Navi Chawla

Professor Holmes

Econometrics

4 December 2025

Determinants of Happiness in the United States

Abstract, references, and appendix not included in total page count

Total page count: 10

Abstract

This paper investigates the socioeconomic, demographic, and health-related determinants of subjective happiness among adults in the United States using the 2024 General Social Survey (GSS). The GSS measures happiness in three categories—very happy, pretty happy, and not too happy—I transform this variable into a binary measure equal to one if a respondent is very or pretty happy and zero if the respondent is not too happy. This binary outcome allows estimation of a logistic regression model, which serves as the main specification. A linear regression using a 1–3 scale is included as a robustness check. The analysis examines the roles of income, self-rated health, marital status, education, and demographic characteristics such as age, gender, race, and region. The analytic sample consists of 2,808 respondents after listwise deletion of missing values. Results show that self-rated health and marital status are the strongest predictors of happiness, while income has weak and inconsistent effects and education is not statistically significant. The OLS robustness check confirms these patterns. I discuss conceptual limitations

related to omitted variables, measurement error, and reverse causality. The paper concludes with reflections on policy implications and future research directions.

Introduction

As economists increasingly recognize that material well-being does not fully capture human welfare, subjective well-being has emerged as a crucial complement to traditional economic indicators. Conventional metrics such as income, consumption, and employment describe external economic circumstances but fail to reflect how individuals experience their lives. In contrast, measures of happiness and life satisfaction capture internal evaluations of well-being that may diverge significantly from objective economic conditions. A society may experience economic growth without corresponding increases in subjective well-being, and individuals with similar incomes may report vastly different levels of happiness.

The United States offers a compelling context for studying happiness. Although it is one of the wealthiest nations in terms of per capita income, levels of self-reported happiness have remained relatively stable or even declined slightly in recent decades, with average U.S. life satisfaction fluctuating only between roughly 6.8 and 7.2 since the mid-2000s (Sachs et al., 2024). Understanding what predicts happiness is therefore important for interpreting the broader social welfare landscape beyond economic output.

In this paper, I investigate the determinants of subjective happiness in the United States using the 2024 General Social Survey (GSS). A key advantage of the GSS is its consistent and longstanding happiness measure, which categorizes respondents as “very happy,” “pretty happy,” or “not too happy.” Although this measure is ordinal, ordered logit requires the proportional-odds assumption, so I use a binary specification. I transform the outcome into a binary measure equal to one if the respondent is very or pretty happy and zero otherwise. This allows me to estimate a binary logistic regression (Model #3).

The research question guiding this paper is: Which factors—among income, health, marital status, education, and demographics—are most strongly associated with happiness among U.S. adults? Following conventional practice, I include controls for age, gender, race, and region. To ensure robustness, I also estimate a linear regression using a 1–3 happiness scale.

The remainder of the paper proceeds as follows. Section 2 provides institutional background on the GSS and reviews the academic literature on happiness. Section 3 describes the data and variable construction. Section 4 outlines the econometric model and its assumptions. Section 5 presents the results. Section 6 discusses the findings and their implications. Section 7 concludes.

Literature Review

The General Social Survey (GSS) is one of the most widely used data sources in U.S. social science research and has been fielded on a nearly biennial basis since 1972. The GSS's consistency in questionnaire design, sampling methodology, and fieldwork makes it a gold standard for studying social trends. One of the longest-running items on the GSS is its simple but powerful happiness question: "Taken all together, how would you say things are these days—would you say that you are very happy, pretty happy, or not too happy?" Because the wording has remained virtually unchanged for more than fifty years, the measure is highly reliable across time and across demographic groups (General Social Survey, 2024). The GSS ensures national representation through probability sampling and applies careful weighting to reflect the U.S. population.

The scholarly literature identifies several major determinants of subjective happiness, which help guide expectations for the empirical analysis. A consistent finding is that self-rated health is one of the strongest predictors of happiness. Individuals who report poor or fair health are substantially less likely to report high levels of well-being, and this association persists after controlling for socioeconomic status and demographics. Diener et al. (1999) show that both physical and mental health contribute significantly to subjective well-being. The 2024 World Happiness Report similarly notes that health indicators are more predictive of well-being than income in most countries, highlighting how chronic illness, disability, and mental health challenges reduce subjective happiness (Sachs et al., 2024).

Marital status emerges as another central determinant of well-being in the literature. Married individuals tend to report higher happiness levels than those who are widowed, divorced, separated, or never married. Helliwell and Putnam (2004) argue that marriage provides emotional support, companionship, and social capital, all of which contribute to higher subjective well-being. The emotional benefits of marriage are particularly salient during stressful life events, where relational support serves as a buffer against psychological distress. Conversely, widowhood and divorce often lead to substantial declines in happiness (Lucas et al., 2003).

Income's relationship with happiness is more nuanced. Kahneman and Deaton (2010) distinguish between life evaluation—a cognitive assessment of one's life—and emotional well-being, which reflects daily affect. Their research finds that emotional well-being rises with income only up to a certain threshold, beyond which additional income does not significantly increase day-to-day happiness. Clark et al. (2008) further suggest that relative income, or income compared to one's peers, may be a stronger predictor of well-being than absolute income. This perspective helps explain the "Easterlin Paradox," which observes that large increases in national income do not necessarily translate into proportional increases in national happiness levels.

Education exhibits mixed effects in the literature. While greater education expands employment opportunities and provides social and cognitive benefits, its direct association with happiness is often small or statistically insignificant once income is controlled for. Powdthavee (2010) suggests that education may raise aspirations, leading individuals to evaluate their lives relative to higher expectations, which may offset potential well-being gains. De Neve et al. (2012) emphasize that personality traits and genetic predispositions shape both educational attainment and subjective well-being, complicating any direct causal interpretation.

Together, these findings suggest that health and marital status should have the strongest associations with happiness, income should have modest or diminishing effects, and education should have limited direct influence. The GSS's rigorous design and longstanding use of the happiness measure make it an ideal dataset to test these expectations.

Data

This analysis uses the 2024 GSS Cross-Section, Release 1. The original dataset includes 3,309 respondents. To construct the analytic sample, I remove observations with missing values for the dependent variable or any independent variable in the model. After listwise deletion, the final analytic sample consists of 2,808 adults, representing approximately 85 percent of the original dataset.

The dependent variable, happiness, is originally coded on a three-point scale: very happy (1), pretty happy (2), and not too happy (3). I dichotomize the variable so that respondents who are very happy or pretty happy receive a value of one and those who are not too happy receive a zero. This creates a binary indicator, *happy_bin*, suitable for estimation using logistic regression. In the analytic sample, 78.51 percent of respondents are coded as happy, and 21.49 percent are not.

To support the robustness check using a linear regression, I create the variable happy3, which assigns values of 3, 2, and 1 corresponding to the original categories of very happy, pretty happy, and not too happy. The mean of happy3 is approximately 1.996, indicating that the average respondent lies close to the “pretty happy” category.

Income is measured across twelve categorical brackets, ranging from under $1,000 to $25,000 or more. These brackets reflect annual family income. These brackets are unevenly spaced and do not map cleanly onto absolute purchasing power, so I include them as a set of dummy variables rather than imposing a linear specification. Approximately 77.64 percent of respondents fall into the highest income category (“$25,000 or more”), reflecting the broadness of GSS income brackets. It is important to note that because the top category is open-ended (“$25,000 or more”), a large share of respondents fall into it, creating a skewed distribution and limiting observable variation in higher income levels.

Self-rated health is measured using four categories: excellent, good, fair, and poor. In the analytic sample, 17.70 percent report excellent health, 53.49 percent report good health, 24.68 percent report fair health, and 4.13 percent report poor health. These proportions are consistent with other U.S. national surveys.

Marital status includes five categories: married (41.38 percent), widowed (7.02 percent), divorced (16.27 percent), separated (3.28 percent), and never married (32.05 percent).

Demographic controls include education (mean of 14.31 years), age (mean of 50.09 years), gender (54.42 percent female), race (70.69 percent White, 17.77 percent Black, 11.54 percent other), and region (based on the nine U.S. Census regions).

Descriptive analyses highlight clear patterns. Happiness is substantially higher among individuals who are married or in better health. By contrast, those in poor health or who are separated or divorced report significantly lower happiness. Income shows only mild bivariate associations with happiness, suggesting the importance of multivariate analysis.

Econometric Model

The main econometric model is a binary logistic regression, which estimates the probability that a given respondent is happy. The model takes the form:

$$Pr(happy\_bin = 1|X) = \Lambda(\beta_0 + \beta_1 Income + \beta_2 Health + \beta_3 Marital + \beta_4 Education + \gamma Controls)$$

where Λ represents the logistic cumulative distribution function. Income, health, marital status, gender, race, and region are included as sets of indicator variables, while age and education enter as continuous variables. Standard errors are estimated robustly to account for heteroskedasticity.

Although the outcome is ordinal, ordered logit relies on the proportional-odds assumption. To avoid relying on this assumption, I instead estimate a binary logistic regression model. The binary logit model provides straightforward interpretation and avoids reliance on the proportional-odds assumption.

To assess robustness, I also estimate an OLS model using the linearized happiness variable. Although OLS imposes cardinal assumptions on an ordinal outcome, it provides a useful check: if effects are similar across logit and OLS, results can be interpreted with greater confidence.

Identification concerns are substantial. Happiness is influenced by numerous unobserved factors, including personality traits (De Neve et al., 2012), optimism, and mental health. These unobservables may also affect predictors such as health, marriage, and income, creating omitted variable bias. Reverse causality is also likely; happier individuals may take better care of themselves, be more successful in relationships, and perform better at work. Because the GSS is cross-sectional, it is impossible to resolve these issues. For this reason, I interpret all coefficients as conditional associations rather than causal effects.

Results

The logistic regression model is estimated using all 2,808 respondents with complete data. As shown in Table 1.1, the pseudo R^2 is 0.0802, and the Wald chi-squared statistic indicates that the model is jointly significant at $p < 0.001$. The corresponding coefficient table for key predictors is reproduced in Table 1.2.

Self-rated health is the strongest predictor of happiness. The average marginal effects reported in Table 2.1 show that, relative to individuals in excellent health, those in good health exhibit a small and statistically insignificant change in happiness probability (−1.6 percentage points). In contrast, respondents in fair health are 12.9 percentage points less likely to be happy ($p < 0.001$), and those in poor health are 31.1 percentage points less likely to be happy ($p < 0.001$). These patterns, summarized in Table 2.2, illustrate the steep and monotonic gradient between declining health and subjective well-being.

Marital status is the second strongest determinant. As shown in Table 2.2, widowed individuals are 10.9 percentage points less likely to be happy compared to married respondents. Divorced individuals are 12.3 percentage points less likely to be happy, separated individuals 19.0 percentage points less likely, and never-married individuals 13.3 percentage points less likely. All effects are large in magnitude and highly statistically significant.

Income is surprisingly weak as a predictor. Most income brackets show no statistically significant association with happiness after controlling for health and marital status. Only the highest income category ("$25,000 or more") shows marginal significance—a 10.7 percentage-point increase in the probability of being happy ($p \approx 0.07$). This pattern is visible in the coefficients reported in Table 1.2 and the marginal effects in Table 2.2.

Education and age are not statistically significant. Gender exhibits a marginally significant but substantively small effect: women are about 2.7 percentage points more likely to be happy ($p \approx 0.07$), as reported in Table 2.2. Race and region do not display consistent or meaningful patterns.

The OLS robustness model, presented in Tables 3.1 and 3.2, confirms these findings. Respondents in fair and poor health score 0.389 and 0.560 points lower on the 1–3 happiness scale, respectively. Marital status coefficients remain large and significant, while income and education continue to show minimal explanatory power. The consistency between the logistic and OLS models strengthens confidence in the empirical patterns identified.

Discussion and Implications

The patterns observed in the regression tables reinforce long-standing findings in the happiness literature. As illustrated in Tables 2.1–2.2, health and marital status display the largest and most statistically robust associations with happiness, whereas income and education contribute relatively little once these core determinants are taken into account.

Similarly, the large differences in happiness by marital status align with Helliwell and Putnam's (2004) arguments regarding the social and emotional value of marriage and stable relationships. The substantial declines associated with separation and divorce are consistent with Lucas et al. (2003), who document that the emotional toll of relationship dissolution produces durable reductions in subjective well-being.

Income's limited effects in this analysis align with a long-standing observation in the literature: while income contributes to financial stability and comfort, its direct impact on day-to-day happiness is modest after a certain threshold is reached (Kahneman & Deaton, 2010). The fact that only the highest income category shows marginal significance suggests diminishing marginal utility of income and the possibility that relative income or expectations matter more than absolute levels (Clark et al., 2008).

The finding that education does not significantly predict happiness once income, health, and relationships are controlled for confirms insights from Powdthavee (2010) and De Neve et al. (2012). Education may indirectly affect well-being through better jobs or higher income, but its direct contribution may be offset by increased stress, expectations, or social comparison.

Despite these insights, the analysis remains constrained by important limitations. The GSS is cross-sectional, making it impossible to disentangle causality from correlation. Happiness may influence health, income, and marital stability as much as these factors influence happiness.

Omitted variables—such as personality traits—pose serious identification issues (De Neve et al., 2012). Measurement error in self-reported variables may also bias the estimates. For these reasons, this paper does not claim causal effects but instead provides descriptive patterns consistent with prior literature and supported by the most recent data.

Conclusion

Using data from the 2024 General Social Survey, this paper examined the determinants of happiness among U.S. adults using a binary logistic regression model and an OLS robustness check. The results reveal that self-rated health and marital status are the strongest predictors of happiness, while income and education play relatively minor roles. These findings are consistent with decades of happiness research and underscore the importance of physical well-being and relational stability in shaping subjective well-being.

Although the analysis cannot identify causal relationships due to omitted variable bias and cross-sectional data limitations, it provides robust descriptive evidence using newly released national data. Future research could employ panel data, causal identification strategies, or richer psychological measures to advance understanding of the determinants of happiness.

References

Clark, A. E., Frijters, P., & Shields, M. A. (2008). Relative income, happiness, and utility: An explanation for the Easterlin paradox. Journal of Economic Literature, 46(1), 95–144. https://doi.org/10.1257/jel.46.1.95

De Neve, J.-E., Diener, E., Tay, L., & Xuereb, C. (2012). The objective benefits of subjective well-being. Applied Psychology: Health and Well-Being, 4(1), 1–30. https://doi.org/10.1111/j.1758-0854.2011.01059.x

Diener, E., Suh, E. M., Lucas, R. E., & Smith, H. L. (1999). Subjective well-being: Three decades of progress. Psychological Bulletin, 125(2), 276–302. https://doi.org/10.1037/0033-2909.125.2.276

General Social Survey. (2024). GSS 2024 Cross-Section dataset and codebook. NORC at the University of Chicago. https://gss.norc.org

Helliwell, J. F., & Putnam, R. D. (2004). The social context of well-being. Philosophical Transactions of the Royal Society B: Biological Sciences, 359(1449), 1435–1446. https://doi.org/10.1098/rstb.2004.1522

Kahneman, D., & Deaton, A. (2010). High income improves evaluation of life but not emotional well-being. Proceedings of the National Academy of Sciences, 107(38), 16489–16493. https://doi.org/10.1073/pnas.1011492107

Lucas, R. E., Clark, A. E., Georgellis, Y., & Diener, E. (2003). Reexamining adaptation and the set point model of happiness. Journal of Personality and Social Psychology, 84(3), 527–539. https://doi.org/10.1037/0022-3514.84.3.527

Powdthavee, N. (2010). How much does money really matter? Estimating the causal effects of income on happiness. Empirical Economics, 39(1), 77–92. https://doi.org/10.1007/s00181-009-0295-5

Sachs, J., Layard, R., Helliwell, J. F., & De Neve, J.-E. (2024). World Happiness Report 2024. Sustainable Development Solutions Network. https://worldhappiness.report/

Appendix: Tables

Table 1.1

```
. logit happy_bin i.income i.health i.marital educ age i.sex i.race i.region, vce(robust)

---------------------------------------------------------------------------
 happy_bin | Coefficient  Std. Err.      z    P>|z|
----------+----------------------------------------------
 $25kplus |   0.6076      0.2922     2.08   0.038
 fair     |  -0.8169      0.1591    -5.13   0.000
 poor     |  -1.6175      0.2353    -6.87   0.000
 widowed  |  -0.7836      0.2172    -3.61   0.000
 divorced |  -0.8652      0.1482    -5.84   0.000
 nevermar |  -0.9186      0.1326    -6.93   0.000
 educ     |   0.0211      0.0189     1.12   0.264
 age      |   0.0034      0.0035     0.99   0.325
 female   |   0.1805      0.1008     1.79   0.073
---------------------------------------------------------------------------
```

Table 1.2

Table 1. Logistic Regression Predicting Happiness (n = 2,808)

Predictor	Coef	SE	z	p
$25,000 or more	0.6076	0.2922	2.08	p=0.038
Fair health	-0.8169	0.1591	-5.13	p<0.001
Poor health	-1.6175	0.2353	-6.87	p<0.001
Widowed	-0.7836	0.2172	-3.61	p<0.001
Divorced	-0.8652	0.1482	-5.84	p<0.001
Never married	-0.9186	0.1326	-6.93	p<0.001
Education	0.0211	0.0189	1.12	p=0.264
Age	0.0034	0.0035	0.99	p=0.325
Female	0.1805	0.1008	1.79	p=0.073

Table 2.1

```
. margins, dydx(*)

Average marginal effects
------------------------------------------------------------------------------
 Predictor |    dy/dx     Std. Err.      z    P>|z|
----------+-----------------------------------------------
 fair      |  -0.1292      0.0238    -5.43   0.000
 poor      |  -0.3106      0.0509    -6.10   0.000
 widowed   |  -0.1087      0.0358    -3.03   0.002
 divorced  |  -0.1236      0.0233    -5.28   0.000
 nevermar  |  -0.1332      0.0199    -6.70   0.000
 $25kplus  |   0.1071      0.0592     1.81   0.070
------------------------------------------------------------------------------
```

Table 2.2

Table 2. Average Marginal Effects from Logit Model (n = 2,808)

Predictor	AME	SE	z	p
Fair health	-0.1292	0.0238	-5.43	p<0.001
Poor health	-0.3106	0.0509	-6.10	p<0.001
Widowed	-0.1087	0.0358	-3.03	p=0.002
Divorced	-0.1236	0.0233	-5.28	p<0.001
Never married	-0.1332	0.0199	-6.70	p<0.001
$25,000 or more	0.1071	0.0592	1.81	p=0.070
Female	0.0274	0.0153	1.79	p=0.074

Table 3.1

```
. reg happy3 i.income i.health i.marital educ age i.sex i.race i.region, vce(robust)

------------------------------------------------------------------------------
  happy3  | Coefficient  Std. Err.      t    P>|t|
----------+----------------------------------------------
 fair     |  -0.3893      0.0407    -9.57   0.000
 poor     |  -0.5603      0.0673    -8.31   0.000
 widowed  |  -0.2491      0.0505    -4.94   0.000
 divorced |  -0.2608      0.0438    -5.95   0.000
 nevermar |  -0.2830      0.0309    -9.13   0.000
 $25kplus |   0.0734      0.0971     0.76   0.450
 female   |   0.0457      0.0236     1.94   0.053
------------------------------------------------------------------------------
```

Table 3.2

```
Table 3. OLS Robustness Model Predicting Happiness (1-3 scale; n = 2,808)

Predictor                         Coef      SE        t      p
----------------------------------------------------------------
Fair health                     -0.3893   0.0407   -9.57   p<0.001
Poor health                     -0.5603   0.0673   -8.31   p<0.001
Widowed                         -0.2491   0.0505   -4.94   p<0.001
Divorced                        -0.2608   0.0438   -5.95   p<0.001
Never married                   -0.2830   0.0309   -9.13   p<0.001
$25,000 or more                  0.0734   0.0971    0.76   p=0.450
Female                           0.0457   0.0236    1.94   p=0.053
```