Coefficient	Standard Error	Z	P>|Z|	95% CI	
Trouble Preparing Hot Meals						
1.yes	-0.59	0.07	-8.62	0.00	-0.72	-0.45
2.can't	-0.53	0.18	-2.88	0.00	-0.88	-0.17
9.don't	-0.38	0.06	-6.02	0.00	-0.50	-0.26
Trouble Using a Map						
1.yes	-0.57	0.05	11.78	0.00	-0.67	-0.48
2.can't	-0.70	0.09	-8.11	0.00	-0.86	-0.53
9.don't	-0.53	0.05	11.48	0.00	-0.62	-0.44
Out of Pocket Medical Expenditures	0.00	0.00	-2.06	0.04	0.00	0.00
Total Medical Expenditures	0.01	0.01	1.01	0.31	-0.01	0.03
Age	1.40	0.39	3.63	0.00	0.65	2.16
Age^2	-0.02	0.01	-3.41	0.00	-0.03	-0.01
Age^3	0.00	0.00	2.47	0.01	0.00	0.00
HIBP * Age						
1.yes	0.02	0.00	4.99	0.00	0.01	0.03
3.disp	0.12	0.06	1.81	0.07	-0.01	0.24
4.disp	-0.02	0.01	-1.34	0.18	-0.04	0.01
5.disp	-0.05	0.04	-1.13	0.26	-0.13	0.04
BMI * BMI	0.00	0.00	-7.06	0.00	0.00	0.00
BMI * Age	0.00	0.00	9.58	0.00	0.00	0.00
Wave 5	-0.30	0.05	-5.67	0.00	-0.40	-0.19
Wave 6	-0.30	0.05	-5.46	0.00	-0.41	-0.19
Wave 7	-0.50	0.07	-7.22	0.00	-0.63	-0.36
Wave 8	-0.47	0.07	-6.58	0.00	-0.61	-0.33
Wave 9	-0.44	0.07	-6.08	0.00	-0.58	-0.30
Wave 10	-0.88	0.07	11.7	0.00	-1.02	-0.73
Wave 11	-0.82	0.08	10.72	0.00	-0.97	-0.67
Wave 12	-0.64	0.08	-8.08	0.00	-0.79	-0.48
Wave 13	-0.57	0.08	-7.14	0.00	-0.73	-0.42
Wave 14	-0.23	0.08	-2.76	0.01	-0.40	-0.07
Mid Atlantic	0.41	0.12	3.43	0.00	0.18	0.65
Northeast Central	0.11	0.12	0.94	0.35	-0.12	0.34
Northwest Central	0.21	0.13	1.63	0.10	-0.04	0.45
South Atlantic	0.05	0.11	0.48	0.63	-0.16	0.27
Southeast Central	-0.24	0.13	-1.77	0.08	-0.50	0.03
Southwest Central	-0.01	0.13	-0.07	0.95	-0.25	0.24
Mountain	0.14	0.14	1.04	0.30	-0.13	0.41
Pacific	0.23	0.12	1.88	0.06	-0.01	0.46
Not in US	-0.21	0.45	-0.48	0.63	-1.08	0.66

Note: Variables that had 2, 3, 4, 5, and 6 in place of an indicator variable were missing and not sure responses. These include cases such as heart disease, has had a stroke, a psychological conditions, lung disease, diabetes, breathing test position, Alzheimer's, and dementia.