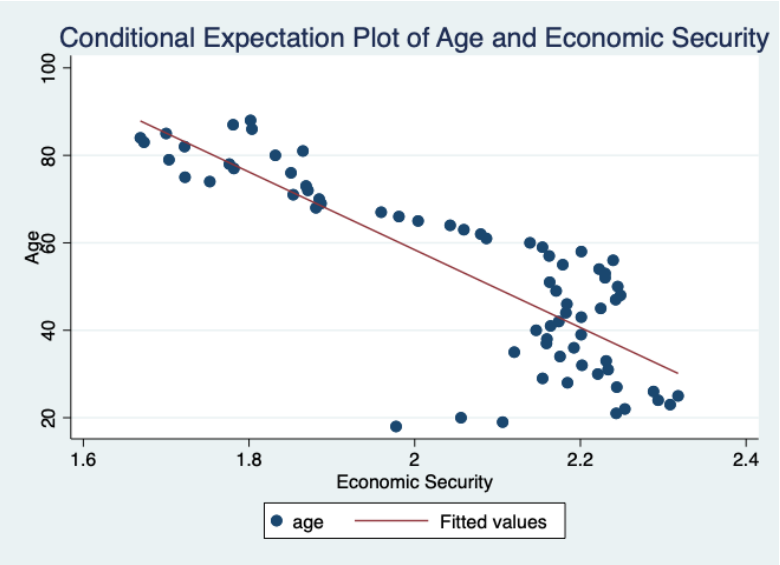
Alessa Somer
Econometrics (Econ 255)
Data Brief #2

Introduction

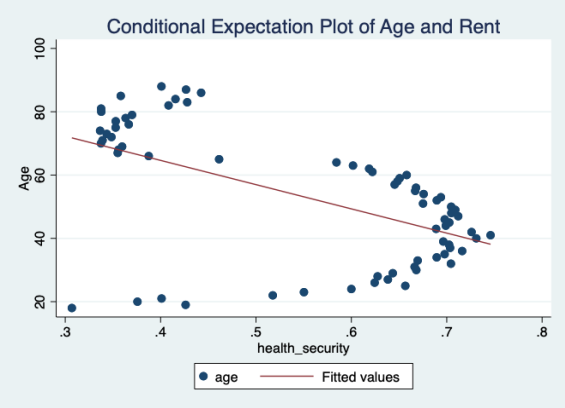
This data brief is utilizes the "Household Pulse Survey: Measuring Social and Economic Impacts duing the Coronavirus Pandemic." The survey is conducted by the U.S. Census Bureau, which is the principal agency for the U.S. Federal Statistical System. "The Household Pulse Survey is a 20-minute online survey studying how the coronavirus pandemic and other emergent issues impacting households across the country from a social and economic perspective. The HPS also asks about core demographic household characteristics, as well as the following topics: Access to infant formula, Children's mental health treatment, COVID-19 vaccinations and long COVID symptoms and impact, Use of antivirals to treat COVID-19, Education, specifically K-12 enrollment, Childcare arrangements, Employment, Food sufficiency, Housing security, etc." The data collection process occurred in several phases from April 2020 - October 2023. The Census Bureau randomly choses houeshold addresses to participate in this survey. This means that only people in the United States with permanent addresses were included in this survey. A limited number of addresses from across the country were scientifically selected to represent the entire population. This data brief will focus on the relationship between the age of the head of the household an economic security in the form of food security, housing security, and healthcare security. This data brief in particular will examine the linear relationship between these variables using OLS. The relationship between age and economic security is important because if there are trends in areas where certain ages have lower economic security, it might indicate to the U.S. government to work to add more economic support for programs targeted at these specific age ranges. The household pulse survey allows us to study the relationship between economic security and age because as mentioned above, it provides very useful data on economic security due to its extensive economic focused questions.

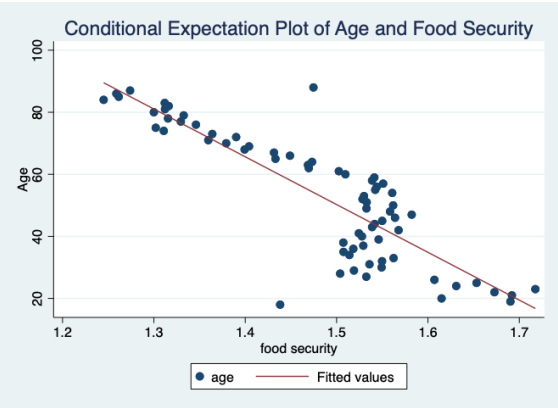


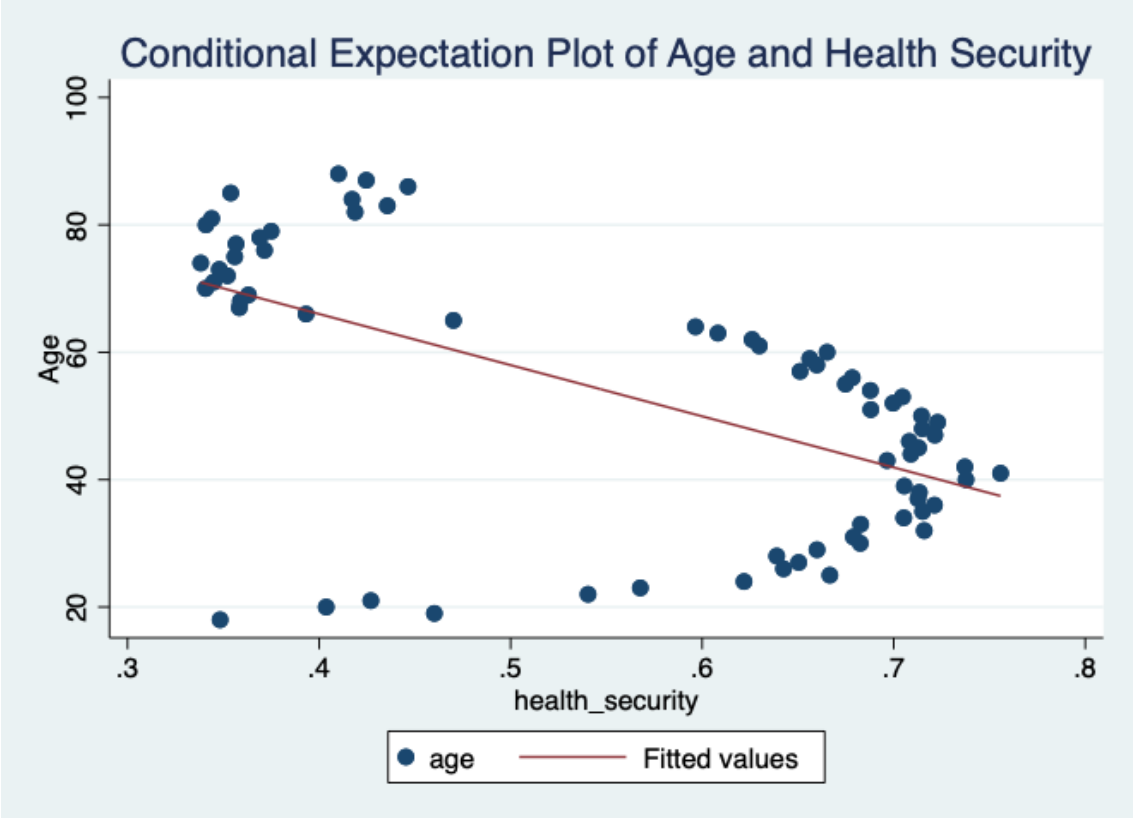
Descriptive Statistics

First, we will examine the relationship between overall economic security and age in this data. In this data brief, "economic security" is gathered from the categorical variable in the data set called 'expns_dif.' In the household survey, people were asked the

question "in the last 7 days, how difficult has it been for your household to pay for usual household expenses, including but not limited to food, rent, or mortgage, car payments, medicail expensises and so on?" Households had the choice to answer with the following options. 1) not at all difficult 2) a little difficult 3)somewhat difficult 4) very difficult -99) question seen but category not selected -88) missing /did not report. The data was cleaned to omit data points with -99 and -88 as answers. As shown by the two way scatter plot, the expectation of age changes with economic security.







The variable in this data brief to measure housing security is represented by a binary "Rent" variable. This variable is derived from 'rentcur' in the census where households were asked if

they “are caught up on rent” and had the following choices to answer - 1 -yes 2 - no. The data set was cleaned to only include data points that had values of 1.

The variable in this data brief to measure food security is represented by a categorical “food security” variable. This variable is derived from 'curfoodsuf' in the census where households were asked about their household food sufficiency for the last 7 days. Households were given the choices of 1- "enough of the kinds of food (I/we wanted to eat)" 2- "Enough but not always the kinds of food (I/we wanted to eat)" 3- "sometimes not enough to eat" 4- "often not enough to eat.” The data was cleaned to only include data points that had values 1-4 (omitting missing data points).
The variable in this data brief to measure health security is represented by a binary “health insurance” variable. This variable is generated from the ‘HLTHINS1’ variable in the census where households were asked if they have insurance through a current or former employer or union (through yourself or another family member). Households could either answer 1) yes or 2) no. The variable was cleaned to only include if households had insurance.

Econometric Analysis

Regression Outputs

	(1)	(2)	(3)	(4)
	age	age	age	age
EXPNS_DIF=1	0	0	0	0
	(.)	(.)	(.)	(.)
EXPNS_DIF=2	-3.290***	-2.738***	-3.203***	-2.914***
	(0.152)	(0.150)	(0.145)	(0.153)
EXPNS_DIF=3	-3.411***	-2.609***	-3.503***	-2.895***
	(0.173)	(0.170)	(0.166)	(0.187)
EXPNS_DIF=4	-5.654***	-4.523***	-6.147***	-4.723***
	(0.189)	(0.186)	(0.182)	(0.229)
rent_yes=0		0	0	0
		(.)	(.)	(.)
rent_yes=1		-7.949***	-7.747***	-7.640***
		(0.154)	(0.149)	(0.150)
health_security=0			0	0
			(.)	(.)
health_security=1			-7.819***	-7.938***
			(0.120)	(0.120)
CURFOODSUF=1				0
				(.)
CURFOODSUF=2				-0.739***
				(0.155)
CURFOODSUF=3				-2.857***
				(0.279)
CURFOODSUF=4				-2.986***
				(0.433)
Constant	55.12***	56.18***	61.32***	61.43***
	(0.0987)	(0.0989)	(0.124)	(0.125)
N	66249	66249	66249	66249
r2	0.0165	0.0545	0.112	0.113

Regression 1: Bivariate OLS regression between age and economic security or 'expns_dif'. Regression 2: Multivariate regression of age on econimic security and housing security (represented by rent). Regression 3: Multivariate Regression of age on economic security, housing security, and health access (represented by access to health insurance). Regression 4: Multivariate

Regression 1 Interpretation: On average, individuals who perceive their household expenses as "a little difficult" are estimated to be 3.290 years younger compared to those who reported, "not

at all difficult". On average, individuals who perceive their household expenses as "somewhat difficult" are estimated to be 3.411 years younger compared to those who reported "not at all difficult". On average, individuals who perceive their household expenses as "very difficult" are estimated to be 5.654 years younger compared to those who reported "not at all difficult".

Regression 2 Interpretation: On average, individuals who perceive their household expenses as "a little difficult" are estimated to be 2.738 years younger compared to those who reported "not at all difficult", given that the household's rent has been paid in the past 7 days. On average, individuals who perceive their household expenses as "somewhat difficult" are estimated to be 2.609 years younger compared to those who reported "not at all difficult", given that the household's rent has been paid in the past 7 days. On average, individuals who perceive their household expenses as "very difficult" are estimated to be 4.523 years younger compared to those who reported "not at all difficult", given that the household's rent has been paid in the past 7 days. On average, individuals who have paid rent are estimated to be 7.949 years younger compared to those who haven't paid rent, controlling for "economic_security" or "expns_dif" in the regression model.

Regression 3 Interpretation: On average, individuals who perceive their household expenses as "a little difficult" are estimated to be 3.203 years younger compared to those who reported "not at all difficult", given that the household's rent has been paid in the past 7 days and the household has health insurance. On average, individuals who perceive their household expenses as "somewhat difficult" are estimated to be 3.503 years younger compared to those who reported "not at all difficult", given that the household's rent has been paid in the past 7 days and the household has health insurance. On average, individuals who perceive their household expenses as "very difficult" are estimated to be 6.147 years younger compared to those who reported "not at all difficult", given that the household's rent has been paid in the past 7 days and the household has health insurance. On average, individuals who have paid rent are estimated to be 7.747 years younger compared to those who haven't paid rent, controlling for "economic_security" or "expns_dif" in the regression model. On average, the average difference in age for individuals with health security compared to those without health security is estimated to be -7.819, while the values for household expense difficulty and if rent has been paid are the same.

Regression 4 Interpretation: On average, individuals who perceive their household expenses as "a little difficult" are estimated to be 2.914 years younger compared to those who reported "not at all difficult", given that the household's rent has been paid in the past 7 days, the household has health insurance, and the household food sufficiency is the same. On average, individuals who perceive their household expenses as "somewhat difficult" are estimated to be 2.895 years younger compared to those who reported "not at all difficult", given that the household's rent has been paid in the past 7 days, the household has health insurance, and the household food

sufficiency is the same. On average, individuals who perceive their household expenses as "very difficult" are estimated to be 4.723 years younger compared to those who reported "not at all difficult", given that the household's rent has been paid in the past 7 days, the household has health insurance, and the household food sufficiency is the same. On average, individuals who have paid rent are estimated to be 7.640 years younger compared to those who haven't paid rent, controlling for "economic_security" or "expns_dif" in the regression model. On average, the average difference in age for individuals with health security compared to those without health security is estimated to be 7.938, while the values for household expense difficulty, if rent has been paid are the same, and the household food sufficiency is the same. On average, individuals who reported 2 for food security have an estimated age that is 0.739 years younger compared to those who reported 1 for food security, controlling for the effects of other variables in the model. On average, individuals who reported 3 for food security have an estimated age that is 2.857 years younger compared to those who reported 1 for food security, controlling for the effects of other variables in the model. On average, individuals who reported 4 for food security have an estimated age that is 2.986 years younger compared to those who reported 1 for food security, controlling for the effects of other variables in the model.

Discussion and Limitations

Based on the results of the linear regression, we can see that the standard error values decrease for each coefficient from Regression 1 to Regression 2. Additionally, the R-squared value increased from 0.0165 to 0.0545 in Regression 1 to Regression 2 respectively. This means that adding the housing security variable in Regression 2 helped us capture more of the variation in "age" in our model. Additionally, the R-squared value increased from 0.0545 to 0.112 in Regression 2 to Regression 3 respectively. Therefore adding the variable "health_security" helps us capture more of the variation in "age" in the model. The R-squared value increased from 0.112 to 0.113 from Regression 3 to Regression 4. Because the R-squared value only increased slightly, it points out that adding the "food security" variable did not help us explain the variation in age too much.

We can't make causal claims about the relationships identified because the data has not been randomly controlled for every factor in the experiment. Or rather, this is a census, not an experiment, so it would be very difficult to be able to make causal inferences about relationships from the study. There are certainly hazards to this data set, and therefore hazards and bias in our interpretations in the data brief. One significant hazard of this data set is that many of the data points have values of "-88" or "-99," meaning that the households didn't have access to certain subsets of questions or chose to not answer them respectively. Therefore, some of the variables that had majority values of -88 should not necessarily be utilized in our OSL and are not as useful as we would have desired. Additionally, it is noteworthy that this census was only seen by people in the United States who have permanent residence. This leads to an omission of subsets

of people such as college students (do their parents still consider them part of their household?), or people currently facing housing insecurity/instability. Another hazard of this data is acknowledging that there may be skewness since missing data points for each variable were omitted. This means that only data points where there were non “-99” or “-88” values for ALL variables mentioned above were examined.

One point of information not included in this analysis is the geographic factor. Although geographic region and state are included in the house survey, I did not look at this information. There could be higher proportions of households included in these regressions from one region of the United States compared to the other that we did not account for. Even though the house survey was sent to random households, my removal of data points that didn’t include values of numerical value could have skewed the data.