Socio-economic Causes of Depression

Cooper Lawrenz

Department of Economics
University of Puget Sound
Aug 18, 2023

Abstract

Depression is expensive to all nations in the world. How the causes of depression are understood is constantly evolving. This paper looks at socioeconomic contributors to depression with care for current understandings of the problem along with recent data. Previous understandings of an inverse relationship between income and depression are bolstered here along with other mental health symptoms. The analysis is conducted with an ordinal logistic regression model assuming proportional odds, implemented in this regression are two unique instruments for personal and family income. The results in this paper are relevant to public policy professionals who aim to minimize depression's cost to their society.

1 Introduction

As numerous papers investigating the interactions between economic conditions and mental health argue, depression incurs significant costs both to the individual and their society. A wide array of studies, conducted in different places and times, have attempted to estimate this cost. Broad calculations have been reported by Chang et al (2012), estimating a total cost of 4,049 million USD annually in South Korea. Similarly, Hu, T. W. (2007) estimated an annual cost of 6,264 million USD in China. König, H. et al (2020) conducted a meta-analysis, stating that "Depression was associated with significantly higher total direct excess costs in all subgroups." Examining an unnamed US corporation, Druss, B.G. et al (2000) concluded that depression is was one of the costliest medical illnesses for that employer. Other approaches, such as that taken by Kessler, R.C. (2012), investigate not only financial costs but also social costs such as marriage status, citing various studies that determine a higher likelihood of divorce for those plagued with depressive conditions.

This extensive justification supports research into causal factors of mental health conditions. While this study is not the first to approach this topic, it is believed that, in this paper, different response variables are used along with a unique combination of controls. Additionally, not all analyses employ instrumental strategies to mitigate a hypothesized bi-directional causality between income and depression. Even though not all studies control for this omitted variable bias, evidence of this bi-directionality can be traced back to Miech RA et al. (1999). Furthermore, as argued by Akhtar-Danesh et al (2007), "Given the significant impact of depression on individuals and society as a whole, a

comprehensive analysis of the prevalence of depression is necessary to ensure that previous findings remain applicable in today's society." Their study was conducted in 2007, and our understanding of mental health has evolved significantly since then.

Appreciation for this topic in the realm of economics can be traced back to an article from the late 1990s written by Ettner SL (1996). Ettner used parental education as an instrument for personal income he additionally used a spouse's education as its own instrument. Ettner concluded that there is a significant relationship between income and mental health, and even given the work's age, its ideas behind instruments for this problem remain relevant.

This being said, the past thirty years have seen monumental changes in both how mental health problems are appreciated and understood. This evolution has been analyzed by G. Schomerus et al. (2012), where it is argued that biological circumstances were not always appreciated prior to the mid-2000s. As such, Many earlier studies did not appreciate non-socioeconomic factors and therefore did not control for them. This study and others more recent, however, attempt to address this.

Moving forward in time, Zimmerman, F. J. et al. (2005) argue that depression and income have a strong correlation on their own; however, when controlling for other factors, this correlation becomes less pronounced, an ever-so idiosyncratic conundrum of econometrics. Of particular interest to the thesis of this paper, Zimmerman F.J. et al. use a large list of instruments for their study, including: logged total inheritance, logged time in the current job, logged mothers' education, the fraction of household income earned by the respondent, hours of television watched per week, rural/suburban/urban residence, and age. Borrowed from this analysis is the use of, the fraction of household income earned by respondents. Notably, both studies prior to this and the results found here invalidate some of Zimmerman's instruments. For example, both Mirowsky, J. et al (1992) and Brodaty, H (1997) demonstrate a correlation between age and depression. Additionally, work done by Wanderman, A. et al (1998) illustrates a connection between neighborhood characteristics and mental health outcomes.

In a 2007 study of a Canadian population using data from 2002, Akhtar-Danesh et al. looked at the socioeconomic impacts of depression. In support of earlier work, they once again observed income to possess a negative correlation, indicating that as income increases, depression decreases. They differentiated between 12-month and lifetime depression as dependent variables, both of which were regressed as binary, based on whether they satisfied the requirements for fitting into the categories. Nothing in their paper suggests the use of an instrument. In their modeling, they demonstrated correlations with: age, gender, marital status, education, household income, and immigration status. In most recent studies of depression, most of these variables are controlled for. Additionally, odds ratio coefficients are used here, a common reporting tool used when studying mental health response variables. Layte, R. (2012) offers an explanation of income inequality being a causal variable. In his study, he observed a wide range of European countries. The analysis took particular care to control for sex, age, and health status. The results of this analysis concluded a

relationship between income inequality and depression, suggesting a degree of external validity of these studies on Western countries while also theorizing a logical connection between income inequality and depression, that individuals who make less, upon comparing themselves to others, could feel inadequate. In line with a greater appreciation for internal factors contributing to rates of depression, Dijkstra-Kersten et al. (2015) provide empirical proof that income and anxiety do have a relationship with one another, particularly that lower-income individuals and families may feel more uncertain about their financial health and well-being in the future. Finally, to the best of my knowledge, there is little to no previous literature studying this problem under the lens of an ordinal logistic regression model with instrumental variables. The most similar work can be seen by Lombardo et al. (2018) who studied mental health response variables as functions of socioeconomic conditions in 2018, they did however not appear to include some form of an instrumental variable.

As demonstrated above, depression is a complicated issue. This analysis provides explanations of income and depression with respect to current understandings of the topic and previous literature. How our societies view this topic is constantly changing, justifying renewed analysis with different techniques and data sets.

2 Data

The data used in this analysis comes from the IPUMS, a harmonized longitudinal microdata set. More specifically, the dataset utilized is the Medical Expenditure Panel Survey (MEPS). This dataset is primarily constructed from data obtained from the US Census Bureau, the Bureau of Labor Statistics, the National Center for Health Statistics, and the Centers for Disease Control. A substantial portion of the funding for IPUMS is provided by the National Institutes of Health, the National Science Foundation, and the Food and Drug Administration. Oversight of the dataset is conducted by the University of Minnesota Institute for Social Research and Data Innovation.

The dataset used in this analysis covers the years 2010 to 2018 which looks at the United States population and comprises slightly over 900,000 observations. Adolescents are not included in this study; thus, any individual under the age of nineteen was excluded from the analysis. Additionally, individuals who refused to answer or were never asked the questions used in these analyses were removed. All said and done the original dataset was reduced to 200,000 observations. It is assumed that this decision added no additional omitted variable bias.

All financial data has been adjusted to the base year of 2009 using a Consumer Price Index and divided by 1000 for both computational and interpretive purposes [1]. Two additional variables used in the analysis were constructed from the dataset. The first variable represents the percentage of annual family income that an individual contributes. Negative values do exist in this dataset and should be interpreted as negative contributions to a family's overall income.

[1] this will be addressed later

The second variable represents capital gains, which is a summation of increased value acquired from an individual's assets. This includes income received from interest, dividends, property sales, rent, and individual retirement accounts.

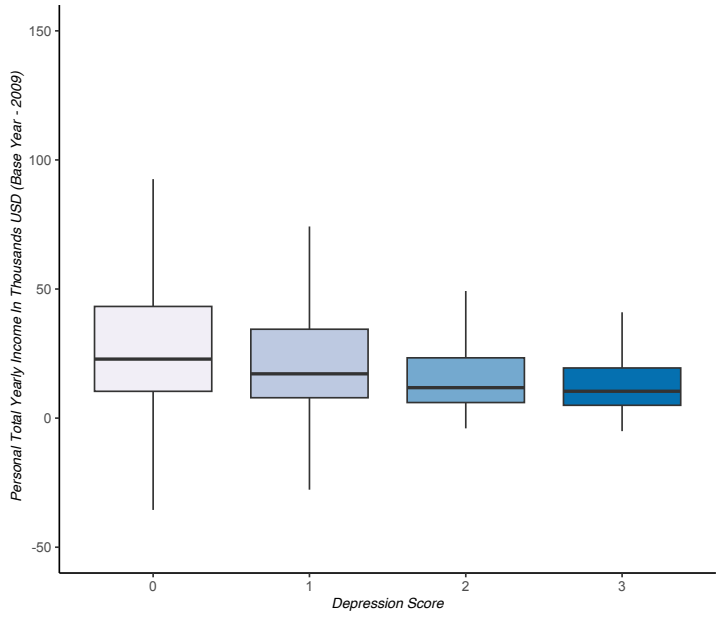
2.1 Surveys and Variables of Interest

To measure the severity of depression survey responses are used. Two main tests are employed here: the PHQ-2 and the K9 Scale. Seven different survey results are utilized to gain a better understanding of interactions with both different tests and questions. The formatting of these questions can be found in Appendix Figure C. Concerns regarding the accuracy of survey data will be addressed later.

I hypothesize there to be an inverse relationship between income and depression. Figures 1 and 2 below Support this hypothesis. Figure 1 Investigates individual personal income while Figure 2 Investigates Family personal income. The particular survey used in both Figures is the PHQDEP [2]. In this analysis, larger numbers are interpreted as higher frequencies of depression. Both Figure 1 and Figure 2 demonstrate that individuals who score higher on the PHQDEP survey also, on average, have lower incomes.

Figure 1

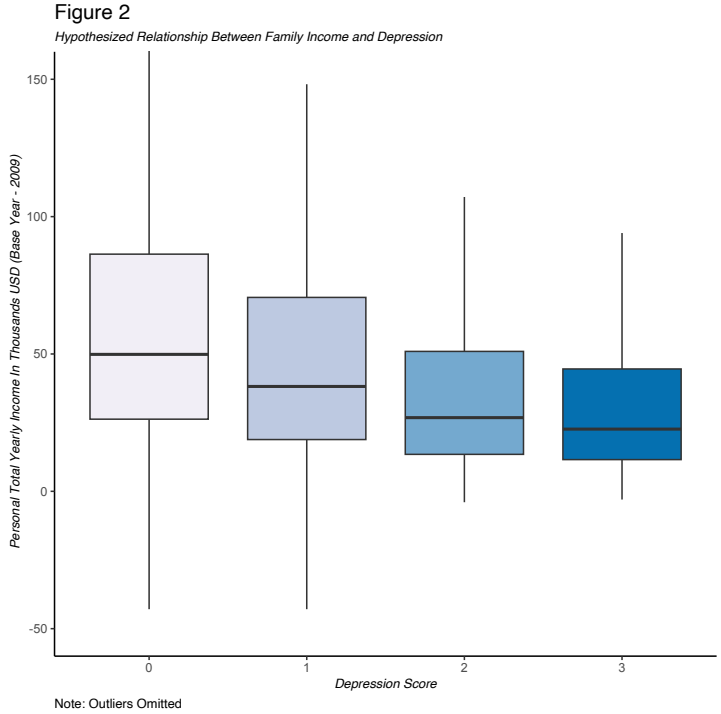
Hypothesized Relationship Between Income and Depression



Note: Outliers Omitted

[2] See appendix section C, question 11 b for exact formatting of the survey question

Figure 2



Two main income variables and seven survey variables are investigated in this analysis. The first two variables pertain to an individual's personal and family income. The survey questions cover a range of mental health symptoms. It should be noted that part c of question 10 is used as a control for anxiety, rather than being treated as its own dependent variable. Additionally, as part of the K9 scale series of questions, it is utilized to control, to the best extent possible, the unobservable differences in how individuals answer survey questions regarding their own mental health problems.

2.2 Controls

As discussed previously, mental health is a complex issue with a multitude of causal variables. Consequently, a host of different controls are employed. Age, gender, marital status, and education have all been extensively studied as significant factors in the depression equation, as demonstrated by Altemus, M. et al (2014) and Droogenback, F. et al. (2018). Additionally, the work by Galobardes and Shaw (2006) indicates a common consensus on the importance of these variables in social studies. Therefore, these variables are included in the model. Rosenquist, Fowler, and J.H. (2011) establish a connection with 'social networks,' showing that individuals with smaller social networks have higher rates of depression; hence, family size is included [3].

[3]This was the best proxy available to the data set

There exists a long-standing consensus regarding the correlation between anxiety and depression, as argued by Dobson (1985), among others. However, anxiety, especially in the field of economics, has not always been used as a control in similar analyses. Given the strongly hypothesized omitted variable bias (OVB) that would occur from its omission, anxiety is included. Fluharty et al. (2016), Paperwalla et al. (2004), and Mendelsohn et al. (2012) all demonstrate a link between smoking and mental health problems. Additionally, Van Ryn et al. (2000), LaVeist et al. (2005), and Cénat et al. (2021) argue that race has correlations with depression, most likely as a result of socioeconomic conditions.

Finally, Spiegel et al. (2003) show that cancer can contribute to mental health problems[4]. In addition, US regions (East, South, Midwest, and West) are used to control for geographical omitted variable bias. Lastly, the language in which the interview was conducted is used to control for potential cultural, socioeconomic, or survey-based differences.

2.3 Instruments

The impact of depression on both employment and productivity is evident in studies by Dooly D. et al. (1994), Lerner, D. et al. (2004), and Jefferis, B. J. et al. (2011). A bidirectional causal relationship is apparent between income and depression [5], justifying the use of instruments. Two instruments are employed in this analysis. The first is the fraction of family income that an individual contributes, a concept traced back to Zimmerman, F. J. This variable is specifically utilized as an instrument for personal income. Statistical tests investigating the relationship between income and the percentage of family total income and depression yielded insignificant results at the 5 percent level. Furthermore, the same tests demonstrated an extremely significant relationship between percentage income totals and personal income totals.

Alternatively, total capital gains are employed as an instrument for family income total. There is not a substantial amount of prior literature on using this as an instrument; however, statistical tests indicated that it was highly correlated with family income totals, and not with depression, making it a valid instrument for this dataset.

3 Model

To gain a better understanding of the data, an ordinal logistic model is fit, assuming proportional odds. [6] this model is utilized due to the ordered nature of the response data, where a clear rank and order exist among the dependent

[4] It's important to note that the 'cancer' variable here refers to ever having been diagnosed with it, as opposed to currently having it, which is a limitation of the chosen dataset

[5] it is assumed that unemployed individuals have lower overall incomes

[6] For a more in-depth exploration of the theory underpinning ordinal logistic proportional odds regression modeling, Frank E. Harrell's book(2001) provides valuable insights

variables. These ranks represent the severity of an individual's symptoms. As Harrell explains, there are alternative strategies for fitting this model. However, the proportional odds variant is chosen for its computational efficiency, which is particularly advantageous for a model that already demands significant computation. Moreover, proportional odds is the most commonly used model in social studies when dealing with ranked dependent variables.

In this model, given that the dependent variables assume J unique values or categories, the relationship between the dependent variable Y and the independent variables X_i can be represented as follows:

$$\Pr[\ Y \geq J \mid X\ ] = \frac{1}{1 + exp[-(\alpha_J\ + X\beta)]}$$

The 'Odds' of being less than or equal to a particular category is thus a ratio between the cumulative probability above or equal to category J and the cumulative probability below category J:

$$Odds = \frac{P(Y \geq J)}{P(Y < J)}$$

Odds values greater than one indicate a higher probability of moving up or staying in the same depression category, whereas odds values less than one indicate a higher probability of moving down a category. Importantly, in this model, J can be interpreted as an intercept. It's crucial to emphasize, both here and later, an assumption that J and X do not interact with each other. Consequently, coefficients represent the marginal impact of X on the probability or odds of moving up a rank, regardless of the specific category J. Proportional odds assumes that the marginal impact of X remains the same when moving between levels, ensuring that the effect of X on moving from one level to the next is consistent. Therefore, the results should be interpreted as the odds of moving up or down a category in Y when there is a one-unit, binary, or categorical change in X, with all other factors held constant. Proportional odds also assumes the use of any intercept α_j beyond predictive tasks is ubiquitous. The validity of the proportional odds assumption will be tested later on.

For the PHQDEP and PHQINT tests, the ordinal values in this model consist of four ranks for J, while for AEFFORT, AHOPELESS, ARESTLESS, ASAD, and ARESTLESS predictions, there are five ranks for J. These differences arise from the nature of each question, with the former having four answer choices and the latter having five.

3.1 Computational and Methodological limitations

Contrary to economic suggestions, income is not logged in this model to illustrate diminishing returns. a consequence of zeros existing within the dataset. Retaining these values was deemed more crucial for this analysis than removing them for computation. Given the frequency of survey takers reporting zero dollars of income, such a decision would introduce significant omitted variable

bias. Alternatively, income variables can be put in quadratic form; however, this made analysis impossible, as non-real numbers were generated by the R package, MASS, when regressing the model. This is a shortcoming of how the model is interpreted in R, which struggles with data possessing both extremely large and small values. Aside from interpretation purposes, this is the main reason why income totals were divided by 1000. That being said, there has already been work done looking at a phenomenon of diminishing returns to income [7].

It is not believed, however, that this limitation disrupts the paper's conclusion: a statistically significant relationship between socioeconomic factors and mental health conditions.

4 Results

Table 1:

	PHQDEP	AEFFORT	AHOPELESS	ARESTLESS	ASAD	AWORTHLESS	PHQINTR
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income Total	0.988***	0.993***	0.987***	0.996***	0.985***	0.986***	0.987***
	(0.0002)	(0.0002)	(0.0002)	(0.0001)	(0.0002)	(0.0003)	(0.0002)
Other Controls	No	No	No	No	No	No	No
N	185,023	184,665	184,900	184,688	184,868	184,693	185,423

Notes:
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

[7] see Akhtar - Danesh et al

Table 2:

	PHQDEP	PHQINTR	AEFFORT	AHOPELESS	ARESTLESS	ASAD	AWORTHLESS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income Total	0.998***	0.997***	0.999***	0.997***	1.000***	0.997***	0.995***
	(0.0004)	(0.0004)	(0.0003)	(0.0004)	(0.0003)	(0.0004)	(0.0004)
Family Income Total	0.998***	0.999***	0.998***	0.997***	0.999***	0.998***	0.998***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Female	1.137***	0.973***	0.975***	0.958***	0.931***	1.117***	0.899***
	(0.014)	(0.014)	(0.012)	(0.014)	(0.011)	(0.015)	(0.016)
Divorced	1.274***	1.247***	1.161***	1.262***	1.100***	1.311***	1.210***
	(0.022)	(0.022)	(0.018)	(0.022)	(0.018)	(0.023)	(0.025)
Poor Health	10.560***	12.759***	11.932***	8.159***	5.462***	7.548***	8.425***
	(0.040)	(0.039)	(0.035)	(0.039)	(0.035)	(0.039)	(0.042)
Bachelors Degree	1.050***	0.947***	1.092***	0.910***	1.185***	0.725***	0.869***
	(0.069)	(0.066)	(0.058)	(0.062)	(0.057)	(0.065)	(0.073)
Currently Smokes	1.424***	1.345***	1.222***	1.331***	1.384***	1.386***	1.328***
	(0.018)	(0.017)	(0.015)	(0.017)	(0.015)	(0.018)	(0.019)
Ever Diagnosed With Cancer	0.978***	1.120***	1.153***	0.922***	0.967***	0.925***	0.961***
	(0.023)	(0.022)	(0.019)	(0.023)	(0.019)	(0.024)	(0.026)
Nervous Some of the Time	4.265***	3.624***	4.426***	6.173***	6.419***	4.705***	5.012***
	(0.018)	(0.017)	(0.013)	(0.018)	(0.013)	(0.019)	(0.022)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	159,988	160,189	159,867	160,163	159,897	160,067	159,949

Notes:

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 3:

	PHQDEP	PHQINTR	AEFFORT	AHOPELESS	ARESTLESS	ASAD	AWORTHLESS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income Total	0.997***	0.996***	0.999***	0.998***	1.000***	0.997***	0.995***
	(0.001)	(0.001)	(0.0005)	(0.001)	(0.0004)	(0.001)	(0.001)
Family Income Total	0.998***	0.999***	0.998***	0.997***	0.999***	0.998***	0.998***
	(0.0003)	(0.0003)	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0003)
Female	1.131***	0.958***	0.976***	0.962***	0.930***	1.111***	0.891***
	(0.015)	(0.014)	(0.012)	(0.015)	(0.012)	(0.015)	(0.016)
Divorced	1.286***	1.265***	1.163***	1.267***	1.097***	1.318***	1.228***
	(0.023)	(0.022)	(0.019)	(0.022)	(0.018)	(0.023)	(0.025)
Poor Health	10.549***	12.718***	12.103***	8.268***	5.479***	7.716***	8.373***
	(0.040)	(0.040)	(0.035)	(0.039)	(0.035)	(0.040)	(0.043)
Bachelors Degree	1.019***	0.962***	1.080***	0.909***	1.189***	0.723***	0.862***
	(0.071)	(0.069)	(0.059)	(0.064)	(0.058)	(0.066)	(0.075)
Currently Smokes	1.428***	1.351***	1.227***	1.336***	1.384***	1.388***	1.334***
	(0.018)	(0.017)	(0.015)	(0.018)	(0.015)	(0.018)	(0.019)
Ever Diagnosed with Cancer	0.974***	1.116***	1.152***	0.925***	0.969***	0.921***	0.964***
	(0.024)	(0.023)	(0.019)	(0.024)	(0.019)	(0.025)	(0.026)
Nervous Some of the Time	4.271***	3.646***	4.459***	6.174***	6.459***	4.735***	5.077***
	(0.018)	(0.017)	(0.014)	(0.018)	(0.013)	(0.019)	(0.022)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	157,613	157,613	157,613	157,613	157,613	157,613	157,613

Notes:

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

Table 4:

	PHQDEP	PHQINTR	AEFFORT	AHOPELESS	ARESTLESS	ASAD	AWORTHLESS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Income Total	0.996***	0.992***	0.992***	1.003***	0.999***	1.000***	0.992***
	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)	(0.004)
Family Income Total	0.999***	1.003***	1.005***	0.992***	1.001***	0.996***	1.001***
	(0.003)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)	(0.003)
Female	1.132***	0.950***	0.945***	0.989***	0.924***	1.132***	0.889***
	(0.021)	(0.021)	(0.017)	(0.021)	(0.016)	(0.023)	(0.024)
Divorced	1.308***	1.352***	1.306***	1.151***	1.123***	1.253***	1.277***
	(0.060)	(0.058)	(0.047)	(0.061)	(0.045)	(0.065)	(0.067)
Poor Health	10.694***	13.149***	12.557***	8.043***	5.522***	7.645***	8.593***
	(0.044)	(0.043)	(0.038)	(0.043)	(0.037)	(0.044)	(0.047)
Bachelors Degree	1.001***	0.904***	1.005***	0.973***	1.172***	0.745***	0.829***
	(0.080)	(0.078)	(0.065)	(0.074)	(0.064)	(0.077)	(0.085)
Currently Smokes	1.436***	1.375***	1.260***	1.309***	1.391***	1.374***	1.351***
	(0.022)	(0.021)	(0.018)	(0.022)	(0.017)	(0.023)	(0.024)
Ever Diagnosed with Cancer	0.973***	1.109***	1.138***	0.934***	0.967***	0.926***	0.960***
	(0.024)	(0.023)	(0.020)	(0.024)	(0.020)	(0.025)	(0.027)
Nervous Some of the Time	4.264***	3.630***	4.425***	6.202***	6.448***	4.746***	5.061***
	(0.019)	(0.018)	(0.014)	(0.019)	(0.013)	(0.019)	(0.023)
Other Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	157,613	157,613	157,613	157,613	157,613	157,613	157,613

Notes:

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

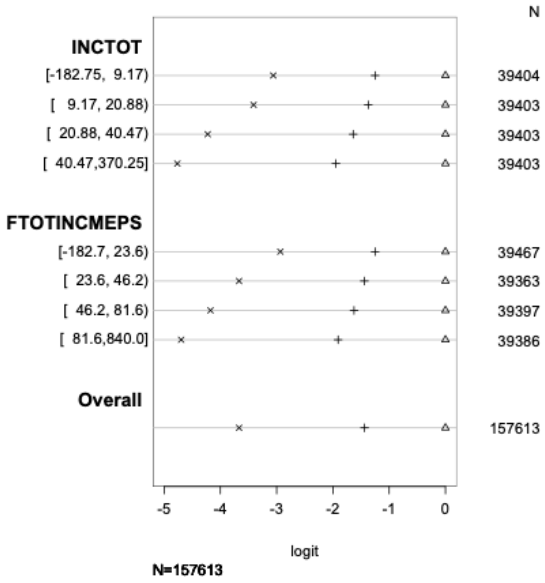
Table 1 demonstrates a fitted model for personal income total without any controls. Here, a statistically significant relationship is observed, indicating that increases in an individual's income make it more likely for the individual to be placed at a lower rank in all tests. In Tables 2, 3, and 4, a statistically significant relationship can be observed between income levels and depression when controlling for omitted variable bias. In the latter two, causal directionality can also be hypothesized. All data points show an extremely strong statistical significance. A value less than one indicates that a per unit change in x would make it more likely to move down a rank [0,1,2,3]. Considering that lower survey scores indicate less severe depressive symptoms, it can be interpreted that coefficient values less than one contribute to individuals being less depressed, and vice versa for coefficient values greater than one. Most results here suggest a negative relationship between income and mental health symptoms. As income decreases, all else held constant, the likelihood of an individual moving down a category in each test increases. This can also be phrased as: a randomly selected individual with a higher income is more likely to be less depressed than a randomly selected individual with a relatively lower income. Notably for personal income, feelings of worthlessness saw the greatest increases in odds with corresponding decreases in income. The lack of impact of income on column five could result from similarities between the restlessness variable and the anxiety variable, which is controlled for there, as individuals who are restless are also very likely anxious and vice versa. However, correlation tests between anxiety and other variables showed no larger correlations than they had with restlessness. The instrument that had the most impact on results was that of capital gains, which for multiple test categories displays a positive relationship. An alternative way to view this relationship can be seen in the appendix in Figures 4 and 5. Figure 5 demonstrates the predicted change in the likelihood of being placed in a respective PHQDEP category, all else constant. More clearly, the graph demonstrates a 100,000 dollar increase in income causing about a 1 percent higher likelihood of being in PHQDEP category zero; for an individual answering that in the past two weeks, they have been feeling down, depressed, or hopeless, 'Not at all'. Results here do seem to support Zimmerman et al.'s claim that initially, a relationship does seem to exist between income and depression; however, once controlling for omitted variables, this relationship becomes less distinct. However, there is still an impactful relationship between income and depression shown here, one which shouldn't be discredited.

4.1 Proportional Odds

The key assumption for an ordinal logistic proportional odds model is, expectedly, proportional odds. This implies that the odds of moving from one category to another are the same across all ranks. Essentially, it assumes that the independent variables X are independent of J, and that an interaction between X and Y would equate to nothing. Although there is no universal method to test this assumption, Harrel outlines certain strategies to analyze it. Figure 3 illustrates a proportional odds test.

Figure 3

Proportional odds



for reference, INCTOT represents an individual's total annual income in thousands of dollars, while FTOTINCMEPS represents a family's total annual income in thousands of dollars. The figure compares multiple binary logistic regression coefficients at different cutoff points for income values. For example, the top row of values indicates that the logit for total income values greater than -182.75 and less than 9.17 is approximately negative one. The triangle points are used to normalize the coefficients to a similar value, which is zero. The crosses illustrate odds ratios for having a PHQDEP score of 2, while the X's illustrate odds ratios for having a PHQDEP score of 3. A model that perfectly adheres to the proportional odds assumptions would have points all maintaining the same odds value, such as the triangles.

From Figure 3, we can derive that proportional odds for this model are relatively strong for PHQDEP scores of 2. However, proportional odds are not as strong for PHQDEP scores of 3. Given this, the model's ability to predict higher levels of PHQDEP depression can be presumed to have less accuracy.

5 Discussion

The analysis conducted in this paper demonstrates a statistically significant relationship between income and mental health symptoms when controlling for both biological and socioeconomic factors. Moreover, the use of instrumental variables in this analysis strengthens such conclusions. These findings align with

prior research in this area. Given the wide range of costs associated with mental health problems and the established causal relationships between income and mental health problems, there is substantial justification for this work in both the public and private sectors. As argued earlier, a government should strive to maximize society's happiness and life satisfaction. This work sheds light on components involved in happiness and well-being, thus guiding the decisions made by public officials.

That being said, there are limitations to this research. Firstly, there is the issue of external validity. The study exclusively focuses on the US population. While the work by Layte. R. justifies a degree of external validity applicable to Western nations, there is little to no proof of external validity in other regions of the world. Secondly, survey data is susceptible to various problems. Historically in America, mental health has been stigmatized for some time, likely causing people to answer mental health questions untruthfully. Additionally, not everyone assigns identical numerical scores to the same mental health severity; two individuals experiencing the same degree of mental trauma may assign different rankings to it. Any systemic problems in how US individuals answer survey data are not captured in this analysis, as there is no universally agreed upon control for such issue. Despite using factors such as sex to attempt to control for these problems, there likely exist immeasurable factors [8]. Thirdly, testing of proportional odds assumptions raises concerns about the validity of coefficients in this analysis. However, Harrell posits that these assumptions can be somewhat relaxed, but there is still no common consensus on to what extent a violation here invalidates results. Nonetheless, the primary understanding to be derived here is an existing relationship between income and depression.

In conclusion, it is evident, as proven here, that income is heavily correlated with variables that play a significant role in mental health problems while also being a determinant itself.

6 Works Cited

Chang, S. M., Hong, J. P., Cho, M. J. (2012). Economic burden of depression in South Korea. Social Psychiatry and Psychiatric Epidemiology, 47, 683–689.

Lombardo, P., Jones, W., Wang, L., Shen, X., Goldner, M. E. (2018). The fundamental association between mental health and life satisfaction: results from successive waves of a Canadian national survey. BMC Public Health, 18.

Akhtar-Danesh, N., Landeen, J. (2007). Relation between depression and sociodemographic factors. International Journal of Mental Health Systems.

Brodaty, H., Peters, K., Boyce, P., Hickie, I., Parker, G., Mitchell, P., Wilhelm, K. (1997). Age and depression. Journal of Affective Disorders,.

Cénat, J. M., Kogan, C., Noorishad, P. G., Hajizadeh, S., Dalexis, R. D., Ndengeyingoma, A., Guerrier, M. (2021). Prevalence and correlates of depression among Black individuals in Canada: The major role of everyday racial discrimination. Depression and Anxiety.

[8] This problem becomes less substantial if entity fixed effects over time were used

Dijkstra-Kersten, S. M., Biesheuvel-Leliefeld, K. E., van der Wouden, J. C., Penninx, B. W., van Marwijk, H. W. (2015). Associations of financial strain and income with depressive and anxiety disorders. J Epidemiol Community Health.

Dobson, K. S. (1985). The relationship between anxiety and depression. Clinical Psychology Review.

Druss, B. G., Rosenheck, R. A., Sledge, W. H. (2000). Health and disability costs of depressive illness in a major US corporation. American Journal of Psychiatry.

Ettner, S. L. (1996). New evidence on the relationship between income and health. Journal of Health Economics.

Fluharty, M., Taylor, A. E., Grabski, M., Munafò, M. R. (2016). The association of cigarette smoking with depression and anxiety: a systematic review. Nicotine and Tobacco Research.

Galobardes, B., Shaw, M., Lawlor, D. A., Lynch, J. W., Davey Smith, G. (2006). Indicators of socioeconomic position (part 1). Journal of Epidemiology and Community Health.

Galobardes, B., Shaw, M., Lawlor, D. A., Lynch, J. W. (2006). Indicators of socioeconomic position (part 2). Journal of Epidemiology and Community Health.

Harrell, F. E. (2001). Regression Modeling Strategies: with Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer.

Helliwell, J., Layard, R., Sachs, J. (2012). World Happiness Report.

Hu, T. W., He, Y., Zhang, M., Chen, N. (2007). Economic costs of depression in China. Social Psychiatry and Psychiatric Epidemiology, 42, 110–116.

Kessler, R. C. (2012). The costs of depression. Psychiatric Clinics, 35(1), 1–14.

Konig, H., Konig, H. H., Konnopka, A. (2020). The excess costs of depression: a systematic review and meta-analysis. Epidemiology and Psychiatric Sciences, 29, e30.

Layte, R. (2012). The association between income inequality and mental health: testing status anxiety, social capital, and neo-materialist explanations. European Sociological Review, 28(4), 498–511.

LaVeist, T. A. (2005). Disentangling race and socioeconomic status: a key to understanding health inequalities. Journal of Urban Health, 82, iii26–iii34.

Mendelsohn, C. (2012). Smoking and depression: a review. Australian Family Physician, 41(5), 304–307.

Mirowsky, J., Ross, C. E. (1992). Age and depression. Journal of Health and Social Behavior, 187–205.

Miech, R. A., Caspi, A., Moffitt, T. E., Wright, B. R. (1999). Low socioeconomic status and mental disorders: a longitudinal study of selection and causation during young adulthood. American Journal of Sociology, 104, 1096–1131.

Paperwalla, K. N., Levin, T. T., Weiner, J., Saravay, S. M. (2004). Smoking and depression. Medical Clinics, 88(6), 1483–1494.

Rosenquist, Fowler, J. H., Christakis, N. A. (2011). Social network determinants of depression. Molecular Psychiatry, 16(3), 273–281. doi:10.1038/mp.2010.13

Schomerus, G., Schwahn, C., Holzinger, A., Corrigan, P. W., Grabe, H. J., Carta, M. G., Angermeyer, M. C. (2012). Evolution of public attitudes about mental illness: A systematic review and meta-analysis. Acta Psychiatrica Scandinavica, 125(6), 440–452.

Van Ryn, M., Burke, J. (2000). The effect of patient race and socio-economic status on physicians' perceptions of patients. Social Science Medicine, 50(6), 813–828.

Wandersman, A., Nation, M. (1998). Urban neighborhoods and mental health: Psychological contributions to understanding toxicity, resilience, and interventions. American Psychologist, 53(6), 647.

Zimmerman, F. J., Katon, W. (2005). Socioeconomic status, depression disparities, and financial strain: what lies behind the income-depression relationship? Health Economics, 14(12), 1197–1215.

Altemus, M., Sarvaiya, N., Epperson, C. N. (2014). Sex differences in anxiety and depression clinical perspectives. Frontiers in Neuroendocrinology, 35(3), 320–330.

Van Droogenbroeck, F., Spruyt, B., Keppens, G. (2018). Gender differences in mental health problems among adolescents and the role of social support: results from the Belgian health interview surveys 2008 and 2013. BMC Psychiatry.

Spiegel, D., Giese-Davis, J. (2003). Depression and cancer: mechanisms and disease progression. Biological Psychiatry.

Dooley, D., Catalano, R., Wilson, G. (1994). Depression and unemployment: panel findings from the Epidemiologic Catchment Area study. American Journal of Community Psychology.

Lerner, D., Adler, D. A., Chang, H., Lapitsky, L., Hood, M. Y., Perissinotto, C. (2004). Unemployment, job retention, and productivity loss among employees with depression. Psychiatric Services.

Jefferis, B. J., Nazareth, I., Marston, L., Moreno-Kustner, B., Bellón, J. Á., Svab, I. (2011). Associations between unemployment and major depressive disorder: evidence from an international, prospective study (the predict cohort). Social Science Medicine.

7 Appendix

A Summary Statistics

Table 1
Summary Statistics

X	N	Mean	SD	Median	Mad	Min	Max	SE
YEAR	157613	4.4323945	2.2539586	4.0000000	2.9652000	1.00000	8.00000	0.0056774
SEX	157613	1.5394796	0.4984405	2.0000000	0.0000000	1.00000	2.00000	0.0012555
MARSTAT	157613	2.5459829	1.7494475	1.0000000	0.0000000	1.00000	5.00000	0.0044066
REGIONMEPS	157613	3.7614600	1.0198603	4.0000000	1.4826000	1.00000	5.00000	0.0025689
FAMSIZE	157613	3.0198651	1.6677545	3.0000000	1.4826000	1.00000	15.00000	0.0042008
RACEA	157613	1.5568893	1.0952040	1.0000000	0.0000000	1.00000	6.00000	0.0027587
INTERVLANG	157613	1.2025341	0.5136434	1.0000000	0.0000000	1.00000	4.00000	0.0012938
EDUC	157613	18.9814229	5.2517030	20.0000000	2.9652000	1.00000	28.00000	0.0132283
FTOTINCMEPS	157613	61.0265633	53.7082403	46.2060000	38.9864496	-182.74620	840.01738	0.1352835
INCTOT	157613	30.1233758	32.4744987	20.8840000	20.6327764	-182.74620	370.25179	0.0817987
HEALTH	157613	2.4662433	1.0234220	2.0000000	1.4826000	1.00000	5.00000	0.0025779
CANCEREV	157613	1.0879813	0.2832686	1.0000000	0.0000000	1.00000	2.00000	0.0007135
SMOKENOW*	157613	1.1619790	0.3684327	1.0000000	0.0000000	1.00000	2.00000	0.0009280
AEFFORT	157613	1.6478527	1.0132856	1.0000000	0.0000000	1.00000	5.00000	0.0025523
AHOPELESS	157613	1.4009695	0.8108780	1.0000000	0.0000000	1.00000	5.00000	0.0020425
ANERVOUS	157613	1.7254351	0.9166298	1.0000000	0.0000000	1.00000	5.00000	0.0023089
ARESTLESS	157613	1.7231003	0.9565274	1.0000000	0.0000000	1.00000	5.00000	0.0024094
ASAD	157613	1.3414566	0.7523006	1.0000000	0.0000000	1.00000	5.00000	0.0018949
AWORTHLESS	157613	1.2919112	0.7432003	1.0000000	0.0000000	1.00000	5.00000	0.0018720
PHQINTR	157613	1.3448256	0.7050404	1.0000000	0.0000000	1.00000	4.00000	0.0017759
PHQDEP	157613	1.3196691	0.6699231	1.0000000	0.0000000	1.00000	4.00000	0.0016874
FRACFAMINC	157613	0.5676183	0.3804846	0.5090662	0.5043148	-15.07143	41.81818	0.0009584
CAPGAIN	157613	1.1755656	7.8477206	0.0000000	0.0000000	-436.08511	577.44467	0.0197673

B Variable Key

Variable.Name	Variable.Description	Variable.type.or.category
YEAR	Year	2010 - 2018
SEX	Biological Sex	Male|Female
MARSTAT	Mariage Status	Married|Widowed|Divorced|Separate|Never Maried
REGIONMEPS	Census Reigon	Northeast|Midwest|South|West
FAMSIZE	Family Size	Count
RACEA	Race	White|Black/African-American|Alaskan Native or American Indian|Asian|Pacific Islander|Multiple Races
INTERVLANG	Interview Language	English|Spanish|English and Spanish|Other
EDUC	Educational Atainment	Various categories such as "Bachelors Degree" or "GED"
FTOTINCMEPS	Total Family Income	Count/1000 * CPI Factor(2009)
INCTOT	Total Personal Income	Count/1000 * CPI Factor(2009)
HEALTH	Health Status	Excelent|Very Good|Good|Fair|Poor
CANCEREV	Ever told had cancer	No|Yes
SMOKENOW	Smokes cigarettes now	No|Yes
AEFFORT	Felt Everything an Effort, Past 30 days	None of the time|A little of the time|Some of the time|Most of the time|All of the time
AHOPELESS	How Often Felt Hopeless, Past 30 days	None of the time|A little of the time|Some of the time|Most of the time|All of the time
ANERVOUS	How Often Felt Nervous, Past 30 days	None of the time|A little of the time|Some of the time|Most of the time|All of the time
ARESTLESS	How Often Felt Restless, Past 30 days	None of the time|A little of the time|Some of the time|Most of the time|All of the time
ASAD	How Often Felt Sad, Past 30 days	None of the time|A little of the time|Some of the time|Most of the time|All of the time
AWORTHLESS	How Often Felt Worthless, Past 30 days	None of the time|A little of the time|Some of the time|Most of the time|All of the time
PHQINTR	Little Interest in Doing Things: Past Two Weeks	Not at all|Several Days|More than half the days|Nearly every day
PHQDEP	Feeling Down, Depressed, or Hopeless: Past Two Weeks	Not at all|Several Days|More than half the days|Nearly every day
FRACFAMINC	Fraction of Family Income made by Individual	Percentage
CAPGAINS	Summation of Individuals Captial Asset Annual Returns	Count

C Survey Questions

10.
The following questions ask about how you have been feeling during **the past 30 days**. For each question, please mark the box that best describes how often you had this feeling.
During the past 30 days, about how often did you feel...

a. nervous?

() All of the time
() Most of the time
() Some of the time
() A little of the time
() None of the time

b. hopeless?

() All of the time
() Most of the time
() Some of the time
() A little of the time
() None of the time

c. restless or fidgety?

() All of the time
() Most of the time
() Some of the time
() A little of the time
() None of the time

d. so sad that nothing could cheer you up?

() All of the time
() Most of the time
() Some of the time
() A little of the time
() None of the time

e. that everything was an effort?

() All of the time
() Most of the time
() Some of the time
() A little of the time
() None of the time

f. worthless?

() All of the time
() Most of the time
() Some of the time
() A little of the time
() None of the time

11.
The following two questions ask about how you have been feeling in the **past 2 weeks.**
Over the last 2 weeks, how often have you been bothered by any of the following problems?

Nearly every day
More than half the days
Several days
Not at all

a. Little interest or pleasure in doing things...

()
()
()
()

b. Feeling down, depressed, or hopeless...

()
()
()
()

D Change in Probabilities

Figure 4

Change in probability

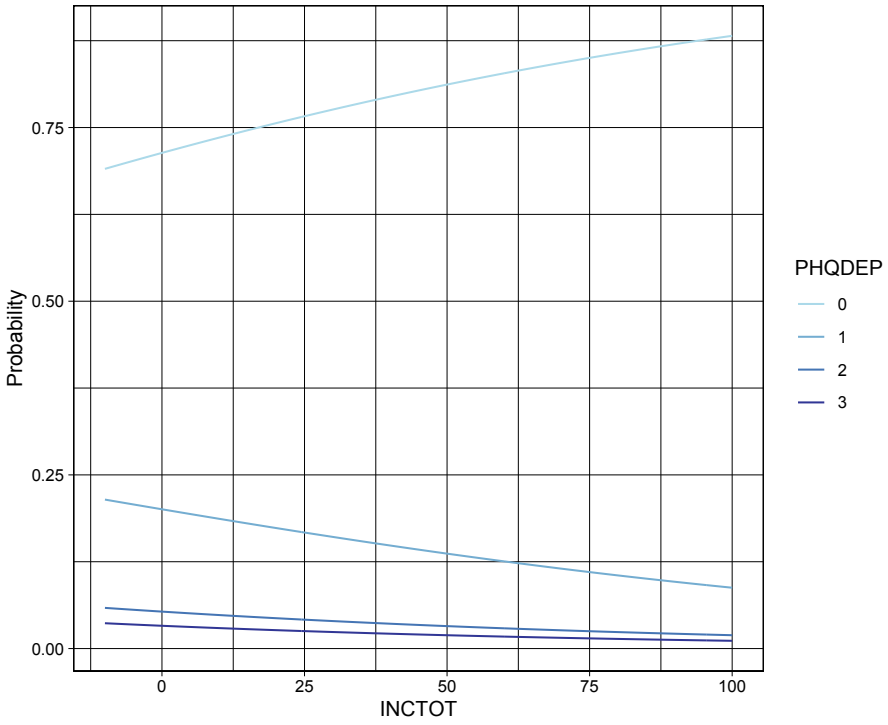


Figure 5

Change in probability all else constant